

UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II



DOTTORATO DI RICERCA IN BIOLOGIA

PhD Course in Biology

XXXVIII CYCLE-PNRR

Design and Development of Molecular Platforms for Targeting Nucleic Acids in Cancer

Sede principale di svolgimento dell'attività sperimentale:

Consiglio Nazionale delle Ricerche -Istituto di Cristallografia — URT Caserta (CNR-IC) —Caserta

Sede secondaria di svolgimento dell'attività sperimentale:

Consiglio Nazionale delle Ricerche -Istituto di Chimica Biomolecolare — Sede di Padova (CNR-ICB) -
Padova

SUPERVISOR

Prof. Angelo Fontana

PhD Student

Martina Rotondo

Co-SUPERVISOR

Dr. Barbara Biondi

Dr. Michele Saviano

2023-2025

TABLE OF CONTENTS

Abbreviations	1
Ammino Acid Residue Abbreviations.....	3
PNA Monomers Abbreviations.....	4
Abstract	5
INTRODUCTION	7
1. Cancer: a broad disease	9
2. Epidemiology and Global Impact	9
2.1. Global Incidence.....	9
2.2. Mortality and Survival.....	10
3. The Cell cycle	11
3.1. Phases of the Cell Cycle	11
3.2. Cell Cycle Control	12
4. Factors Influencing the Cell Cycle.....	13
4.1. PI3K/Akt/mTOR pathway	13
4.2. Ras/Raf/MEK/MAPK pathway	14
5. Carcinogenesis	15
5.1. Oncogenes and Tumour Suppressor Genes	16
6. Hallmarks of Cancer	17
6.1. Promotion of Proliferative Signal.....	18
6.2. Insensitivity to Growth-Inhibitory Signals.....	19
6.3. Resistance to Programmed Cell Death.....	20
6.4. Contact Inhibition and Its Evasion	21
6.5. Acquisition of Unlimited Replicative Potential	21
6.6. Activation of Angiogenesis	22
6.7. Activation of Cellular Invasion and Metastasis	22
6.8. Genomic Instability and Mutation.....	24
6.9. An Enabling Characteristic: Tumour-Promoting Inflammation.....	24
6.10. Reprogramming of Energy Metabolism	25
6.11. Evading Immune Destruction	26
6.12. Tumour Classification.....	26
7. Causes of Cancer	27

7.1. Exogenous risk factors	27
8. Aim of the Thesis	30
References.....	32
CHAPTER I.....	39
Introduction	41
1.1. History	41
1.2. Genomic organisation and classification of miRNAs	42
1.3. Canonical miRNA biogenesis.....	42
1.4. Mechanism of Action and Function of Human miRNAs	44
1.5. microRNAs in Cancer: Oncogenes or Tumour Suppressor	46
1.6. microRNA-221 and Its Role in Cancer	47
1.7. Peptide Nucleic Acids as a tool for silencing oncogenic miRNAs	49
1.8. Cell-Penetrating Peptides (CPPs): Mechanisms and Applications	55
1.9. Aim of the Chapter	61
2. Results and Discussion	63
2.1. Solid Phase Peptide Synthesis	63
2.2. Click Chemistry	96
2.3. Disulfide Bond Formation	105
2.4. Biological Evaluation of CPP-PNA conjugates	112
2.5. Conformational Studies by Circular Dichroism.	129
2.6. Thermal Denaturation (Melting) Studies of miRNA/PNA Duplexes	132
3. Conclusion	136
4. Experimental Part	138
4.1. General Procedure for Solid-Phase Synthesis of PNAs and Peptides.....	138
4.2. Coupling of the First Amino Acid onto 2-Chlorotrityl Chloride Resin.....	139
4.3. Determination of Resin Loading after Fmoc Deprotection by UV-Vis Spectroscopy	139
4.4. Cleavage from the Resin	140
4.5. Analytical Cleavage.....	140
4.6. Synthesis.....	141
4.7. Optimized Procedure for Click Chemistry Reactions	152
4.8. General Procedure for Conjugation via Disulfide Bridge Formation	154
4.9. Culture Cell	155

4.10.	Determination of the Linear Range for CCK-8 Assay	155
4.11.	Cell Proliferation Assay with CF-Conjugated Cell-Proliferation Peptides	156
4.12.	Cell Proliferation Assay General Procedure	156
4.13.	Internalization Studies Protocol.....	157
4.14.	Conformational Studies by CD Spectroscopy	158
4.15.	UV-melting Analysis.....	159
References.....		160
CHAPTER II		170
1.	Introduction.....	172
1.1.	Structural Proteins and Cytoskeleton	172
1.2.	Intermediate Filaments and Vimentin	174
1.3.	Vimentin in Cell Migration and EMT	176
1.4.	Vimentin in Metastasis and Tumour Progression.....	177
1.5.	Interplay between Vimentin and G-Quadruplexes Structures	178
1.6.	G-Quadruplexes and Their Biological Role	179
1.7.	Aim of the chapter	185
2.	Results and Discussion	186
2.1.	Diketopiperazine Formation and Strategies for Its Prevention	186
2.2.	Methionine Substitution with Norleucine for Enhanced Stability	199
2.3.	Alanine Scanning.....	201
2.4.	Fragment-Based and Convergent Peptide Synthesis.....	202
2.5.	Circular Dichroism Characterization of Vimentin-derived Peptides	224
2.6.	NMR Spectroscopy Characterization of Vimentin Fragment	227
2.7.	Interaction Studies	229
3.	Conclusion	242
4.	Experimental Part	244
4.1.	General Procedure for Solid-Phase Synthesis of Peptides	244
4.2.	Acetylation.....	244
4.3.	Convergent synthesis of peptide fragments.....	245
4.4.	Cleavage from the resin.....	245
4.5.	Synthesis.....	246
4.6.	Conformational Studies by CD Spectroscopy	266
4.7.	NMR Spectroscopy	266

4.8. Electrophoretic Mobility Shift Assay	266
4.9. Fluorescence Thermal Assay	267
4.10. Surface Plasmon Resonance	267
References.....	268
CHAPTER III.....	274
1. Introduction.....	276
1.1. Circular Dichroism: Physical Principles and Theoretical Foundations.....	276
1.2. Applications of CD in Secondary Structure Analysis	278
1.3. Synchrotron Radiation: Principles and Properties.....	282
1.4. HT-SRCD/Beamline B23 at Diamond Light Source.....	283
1.5. Aim of the Chapter	287
2. Results and Discussion	288
2.1. Ligand Design and Synthesis	288
2.2. HT-SRCD Analysis of G4-Ligand Interaction	289
3. Conclusion	297
4. Experimental Part	298
4.1. General Procedure for the Synthesis of Rhau25 peptide	298
4.2. Preparation of G4-Forming Sequences	298
4.3. Preparation of small-molecule ligands	298
4.4. High-Throughput Synchrotron Radiation Circular Dichroism	299
References.....	300
Conclusion.....	303
Materials and Methods	307
1. Materials.....	307
2. Instruments	307
Appendix	309
Chapter I.....	309
Chapter II	319

Abbreviations

IARC: International Agency for Research on Cancer

ISS: Italy's National Institute Health

CDK: Cyclin-Dependent Kinase

Rb: Retinoblastoma protein

MPF: Maturation-Promoting Factor

GSK3: Glycogen Synthase Kinase 3

NSCLC: Non-Small Cell Lung Cancer

EGFR: Epidermal Growth Factor Receptor

HER2: Human Epidermal growth factor Receptor 2

PDGF: Platelet-Derived Growth Factor

TGF α : Transforming Growth Factor α

TGF β : Transforming Growth Factor β

VEGF-A: Vascular Endothelial Growth Factor A

TSP-1: Thrombospondin

EMT: Epithelial-Mesenchymal Transition

PET: Positron Emission Tomography

NK: Natural Killer

MiRNA: microRNA

MiR-221: microRNA-221

siRNA: Small Interfering RNA

mRNA: messenger RNA

lncRNA: Long Non-Coding RNA

pre-miRNA: Precursor microRNA

AGO: Argonaute protein

RISC: RNA-Induced Silencing Complex

MiRNA-RNP: microRNA-RNA-Binding Protein

LNA: locked Nucleic Acid

PNA: Peptide Nucleic Acid

CPP: Cell Penetrating Peptide

NLS: Nuclear Signal

PNT: Penetratin

MAP: Model Amphipathic Peptide

PSA: α -2,8-Polysialic Acid

Abbreviations

CD: Circular Dichroism

NMR: Nuclear Magnetic Resonance

CIP rules: Cahn-Ingold-Prelog rules

SPPS: Solid Phase Peptide Synthesis

DVB: Divinylbenzene

DMF: Dimethylformamide

DCM: Dichloromethane

NMP: N-methylpyrrolidone

CTC resin: 2-chlorotrityl chloride resin

TFA: Trifluoroacetic Acid

Boc: *tert*-butyloxycarbonyl

HF: hydrofluoric acid

Fmoc: fluorenylmethyloxycarbonyl

DBU: 1,8-Diazabicyclo[5.4.0]undec-7-ene

DBF: dibenzofulvene

HBTU: Hexafluorophosphate Benzotriazole Tetramethyl Uronium

HOBt: Hydroxybenzotriazole

HATU: O-(7-Azabenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate

HOAt: 1-Hydroxy-7-azabenzotriazole

CF: carboxyfluorescein

DIC: N,N'-Diisopropylcarbodiimide

Oxyme Pure: ethyl-2-cyano-2-(hydroxyimino)acetate

DIPEA: N,N-Diisopropylethylamine

NCA: N-carboxyanhydrides

DKP: diketopiperazine

TIS: triisopropylsilane

EDT: Ethane-1,2-dithiol

DODT: dithiodiethanethiol

CuACC: Cu(I)-catalysed azide-alkyne cycloaddition

DTNP: 2,2'-dithiodipyridine

CQ: Chloroquine

FBS: Fetal Bovine Serum

GFAP: Glial Fibrillary Acidic Protein

IF: Intermediate Filaments

ECM: Extracellular Matrix

MMP: Matrix Metalloproteinase

G4: G-Quadruplex

G4rp: G-Quadruplex repeat

Ala Scan: Alanine Scanning

HFPI: 1,1,1,3,3,3-hexafluoro-2-propanol

EMSA: Electrophoretic Mobility Shift Assay

FM: Fluorescence Melting

ssDNA: Single-Stranded DNA

SPR: Surface Plasmon Resonance

SRCD: Synchrotron Radiation Circular Dichroism

HT-SRCD: High-Throughput Synchrotron Radiation Circular Dichroism

CPL: Circularly Polarised Light

NPL: Non-Polarised Light

LPL: Linearly Polarised Light

PEM: Photoelastic Modulator

MMP: Mueller Matrix Polarimeter

PMT: Photomultiplier Tube

Amino Acid Residue Abbreviations

A= Ala: Alanine

C = Cys: Cysteine

D = Asp: Aspartic Acid

E = Glu: Glutamic Acid

F= Phe: Phenylalanine

G = Gly: Glycine

H = His: Histidine

I = Ile: Isoleucine

K= Lys: Lysine

L = Leu: Leucine

nL= nLeu: norLeucine

M = Met: Methionine

N = Asn: Asparagine

P = Pro: Proline

Abbreviations

Q = Gln: Glutamine

R = Arg: Arginine

T = Thr: Threonine

V = Val: Valine

W = Trp: Tryptophan

Y = Tyr: Tyrosine

Ahx: 6-Aminohexanoic Acid

β -Ala: β -Alanine

PNA Monomers Abbreviations

Fmoc-AEG(NB)-OH: Fmoc-AEG(Nucleobase)-OH

Fmoc-AEG(A)-OH: Fmoc-AEG(Adenine)-OH, Fmoc-*A*-OH

Fmoc-AEG(C)-OH: Fmoc-AEG(Cytosine)-OH, Fmoc-*C*-OH

Fmoc-AEG(G)-OH: Fmoc-AEG(Guanine)-OH, Fmoc-*G*-OH

Fmoc-AEG(T)-OH: Fmoc-AEG(Thymine)-OH, Fmoc-*T*-OH

Abstract

This thesis investigates the development and characterisation of peptide- and small molecule-based systems designed to recognise and modulate nucleic acids. The unifying focus is on molecular mechanisms relevant to cancer biology and therapeutic innovation. An integrated strategy combining synthetic organic chemistry, peptide chemistry, circular dichroism (CD), fluorescence spectroscopy, surface plasmon resonance, nuclear magnetic resonance (NMR) and cellular assays is employed to explore how molecular architecture and environmental conditions dictate the recognition of oncogenic nucleic acid targets.

The first section of the thesis addresses the synthesis of Peptide Nucleic Acid (PNA) conjugates functionalised with Cell-Penetrating Peptides (CPPs) for modulating the activity of the oncogenic microRNA miR-221-5p, which is a key regulator involved in tumour progression and resistance pathways. CPP-PNA conjugates were obtained via CuACC and disulfide coupling strategies and evaluated in terms of internalisation and biological activity. Although hybridisation capability was preserved, internalisation assays revealed limitations related to aggregation that highlight the need for improved formulation. Nevertheless, these results establish a solid methodological basis for the future refinement of PNA-based gene-silencing tools.

The second part of the study examined peptide fragments from Vimentin, an intermediate filament protein that plays a central role in epithelial-mesenchymal transition (EMT), cell motility, and metastatic dissemination. CD and NMR analyses revealed that Vimentin-derived peptides adopt different conformational states under membrane-mimetic conditions. Alanine-scanning variants revealed the critical importance of aromatic and basic residues in mediating interactions with G-quadruplex (G4) DNA – noncanonical structures associated with the transcriptional regulation of oncogenes – thus offering insight into how the intrinsically disordered regions of Vimentin may recognise nucleic acids in contexts relevant to cancer.

The third part extends the investigation of G4 recognition to tetrazole-based small molecules and the peptide Rhau25, employing high-throughput synchrotron radiation circular dichroism (HT-SRCD) – a label-free technique ideally suited to rapid ligand screening. These studies revealed that ligand accessibility and induced structural changes are strongly influenced by ionic conditions, with Na⁺ promoting more flexible G4 conformations conducive to binding. Several tetrazole derivatives were found to selectively perturb telomeric and regulatory G4s, while Rhau25 robustly remodelled G4 topologies. These findings identify promising hit compounds for the future development of G4-targeting anti-cancer ligands and validate HT-SRCD as an efficient discovery tool.

Abstract

Overall, this thesis integrates synthetic chemistry with structural and biophysical analysis to provide new mechanistic insights and methodological foundations for developing next-generation nucleic acid-targeting agents for use in cancer therapeutics.

Introduction

Cancer: Origins, Hallmarks, and Clinical Implications



1. Cancer: a broad disease

The term ‘cancer’ does not refer to a single disease, but rather to a diverse group of disorders that differ significantly in terms of symptoms, location, and affected tissues or organs. However, they all have one key feature in common: abnormal and uncontrolled cell growth that evades normal regulatory mechanisms.[1]

Three terms are commonly used when discussing these conditions: ‘neoplasia’, ‘tumour’, and ‘cancer’.

‘Neoplasia’ refers to new cellular growth that replaces pre-existing tissue in a way that deviates from normal physiology. Depending on its characteristics, a neoplasm may be benign or malignant.[2]

The term ‘tumour’ describes the development of a neoplasm into a mass of cells that presents as a palpable lump.

Finally, ‘cancer’ specifically refers to a malignant neoplasm, which is characterised by its ability to infiltrate the body and spread through metastasis.[1]

2. Epidemiology and Global Impact

Oncological epidemiology investigates the distribution and determinants of cancer within populations. This provides essential data for cancer prevention, healthcare planning and research. The GLOBOCAN 2024 report, published by the International Agency for Research on Cancer (IARC), is one of the most comprehensive references worldwide and offers an in-depth overview of the global cancer burden.[3]

2.1. Global Incidence

In 2024, an estimated 20.6 million new cancer cases were recorded, resulting in 10.3 million cancer-related deaths. These striking figures are expected to rise further; according to GLOBOCAN 2024 projections, there will be more than 30 million new cases annually by 2045, mainly due to population ageing and increased exposure to risk factors in low- and middle-income countries.[3]

The most frequently diagnosed cancers worldwide are:[4]

- Breast cancer: 2.48 million new cases (12.1% of all diagnoses). It is now the most commonly diagnosed cancer globally, surpassing lung cancer.
- Lung cancer: 2.45 million cases (11.9%). Despite having a slightly lower incidence rate than breast cancer, lung cancer remains the leading cause of cancer-related deaths due to its high fatality rate.

Introduction



- Colorectal cancer: 2.02 million cases (9.8%). Incidence is rising in both high-income countries and regions that previously had low rates.
- Prostate cancer: 1.55 million cases (7.5%). It shows a significant geographical variability linked to genetic factors and the widespread use of PSA screening.

2.2. Mortality and Survival

The incidence of cancer does not always align with mortality rates. For example, lung cancer has a similar incidence rate to breast cancer but causes nearly twice as many deaths. This is largely due to late diagnosis and the limited effectiveness of therapies in advanced stages.[5]

Survival outcomes vary widely and are influenced by[6]:

- Access to early detection programs (e.g. mammography and colonoscopy);
- The availability of advanced treatments (such as targeted therapies and immunotherapy);
- Socioeconomic and geographic disparities.

In high-income countries, the five-year survival rate for breast cancer exceeds 90%, whereas in many parts of sub-Saharan Africa, it is below 40%.[7]

The latest data from Italy's National Institute Health (ISS) highlights a positive trend: in 2023, the five-years survival rate was 60.8% for men and 66.5% for women and has steadily improved over time. Moreover, Italy reports survival rates higher than the European average for several cancers:[8]

- Breast cancer: 88% (Italy) vs 83% (Europe)
- Colorectal cancer: 63% vs 59%
- Prostate cancer: 91% vs 85%
- Kidney cancer: 69% vs 62%
- Stomach cancer: 34% vs 26%

Some cancers show particularly favourable prognoses. For example, testicular cancer has cure rates above 95%, and in localised cases, five-year survival can exceed 99%.[8]

Overall, Italy ranks among the countries with the highest five-year survival rates in Europe, largely thanks to effective prevention strategies, early detection, and increasingly advanced treatments.[9]



3. The Cell cycle

The cell cycle is the orderly sequence of events by which a single parent cell grows and divides to produce two genetically identical daughter cells. The cycle is divided into distinct phases, each of which ensures the accurate duplication of cellular components. This subdivision also facilitates strict regulatory control. If these control mechanisms fail, however, abnormal proliferation may occur, potentially leading to cancer.[10]

3.1. Phases of the Cell Cycle

The cell cycle is classically divided into five main phases: G_0 , G_1 , S, G_2 , and M (Figure 1).

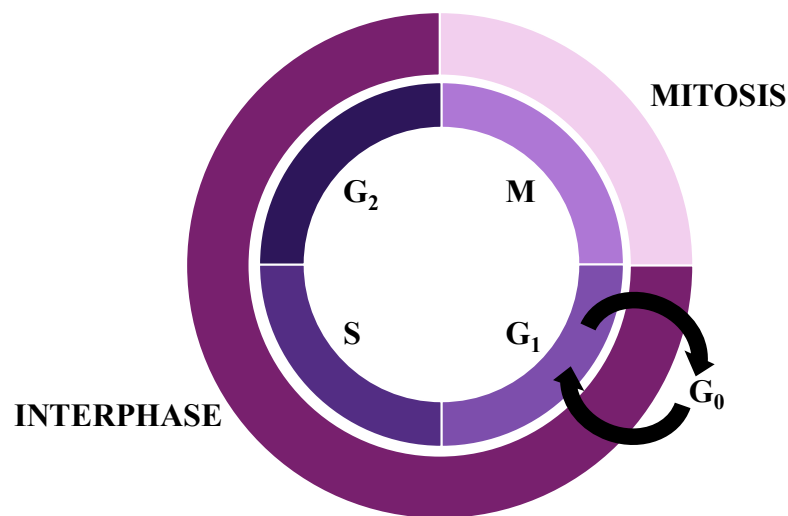


Figure 1. Schematic representation of the cell cycle.

- **G_0 Phase (quiescence).** Although often referred to as a resting phase, G_0 is more accurately described as a state of quiescence.[11] During this phase, the cell does not proliferate but continues to perform essential functions such as nutrient uptake and protein synthesis. This phase can be reversible, allowing the cell to re-enter the cycle, or irreversible, as in the case of senescent or terminally different cells.[11]
- **G_1 Phase (cell growth).** During this phase, the cell undergoes extensive biosynthetic activity, resulting in an increase in size and organelle content. At the end of this phase, the cell has three options: it can exit the cycle into G_0 for differentiation; progress into the S phase; or enter a reversible G_0 state.[12] The transition from G_1 to S is tightly regulated by the G_1/S checkpoint and is controlled by cyclins and cyclin-dependent kinases (CDKs). These ensure that conditions are favourable before DNA replication begins.[12]
- **S Phase (DNA synthesis).** Once the checkpoint is passed, DNA replication begins. Each chromosome is duplicated to form sister chromatids. Protein synthesis is largely downregulated

Introduction



during this phase, except for the production of histones, which are essential for packaging the newly replicates DNA into nucleosomes.[13]

- **G₂ Phase (preparation for mitosis).** During this phase, the cell resumes growth and synthesises the proteins and other components required for division. This phase includes a crucial DNA damage checkpoint where errors arising during replication are detected. The protein p53 plays a central role in this surveillance, activating DNA repair pathways when needed. Entry into mitosis requires the activation of the cyclin B1/CDK1 complex once a threshold level is reached.[14]
- **M Phase (mitosis and cytokinesis).** This is the stage at which the duplicated chromosomes are accurately segregated into the two daughter nuclei. Mitosis is subdivided into prophase, prometaphase, metaphase, anaphase, and telophase.[15] The chromosomes are aligned and then pulled apart to opposite poles by the mitotic spindle, which is composed of microtubules. Finally, cytokinesis completes the process: a contractile actin ring constricts the parent cell to produce two daughter cells with identical genetic material and roughly equal organelle distribution. Errors in mitosis can trigger apoptosis or contribute to tumour development.[15]

3.2. Cell Cycle Control

The cell cycle is a tightly regulated process whose progression depends on passing specific checkpoints. These control mechanisms ensure accurate cell division, preventing errors that could compromise cell viability or genetic stability. There are three major checkpoints: the G₁ checkpoint, the G₂ checkpoint, and the mitosis (M phase) checkpoint.

- **G₁ checkpoint (G₁/S restriction point).** At this stage, the cell assesses its internal and external conditions before committing to DNA replication. It checks for nutrient availability, growth factors, proper protein synthesis and DNA integrity. Depending on the outcome, the cell may enter G₀, which involves quiescence or differentiation, or proceed to the S phase. This is regulated by oscillating levels of cyclins and the activity of cyclin-dependent kinases (CDKs).[16]

A key regulatory axis here is the Rb-E2F complex:

- E2F transcription factors activate the expression of cyclins, checkpoint regulators and DNA repair proteins.[17]
- Rb (the retinoblastoma protein) binds and inhibits E2F. However, E2F is only released when Rb is phosphorylated by cyclin D-CDK4/6 and subsequently by cyclin E-CDK2, enabling the cell to enter the S phase. This marks the late restriction point in G₁. [17]



- **G₂ checkpoint (pre-mitotic control).** During the G₂ phase, the cell verifies that DNA replication and organelle duplication have been successfully completed. Entry into mitosis is permitted only under optimal conditions.[18] This transition is driven by the cyclin B-CDK1 complex, also known as the maturation-promoting factor (MPF). Cyclin B levels gradually increase from early G₁ until they reach a threshold. At this point, binding to CDK1 activates a cascade that triggers mitotic entry.[18]
- **M checkpoint (spindle assembly checkpoint).** During mitosis, specifically between metaphase and anaphase, a control mechanism ensures that all chromosomes are properly aligned and attached to the mitotic spindle. Only when this requirement is met does the cell proceed with chromatid separation. This checkpoint is crucial in preventing chromosome mis-segregation, which could lead to aneuploidy and contribute to tumour development.[19, 20]

4. Factors Influencing the Cell Cycle

Cell proliferation is controlled by external stimuli that signal to the cell whether to grow or remain quiescent. Mutations in genes that encode proteins involved in these signalling pathways can disrupt this regulation, often leading to uncontrolled growth and tumour development.[21]

The two major signalling cascades that regulate proliferation are:

- The PI3K/Akt/mTOR pathway;
- The Ras/Raf/MEK/MAPK pathway.

4.1. PI3K/Akt/mTOR pathway

The PI3K/Akt/mTOR signalling pathway (Figure 2) begins with the activation of PI3K by a ligand-bound receptor tyrosine kinase (RTK).[22] Once activated, PI3K phosphorylates membrane phosphatidylinositols to generate PIP3, which acts as a docking site for PDK1 and Akt. Akt is then phosphorylated by mTORC2. In its active state, Akt modifies several intracellular substrate, including glycogen synthase kinase 3 (GSK3), FOXO transcription factors and the pro-apoptotic protein BAD.[23] Through these actions, Akt suppresses apoptosis and promotes cell-cycle progression. A key step in this cascade is the indirect activation of mTORC1. Akt phosphorylates the TSC1/TSC2 complex, which normally acts as a negative regulator. This inhibition allows the small GTPase-Rheb-GTP to activate mTORC1.[24] Acting as a metabolic sensor, mTORC1 integrates signals from growth factors, nutrient availability and cellular energy status. In response, it promotes anabolic processes, such as the synthesis of proteins, lipids, and organelles, while simultaneously inhibiting catabolic processes, such as autophagy.[24]

Introduction



The pathway is terminated through negative feedback mechanisms: phosphorylated mTORC1 suppresses PI3K activity, while the phosphatase PTEN converts PIP3 back in PIP2, thereby preventing further PI3K activation. Due to its pivotal role in controlling cell growth and survival, this pathway is often hyperactivated in various cancers, including non-small cell lung cancer (NSCLC).[22]

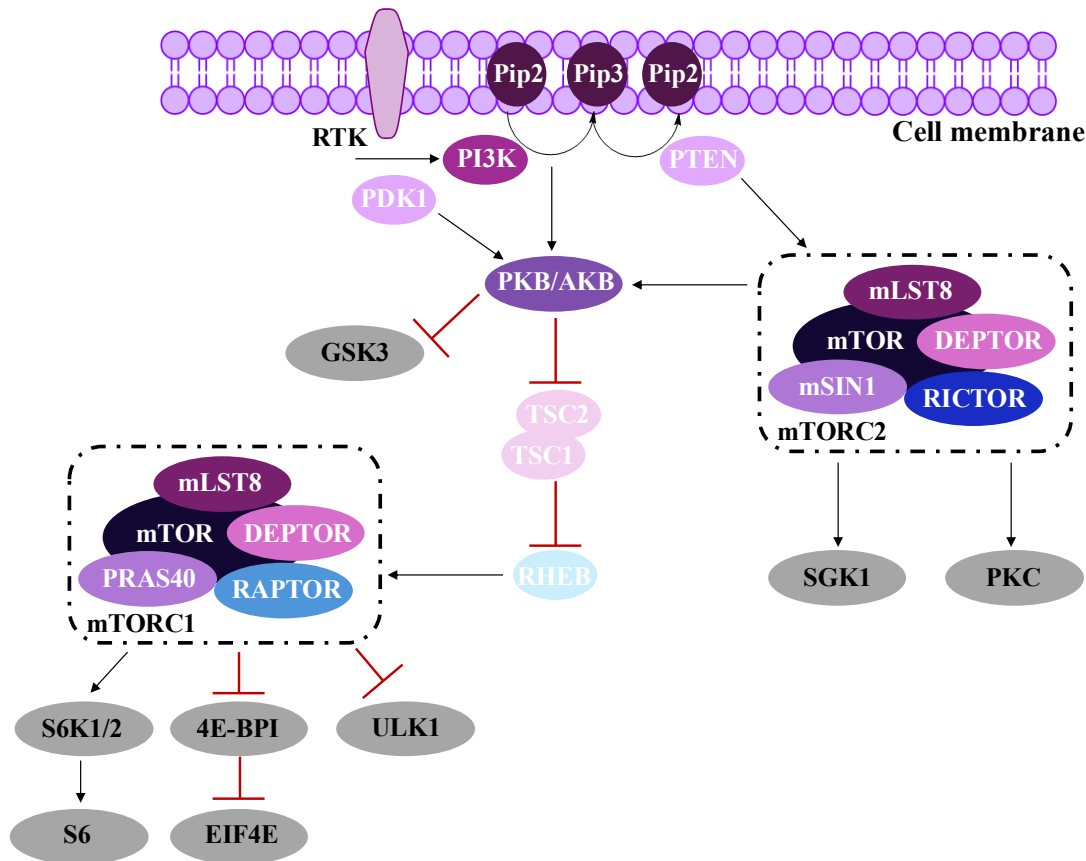


Figure 2. PI3K/Akt/mTOR pathway.

4.2. Ras/Raf/MEK/MAPK pathway

The Ras/Raf/MEK/MAPK pathway (Figure 3) is a key intracellular signalling cascade that stimulates cell proliferation. It begins when an external stimulus, such as the growth factor TGF α , binds to the receptor tyrosine kinase EGFR. This triggers the receptor to undergo autophosphorylation at specific tyrosine residues, which then serve as docking sites for adaptor proteins such as GRB2. GRB2 then recruits the guanine nucleotide exchange factors SOS through its SH2 and SH3 domains.[25]

Formation of the phosphorylated EGFR-GRB2-SOS complex activates Ras, a small GTPase that plays a central role in multiple signalling pathways. Ras cycles between an active state when bound to GTP and inactive state when bound to GDP. Once activated, Ras phosphorylates RAF, which activates MEK. Activated MEK then stimulates MAPK, which translocates into the nucleus and activates several transcription factors, including those responsible for expressing the oncogene s-myc.[26]

This cascade induces the cell to synthesise new proteins, enhance survival mechanisms and reduce apoptosis. Furthermore, MAPK activation is crucial for regulating the cell cycle, specifically promoting the production of cyclin D during G₁ phase. This provides the necessary signal to overcome the restriction point, allowing progression into cell division.[27]

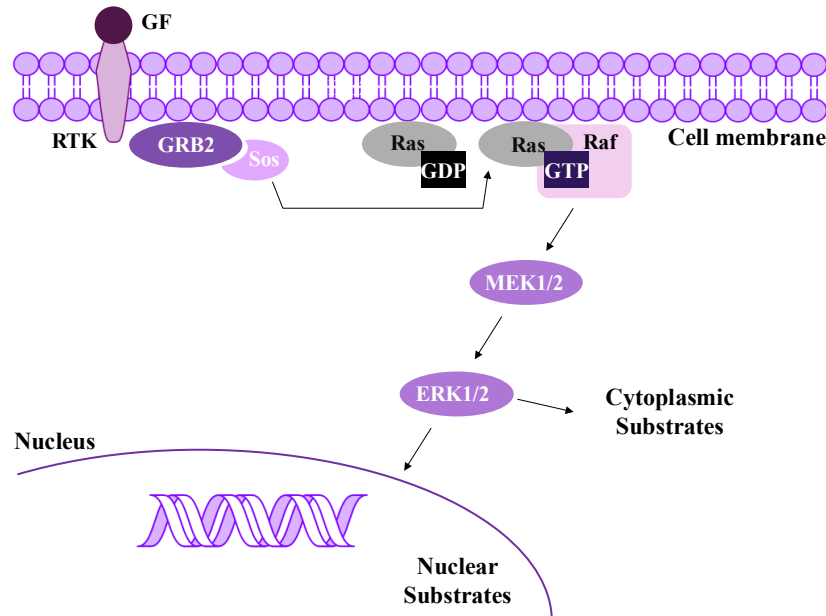


Figure 3. *Ras/Raf/MEK/MAPK pathway.*

5. Carcinogenesis

In essence, cancer is a genetic disease. Multiple DNA mutations (Figure 4) accumulate to cause neoplastic transformation, giving the cell distinctive traits such as uncontrolled proliferation, evasion of programmed cell death, the ability to invade surrounding tissues, and the potential to form distant metastases. These mutations may be caused by external factors, such as tobacco smoke, or they may be hereditary. However, a single mutation is insufficient to transform a normal cell into a tumour cell. Rather, a gradual accumulation of mutations is required, particularly in genes that regulate growth, proliferation, and cell death, as well as those responsible for DNA repair and immune evasion (Figure 4).[1]

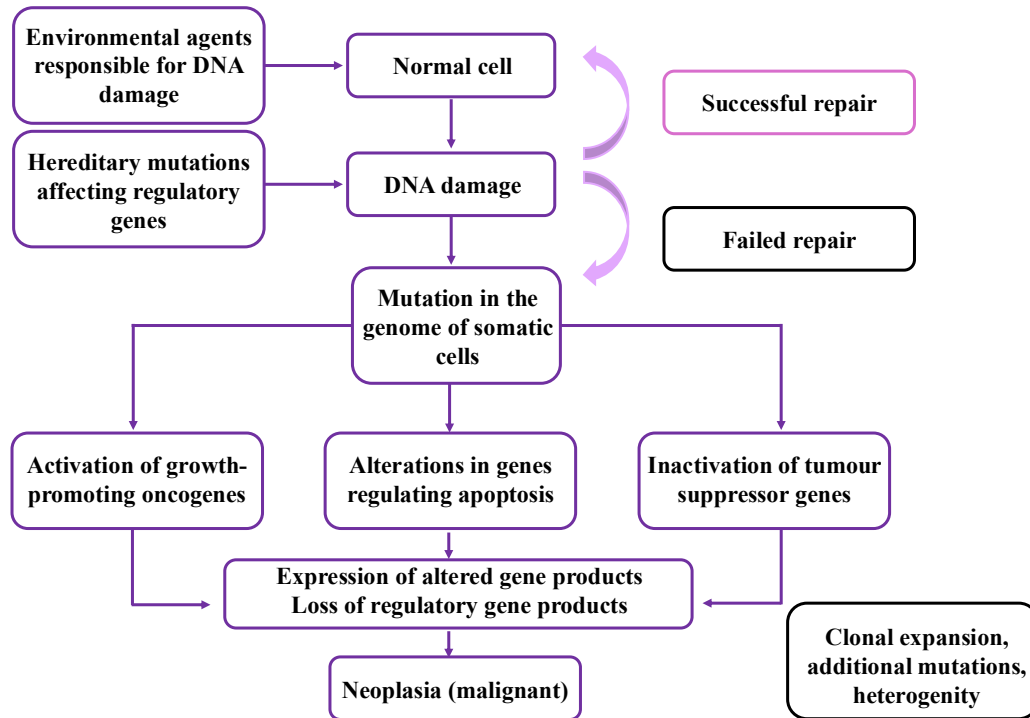


Figure 4. Carcinogenesis pathway.

The genes involved in this process can be broadly divided into two major categories:

1. **Oncogenes.** They typically promote cell growth, but when mutated, they trigger uncontrolled proliferation.[28]
2. **Tumour suppressor genes.** They safeguard genomic integrity and restrain excessive cell division, and their inactivation removes critical protective barriers.[28]

Other relevant gene classes include those that regulate apoptosis and those involved in DNA replication and repair.[29]

The progression from a normal cell to a tumour cell is therefore gradual and highly complex. Each time genetic damage occurs, the organism attempts to correct the error through repair mechanisms. If this is not possible, the damaged cell is usually eliminated by apoptosis. It is only when these defences fail, and mutations continue to accumulate unchecked that the cell progressively acquires the defining features of the tumour phenotype.[30]

5.1. Oncogenes and Tumour Suppressor Genes

Oncogenes originate from mutated or hyperactivated forms of proto-oncogenes, which are a family of genes that normally regulate essential physiological processes, such as cell proliferation, survival, differentiation, angiogenesis, and migration. When these functions are altered, they contribute to the hallmarks of cancer.[28] Proto-oncogenes can encode various proteins, including growth factors such



as EGF and PDGF; growth factor receptors such as EGFR and VEGFR; cell-cycle regulators including cyclins; transcriptional regulators; and signal transduction proteins such as Ras.[31]

In cancer, proto-oncogenes acquire mutations that alter their expression and convert them into oncogenes. This transformation can occur through several mechanisms: gene overexpression, which leads to the excessive production of the encoded protein; gene amplification, which increases the number of gene copies; or constitutive activation, which results in a protein that remains active regardless of external stimuli. Prominent examples include ErbB, c-myc and Bcl-2, which all play a critical role in regulating cell growth, survival and inhibition of apoptosis.[32]

In contrast, tumour suppressor genes act as fundamental barriers against neoplastic transformation. They encode proteins that perform various functions, such as regulating the cell cycle, preventing uncontrolled proliferation, repairing damaged DNA, maintaining cell adhesion and inducing apoptosis in response to irreparable damage.[33]

One of the best-characterised tumour suppressors is the Rb protein, which is a key regulator of DNA replication and cell-cycle progression. Initially associated with retinoblastoma, Rb inactivation has also been observed in other cancers, including bladder, breast, and lung carcinomas. This highlights the importance of tumour cells bypassing this critical negative control mechanism.[33] Another key player is p53, often referred to as the ‘guardian of the genome’. This protein, encoded by the TP53 gene, responds to genomic stress, DNA damage and metabolic imbalances by activating transcription factors that increase its levels and enhance its activity. Depending on the context, p53 can either halt the cell cycle until normal conditions are restored, or trigger apoptosis when the damage is beyond repair.[34]

In tumour cells, however, this control system is often disrupted. Gene deletions can completely abolish tumour suppressor activity, while inactivating mutations often produce altered non-functional proteins. Additionally, tumour suppressors also contribute to this regulatory network. For example, PTEN counteracts the PI3K/Akt/mTOR pathway through its phosphatase activity, while proteins such as BCL2 and BRAF are involved in apoptosis regulation and cell-cycle control, respectively.[35]

6. Hallmarks of Cancer

In 2000, six defining features of tumour cells were identified and named the hallmarks of cancer (Figure 5). These traits represent the capabilities acquired by cells during neoplastic transformation and help to explain their abnormal behaviour.[36] In fact, tumour cells display a persistent ability to promote proliferative signalling while simultaneously becoming insensitive to the growth-inhibitory mechanisms that normally regulate cell division. They develop resistance to programmed cell death, acquire limitless replicative potential and stimulate the formation of new blood vessels through

Introduction



angiogenesis, thereby ensuring a continuous supply of nutrients and oxygen.[37] Another crucial feature is their capacity to invade surrounding tissues and spread to distant organs – a process that underlies the development of metastases, making one of the most dangerous aspects of cancer progression.[38]

With advances in cancer research, this framework was expanded in 2011 to include four additional features. It has been demonstrated that tumour cells can reprogram their energy metabolism to sustain accelerated growth and can evade detection and elimination by the immune system. Furthermore, they exhibit high levels of genomic instability, fostering the accumulation of mutations and frequently promote chronic inflammation, a condition that contributes to tumour progression.[39]

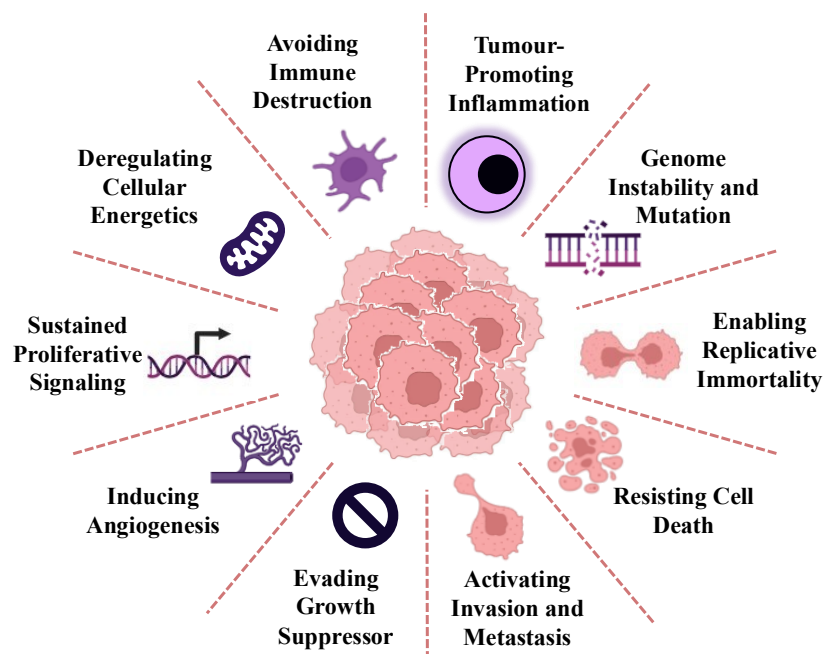


Figure 5. Schematic representation of Hallmarks of Cancer.

6.1. Promotion of Proliferative Signal

In normal cells, proliferation is tightly controlled by a balance of growth-promoting and inhibitory factors. The former drive progression through the cell cycle and support differentiation, while the latter limit these processes.[40] This equilibrium ensures that cell division only occurs when necessary for tissue maintenance and function. Growth factors act by binding to receptors located either on the cell membrane – often receptor tyrosine kinases – or inside the cell. This binding triggers complex intracellular signalling cascades that ultimately stimulate cell growth, survival and replication.[40]

In cancer cells, however, this regulatory system becomes disrupted. One of the earliest hallmarks of malignancy is the acquisition of proliferative autonomy. This can result from genetic alterations leading to the overexpression of growth factor receptors, which renders cells hypersensitive to external stimuli. Consequently, a single molecule of growth factor, which would normally elicit a



modest response, can provoke an amplified proliferative signal.[35] Notable examples include the epidermal growth factor receptor (EGFR/erbB), which is often overexpressed in breast, brain, and gastric cancers,[41] and the human epidermal growth factor receptor 2 (HER2/neu), which is frequently upregulated in gastric and breast tumours.[42]

Cancer cells may also bypass external signalling altogether by producing their own growth factors, thereby creating an autocrine loop that sustains continuous proliferation. Platelet-derived growth factor (PDGF), for example, is produced in glioblastoma, while transforming growth factor α (TGF α) is found in sarcomas.[43]

Another mechanism involves mutations that cause the constitutive activation of signalling proteins, such as Ras.[44] In this state, Ras remains active even in the absence of GTP binding, continuously driving downstream signalling. Such mutations are present in a wide range of cancers, particularly melanoma, where they occur around 40% of cases. Together, these changes give tumour cells an ongoing urge to replicate and differentiate, which is one of the key features of cancer.[44]

6.2. Insensitivity to Growth-Inhibitory Signals

In healthy tissues, cell proliferation is not only kept in check by growth-promoting factors, but also by a variety of inhibitory signals that safeguard tissue homeostasis. These signals can be soluble molecules, components of the extracellular matrix or proteins expressed on the surface of neighbouring cells. By binding to specific surface receptors, they transmit signals that limit excessive cell division.[45] There are two possible outcomes: cells may be pushed into a reversible quiescent state (G_0), from which they can re-enter the cell cycle when conditions become favourable; or they may be induced to permanently lose their proliferative capacity.[46]

For cancer cells to replicate unrestrictedly, they must evade these inhibitory mechanisms. At the molecular level, most antiproliferative signals converge on the retinoblastoma protein (pRb) and its close relatives, p107 and p130. Under normal conditions, pRb binds to transcription factors of the E2F family, preventing them from activating genes required for DNA replication.[47] In tumours, however, mutations or dysregulation of the phosphorylation state of pRb lead to the uncontrolled release of E2F. This allows E2F to move into the nucleus and drive transcriptional programmes that favour proliferation.[48]

Another crucial player is the transforming growth factor β (TGF β), a potent suppressor of cell division. One of its main actions is to prevent the inactivation of pRb by inducing the expression of proteins such as p15^{INK4A/B} and p21. These proteins then inhibit the cyclin-CDK complexes that would otherwise phosphorylate and deactivate pRb.[49] However, many tumour cells acquire resistance to TGF β , either by downregulating its receptors or through mutations that render these



receptors non-functional.[50] Furthermore, mutations or deletions affecting genes that encode key regulators of cell cycle have been observed in several cancers. This includes the RB1 gene itself. Inactivating alterations to this gene produce a defective pRb protein that is unable to perform its crucial role of preventing excessive cell growth.[51]

6.3. Resistance to Programmed Cell Death

Programmed cell death, also known as apoptosis, acts as a natural defence against cancer development. This process is not random, but rather the outcome of a delicate balance between signals that drive the cell towards apoptosis (such as DNA damage) and signals that support survival and proliferation instead. When pro-apoptotic stimuli prevail, the cell is eliminated; when survival signals dominate, the cell continues to divide.[52]

The apoptotic programme is controlled by a complex network of upstream regulators, which receive and integrate signals, and downstream effectors, which execute cell death.[53] There are two main pathways that can initiate apoptosis: the extrinsic pathway, which is triggered by extracellular death signals;[54] and the intrinsic pathway, which responds to intracellular stresses, such as genomic damage or metabolic imbalances.[55] Ultimately, both pathways converge on the activation of a cascade of enzymes known as caspases, which orchestrate the orderly dismantling of the cell. This allows neighbouring cells or phagocytes to clear the remains without causing inflammation.[53]

The intrinsic pathway plays the most crucial role in this process as a barrier to cancer development. Many studies have shown that apoptosis is strongly attenuated in advanced tumours, which grow uncontrollably and develop resistance to therapy.[56]

Cancer cells adopt a variety of strategies to evade apoptosis. One of the most common is the loss of the function of the tumour suppressor TP53, which removes a key control point in apoptotic signalling, allowing the accumulation of mutations that drive tumour progression.[57] Alternatively, tumours may increase the expression of anti-apoptotic proteins such as Bcl-2 and Bcl-xL, or enhance survival signalling through factors such as IGF1/2. At the same time, they may reduce levels of pro-apoptotic factors or block the extrinsic pathway by altering signal transduction proteins. This prevents death receptor signalling from triggering apoptosis.[58]

By evading this fundamental defence mechanism, cancer cells ensure their survival and enhance their resistance to conventional therapies.[59]



6.4. Contact Inhibition and Its Evasion

Decades of research have demonstrated that normal cells can regulate their growth through a process known as contact inhibition. When cells proliferate and reach confluence, they form organised monolayers and stop dividing due to their physical contact with neighbouring cells. This mechanism ensures tissue homeostasis and prevents uncontrolled growth.[60] However, this behaviour is not observed in cancer cells, which continue to divide even after forming dense layers and thus lose this essential growth control.[61]

The product of the NF2 gene, which has long been recognised as a tumour suppressor, is a key molecular player in contact inhibition. Loss of NF2 function is associated with the development of neurofibromatosis, which highlights the critical role of the gene in controlling growth. The NF2 gene encodes Merlin, a protein that enforces contact inhibition by linking cell surface adhesion molecules, such as E-cadherin, with transmembrane receptor tyrosine kinases, including EGFR. By doing this, Merlin strengthens cadherin-mediated cell-cell junctions, thereby stabilising tissue architecture.[62] Furthermore, Merlin acts as a regulator of signalling by ‘masking’ growth factor receptors. This reduces their ability to efficiently activate proliferative pathways, even in the presence of external stimuli. In tumour cell, loss or inactivation of Merlin disrupts this control system, weakening intercellular adhesion and allowing growth factor receptors to signal unchecked. This ultimately promotes uncontrolled proliferation.[62]

6.5. Acquisition of Unlimited Replicative Potential

In most cases, normal cells have a finite replicative lifespan, except for stem cells. Once this limit is reached, the cells undergo senescence or programmed cell death. This mechanism acts as a physiological barrier that prevents uncontrolled proliferation.[63]

However, cancer cells are able to bypass this restriction. Numerous *in vitro* studies have shown that they behave as ‘immortal’, continuing to divide indefinitely. This suggests that acquiring unlimited replicative potential is not just an *in vitro* phenomenon, but a fundamental step in tumour progression.[64]

The primary factor underlying the limited lifespan of healthy cells is the progressive shortening of telomeres – the protective structures at the ends of chromosomes. With each cell division, the telomeres become shorter until they can no longer protect chromosomal DNA, ultimately triggering cellular senescence.[65] Hayflick famously described this process as a ‘replicometer’, a biological counter that determines the replicative potential of normal cells and represents a critical hurdle in carcinogenesis.[66]



Telomerase plays a central role in overcoming this barrier; it is a specialised DNA polymerase capable of adding repetitive nucleotide sequences to telomeres, thereby counteracting their progressive erosion. In differentiated somatic cells, telomerase activity is virtually absent, whereas it is strongly reactivated in approximately 90% of tumour cells.[67] This reactivation enables cancer cells to maintain telomere length, allowing them to proliferate indefinitely and supporting their defining feature of cellular immortality.[67]

6.6. Activation of Angiogenesis

In order to grow and survive, tumours require a constant supply of oxygen and nutrients, as well as efficient mechanism for removing metabolic waste and carbon dioxide. These needs are met through angiogenesis, or the formation of new blood vessels.[68] In normal adult tissues, angiogenesis is usually only temporary, occurring in specific situations such as wound healing. In tumours, however, the so-called ‘angiogenic switch’ remains persistently activated, resulting in continuous vascular growth that supports the expansion of the neoplasm.[69]

This process is tightly regulated by the balance between pro-angiogenic and anti-angiogenic signals. One of the best-known inducers is vascular endothelial growth factor A (VEGF-A), which is essential for embryonic and postnatal development, as well as for the homeostatic survival of endothelial cells under physiological and pathological conditions.[70] VEGF signalling can be upregulated by hypoxia or through oncogenic activation. Conversely, thrombospondin-1 (TSP-1) acts as a major inhibitor by binding to specific receptors on endothelial cells and triggering inhibitory cascades that counteract pro-angiogenic signalling.[70]

The vasculature generated by tumours is clearly distinguishable from that of healthy tissues. New vessels often sprout prematurely, showing excessive and convoluted branching. These vessels tend to be dilated, distorted and structurally unstable, resulting in irregular blood flow and frequent microhaemorrhages. Additionally, endothelial cells exhibit abnormal proliferation and apoptosis rates, further compromising the integrity of the vascular network.[71]

Notably, angiogenesis is now recognised as a process that can be activated very early in tumour development, even when the mass is still microscopic. This finding contrasts with earlier assumptions that angiogenesis only occurs once tumours had reached a macroscopic size.[72]

6.7. Activation of Cellular Invasion and Metastasis

The invasion and metastasis process does not occur in a single step but rather follows a complex sequence of events often referred to as the ‘invasion-metastasis cascade’.[72] This model outlines the successive biological changes that tumour cells undergo during their progression:



1. Transformation of a normal cell into a malignant one;[72]
2. Activation of angiogenesis;[68]
3. Acquisition of motility and infiltration of nearby blood and lymphatic vessels by tumour cells;[73]
4. Formation of aggregates composed of tumour cells, platelets and lymphocytes. Rather than neutralising them, these aggregates protect cancer cells from immune surveillance and facilitate their adhesion to the capillary endothelium;[74]
5. Transport of malignant cells through the circulatory and lymphatic systems;[74]
6. Arrest of tumour cells within capillaries, where blood flow is slower and vessel walls are thinner;[73]
7. Extravasation, *i.e.* the passage of tumour cells from bloodstream into the surrounding tissue, is aided by platelets, which release adhesion molecules, and lymphocytes, which secrete chemokines and inflammatory mediators that increase vascular permeability;[75]
8. Proliferation of tumour cells in response to a favourable microenvironment rich in growth factors, nutrients and space, followed by angiogenesis to sustain further expansion;[76]
9. Formation of micrometastases, which are small clusters of malignant cells;[77]
10. Outgrowth of micrometastases into macroscopic tumours (the colonisation phase), which in turn may give rise additional metastatic lesions.[77]

One of the most significant strategies that enables tumour cells to invade surrounding tissues is the epithelial-mesenchymal transition (EMT). This process may be transient or persistent and induces profound changes in gene expression, including loss of adherens junctions, morphological conversion from a polygonal to a fibroblast-like shape, production of matrix-degrading enzymes (such as metalloproteinases), increased motility and enhanced resistance to apoptosis.[78] A hallmark of EMT is the downregulation of E-cadherin, a protein that is essential for cell-cell adhesion and is often associated with tumour progression.[79] Consequently, cells acquire a ‘slenderer’ and more mobile phenotype, enabling them to infiltrate the surrounding, already compromised tissues. The induction of EMT-related transcription factors has also been observed in non-epithelial tumours, such as sarcomas and neuroectodermal malignancies.[79]

As early as 1889, Paget proposed the famous ‘seed and soil’ theory in *The Lancet*. According to this concept, just as a seed, cells can only proliferate within certain tissues. This theory remains relevant today. Subsequent studies have clarified that the preferential localisation of metastases depends on two main factors: the microenvironment of the target organ, and the interactions between tumour cells and endothelial cells which dictate their selective adhesion to the capillary network of specific organs.[80]



6.8. Genomic Instability and Mutation

The acquisition of hallmark traits by cancer cells largely depends on genetic alterations to the tumour cell genome. Certain mutant phenotypes confer a selective advantage, enabling specific clones to grow more rapidly and colonise surrounding tissues. Therefore, tumour progression can be understood as a series of clonal expansions, each of which is initiated by the random emergence of a new mutant phenotype.[81] These alterations are not solely the result of DNA mutations but can also arise from epigenetic modifications such as DNA methylation or histone modifications. While these changes do not directly alter the nucleotide sequence, they can have a significant impact on the regulation of gene expression.[82]

To acquire the set of genes necessary for neoplastic transformation, tumour cells must accumulate DNA alterations. This occurs through increased sensitivity to mutagenic agents and defects in genomic maintenance mechanisms. The situation is further exacerbated by frequent impairment of DNA surveillance systems.[83] TP53 mutations are a classic example which prevent the protein from detecting DNA damage and triggering protective responses such as apoptosis or senescence.[84]

Loss of telomeric DNA also contributes to genomic instability by promoting chromosomal rearrangements, such as amplifications or deletions of entire segments. In this scenario, telomerase plays a dual role: it counteracts genomic instability by restoring telomeres, but it also grants tumour cells unlimited replicative potential.[85]

Consequently, cancer progression tends to select for cells with unstable genomes, as these are better able to accumulate mutations, thereby accelerating the transition from precancerous lesions to fully developed tumours.[86]

6.9. An Enabling Characteristic: Tumour-Promoting Inflammation

For a long time, it was thought that the presence of immune cells within tumours and their release of signalling molecules served only one purpose: to fight the tumour by treating it as a foreign entity.[87]

While this view is partly correct in that the immune system would indeed be capable of eliminating many transformed cells if not evaded, more recent studies have revealed a paradox: the immune response associated with tumours can also actively promote their development.[88]

Chronic inflammation contributes to both tumour initiation and malignant progression by creating a microenvironment that favours the acquisition of hallmark traits.[89]

The inflammatory microenvironment supports cancer growth by:

- Providing cytokines and survival factors that inhibit apoptosis;[90]



- Releasing enzymes (such as metalloproteinases) that remodel the extracellular matrix, thus facilitating angiogenesis, invasion, and metastasis;[91]
- Generating signals that promote EMT.[92]

This inflammation is most often chronic, arising from persistent viral infections or autoimmune diseases where immune cells remain continuously activated. Over time, the constant production of these mediators, which were originally intended to protect the host, ends up fuelling tumour formation. In addition, inflammatory cells can release reactive oxygen species (ROS), which are well known for their mutagenic potential and can accelerate the transition from normal to malignant cells.[93]

For these reasons, inflammation is not considered an intrinsic hallmark of cancer cells, but rather an enabling characteristic since it substantially contributes to the acquisition of malignant traits.[94]

6.10. Reprogramming of Energy Metabolism

The chronic and often uncontrolled growth that characterises tumour cells involves alterations not only in proliferation, but also in energy metabolism to ensure a sufficient energy supply for cellular division and survival.[37]

Under normal conditions, healthy cells rely on two main strategies for producing energy: in the presence of oxygen, glucose is metabolised through glycolysis, the Krebs cycle and oxidative phosphorylation, resulting in a high yield of ATP; in the absence of oxygen, glycolysis predominates, generating much less energy and producing pyruvate, only a small amount of which enters the mitochondria.[95]

In the 1920s, Otto Warburg first described an unusual feature of cancer cell metabolism: even in the presence of oxygen, tumour cells tend to favour glycolysis. This phenomenon is known as ‘aerobic glycolysis’.[96] Although this pathway is energetically less efficient, tumour cells overcome this disadvantage by dramatically increasing glucose uptake, often through the overexpression of GLUT1 transporters. This metabolic adaptation has been documented in many types of human tumours, where increased glucose consumption can be detected non-invasively using positron emission tomography (PET) with radiolabelled glucose analogue 18F-fluorodeoxyglucose.[96]

Reliance on glycolysis becomes even more pronounced under hypoxic conditions, which are common in the tumour microenvironment. Hypoxia-response systems act pleiotropically to upregulate both glucose transporters and several glycolytic enzymes, thereby supporting tumour growth even when oxygen is scarce.[97]



6.11. Evading Immune Destruction

Another crucial factor in tumour progression is interaction with the immune system. Under normal circumstances, immune cells can detect and eliminate abnormal cells, including cancer cells.[98] Therefore, in order to grow and survive, tumours must develop strategies to evade immune surveillance.[59] Clinical evidence supports this concept: immunocompromised patients, such as those with HIV, have a higher incidence of cancer. Specifically, a deficiency in cytotoxic CD8⁺ T lymphocytes, T helper cells or natural killer (NK) cells significantly increases the risk of tumour development. These immune cells play a vital role in recognising tumour-associated antigens and initiating targeted cytotoxic responses.[59]

However, during tumour evolution, highly immunogenic clones are rapidly eliminated while less immunogenic populations are positively selected, enabling them to proliferate. Furthermore, tumours can actively suppress immune activity by releasing immunosuppressive factors, such as TGFβ, which paralyses infiltrating T cells and NK cells, thereby weakening the antitumour response.[99]

The ability to evade immune elimination is a significant obstacle that cancer must overcome during its progression.[64]

6.12. Tumour Classification

Tumours can be classified according to various criteria, including site of origin, tissue type, differentiation grade and TNM system.[100]

- **By site of origin:** tumours are identified based on the organ or tissue in which they arise. Examples include breast, lung, bladder and liver cancer.[100]
- **By tissue type:** this classification requires a biopsy and provides information about the nature of tumour cells.[101] The main type include:
 1. *Carcinoma:* originates from epithelial cells and is the most common tumour, representing about 80-90% of all cancers.[102] Main subtypes are:
 - Adenocarcinoma, which develops in organs or glands (e.g. lung or breast carcinoma);[102]
 - Squamous cell carcinoma, which originates from squamous epithelium.[102]
 2. *Sarcoma:* originates from connective tissue, muscle, bone or cartilage. Examples include osteosarcoma (bones) and chondrosarcoma (cartilage).[103]
 3. *Myeloma:* arises from mutations in plasma cells which produce antibodies.[104]



4. *Leukaemia*: results from the uncontrolled proliferation of undifferentiated stem cells in the bone marrow. Examples include acute or chronic myeloid leukaemia and acute or chronic lymphocytic leukaemia.[104]
5. *Lymphoma*: involves lymph nodes or lymphatic glands and are solid tumours, unlike myeloma and leukaemia.[105] Subtypes include:
 - Hodgkin lymphoma (from B lymphocytes);[105]
 - Non-Hodgkin lymphoma (from B or T lymphocytes).[105]
- **By grade**: tumour cells are classified according to their level of differentiation. The less they resemble normal cells of the tissue of origin, the higher the grade, numbered from 1 to 4.[106]
- **TNM classification**: widely used in clinical practice for solid tumours, it helps provide initial diagnosis, prognosis, and treatment selection.[107] It is based on three parameters:
 1. T (tumour): indicates tumour size, from 1 (small) to 4 (large).[107]
 2. N (nodes): indicates involvement of nearby lymph nodes; from 0 (no nodes involved) to 3 (multiple nodes involved).[107]
 3. M (metastasis): indicates the presence or absence of metastases. The number 1 is used if at least one metastasis is present, or 0 if no metastases are detected.[107]

7. Causes of Cancer

Cancer does not result from a single cause, but from a combination of risk factors which may sometimes be absent. These risk factors are classified as either exogenous (modifiable or avoidable, such as lifestyle and exposure to toxic substances) or endogenous (intrinsic to the body and unavoidable).[108]

7.1. Exogenous risk factors

Exogenous risk factors include:

- **Cigarette smoking.** Smoking contains over 8000 chemical compounds, including carcinogens such as nitrosamines and aromatic hydrocarbons, and is a major risk factor. These compounds can damage DNA, promoting tumour formation.[109]
- **Alcohol consumption.** Ethanol is metabolised into acetaldehyde, which damages DNA. Excessive alcohol consumption increases the risk of cancers of the digestive tract, liver, colon and breast. The combination of alcohol and smoking has a synergistic effect, particularly in relation to cancers of the mouth and throat.[109]



- **Diet and nutrition.** Certain foods, such as red and processed meats, increase the risk of cancer, whereas fruits, vegetables, vitamins and fibre have a protective effect.[110]
- **Sedentary lifestyle and being overweight.** A lack of physical activity can promote insulin-dependent tumours (in the colon, breast and uterus) and increase systemic inflammation.[111]
- **Sun exposure and UV radiation.** Excessive exposure to UVA and UVB rays damages DNA, promotes mutations and oxidative stress, and contributes to the developments of melanomas and skin carcinomas.[111]

These factors can often act in combination, amplifying the risk of cellular transformation and cancer development.

7.1.1. Exogenous risk factors: Environmental

- **Ionising radiation.** X-rays, gamma rays, alpha and beta particles, and high-energy UV can damage DNA by causing strand breaks and mutations. The organs most sensitive to this damage are the thyroid and bone marrow, while solid tumours may take decades to develop after exposure.[111]
- **Air pollution.** Exposure to fine particulate matter (PM10 and PM2.5) has been linked to an increased risk of lung cancer; outdoor particulate matter is classified as a confirmed human carcinogen by the IARC.[112]
- **Chemical agents.** Substances such as benzene, dioxins, pesticides, vinyl chloride, polycyclic aromatic hydrocarbons and asbestos can cause mutations or chronic inflammation, increasing the risk of leukaemia, skin cancers, mesothelioma, and other neoplasms.[113]
- **Hormones.** hormonal imbalances due to obesity medication or pathological conditions can promote hormone-dependent cancers, such as breast, prostate and cervical cancers.[114]

7.1.2. Endogenous risk factors

Endogenous risk factors include:

- **Age.** Aging increases the accumulation of mutations and cancer risk, with most diagnoses occurring in older age.[115]
- **Genetic predisposition.** Inherited mutations, such as those in the BRCA genes, increase susceptibility to specific cancers (e.g. breast, ovarian, and prostate).[116]



- **Infectious agents.** Viruses such as HPV, hepatitis B and C, and Epstein-Barr, as well as bacteria such as *Helicobacter pylori*, can insert genetic material into cells or cause chronic inflammation, promoting cancerous transformations.[117]

These factors often act together, further increasing the risk of cancer.



8. Aim of the Thesis

The aim of this thesis is to advance the field of cancer-oriented molecular therapeutics by taking a multidisciplinary approach that combines synthetic nucleotide chemistry, peptide engineering, structural biology and high-resolution biophysical analysis. Combining these complementary approaches enables to address three fundamental biological dimensions tightly connected to tumour development, progression, and therapeutic response (Figure 6). Each section of the thesis offers a distinct perspective – chemical, structural or functional – to contribute to the goal of designing and validating novel molecular tools capable of modulating pathological pathways in oncology.

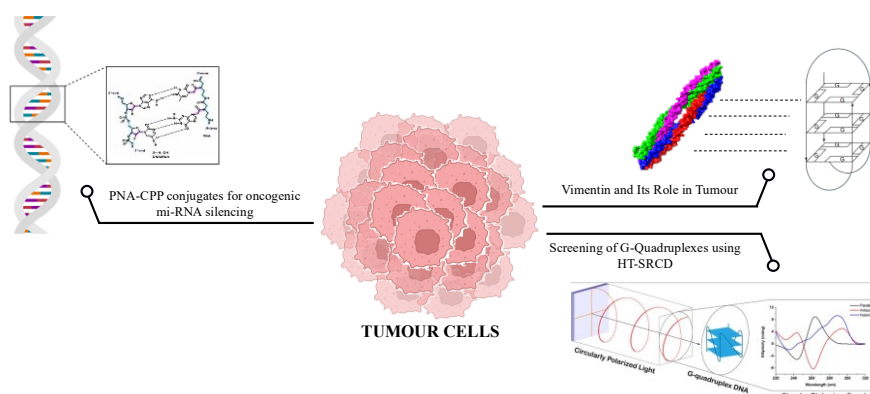


Figure 6. Schematic representation of the aims of the thesis.

The first research objective is to focus on the chemical synthesis and biological evaluation of Peptide Nucleic Acid (PNA)-Cell Penetrating Peptide (CPP) conjugates as next-generation antisense tools against oncogenic microRNAs. The microRNA-221-5p has been selected as a paradigmatic target due to its well-established role in the proliferation of cells, evasion of apoptosis and drug resistance in numerous solid tumours.[118] Although PNAs offer unparalleled hybridisation specificity and enzymatic stability, their therapeutic applicability is limited by poor intracellular transport and a tendency to aggregate in physiological media. This thesis therefore explores two orthogonal conjugation methodologies – CuAAC click chemistry and thiol-disulfide exchange – to covalently integrate PNAs with selected CPPs. A significant part of the work addresses the synthetic complexity of these conjugates, optimisation of coupling efficiencies, avoidance of aggregation-prone intermediates and rigorous structural verification. Furthermore, biological assays conducted on cells examine internalisation mechanisms and the early functional impact on cell viability under different conditions, thereby establishing the basis for the future development of a new conjugate as a next-generation delivery platform.



The second objective is to investigate the structure-function relationship of peptide fragments derived from Vimentin, a hallmark of EMT and a key driver of metastatic dissemination. Vimentin is characterized by extensive intrinsic disorder interspersed with structured coiled-coil domain, which enable dynamic interactions with nucleic acids, cytoskeletal partners and regulatory proteins. Despite its centrality in cancer biology, the molecular determinants underlying Vimentin's recognition of G-quadruplexes (G4s),[119] remain poorly understood. To address this, the thesis employs an extensive synthetic campaign incorporating alanine scanning, to investigate the role of aromatic and cationic residues in structural propensity in G4 binding. These studies provide an insight into how intrinsically disordered protein regions may participate in nucleic acid binding.

The third objective broadens the scope of the thesis to include the systematic screening characterisation of small-molecule and peptide ligands of G4 structures. Given the emerging role of G4s in transcriptional regulation, telomere maintenance, genome stability and oncogene expression, identifying molecules that can selectively bind to or remodel these structures represents a promising therapeutic approach. Using high-throughput synchrotron radiation circular dichroism (HT-SRCD),[120] the interaction of G4s with tetrazole derivatives and G4-binding peptide Rhau25 was monitored with high sensitivity. This study highlights that G4 topology, and therefore ligand accessibility, is significantly influenced by the ionic microenvironment, revealing the conditions under which selective ligand recognition or topology switching occurs.

This work advances a unified framework spanning molecular design, synthetic development, structural determination and functional validation. By integrating various methodologies, such as organic synthesis, peptide chemistry, nucleic acids biophysics and high-resolution spectroscopic screening, the work wants to provide new insight into how biomolecules recognise and modulate each other in oncogenic contexts. Furthermore, the approaches developed here lay the groundwork for the synthesis of selective modulators of RNA, DNA and protein targets implicated in cancer progression, paving the way for future therapeutic and diagnostic innovations.

References

References

1. Brown, J.S., et al., *Updating the Definition of Cancer*. Molecular Cancer Research, 2023. **21**(11): p. 1142-1147.
2. Liu, X. and P. Raju, *5.42 - In Vitro Cancer Model for Drug Testing*, in *Comprehensive Biotechnology (Second Edition)*, M. Moo-Young, Editor. 2011, Academic Press: Burlington. p. 543-549.
3. Begum, H., et al., *International Cancer Burden Analysis 2020-2024: GLOBOCAN-Derived Estimates of Incidence and Mortality for 30 Malignancies in 190 Geographic Regions*. American Journal of BioMedicine, 2025. **13**(1): p. 1-18.
4. Bray, F., et al., *Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries*. 2024. **74**(3): p. 229-263.
5. Marmot, M., et al., *WHO European review of social determinants of health and the health divide*. Lancet, 2012. **380**(9846): p. 1011-29.
6. Torre, L.A., et al., *Global Cancer Incidence and Mortality Rates and Trends--An Update*. Cancer Epidemiol Biomarkers Prev, 2016. **25**(1): p. 16-27.
7. Anyigba, C.A., G.A. Awandare, and L. Paemka, *Breast cancer in sub-Saharan Africa: The current state and uncertain future*. Exp Biol Med (Maywood), 2021. **246**(12): p. 1377-1387.
8. Lorusso, S., et al., *Italian health data system: current data interconnection and digital health ecosystem evolution*. The Lancet Regional Health – Europe, 2025. **51**.
9. Heidari, A., *The Progress of Cancer Survival Information in Europe: A Review*. THE PROGRESS: A Journal of Multidisciplinary Studies, 2025. **6**(3): p. 51-83.
10. Stubbs, S. and N. Thomas, *[1] - Dynamic Green Fluorescent Protein Sensors for High-Content Analysis of the Cell Cycle*, in *Methods in Enzymology*, J. Inglese, Editor. 2006, Academic Press. p. 1-21.
11. Shi, W. and S.F. Dowdy, *Cell Cycle Controls in G1 and G0*, in *Encyclopedia of Biological Chemistry*, W.J. Lennarz and M.D. Lane, Editors. 2004, Elsevier: New York. p. 328-331.
12. Pelley, J.W., *15 - Organization, Synthesis, and Repair of DNA*, in *Elsevier's Integrated Biochemistry*, J.W. Pelley, Editor. 2007, Mosby: Philadelphia. p. 123-133.
13. Limas, J.C. and J.G. Cook, *Preparation for DNA replication: the key to a successful S phase*. FEBS Lett, 2019. **593**(20): p. 2853-2867.
14. Davy, C. and J. Doorbar, *G2/M cell cycle arrest in the life cycle of viruses*. Virology, 2007. **368**(2): p. 219-226.
15. Yang, V.W., *Chapter 15 - The Cell Cycle*, in *Physiology of the Gastrointestinal Tract (Fifth Edition)*, L.R. Johnson, et al., Editors. 2012, Academic Press: Boston. p. 451-471.
16. Weitzman, M.D. and J.Y.J. Wang, *Cell Cycle: DNA Damage Checkpoints*, in *Encyclopedia of Biological Chemistry (Second Edition)*, W.J. Lennarz and M.D. Lane, Editors. 2013, Academic Press: Waltham. p. 410-416.
17. Fong, B.C., et al., *The Rb/E2F axis is a key regulator of the molecular signatures instructing the quiescent and activated adult neural stem cell state*. Cell Rep, 2022. **41**(5): p. 111578.

18. Zámboorszky, J., *Cell Cycle Transitions, G2/M*, in *Encyclopedia of Systems Biology*, W. Dubitzky, et al., Editors. 2013, Springer New York: New York, NY. p. 329-333.
19. McAinsh, A.D. and G.J.P.L. Kops, *Principles and dynamics of spindle assembly checkpoint signalling*. Nature Reviews Molecular Cell Biology, 2023. **24**(8): p. 543-559.
20. Hervé, S., et al., *Chromosome mis-segregation triggers cell cycle arrest through a mechanosensitive nuclear envelope checkpoint*. Nat Cell Biol, 2025. **27**(1): p. 73-86.
21. Duronio, R.J. and Y. Xiong, *Signaling pathways that control cell proliferation*. Cold Spring Harb Perspect Biol, 2013. **5**(3): p. a008904.
22. Porta, C., C. Paglino, and A. Mosca, *Targeting PI3K/Akt/mTOR Signaling in Cancer*. Front Oncol, 2014. **4**: p. 64.
23. Follo, M.Y., et al., *PLC and PI3K/Akt/mTOR signalling in disease and cancer*. Adv Biol Regul, 2015. **57**: p. 10-6.
24. Li, X., et al., *PI3K/Akt/mTOR signaling pathway and targeted therapy for glioblastoma*. Oncotarget, 2016. **7**(22): p. 33440-50.
25. Bahar, M.E., H.J. Kim, and D.R. Kim, *Targeting the RAS/RAF/MAPK pathway for cancer therapy: from mechanism to clinical studies*. Signal Transduction and Targeted Therapy, 2023. **8**(1): p. 455.
26. Lin, C.W., et al., *A two-component protein condensate of the EGFR cytoplasmic tail and Grb2 regulates Ras activation by SOS at the membrane*. Proc Natl Acad Sci U S A, 2022. **119**(19): p. e2122531119.
27. Parrish, A.B., C.D. Freel, and S. Kornbluth, *Cellular mechanisms controlling caspase activation and function*. Cold Spring Harb Perspect Biol, 2013. **5**(6).
28. Dakal, T.C., et al., *Oncogenes and tumor suppressor genes: functions and roles in cancers*. MedComm (2020), 2024. **5**(6): p. e582.
29. Kiraz, Y., et al., *Major apoptotic mechanisms and genes involved in apoptosis*. Tumour Biol, 2016. **37**(7): p. 8471-86.
30. Arvelo, F., F. Sojo, and C. Cotte, *Tumour progression and metastasis*. Ecancermedicalscience, 2016. **10**: p. 617.
31. Newkirk, K.M., E.M. Brannick, and D.F. Kusewitt, *Chapter 6 - Neoplasia and Tumor Biology1*, in *Pathologic Basis of Veterinary Disease (Sixth Edition)*, J.F. Zachary, Editor. 2017, Mosby. p. 286-321.e1.
32. Brown, G., *Oncogenes, Proto-Oncogenes, and Lineage Restriction of Cancer Stem Cells*. Int J Mol Sci, 2021. **22**(18).
33. Singh, S.R., et al., *Exploring the Genetic Orchestra of Cancer: The Interplay Between Oncogenes and Tumor-Suppressor Genes*. 2025. **17**(7): p. 1082.
34. Zhang, H., et al., *Unraveling the Guardian: p53's Multifaceted Role in the DNA Damage Response and Tumor Treatment Strategies*. Int J Mol Sci, 2024. **25**(23).
35. Glaviano, A., et al., *Cell cycle dysregulation in cancer*. Pharmacol Rev, 2025. **77**(2): p. 100030.
36. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer*. Cell, 2000. **100**(1): p. 57-70.

References

37. Feitelson, M.A., et al., *Sustained proliferation in cancer: Mechanisms and novel therapeutic targets*. Semin Cancer Biol, 2015. **35 Suppl**(Suppl): p. S25-s54.
38. Li, Y., et al., *Invasion and metastasis in cancer: molecular insights and therapeutic targets*. Signal Transduction and Targeted Therapy, 2025. **10**(1): p. 57.
39. DeBerardinis, R.J. and N.S. Chandel, *Fundamentals of cancer metabolism*. Sci Adv, 2016. **2**(5): p. e1600200.
40. Liu, X., J. Yan, and M.W. Kirschner, *Cell size homeostasis is tightly controlled throughout the cell cycle*. PLoS Biol, 2024. **22**(1): p. e3002453.
41. Tito, C., et al., *EGF receptor in organ development, tissue homeostasis and regeneration*. J Biomed Sci, 2025. **32**(1): p. 24.
42. Iqbal, N. and N. Iqbal, *Human Epidermal Growth Factor Receptor 2 (HER2) in Cancers: Overexpression and Therapeutic Implications*. Mol Biol Int, 2014. **2014**: p. 852748.
43. Saha, S.K. and A.R. Khuda-Bukhsh, *Molecular approaches towards development of purified natural products and their structurally known derivatives as efficient anti-cancer drugs: Current trends*. European Journal of Pharmacology, 2013. **714**(1): p. 239-248.
44. Yang, X. and H. Wu, *RAS signaling in carcinogenesis, cancer therapy and resistance mechanisms*. Journal of Hematology & Oncology, 2024. **17**(1): p. 108.
45. Zhu, J. and C.B. Thompson, *Metabolic regulation of cell growth and proliferation*. Nat Rev Mol Cell Biol, 2019. **20**(7): p. 436-450.
46. Marescal, O. and I.M. Cheeseman, *Cellular Mechanisms and Regulation of Quiescence*. Dev Cell, 2020. **55**(3): p. 259-271.
47. Galassi, C., et al., *The hallmarks of cancer immune evasion*. Cancer Cell, 2024. **42**(11): p. 1825-1863.
48. Nisar, S., et al., *Exploring Dysregulated Signaling Pathways in Cancer*. Curr Pharm Des, 2020. **26**(4): p. 429-445.
49. Zhang, Y., P.B. Alexander, and X.F. Wang, *TGF- β Family Signaling in the Control of Cell Proliferation and Survival*. Cold Spring Harb Perspect Biol, 2017. **9**(4).
50. Cecerska-Heryć, E., et al., *TGF- β Signaling in Cancer: Mechanisms of Progression and Therapeutic Targets*. 2025. **26**(15): p. 7326.
51. Darwiche, N., *Epigenetic mechanisms and the hallmarks of cancer: an intimate affair*. Am J Cancer Res, 2020. **10**(7): p. 1954-1978.
52. Yuan, J. and D. Ofengeim, *A guide to cell death pathways*. Nature Reviews Molecular Cell Biology, 2024. **25**(5): p. 379-395.
53. Mustafa, M., et al., *Apoptosis: A Comprehensive Overview of Signaling Pathways, Morphological Changes, and Physiological Significance and Therapeutic Implications*. Cells, 2024. **13**(22).
54. Teng, X. and J.M. Hardwick, *Apoptosis: the extrinsic pathway*, in *Molecular Oncology: Causes of Cancer and Targets for Treatment*, E.P. Gelmann, C.L. Sawyers, and I.I.I.F.J. Rauscher, Editors. 2013, Cambridge University Press: Cambridge. p. 353-366.

55. Hotamisligil, G.S. and R.J. Davis, *Cell Signaling and Stress Responses*. Cold Spring Harb Perspect Biol, 2016. **8**(10).
56. Pfeffer, C.M. and A.T.K. Singh, *Apoptosis: A Target for Anticancer Therapy*. Int J Mol Sci, 2018. **19**(2).
57. Fernald, K. and M. Kurokawa, *Evading apoptosis in cancer*. Trends Cell Biol, 2013. **23**(12): p. 620-33.
58. Chaudhary, B., et al., *Targeting anti-apoptotic mechanisms in tumour cells: Strategies for enhancing Cancer therapy*. Bioorganic Chemistry, 2025. **159**: p. 108388.
59. Kim, S.K. and S.W. Cho, *The Evasion Mechanisms of Cancer Immunity and Drug Intervention in the Tumor Microenvironment*. Front Pharmacol, 2022. **13**: p. 868695.
60. Sumida, T.S., N.T. Cheru, and D.A. Hafler, *The regulation and differentiation of regulatory T cells and their dysfunction in autoimmune diseases*. Nature Reviews Immunology, 2024. **24**(7): p. 503-517.
61. Matthews, H.K., C. Bertoli, and R.A.M. de Bruin, *Cell cycle control in cancer*. Nature Reviews Molecular Cell Biology, 2022. **23**(1): p. 74-88.
62. Vlashi, R., et al., *The molecular biology of NF2/Merlin on tumorigenesis and development*. Faseb j, 2024. **38**(13): p. e23809.
63. Gillooly, J.F., et al., *Explaining differences in the lifespan and replicative capacity of cells: a general model and comparative analysis of vertebrates*. Proc Biol Sci, 2012. **279**(1744): p. 3976-80.
64. Tufail, M., C.-H. Jiang, and N. Li, *Immune evasion in cancer: mechanisms and cutting-edge therapeutic approaches*. Signal Transduction and Targeted Therapy, 2025. **10**(1): p. 227.
65. Victorelli, S. and J.F. Passos, *Telomeres and Cell Senescence - Size Matters Not*. EBioMedicine, 2017. **21**: p. 14-20.
66. de Bardet, J.C., et al., *Cell immortalization: In Vivo Molecular Bases and In Vitro Techniques for Obtention*. BioTech (Basel), 2023. **12**(1).
67. Wang, X., et al., *Distinct roles of telomerase activity in age-related chronic diseases: An update literature review*. Biomedicine & Pharmacotherapy, 2023. **167**: p. 115553.
68. Jiang, X., et al., *The role of microenvironment in tumor angiogenesis*. Journal of Experimental & Clinical Cancer Research, 2020. **39**(1): p. 204.
69. DiPietro, L.A., *Angiogenesis and wound repair: when enough is enough*. J Leukoc Biol, 2016. **100**(5): p. 979-984.
70. Kazerounian, S. and J. Lawler, *Integration of pro- and anti-angiogenic signals by endothelial cells*. J Cell Commun Signal, 2018. **12**(1): p. 171-179.
71. Ribatti, D. and F. Pezzella, *Overview on the Different Patterns of Tumor Vascularization*. Cells, 2021. **10**(3).
72. Sobczyńska-Rak, A., et al., *Role and Mechanisms of Angiogenesis in Tumours*. 2025. **14**(7): p. 756.
73. Rajput, S., P. Kumar Sharma, and R. Malviya, *Fluid mechanics in circulating tumour cells: Role in metastasis and treatment strategies*. Medicine in Drug Discovery, 2023. **18**: p. 100158.

References

74. Gan, J., X. Zhang, and J. Guo, *The role of platelets in tumor immune evasion and metastasis: mechanisms and therapeutic implications*. Cancer Cell International, 2025. **25**(1): p. 258.
75. Dupas, A., J.G. Goetz, and N. Osmani, *Extravasation of immune and tumor cells from an endothelial perspective*. J Cell Sci, 2024. **137**(21).
76. Xiao, Z. and E. Puré, *The fibroinflammatory response in cancer*. Nature Reviews Cancer, 2025. **25**(6): p. 399-425.
77. Sun, J., et al., *Lymph node micrometastasis in non-small cell lung cancer*. Biomedicine & Pharmacotherapy, 2022. **149**: p. 112817.
78. Dong, A. and C. Blanpain, *Identification, functional insights and therapeutic targeting of EMT tumour states*. Nature Reviews Cancer, 2025.
79. Aban, C.E., et al., *Downregulation of E-cadherin in pluripotent stem cells triggers partial EMT*. Sci Rep, 2021. **11**(1): p. 2048.
80. Akhtar, M., et al., *Paget's "Seed and Soil" Theory of Cancer Metastasis: An Idea Whose Time has Come*. Adv Anat Pathol, 2019. **26**(1): p. 69-74.
81. Hanahan, D. and Robert A. Weinberg, *Hallmarks of Cancer: The Next Generation*. Cell, 2011. **144**(5): p. 646-674.
82. Liyanage, V.R., et al., *DNA modifications: function and applications in normal and disease States*. Biology (Basel), 2014. **3**(4): p. 670-723.
83. Casalino, L. and P. Verde, *Multifaceted Roles of DNA Methylation in Neoplastic Transformation, from Tumor Suppressors to EMT and Metastasis*. Genes (Basel), 2020. **11**(8).
84. Wang, H., et al., *Targeting p53 pathways: mechanisms, structures and advances in therapy*. Signal Transduction and Targeted Therapy, 2023. **8**(1): p. 92.
85. Cao, T., et al., *Insufficient telomeric DNA damage response promotes chromosomal instability in aged oocytes*. Sci Bull (Beijing), 2025.
86. Zhang, D., S.-Y. Lin, and M. Lee, *Editorial: Exploring genomic instability of cancers: applications in diagnosis and treatment*. 2025. **Volume 13 - 2025**.
87. Peña-Romero, A.C. and E. Orenes-Piñero, *Dual Effect of Immune Cells within Tumour Microenvironment: Pro- and Anti-Tumour Effects and Their Triggers*. Cancers (Basel), 2022. **14**(7).
88. Zach, M. and G.P. Greslehner, *Understanding immunity: an alternative framework beyond defense and strength*. Biol Philos, 2023. **38**(1): p. 7.
89. Yang, L., et al., *Intratumoral microbiota: roles in cancer initiation, development and therapeutic efficacy*. Signal Transduction and Targeted Therapy, 2023. **8**(1): p. 35.
90. Nishida, A. and A. Andoh, *The Role of Inflammation in Cancer: Mechanisms of Tumor Initiation, Progression, and Metastasis*. 2025. **14**(7): p. 488.
91. Cabral-Pacheco, G.A., et al., *The Roles of Matrix Metalloproteinases and Their Inhibitors in Human Diseases*. Int J Mol Sci, 2020. **21**(24).
92. Allgayer, H., et al., *Epithelial-to-mesenchymal transition (EMT) and cancer metastasis: the status quo of methods and experimental models* 2025. Molecular Cancer, 2025. **24**(1): p. 167.

93. Yacine, A., et al., *Chronic Inflammation: A Multidisciplinary Analysis of Shared Pathways in Autoimmune, Infectious, and Degenerative Diseases*. Cureus, 2025. **17**(4): p. e82579.
 94. Hibino, S., et al., *Inflammation-Induced Tumorigenesis and Metastasis*. 2021. **22**(11): p. 5421.
 95. Flood, D., E.S. Lee, and C.T. Taylor, *Intracellular energy production and distribution in hypoxia*. J Biol Chem, 2023. **299**(9): p. 105103.
 96. Bononi, G., et al., *Historical perspective of tumor glycolysis: A century with Otto Warburg*. Seminars in Cancer Biology, 2022. **86**: p. 325-333.
 97. Kierans, S.J. and C.T. Taylor, *Regulation of glycolysis by the hypoxia-inducible factor (HIF): implications for cellular physiology*. J Physiol, 2021. **599**(1): p. 23-37.
 98. Gonzalez, H., C. Hagerling, and Z. Werb, *Roles of the immune system in cancer: from tumor initiation to metastatic progression*. Genes Dev, 2018. **32**(19-20): p. 1267-1284.
 99. Roerden, M. and S. Spranger, *Cancer immune evasion, immunoediting and intratumour heterogeneity*. Nature Reviews Immunology, 2025. **25**(5): p. 353-369.
 100. Carbone, A., *Cancer Classification at the Crossroads*. 2020. **12**(4): p. 980.
 101. Kumar, R., R. Srivastava, and S. Srivastava, *Detection and Classification of Cancer from Microscopic Biopsy Images Using Clinically Significant and Biologically Interpretable Features*. J Med Eng, 2015. **2015**: p. 457906.
 102. Yan, W., et al., *Squamous Cell Carcinoma - Similarities and Differences among Anatomical Sites*. Am J Cancer Res, 2011. **1**(3): p. 275-300.
 103. Hall, F., V. Villalobos, and B. Wilky, *Future directions in soft tissue sarcoma treatment*. Current Problems in Cancer, 2019. **43**(4): p. 300-307.
 104. Abduh, M.S., *An overview of multiple myeloma: A monoclonal plasma cell malignancy's diagnosis, management, and treatment modalities*. Saudi J Biol Sci, 2024. **31**(2): p. 103920.
 105. Campo, E., G. Frigola, and S.A. Pileri, *The classification of lymphomas*, in *Hoffbrand's Postgraduate Haematology*. 2025. p. 361-394.
 106. Waldum, H.L., et al., *Classification of tumours*. J Exp Clin Cancer Res, 2008. **27**(1): p. 70.
 107. Rosen, R.D. and A. Sapra, *TNM Classification*, in *StatPearls*. 2025, StatPearls Publishing
- Copyright © 2025, StatPearls Publishing LLC.: Treasure Island (FL) ineligible companies. Disclosure: Amit Sapra declares no relevant financial relationships with ineligible companies.
108. Wu, S., et al., *Evaluating intrinsic and non-intrinsic cancer risk factors*. Nature Communications, 2018. **9**(1): p. 3490.
 109. Li, Y. and S.S. Hecht, *Carcinogenic components of tobacco and tobacco smoke: A 2022 update*. Food Chem Toxicol, 2022. **165**: p. 113179.
 110. Qian, F., et al., *Red and Processed Meats and Health Risks: How Strong Is the Evidence?* Diabetes Care, 2020. **43**(2): p. 265-271.
 111. Goyal, J. and G. Rakhra, *Sedentarism and Chronic Health Problems*. Korean J Fam Med, 2024. **45**(5): p. 239-257.

References

112. Orellano, P., et al., *Short-term exposure to particulate matter (PM10 and PM2.5), nitrogen dioxide (NO2), and ozone (O3) and all-cause and cause-specific mortality: Systematic review and meta-analysis*. Environment International, 2020. **142**: p. 105876.
113. Kharrazian, D., *Exposure to Environmental Toxins and Autoimmune Conditions*. Integr Med (Encinitas), 2021. **20**(2): p. 20-24.
114. Ding, S., C.O. Madu, and Y. Lu, *The Impact of Hormonal Imbalances Associated with Obesity on the Incidence of Endometrial Cancer in Postmenopausal Women*. J Cancer, 2020. **11**(18): p. 5456-5465.
115. White, M.C., et al., *Age and cancer risk: a potentially modifiable relationship*. Am J Prev Med, 2014. **46**(3 Suppl 1): p. S7-15.
116. Pal, M., D. Das, and M. Pandey, *Understanding genetic variations associated with familial breast cancer*. World J Surg Oncol, 2024. **22**(1): p. 271.
117. Ofoezie, E.F., et al., *Role of infectious agents in cancer pathogenesis and therapy*. The Microbe, 2025. **6**: p. 100284.
118. Di Martino, M.T., et al., *miR-221/222 as biomarkers and targets for therapeutic intervention on cancer and other diseases: A systematic review*. Mol Ther Nucleic Acids, 2022. **27**: p. 1191-1224.
119. Ceschi, S., et al., *Vimentin binds to G-quadruplex repeats found at telomeres and gene promoters*. Nucleic Acids Res, 2022. **50**(3): p. 1370-1381.
120. Hussain, R., et al., *High-throughput SRCD using multi-well plates and its applications*. Scientific Reports, 2016. **6**(1): p. 38028.

CHAPTER I

**microRNA-221 as a Therapeutic Target: Molecular Basis, Delivery
Approaches, and Inhibition Strategies**



Introduction

MicroRNAs (miRNAs) have emerged as key post-transcriptional regulators of gene expression and are now recognised as important factors in the development and progression of various human cancer. Among them, miR-221 has attracted particular attention due to its strong oncogenic activity and recurrent overexpression in various solid tumours, including lung cancer. Persistent deregulation of miR-221 contributes to uncontrolled proliferation, metastatic dissemination, and resistance to therapy, making it a promising therapeutic target. This chapter focuses on the molecular rationale for targeting miR-221-5p and introduces the design principles and delivery challenges associated with microRNA-directed therapeutic strategies.

1.1. History

Within the context of the traditional central dogma of molecular biology, RNA molecules were initially considered to be merely transient intermediates in the flow of genetic information, connecting DNA and protein synthesis.[1] However, this linear perspective was revised following the discovery of RNA interference (RNAi), which revealed that RNA molecules are actively involved in regulating gene expression.[2]

It is now well established that small RNA molecules can interact with specific other RNA species to modulate their activity, either repressing or enhancing their function.[3] A classic example is provided by small interfering RNAs (siRNAs), which bind to complementary messenger RNAs (mRNAs), thereby reducing their translation into proteins.[2]

A landmark advancement occurred in 1993 with the discovery of an endogenous small regulatory RNA in the nematode *Caenorhabditis elegans* (*C. elegans*). This provided the first evidence of a broad class of short (approximately 22 nucleotides), single-stranded ribonucleic acids, which were later termed microRNAs.[4] Since then, a large number of miRNAs have been identified in most eukaryotic organisms, and they are now recognised as fundamental post-transcriptional regulators of gene expression in complex biological systems.[5]

Given their pivotal role, it is not surprising that miRNAs themselves are subjected to strict regulatory control. Their finely tuned expression patterns enable precise temporal and tissue-specific shaping of gene expression, which is indispensable for orchestration of processes involved in development and growth of organisms.[6] To date, numerous studies have demonstrated that miRNAs are involved in a wide variety of cellular pathways. Their discovery has expanded our understanding of gene regulation and filled critical gaps in our knowledge of cellular networks, providing key insights into previously elusive aspects of biological regulation.[5]



1.2. Genomic organisation and classification of miRNAs

MicroRNAs are a conserved class of non-coding RNAs that range in length from 17 to 22 nucleotides. First discovered in the late 1990s, they have since emerged as key regulators of gene expression at the post-transcriptional level, playing an essential role in a variety of biological processes.[7]

According to the latest release of the miRNA database (v22), the human genome contains about 2300 mature miRNAs.[8] Each miRNA is assigned a prefix of 'miR' followed by a progressive number reflecting the order of discovery, and miRNAs that share an identical sequence in position 2-8 of the mature strand are grouped into the same family. Small nucleotide variations are denoted by a lowercase letter, as in the case of miR-125a and miR-125b. When the same mature miRNA originates from different genetic loci, a numerical suffix is added; for example, miR-125b-1 and miR-125b-2.[9] Each genetic locus can produce two mature miRNAs: one derived from the 5'→3' strand of the precursor and the other from the 3'→5' strand. These are designated with the suffixes '-5p' and '-3p', respectively. This updated nomenclature replaced the earlier use of the asterisk (*) to indicate the less abundant strand, since it has been demonstrated that so-called 'passenger' or secondary strand can also exert functional roles.[10]

From a genomic perspective, microRNA genes are most frequently located within intronic regions of protein-coding genes or long non-coding RNAs (lncRNAs), and less commonly within exonic regions.[11] In many species, multiple miRNA loci with related sequences have been identified, often originating from gene duplication events and forming clusters or polycistronic transcription units; miRNAs within the same cluster are generally co-transcribed, though they can undergo further post-transcriptional regulation.[12]

In contrast, intergenic miRNAs are transcribed as independent units, each with its own promoter. The precise location of these regulatory regions is not fully characterised for most miRNAs but can often be inferred from CpG island analysis. Finally, some miRNA genes are embedded within the introns of protein-coding genes and share the same promoter as their host gene, highlighting the remarkable complexity of miRNA regulation.[13]

1.3. Canonical miRNA biogenesis

The biogenesis of microRNAs is a multistep process that begins in the nucleus and continues in the cytoplasm (Figure 1.1).[14] miRNAs genes are primarily transcribed by RNA polymerase II, generating long primary transcripts known as pri-miRNAs. These molecules share the same feature as messenger RNAs, such as a 5' cap and a 3' poly(A) tail and are usually around 1000 nucleotides



long. Their secondary structure is characterised by one or more hairpin loops, which are essential for subsequent processing events.[15]

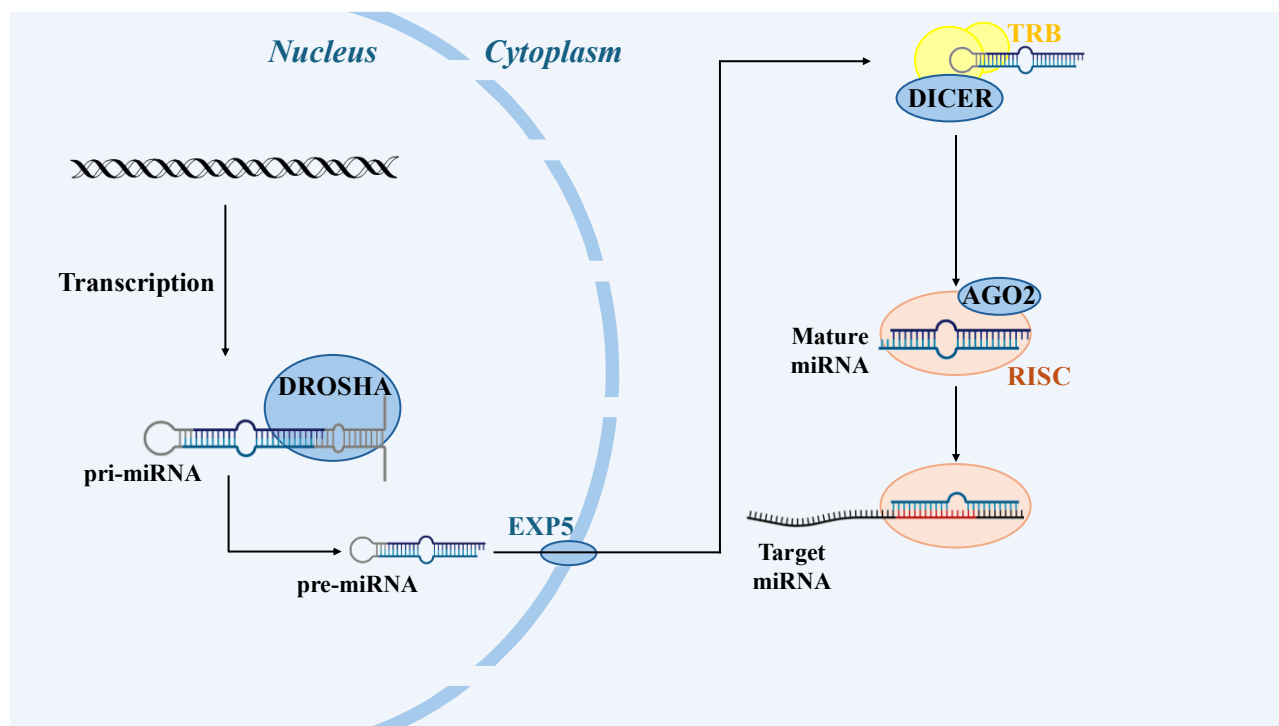


Figure 1.1. *miRNA biogenesis.*

In the nucleus, the pri-miRNA undergoes a first maturation step (cropping), which is mediated by the microprocessor complex. This complex consists of the RNase III endonuclease Drosha and its essential cofactor DGCR8 (also known as *Pasha* in *C. elegans* and *D. melanogaster*), which is a double-stranded RNA-binding protein that provides structural specificity.[16] DGCR8 acts as a molecular ruler, positioning Drosha to cleave approximately 11 base pairs away from the base of the hairpin. This produces a precursor miRNA (pre-miRNA) of 60-100 nucleotides with a characteristic 3' overhang.[17]

The pre-miRNA is then exported from the nucleus by Exportin-5 (Exp5) in complex with Ran GTP, which recognises the 3' overhang left by Drosha specifically. Once in the cytoplasm, a second processing step (dicing) is carried out by Dicer, another RNase III enzyme, in association with cofactors such as TRBP. Dicer cleaves near the terminal loop of the pre-miRNA hairpin, releasing a 20-22 nucleotide RNA duplex with 2 nucleotide 3' overhangs.[18] Dicer contains two RNase III catalytic domains, a PAZ domain that recognises the 5' phosphate of the pre-miRNA and a DUF283 domain that contributes to substrate orientation.[19]

At this stage, the miRNA/miRNA* duplex is loaded onto Argonaute (AGO) protein, which is the core component of the RNA-induced silencing complex (RISC).[20] Based on thermodynamic stability, one strand is selected as the guide strand and becomes the mature, functional miRNA, while



the other strand, historically known as the miRNA* or passenger strand, was initially thought to degrade. However, accumulating evidence shows that these passenger strands can also be functional and participate in gene regulation.[10] To reflect this, modern nomenclature distinguishes between the two strands, referring to them as 5p and 3p depending on whether they originate from the 5' or 3' arm of the precursor.[21]

Among the four human Argonaute proteins (AGO1-4), only AGO2 possesses catalytic activity. Once loaded with the guide miRNA, AGO forms the minimal silencing complex, often termed the miRISC, which includes additional cofactors such as GW182. The miRISC recognises target mRNAs primarily through imperfect base pairing within their 3' untranslated regions (3'UTRs). A critical determinant of target recognition is the seed region of the miRNA (nucleotides 2-7 from the 5' end).[22]

Gene silencing mediated by miRISC can proceed through multiple mechanisms. One major pathway involves mRNA degradation, initiated by deadenylation of the poly(A) tail via the CCR4-NOT and PAN2-PAN3 complexes, followed by decapping via DCP1/2 complex. The resulting transcript is then degraded by either 5'→3' exonuclease XRN1 or the 3'→5' exosome complex.[23] Alternatively, miRNAs can suppress translation without necessarily degrading the mRNA. This is achieved by either blocking the recruitment of the translation initiation complex – particularly the interaction between eIF4E and the 5' cap – or interfering with the elongation and termination steps through interactions with effector proteins.[24]

These mechanisms demonstrate that mature miRNAs, when incorporated into RISC, act as versatile post-transcriptional regulators. They fine-tune gene expression by guiding the complex to complementary sequences on target mRNA, which ultimately leads to translational inhibition or mRNA destabilisation. These processes are essential for development, differentiation and cellular homeostasis.[14]

1.4. Mechanism of Action and Function of Human miRNAs

Once microRNA processing is complete, the mature single-stranded miRNA is incorporated into the microRNA-RNA-binding protein (miRNA-RNP) complex, where it facilitates gene silencing by pairing with complementary sequences in the 3' untranslated region (3' UTR) of target mRNAs.[25] Crucially, only a short segment of the miRNA, known as the 'seed region' (nucleotides 2-8, occasionally extending to positions 9 or 10), is necessary for target recognition. This sequence plays a pivotal role in determining target specificity and function.[26] Interaction between the seed region and the 3' UTR leads to post-transcriptional repression of gene expression through three primary mechanisms:



1. Site-specific cleavage, which is exclusively mediated by AGO2 and requires perfect or near-perfect complementarity between the miRNA and its target. This process is well documented in plants, but rare in mammals.[27]
2. mRNA destabilisation and enhanced degradation, a widespread mechanism in mammals that often involves partial mismatches between the miRNA and its target sequence.[28]
3. Translational repression, which inhibits protein synthesis without necessarily reducing mRNA levels. Although this mechanism is common, the precise molecular pathway remains under debate.[29]

The latter two mechanisms are generally grouped under the term ‘non-cleavage repression’, since they do not involve the direct slicing of mRNA and can be mediated by any of the four Argonaute proteins (AGO1-4).[30]

In addition to canonical regulation, further layers of control influence microRNA-mediated gene silencing. For example, the secondary structures of the target mRNA may hinder access to the binding site, while RNA-binding proteins can mask complementary regions, preventing effective pairing with the miRNA seed sequence.[31]

Interestingly, emerging evidence suggests that miRNAs can bind to non-canonical sites, including the 5' UTR and coding regions (ORFs), of target mRNAs. Even more remarkably, under specific cellular conditions such as growth arrest, some miRNAs have been reported to enhance translation rather than repress it.[32]

The regulatory capacity of miRNAs is extensive: each miRNA can potentially target hundreds of different mRNAs, while individual mRNAs often contain binding sites for multiple distinct miRNAs. This combinatorial regulation enables fine-tuned, highly dynamic control of gene expression.[33] Current bioinformatic analyses suggests that up 60% of mammalian mRNAs harbour evolutionarily conserved miRNA-binding sites, highlighting the pervasive impact of miRNAs across gene networks.[34]

Consequently, miRNAs are now recognised as versatile and multifunctional regulators of post-transcriptional gene expression, exerting their influence through a broad spectrum of canonical and non-canonical mechanisms that contribute to physiological and pathological processes.[35]



1.5. microRNAs in Cancer: Oncogenes or Tumour Suppressor

Since their discovery in 1993, microRNAs have progressively been implicated in a wide range of cellular pathways and human disease. For example, the let-7 family of miRNAs has been extensively studied and found to play a role in conditions such as asthma, cardiac hypertrophy, inflammation, Parkinson's disease and various types of cancer (Figure 1.2).[36-38]

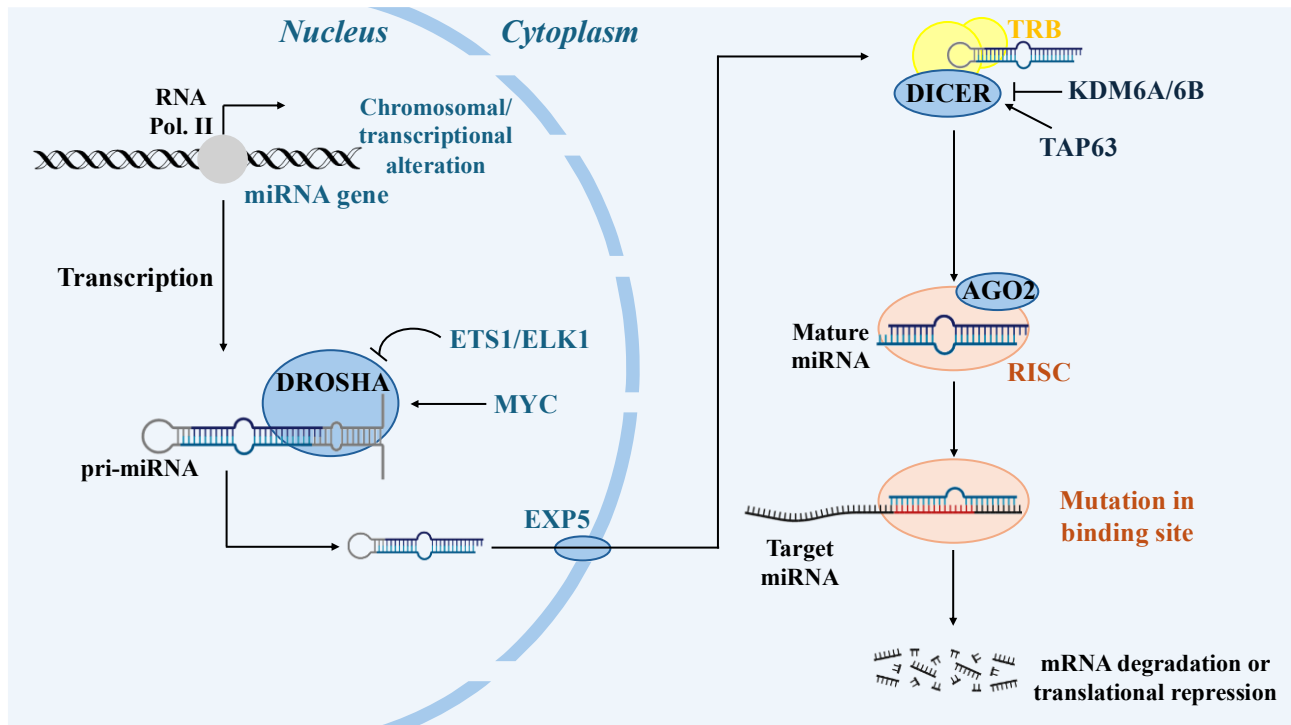


Figure 1.2. Alteration throughout microRNA biogenesis can affect availability of target mRNA.

Over time, cancer cells acquire molecular networks that sustain their growth and, in some cases, enable resistance to therapy. Within this framework, dysregulated miRNAs can act as either tumour suppressors or oncogenes (oncomiRs). Their oncogenic or suppressive activity depends on the target genes and pathways they regulate, which often involve one of the hallmarks of cancer.[39]

Functional studies using gain- and loss-of-function approaches have provided critical evidence of the roles of miRNAs in carcinogenesis.[40] For instance, tumour suppressor miRNAs have been demonstrated to impede cancer progression by targeting pivotal oncogenes[41]: miRNA-15a and miRNA-16-1 regulate BCL2,[42] let-7 target MYC and RAS,[43] and miRNA-29b modulates Mcl1.[44] Conversely, silencing studies using antisense oligonucleotides have revealed that oncogenic miRNAs, which are often overexpressed in cancer, inhibit apoptosis by targeting pro-apoptotic genes.[45] Examples include miRNA-221,[46] miRNA-222,[46] miRNA-21,[47] miRNA-24,[48] miRNA-133,[49] miRNA-17-5p,[50] and miRNA-19,[51] all of which impair programmed cell death when deregulated.



Some of the earliest *in vivo* demonstrations of miRNAs with oncogenic activity came from studies of the miRNA-17-92 cluster and miRNA-115, both of which are found to be overexpressed in lymphoproliferative disorders such as lymphomas and leukaemias.[52] The infection of murine haematopoietic stem cells with a retrovirus carrying the miRNA-17-92 cluster accelerated the development of lymphoma in Myc transgenic mice.[53] Similarly, the transgenic overexpression of the miRNA-17-92 cluster in B cells resulted in lymphoproliferative disease and autoimmunity. This was associated with increased lymphocyte proliferation and reduced apoptosis, which was partially due to the direct targeting of the pro-apoptotic genes BIM and PTEN.[54] Ventura and colleagues demonstrated that knockout of the miRNA-17-92 cluster resulted in neonatal mortality in mice due to lung hypoplasia and ventricular septal defects, whereas deletion of the cluster reduced Myc-induced oncogenesis.[55]

In contrast, overexpression of miRNA-155 in lymphoid tissues was sufficient to induce cancer in the absence of cooperative mutations or Myc expression. Croce *et al.* generated miRNA-155 transgenic mice that exhibited polyclonal lymphoid proliferation and eventually progressed to acute lymphoblastic lymphoma or leukaemia. This study provided one of the first direct demonstrations that the dysregulation of a single microRNA (miRNA) can drive malignant transformation.[56]

However, despite considerable advances in research, it is still unlikely that individual miRNAs will be found to drive a distinct cancer phenotype through a single target gene.[57] Instead, miRNAs are increasingly recognised as master regulators that engage in complex networks, simultaneously modulating multiple mRNA and interacting with transcriptional and post-transcriptional regulatory machinery.[35] This multi-targeting capacity is one of the most compelling reasons why miRNAs are being actively investigated as potential anticancer therapeutics.

1.6. microRNA-221 and Its Role in Cancer

miRNA221 is part of the miRNA221/222 cluster, which comprises two highly homologous miRNAs located on the Xp11.3 chromosome that are widely described as oncogenic regulators in multiple human cancers. These miRNAs are transcribed as a single primary transcript and processed into the mature miR-221-5p and miR-221-3p strands via canonical Drosha- and Dicer-dependent pathways.[58] Both strands derive from opposite arms of the same precursor hairpin, with the 5p form arising from the 5' arm and the 3p form from the 3' arm. While earlier studies emphasised the abundance of the 3p strand in certain tumour contexts, recent findings suggest that the 5p strand also plays an independent and biologically relevant regulatory role, particularly in solid tumours such as non-small-cell lung cancer (NSCLC).[59] Both isoforms can bind to complementary sequences in



target mRNAs, thereby repressing translation or promoting degradation and acting as post-transcriptional regulators of multiple oncogenic and tumour-suppressive pathways.[59]

According to Abak *et al.*, miRNA-221 is a key modulator of tumour progression in epithelial cancers. These miRNAs are often overexpressed and function as oncogenes by repressing crucial tumour suppressor genes, such as PTEN, CDKN1B (p27^(Kip)), CDKN1C (p57^(Kip2)) and SOCS3. This results in the activation of proliferative and anti-apoptotic signalling networks, such as PI3K/Akt and JAK/STAT. The review also highlights that elevated expression of this miRNA enhances angiogenesis, EMT and metastatic dissemination, thus confirming their critical contribution to malignant transformation.[58]

Building on this, Di Martino *et al.* shed light on the intricate post-transcriptional regulatory network involving miR-221-5p and miR-221-3p. Both microRNAs are regulated by a network of long non-coding RNAs (lncRNAs) acting as competing endogenous RNAs (ceRNAs), which sequester or release miR-221 depending on the cellular context. In NSCLC, several lncRNAs, including GAS5, TUG1 and ZNF281, directly modulate the activity of miR-221. GAS5 binds to miR-221-3p, thereby suppressing its oncogenic activity and restoring IRF2 expression, which reduces tumour cell proliferation, migration, and invasion. Similarly, TUG1 acts as a molecular sponge that inhibits miR-221, leading to the upregulation of PTEN expression, contributing to lung cancer progression and drug resistance. Together, these findings emphasise the dual influence of miR-221 strands and their ceRNA interactors in determining tumour aggressiveness and response to therapy.[46]

Further evidence supports the immunomodulatory role of miR-221-5p in cancer biology. The study showed that tumour-derived exosomes containing high levels of miR-221-5p can reduce the effectiveness of natural killer (NK) cells by targeting important immune-activating genes such as CD96 and perforin (PRF1). While this research was conducted in the context of bladder cancer, the identified mechanism is likely to be conserved across various epithelial tumours. This exosomal communication pathway is consistent with the increasing recognition that circulating miR-221 acts as both an intracellular regulator and an extracellular messenger that reshapes the tumour microenvironment and promotes metastasis.[60]

In the context of lung cancer, both strands of miR-221 have been associated with tumour growth, invasion and chemoresistance. Overexpression of both strands of miR-221 in NSCLC correlates with a poorer prognosis, a higher metastatic potential and a reduced sensitivity to platinum-based chemotherapy. They act by repressing tumour suppressors, including PTEN, SOCS1/3 and IRF2, which leads to the sustained activation of pro-survival and proliferative pathways, such as PI3K/Akt, JAK/STAT and NF- κ B. These signalling perturbations have combined effect of enhancing cell cycle



progression, inhibiting apoptosis, and promoting EMT-related phenotypes, ultimately contributing to tumour aggressiveness and metastatic dissemination.[61]

Overall, miR-221-5p and miR-221-3p emerge as pivotal molecular nodes that integrate oncogenic signalling, epigenic regulation and immune evasion in lung cancer. Their dual role as both intracellular effectors and extracellular mediators highlights their potential as both diagnostic biomarkers and therapeutic targets. Inhibiting the activity of miR-221 – using antisense oligonucleotides, long non-coding RNA modulation, or exosome interference – may be an effective way to counteract tumour growth, overcome drug resistance, and restore immune surveillance in solid malignancies.[46]

1.7. Peptide Nucleic Acids as a tool for silencing oncogenic miRNAs

Various antisense platforms have been developed to counteract the oncogenic activity of miRNAs, including modified oligonucleotides, locked nucleic acids (LNAs),[62] small interfering RNAs (siRNAs)[63] and, more recently, peptide nucleic acids (PNAs).[64]

1.7.1. Peptide Nucleic Acids

Several antisense chemistries have been developed to target oncogenic microRNAs. However, significant limitations remain in terms of stability, specificity, and intracellular delivery. For instance, LNAs exhibit a high binding affinity and improved nucleases resistance compared to unmodified oligonucleotides. However, they retain a charged phosphodiester backbone, which may lead to off-target interactions, immune activation and sequence-dependent toxicity. Furthermore, although LNA-based anti-miRNAs have shown, their long-term stability and biodistribution profiles are still being optimised.[62]

Peptide Nucleic acids offer a conceptually distinct alternative. PNAs are synthetic analogues of DNA and RNA (Figure 1.3). In PNAs, the natural sugar-phosphate backbone is replaced by a polyamidic N-(2-aminoethyl)glycine backbone and the nucleobases are attached via a carboxymethyl linker.[65] Despite this structural modification, PNAs can still form sequence specific hydrogen bonds with complementary nucleic acids, following both Watson and Crick and Hoogsteen pairing rules.[66]

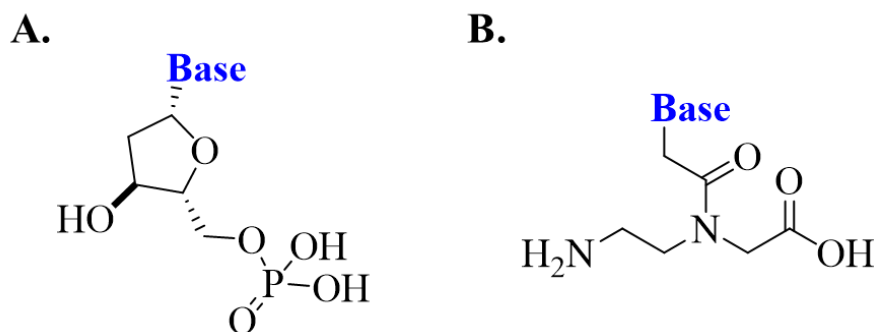


Figure 1.3. Chemical structure of **A)** DNA and **B)** PNA.

The neutral, uncharged backbone eliminates the electrostatic repulsion present in DNA/DNA duplexes. This results in a higher binding affinity and enhanced thermal stability of PNA/DNA or PNA/RNA duplexes, even under low-salt conditions.[67]

Due to their artificial backbone, PNAs are resistant to degradation by enzymes such as nucleases and proteases making them highly stable in biological fluids. This stability, together with the robustness of the polyamide backbone, enables PNAs to withstand a broad range of temperature and pH values. Unlike DNA, PNAs do not depurinate in acidic conditions, although rearrangements can occur at a pH greater than 11.[65] The structural and chemical properties of PNAs enable strand invasion, whereby a PNA oligomer displaces one strand of a DNA duplex to form a PNA/DNA hybrid. This property underpins their use in antisense and molecular diagnostic applications.[68]

PNAs are achiral, non-ionic molecules that can hybridise with complementary strands in parallel or antiparallel orientations; the antiparallel arrangement is generally more thermodynamically stable. The neutral backbone also reduces the formation of secondary structures in target DNA or RNA, thereby increasing the accessibility of complementary PNA oligomers (Figure 1.4).[68] Short PNA probes (8-15 nucleotides long) demonstrate excellent sequence specificity and are widely employed in microarrays and affinity-based assays. However, longer mixed-based PNAs (polymer nucleic acids) containing more than 16 nucleotides can display decreased sequence specificity due to an inverse correlation between binding affinity and mismatch discrimination.[69]

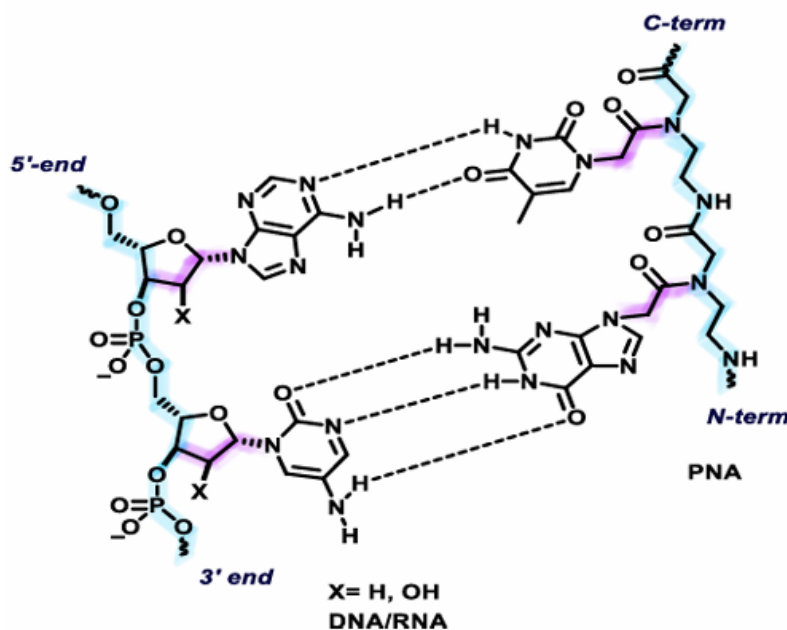


Figure 1.4. Preferred orientation of PNA/DNA or PNA/RNA duplex.

First developed in the 1980s by Ole Buchardt and Peter Nielsen, PNAs combine the recognition properties of nucleic acids with the chemical stability and synthetic versatility of peptides. This allows PNA strands to be generated via standard solid-phase peptide synthesis.[70] The strong binding affinity, high sequence and resistance to enzymatic degradation of PNAs make them a powerful tool in gene regulation studies, diagnostics and therapeutic applications.

1.7.2. PNAs Applications

Thanks to their high thermal stability and resistance to nucleases and proteases, PNAs are ideal for the artificial modulation of gene expression. Currently, PNAs are used as potential therapeutic agents and as powerful tools in molecular biology and diagnostic (Figure 1.5).[65]

Therapeutic Applications. PNAs show promise as gene-targeted therapeutic drugs, acting through several mechanisms, including antisense, antigene, and decoy strategies.

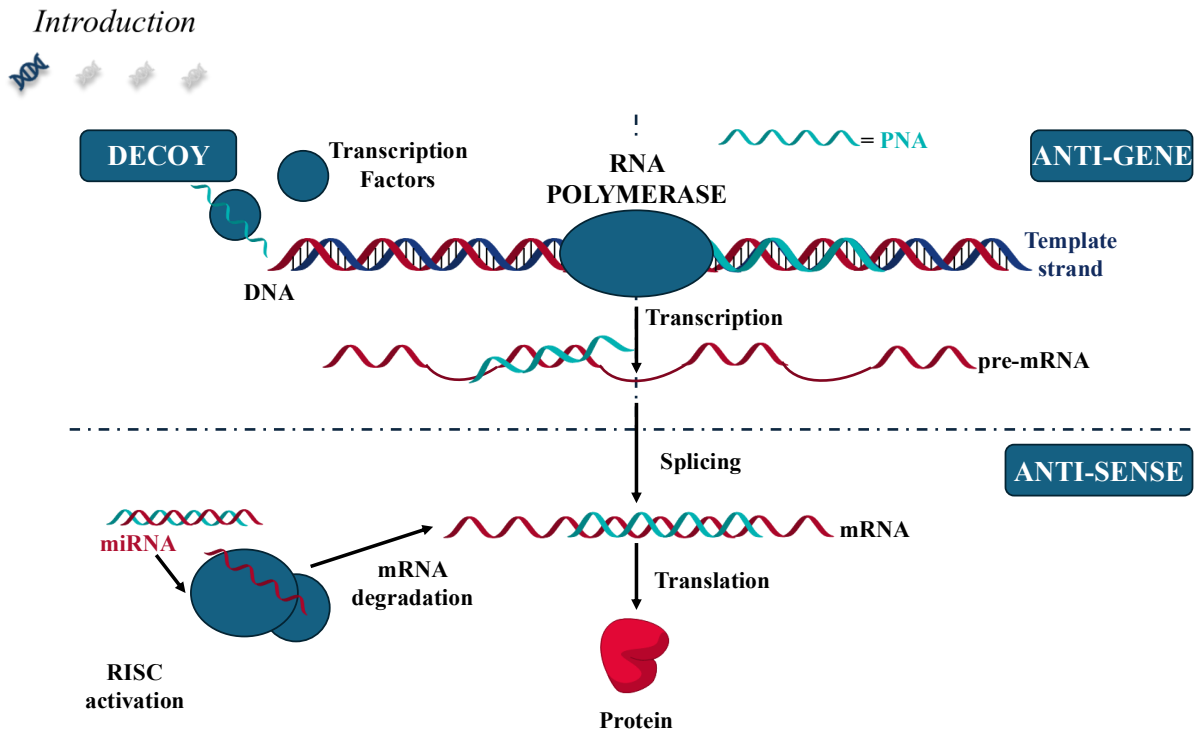


Figure 1.5. Different strategies for gene regulation using PNAs: antisense, blocking translation from mRNA into protein or regulating splicing of pre-mRNA; anti-gene, blocking transcription from DNA to mRNA; decoy, scavenging transcription factors; anti-miRNA, preventing micro-RNA activity.

1. *Antisense Strategy.* Antisense oligonucleotides bind to complementary regions of target mRNA via Watson-Crick base pairing, thereby disrupting translation.[71] The exact mechanism of action is still debated, but two main models have been proposed:
 - *Steric hindrance.* The high-affinity binding of the oligonucleotide to the mRNA physically blocks ribosome access, thereby preventing translation.[72]
 - *RNase H-mediated degradation.* Typical antisense oligonucleotide recruit RNase H to degrade the RNA strand of the RNA-DNA duplex.[73]

However, PNAs do not activate RNase H; their antisense activity relies on steric blocking of RNA processing, nuclear export or translation. Both duplex-forming and triplex-forming PNAs can inhibit translation. Triplex forming PNAs, which target homopurine regions of 10-15 bases, interfere with translation elongation, whereas duplex PNAs generally act at the initiation codon.[74]

Antisense PNAs have been shown to modulate pre-miRNA splicing and inhibit specific genes (e.g. the retinoic acid receptor). When conjugated to cell-penetrating peptides (CPPs), they can efficiently target bacterial mRNA, viral RNA and splicing machinery. These properties also enable *in vivo* applications, including antimicrobial strategies.[75]
2. *Antigene strategy.* Antigene PNAs inhibit gene transcription by forming triplex structures or strand-invasion complexes with double-stranded DNA. These complexes can either sterically block RNA polymerase progression or prevent transcription factors from binding



to promoter regions.[76] Homopyrimidine PNAs are particularly effective at targeting promoters or coding regions. Although strand invasion can be slow under physiological salt conditions, chemical modifications such as linking Watson-Crick and Hoogsteen strands, incorporating pseudoisocytosines, intercalators or positively charged residues enhance binding kinetics.[77] Nuclear-targeted PNAs directed against the C-Myc or N-Myc genes have demonstrated potent transcriptional inhibition and antiproliferative effects in human tumour cells.[78]

3. *Decoy Strategy*. In decoy applications, transcription factors bind synthetic PNA sequences rather than endogenous DNA, thereby preventing transcriptional activation. PNAs alone are often unrecognised by transcription factors due to structural differences.[79] However, PNA-DNA chimeras, in which a PNA recognition sequence is covalently linked to a DNA segment, have proven effective in allowing transcription factor binding while blocking access to native genomic targets.[80]

1.7.3. PNAs Delivery

One of the major challenges in PNA-based therapy is the low cellular uptake of PNAs. The absence of charged groups in the PNA backbone results in low solubility and a tendency to aggregate, a problem that worsens with increasing oligomer length or higher purine content.[81] Strategies have been developed to address these limitations, such as conjugating charged molecules or amino acids to the ends of the PNA or incorporating modified monomers with charged side chains, to enhance solubility and delivery.[82] Furthermore, PNA-DNA chimeras have been designed to enhance solubility and cellular uptake whilst maintaining DNA-like structural and biological properties.[80] Several delivery systems have been explored and can be broadly divided into two categories: unmodified PNAs and modified PNAs conjugated to carriers or guide ligands.

1. *Unmodified PNA Delivery*. Unmodified PNAs can be internalised via liposomes, which are phospholipid vesicles that can be either neutral or cationic. The nucleic acid can be encapsulated inside the vesicle or anchored to its surface through electrostatic interactions. Positively charged liposomes interact efficiently with negatively charged cell membranes and are internalised via endocytosis.[83] Other methods of delivering unmodified PNAs include electroporation, microinjection or diffusion at high concentrations.[84] However, the latest often carries a high risk of cytotoxicity.
2. *Modified PNA Delivery*. To improve specificity and efficiency of delivery, PNAs can be conjugated to cell-recognised molecules, including ligand-guided or receptor-targeted approaches. These approaches facilitate uptake through receptor-mediated endocytosis.



Although this approach is highly specific, it is sometimes limited by receptor availability and endosomal/lysosomal entrapment.[85]

An alternative, receptor-independent strategy involves the use of cell-penetrating peptides (CPPs), which physically interact with biological membranes in order to transport conjugated cargo into cells. These short, positively charged peptides include sequences derived from HIV Tat and *Antennapedia homeodomain*. [86] They are capable of crossing membranes and facilitating intracellular delivery. PNAs can be linked to CPPs either covalently or non-covalently, forming nanoparticles in the latter case. Cellular uptake mainly occurs through endocytosis, although direct membrane penetration may also contribute. Once internalised, PNAs must reach their intracellular targets, which may require endosomal escape.[87]

Conjugation to CPPs has been successfully applied to multiple cell types. Several studies involving penetratin- or transportan-conjugated PNAs revealed diffuse cytoplasmatic and nuclear localisation, which is consistent with direct membrane penetration mechanisms.[88] To improve peptide stability in biological fluids, D-amino acids or retro-inverso peptides are employed, as these resist enzymatic degradation while maintaining biological activity.[89]

3. *Nuclear Delivery*. Efficient nuclear delivery requires two key elements: effective cytoplasmatic transport and the presence of a nuclear signal (NLS). NLSs are short, highly basic peptide sequences that interact with nuclear pore receptors, enabling ATP-dependent transport into the nucleus.[90] For example, a dual system combining a PNA conjugated to a membrane-penetrating peptide with another PNA linked to an NLS peptide has been shown to enable cytoplasmatic entry within one hour and nuclear localisation within three hours in human prostate cancer cells.[91] Chemical modifications of first-generation PNAs are also being explored to further enhance cellular uptake and nuclear delivery.



1.8. Cell-Penetrating Peptides (CPPs): Mechanisms and Applications

1.8.1. CPPs Discovery

Polypeptides and oligonucleotides are generally regarded as molecule with low therapeutic potential due to their limited ability to cross biological membranes and their rapid degradation. These limitations pose a significant challenge in both biomedical research and pharmaceutical industry. The ability to deliver macromolecules into cells could significantly expand the range of potential biological targets.[92]

However, a major conceptual breakthrough occurred when a peptide derived from the homeodomain of the transcription factor encoded by the *Antennapedia gene* was identified. This peptide was shown to cross cellular membranes and mediate the intracellular transport of conjugated oligopeptides or oligonucleotides. This discovery opened up a whole new area of research in the field of molecular delivery.[93]

The concept of protein transduction was independently described for the first time in 1988 by Frank and Pabo [94] and by Green and Loewenstein, [95] who demonstrated that HIV trans-activator of transcription (Tat) protein could be internalised *in vitro* and activate viral gene expression. In 1997, Vives further advanced this concept by identifying a truncated fragment of the Tat protein that retained efficient cell-penetrating activity.[96]

Soon afterwards, a 16-amino acid amphipathic peptide named Penetratin (pAntp), derived from the homeodomain of *Drosophila Antennapedia*, was reported as another efficient cell-penetrating sequence.[97] Since then, more than one hundred peptides, consisting of 5 to 40 amino acids have been identified as capable of traversing biological membranes in a receptor- and energy independent manner. This enables the intracellular delivery of a wide variety of bioactive cargoes, including small molecules, nucleic acids, and therapeutic agents.[98]

The term ‘Cell-Penetrating Peptides’ has been coined to refer to this class of short, cationic peptides (typically fewer than 20 amino acids) that can cross the plasma membrane of many mammalian cells.[99]

1.8.2. Structural and Functional Features of CPPs

The therapeutic relevance of CPPs lies in their ability to facilitate the uniform delivery of large biologically active molecules, such as proteins and nucleic acids, into cells. In recent years, CPPs have attracted growing interest due to their potential to deliver recombinant proteins to various tissues in animal models. These advances have reinforced the idea that intracellularly active macromolecules,



which were once considered impossible to deliver, could become clinically exploitable if they are conjugated to suitable CPP carriers.[100]

CPPs can be classified according to their origin or sequence characteristics. Most known CPPs are cationic, non-cell-specific peptides that are typically enriched in basic amino acids, such as arginine and lysine. Their cellular uptake is primarily initiated by electrostatic interactions between the positively charged CPP and anionic components of the plasma membrane (phospholipids, cell-surface glycoproteins).[101] Arginine is particularly effective due to its guanidinium group, which forms bidentate hydrogen bonds with phosphate, carboxylate or sulphate groups on the plasma membrane. Lysine lacks this structural feature and is therefore generally less efficient when active alone.[102] The sequence length and amino acid arrangement of CPPs determine their translocation efficiency. Low-amphipathic peptides in this class include Penetratin, Tat, and polyarginines, whereas a distinct class of highly amphipathic CPPs – such as the Model Amphipathic Peptide (MAP), Transportan and Pep1 – derive their positive charge predominantly from lysine residues. A third structural category comprises amphipathic peptides in which hydrophobic and charged residues are segregated along the sequence, as in the pVEC and MPG peptides.[99]

One of the most widely studied internalising peptides is Penetratin, which is derived from the homeodomain of *Drosophila Antennapedia* transcription factor. [97] Homeoproteins are DNA-binding transcription factors characterised by 60 amino acids homeodomain. Early discoveries revealed that the *Antennapedia* homeodomain could be secreted and internalised by neighbouring cells, suggesting that intercellular protein transfer is not restricted to classical secretion pathways.[103]

Structure-function analyses, including site-directed mutagenesis, have demonstrated the third helix of the homeodomain (residues 43-58) is both necessary and sufficient for cell translocation. This has led to the development of Penetratin as a minimal CPP. [104]

Penetratin internalises efficiently at both 37°C and 4°C in a non-saturable manner, indicating that its uptake does not require interaction with stereospecific receptor. Although uptake is increased in cells expressing α -2,8-polysialic acid (PSA) – likely due to an increased local concentration or membrane stabilisation – Penetratin also penetrates PSA-negative cells efficiently, suggesting that multiple classes of negatively charged surface molecules may facilitate uptake.[104]

Further support for a receptor independent mechanism comes from synthetic analogues of Penetratin. Both the all-D enantiomer and the retro-inverso peptide penetrate cells with a similar efficiency to that of the native sequence.[105] Conformational studies using circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopies have revealed that Penetratin adopts predominantly a



helical conformation in TFE:H₂O (9:1) solution resembling a 3₁₀-helix, with β -turns at both termini, and featuring a hydrophobic core that may promote membrane association.[106]

In addition to naturally occurring CPPs, numerous synthetic analogues have been engineered from short synthetic sequences, including Transportan, the Model Amphipathic Peptide and homoarginine-based vectors. Remarkably, peptides composed entirely of arginine residues have been reported to enter cells via an energy-dependent yet non-endocytic process, thereby expanding the structural and mechanistic diversity of CPP-mediated delivery systems even further.[87]

1.8.3. Mechanisms of Cellular Uptake

In order to understand how CPPs achieve intracellular transport, it is first essential to consider the general mechanism by which cells internalise foreign molecules (Figure 1.6). Cells rely on three main types of uptake mechanisms: endocytosis, membrane transport and passive diffusion. The size, charge, and physicochemical properties of the incoming compound dictate which mechanism is used. However, CPPs represent a mechanistically distinct class of delivery vector as they exploit both endocytic pathways and a unique, energy-independent mechanism involving the transient perturbation of the lipid bilayer. This duality distinguishes CPPs from all other biomolecules and highlights their potential as delivery systems for therapeutic cargoes.[87]

These pathways are fundamental mechanisms of cellular internalization that enable the transport of a wide variety of substances from the extracellular environment into the intracellular space. These processes are essential for maintaining cellular homeostasis and regulating nutrient uptake, receptor signalling, and defence against pathogens.[107]

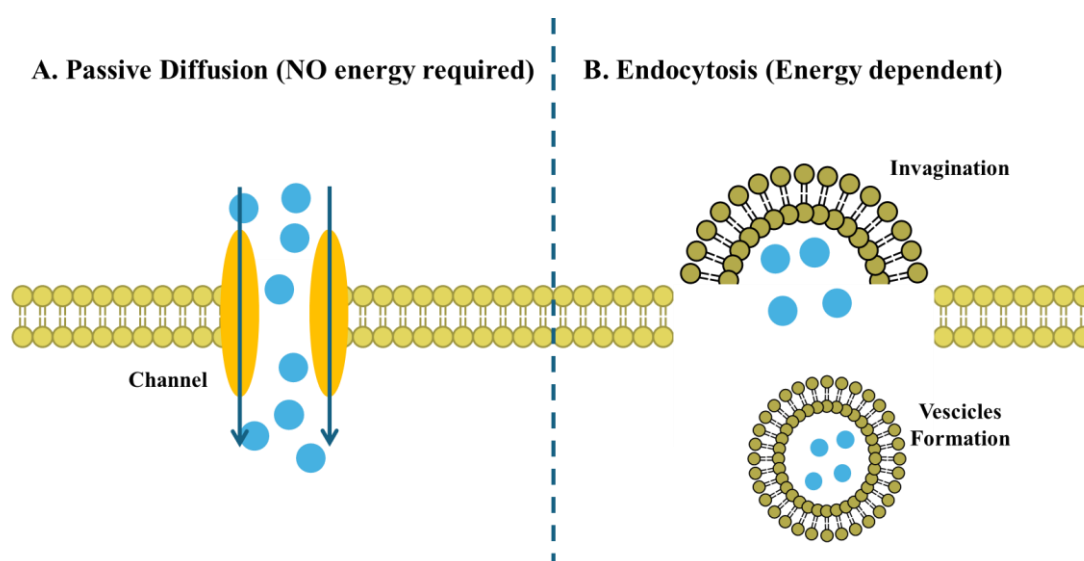


Figure 1.6. Internalisation pathways: *A) Passive Diffusion; B) Endocytosis.*



Endocytosis is an energy-dependent process involving the active remodelling of the plasma membrane to form vesicular structures that can engulf extracellular materials. This mechanism enables cells to internalise macromolecules, large particles and fluids that cannot cross the lipid bilayer directly.[108] There are several distinct forms of endocytosis:

- Phagocytosis, often referred to as ‘cell eating’, is the process by which specialised cells, such as macrophages and neutrophils, engulf large particles, including microorganisms and cellular debris. The plasma membrane extends around the target particle, enclosing it in a vesicle known as a phagosome. This subsequently fuses with lysosomes, which carry out enzymatic degradation.[109]
- Pinocytosis, or ‘cell drinking’, involves the non-selective uptake of extracellular fluid and dissolved solutes through small vesicles. This form of endocytosis occurs continuously in many cell types and plays a crucial role in sampling the extracellular environment.[110]
- Receptor-mediated endocytosis provides a highly selective means of internalisation. In this process, specific molecules bind to complementary receptors on the cell surface, triggering the formation of clathrin-coated vesicles that carry the bound ligands into the cytoplasm. This mechanism is essential for the regulated uptake of hormones, growth factors and lipoproteins.[111]

On the other hand, passive diffusion refers to the passage of small molecules across the plasma membrane, often occurs without the involvement of vesicular transport:[112]

- Small, nonpolar molecules such as oxygen, carbon dioxide and certain lipophilic compounds may undergo this process via passive diffusion, moving along their concentration gradient without expending energy.[113]
- Polar or charged molecules that cannot cross the hydrophobic core of the membrane freely use facilitated diffusion and active transport mechanisms. These involve specific carrier or channel proteins that facilitate the movement of ions, sugars, and amino acids either down or against their concentration gradients.[114]

When molecules cannot cross the lipid bilayer by simple diffusion, cells rely on a sophisticated network of membrane transport proteins to mediate the selective and regulated exchange of ions and metabolites. These systems overcome the hydrophobic membrane’s intrinsic permeability barrier, ensuring tight control over the composition of the interior of the cell and maintaining essential physiological functions.[114] One major class of these transport systems comprises ion channels: pore-forming proteins that allow the rapid movement of inorganic ions such as Na^+ , K^+ , Ca^{2+} , and Cl^- . Their opening and closing are precisely regulated by chemical ligands, changes in membrane potential, or mechanical forces. As they conduct ions down their electrochemical gradients, ion channels support processes requiring fast, coordinated responses, including action potential



propagation, neurotransmission, muscle contraction and osmotic regulation.[115] In parallel, the membrane contains a wide range of carrier proteins that mediate the translocation of larger or more polar molecules, such as sugars, amino acids and nucleosides, through a series of conformational changes. Unlike channels, carriers bind their substrates through specific recognition sites and subsequently undergo structural rearrangements that alternately expose the binding pocket to either side of the membrane. This process, known as facilitated diffusion, occurs without the expenditure of energy, enabling the efficient and selective uptake of essential nutrients that cannot diffuse across the lipid bilayer.[116] For molecules that must be transported against their concentration gradient, cells employ active transporters, which require an external energy source. Primary active transporters, such as the Na^+/K^+ -ATPase, directly consume ATP to pump ions uphill, thereby establishing electrochemical gradients that are indispensable for cellular excitability and volume control. Secondary active transporters, on the other hand, harness the energy stored in these pre-existing gradients to drive the cotransport or antiport of molecules (e.g. the Na^+ /glucose symporter), thereby integrating nutrient uptake with ion homeostasis.[117]

Together, these complementary strategies ensure the efficient exchange of materials between the cell and its environment. While endocytosis provides a means of internalising macromolecules selectively or in bulk, channels or transporters allow for the controlled passage of smaller molecules and ions that are critical for metabolic and signalling processes. The coordination of these mechanisms reflects the adaptability of the cell membrane, which serves as a protective barrier and a dynamic interface for communication and exchange with the extracellular environment.[118]

1.8.4. Interplay between Direct Membrane Translocation and Endocytic Uptake in CPP Internalisation.

CPPs and CPP-cargo conjugates have a unique ability to cross the plasma membrane, using mechanisms that are fundamentally different from the standard uptake routes used by most biological molecules. A key feature of CPPs is their ability to transiently and non-deleteriously perturbate the lipid bilayer, enabling direct access to the cytosol without the involvement of vesicular internalisation. This phenomenon is often referred to as ‘direct translocation’ and has been explained by several biophysical models:[119]

- The transient pore model suggests that strong electrostatic interactions between highly cationic CPPs, such as TAT or poly-arginines, and negatively charged phospholipids promote the formation of short-lived toroidal or barrel-stave pores. These pores destabilise the membrane momentarily, allowing the peptide or its cargo to slip through before the lipid bilayers releases rapidly.[120]



- Alternatively, the inverted micelle model proposes that CPPs accumulate at the membrane interface to form micelle-like structures that engulf the peptide, invert across the bilayer and release their contents directly into the cytosol.[121]
- A third mechanism, often described as a carpet-like model, suggests that CPPs adsorb extensively onto the membrane tension and facilitating partial insertion or transbilayer passage.[122]

These receptor-independent processes generally occur more frequently at higher peptide concentrations, particularly for CPPs enriched in cationic residues. The charge density of these residues drives strong electrostatic interactions with anionic lipid headgroups.

However, CPPs internalisation does not rely exclusively on direct membrane perturbation. At lower concentrations of the peptide, where the energetic threshold for bilayer destabilisation may not be reached, or when bulky or hydrophilic cargoes are attached, CPPs predominantly exploit classical endocytic pathways. These include clathrin-mediated uptake, caveolin-dependent internalisation and micropinocytosis. Each of these processes transports CPPs into intracellular vesicles rather than directly into the cytosol. Although efficient, these endocytic routes often lead to endosomal sequestration, which significantly limits bioavailability and biological activity.[123]

This is a particular challenge for CPP-PNA conjugates, where endosomal entrapment can prevent the antisense oligonucleotide from reaching its target miRNA, as has been observed experimentally. The predominance of one or the other internalisation mechanism is influenced by factors such as peptide sequence and charge distribution, local peptide concentration, cargo size and hydrophobicity, and the physicochemical properties of the target membrane. Therefore, understanding the interplay between these routes is crucial for the rational design of effective CPP-PNA conjugates and optimising their performance in therapeutic settings.[123]

1.8.5. Cargo Conjugation and Delivery Strategies

CPP conjugation can sometimes reduce the biological activity of the cargo molecule. Therefore, effective linker strategies are required to enable the successful delivery of the cargo to its target site, and ideally its release there too.[124] Several approaches have been developed to link CPPs and their cargoes:

- Non-covalent complex formation between the carrier and the cargo (though this is rarely used);[124]
- Sequential synthesis of the peptide chain when the cargo is also a peptide;[125]



- Conjugation through disulfide (S-S) bond formation between the carrier and the cargo. This facilitates both cellular uptake and intracellular reduction of the bond, enabling cargo release;[126, 127]
- Click chemistry, which involves the formation of a stable triazole bond that is resistant to physiological conditions and enzymatic degradation.[128]

1.9. Aim of the Chapter

This chapter explores the design, synthesis and biological evaluation of CPP-PNA conjugates that have been developed to serve as intracellular delivery tools for antisense applications. PNAs have an excellent affinity for hybridisation and are highly stable, but they have low intrinsic membrane permeability. This means that their functional use in cells depends on the choice of an efficient carrier. To this end, a panel of CPPs was selected and compared to identify the most suitable delivery system for the specific PNA sequence targeting miRNA-221-5p.

The efficiency of a CPP is not an intrinsic property of the peptide alone; it is also influenced by the nature, size and polarity of the cargo it transports. Therefore, rather than assuming that a CPP known to internalise efficiently on its own would automatically perform well when conjugated to a PNA, a comparative evaluation was necessary. This study selected several CPPs with different structural and physicochemical features – including TAT, Penetratin and P14 – representing distinct families of cationic, amphipathic and arginine-rich peptides. Their different charge distribution, secondary structure propensity, and membrane interaction behaviour provided a rational basis for identifying the most promising scaffold for PNA delivery.

Two complementary conjugation strategies were implemented and compared to covalently attach the CPPs to the PNA:

- CuAAC chemistry, which generates a stable triazole linkage,
- Disulfide bond formation, yielding a reversible, reduction sensitive linker.

These two approaches were chosen because they are robust and compatible with both CPP and PNA chemistry, typically providing high coupling yields.

However, their biological behaviour is expected to differ significantly. The triazole linkage is chemically stable inside cells, ensuring that the CPP remains permanently attached to the PNA. While this guarantees the delivery of an intact conjugate, there is a risk that CPP moiety will interfere with hybridisation to the target microRNA due to steric hindrance. In contrast, the disulfide linkage offers a cleavable alternative by exploiting the strongly reducing environment of the cytosol, where redox proteins can promote the rapid release of the free PNA. While this may enhance target binding, it also

Introduction



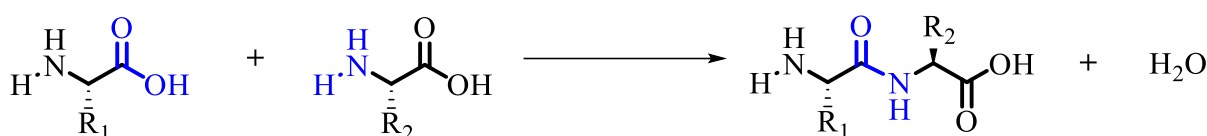
introduces potential drawbacks, such as premature cleavage in extracellular media or endosomes. This could prevent the CPP from transporting its cargo to the intended intracellular compartment. These considerations also explain why the CPP-PNA conjugates could not be synthesised as a single, continuous chain using Fmoc-SPPS. Although direct synthesis is conceptually simpler, it often suffers from low yields, aggregation on the resin and incompatibility between PNA monomers and peptide elongation conditions. Furthermore, the requirement to introduce orthogonal functional groups for selective coupling renders a convergent post-synthetic conjugation approach more flexible and reliable.

Overall, this chapter aims to establish a comprehensive framework linking chemical design, conjugation strategy and biological performance. By systematically comparing different CPPs, linker chemistries, and internalisation behaviours, the work provides the experimental basis for developing efficient CPP-PNA systems for intracellular delivery and microRNA modulation.



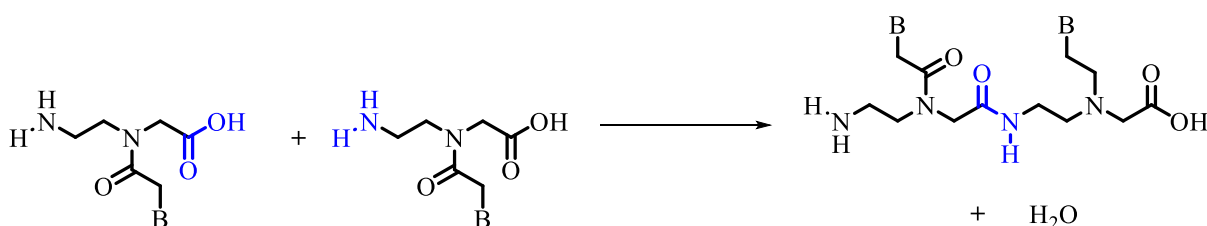
2. Results and Discussion

Peptides are molecules composed of amino acids that are linked together by peptide (amide) bonds (Scheme 1.1). They can range in length from just two amino acids to around fifty residues. In nature, peptides and proteins are usually made up of the 20 standard α -amino acids. An α -amino acid contains an amino group attached to the α -carbon of a carboxyl group. Amino acids that occur naturally are almost exclusively found in the L-configuration, except for glycine, which is achiral. D-amino acids, the mirror-image forms, are rare in mammals but more commonly found in peptides produced by bacteria, fungi, and other microorganisms. According to Cahn-Ingold-Prelog (CIP) rules, L-amino acids generally have the S configuration, except for cysteine, which has the R configuration.[129]



Scheme 1.1. Formation of a dipeptide from two L- α -amino acids via condensation reaction.

Peptide Nucleic Acids are synthetic analogues of peptides in which the natural peptide backbone is replaced by a pseudopeptide scaffold, typically composed of N-(2-aminoethyl)-glycine units, to which nucleobases are attached. (Scheme 1.2) [130]



Scheme 1.2. Formation of a PNA dimer via condensation reaction.

In both peptides and PNAs, a key challenge in synthesising a defined sequence is achieving regioselective coupling – ensuring that each monomer connects exclusively through the desired functional groups. To enable this level of control, the reactive amino and carboxyl groups that should not participate in the coupling step must be temporarily protected using suitable protecting groups.

2.1. Solid Phase Peptide Synthesis

Although solution-phase synthesis remains useful for the large-scale production of specific peptide sequences, solid-phase peptide synthesis (SPPS) is now the most widely used strategy for preparing peptides – and, by extension, PNAs.

This technique was first introduced in 1963 by Bruce Merrifield, who revolutionised peptide chemistry by proposing the assembly of peptide chains on an insoluble solid support.[131] This approach dramatically reduced synthesis time by eliminating the laborious purification steps required

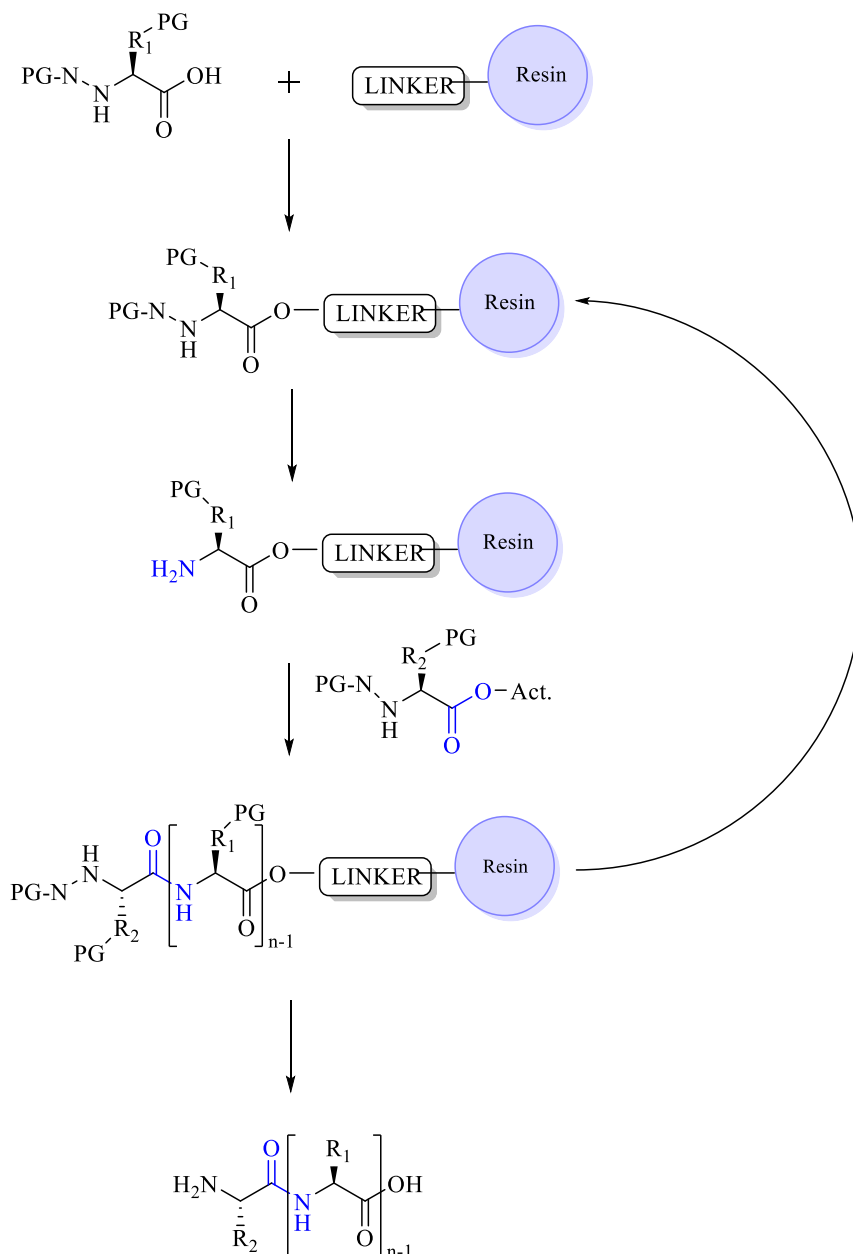


at each stage of solution-phase synthesis and paved the way for the full automation of the process. Another advantage is that all steps can be carried out in a single reaction vessel, eliminating the need of material transfer.

In SPPS (Scheme 1.3), the peptide or PNA chain is anchored to a solid support, typically a cross-linked polystyrene resin containing 1-2% divinylbenzene. The resin is functionalised with a suitable linker that enables the first monomer to be covalently attached through its carboxyl (or pseudopeptide) group. This gives access to either an ester- or amide- bound intermediate, ultimately leading to a C-terminal acid or amide product. Prior to synthesis, the resin is swollen – typically in DMF – to maximise accessibility to reactive sites of the polymer matrix.[132]

The peptide or PNA is then assembled in a stepwise C→N direction via iterative cycles of N α -deprotection and coupling of the incoming amino acid or PNA monomer. The carboxyl group of the amino acid is activated to increase its electrophilicity. After each coupling step, excess reagents and by-products are easily removed by filtration and washing. This sequence of deprotection followed by coupling is repeated until the desired oligomer is obtained. The final product is then cleaved from the resin.[133]

To ensure high overall yields – which are essential when synthesising medium- to long-chain peptides – it is crucial to use orthogonal protecting group strategies. The α -amino group of each incoming residue is temporarily protected, typically with a urethane-type group. This can be removed under mild conditions, minimising the risk of racemisation via α -CH deprotonation or 5(4H)-oxazolone formation. Side-chain functional groups, if present, are masked with semi-permanent protecting groups, that remain stable during chain elongation, but are removed during the final global deprotection step, alongside resin cleavage.[133]

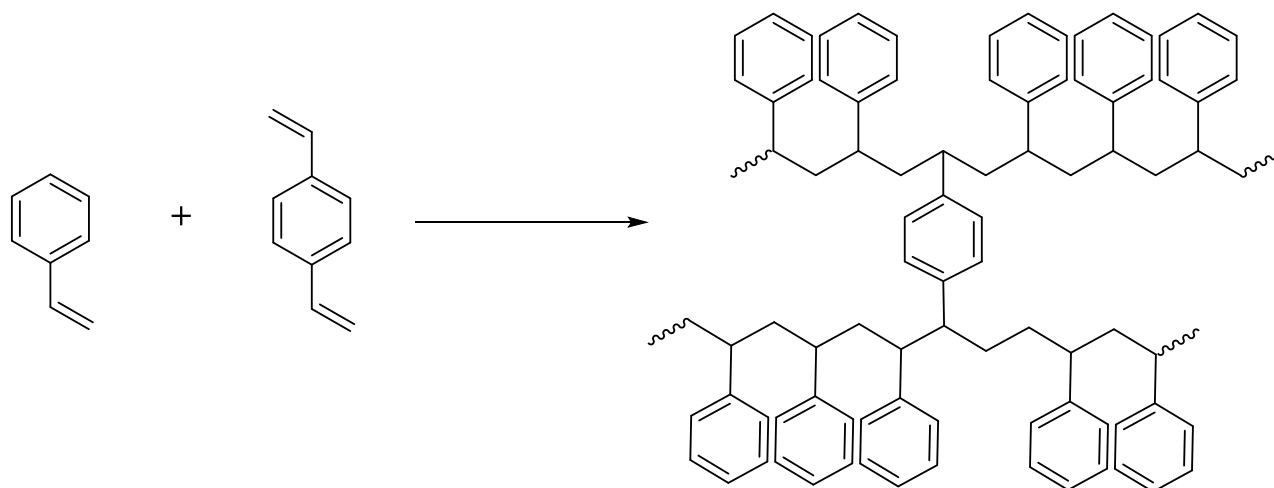


Scheme 1.3. SPPS procedure.

The same fundamental logic and workflow is applied to PNA synthesis, with the main difference being the nature of the monomers and backbone. Nevertheless, PNAs fully benefit from the advantages of the SPPS platform, enabling rapid, high-fidelity assembly under highly controlled conditions.

2.1.1. The Solid Support and the Resin

The solid support usually consists of a cross-linked polystyrene polymer that contains 1-2% divinylbenzene (DVB) (Scheme 1.4). The DVB units create branching points within the polymer structure, providing excellent mechanical stability and enabling the resin to efficiently absorb organic solvents.[134]

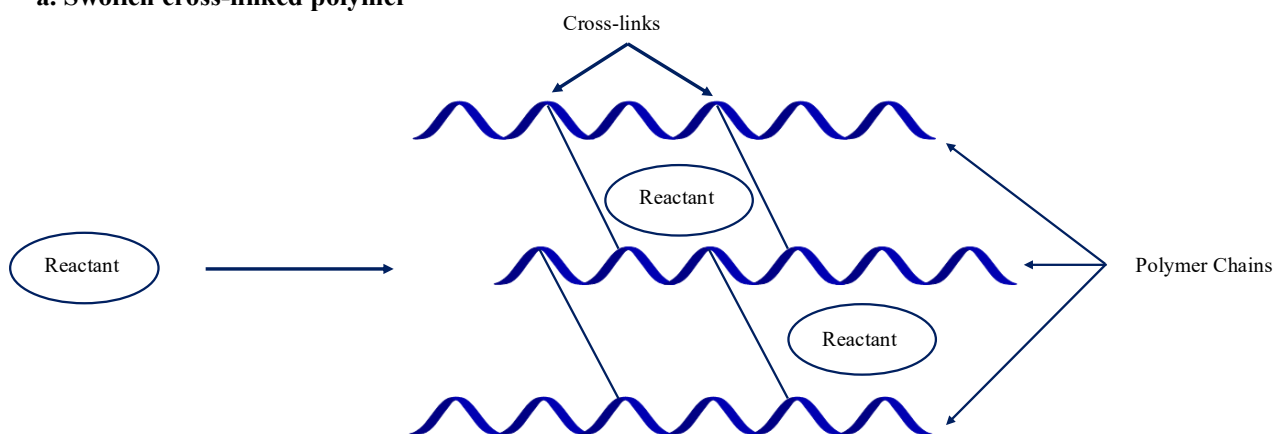


Scheme 1.4. Polymerization of styrene and DVB.

The support must be insoluble in the reaction medium and chemically inert towards all reagents used during synthesis. It must also be capable of withstanding mechanical stress and composed of particle large enough to allow rapid filtration (Figure 1.7). Prior to synthesis, the resin is swollen in the selected solvent – most commonly dimethylformamide (DMF), dichloromethane (DCM), or N-methylpyrrolidone (NMP) – for approximately 40 minutes. This swelling process is essential to ensure efficient penetration of the reagents throughout the polymer matrix and grant access to all functionalised sites where the peptide chain is assembled.[134]



a. Swollen cross-linked polymer



b. Nonswollen cross-linked polymer

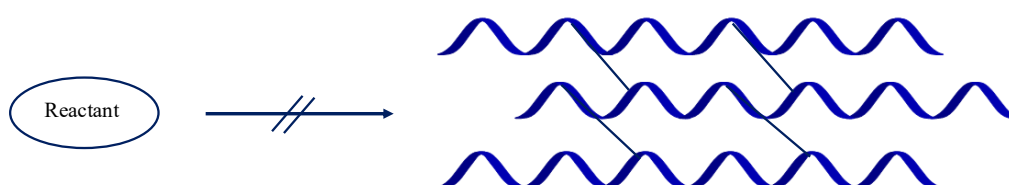


Figure 1.7. Swollen effect on SPPS. *A)* swollen polymer can accommodate reactants inside its network resulting in reaction occurrence. *B)* reactants cannot enter the narrow network of non-swollen polymer; and the reaction does not take place.

The so-called ‘resin’ is obtained by functionalising the polymer with an organic linker (Figure 1.8). Since linkers can differ significantly in length and chemical nature, there are now hundreds of commercially available polystyrene-based resins accessible for SPPS, which can be tailored to generate different C-terminal functionalities upon cleavage.[133]

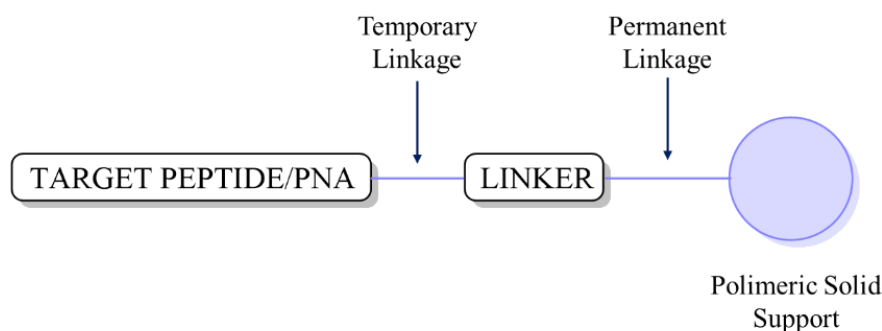


Figure 1.8. Immobilization of a target peptide or PNA to a solid support via a linker.

In this study, two different polystyrene-DVB resins were used. Rink Amide (Figure 1.9) resin was used for synthesising both peptides and PNAs, since its acidolytic cleavage produces the desired C-terminal amide ($-\text{NH}_2$), which is the native terminus found in many biological systems. Furthermore, C-terminal amidation is widely used in synthetic peptides because it increases metabolic stability and resistance to proteolytic degradation.[135]

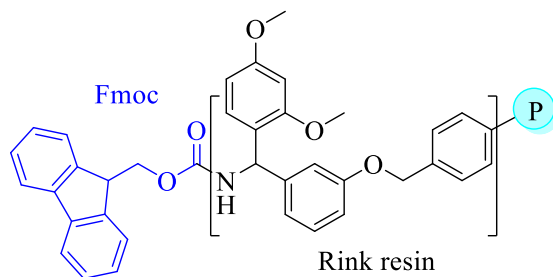


Figure 1.9. Rink Amide resin.

Conversely, 2-chlorotrityl chloride (2-Cl-Trt) resin (Figure 1.10) was used exclusively for peptide synthesis to produce C-terminal carboxylic acids. It is worth noting that 2-Cl-Trt resin is classified as super acid-labile, allowing cleavage under mild conditions such as treatment with low percentages of TFA in DCM (often as low as 1%) without affecting the side chain protecting groups. In this study, however, the traditional cleavage conditions were applied to ensure the complete release of the peptide from the solid support.[136]

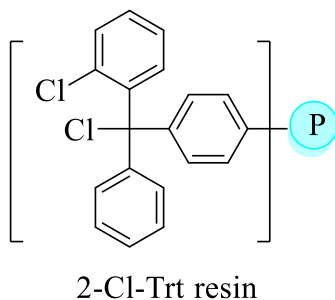


Figure 1.10. 2-chlorotrityl chloride resin.

2.1.2. The Fmoc protecting group for the amino acids

In SPPS, two distinct protecting group strategies have historically been used: the Boc strategy (Figure 1.11, B) and the Fmoc strategy (Figure 1.11, A).

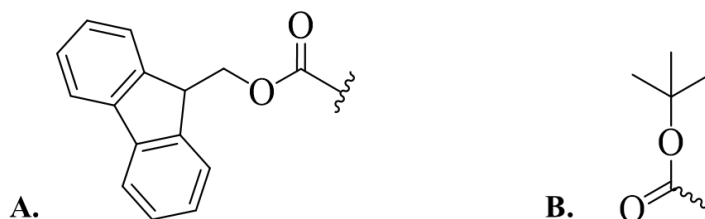


Figure 1.11. Structure of the A) Fmoc and the B) Boc protecting groups.

In the Boc approach, the *tert*-butyloxycarbonyl (Boc) group is used to temporarily protect the α -amino group. Boc removal requires strong acidic conditions, typically 50% trifluoroacetic acid (TFA) in dichloromethane (DCM), whereas final deprotection of the side chains and cleavage from the resin demand extremely strong and hazardous acids such as hydrofluoric acid (HF) or trifluoromethanesulfonic acid.[137]



While the Boc strategy prevents undesired side reactions by releasing the peptide with a protonated N-terminus, the handling of highly toxic reagents and the need for specialised equipment have significantly limited its use in typical laboratory settings.

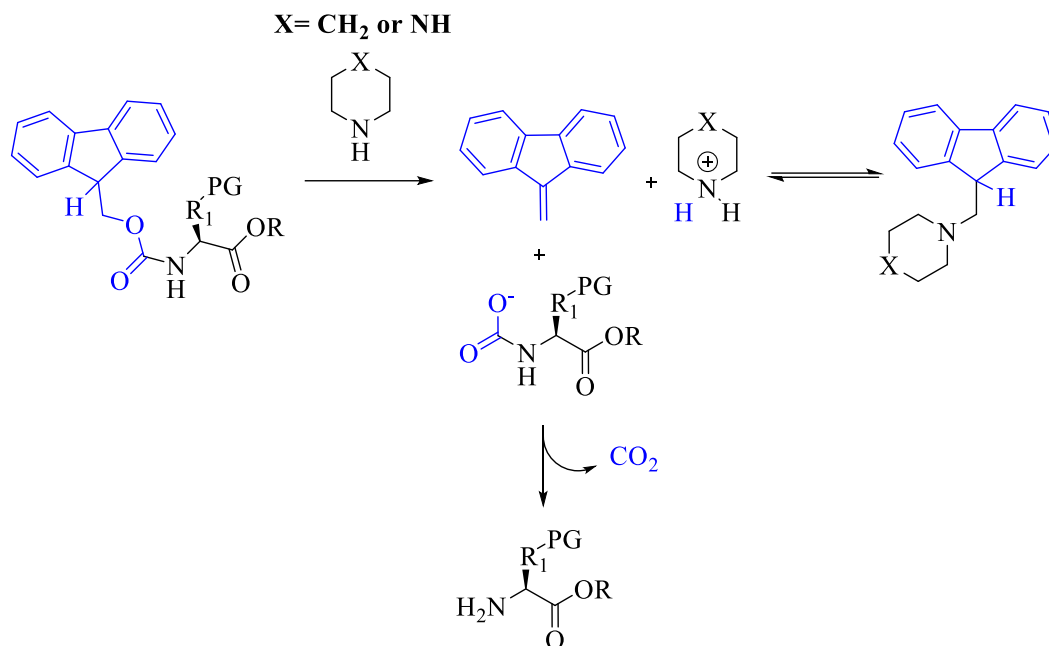
In contrast, the Fmoc strategy – now the method of choice in most peptide laboratories – uses the base-labile fluorenylmethyloxycarbonyl (Fmoc) group to reduce the nucleophilicity of the α -amino function. In this approach, mild basic conditions are used for temporary N-terminal deprotection (Table 1.1). In fact, the Fmoc group can be easily removed when primary (cyclohexylamine or ethanolamine) or secondary (piperidine and piperazine) amines are present. Furthermore, it is quickly removed in the presence of F^- ions (TBAF) and DBU but slowly removed in the presence of tertiary amines. The N-terminus is released in its deprotonated form, which increase the risk of side reactions. However, the method offers vastly greater safety, practicality and compatibility with automation.[138] The Fmoc removal conditions are summarised in Table 1.1.

**Table 1.1.** Different methodologies for the removal of the Fmoc group.

Base strength	Base	Time necessary for the complete removal of the Fmoc group
Strong Base	DBU	< 5 min
	Pyrrolidine	
	Cyclohexylamine	
	Piperidine	
Medium Base	Ethanolamine	15 min
	Piperazine	1h
	Morpholine	2h
	Diethylamine	3h
Weak Base	Triethylamine	> 18 h
	Ammonia	
	Triethylamine	
	Dicyclohexylamine	

In SPPS, an orthogonal protection strategy is employed: the α -amino group is protected using a temporary protecting group that is removed by base, while the sidechain protecting groups are stable under base conditions and are cleaved only under strong acidic conditions at the of the synthesis, simultaneously with resin cleavage.

In this study, the Fmoc group was selected for protecting the N-terminus in both peptide and PNA synthesis. The Fmoc group is typically removed through E_{1cb} elimination (Scheme 1.5), whereby a secondary amine abstracts the acidic proton adjacent to the carbonyl. This generates a resonance-stabilised anion and liberates dibenzofulvene (DBF), which is then rapidly quenched by an excess of base to prevent side reactions.



Scheme 1.5. Mechanism for the removal of the Fmoc group.

Two different deprotection conditions were employed, depending on the nature of the growing chain. For peptides, Fmoc was removed using a 20% piperidine solution in DMF, ensuring rapid and efficient deprotection. However, early syntheses revealed significant aggregation and formation of side products when using this standard condition for PNAs. Therefore, a milder and more controlled protocol was adopted: 5% piperazine in DMF supplemented with 2% DBU. The amidine structure of DBU makes it a stronger base than the secondary amines that are usually employed. However, it cannot react with DBF. Therefore, to prevent secondary reactions by the DBU, released during the removal of the Fmoc group, piperazine is used. Piperazine also acts as a chaotropic molecule, reducing the aggregation effect that can occur in growing peptide/PNA chains.

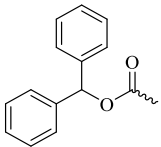
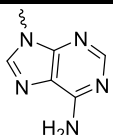
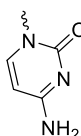
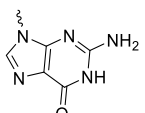
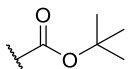
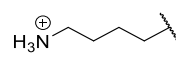
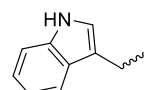
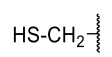
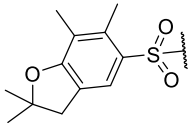
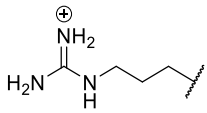
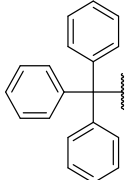
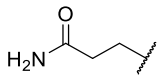
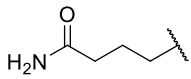
In both cases, the DBF by-products are efficiently removed through simple washing and filtration steps after each deprotection cycle. This leaves the free, non-protonated α -amino functionality ready for subsequent coupling reactions.



2.1.3. Side chains protection

As mentioned above, the side chains of the various amino acids and PNA monomers must also be protected orthogonally to the amino function protecting group to avoid secondary reactions.[138] The protecting groups used are shown in Table 1.2.

Table 1.2. Most common protecting groups in Fmoc SPPS for the side chains of the monomers used in this chapter.

Protecting Group	Amino Acid/PNA monomer	Side Chain
 Bhoc	Adenine	
	Cytosine	
	Guanine	
 Boc	Lys	
	Trp	
	Cys	
 Pbf	Arg	
 Trt	Asn	
	Cys	
	Gln	



2.1.4. Activation of the carboxylic group

Different coupling strategies were adopted in this work depending on the nature of the system to be synthesised and the chemical requirements of each condensation reaction (Figure 1.12). Specifically, HBTU in the presence of HOBT was used for stepwise peptide chain elongation on a solid support; HATU was used exclusively to couple carboxyfluorescein to N-terminus of peptides; and Oxyma Pure and DIC were combined for PNA oligomer synthesis.

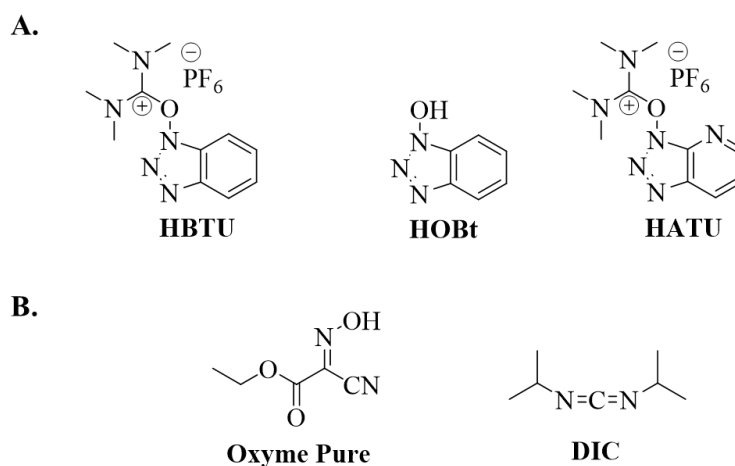


Figure 1.12. Structure of the activators **A)** HBTU, HOBT, HATU and **B)** Oxyma Pure and DIC.

Each system activates the α -carboxylic function through a distinct, yet conceptually related, mechanism based on the transient formation of an electrophilic acylating species. The reactivity and stability of this species are finely tuned by the electronic properties of the activating reagent and the leaving group.

Two closely related uronium-based activation strategies were used in peptide synthesis. The first uses HBTU and HOBT together, while the second relies solely on HATU. In the HBTU/HOBT protocol, HOBT acts as both an additive and as a transient intermediate that, according to Le Châtelier's principle, drives the reaction to completion. It nucleophilically attacks the acyl-uronium intermediate to form the more reactive benzotriazolyl ester, accelerating coupling and suppressing racemisation. Conversely, HATU contains a 7-aza-benzotriazole leaving group that enables self-activation without the need for external additives. The additional nitrogen atom enhances the electron-withdrawing properties of the ring, improving the HOAt leaving group's ability. It also participates in intramolecular base catalysis, providing a 'double-neighbouring group effect'. This dual assistance explains the superior coupling efficiency and reduced racemisation by around 50% typically observed for HATU, despite the reagent being more expensive. Crystallographic and mechanistic studies have shown that, in the benzotriazolyl active esters produced by these systems, the carbonyl group adopts an orthogonal orientation with respect to the α -carboxylic plane. Meanwhile, the triazolic nitrogen engages in intramolecular hydrogen bonding and anchimeric assistance. This arrangement of



geometry and electrons orients the electrophilic carbon for nucleophilic attack along the Bürgi-Dunitz trajectory and facilitates proton transfer from the incoming amine to the leaving group (Figure 1.13). [139, 140]

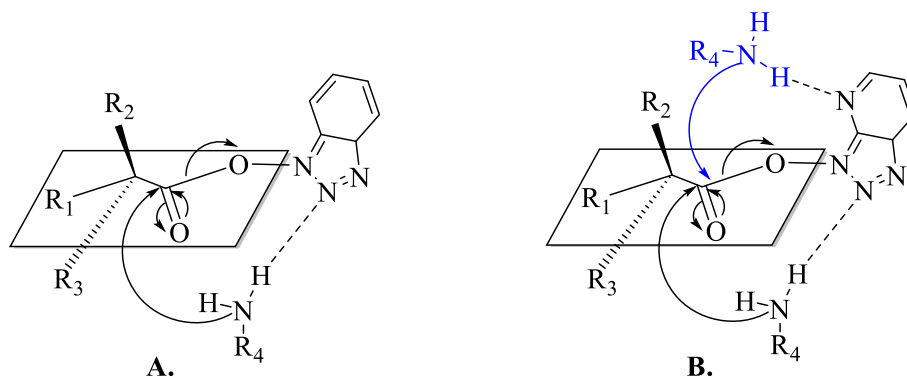
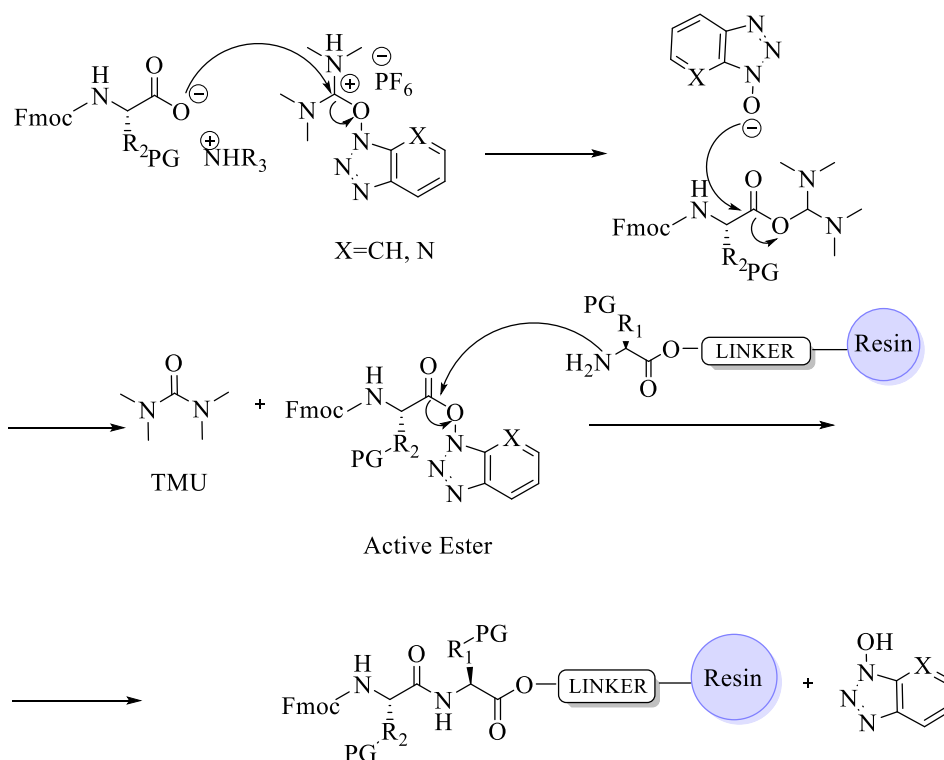


Figure 1.13. Neighbouring effect for **A)** HBTU and **B)** HATU.

Upon deprotonation of the amino acid carboxyl group by the tertiary base diisopropylethylamine (DIPEA), the carboxylate attacks the cationic centre of the uranium intermediate. The reactive intermediate then transfers the acyl group to the amine of the growing peptide chain, resulting in peptide bond formation (Scheme 1.6).[141]



Scheme 1.6. Activation of the carboxylic group.

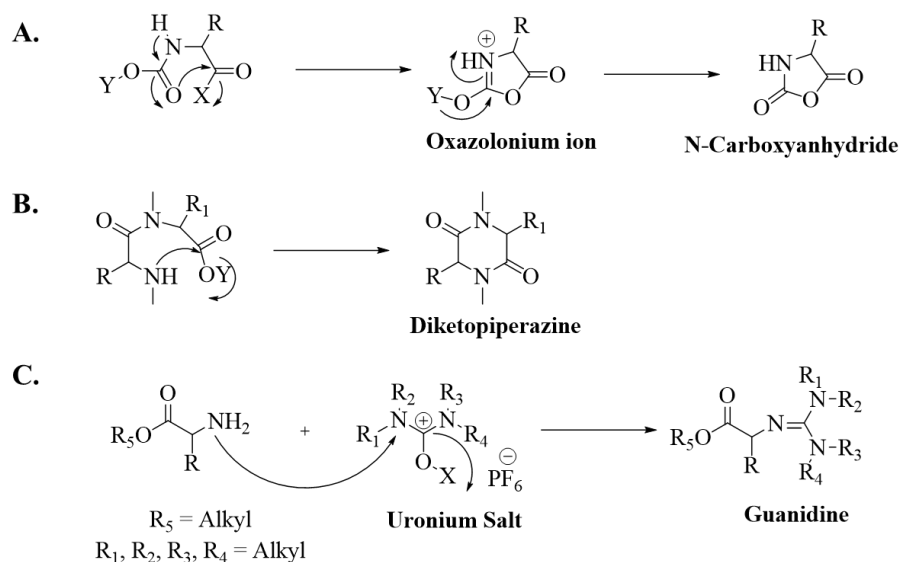
The distinction between HBTU/HOBt and HATU lies not only in the composition of the reagents but also in the balance between external and internal catalysis. The former depends on the *in situ* formation of an active HOBt ester to sustain reactivity, whereas the latter incorporates the catalytic function directly into the heterocyclic scaffold of the reagent. Both strategies provide high coupling



efficiency and stereochemical control when used with the correct stoichiometry (in this work, three equivalents of amino acid and activator, and six equivalents of DIPEA).[140]

Besides, using uronium-based activators such as HBTU and HATU can significantly reduce the occurrence of several side reactions associated with peptide coupling (**Scheme 1.7.**). One such undesirable process is the formation of N-carboxyanhydrides (NCAs), which can occur when the carboxylic function is excessively activated – for instance, when it is present as an acyl chloride – leading to intramolecular cyclisation and loss of the formation of the reactive species. In this context, the oxazolones acts as a transient intermediate in the formation of NCAs, arising from the intramolecular condensation of activated amino acids or dipeptides. These cyclic intermediates are prone to racemization through enolization at the α -carbon, particularly when activation is prolonged or nucleophilic additives such as HOBt are absent, as these efficiently trap the acyl intermediate to form a more stable active ester (Scheme 1.7, A).[141]

Another common side reaction is the formation of diketopiperazine (DKP), which is typically observed in dipeptides esters or when the terminal amine attacks the preceding activated carbonyl. Its formation is favoured when the peptide bond adopts a *cis* configuration, particularly in sequences containing proline or N-methylated amino acids (Scheme 1.7, B). Finally, when uronium salts are used, guanidine-type by-products may result from the direct reaction between the uronium reagent and the nucleophilic amine of the peptide chain. This side process usually occurs upon slow pre-activation of the carboxylic acid or when excess activator is used (Scheme 1.7, C).[141]

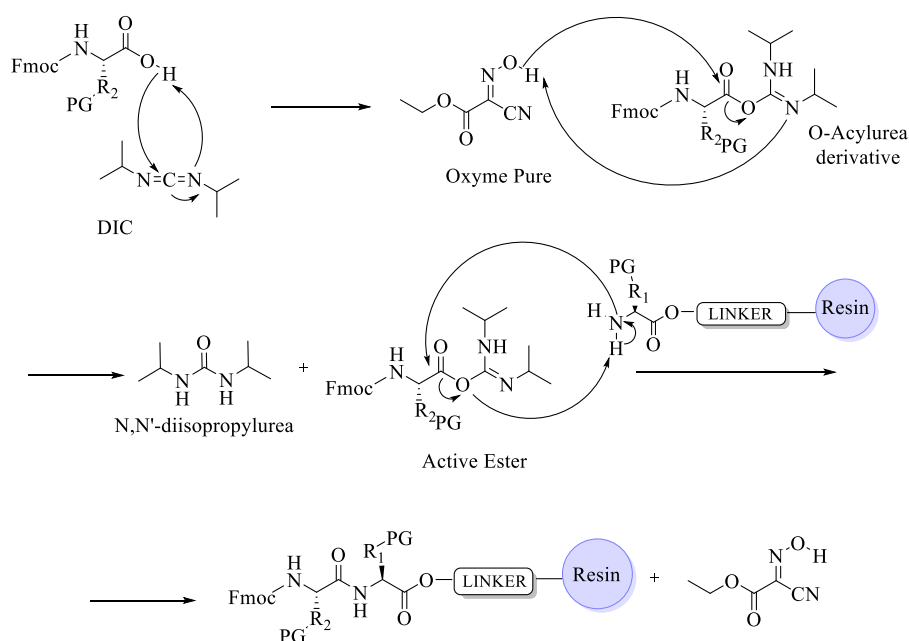


Scheme 1.7. Side reactions in peptide synthesis.

Under the optimised conditions employed – rapid *in situ* activation and controlled reagent stoichiometry – these pathways were suppressed, ensuring coupling efficiency and preservation of stereochemical integrity.

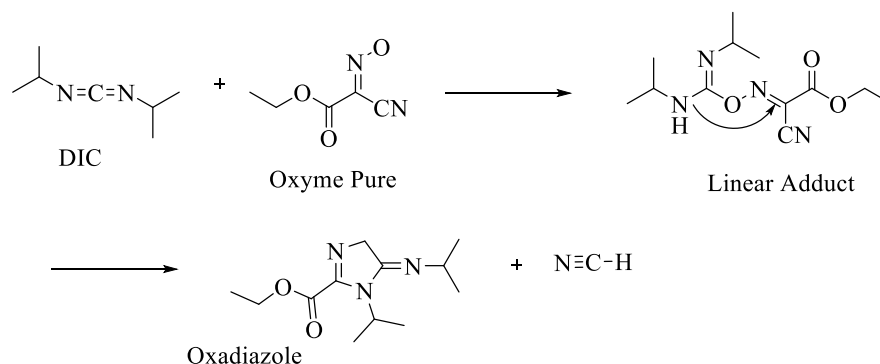


For PNA synthesis, where the backbone lacks chiral centres and differs electronically from the peptide framework, activation was achieved through carbodiimide-mediated condensation using *N,N*-diisopropylcarbodiimide (DIC) in conjunction with the safer additive Oxyma Pure (ethyl-2-cyano-2-(hydroxyimino)acetate). This system replaces the traditional HBTU/HOBt while maintaining high coupling efficiency and low racemisation potential. Mechanistically, DIC activates the carboxylic acid by forming an *O*-acylisourea intermediate, which then reacts with Oxyma Pure to produce a reactive oxyme ester that acts as the acylating agent. The electron-withdrawing cyano group of Oxyma Pure stabilises the developing negative charge during the transition state, ensuring clean amide or amide-like bond formation under mild conditions (Scheme 1.8).[142]



Scheme 1.8. Activation of the carboxylic group.

Nevertheless, DIC and Oxyma Pure may undergo a competing side reaction, whereby Oxyma Pure reacts directly with DIC to form a linear adduct. This adduct may then undergo intramolecular 5-endo-trig cyclisation to produce a 1,2,4-oxadiazole, accompanied by the formation of hydrogen cyanide (HCN). Although this reaction occurs only at trace levels (typically below 1% conversion of Oxyma Pure), it can be minimised further by optimising the sequence of reagent addition. Empirical and mechanistic studies have demonstrated that pre-activating the Fmoc-amino acid with DIC alone for 2-5 minutes, then adding it to resin and subsequently introducing Oxyma Pure, almost completely suppresses oxadiazole formation and prevents unproductive DIC consumption, with less than 0.2% of side products detected. The reactive pathway can therefore be described as a stepwise activation – transfer mechanism, whereby the acylisourea acts as a short-lived electrophile that is intercepted by the nucleophile on the resin rather than by Oxyma Pure itself (Scheme 1.9).[142]



Scheme 1.9. Side reactions during SPPS.

Overall, the activation strategies adopted in this work combine distinct yet complementary principles: uronium salts such as HBTU/HOBt and HATU, offer tuneable intramolecular catalysis and high coupling reactivity, which is ideal for peptide synthesis and functionalisation. In contrast, the DIC/Oxyma Pure system provides a safe and efficient approach, which is particularly suited to PNA assembly. Deliberately choosing each activation mode ensures the optimal electrophilicity of the carbonyl centre, controlled reaction kinetics and the preservation of stereochemical and chemical integrity throughout all coupling steps.

2.1.5. Cleavage from the resin

During the final cleavage step, treating the peptide or PNA bound to the resin with an acidic cleavage cocktail simultaneously releases the target molecule from the solid support and removes the side chain protecting groups.[143] This process can generate transient reactive intermediates, such as carbocations, which may induce undesired covalent modifications to nucleophilic residues. To suppress these side reactions, nucleophilic additives known as scavengers are incorporated into the trifluoroacetic acid (TFA) cleavage mixture. These compounds act as chemical traps for reactive species, effectively minimising acid-induced side reactions.[144]

Therefore, scavengers such as water, phenol, EDT, and thioanisole are added to neutralise these electrophiles. However, with the development of improved linker systems and acid-labile protecting groups, the need for strongly nucleophilic and malodorous scavengers such as EDT and thioanisole in sequences containing cysteine residues has greatly decreased. In most cases, these can be replaced by the less toxic trialkylsilanes (*e.g.* TIS, triisopropylsilane) or selective thiol-based agents (*e.g.* DODT, 2,2'-dithiodiethanethiol), depending on the amino acid composition and protection scheme. The optimal cleavage time depends primarily on the nature and number of the side chain protecting groups, as well as the composition of the cleavage cocktail employed.



2.1.6. Alkyne-AntimiRNA-221-5p Synthesis

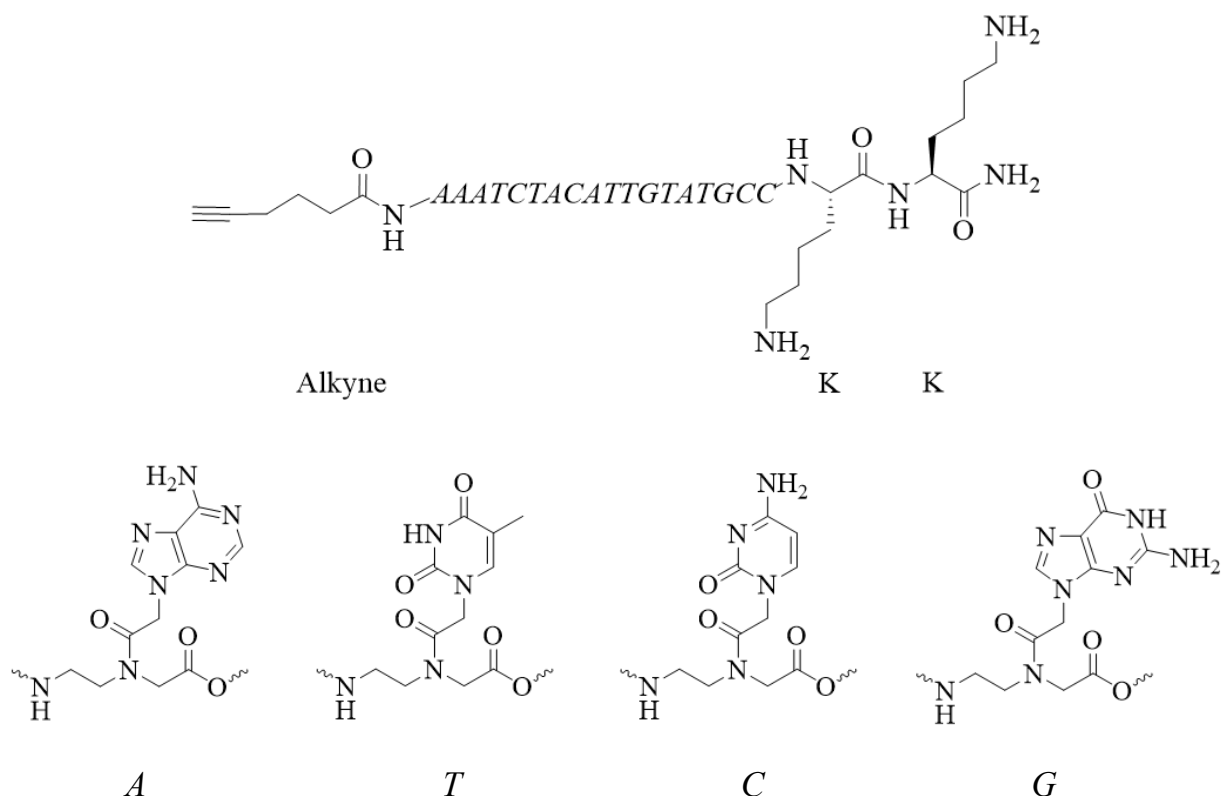


Figure 1.14. Schematic structure of Alkyne-AntimiRNA-221-5p.

The anti-miRNA-221-5p PNA (Figure 1.14) was synthesised manually using standard Fmoc chemistry and solid-phase peptide synthesis methodology. The coupling reactions employed Oxyma Pure and DIC as activation reagents, providing high coupling efficiency.

Two Lys residues were introduced at the C-terminus of the sequence from the outset to enhance aqueous solubility and compensate for the intrinsic hydrophobicity and aggregation tendency commonly observed in PNA oligomers. The 5-hexynoic acid bearing an alkyne group was also incorporated at the N-terminus to enable subsequent Cu(I)-catalysed azide-alkyne cycloaddition (CuACC) conjugation reactions with azide-functionalised CPPs.

During synthesis, Fmoc deprotection was initially carried out using just 20% v/v piperidine solution in DMF. Subsequently, from the fourteenth monomer onwards – corresponding to a thymine residue – deprotection was performed using a 5% w/v piperazine solution in DMF in presence of 2% of DBU. This modification was introduced to mitigate aggregation phenomena arising from base-stacking interactions between nucleobases in the growing chain. Piperazine acts as a chaotropic base capable of disrupting intermolecular interactions, [145] thus preventing the formation of truncated or deleted sequences that can result from steric hindrance within aggregated resin-bound PNAs.



Due to the considerable length of the target sequence, intermediate analytical controls were performed during synthesis to monitor chain elongation and coupling efficiency. Small aliquots of the resin-bound PNA were periodically sampled and subjected to acidic cleavage using 95% TFA solution. The cleaved material was analysed by LC-ESI-MS to verify the progress of the reaction. These analyses revealed an increase in side products and deletion sequences as the chain length increased. This is consistent with partial coupling inefficiencies and aggregation events. Consequently, the crude product exhibited a lower overall yield and reduced purity, complicating the subsequent purification process (Figure 1.15, A).

To address these issues, the synthesis was repeated under optimised conditions. In addition to the modified deprotection strategy, the reaction was conducted under high dilution in DMF to limit chain aggregation. Coupling times were also extended – from 50 minutes to 75 minutes up to 90 minutes for the most sterically hindered nucleobases guanine – to ensure complete coupling. These adjustments led to a significant improvement in the synthetic yield and the purity of the final product (Figure 1.15, B).

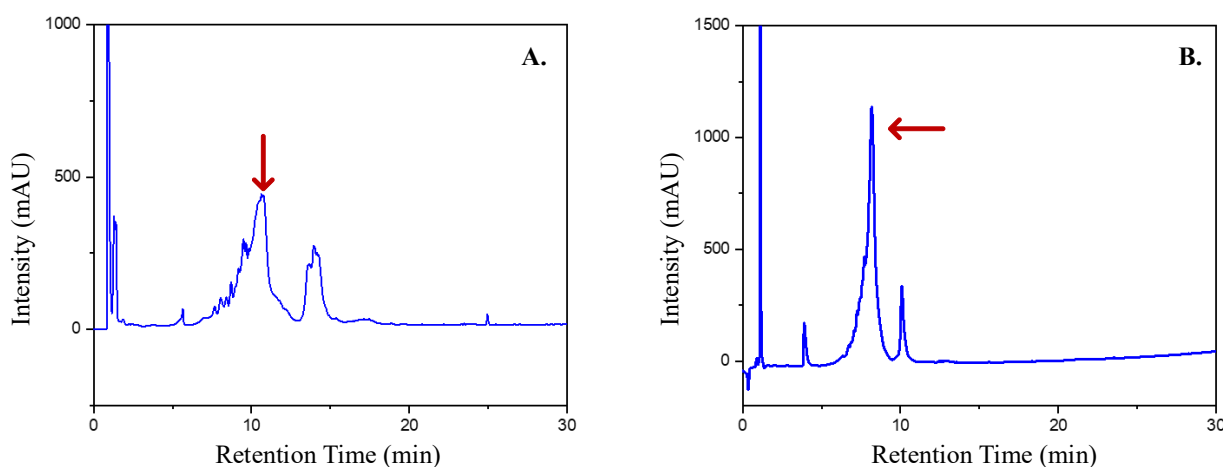


Figure 1.15. *A) Analytical HPLC chromatogram of the Alkyne-PNA-KK-NH₂ crude, after the first synthesis, recorded using a Kinetex C18 column, 1 mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 50% over 30 minutes. The product peak is indicated by the arrow. B) Analytical HPLC chromatogram of the Alkyne-PNA-KK-NH₂ crude, after the optimised synthesis, recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes. The product peak is indicated by the arrow.*

The crude PNA was purified using preparative reverse-phase HPLC on a DIASOGEL SP120 C18 column (300 Å, 10µm) with a binary solvent gradient ranging from 10% to 40% solvent B over 30 minutes. Fractions corresponding to the main product were collected and lyophilised. The identity and purity of the final compound were confirmed by analytical HPLC (Figure 1.16) and ESI-MS analysis (Figure 1.17).

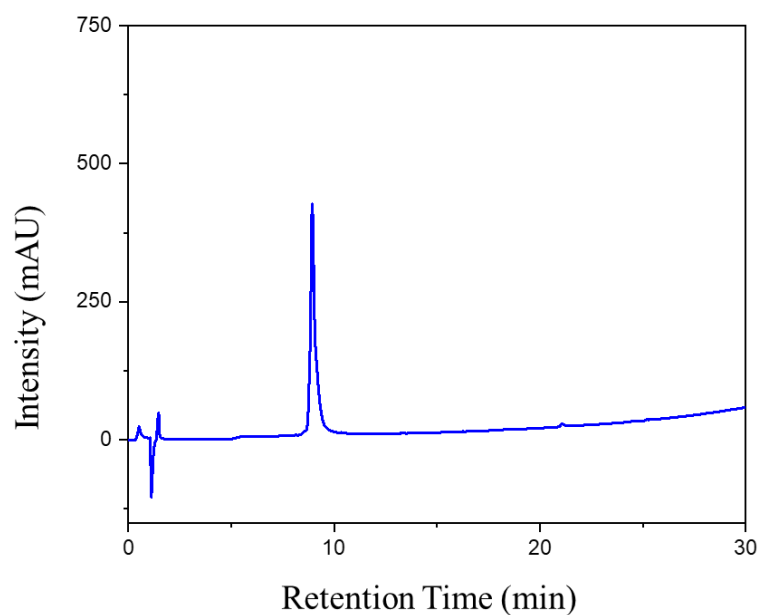


Figure 1.16. Analytical HPLC chromatogram of the Alkyne-PNA-KK-NH₂ PNA recorded using a Kinetex C18 column, 1 mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

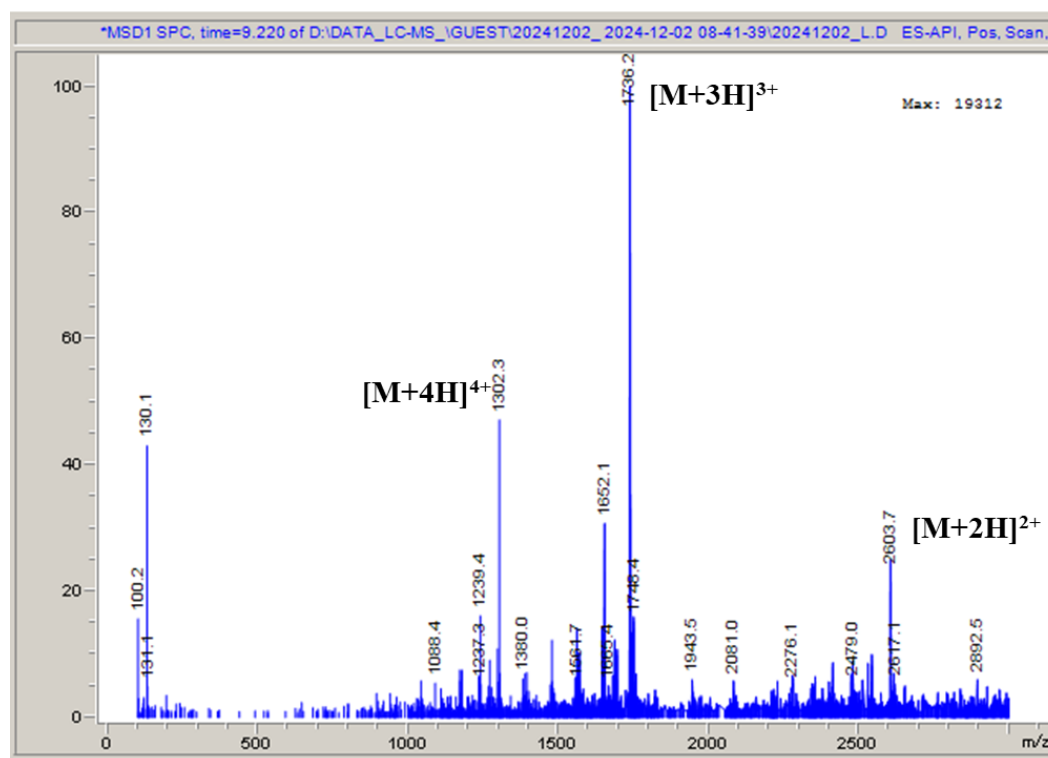


Figure 1.17. ESI-MS spectrum of the Alkyne-PNA-KK-NH₂ PNA. Molecular formula: C₂₁₂H₂₇₃N₁₀₅O₅₇, [M+H]¹⁺ Calculated: 5204.7 Da; Experimental mass: [M+2H]²⁺: 2603.2 Da, [M+3H]³⁺: 1735.8 Da, [M+4H]⁴⁺: 1302.1 Da.

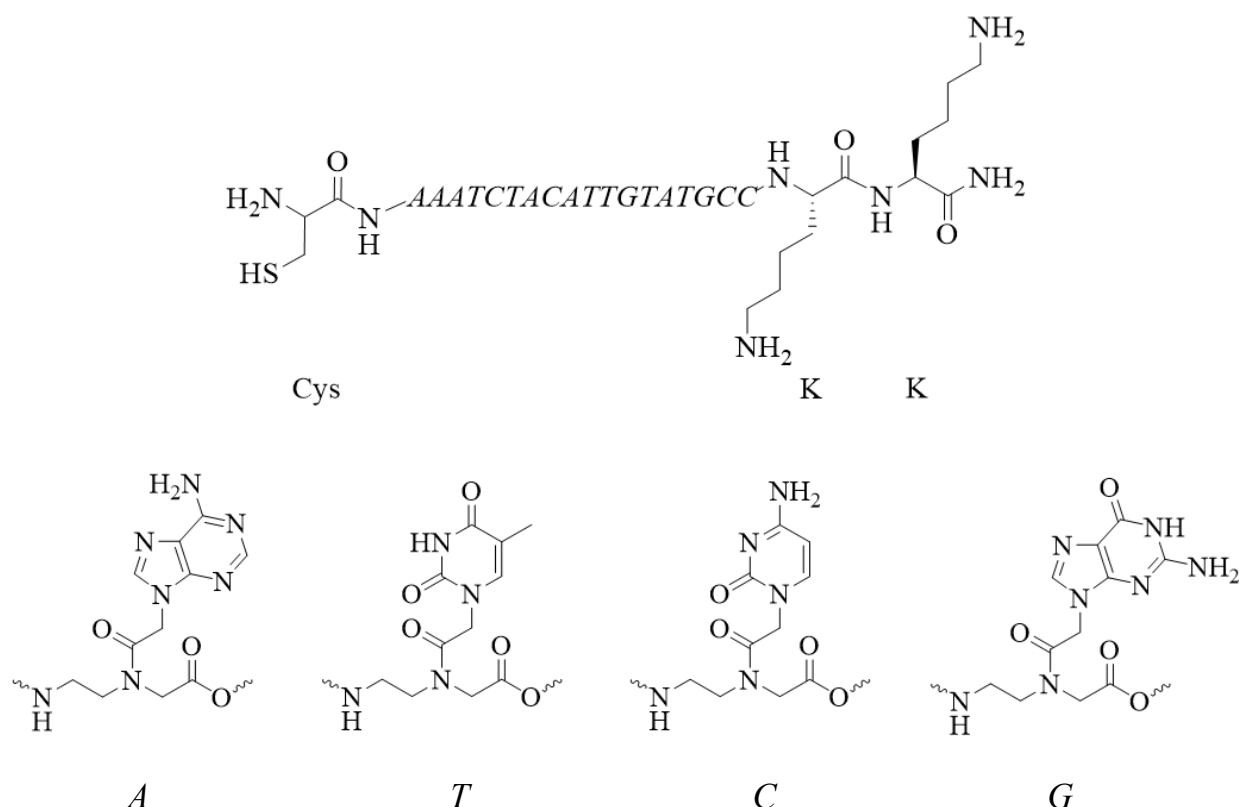


Figure 1.18. Schematic Structure of Cys-AntimiRNA221-5p.

The Cys-anti-miRNA-221-5p (Figure 1.18) was synthesised manually using SPPS, following the optimised procedure that had previously been established for the alkyne-functionalised analogue. All the same precautions were adopted as in the previous synthesis, including the use of Oxyma Pure/DIC as coupling reagents, high dilution in DMF to minimise aggregation, and the use of a solution 5% w/v piperazine in DMF in presence of 2% DBU. Extended coupling times were also used for guanine residues to ensure complete reaction.

In this case, the Boc-protected cysteine was introduced at the N-terminus in order to generate the required free thiol functionality for subsequent disulfide bridge formation with cysteine-containing CPPs. The Boc group was chosen to eliminate the need for an additional Fmoc-deprotection step, since it is removed directly under acidic TFA cleavage conditions.

However, analytical cleavages revealed the formation of two major species: the expected PNA and a secondary product displaying a +162 Da mass shift. This substantially reduced the reaction yield.

Initially, the possibility of a contribution from the protecting groups used in PNA synthesis was considered. However, this hypothesis was dismissed for two main reasons. Firstly, the Bhoc protecting group would generate adducts with different mass increments (~211 Da) and therefore cannot account for the +162 Da shift. Secondly, no species with a mass increment of +162 Da was



observed during the stepwise analytical cleavages routinely performed throughout PNA synthesis, indicating that this modification is neither intrinsic to the PNA monomers nor to their deprotection chemistry. Therefore, this side product is likely to be associated with unwanted alkylation events involving the Boc-protecting cysteine during TFA cleavage. The *t*-butyl cation released upon Boc deprotection can react with nucleophilic sites within the molecule, particularly the cysteine thiol, leading to the formation of S-alkylated adducts.

Cleavage was performed using both DODT (Figure 1.17, A) and EDT (Figure 1.17, B) as scavengers, and identical results were obtained in both cases. This confirms that the side reactions originated from the intrinsic instability of Boc-Cys(Trt)-OH, rather than insufficient scavenging. For subsequent cleavages, DODT was preferred over EDT due to its reduced odour.

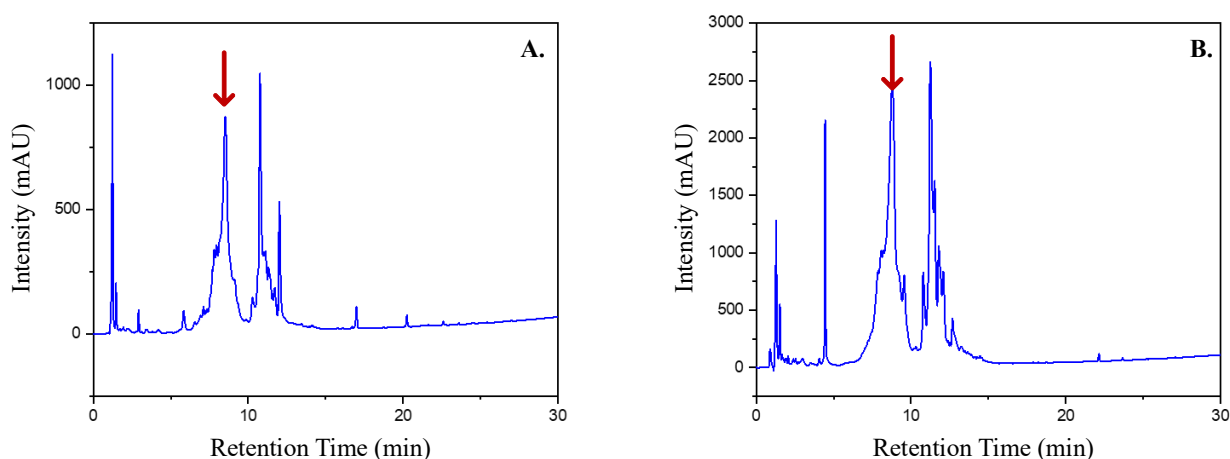


Figure 1.19. A) Analytical HPLC chromatogram of the Cys-PNA-KK-NH₂ crude, after the cleavage in presence of DODT, recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes. The product peak is indicated by the arrow. B) Analytical HPLC chromatogram of the Cys-PNA-KK-NH₂ crude, after the cleavage in presence of EDT, recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes. The product peak is indicated by the arrow.

The crude product was purified by preparative reverse phase HPLC using a DIASOGEL SP120 C18 column (300 Å, 10µm) with a binary solvent gradient ranging from 10% to 40% solvent B over 30 minutes. Fractions corresponding to the main product were collected and lyophilised. The identity and purity of the final compound were confirmed by analytical HPLC (Figure 1.20) and ESI-MS analysis (Figure 1.21).

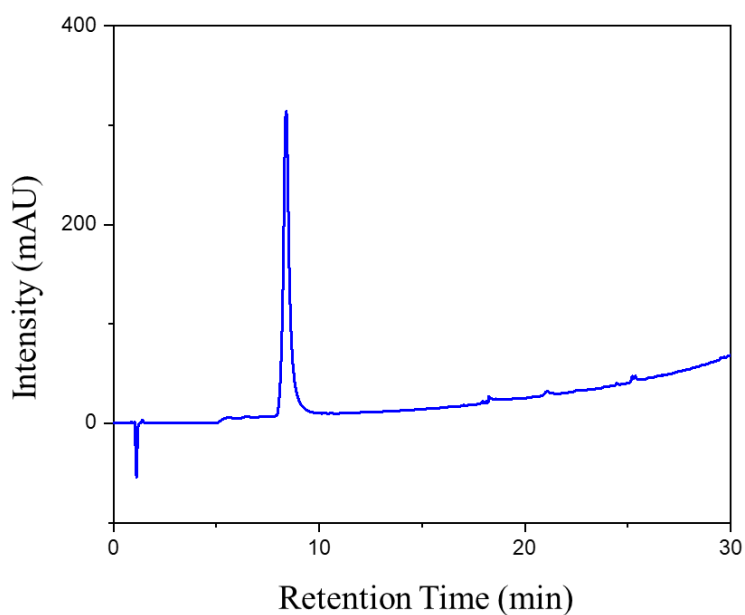


Figure 1.20. Analytical HPLC chromatogram of the Cys-PNA-KK-NH₂ PNA recorded using a Kinetex C18 column, 1 mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

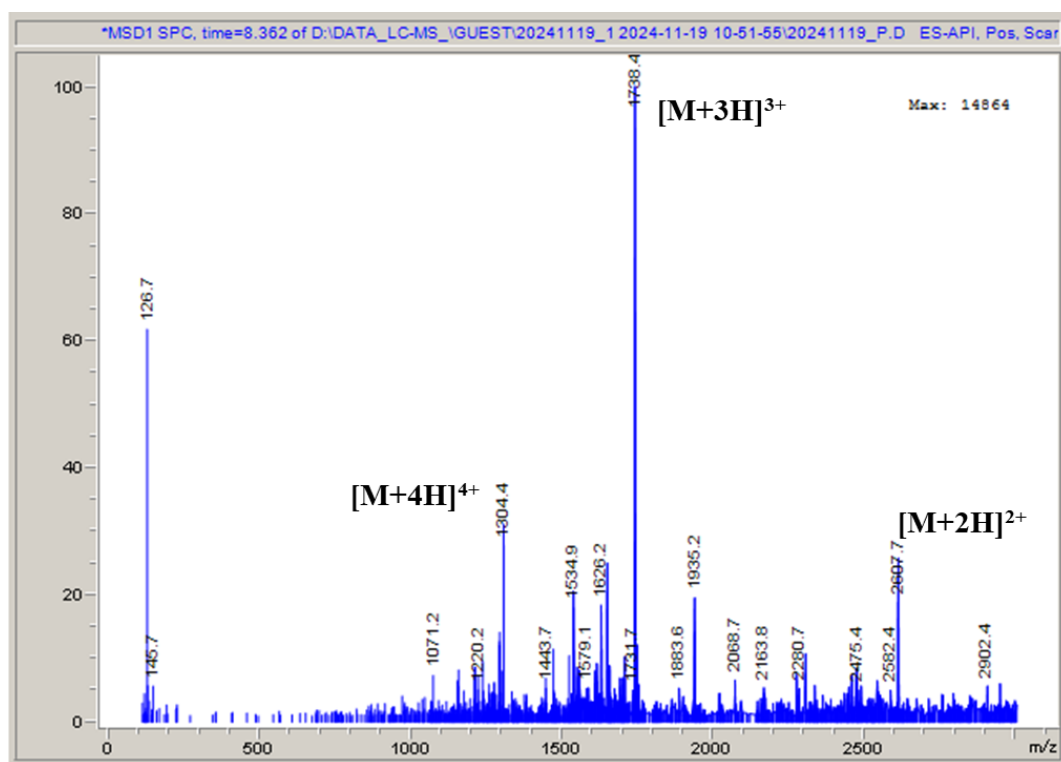


Figure 1.21. ESI-MS spectrum of the Cys-PNA-KK-NH₂ PNA. Molecular formula: C₂₀₉H₂₇₂N₁₀₆O₅₇S, [M+H]¹⁺ Calculated: 5212.8 Da; Experimental mass: [M+2H]²⁺: 2606.7 Da, [M+3H]³⁺: 1738.1 Da, [M+4H]⁴⁺: 1303.8 Da.



2.1.8. Scramble PNA Sequence Synthesis

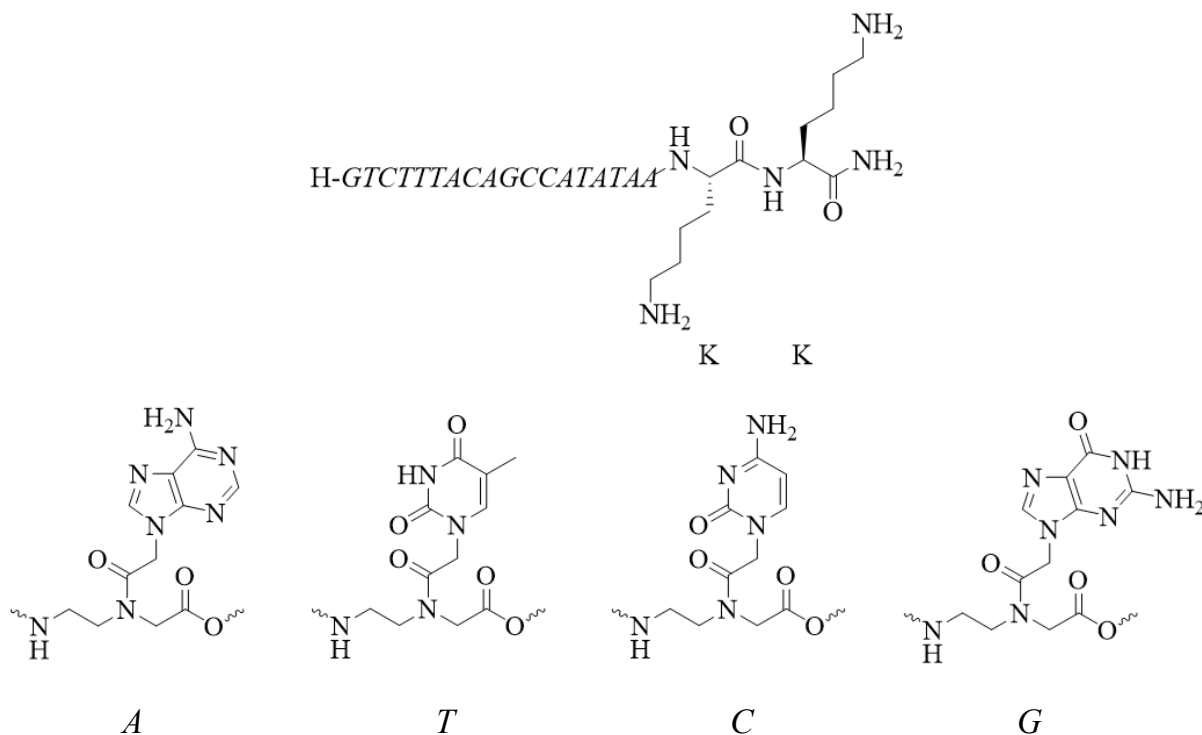


Figure 1.22. Schematic Structure of the Scramble PNA.

The scrambled PNA sequence (Figure 1.22), which was designed as a negative control for the biological assays, was synthesised manually using SPPS under the optimised conditions previously described for the anti-miRNA-221-5p PNA. The same precautions were taken through the synthesis, including using Oxyma Pure/DIC as the coupling reagents, using high dilution in DMF and the Fmoc deprotection solution 5% w/v piperazine in DMF in presence of 2% of DBU to minimise aggregation, and using extended coupling times to ensure complete reactions.

The synthesis proceeded smoothly, with no significant side reactions or deletion sequences occurring. The scramble PNA was then purified using preparative reverse-phase HPLC using a DIASOGEL SP120 C18 column (300 Å, 10µm) with a binary solvent gradient ranging from 10% to 40% solvent B over 30 minutes. Fractions corresponding to the main product were collected and lyophilised. The identity and purity of the final compound were confirmed by analytical HPLC (Figure 1.23) and ESI-MS analysis (Figure 1.24).

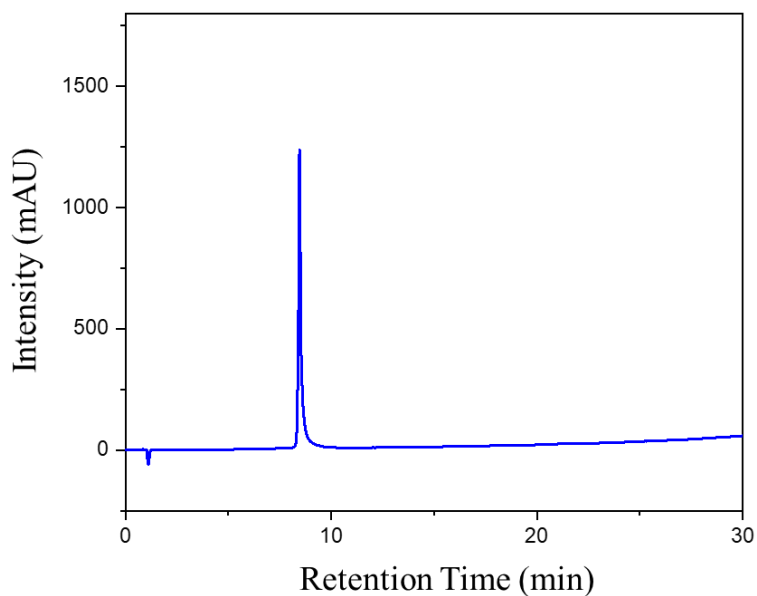


Figure 1.23. Analytical HPLC chromatogram of the H-Scramble-KK-NH₂ PNA recorded using a Kinetex C18 column, 1 mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

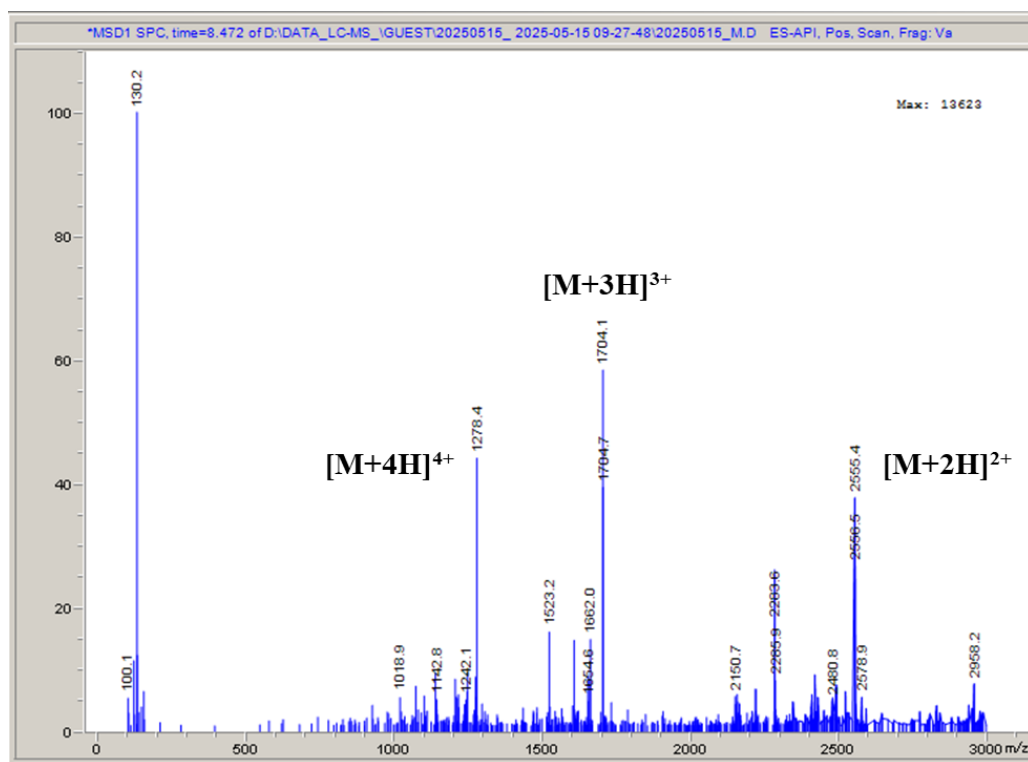


Figure 1.24. ESI-MS spectrum of the H-Scramble-KK-NH₂ PNA. Molecular formula: C₂₀₆H₂₆₇N₁₀₅O₅₆, [M+H]¹⁺ Calculated: 5108.8 Da; Experimental mass: [M+2H]²⁺: 2554.8 Da, [M+3H]³⁺: 1703.5 Da, [M+4H]⁴⁺: 1277.9 Da.



2.1.9. CF-Tat-Ahx-N₃-NH₂ Synthesis

The H-Tat-Ahx-N₃-NH₂ peptide was synthesised using automated SPPS to optimise reaction time and ensure high reproducibility. To improve coupling efficiency and overall yield, double coupling reaction were performed for each amino acid during chain elongation.

An azide group was introduced at the C-terminus of the Tat sequence to enable subsequent Cu(I)-catalysed azide-alkyne cycloaddition with the alkyne-functionalised PNA. To enable fluorescence-based monitoring of cellular uptake, the N-terminus of the CPP was labelled with carboxyfluorescein (CF), using HATU as the coupling reagent and DIPEA as the base. The reaction was then left to proceed overnight at room temperature.

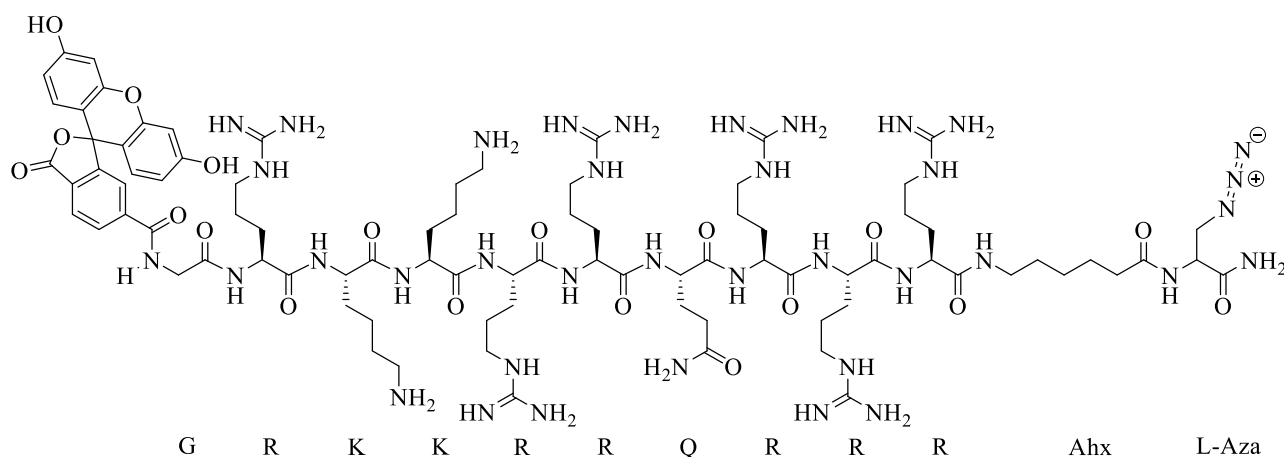


Figure 1.25. Structure of Structure of CF-Tat-Ahx-N₃-NH₂ peptide.

The CF-Tat-Ahx-N₃-NH₂ peptide (Figure 1.25) was then purified by low-pressure liquid chromatography using a Biotage Isolera Prime system equipped with a SNAP Ultra C18 reversed-phase column with an elution gradient 0-50% B over 50 minutes. The identity and purity of the final peptide were confirmed by HPLC-MS analysis (Figure 1.26, Figure 1.27), which showed a single major peak corresponding to the expected molecular mass.

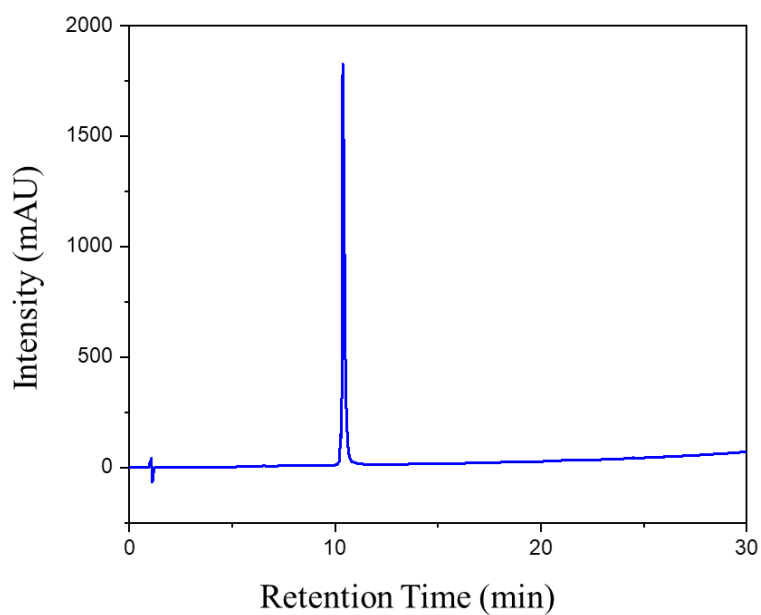


Figure 1.26. Analytical HPLC chromatogram of CF-Tat-N₃-NH₂ recorded using a Kinetex C18 column, 1 mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

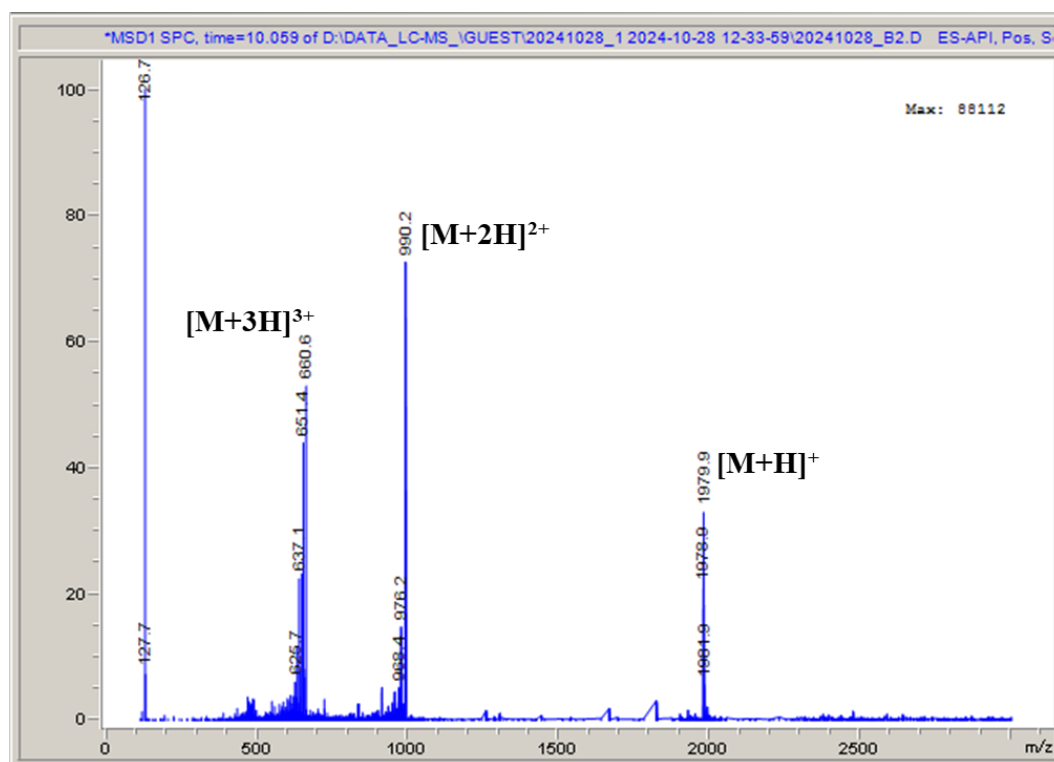


Figure 1.27. ESI-MS spectrum of CF-Tat-N₃-NH₂. Molecular formula: C₈₅H₁₃₅N₃₇O₁₉, [M+H]¹⁺ Calculated: 1979.1 Da; Experimental mass: [M+H]⁺: 1977.9 Da, [M+2H]²⁺: 989.8 Da, [M+3H]³⁺: 660.3 Da.



2.1.10. CF-PNT-Ahx-N₃-NH₂ Synthesis

The H-PNT-Ahx-N₃-NH₂ peptide (Figure 1.28) was synthesised using automated SPPS, under the same optimised conditions as CF-Tat-Ahx-N₃-NH₂. This included double coupling reactions to ensure complete chain assembly. An azide moiety was introduced at the C-terminus to enable click chemistry conjugation with alkyne-bearing PNAs.

For cellular tracking experiments, the N-terminal amine was reacted with CF using HATU/DIPEA activation. The reaction mixture was stirred overnight at room temperature to achieve complete labelling.

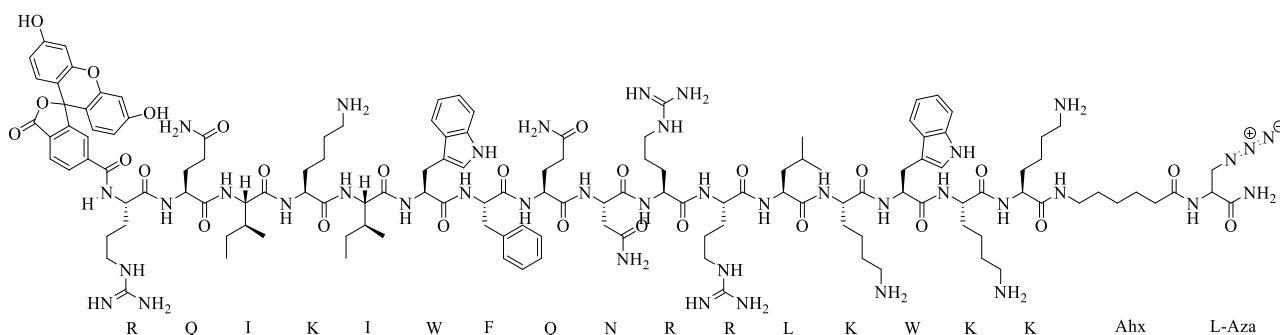


Figure 1.28. Structure of CF-PNT-Ahx-N₃-NH₂ peptide.

Purification was performed using a Biotage Isolera Prime system and a SNAP Ultra C18 reversed-phase column with low-pressure liquid chromatography with an elution gradient 0-50% B over 50 minutes. The CF-PNT-Ahx-N₃-NH₂ peptide obtained was characterized by HPLC-MS (Figure 1.29, Figure 1.30) and was used for subsequent click conjugation reactions with PNAs.

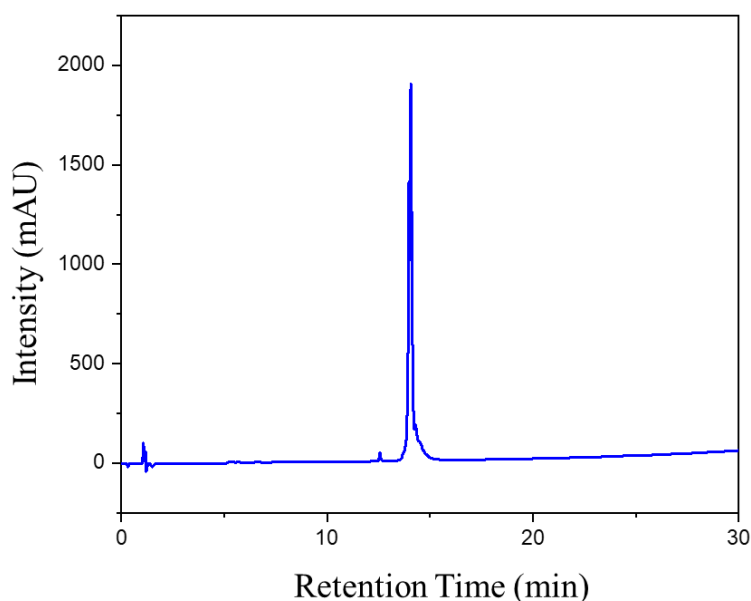


Figure 1.29. Analytical HPLC chromatogram of CF-PNT-Ahx-N₃-NH₂ recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

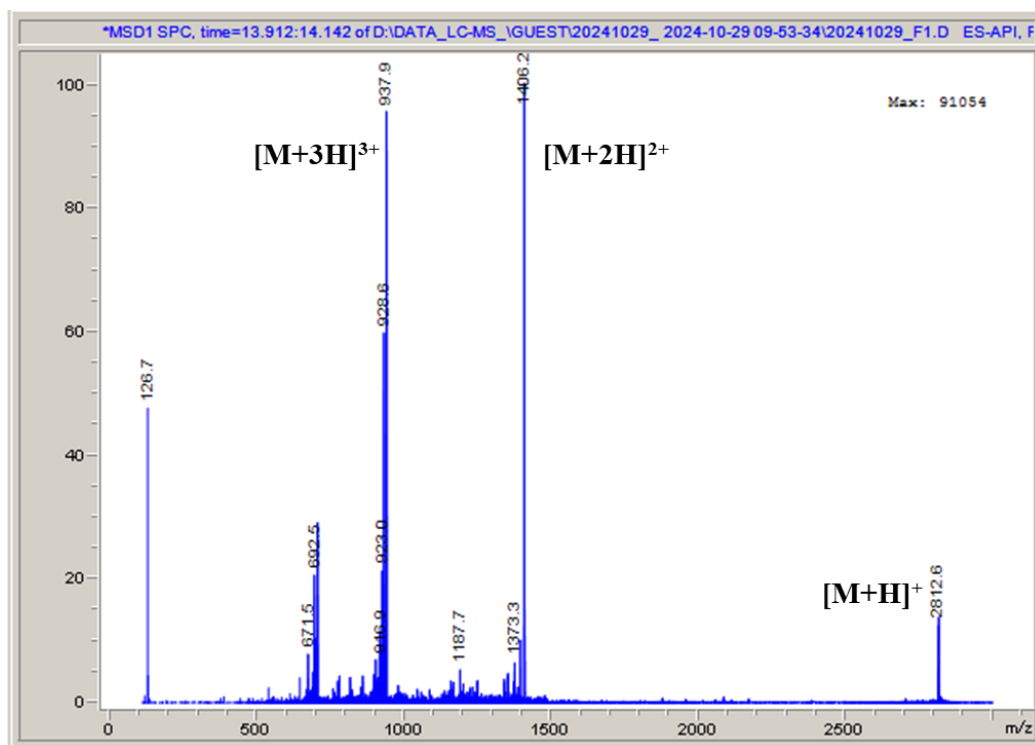


Figure 1.30. ESI-MS spectrum of CF-PNT-Ahx-N₃-NH₂. Molecular formula: C₁₃₅H₁₉₅N₃₉O₂₈, [M+H]¹⁺ Calculated: 2813.3 Da; Experimental mass: [M+H]¹⁺: 2811.6 Da, [M+2H]²⁺: 1405.7 Da, [M+3H]³⁺: 937.6 Da.

2.1.11. CF-P14-Ahx-N₃-NH₂ Synthesis

The H-P14-Ahx-N₃-NH₂ peptide (Figure 1.31) is a hybrid cell-penetrating peptide (CPP), combining the structural features of Tat and Penetratin to improve the efficiency of internalisation while maintaining stability and solubility. The peptide was synthesised using automated SPPS, with double coupling steps applied to each residue to maximise yield.

An azide group was introduced at the C-terminus to enable click chemistry coupling with alkyne-modified PNAs. Meanwhile, the N-terminus was functionalised with CF using HATU/DIPEA as activation reagents. The reaction was then left overnight at room temperature to ensure complete incorporation of the fluorophore for fluorescence-based cellular visualisation.

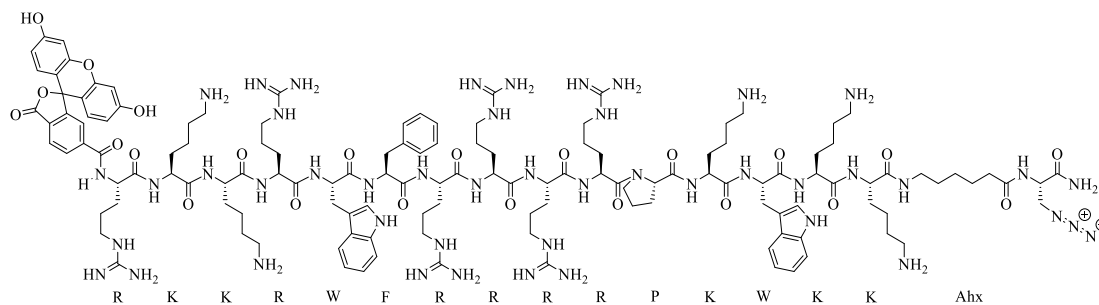


Figure 1.31. Structure of CF-P14-Ahx-N₃-NH₂ peptide.





2.1.12. Optimization and Change of Solid Support

Several trials were performed during the optimisation of Tat-derived synthesis (and the subsequent synthesis of PNT-derived synthesis) to evaluate the effect of different spacers on coupling efficiency and product stability. Three test sequences were synthesised on Rink Amide resin:

- Fmoc-RRR-Ahx-Cys-NH₂ (Figure 1.34, A)
- Fmoc-RRR-βAla-Cys-NH₂ (Figure 1.34, B)
- Fmoc-RRR-Cys-NH₂ (Figure 1.34, C)

Manual synthesis was used for these short sequences.

However, synthesising all three peptides on Rink Amide resin proved problematic. LC-MS analyses revealed multiple fragmentation peaks and low-mass side products, indicating incomplete coupling and side-chain degradation during chain elongation or cleavage. Notably, the expected molecular peaks were accompanied by minor species corresponding to Fmoc or cysteine-related fragments. These issues were consistent across all tested spacers, implying that the problem originated primarily from the solid support rather than the spacer itself.[146]

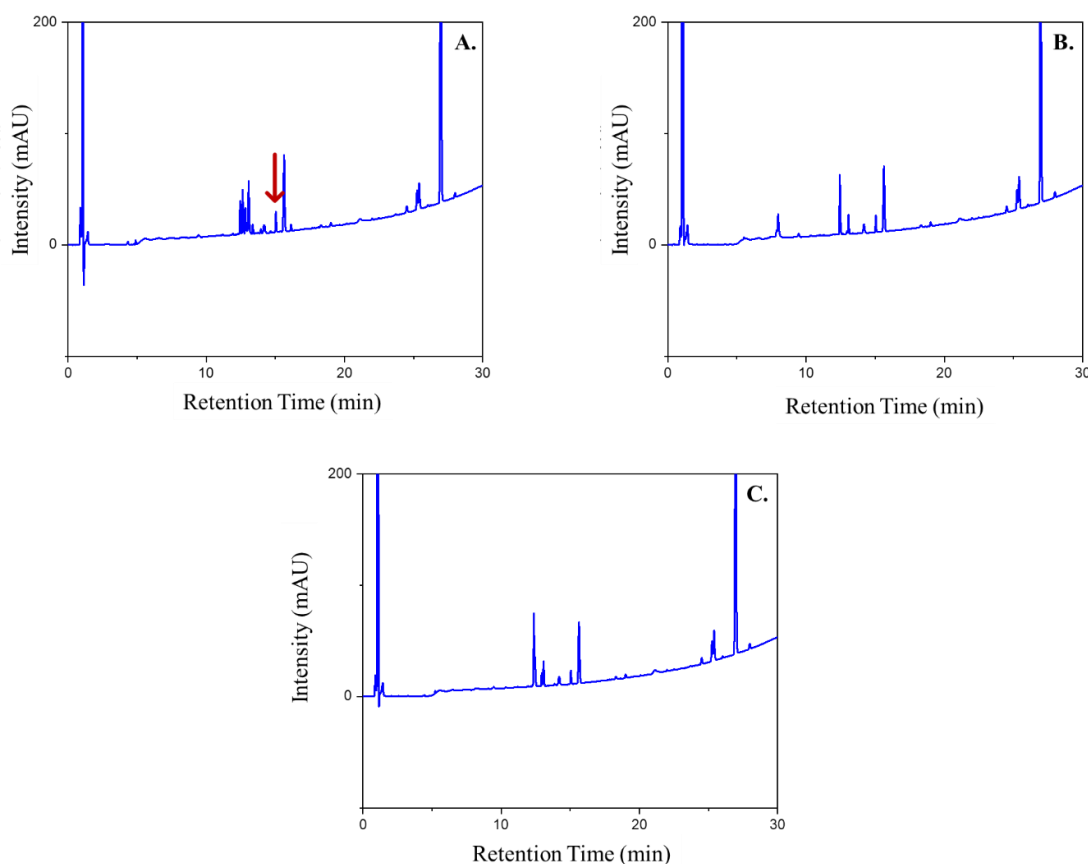


Figure 1.34. Analytical HPLC chromatogram of the A) Fmoc-RRR-Ahx-C-NH₂ B) Fmoc-RRR-βAla-C-NH₂ C) Fmoc-RRR-C-NH₂ crude, after the cleavage, recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes. The product peak, if present, is indicated by the arrow.



Consequently, the synthesis was switched to 2-CTC resin. This change improved the synthetic outcome and facilitated the subsequent preparation of the Cys(NTP) peptides.

2.1.10.1. CF-Tat-Ahx-Cys(TNP)-OH Synthesis

The H-Tat-Ahx-Cys-OH peptide was synthesised using automated SPPS to optimise reaction time and improve overall yield. To ensure complete coupling efficiency, double coupling steps were performed for each amino acid during chain elongation.

A cysteine residue was introduced at the C-terminal end of the native Tat sequence to enable disulfide bond formation with the cargo PNAs in subsequent conjugation steps. Additionally, an aminohexanoic acid (Ahx) spacer was incorporated to provide structural flexibility and facilitate subsequent conjugation.

To enable fluorescence-based visualisation of the conjugates in cellular assays, a carboxyfluorescein (CF) moiety was coupled to the N-terminus of the peptide using HATU as the activating agent and DIPEA as the base. The coupling reaction was carried out overnight at room temperature to ensure efficient labelling of the peptide with the fluorophore.

During cleavage from the resin, the C-terminal cysteine of the peptide was activated *in situ* by adding 2,2'-dithiodipyridine (DTNP) [126] to the cleavage mixture in a 5-fold molar excess relative to the peptide. This treatment yielded the activated disulfide form of CF-Tat-Ahx-Cys-OH, which was then ready for thiol-disulfide exchange with cysteine-containing PNAs.

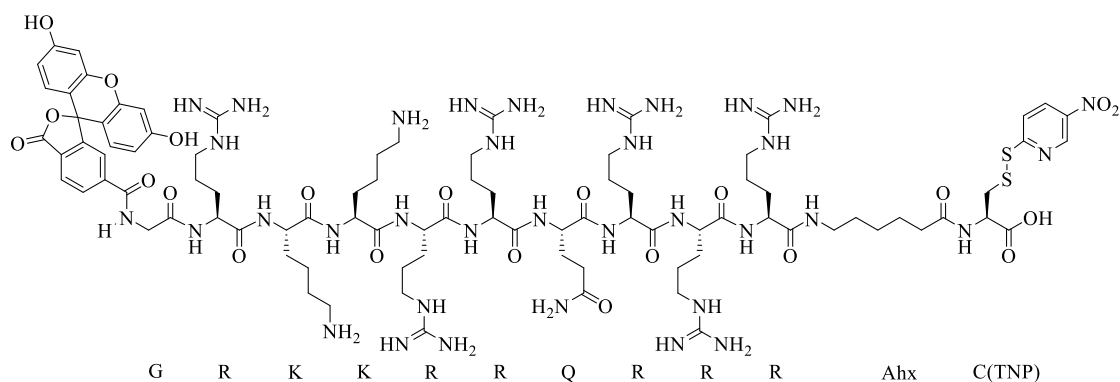


Figure 1.35. Structure of CF-Tat-Ahx-Cys(TNP)-OH peptide.

The activated CF-Tat-Ahx-Cys(TNP)-OH peptide (Figure 1.35) was obtained with good purity, as confirmed by HPLC-MS analysis (Figure 1.36, Figure 1.37). This allowed it to be used directly in conjugation reactions, without the need for further preparative HPLC purification.

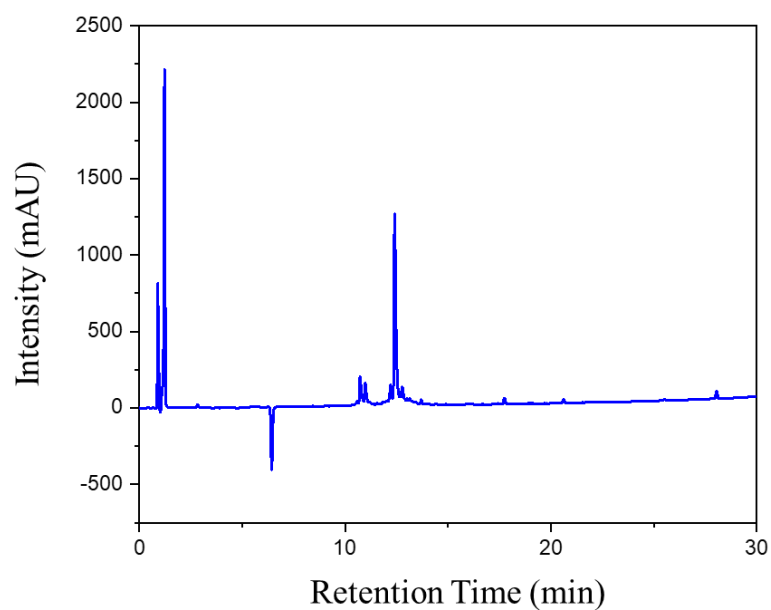


Figure 1.36. Analytical HPLC chromatogram of CF-Tat-Ahx-Cys(TNP)-OH recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

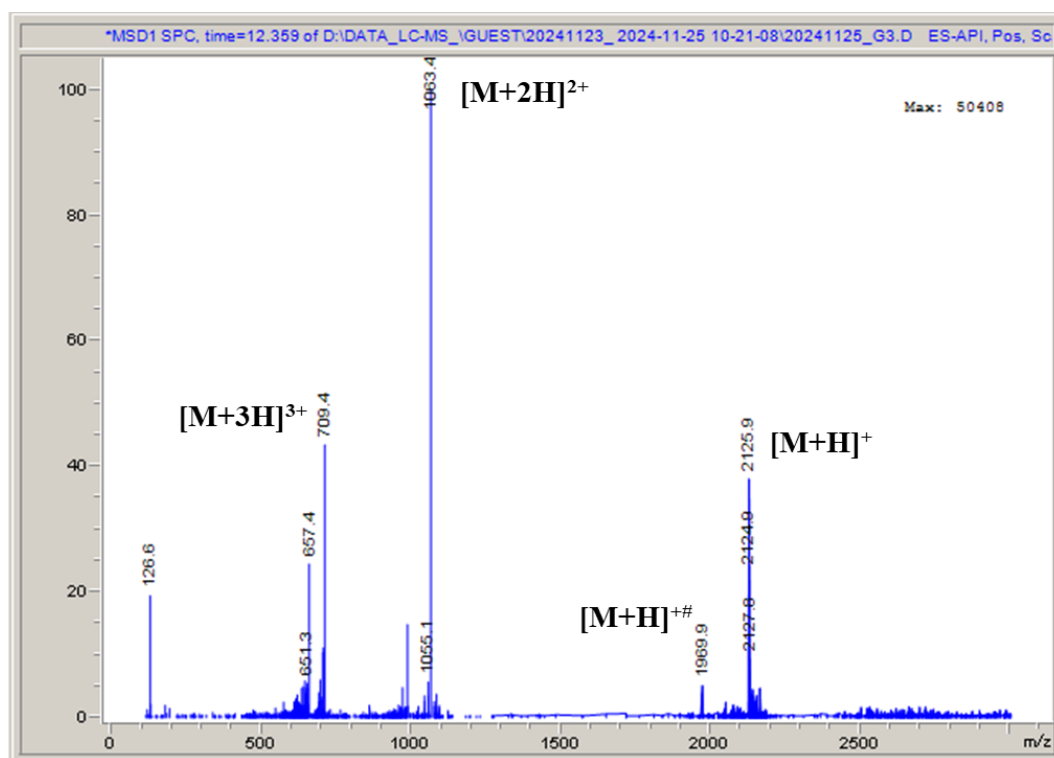


Figure 1.37. ESI-MS spectrum of CF-Tat-Ahx-Cys(TNP)-OH. Molecular formula: $C_{90}H_{137}N_{35}O_{22}S_2$, $[M+H]^+$ Calculated: 2125.3 Da; Experimental mass: $[M+H]^+$: 2124.9 Da, $[M+2H]^{2+}$: 1062.9 Da, $[M+3H]^{3+}$: 709.0 Da, $[M+H]^{\#}$: 1968.9 Da represents the peptide experimental mass after the loss of the activation on the Cys residue.



2.1.10.2. CF-PNT-Ahx-Cys(TNP)-OH Synthesis

The H-PNT-Ahx-Cys-OH peptide was synthesised using automated SPPS to ensure efficient assembly and high reproducibility. To maximise coupling yield, double coupling reactions were performed for each amino acid.

A cysteine residue was introduced at the C-terminus of the natural Penetratin sequence to enable disulfide bond formation with cysteine-containing cargo PNAs. The Ahx spacer was also added to provide flexibility and spatial separation between the peptide backbone and the reactive thiol group, thereby facilitating the disulfide exchange reaction.

For subsequent cellular uptake and imaging studies, the N-terminus of the peptide was labelled with CF. This was achieved by reacting the resin-bound peptide with CF in the presence of HATU and DIPEA, with the reaction proceeding overnight at room temperature.

During cleavage from the resin, the C-terminal cysteine was activated in situ by adding DTNP to the cleavage cocktail at a molar ratio of 5:1 relative to the peptide. This step produced the activated CF-PNT-Ahx-Cys(TNP)-OH (Figure 1.38), which was then ready for conjugation with PNA cargo.

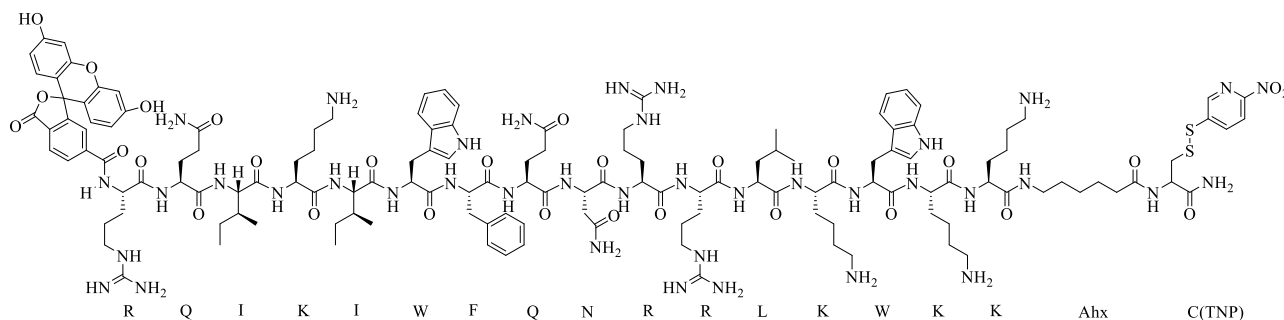


Figure 1.38. Structure of CF-PNT-Ahx-Cys(TNP)-OH peptide.

The resulting peptide was obtained with satisfactory purity, as confirmed by HPLC-MS analysis (Figure 1.39, Figure 1.40), and used directly for conjugation without the need for preparative HPLC purification.

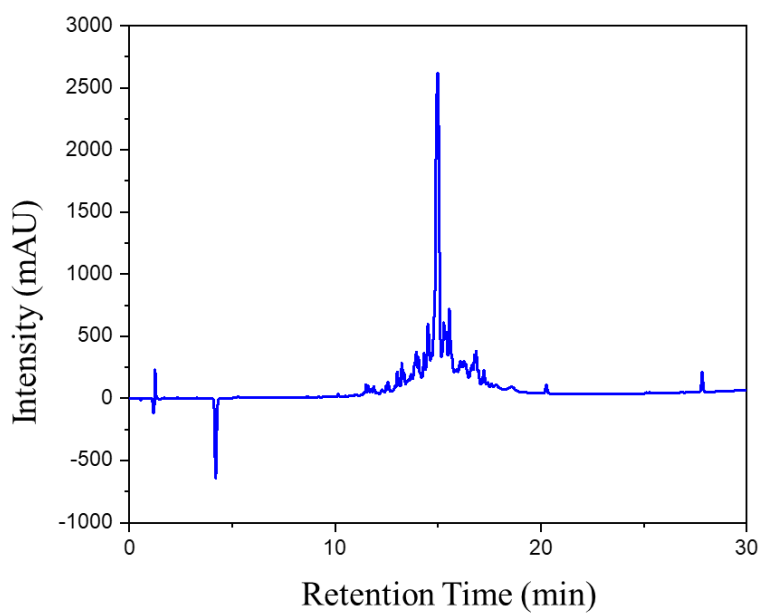


Figure 1.39. Analytical HPLC chromatogram of CF-PNT-Cys(TNP)-OH recorded using a Kinetex C18 column, 1 mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

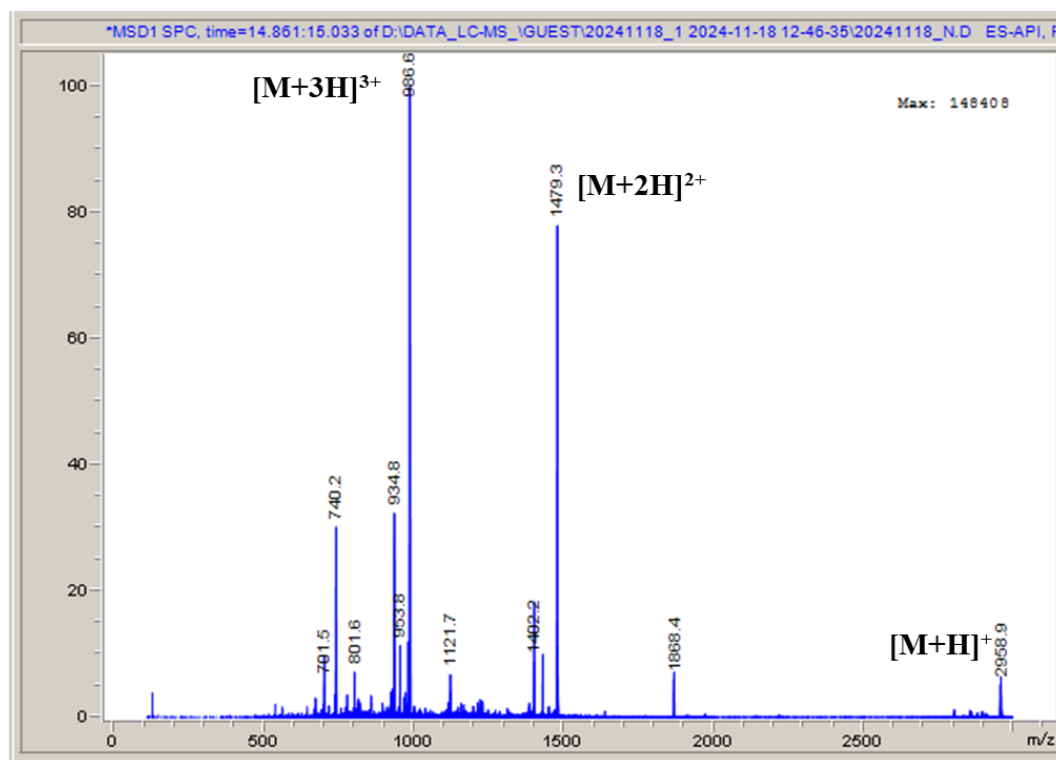


Figure 1.40. ESI-MS spectrum of CF-PNT-Cys(TNP)-OH. Molecular formula: $C_{140}H_{198}N_{38}O_{30}S_2$, $[M+H]^+$ Calculated: 2958.2 Da; Experimental mass: $[M+H]^+$: 2957.9 Da, $[M+2H]^{2+}$: 1478.8 Da, $[M+3H]^{3+}$: 986.3 Da.



2.2. Click Chemistry

Conjugated peptide-nucleic acid systems could have a wide range of applications. These applications require highly efficient and specific synthetic methodologies that can provide high yields at a low production cost. In 2003, K. B. Sharpless introduced click chemistry, a new synthetic concept. This approach involves constructing complex molecular architectures by reliably connecting small, highly functionalised building blocks. Click reactions can be a strategy for assembling peptide-nucleic acid conjugates because they are simple, reproducible and robust. According to Sharpless *et al.*, reactions that fall within the ‘click chemistry’ paradigm must fulfil the following criteria:

- Quantitative or near-quantitative yields;
- Mild and biocompatible reaction conditions (an aqueous medium at room temperature and physiological pH);
- Minimal or no by-product formation.

A prototypical example of a click reaction is the formation of a 1,2,3-triazole ring via Huisgen’s 1,3-dipolar cycloaddition of an azide with a terminal alkyne. Although azides and alkynes are thermodynamically unstable, they are kinetically inert and remarkably stable towards biomolecules under physiological conditions, showing no undesired reactivity in aqueous or reducing environments. However, the original thermal cycloaddition between these partners proceeds slowly due to the high activation barrier associated with concerted transition states.[147]

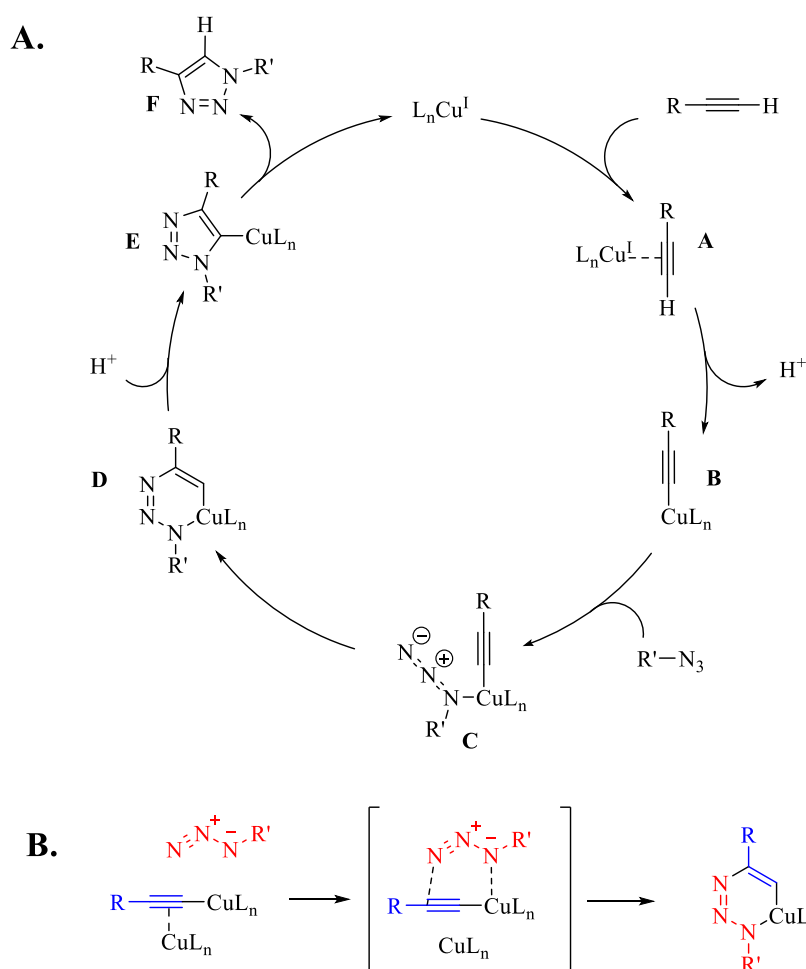
Sharpless revolutionised this process by introducing the copper(I)-catalysed azide-alkyne cycloaddition (CuAAC) reaction, which dramatically increases the reaction rate by stabilising a copper-acetylide intermediate in an aqueous medium. The CuAAC reaction proceeds regioselectively to yield 1,4-disubstituted triazoles under mild and tolerant conditions. The resulting triazole linkage is chemically and enzymatically stable; unlike amide bonds, it is not hydrolysed and unlike aromatic or heteroaromatic systems, it is resistant to oxidation and reduction. Furthermore, the geometric rigidity of triazole makes it an ideal structural linker, providing defined spatial orientation between the connected fragments.

Despite its versatility, the CuAAC reaction has one inherent limitation: the need for Cu(I) as a catalyst. Copper(I) salts are cytotoxic,[148] primarily due to the generation of reactive oxygen species and their interaction with nucleic acids and proteins. This leads to DNA strand cleavage and protein denaturation.[149] Therefore, when preparing bioconjugates for biological assays, it is essential to efficiently remove residual copper using purification techniques such as dialysis, microfiltration or preparative HPLC. To prevent oxidation of Cu(I) to Cu(II), the reaction is typically carried out under



an inert atmosphere with the addition of a reducing agent such as sodium ascorbate, which regenerates Cu(I) continuously *in situ* (Scheme 1.10, B).

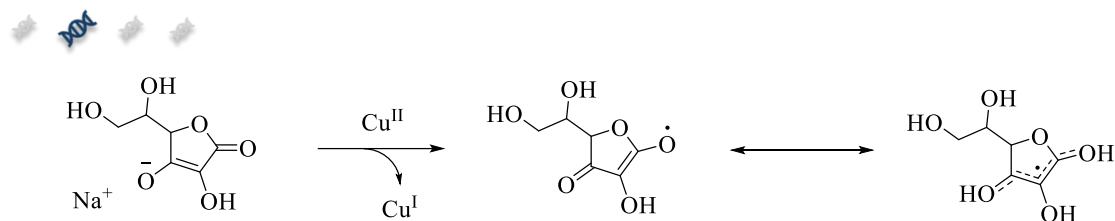
The mechanism of the Cu(I)-catalysed azide-alkyne cycloaddition (Scheme 1.10, A) involves the initial π -coordination of the terminal alkyne to Cu(I), followed by deprotonation to form a σ -acetylide intermediate. The azide then coordinates to the metal centre, aligning the reacting orbitals for nucleophilic attack. This leads to the formation of a six-membered metalacyclic intermediate, which then rearranges to yield the copper-triazolide complex. Protonolysis of this complex releases the 1,4-disubstituted 1,2,3-triazole product and regenerates the catalyst for subsequent cycles.[150] Scheme 1.10 shows the conjugation of PNA to CPP via the CuAAC catalytic cycle.



Scheme 1.10. **A)** Catalytic cycle of the CuAAC reaction to form 1,2,3-triazole **B)** formation of the metallocycle (D).

In this work, the CuAAC click reaction was used to synthesise PNA-CPP conjugates, enabling the intracellular delivery of PNA sequences. Specifically, alkyne-functionalised PNAs were reacted with azide-functionalised CPPs in a reducing aqueous medium. The reaction was catalysed by $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, which was reduced *in situ* to Cu(I) by sodium L-ascorbate. This generated ascorbyl radicals that maintained the catalytic cycle under mild, biocompatible conditions.

Results and Discussion



Scheme 1.11. Oxidation mechanism of the reducing agent sodium *L*-ascorbate.

This strategy enabled the efficient and selective formation of a stable triazole linkage between the PNA and peptide components, ensuring the structural integrity and reproducibility of the conjugation process.

2.2.1. Click Reaction of CF-Tat-(Triazole)-PNA-NH₂

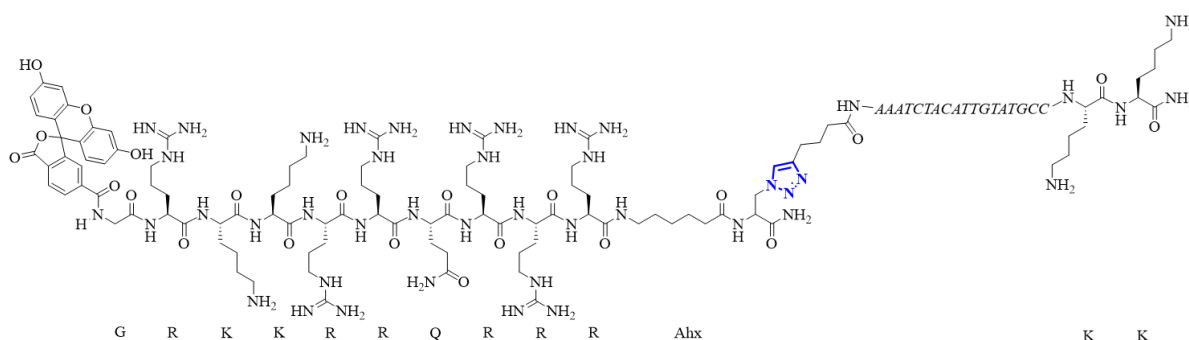


Figure 1.41. Structure of CF-Tat(Triazole)PNA-NH₂

The CF-Tat-N₃-NH₂ CPP was conjugated to the alkyne-functionalised PNA via a CuAAC reaction (Figure 1.41). Both components were dissolved in an H₂O:*t*-BuOH (1:3 v/v) mixture to ensure adequate solubility. The reaction was initiated by sequentially adding CuSO₄·5H₂O and sodium ascorbate at a molar ratio of 1:4, which allowed the *in situ* reduction of Cu(II) to Cu(I). The mixture was stirred at room temperature for 24 hours.

Analytical HPLC monitoring confirmed the formation of the desired conjugate, showing complete consumption of the CPP and the presence of residual, unreacted PNA (Figure 1.42).

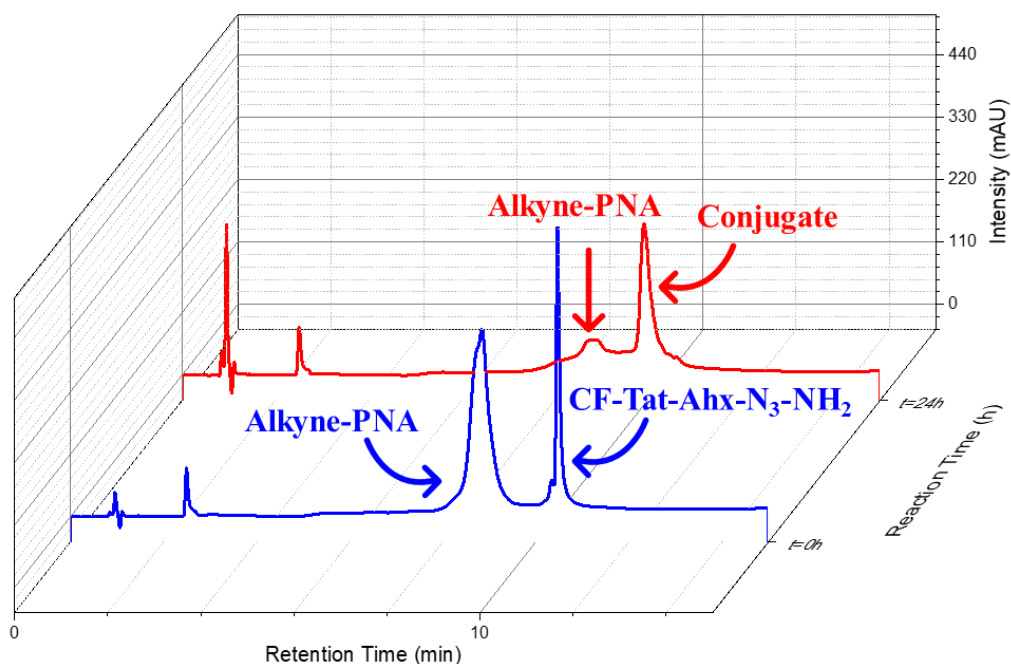


Figure 1.42. Monitoring of the click conjugation reaction between CF-Tat-N₃-NH₂ and the alkyne-functionalised PNA by analytical HPLC at different time points. The chromatogram is recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

The crude product was then purified using reverse-phase HPLC to obtain the pure conjugate (Figure 1.43, Figure 1.44).

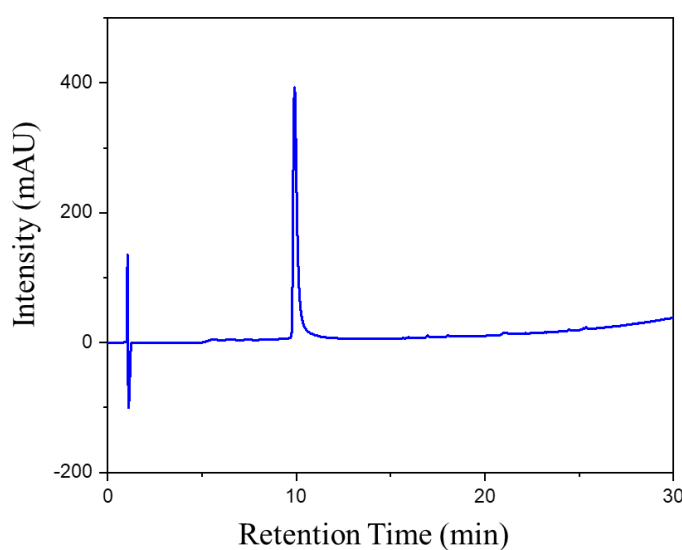


Figure 1.43. Analytical HPLC chromatogram of CF-Tat-(triazole)-PNA-NH₂ recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

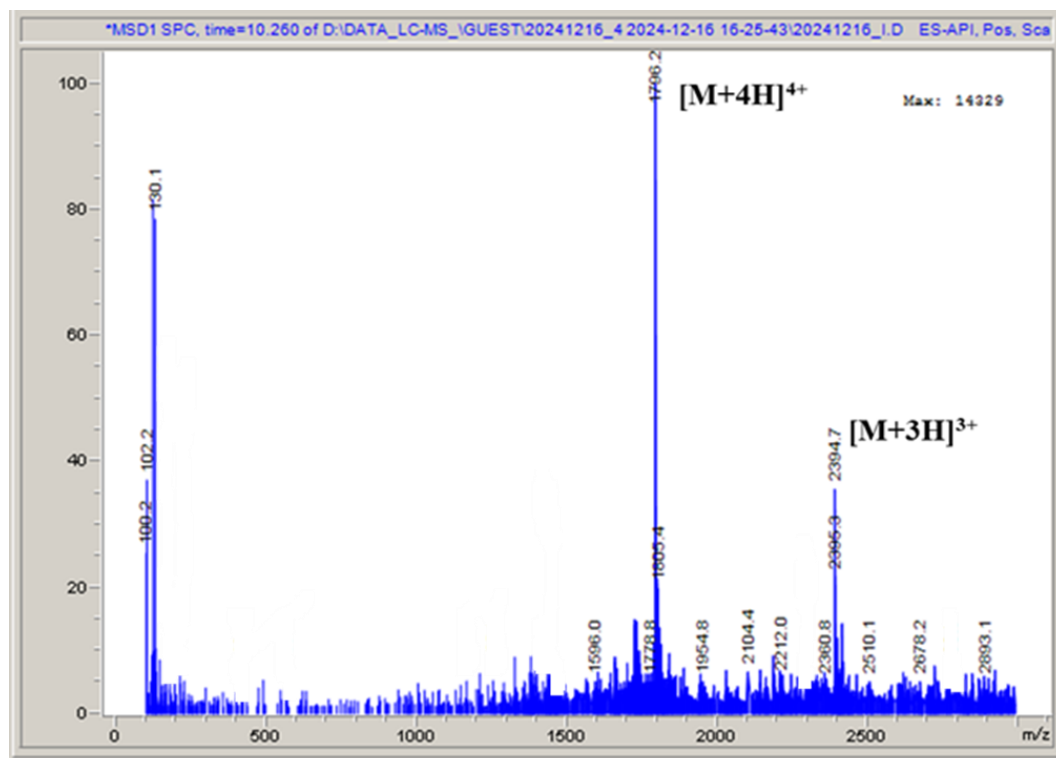


Figure 1.44. ESI-MS spectrum of CF-Tat(triazole)PNA-NH₂. Molecular formula: C₂₉₆H₄₀₆N₁₄₂O₇₆, [M+H]¹⁺ Calculated: 7181.5 Da; Experimental mass: [M+3H]³⁺: 2396.1 Da, [M+4H]⁴⁺: 1795.8 Da.

2.2.2. Click Reaction of CF-PNT-(Triazole)-PNA-NH₂

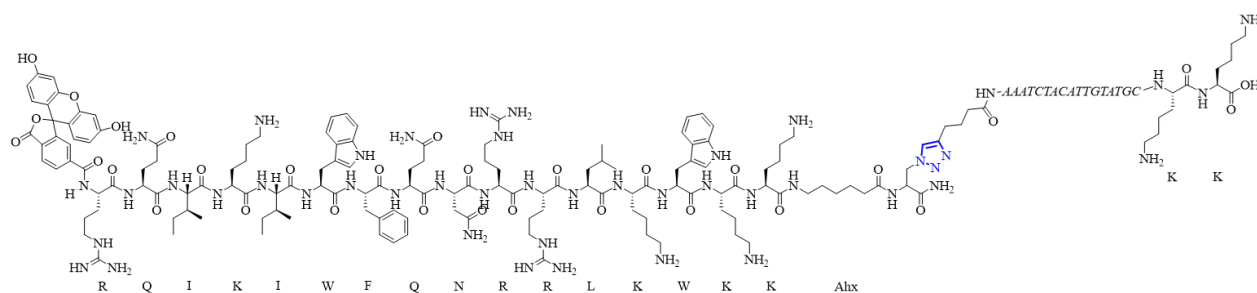


Figure 1.45. Structure of CF-PNT-(Triazole)-PNA-NH₂.

Conjugation between CF-PNT-N₃-NH₂ CPP and alkyne-functionalised PNA was achieved via a CuAAC reaction (Figure 1.45). The alkyne-PNA (1.0 equiv.) and the azide-functionalized CPP (1.3 equiv.) were dissolved in a 1:3 v/v H₂O:*t*-BuOH mixture to ensure the components were fully soluble and to promote the coupling reaction. To generate Cu(I) catalyst *in situ*, CuSO₄·5H₂O (4 equiv.) and sodium ascorbate (16 equiv.) were added sequentially to the reaction mixture. The solution was stirred at room temperature for 24 hours. The slight excess of CPP was intentional, to drive the reaction towards the complete consumption of the alkyne-PNA, the most valuable and limiting component.



Reaction progress was monitored using analytical HPLC, confirmed the formation of the desired conjugate, showing almost complete consumption of the CPP and the PNA (Figure 1.46).

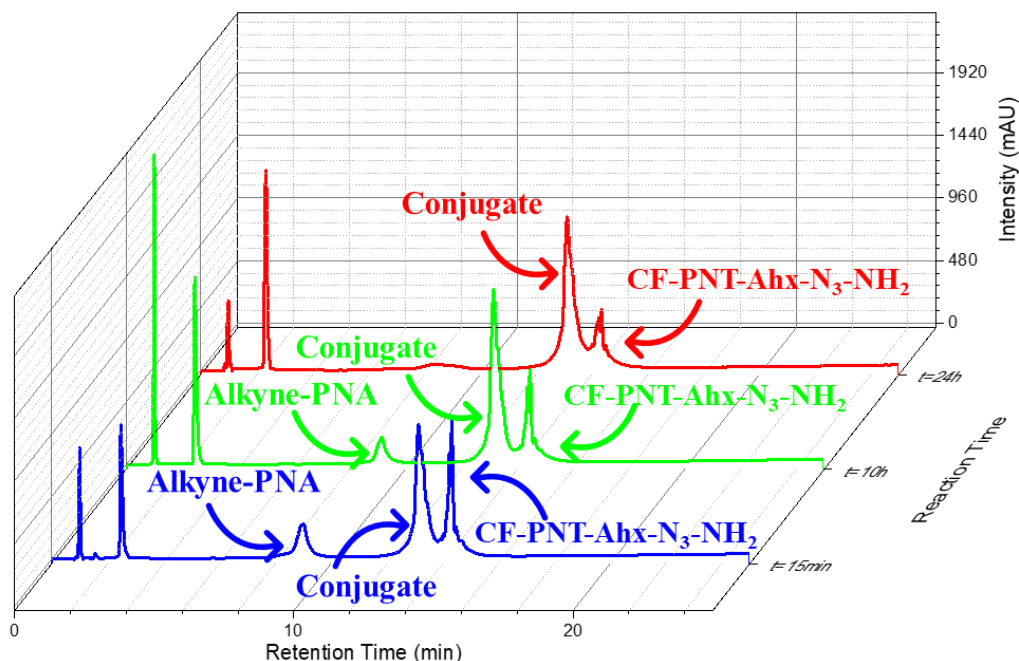
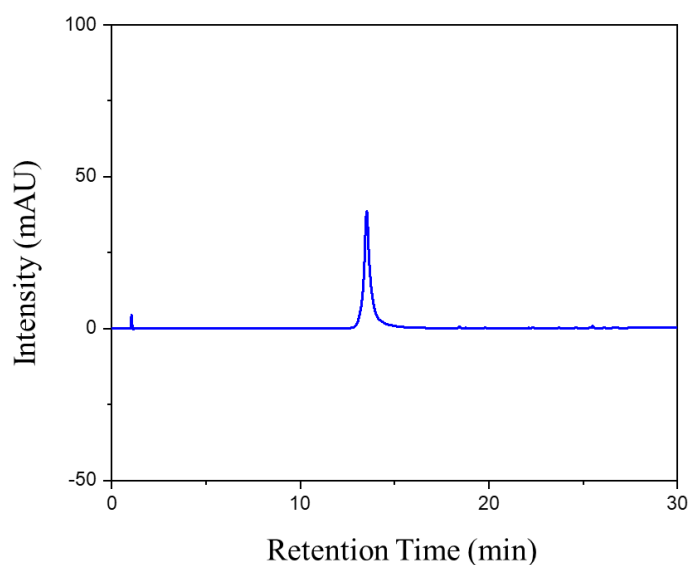


Figure 1.46. Monitoring of the click conjugation reaction between CF-PNT-N₃-NH₂ and the alkyne-functionalised PNA by analytical HPLC at different time points. The chromatogram is recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

The crude product was then purified using reverse-phase HPLC to obtain the pure conjugate (Figure 1.47, Figure 1.48).

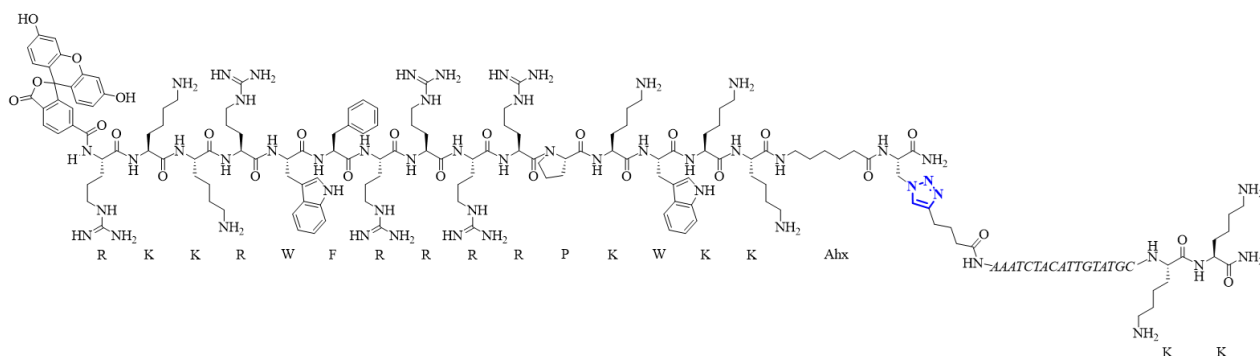


*MSD1 SPC, time=13.422 of D:\DATA_LC-MS\GUEST\20241205_2 2024-12-05 17-15-08\20241205_E1.D ES-API, Pos, S

Mass spectrum showing relative intensity (0 to 100) versus m/z (0 to 3000). The base peak is at m/z 2004.5, labeled $[M+4H]^{4+}$. Other significant peaks are labeled with their m/z values and charge states:

- m/z 102.2
- m/z 130.2
- m/z 1336.7
- m/z 1493.7
- m/z 1599.9
- m/z 1603.5, labeled $[M+5H]^{5+}$
- m/z 1866.4
- m/z 1935.8
- m/z 2683.0
- m/z 2689.2
- m/z 2672.0, 2672.6, 2672.0, labeled $[M+3H]^{3+}$

Max: 21640



jugation of CF-P14-N₃-NH₂ peptide with alkyne-functionalised PNA was achieved via a



temperature for 24 hours. As in the previous case, an excess of CPP was used to maximise PNA conversion and minimise the presence of unreacted alkyne-PNA at the end of the reaction.

After 24 hours, analytical HPLC confirmed the formation of the desired conjugate, indicating near-complete consumption of the CPP and the PNA (Figure 1.50).

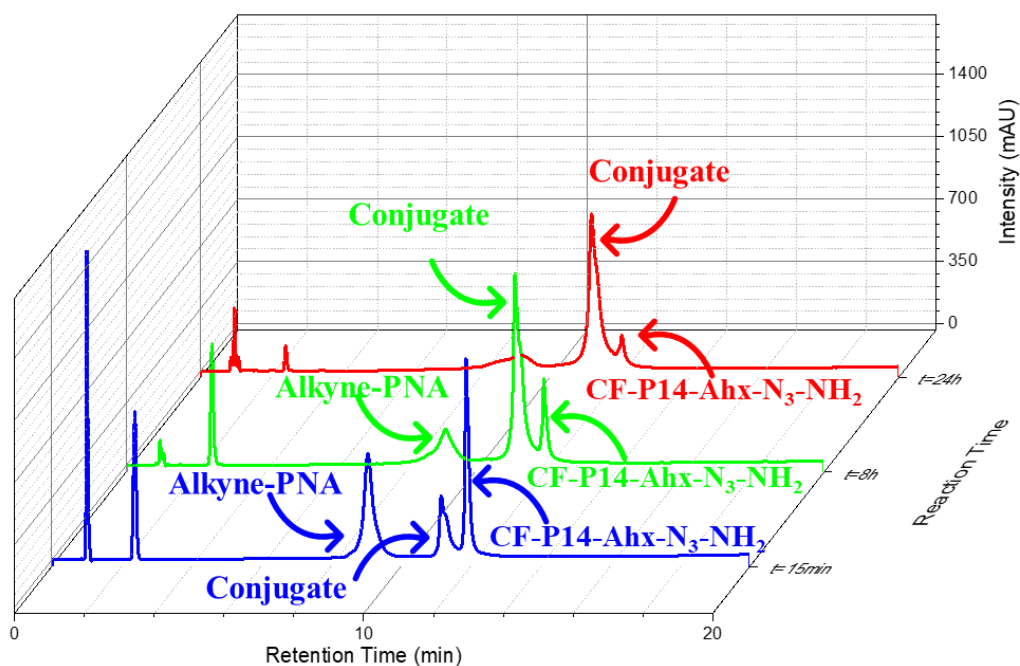


Figure 1.50. Monitoring of the click conjugation reaction between CF-P14-N₃-NH₂ and the alkyne-functionalised PNA by analytical HPLC at different time points. The chromatogram is recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

The crude product was then purified using reverse-phase HPLC to obtain the pure conjugate (Figure 1.51, Figure 1.52).

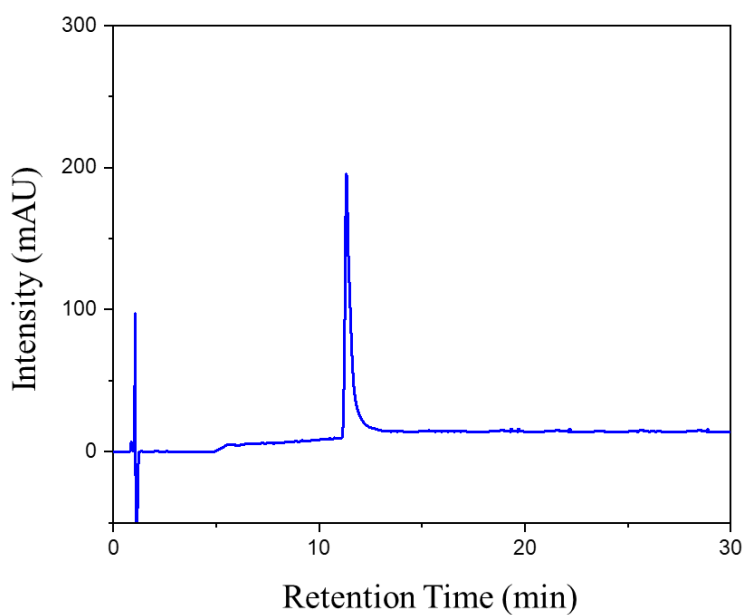


Figure 1.51. Analytical HPLC chromatogram of CF-P14-(triazole)-PNA-NH₂ recorded using a Kinetex C18 column, 1 mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

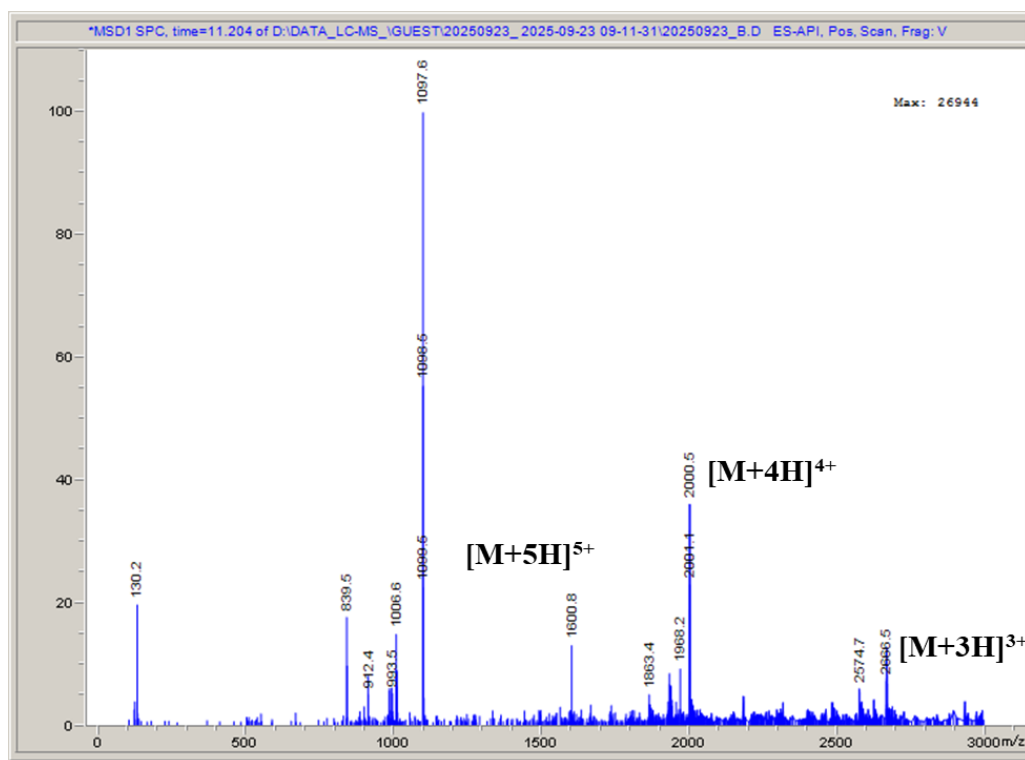


Figure 1.52. ESI-MS spectrum of CF-P14(triazole)PNA-NH₂. Molecular formula: C₃₄₄H₃₄₂N₁₅₀O₈₁, [M+H]¹⁺ Calculated: 8000.2 Da; Experimental mass: [M+3H]³⁺: 2666.3 Da, [M+4H]⁴⁺: 2000.1 Da, [M+5H]⁵⁺: 1600.3 Da.



2.3. Disulfide Bond Formation

Conjugation through disulfide bond formation had already been investigated in previous works. However, when both reacting partners contained free Cys residues, the process was difficult to control. The random oxidation of thiol groups in solution often resulted in the formation of undesired homodimers rather than the intended heteroconjugate. Consequently, the overall yield and selectivity towards the target disulfide-linked product were significantly reduced. To overcome these limitations, a different synthetic strategy was used in this study. This approach involves the selective activation of the cysteine residue on the CPP before conjugation to prevent self-dimerization and favour the formation of the disulfide bond between the CPP (carrier) and the PNA (cargo).

The C-terminal Cys of the CPP was activated using 2,2-dithiodi(5-nitropyridine) (DTNP), which was added in excess to the peptide cleavage solution.[126] DTNP reacts with the thiol group of the Cys residue to form a mixed disulfide intermediate, generating an activated CPP. The 5-nitropyridyl disulfide moiety acts as an excellent leaving group in subsequent thiol-disulfide exchange reactions because the by-product released upon exchange is resonance-stabilised within the nitropyridyl ring. The activated CPP-S-S-Py(NO₂) intermediate was then reacted with reduced PNA bearing a free cysteine residue at its N-terminus in a subsequent step, leading a thiol-disulfide exchange reaction and formation of the final CPP-S-S-PNA conjugate. Since disulfide exchange is an equilibrium process, pre-activation of the CPP effectively shifts the equilibrium towards formation of the heteroconjugate, thereby minimising formation of undesired homodimeric species. Scheme 1.12 illustrates the general reaction scheme for the formation of CPP-PNA conjugates through disulfide linkage.



The chemical structure of Aa100 is a complex polypeptide chain of 100 residues. It features a highly folded backbone with numerous disulfide bonds (S-S) connecting different parts of the chain. The structure is decorated with various post-translational modifications, including hydroxyl groups (OH), guanidino groups (HN=NH), and a large aromatic system on the left. The amino acid sequence is indicated by letters below the chain: G, R, K, K, R, R, O, R, R, Ahx, and C-C. A specific segment of the sequence, AAATCTACATTGLATGCC, is highlighted in blue. The structure is shown in a ribbon representation, with the backbone in black and side chains in various colors.

Conjugation between CF-Tat-Cys(TNP)-OH and cysteine-functionalised PNA was achieved via a thiol-disulfide exchange reaction in an ammonium acetate buffer at pH 4.5 (Figure 1.53). The reaction proceeded at room temperature under mild conditions and was monitored by analytical HPLC (Figure 1.54). During the reaction, the intensity of the peaks corresponding to the starting material gradually decreased, while a new distinct peak corresponding to the desired conjugate appeared at a different retention time.

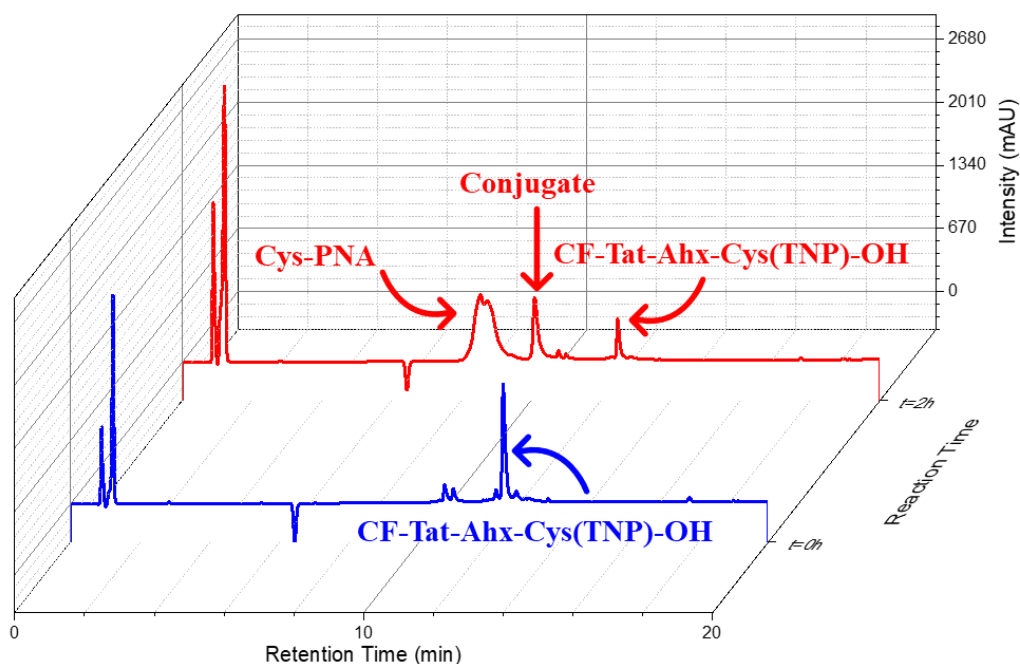


Figure 1.54. Monitoring of the disulfide-exchange conjugation reaction between CF-Tat-Cys(NTP)-OH and the cysteine-functionalised PNA by analytical HPLC at different time points. The chromatogram is recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

The crude mixture was purified using RP-HPLC with a DAISOGEL C18 column (300 Å, 10 µm), using a linear gradient of 10-40% B over 30 minutes. Analytical HPLC (Figure 1.55) and ESI-MS (Figure 1.56) analyses confirmed the formation of the expected conjugate with the correct mass.

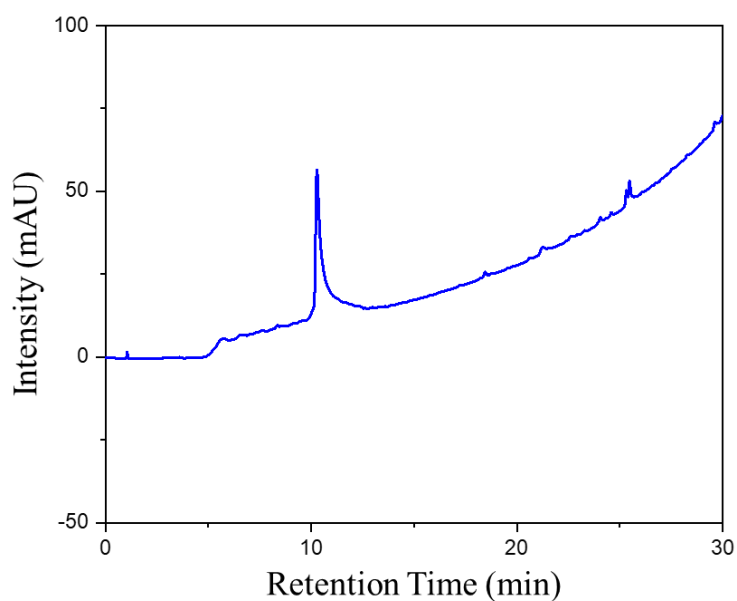


Figure 1.55. Analytical HPLC chromatogram of CF-Tat-S-S-PNA-NH₂ recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

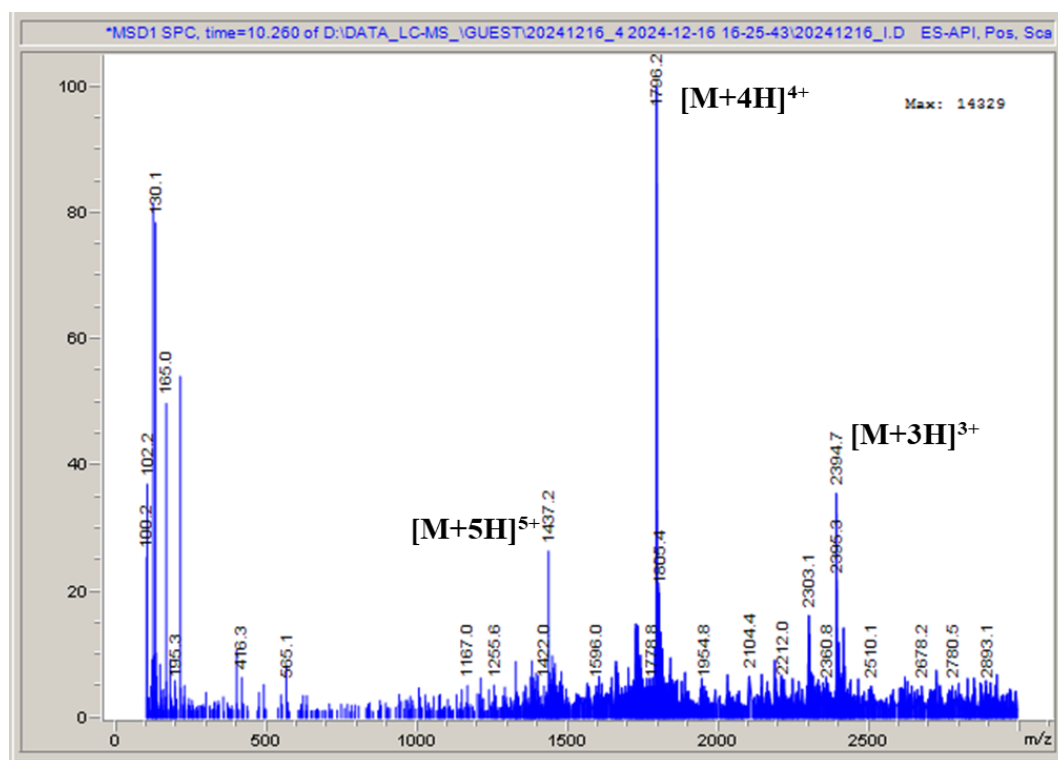


Figure 1.56. ESI-MS spectrum of CF-Tat-S-S-PNA-NH₂. Molecular formula: C₂₉₄H₄₀₆N₁₄₀O₇₆S₂, [M+H]¹⁺ Calculated: 7181.7 Da; Experimental mass: [M+3H]³⁺: 2396.1 Da, [M+4H]⁴⁺: 1795.8 Da, [M+5H]⁵⁺: 1437.0 Da.



2.3.2. Formation of the Conjugate CF-PNT-S-S-PNA-OH

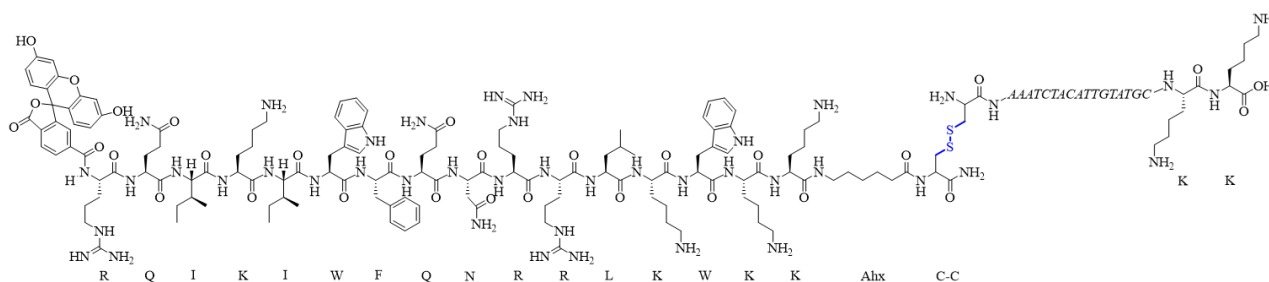


Figure 1.57. Structure of CF-PNT-S-S-PNA-NH₂.

Conjugation of CF-PNT-Cys(TNP)-OH and cysteine-bearing PNA was performed under the same conditions as for the Tat conjugate (Figure 1.57). An ammonium acetate buffer at pH 4.5 was used to facilitate selective thiol-disulfide exchange and limits the oxidation of the free Cys group of the PNA. The reaction proceeded for two hours at room temperature with gentle stirring. As in the previous case, analytical HPLC was used to monitor the reaction, revealing a progressive reduction in the CPP peak consistent with the formation of the desired product (Figure 1.58).

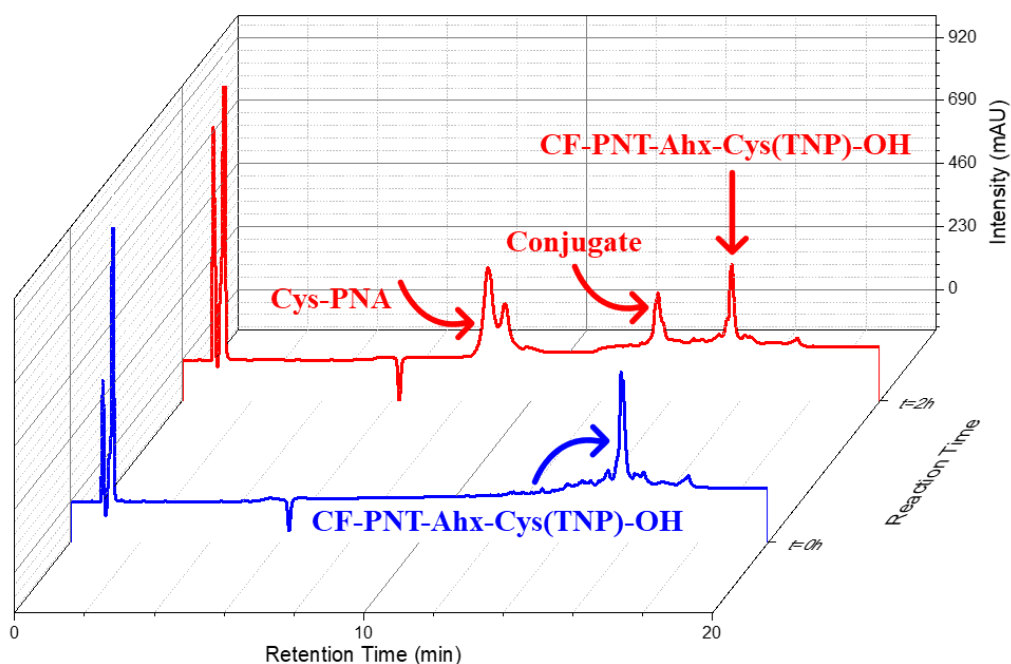


Figure 1.58. Monitoring of the disulfide-exchange conjugation reaction between CF-PNT-Cys(TNP)-OH and the cysteine-functionalised PNA by analytical HPLC at different time points. The chromatogram is recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

Results and Discussion



The reaction mixture was purified using RP-HPLC on a DAISO GEL C18 column (300 Å, 10 µm), using a linear gradient of 10-40% B over 30 minutes. Analytical HPLC (Figure 1.59) and ESI-MS (Figure 1.60) analyses confirmed the formation of the expected conjugate with the correct mass.

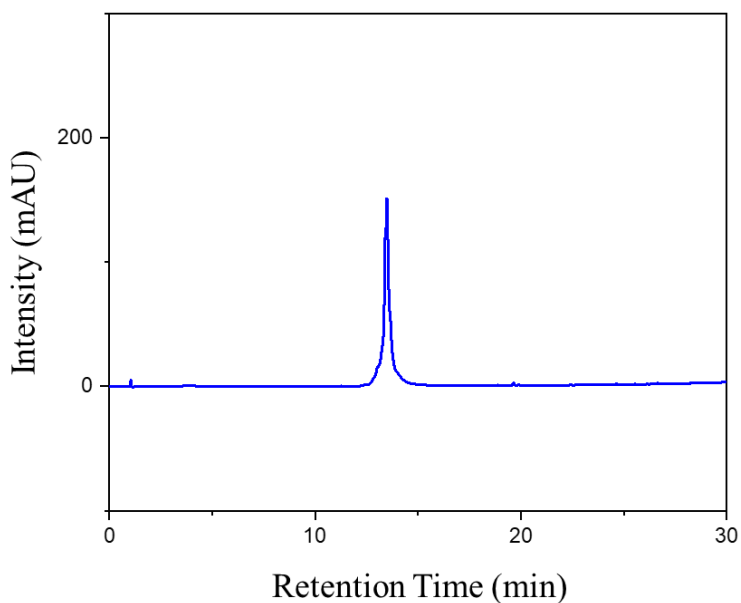


Figure 1.59. Analytical HPLC chromatogram of CF-PNT-S-S-PNA-NH₂ recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

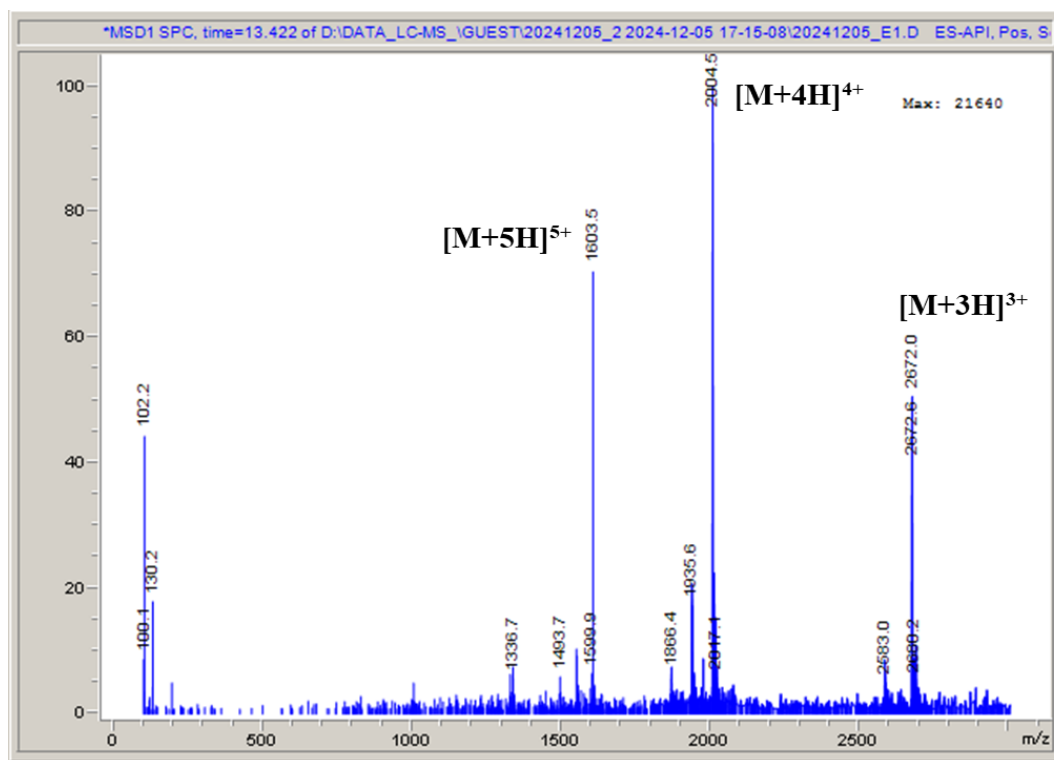


Figure 1.60. ESI-MS spectrum of CF-PNT-S-S-PNA-NH₂. Molecular formula: C₃₄₄H₄₆₇N₁₄₃O₈₄S₂, [M+H]¹⁺ Calculated: 8014.4 Da; Experimental mass: [M+3H]³⁺: 2671.7 Da, [M+4H]⁴⁺: 2004.1 Da, [M+5H]⁵⁺: 1603.3 Da.



2.3.3. Summary of Synthetic Optimisation and Rationale for Conjugate Selection

Before moving on to biological evaluation, it is helpful to summarise the optimisation steps that guided the design of the CPP-PNA systems examined in this study. Throughout the synthetic phase, various conjugation strategies, linker chemistries, and peptide delivery domains were tested to identify the most reliable and versatile platform for intracellular transport. A total of five CPP-PNA conjugates were successfully synthesised using two orthogonal coupling approaches: the CuAAC ‘click’ reaction and disulfide exchange. From a synthetic standpoint, the CuACC strategy provided higher coupling yields, cleaner reactions profiles and greater robustness, particularly when working with sterically demanding PNAs. The triazole linker also offers excellent stability under physiological conditions, ensuring the CPP remains covalently tethered during cellular uptake. In contrast, the disulfide-based approach was more sensitive to reaction conditions. Although it is attractive due to its intracellular cleavability, it suffers from lower overall yields and occasional by-product formation, particularly when C-terminal cysteine-containing constructs were involved. Nevertheless, the reversible nature of the disulfide bond is of interest from a biological standpoint, as intracellular reduction could facilitate cargo release in the cytosol.

These synthetic considerations also influenced the choice of subsequent experiments. For example, although the scrambled PNA sequence had been synthesised and characterised, it was not initially conjugated to CPPs. At this stage of the project, the focus was on identifying the most effective CPP-linker combination for delivering the active PNA sequence efficiently. It only became meaningful to produce the corresponding scrambled control under identical conditions once the optimal stability, purity, and uptake efficiency of the conjugation strategy had been determined. The stepwise approach ensures that subsequent biological comparison will be mechanistically interpretable, rather than being affected by differences in synthetic efficiency or linker behaviour.

Overall, the optimisation workflow described here establishes a clear chemical rationale for the constructs that are taken forward for biological testing. The insight obtained regarding linker stability, coupling efficiency, and CPP compatibility provide a solid foundation for the systematic evaluation of intracellular delivery and antisense activity, which is presented in this chapter.



2.4. Biological Evaluation of CPP-PNA conjugates

Following the synthesis and physicochemical characterisation of the PNA sequences and their peptide conjugates, a series of biological studies were performed to evaluate cell viability and cellular uptake efficiency. The ultimate therapeutic goal of this project is to inhibit the oncogenic microRNA-2215p in cancer cells, thereby arresting cell-cycle progression and reducing tumour proliferation. The biological assays aimed to identify the most efficient delivery systems among the different CPP/linker combinations synthesised.

This phase of the work was carried out during a four-month research stay at the CPCV Laboratory (Sorbonne Université, Paris). During this period, I received training in mammalian cell culture techniques and learnt how to perform quantitative assays to measure cell proliferation and cellular internalisation. Strict adherence to sterile procedures is required for these methods: any experiments lasting 24 hours or more must be performed under controlled aseptic conditions to avoid microbial contamination, and maintaining healthy cultures demands continuous monitoring and careful handling. Acquiring these skills was an essential step in enabling the biological evaluation of the synthetic molecules prepared.

In the CPCV Laboratory, CHO-K1 cells are routinely used for CPP studies due to their robustness and reproducibility. For this reason, the first optimisation tests, concentration ranges and additives effects, were performed on CHO cells. This allowed the experimental parameters to be refined under reliable, well-established conditions. A549 human lung carcinoma cells were then introduced to evaluate the biological activity of the PNA-containing conjugates, as this cell line is a more relevant model for studying the antiproliferative effects associated with miRNA-221 inhibition.

Cell viability and cell proliferation were quantified using the CCK-8 metabolic colourimetric assay, while cellular internalisation was assessed using fluorescence spectroscopy. These assays enabled us to compare the performance of the synthesised delivery systems and identify the most suitable CPP/linker combination for driving PNA entry into cells and achieving a measurable biological response.

2.4.1. CHO K1 cell line

The CHO-K1 (Chinese Hamster Ovary) cell line is a widely used mammalian cell models in biological and pharmacological research. Originally derived from the ovary of the Chinese hamster (*Cricetulus Griseus*), the cells were first cultivated in 1957 and have since become the gold standard for studying cellular physiology, cytotoxicity, and recombinant protein expression.[151] The CHO-K1 subline is characterised by stable growth, strong adherence and high transfection efficiency, making it particularly suitable for reproducible *in vitro* experiments. They exhibit a fibroblast-like



morphology and grow as an adherent monolayer under standard culture conditions in DMEM-F12 medium supplemented with fetal bovine serum.[151]

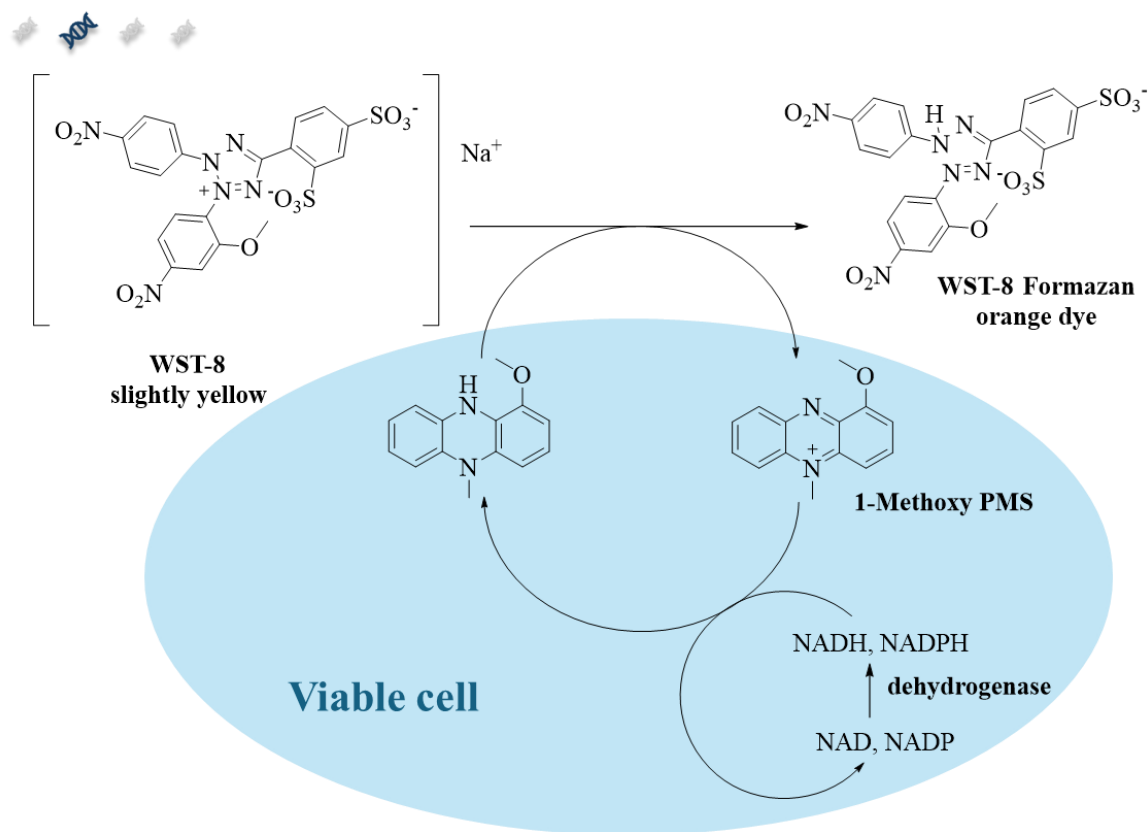
2.4.2. A549 cell line

The A549 cell line, a human lung adenocarcinoma model, is one of the most extensively used in lung cancer research. These cells were originally derived in 1972 from a 58-year-old male patient and exhibit an epithelial morphology. They also share important functional characteristics with alveolar type II pneumocytes, including the ability to synthesise and secrete surfactant phospholipids.[152]

A549 cells are adherent and grow in a monolayer, displaying rapid proliferation. This makes them particularly suitable for *in vitro* experiments exploring tumour growth, signalling pathways and drug responses.[153] Due to their genetic and phenotypic stability, A549 cells are widely used in studies investigating the mechanisms of oncogenesis, apoptosis, autophagy and resistance to chemotherapy or radiotherapy. They have also served as a reference model for investigating miRNA function in lung cancer, with both gain- and loss-of-function experiments being efficiently performed to analyse how specific miRNAs, such as miRNA 221, influence cell proliferation, invasion and survival.[154] Beyond cancer research, A549 cells are used in pulmonary toxicology and virology studies as they provide a reliable model for alveolar epithelial biology. Their versatility and reproducibility make them a cornerstone of lung cancer research, particularly in projects focusing on the functional impact of microRNA dysregulation.[152]

2.4.3. Cell Proliferation Assay (CCK8 Test)

Cell Proliferation and viability were assessed using the Cell Counting Kit-8 (CCK-8) colorimetric assay, which quantitatively measures cellular metabolic activity as an indirect indicator of viable cell number (Scheme 1.13). The assay is based on the principle that WST-8 (2-(2-methoxy-4-mitrophenyl)-3-(4-nitrophenyl)-5-(2,4-disulfophenyl)-2H-tetrazolium monosodium salt), a highly water-soluble tetrazolium compound, is enzymatically reduced.[155] In metabolically active cells, mitochondrial dehydrogenase enzymes catalyse the transfer of electrons from cellular NADH or NADPH to WST-8, reducing it to a stable, water-soluble formazan dye. The reaction can be summarized as follows:



Scheme 1.13. Principle of the CCK-8 detection.

The amount of formazan produced is directly proportional to the number of living cells, since only metabolically active mitochondria can catalyse this reduction. The formazan dye exhibits a strong absorption peak at 450 nm, which can be measured spectrophotometrically to quantify cell viability. Compared with earlier tetrazolium assays, such as MTT, the CCK-8 test offers several advantages. The formazan product is water-soluble, so there is no need for organic solvents to dissolve the crystal. Furthermore, the assay provides higher sensitivity and reproducibility, even at low cell densities. Besides, the absence of washing or solubilisation steps minimises experimental variability, making it particularly suitable for high-throughput applications.[155]

In this work, the CCK-8 assay was used to evaluate the cytocompatibility and potential cytotoxic effects of synthesised CPP-PNA conjugates. In addition, the assay was used to evaluate the functional activity of the conjugates, since effective intracellular delivery of the anti-miRNA-221 PNA is expected to downregulate this oncogenic microRNA, thereby slowing down cell proliferation by inducing cell-cycle arrest. Therefore, the CCK-8 readout reflects the toxic effects of the conjugates and their ability to modulate cell viability through microRNA targeting. Absorbance readings at 450 nm were collected after defined incubation times, and relative cell viability was calculated by comparing treated and untreated samples. This analysis provided quantitative information.

Before evaluating the effects of the CPP-PNA conjugates, a series of preliminary optimisation experiments were conducted to determine the conditions that would allow for the most stable and reproducible cell proliferation assay.



2.4.3.1. Determination of Linear Range for the CCK-8 Assay

Before optimising the incubation protocol or testing the CPP-PNA conjugates, it was essential to determine the appropriate cell seeding density to ensure a linear, unsaturated and sensitive response in the CCK-8 assay. This step enables the selection of an initial cell count that, after 72 hours of proliferation, remains within the assay dynamic range, providing an optimal signal-to-noise ratio from the outset.

First a calibration curve was obtained using CHO-K1 cells seeded at various. After an overnight period for adhesion to occur, the medium was replaced with a 10% CCK-8 solution and absorbance at 450 nm was measured (Table 1.3).

Table 1.3. Determination of the linear detection range of the CCK-8 assay for CHO-K1 cells (positive control, no cells: 0.093 absorbance units).

Number of cells during analysis	Absorbance
5000	0.215333
10000	0.218
20000	0.355583
40000	1.15325

The results showed that, at low densities, the absorbance values were only slightly higher than the background. This was likely due to partial detachment of poorly adherent cells during washing. A clear increase in signal was observed from 20000 to 40000 cells/well, defining the region in which the assay is sufficiently sensitive for CHO-K1 cells. These results helped to determine the minimum number of cells required for reliable measurements in subsequent proliferation experiments.

A similar experiment was conducted with A549 cells to determine the density range at which the CCK-8 signal remains proportional to the number of cells and below saturation.

Table 1.4. Determination of the linear detection range of the CCK-8 assay for CHO-K1 cells (positive control, no cells: 0.12525 absorbance units).

Number of cells during analysis	Absorbance
5000	0.389333
10000	1.146333
20000	3.306
30000	4.047667
40000	4.098333



As shown in Table 1.4, absorbance increased proportionally between 5000 and 30000 cells per well, demonstrating linearity and confirming that the assay accurately reflects metabolic activity between this range. However, at 40000 cells/well, the signal plateaued, indicating reaction saturation and reduced reliability at higher densities. It is also important to note that A549 cells approach confluence at ~30000 cells/well, making these densities suitable for the 72 hours point in proliferation assays. This ensures signal linearity over 72 hours of incubation without reaching saturation.

Thus 1875 cells/well for A549 was selected as the optimal seeding density, enabling reliable absorbance measurements throughout the entire experiment:

- Reliable absorbance measurements throughout the entire experiment
- Linearity of the CCK-8 response at all time points
- Avoidance of signal saturation at the end of the assay

This calibration step was therefore critical for ensuring the accuracy and reproducibility of all subsequent proliferation studies performed on CHO-K1 and A549 cells.

2.4.3.2. *Cell Proliferation Assay with CF-Conjugated Cell-Penetrating Peptides*

This preliminary study aimed to optimise the experimental conditions for the reliable and reproducible evaluation of cell proliferation over 72 hours in preparation for subsequent PNA-CPP conjugate assays. The aim was to identify the most suitable incubation and feeding protocol to maintain consistent cell growth and minimise the effects of stress or detachment induced by medium handling or peptide exposure. This optimization was carried out on CHO-K1 cells using the unconjugated CPPs which are not expected to affect cell proliferation. Preliminary tests revealed that peptides containing an activated Cys residue exhibit strong quenching of the attached fluorophore, even after treatment with TCEP (see graphs reported in Appendix), making quantitative fluorometric analysis unreliable. For this reason, the azido derivatives – whose fluorescence properties remained unaffected – were selected for all subsequent optimisation steps and used as reference CPPs throughout the study.

A clear experimental strategy was defined before assessing the biological activity of the CPP-PNA conjugates. The simplest way to determine whether an anti-miRNA PNA inhibits cell proliferation is to compare the number of cells grown in the absence or presence of the conjugates over cell cycles. To quantify this change, the CCK-8 metabolic assay was selected as it provides a sensitive measurement of cell number over time. However, several parameters required optimisation prior to test the conjugates:

- The initial number of cells seeded per well

- The incubation schedule
- The conditions for refreshing the medium.

These experiments were performed on adherent cells cultured in 96-well plates and the cells must be seeded the day before the experiment to allow proper attachment and recovery. Both CHO-K1 and A549 cells, used successively, exhibit a doubling time of between 18-22 hours, meaning that proliferation assays must be designed to accommodate exponential growth while avoiding overconfluence. For practical reasons and to ensure reproducibility, the tests were limited to a maximum of four days of proliferation (five days including the seeding day).

Taken together, these optimisation steps were essential to establish robust and interpretable conditions for the subsequent evaluation of CPP-PNA conjugates. This ensured that any observed decrease in cell proliferation could be attributed to miRNA targeting rather than experimental artefacts or suboptimal culture conditions.

Cell proliferation assays using CF-conjugated CPPs (CF-Tat-N₃-NH₂ and CF-PNT-N₃-NH₂) revealed significant differences in cell viability, which depended heavily on the adopted incubation and feeding regimen. The proliferation curves obtained in the presence of CPPs were systematically compared with those of control cells treated under identical conditions without CPPs to allow relative cell viability to be calculated. Figure 1.61. therefore, expresses the percentage of viable cells relative to the corresponding control wells, which were incubated and fed in the same way and for the same duration. This ensures that any observed variation can be attributed solely to the effect of the CPPs, rather than to differences in handling or medium replacement.

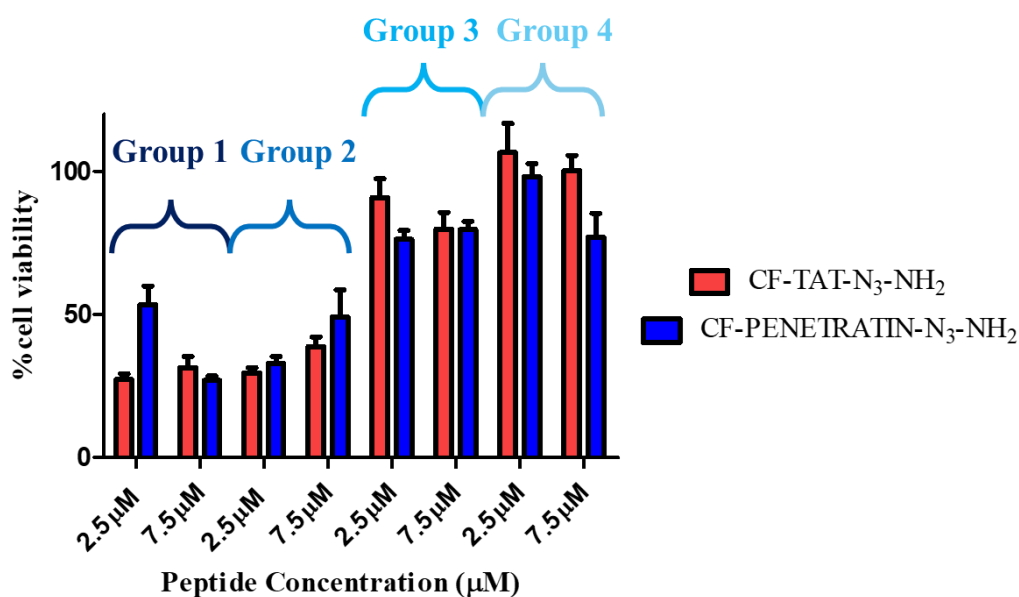


Figure 1.61. Evaluation of CHO-K1 cell proliferation under different incubation and feeding conditions using CF-Tat-N₃-NH₂ (red) and CF-PNT-N₃-NH₂ (blue). CHO-K1 cells were seeded in 96-well plates and treated at 2.5 or 7.5 μM peptide concentrations for 72h under four distinct conditions (Group 1: 5h peptide exposure and daily FBS

Results and Discussion



supplementation; Group 2: 5h peptide exposure, change of the medium after 48h and daily FBS supplementation; Group 3: 72h peptide exposure and daily FBS supplementation; Group 4: 48h peptide exposure, change of the medium after 48h and daily FBS supplementation). Cell viability was quantified using the CCK-8 assay, and absorbance values (450 nm) are reported relative to control cells treated under the same conditions without peptides.

Four incubation regimens were designed to identify the most suitable conditions for reliably monitoring cell proliferation over 72 hours. These regimens differed in terms of the duration of peptide exposure, the timing of medium replacement, and the feeding protocol:

- **Group 1.** Cells were incubated with the CF-labelled CPPs for 5 hours, after which the medium was completely replaced. From that point onwards, the cultures received daily supplementation with serum-containing medium (FBS) without any further changes of the medium.
- **Group 2.** As in Group 1, cells were exposed to peptides for 5 hours, after which the medium was completely replaced. They then received daily FBS supplementation and an additional complete medium change was performed after 48 hours.
- **Group 3.** Cells were exposed to the peptide-containing medium for the entire 72-hours period with no medium replacement and received only daily FBS supplementation.
- **Group 4.** The cells remained in contact with the peptides for the 48 hours, until a complete medium replacement was performed, followed by daily FBS supplementation.

Groups 1 and 2, in which the peptide-containing medium was replaced after just 5 hours, exhibited a significant reduction in cell viability compared to the control group. The most plausible explanation for this is that the early change in the medium destabilised cell adhesion, resulting in partial detachment and impairing subsequent proliferation during the initial growth phase. The additional medium change performed in Group 2 at 48 hours did not compensate for this early disturbance, and viability remained low and unreliable.

In contrast, Groups 3 and 4, in which the peptides remained in contact with the cells for a longer incubation period, displayed markedly higher absorbance values, consistent with better cell viability and more robust proliferation profiles. Of these groups, Group 4 produced the most consistent and reproducible results, demonstrating that a full medium replacement at 48 hours does not induce artefactual cell loss (unlike medium changes performed on day 1), restores nutrients, and improves culture homogeneity. It also supports stable cell growth while maintaining continuous peptide exposure. These results strongly suggests that, by 48 hours, cells are sufficiently adherent to tolerate medium manipulation without detachment. This explains why only the early medium change negatively affected viability.

Overall, these optimisation experiments demonstrate that the most reliable protocol for 72-hour proliferation assays is as follows:



- 48 hours exposure to the CF-CPPs
- Daily FBS supplementation
- A single complete medium replacement at 48 hours

These studies also confirmed that fluorescein labelled Tat and Penetratin do not affect CHO-K1 proliferation at 2.5 μ M and 7.5 μ M over three days under the optimised conditions – an essential prerequisite to ensuring that any future antiproliferative effect can be attributed to the PNA cargo rather than the CPP itself.

2.4.3.3. *CF-CPP-PNA-NH₂ Cell Proliferation Analysis*

Having established the experimental conditions with CF-TAT-N₃-NH₂ and CF-PNT-N₃-NH₂, the evaluation was extended to CPP-PNA conjugates. This phase aimed to evaluate their potential short-term cytotoxicity and their effects on A549 cell proliferation over time, to determine whether the conjugates exert the expected antisense biological activity – namely, inhibition of cell division by targeting miRNA-221 – rather than inducing non-specific toxic responses.

It was therefore essential to distinguish between two fundamentally different biological outcomes: true-antiproliferative activity arising from the sequence-specific engagement of the PNA with its intracellular microRNA target, which leads to the controlled slowing of cell division, and non-specific cytotoxicity, which would manifest as rapid loss of viability due to, for example, to membrane destabilisation by the CPP moiety (e.g. pore formation or membrane rupture) rather than antisense action.

By monitoring cell metabolic activity and proliferation over 72 hours, these assays determine whether their effects are dominated by acute toxicity unrelated to microRNA targeting. This distinction is critical for identifying CPP/linker combinations that support efficient PNA delivery while preserving the mechanistic integrity of the antisense therapeutic strategy.

In an initial screening, the CPP-PNA conjugates were evaluated at a concentration of 2.5 μ M to assess their short-term impact on the proliferation of A549 cells. To determine whether the absence of detectable biological activity was due to inefficient endosomal escapes, a common limitation of CPP-delivered oligonucleotides, the conjugates were tested both alone and in presence of either chloroquine (CQ, 150 μ M) or calcium ions (Ca²⁺, 5 mM).

Like many macromolecular cargoes, CPP-PNA conjugates can enter cells via endocytosis, a pathway which often traps the conjugates inside endosomes, thereby preventing PNAs from reaching their cytosolic microRNA targets. For this reason, CQ is often used as a chemical enhancer of endosomal escape. CQ is a weak base that accumulates inside endosomes, where it becomes protonated and increases the pH of the vesicle interior. This pH buffering leads to osmotic swelling of the vesicles,



destabilisation of the membrane and, in some cases, partial rupture, facilitating the release of internalised material into the cytosol. [156]

Ca^{2+} ions were also tested as an alternative strategy to promote endosomal escape. High concentrations of Ca^{2+} are known to destabilise anionic phospholipid membranes, inducing membrane fusion or transient permeabilization events. In the context of CPP-mediated delivery, Ca^{2+} has been shown to potentiate escape from endocytic vesicles by altering membrane curvature, enhancing vesicle-membrane fusion dynamics and modulating electrostatic interactions that govern CPP-membrane binding. [157, 158]

Both additives had previously been reported to improve the cytosolic release of the CPP conjugates, and preliminary work in the CPCV laboratory supported their potential utility. Their inclusion in the screening was therefore justified in order to determine whether endosomal entrapment represented a limiting step for PNA-mediated microRNA inhibition.[159]

In addition, unconjugated CPPs were co-incubated with CPP-PNA conjugates. This strategy was inspired by the work of Hackenberger *et al.* [160] who demonstrated that adding an excess of free CPP can enhance the delivery of proteins to the cytosol through cooperative interactions at the membrane. Preliminary experiments were performed to evaluate whether a similar effect could improve PNA delivery by co-incubating the CF-Tat-(triazole)-PNA- NH_2 and CF-Tat-S-S-PNA- NH_2 conjugates with free CF-Tat-Cys(TNP)-OH (10 mM). Under these conditions, no enhancement in the biological response was detected. After 48 hours, cell viability remained close to 100%, indicating that the co-administration of unconjugated CPPs neither improved intracellular delivery nor induced any detectable proliferative effect (data not shown). Given the absence of a measurable benefit, this additive strategy was not pursued further.

Figure 1.62 shows an example of the results obtained for Tat conjugates with triazole or disulfide linkers. Analogue experiments on PNT conjugates are presented in the Appendix and gave the same results. Across all conditions – conjugate alone, with CQ, with Ca^{2+} or with free CPP – no impact on cell proliferation was observed at the analysed time points (24, 48, and 72 hours). This suggested that either the uptake efficiency of the conjugates at this concentration was limited or the intracytosolic concentration of the conjugates was too low to elicit a measurable biological effect.

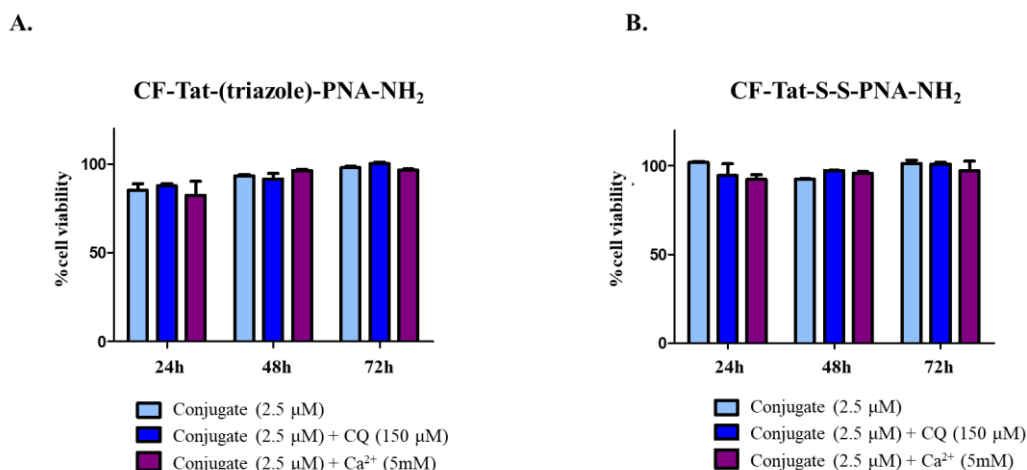


Figure 1.62. **A)** CF-Tat-(triazole)-PNA-NH₂ and **B)** CF-Tat-S-S-PNA-NH₂ were tested at a final concentration of 2.5 μM, either alone or in combination with chloroquine (CQ, 150 μM) or calcium ions (Ca²⁺, 5 mM). Cell viability was evaluated after 24h, 48h and 72h using the CCK-8 assay. The percentage of viable cells is expressed relative to the corresponding control, consisting of cells incubated for the same duration under identical conditions (complete culture medium and serum), but without peptide or additives.

To better assess the cellular response to the CPP-PNA conjugates, the proliferation assays were repeated using a higher concentration of the conjugates (5 μM) and greater amount of chloroquine (500 μM). This second set of experiments included an incubation time of 1 hour to evaluate potential acute membrane toxicity. The obtained data are shown in the figures below.

Importantly, no decrease in cell viability was observed after one hour of incubation with any of the conjugates or additives. This rules out that the conjugates do not damage the plasma membrane. Therefore, any biological effects observed at later time points must arise from intracellular activity following entry into the cells.

After 24 hours of incubation, a small but reproducible reduction in cell viability (approximately from 76% to 90% of cell viability) was detected for two conjugates – CF-Tat-(triazole)PNA-NH₂ and CF-PNT-(triazole)-PNA-NH₂ – when administered alone. This early decrease suggests that these conjugates begin to act shortly after internalisation, most likely through the rapid hybridisation of the PNA with its intracellular microRNA target.

At later points (48 and 72 hours), cell viability stabilised and reach a plateau, with no further decrease despite continued exposure to the conjugates. This behaviour is consistent with a biological mechanism in which:

- The conjugates exert their maximal effect early after internalisation,
- Intracellular PNA levels may gradually diminish due to degradation, sequestration or limited endosomal escape,
- Alternatively, the cells may adapt to the treatment after an initial perturbation of proliferative signalling pathways.



In this second set of experiments, the impact of two additives on cell viability was evaluated specifically in relation to the optimised 72-hour incubation protocol.

- **Chloroquine.** At a concentration of 150 μM , CQ alone exhibited slightly but measurable cytotoxicity after 48 hours (approximately 14% decrease in cell viability), a phenomenon that was not observed in the initial optimisation experiments conducted under different culture conditions. Notably, the presence of both CQ and the CPP-PNA conjugate resulted in a more pronounced reduction in viable cell number after 48 hours compared to CQ alone. This behaviour is consistent with CQ-induced enhancement of endosomal escape, allowing a larger fraction of the internalised conjugate to reach the cytosol and hybridise with its intracellular microRNA target. A partial recovery was noted after 48 hours, likely due to the medium replacement step. In this protocol, the conjugate was not added again at this time, meaning that the intracellular concentration gradually decreased after the first day.

When the concentration of CQ was increased to 500 μM , a marked cytotoxic effect was observed (approximately 76% decrease after 24 hours), which prevented any meaningful interpretation of the activity of the conjugate. At this dose, the toxicity was clearly attributable to CQ itself rather than PNA action.

- **Calcium ions.** At 5mM, Ca^{2+} alone exhibited intrinsic toxicity, leading to a decrease in viability of approximately 57% after 48 hours. Interestingly, when Ca^{2+} was co-administrated with the CPP-PNA conjugates, the resulting cell viability values were slightly higher than those observed for Ca^{2+} alone. This suggests that the conjugates have a protective or stabilising effect on the plasma membrane. However, this co-incubation did not improve the biological activity of the conjugates, no additional reduction in cell proliferation was observed compared to the conjugate alone. Therefore, although Ca^{2+} altered membrane properties and induced some toxicity, it did not enhance cytosolic delivery or antisense activity under the tested conditions.

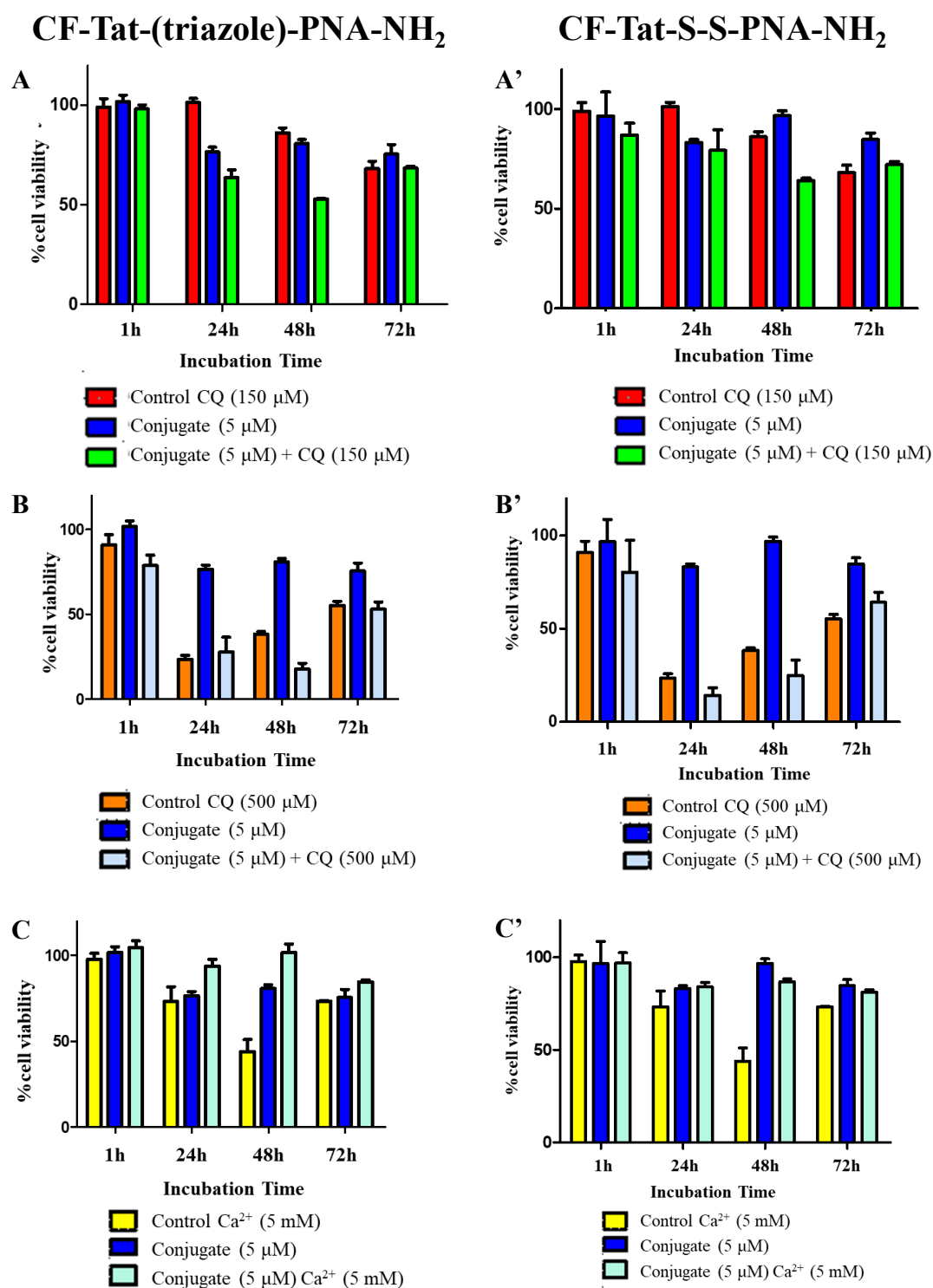


Figure 1.63. CF-Tat-(triazole)-PNA (left) and CF-Tat-S-S-PNA (right) were tested at a final concentration of 5 μM, either alone or in combination with **A, A'** chloroquine (CQ) 150 μM, **B, B'** CQ 500 μM or **C, C'** calcium ions (Ca²⁺, 5 mM). Cell viability was evaluated after 1h, 24h, 48h and 72h using the CCK-8 assay. The percentage of viable cells is expressed relative to the corresponding control, consisting of cells incubated for the same duration under identical conditions (complete culture medium and serum), but without peptide or additives.

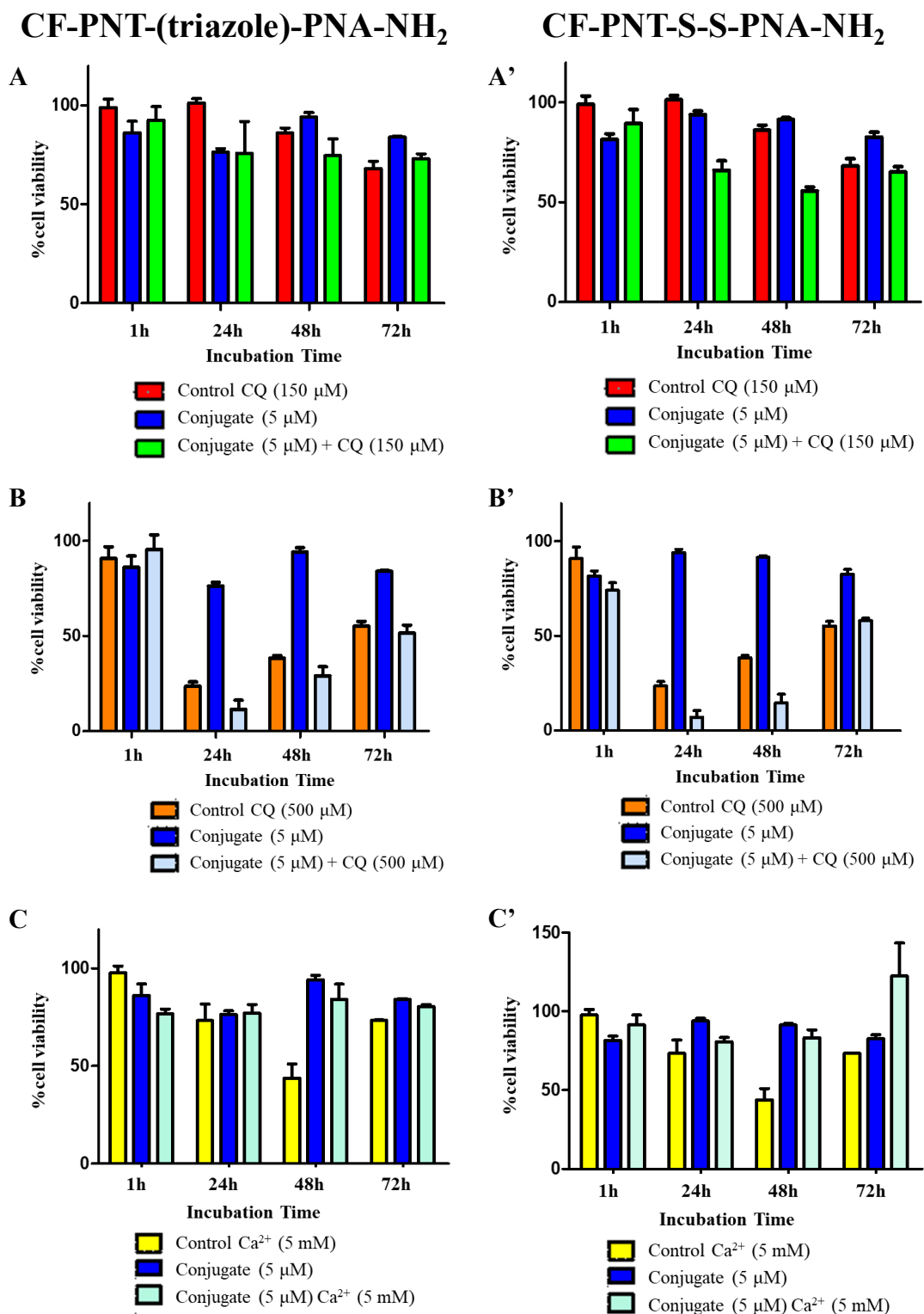


Figure 1.64. CF-PNT-(triazole)-PNA (left) and CF-PNT-S-S-PNA (right) were tested at a final concentration of 5 μ M, either alone or in combination with **A, A'** chloroquine (CQ) 150 μ M, **B, B'** CQ 500 μ M or **C, C'** calcium ions (Ca^{2+} , 5 mM). Cell viability was evaluated after 1h, 24h, 48h and 72h using the CCK-8 assay. The percentage of viable cells is expressed relative to the corresponding control, consisting of cells incubated for the same duration under identical conditions (complete culture medium and serum), but without peptide or additives.

A partial recovery in cell viability was observed after 72 hours in the case of the conjugates with or without CQ, but also of calcium alone, likely reflecting the medium change performed after 48 hours,



which provided fresh nutrients and mitigated the effects of stress. To optimise the experimental setup further, future assays could either avoid the medium change at 48 hours or perform it while reintroducing the conjugate and the selected additive to ensure consistent exposure throughout the 72-period.

A comparison of the four conjugates revealed clear differences in biological activity. The two CuAAC-synthesised triazole linked conjugates consistently showed a measurable decrease in cell proliferation, whereas the disulfide linked analogues exhibited minimal activity. Although disulfide linkages have the potential to be released intracellularly under reducing conditions, they may also undergo premature cleavage in the extracellular environment, in endosomes, or during the initial stages of uptake. This reduces the effective intracellular concentration of the intact CPP-PNA conjugates. In contrast, the triazole linkage is chemically inert and resistant to intracellular reduction, which likely improves stability, intracellular persistence, and overall functional delivery. Taken together, these results indicate that triazole-based CPP-PNA conjugates are the most promising candidates for further biological evaluation, demonstrating reproducible activity and robustness across the tested conditions.

2.4.4. Internalisation Studies by Fluorimetry

The quantitative evaluation of CPP-PNA conjugate internalisation was carried out using the fluorometric protocol illustrated schematically in Figure 1.65. This method developed in the CPCV laboratories, has been applied before to quantify the entry inside cells of unconjugated peptides or peptides grafted at the surface of nanoparticles. [161, 162] It enables the differentiation and quantification of membrane-bound versus internalised fluorescent peptide derivatives by combining controlled enzymatic digestion, cell lysis, and spectrofluorometric measurement.[161]

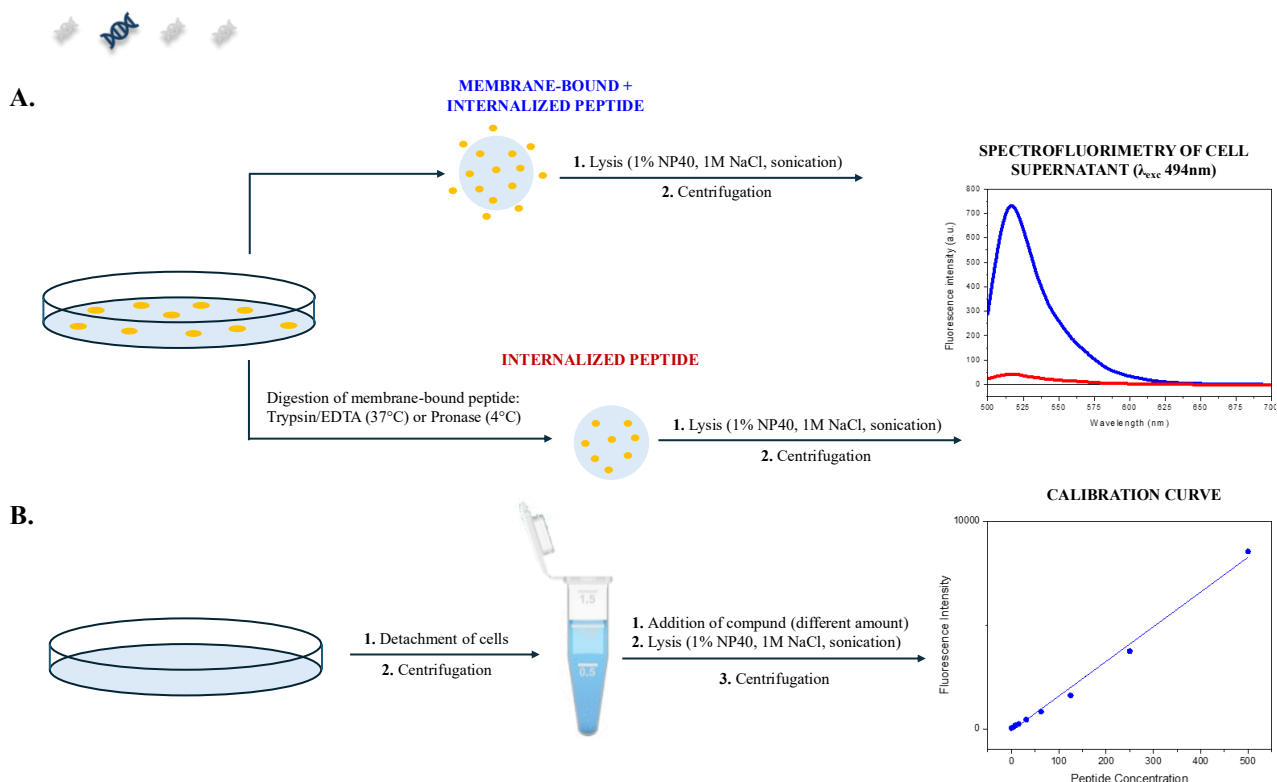


Figure 1.65. Protocol for compounds quantification by fluorimetry. **A)** Sample preparation and **B)** calibration sample preparation ran in parallel.

It was applied here for the first time to the evaluation of peptide-PNA conjugates. Following the incubation of cells with the carboxyfluorescein (CF)-labelled CPP-PNA conjugate, any unbound material was removed by extensive washing. At this stage, the total cell-associated fluorescence includes both membrane-bound and internalised species. To measure only the internalized fraction, the membrane-bound conjugate is removed by enzymatic digestion using either trypsin/EDTA (5 min at 37°C) or pronase (5 min at 4°C). These proteases cleave peptides bound to the extracellular membrane without permeabilising the cells, thereby leaving only the internalised material intact.

Following digestion, the cells were lysed in a solution of 1% NP-40 and 1M NaCl with brief sonication to ensure complete disruption of the cellular membranes and release of the internalised fluorescent conjugates into the supernatant. The lysates are then centrifuged to remove cell debris, after which the fluorescent intensity of the clear supernatant was recorded using a spectrofluorometer (excitation at 494 nm and emission between 515 and 600 nm).

In parallel, a calibration curve (Figure 1.67, B) is created by lysing untreated cells and adding known quantities of CF-labelled conjugates directly to the lysate. Fluorescence intensity was measured under identical conditions to correlate signal strength with conjugates concentration. The quantity of the internalised conjugate could be accurately determined by comparing the fluorescence of the experimental samples with this calibration curve, (in pmoles or molecules per cell).

This method provides a robust, quantitative assessment of peptide internalisation, eliminating overestimation due to surface adsorption. Furthermore, using both enzymatic digestion and detergent-



based lysis ensures precise discrimination between extracellularly bound and intracellularly localised fluorescent species, providing a reliable measure of true cellular uptake efficiency.

2.4.4.1. Internalisation of CF-CPP- N_3 -NH₂

The cellular uptake of the CF-labelled CPPs (CF-Tat- N_3 -NH₂ and CF-PNT- N_3 -NH₂) was evaluated in CHO-K1 cells at 4°C and 37°C to distinguish between energy-dependent and energy-independent internalisation mechanisms (Figure 1.66). CPPs can enter cells through both endocytosis and direct translocation across the plasma membrane at 37°C, whereas endocytosis is strongly inhibited at 4°C and only energy-independent translocation can occur. Following incubation, the surface-bound peptides were removed by enzymatic digestion with trypsin or pronase, allowing the internalised fluorescent fraction to be quantified. CPPs containing terminal Cys residue were excluded because the TNP protecting group was found to significantly quench the fluorescence of the attached CF dye. This quenching effect persisted even after treatment with TCEP, making fluorescence-based quantification unreliable. Consequently, only the azido-terminated analogues were used for internalisation assays.

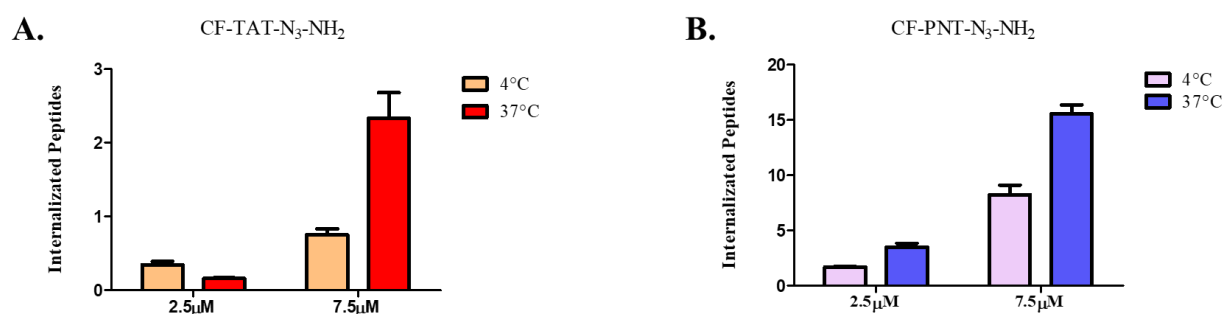


Figure 1.66. Quantification of internalized CF-CPPs in CHO-K1 cells under different enzymatic and temperature conditions.

At low CF-Tat- N_3 -NH₂ (2.5 μM) concentration (Figure 1.68, A), internalisation predominantly occurs through an energy-independent translocation mechanism, as demonstrated by the presence of detectable intracellular fluorescence even at 4°C. This behaviour suggests that some of the peptide can cross the plasma membrane directly, bypassing classical vesicular uptake. Endocytosis is triggered at higher concentration (7.5 μM), possibly because it requires an accumulation of the CPP on the cell membrane, and it then becomes the major route of entry. CF-PNT- N_3 -NH₂ (Figure 1.65, B), was found to enter cells more efficiently than the Tat derivative and was also able to use both translocation and endocytosis, in agreement with previous data obtained at the CPCV lab.[161]

Crucially, these internalisation assays were performed in triplicate, with each experiment repeated twice to yield highly consistent and reproducible results. This strong reproducibility contrasts sharply



with the variability observed for the CPP-PNA conjugates (see the next paragraph), where the presence of the PNA cargo significantly impacted internalisation behaviour, resulting in less uniform uptake profiles.

2.4.4.2. Internalisation of CF-CPP-PNA-NH₂

The internalisation of CPP-PNA conjugates was evaluated in A549 cells at 4°C and 37°C in order to investigate their uptake mechanisms efficiency.

However, unlike the experiments performed with free CPPs, the results obtained for the conjugates were not reproducible. Even under identical conditions, significant variation in fluorescence signals was observed between replicates Figure 1.69 shows a representative example for the Tat-triazole conjugate, while the analogous results for the other conjugates are reported in Appendix.

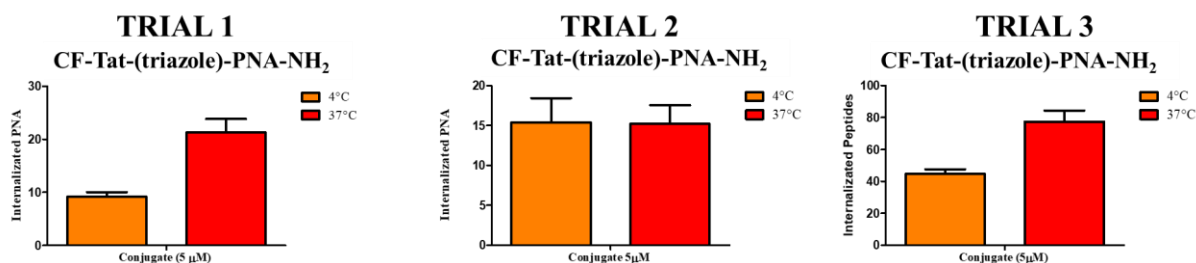


Figure 1.67. Internalization of CF-TAT(triazole)PNA-NH₂ alone and in the presence of CQ (150 μM) in A549 cells.

This lack in reproducibility may be due to an aggregation phenomenon involving the PNA-CPP conjugates in aqueous media or in presence of cells. PNAs are prone to hydrophobic stacking and self-association, [163] and when they are linked to cationic CPPs, the resulting amphiphilic structures can form aggregates or micelle-like assemblies. These aggregates could reduce the effective concentration of the monomeric conjugate available to interact with the cell membrane and cause heterogeneous uptake depending on the size and surface charge of the aggregates. In addition, aggregation may lead to sedimentation during incubation, further complicating quantitative comparisons between samples.

To mitigate these issues, several adjustments could improve sample homogeneity and reproducibility, such as:

- Avoid mechanical agitation (e.g. vortexing) before incubation, as this can promote aggregation of conjugates. It is preferable to gently invert or pipette the solution to achieve proper mixing.
- Prepare fresh stock solutions immediately before use.

These observations suggest that variability in internalisation data primarily arises from the behaviour of the conjugates in solution rather than to differences in intrinsic uptake mechanisms. Future



optimisation should therefore focus on solution-state stabilisation to enable a more accurate correlation between internalisation efficiency and biological response.

2.5. Conformational Studies by Circular Dichroism.

Circular Dichroism (CD) spectroscopy is a powerful technique for investigating the secondary structure and conformational properties of chiral biomolecules such as peptides, proteins, and nucleic acid analogues. This method is based on the differential absorption of left- and right- circularly polarised light, depending on the chiral arrangement of chromophores within the molecule. CD spectra provide valuable information about overall folding and the presence of structural motifs, such as α -helices and β -sheets in peptides and helical conformations in nucleic acids. In this chapter, CD spectroscopy was used to evaluate potential structural modifications resulting from conjugation or hybridisation with complementary nucleic acids strands.[136]

A more detailed discussion of the CD analyses and their interpretation can be found in Chapter 3.

2.5.1. CD Spectroscopic Analysis of miRNA/anti-miRNA and miRNA/Scramble Complexes

CD spectroscopy was employed to investigate the structural behaviour of miRNA-221-5p and its interaction with the complementary anti-miRNA-221-5p PNA and the non-complementary scrambled sequence. The CD spectrum of miRNA-221-5p alone exhibited the typical features of a B-type right-handed helix, including a positive band at 266 nm and a negative one at around 205 nm. This is consistent with the expected secondary structure of RNA oligonucleotides in an aqueous buffer.

When the anti-miRNA-221-5p was added, the CD spectrum showed a clear increase in the intensity of the positive band and a slight red shift of the maximum. This was also accompanied by an increase in ellipticity of the band at 237 nm. These spectral changes indicate the formation of a stable heteroduplex associated with increased base stacking and helicity. The resulting spectrum differs significantly from the arithmetic sum of the individual strands, confirming that specific hybridisation occurs between the complementary sequences (Figure 1.68).

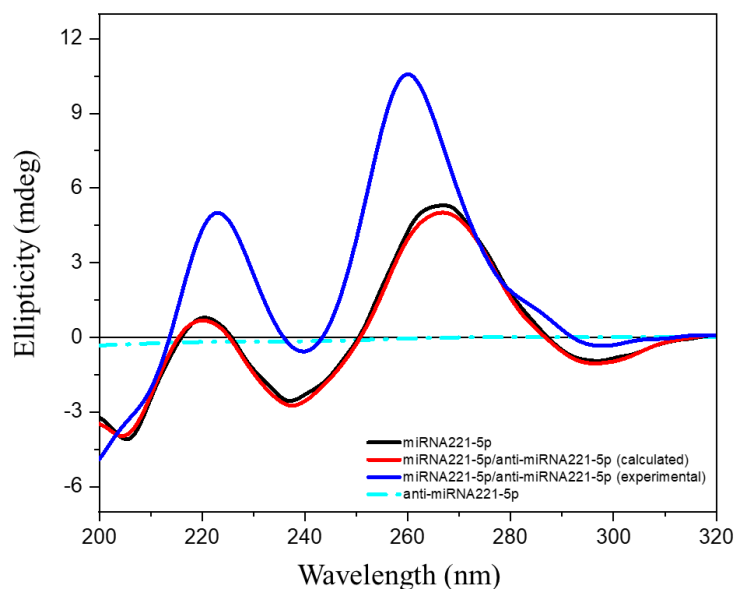


Figure 1.68. CD spectra show the spectrum of the miRNA alone (in black), the experimental spectrum (in blue) and the calculated spectrum (in red) of the miRNA/anti-miRNA duplex. The dashed line represents the spectrum of the anti-miRNA PNA alone.

By contrast, the spectrum of miRNA-221-5p incubated with scrambled PNA remained almost identical to that of free miRNA-221-5p, with only minor intensity variations and no significant wavelength shifts. The experimental trace closely overlapped with the calculated sum of the individual components, demonstrating the absence of stable duplex formation. This confirms that the observed structural changes in the miRNA/anti-miRNA system arise from specific base pairing rather than non-specific electrostatic or stacking interactions (Figure 1.69).

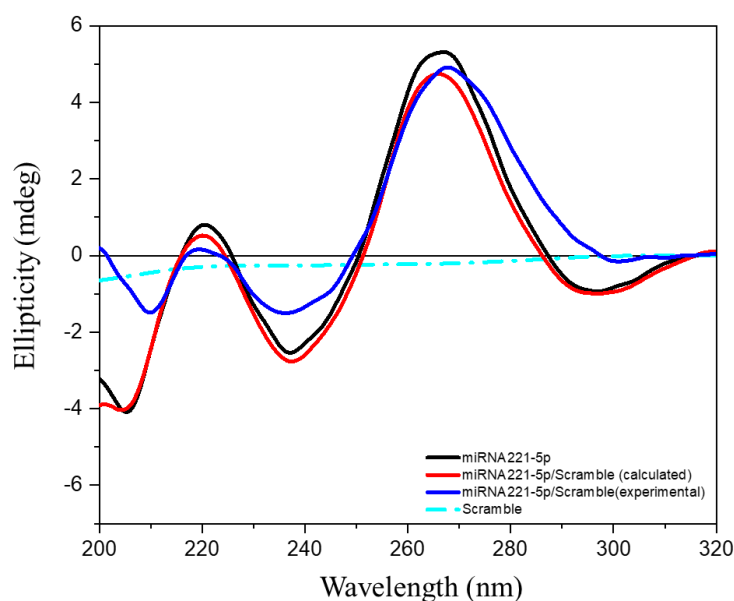


Figure 1.69. CD spectra show the spectrum of the miRNA alone (in black), the experimental spectrum (in blue) and the calculated spectrum (in red) of the miRNA/scramble PNA complex. The dashed line represents the spectrum of the scramble PNA alone.

Overall, CD analysis provides strong evidence that anti-miRNA-221-5p PNA effectively hybridises with the complementary miRNA target, promoting a well-defined, ordered duplex conformation. In contrast, the scrambled control fails to induce any measurable alteration to the secondary structure. These findings are consistent with the selectivity and base-pairing fidelity of PNA-RNA interactions (Figure 1.70).

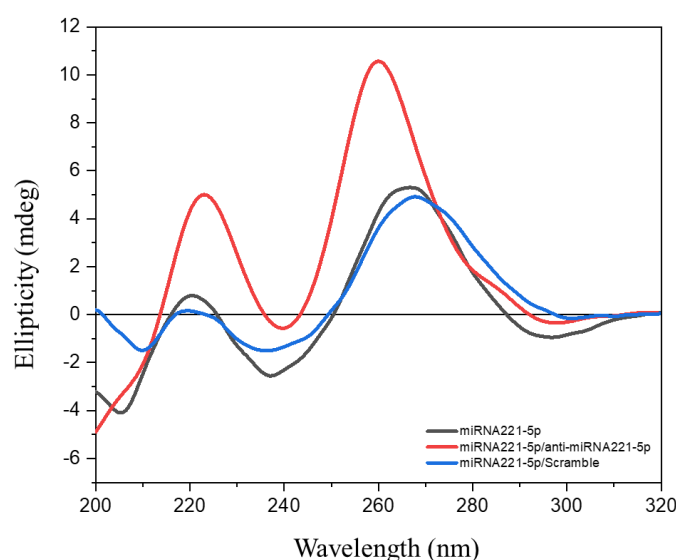


Figure 1.70. Comparison of the CD spectra of the miRNA alone (black line), miRNA/anti-miRNA PNA duplex (red line) and miRNA/scramble PNA complex (black line).



2.6. Thermal Denaturation (Melting) Studies of miRNA/PNA Duplexes

Thermal denaturation, also referred to as UV-melting analyses, is a well-established technique for studying the thermodynamic stability of nucleic acid duplexes. The method is based on the fact that the aromatic bases of nucleic acids absorb ultraviolet light at around 260 nm and that this absorption increases as the duplex denatures into single strands – a phenomenon known as hyperchromicity. Monitoring the change in absorbance as a function of temperature makes it possible to follow the helix-to-coil transition and determine the melting temperature (T_m), which is defined as the point at which half of the duplex molecules are denatured.[164]

The value of T_m provides a direct measure of duplex stability, which depends on the complementarity and sequence composition of the two strands, as well as the ionic strength and chemical environment of the medium. In the case of miRNA/PNA hybrids, the neutral backbone of the PNA enhances binding affinity by eliminating electrostatic repulsion, resulting in higher T_m values compared to analogous DNA or RNA duplexes. Analysis of the shape of the melting curve also offers insight into the cooperativity of the denaturation process and the presence of possible secondary binding events or mismatches.

In this chapter, UV melting experiments were performed to evaluate the binding affinity and thermal stability of miRNA/PNA duplexes. Measurements were taken by monitoring absorbance at 260 nm as a function of temperature under controlled buffer conditions, enabling quantitative comparisons to be made between different PNA sequences and conjugates.

2.6.1. Thermal Denaturation Studies

Thermal denaturation experiments were performed by monitoring UV absorbance at 260 nm as a function of temperature in order to evaluate the stability of the miRNA-221-5p duplexes formed with either the complementary anti-miRNA PNA or the non-complementary scrambled sequence (Table 1.5). The melting temperature (T_m), defined as the midpoint of the duplex-to-single-strand transition, was determined by fitting the experimental data to a sigmoidal melting curve.

Table 1.5. Summary of the T_m found.

Sample	T_m (°C)
miRNA-221-5p	57
miRNA-221-5p/anti-miRNA	75
miRNA-221-p/scramble	61



Free miRNA-221-5p exhibited a T_m of 57°C, which is consistent with the moderate stability typically observed for single-stranded RNA sequences that are capable of partial self-pairing (Figure 1.71).

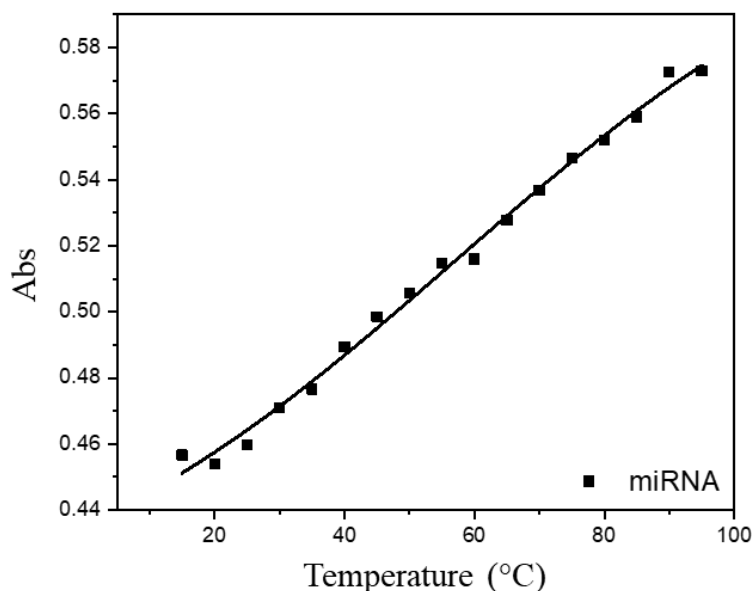


Figure 1.71. UV-melting curves at 260 nm of miRNA 221-5p at 2.5 μ M concentration in 10 mM NaCl, 10 mM phosphate buffer pH=7.4.

Upon hybridisation with anti-miRNA-221-5p PNA, however, the melting temperature increased markedly to 75°C, reflecting the formation of a highly stable, fully complementary heteroduplex. This substantial thermal stabilisation confirms the strong Watson-Crick base pairing and efficient base stacking interactions that are characteristic of PNA-RNA hybrids (Figure 1.72).

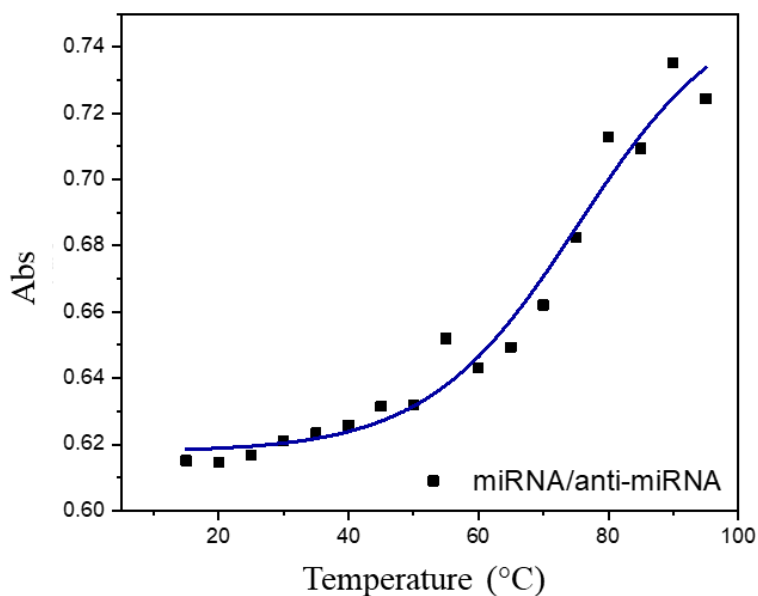


Figure 1.72. UV-melting curves at 260 nm of miRNA 221-5p/anti-miRNA-221-5p duplex at 2.5 μ M concentration in 10 mM NaCl, 10 mM phosphate buffer pH=7.4

In contrast, the duplex formed between miRNA-221-5p and the scrambled PNA had a lower T_m of 61°C, which was only slightly higher than that of the miRNA alone. This limited stabilisation suggests the absence of a fully matched duplex, possibly arising from partial complementarity: approximately five monomers in the scrambled sequence occupy positions capable of transient base pairing with the miRNA. These local interactions could explain the slight increase in melting temperature, but are insufficient to produce stable, ordered hybrid structure (Figure 1.73).

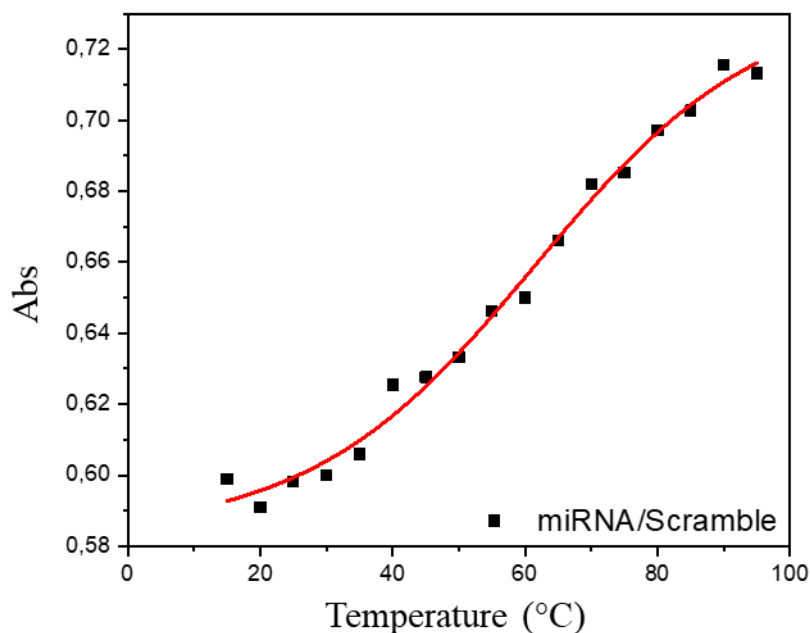


Figure 1.73. UV-melting curves at 260 nm of miRNA 221-5p/scramble duplex at 2.5 μ M concentration in 10 mM NaCl, 10 mM phosphate buffer pH=7.4

Overall, melting curve analysis corroborates CD spectroscopy data, confirming that anti-miRNA-221-5p PNA forms a specific, thermodynamically robust and sequence-selective duplex with its target miRNA. In contrast, the scrambled control engages in only weak, non-specific or partially complementary interactions.

These observations further validate the sequence specificity and high affinity of the designed anti-miRNA-221-5p PNA, which can effectively discriminate between complementary and mismatched RNA targets.



3. Conclusion

This chapter presents a study that provides an integrated chemical and biological framework for developing PNA conjugates functionalised with CPPs. The synthetic strategy combines the precision of SPPS with the efficiency of click chemistry and disulfide mediated coupling. This enables the preparation of a range of conjugates with different structural and functional properties. Both CuAAC and disulfide bond formation were effective conjugation methodologies: the triazole-linked conjugation provided excellent chemical stability, and the disulfide-linked species allowed for potential intracellular release under reductive conditions.

Spectroscopic characterization confirmed the expected structural features of the PNA strands and demonstrated successful hybridisation with the complementary microRNA-221-5p sequence. CD and UV-melting studies revealed stable duplexes, confirming sequence specificity and base pairing fidelity. This conclusion was further supported by control experiments performed with the Scramble PNA sequence, which did not exhibit any detectable interaction with miRNA-221-5p under the same experimental conditions, indicating that the observed binding is sequence-specific.

Preliminary insights into both proliferation and cellular uptake were obtained through biological assays performed on CHO-K1 and A549 cells. CF-labelled CPPs displayed efficient and reproducible internalisation, whereas CPP-PNA conjugates exhibited more variable behaviour. These inconsistent uptake profiles were attributed to aggregation phenomena in aqueous solution, which can reduce entry efficiency and alter the sensitivity with which surface-bound conjugates are removed by proteolysis. This highlights the critical importance of formulation stability when assessing intracellular delivery. Notably, proliferation studies revealed initial indication of biological activity for two conjugates, CF-Tat-(triazole)-PNA-NH₂ and CF-PNT-(triazole)-PNA-NH₂, which produced a measurable reduction in A549 cell proliferation at 5 μ M. These effects were not due to membrane permeabilization or acute cytotoxicity, which supports the hypothesis that they result from the intracellular targeting of miRNA-221-5p by PNA. Although these results are preliminary, they show that sufficient amounts of some conjugates reach the cytosol to influence cell cycle progression, thus validating the overall delivery strategy.

Future studies will build upon these findings by refining the key experimental parameters. For example, optimising the incubation protocol by reintroducing the conjugate after 48-hour medium change may ensure sustained intracellular exposure. It will be essential to test conjugates containing a scrambled PNA sequence in order to verify that the observed biological effects are sequence-specific and arise from microRNA recognition. Further work will also involve exploring alternative CPPs, such as the hybrid peptide CF-P14(triazole)-PNA, a CPP that combines the structural features of Tat



and Penetratin and extended the analysis to include complementary quantitative internalisation techniques, such as confocal microscopy and flow cytometry. Formulation optimisation to minimise aggregation will also be critical for improving reproducibility and therapeutic potential.

In conclusion, this study establishes a methodological basis for the synthesis, characterisation and preliminary biological assessment of CPP-PNA conjugates. It identifies the strengths and limitations of these delivery systems and outlines the next steps required to advance their application in nucleic acid delivery and targeted modulation of oncogenic microRNA pathways.



4. Experimental Part

4.1. General Procedure for Solid-Phase Synthesis of PNAs and Peptides

Both manual and automated approaches to solid-phase synthesis were used to prepare the PNA and peptide sequences described in this thesis. The process of solid-phase synthesis can be summarized as follows:

1. **Resin swelling.** The loaded resin was placed in the syringe reactor and swollen by the addition of DMF, followed by 40 minutes of mechanical agitation. The solvent is then removed by filtration directly through the syringe.
2. **Fmoc Deprotection.** The Fmoc protecting group is removed by adding either a 20% piperidine or a 5% w/v piperazine solution containing 2% DBU in DMF. The mixture is mechanically agitated for 5 minutes, followed by a second addition of the same solution and agitation for an additional 15 minutes. The solution is removed by filtration.
3. **Washing step.** After each operation, the solution is removed by filtration, and the resin is washed with five portions of DMF. Each wash consists of adding DMF, agitating for a few minutes and filtering off the solvent.
4. **Coupling reaction.** The coupling mixture is prepared using 3 equiv. of Fmoc-AminoAcid-OH or 2 equiv. of Fmoc-AEG(NB)-OH, together with an equimolar amount of the chosen activator in anhydrous DMF. This solution is then transferred to the reactor containing the resin. When using HBTU/HOBt or HATU as activators, 6 equiv. of DIPEA are added, or 2 equiv. of DIC are added when using Oxyma Pure. The mixture is mechanically agitated for 50 minutes for amino acids or 75-90 minutes for PNA monomers. Then, the resin is washed with five portions of DMF (for double coupling, after filtering the reaction mixture, a second portion of the activated monomer is added, and the resin is agitated before washing).
5. **Chain Elongation.** Steps 2 to 4 are then repeated for the incorporation of subsequent monomers until the desired sequence is complete.

If a fluorescent dye is to be introduced at N-terminus, a coupling solution is prepared by replacing the amino acid with the appropriate fluorescent probe. The same solid-phase coupling procedure is then applied, except the reaction is allowed to proceed overnight under mechanical agitation.



4.2. Coupling of the First Amino Acid onto 2-Chlorotrityl Chloride Resin

The first amino acid is attached to the resin in a procedure that is specific to this step and distinct from those used for subsequent coupling steps.

1. **Resin swelling.** The resin was swollen by adding DMF (10 mL for each gram of resin) and gently agitated under a nitrogen stream for 40 minutes. After swelling, the solvent was removed by filtration, either under vacuum in automated synthesis or by nitrogen pressure in manual synthesis.
2. **Coupling reaction.** The coupling mixture was prepared by dissolving 0.8 equivalents of the Fmoc-protected amino acid (Fmoc-Cys(Trt)-OH) and two equivalents of DIPEA (relative to the amino acid) in 10 mL of anhydrous DMF per gram of resin. The swollen resin was then suspended in this solution and agitated under nitrogen for 1 hour to allow the amino acid to be covalently attached to the resin via the formation of the trityl ester linkage.
3. **Washing.** At the end of the reaction, the coupling solution was removed by filtration, and the resin was washed five times with DMF to eliminate any unreacted reagents or by-products.
4. **Capping of unreacted sites.** To deactivate any unreacted chlorotrityl sites, the resin was treated with a DMF/MeOH/DIPEA mixture (80/15/5, v/v/v) and agitated under nitrogen for 30 minutes. This step ensures, that all the remaining reactive sites are capped, thereby preventing truncated sequences during subsequent elongation steps.
5. **Peptide chain elongation.** The standard coupling-deprotection-washing sequence described in the previous paragraph was then repeated for the addition of each subsequent amino acid until the desired peptide sequence was completed.

4.3. Determination of Resin Loading after Fmoc Deprotection by UV-Vis Spectroscopy

Resin loading was determined spectrophotometrically following Fmoc deprotection.

1. **Fmoc removal.** The protected resin was treated with 20% piperidine in DMF and stirred for 5 minutes. A second aliquot of the same solution was then added, after which the suspension was agitated for a further 15 minutes to ensure complete cleavage of the Fmoc.
2. **Filtration and washing.** After each step, the solution was removed by filtration, and the resin was washed five times with DMF to eliminate piperidine.



3. **Collection of deprotection solutions.** All the resulting filtrate and washing solutions were combined in a volumetric flask and diluted with MeOH to the required volume.
4. **Sample dilution.** An aliquot of the combined solution was further diluted tenfold with MeOH prior to UV-Vis analysis.
5. **Spectrophotometric determination.** The absorbance of the diluted solution was measured at 300 nm which corresponds to the characteristic absorption of the dibenzofulvene-piperidine adduct. Using Beer-Lambert's law, the concentration of the Fmoc group was calculated. The resin loading (mmol of amino acid per gram of resin) was then determined from this value.

4.4. Cleavage from the Resin

At the end of the solid-phase synthesis process, the peptide or PNA is cleaved from the resin. The cleavage conditions are selected according to the amino acid composition of the sequence. A standard cleavage cocktail consisting of TFA/TIS/H₂O (95:2.5:2.5 v/v/v) or TFA/DODT/H₂O/TIS (95:2.5:2.5:1 v/v/v/v) is typically employed. If activation of Cys is required, 5 equiv. of the activator 2,2'-dithiobis(5-nitropyridine) are added. The cleavage reaction is carried out under mechanical agitation for 3-4 hours. The crude reaction solution is then collected by filtration and most of the solvent is evaporated under a gentle nitrogen stream. The PNA or peptide is then precipitated by adding cold diethyl ether (Et₂O). The resulting suspension is then centrifuged three times at 6500 rpm for six minutes. The supernatant is removed and replaced with a fresh aliquot of cold diethyl ether after each cycle.

4.5. Analytical Cleavage

An analytical cleavage is performed by removing a small portion of the resin directly from the reactor and applying the above-described protocol using a reduced volume of the cleavage cocktail. The resulting sample is analysed by HPLC-MS to confirm the presence of the chromatographic peak corresponding to the expected molecular mass of the product.



4.6. Synthesis

4.6.1. Alkyne-AntimiRNA-221-5p Synthesis

Sequence:

Alkyne-AAATCTACATTGTATGCC-KK-NH₂

Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)	
Rink Amide		0.5		0.150		0.075	
Step	Residue	Exc. (equiv.)	Fmoc-AA/AEG(NB)-OH Coupling (mg)	Activating Agent 1 (g)		Activating Agent 2 (μL)	
1	K (Lys)	3	105.41	Oxyma	32.0	DIC	34.8
2	K (Lys)	3	105.41	Oxyma	32.0	DIC	34.8
3	C (Cytosine)	2	105.26	Oxyma	21.3	DIC	23.2
4	C (Cytosine)	2	105.26	Oxyma	21.3	DIC	23.2
5	G (Guanine)	2	111.26	Oxyma	21.3	DIC	23.2
6	T (Thymine)	2	75.98	Oxyma	21.3	DIC	23.2
7	A (Adenine)	2	108.86	Oxyma	21.3	DIC	23.2
8	T (Thymine)	2	75.98	Oxyma	21.3	DIC	23.2
9	G (Guanine)	2	111.26	Oxyma	21.3	DIC	23.2
10	T (Thymine)	2	75.98	Oxyma	21.3	DIC	23.2
11	T (Thymine)	2	75.98	Oxyma	21.3	DIC	23.2
12	A (Adenine)	2	108.86	Oxyma	21.3	DIC	23.2
13	C (Cytosine)	2	105.26	Oxyma	21.3	DIC	23.2
14	A (Adenine)	2	108.86	Oxyma	21.3	DIC	23.2
15	T (Thymine)	2	75.98	Oxyma	21.3	DIC	23.2
16	C (Cytosine)	2	105.26	Oxyma	21.3	DIC	23.2
17	T (Thymine)	2	75.98	Oxyma	21.3	DIC	23.2
18	A (Adenine)	2	108.86	Oxyma	21.3	DIC	23.2
19	A (Adenine)	2	108.86	Oxyma	21.3	DIC	23.2
20	A (Adenine)	2	108.86	Oxyma	21.3	DIC	23.2
21	Alkyne	3	24.83 μL	Oxyma	32.0	DIC	34.8

Experimental Part



The PNA was synthesised manually using standard solid-phase procedures.

Fmoc deprotection was initially performed using a 20% piperidine solution in DMF up to the Fmoc-AEG(T)-OH at position 14. From that point onwards, a 5% piperazine solution containing 2% DBU in DMF was used.

The alkyne moiety was introduced using 5-hexyno-1-oic acid.

An analytical cleavage was performed every six nucleobases to monitor chain elongation.

Crude reaction mixtures were analysed by HPLC-MS under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5 μ m, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using DAISOGEL SP120 C18, 300Å, 10 μ m column. Increasing amounts of the crude product were injected progressively over the course of five consecutive runs. The elution was conducted with a gradient of 10-40% B over 30 minutes.

The fractions displaying the highest purity were pooled, the acetonitrile was removed under reduced pressure and the PNA was subsequently lyophilised.

Final yield after purification: 5%

Purity (HPLC): 98%

Retention Time: 8.94 min

Molecular formula: C₂₁₂H₂₇₃N₁₀₅O₅₇, [M+H]¹⁺ Calculated: 5204.7

MS: [M+2H]²⁺=2603.2; [M+3H]³⁺=1735.8; [M+4H]⁴⁺=1302.1



4.6.2. Cys-AntimiRNA-221-5p Synthesis

Sequence:



Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)	
Rink Amide		0.5		0.150		0.075	
Step	Residue	Exc. (equiv.)	Fmoc-AA/AEG(NB)-OH Coupling (mg)	Activating Agent 1 (g)		Activating Agent 2 (μL)	
1	K (Lys)	3	105.41	Oxyma	32.0	DIC	34.8
2	K (Lys)	3	105.41	Oxyma	32.0	DIC	34.8
3	C (Cytosine)	2	105.26	Oxyma	21.3	DIC	23.2
4	C (Cytosine)	2	105.26	Oxyma	21.3	DIC	23.2
5	G (Guanine)	2	111.26	Oxyma	21.3	DIC	23.2
6	T (Thymine)	2	75.98	Oxyma	21.3	DIC	23.2
7	A (Adenine)	2	108.86	Oxyma	21.3	DIC	23.2
8	T (Thymine)	2	75.98	Oxyma	21.3	DIC	23.2
9	G (Guanine)	2	111.26	Oxyma	21.3	DIC	23.2
10	T (Thymine)	2	75.98	Oxyma	21.3	DIC	23.2
11	T (Thymine)	2	75.98	Oxyma	21.3	DIC	23.2
12	A (Adenine)	2	108.86	Oxyma	21.3	DIC	23.2
13	C (Cytosine)	2	105.26	Oxyma	21.3	DIC	23.2
14	A (Adenine)	2	108.86	Oxyma	21.3	DIC	23.2
15	T (Thymine)	2	75.98	Oxyma	21.3	DIC	23.2
16	C (Cytosine)	2	105.26	Oxyma	21.3	DIC	23.2
17	T (Thymine)	2	75.98	Oxyma	21.3	DIC	23.2
18	A (Adenine)	2	108.86	Oxyma	21.3	DIC	23.2
19	A (Adenine)	2	108.86	Oxyma	21.3	DIC	23.2
20	A (Adenine)	2	108.86	Oxyma	21.3	DIC	23.2
21	C (Cys)	3	131.78	Oxyma	32.0	DIC	34.8

Experimental Part



The PNA was synthesised manually using standard solid-phase procedures.

Fmoc deprotection was initially performed using a 20% piperidine solution in DMF up to the Fmoc-AEG(T)-OH at position 14. From that point onwards, a 5% piperazine solution containing 2% DBU in DMF.

An analytical cleavage was performed every six nucleobases to monitor chain elongation.

Crude reaction mixtures were analysed by HPLC-MS under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5 μ m, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using DAISOGEL SP120 C18, 300Å, 10 μ m column. Increasing amounts of the crude product were injected progressively over the course of five consecutive runs. The elution was conducted with a gradient of 10-40% B over 30 minutes.

The fractions displaying the highest purity were pooled, the acetonitrile was removed under reduced pressure and the PNA was subsequently lyophilised.

Final yield after purification: 5%

Purity (HPLC): 96%

Retention Time: 8.67 min

Molecular formula: C₂₀₉H₂₆₇N₁₀₆O₅₇S, [M+H]¹⁺ Calculated: 5211.8

MS: [M+2H]²⁺=2606.7; [M+3H]³⁺=1738.1; [M+4H]⁴⁺=1303.8



4.6.3. Scramble PNA Sequence Synthesis

Sequence:



Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)	
Rink Amide		0.5		0.050		0.025	
Step	Residue	Exc. (equiv.)	Fmoc-AA/AEG(NB)-OH Coupling (mg)	Activating Agent 1 (g)		Activating Agent 2 (μL)	
1	K (Lys)	3	35.14	Oxyma	10.7	DIC	11.6
2	K (Lys)	3	35.14	Oxyma	10.7	DIC	11.6
3	A (Adenine)	2	36.29	Oxyma	7.1	DIC	7.7
4	A (Adenine)	2	36.29	Oxyma	7.1	DIC	7.7
5	T (Thymine)	2	25.33	Oxyma	7.1	DIC	7.7
6	A (Adenine)	2	36.29	Oxyma	7.1	DIC	7.7
7	T (Thymine)	2	25.33	Oxyma	7.1	DIC	7.7
8	A (Adenine)	2	36.29	Oxyma	7.1	DIC	7.7
9	C (Cytosine)	2	35.09	Oxyma	7.1	DIC	7.7
10	C (Cytosine)	2	35.09	Oxyma	7.1	DIC	7.7
11	G (Guanine)	2	37.09	Oxyma	7.1	DIC	7.7
12	A (Adenine)	2	36.29	Oxyma	7.1	DIC	7.7
13	C (Cytosine)	2	35.09	Oxyma	7.1	DIC	7.7
14	A (Adenine)	2	36.29	Oxyma	7.1	DIC	7.7
15	T (Thymine)	2	25.33	Oxyma	7.1	DIC	7.7
16	T (Thymine)	2	25.33	Oxyma	7.1	DIC	7.7
17	T (Thymine)	2	25.33	Oxyma	7.1	DIC	7.7
18	C (Cytosine)	2	35.09	Oxyma	7.1	DIC	7.7
19	T (Thymine)	2	25.33	Oxyma	7.1	DIC	7.7
20	G (Guanine)	2	37.09	Oxyma	7.1	DIC	7.7

Experimental Part



The PNA was synthesised manually using standard solid-phase procedures.

Fmoc deprotection was initially performed using a 20% piperidine solution in DMF up to the Fmoc-AEG(A)-OH at position 14. From that point onwards, a 5% piperazine solution containing 2% DBU in DMF.

An analytical cleavage was performed every six nucleobases to monitor chain elongation.

Crude reaction mixtures were analysed by HPLC-MS under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5 μ m, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using DAISOGEL SP120 C18, 300Å, 10 μ m column. Increasing amounts of the crude product were injected progressively over the course of three consecutive runs. The elution was conducted with a gradient of 10-40% B over 30 minutes.

The fractions displaying the highest purity were pooled, the acetonitrile was removed under reduced pressure and the PNA was subsequently lyophilised.

Final yield after purification: 6%

Purity (HPLC): 98%

Retention Time: 8.43 min

Molecular formula: C₂₀₆H₂₆₇N₁₀₅O₅₆, [M+H]¹⁺ Calculated: 5106.8

MS: [M+2H]²⁺= 2554.8; [M+3H]³⁺= 1703.5; [M+4H]⁴⁺= 1277.9



4.6.4. CF-Tat-Ahx-N₃-NH₂ Sequence Synthesis

Sequence:

CF-GRKKRRQRRR-Ahx-Aza-NH₂

Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)
Rink Amide		0.65		0.250		0.1625
Residue		MW (g/mol)	Exc. (equiv..)	N° coupling	Mass (mg)	Volume DMF (mL)
Aza	L-Aza	352.3	3	2	429.41	2.5
X	Ahx	353.4	3	2	344.57	2.5
R	Arg	648.8	3	12	3527.85	12.5
Q	Gln	610.7	3	2	664.14	2.5
K	Lys	468.5	3	4	917.09	4.5
G	Gly	297.3	3	2	323.31	2.5
HBTU		379.3	3	22	3712.40	22.5
HOBt		153.3	3	22	1498.47	
Residue		Exc. (equiv.)	Mass (mg)	Activating Agent 1 (mg)		
CF		3	183.46	HATU 185.36		

The peptide was synthesised automatically using standard solid-phase procedures. All the couplings were performed as double couplings.

Fmoc deprotection was performed using a 20% piperidine solution in DMF.

Crude reaction mixtures were analysed by HPLC-MS under the following conditions: 5-95% B in 30 min or 5-45% in 30 min on a Phenomenex Kinetex C18, 3.5µm, 100Å, 4.6x100 mm column.

Purification was performed by low pressure LC, Biotage Isolera Prime, using SNAP Ultra C18 inverse phase column. 100mg of product were injected over the course of two consecutive runs. The elution was conducted with a gradient of 0-50% B over 50 minutes.

The fractions displaying the highest purity were pooled, the acetonitrile was removed under reduced pressure, and the peptide was subsequently lyophilised.

Final yield after purification: 20%

Purity (HPLC): 98%

Retention Time: 10.36 min

Molecular formula: C₈₅H₁₃₅N₃₇O₁₉, [M+H]¹⁺Calculated: 1979.1

MS: [M+H]⁺= 1977.9; [M+2H]²⁺= 989.8; [M+3H]³⁺= 660.3



4.6.5. CF-PNT-Ahx-N₃-NH₂ Sequence Synthesis

Sequence:

CF-RQIKIWFQNRRNleKWKK-Ahx-Aza-NH₂

Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)
Rink Amide		0.65		0.250		0.1625
Residue		MW (g/mol)	Exc. (equiv..)	N° coupling	Mass (mg)	Volume DMF (mL)
Aza	L-Aza	352.3	3	2	429.41	2.5
X	Ahx	353.4	3	2	344.57	2.5
K	Lys	468.5	3	8	1941.35	8.5
R	Arg	648.8	3	6	2055.89	6.5
Q	Gln	610.7	3	4	1339.72	4.5
W	Trp	526.6	3	4	1155.23	4.5
Nle	Nle	353.41	3	2	430.72	2.5
F	Phe	387.4	3	2	472.1	2.5
I	Ile	353.4	3	4	775.27	4.5
N	Asn	596.7	3	2	727.23	2.5
HBTU		379.3	3	36	6749.17	36.5
HOBt		153.3	3	36	2724.22	
Residue		Exc. (equiv.)	Mass (mg)	Activating Agent 1 (mg)		
CF		3	183.46	HATU 185.36		

The peptide was synthesised automatically using standard solid-phase procedures. All the couplings were performed as double couplings.

Fmoc deprotection was initially performed using a 20% piperidine solution in DMF.

Crude reaction mixtures were analysed by HPLC-MS under the following conditions: 5-95% B in 30 min or 5-45% B in 30 min on a Phenomenex Kinetex C18, 3.5µm, 100Å, 4.6x100 mm column.

Purification was performed by low pressure LC, Biotage Isolera Prime, using SNAP Ultra C18 inverse phase column. 100mg of product were injected over the course of two consecutive runs. The elution was conducted with a gradient of 0-50% B over 50 minutes.

The fractions displaying the highest purity were pooled, the acetonitrile was removed under reduced pressure, and the peptide was subsequently lyophilised.

Final yield after purification: 22%

Purity (HPLC): 98%

Retention Time: 14.06 min

Molecular formula: C₁₃₅H₁₉₅N₃₉O₂₈, [M+H]¹⁺ Calculated: 2812.6

MS: [M+H]⁺ = 2811.6; [M+2H]²⁺ = 1405.7; [M+3H]³⁺ = 937.6



4.6.6. CF-P14-Ahx-N₃-NH₂ Sequence Synthesis

Sequence: CF-RKKRWFRRRRPKWKK-Ahx-Aza-NH₂

Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)
Rink Amide		0.65		0.250		0.1625
Residue		MW (g/mol)	Exc. (equiv..)	N° coupling	Mass (mg)	Volume DMF (mL)
Aza	L-Aza	352.3	3	2	429.41	2.5
X	Ahx	353.4	3	2	344.57	2.5
K	Lys	468.5	3	10	2398.13	10.5
W	Trp	526.6	3	4	1155.23	4.5
P	Pro	337.4	3	2	411.21	2.5
R	Arg	648.8	3	12	3953.63	12.5
F	Phe	387.4	3	2	472.14	2.5
HBTU		379.3	3	30	5639.72	30.5
HOBt		153.3	3	30	2276.41	
Residue		Exc. (equiv.)	Mass (mg)	Activating Agent 1 (mg)		
CF		3	183.46	HATU 185.36		

The peptide was synthesised automatically using standard solid-phase procedures. All the couplings were performed as double couplings.

Fmoc deprotection was performed using a 20% piperidine solution in DMF.

Crude reaction mixtures were analysed by HPLC-MS under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5µm, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using Luna C18, 300Å, 10µm column. Increasing amounts of the crude product were injected progressively over the course of five consecutive runs.

The elution was conducted with a gradient of 10-40% B over 30 minutes.

The fractions displaying the highest purity were pooled, the acetonitrile was removed under reduced pressure, and the product was subsequently lyophilised.

Final yield after purification: 8%

Purity (HPLC): 98%

Retention Time: 11.72 min

Molecular formula: C₁₃₂H₁₉₅N₄₅O₂₄, [M+H]¹⁺ Calculated: 2796.3

MS: [M+H]⁺ = 2795.7; [M+2H]²⁺ = 1398.0; [M+3H]³⁺ = 932.2



4.6.7. CF-Tat-Ahx-Cys(TNP)-OH Sequence Synthesis

Sequence:



Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)
2-CTC		0.65		0.250		0.1625
Residue		MW (g/mol)	Exc. (equiv..)	N° coupling	Mass (mg)	Volume DMF (mL)
C	Cys	585.7	3	2	713.82	2.5
X	Ahx	353.4	3	2	344.57	2.5
R	Arg	648.8	3	12	3527.85	12.5
Q	Gln	610.7	3	2	664.14	2.5
K	Lys	468.5	3	4	917.09	4.5
G	Gly	297.3	3	2	323.31	2.5
HBTU		379.3	3	22	3712.40	22.5
HOBt		153.3	3	22	1498.47	
Residue		Exc. (equiv.)	Mass (mg)	Activating Agent 1 (mg)		
CF		3	183.46	HATU 185.36		

The peptide was synthesised automatically using standard solid-phase procedures. All the couplings were performed as double couplings.

Fmoc deprotection was initially performed using a 20% piperidine solution in DMF.

The cleavage solution used is TFA+TIS+H₂O (95:2.5:2.5) + 5 equivalents of DTNP activator.

Crude reaction mixtures were analysed by HPLC-MS under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5µm, 100Å, 4.6x100 mm column.

Final yield: 78%

Purity (HPLC): 92%

Retention Time: 12.35 min

Molecular formula: C₉₀H₁₃₇N₃₅O₂₂S₂, [M+H]¹⁺ Calculated: 2125.9

MS: [M+H]⁺ = 2124.9; [M+2H]²⁺ = 1062.9; [M+3H]³⁺ = 709.0



4.6.8. CF-PNT-Ahx-Cys(TNP)-OH Sequence Synthesis

Sequence:



Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)
2-CTC		0.65		0.250		0.1625
Residue		MW (g/mol)	Exc. (equiv..)	N° coupling	Mass (mg)	Volume DMF (mL)
C	Cys	585.7	3	2	713.82	2.5
X	Ahx	353.4	3	2	344.57	2.5
K	Lys	468.5	3	8	1941.35	8.5
R	Arg	648.8	3	6	2055.89	6.5
Q	Gln	610.7	3	4	1339.72	4.5
W	Trp	526.6	3	4	1155.23	4.5
Nle	Nle	353.41	3	2	430.72	2.5
F	Phe	387.4	3	2	472.1	2.5
I	Ile	353.4	3	4	775.27	4.5
N	Asn	596.7	3	2	727.23	2.5
HBTU		379.3	3	36	6749.17	36.5
HOBt		153.3	3	36	2724.22	
Residue		Exc. (equiv.)	Mass (mg)	Activating Agent 1 (mg)		
CF		3	183.46	HATU 185.36		

The peptide was synthesised automatically using standard solid-phase procedures. All the couplings were performed as double couplings.

Fmoc deprotection was performed using a 20% piperidine solution in DMF.

The cleavage solution used is TFA+TIS+H₂O (95:2.5:2.5) + 5 equiv. of DTNP activator.

Crude reaction mixtures were analysed by HPLC-MS under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5µm, 100Å, 4.6x100 mm column.

Final yield: 42%

Purity (HPLC): 82%

Retention Time: 15.00 min

Molecular formula: C₁₄₀H₁₉₈N₃₈O₃₀S₂, [M+H]¹⁺ Calculated: 2958.2

MS: [M+H]⁺ = 2957.9; [M+2H]²⁺ = 1478.8; [M+3H]³⁺ = 986.3



4.7. Optimized Procedure for Click Chemistry Reactions

The alkyne-functionalised PNA (1.0 equiv.) and the azide-functionalised CF-CPP-N₃-NH₂ (1.3 equiv.) were dissolved in a 1:3 (v/v) mixture of H₂O and *t*-BuOH to afford a homogeneous reaction solution. CuSO₄·H₂O (4.0 equiv.) and sodium L-ascorbate (16.0 equiv.) were then added to generate the Cu(I) catalyst *in situ*. The reaction mixture was stirred overnight at room temperature under gentle agitation.

The reaction progress was monitored by HPLC-MS using a linear gradient of 5-95% solvent B over 30 minutes on a Phenomenex Kinetex C18 column, 3.5μm, 100 Å, 4.6x100 mm.

After confirming the formation of the desired conjugate, the reaction mixture was lyophilised to remove solvents and the crude product was purified by preparative HPLC to produce the pure CF-CPP(triazole)PNA conjugate.

4.7.1. Click Reaction of CF-Tat(Triazole)PNA-NH₂

The CF-Tat-N₃-NH₂ peptide (1.0 equiv.) and then alkyne-functionalised PNA (1.0 equiv.) were dissolved in a 1:3 (v/v) mixture of water and *t*-BuOH. CuSO₄·H₂O (4.0 equiv.) and sodium L-ascorbate (16.0 equiv.) were then added. The reaction mixture was stirred overnight at room temperature under gentle agitation to ensure complete conjugation.

Once complete, the mixture was directly purified by preparative HPLC using a DAISOGEL SP120 C18, 300 Å, 10 μm column. Increasing amounts of the crude product were injected progressively over three consecutive purification runs. Elution was performed using a linear gradient of 10-40% solvent B over 30 minutes.

The fractions displaying the highest purity were pooled and the acetonitrile removed under reduced pressure. The resulting aqueous solution was lyophilised to yield the purified PNA conjugate.

Final yield after purification: 30%

Purity (HPLC): 99%

Retention Time: 9.90 min

Molecular formula: C₂₉₆H₄₀₆N₁₄₂O₇₆, [M+H]¹⁺ Calculated: 7181.5

MS: [M+3H]³⁺ = 2396.1; [M+4H]⁴⁺ = 9

4.7.2. Click Reaction of CF-PNT(Triazole)PNA-NH₂

The CF-PNT-N₃-NH₂ peptide (1.3 equiv.) and then alkyne-functionalised PNA (1.0 equiv.) were dissolved in a 1:3 (v/v) mixture of water and *t*-BuOH. CuSO₄·H₂O (4.0 equiv.) and sodium L-



ascorbate (16.0 equiv.) were then added. The reaction mixture was stirred overnight at room temperature under gentle agitation to ensure complete conjugation.

Once complete, the mixture was directly purified by preparative HPLC using a DAISOGEL SP120 C18, 300 Å, 10 µm column. Increasing amounts of the crude product were injected progressively over three consecutive purification runs. Elution was performed using a linear gradient of 10-40% solvent B over 30 minutes.

The fractions displaying the highest purity were pooled and the acetonitrile removed under reduced pressure. The resulting aqueous solution was lyophilised to yield the purified PNA conjugate.

Final yield after purification: 49%

Purity (HPLC): 97%

Retention Time: 13.49 min

Molecular formula: $C_{346}H_{467}N_{145}O_{76}$, $[M+H]^+$ Calculated: 8020.0

MS: $[M+3H]^{3+} = 2671.7$; $[M+4H]^{4+} = 2004.0$; $[M+5H]^{5+} = 1603.3$

4.7.3. Click Reaction of CF-P14(Triazole)PNA-NH₂

The CF-P14-N₃-NH₂ peptide (1.3 equiv.) and then alkyne-functionalised PNA (1.0 equiv.) were dissolved in a 1:3 (v/v) mixture of water and *t*-BuOH. CuSO₄ · H₂O (4.0 equiv.) and sodium L-ascorbate (16.0 equiv.) were then added. The reaction mixture was stirred overnight at room temperature under gentle agitation to ensure complete conjugation.

Once complete, the mixture was directly purified by preparative HPLC using a DAISOGEL SP120 C18, 300 Å, 10 µm column. Increasing amounts of the crude product were injected progressively over three consecutive purification runs. Elution was performed using a linear gradient of 10-40% solvent B over 30 minutes.

The fractions displaying the highest purity were pooled and the acetonitrile removed under reduced pressure. The resulting aqueous solution was lyophilised to yield the purified PNA conjugate.

Final yield after purification: 53%

Purity (HPLC): 98%

Retention Time: 11.34 min

Molecular formula: $C_{344}H_{342}N_{150}O_{81}$, $[M+H]^+$ Calculated: 8000.2

MS: $[M+3H]^{3+} = 2666.3$; $[M+4H]^{4+} = 2000.1$; $[M+5H]^{5+} = 1600.3$



4.8. General Procedure for Conjugation via Disulfide Bridge Formation

First, an ammonium acetate buffer solution (NH_4COOH , pH 4.5) was prepared to create a mildly acidic environment suitable for thiol-disulfide exchange. The activated internalising peptide, obtained during the resin cleavage process, was then dissolved in 2 mL of this solution. The PNA cargo was then added gradually to the peptide solution under mechanical stirring. The two components were mixed in equimolar amounts and stirred for two hours at room temperature to allow complete disulfide bond formation between the cysteine residue of the two molecules.

Reaction progress was monitored by analytical HPLC, with small aliquots of the reaction mixture collected. The formation of the expected conjugate was confirmed by HPLC-MS analysis. Once the reaction was complete, the mixture was purified using preparative HPLC to isolate the desired conjugate.

4.8.1. Reaction of CF-Tat-S-S-PNA-NH₂

The CF-Tat-Cys(TNP)-OH peptide (1.0 equiv.) was dissolved in a NH_4COOH buffer solution at pH 4.5. The cysteine-functionalised PNA, also dissolved in the same buffer, were then added. The reaction mixture was stirred for 2 hours at room temperature under gentle agitation to ensure complete conjugation.

Once complete, the mixture was directly purified by preparative HPLC using a DAISOGEL SP120 C18, 300 Å, 10 µm column. Increasing amounts of the crude product were injected progressively over two consecutive purification runs. Elution was performed using a linear gradient of 10-40% solvent B over 30 minutes.

The fractions displaying the highest purity were pooled and the acetonitrile removed under reduced pressure. The resulting aqueous solution was lyophilised to yield the purified PNA conjugate.

Final yield after purification: 28%

Purity (HPLC): 97%

Retention Time: 10.08 min

Molecular formula: $\text{C}_{294}\text{H}_{406}\text{N}_{140}\text{O}_{76}\text{S}_2$, $[\text{M}+\text{H}]^{1+}$ Calculated: 7181.7

MS: $[\text{M}+3\text{H}]^{3+} = 2396.1$; $[\text{M}+4\text{H}]^{4+} = 1795.8$; $[\text{M}+5\text{H}]^{5+} = 1437.0$

4.8.2. Reaction of CF-PNT-S-S-PNA-NH₂

The CF-PNT(TNP)-Cys-OH peptide (1.0 equiv.) was dissolved in a NH_4COOH buffer solution at pH 4.5. The cysteine-functionalised PNA, also dissolved in the same buffer, were then added. The reaction mixture was stirred for 2 hours at room temperature under gentle agitation to ensure complete conjugation.



Once complete, the mixture was directly purified by preparative HPLC using a DAISOGEL SP120 C18, 300 Å, 10 µm column. Increasing amounts of the crude product were injected progressively over two consecutive purification runs. Elution was performed using a linear gradient of 10-40% solvent B over 30 minutes.

The fractions displaying the highest purity were pooled and the acetonitrile removed under reduced pressure. The resulting aqueous solution was lyophilised to yield the purified PNA conjugate.

Final yield after purification: 27%

Purity (HPLC): 97%

Retention Time: 13.48 min

Molecular formula: $C_{344}H_{467}N_{143}O_{84}S_2$, $[M+H]^+_{\text{Calculated}}: 8013.4$

MS: $[M+3H]^{3+} = 2671.7$; $[M+4H]^{4+} = 2004.1$; $[M+5H]^{5+} = 1603.3$

4.9. Culture Cell

CHO-K1 cells were cultured in DMEM-F12 medium and supplemented with 10% fetal bovine serum (FBS), penicillin (100,000 IU/L), streptomycin (100,000 IU/L), and amphotericin B (1 mg/L) in a humidified atmosphere containing 5% CO₂ at 37 °C.

A549 cells were cultured in DMEM medium and supplemented with 10% fetal bovine serum (FBS), penicillin (100,000 IU/L), streptomycin (100,000 IU/L), and amphotericin B (1 mg/L) in a humidified atmosphere containing 5% CO₂ at 37 °C.

4.10. Determination of the Linear Range for CCK-8 Assay

In order to determine the linear detection range of the CCK-8 colorimetric assay for the reliable quantification of cell proliferation, CHO-K1 and A549 cells were seeded in a 96-well plate at various densities (2500, 5000, 10000, 20000, 30000, and 40000 cells/well) in a 96-well plate, with each well containing 100 µL of complete culture medium (DMEM-F12 or DMEM supplemented with 10% FBS).

The cells were then allowed to adhere and recover overnight at 37°C in a 5% CO₂ incubator. Wells without cells were included as positive controls.

The following day, the culture medium was aspirated, and the cells were gently washed once with HBSS. Then, a 10% v/v solution of CCK-8 reagent in DMEM-F12 or DMEM (100µL/well) was added, and the plate was incubated for 3 hours at 37°C.

Absorbance was measured at 450 nm (signal) and 620 nm (background) using a microplate reader. All measurements were reported as mean absorbance values after subtraction of the blank signal.



4.11. Cell Proliferation Assay with CF-Conjugated Cell-Proliferation Peptides

The CHO-K1 cells were seeded in 96-well plates at a density of 2500 cells per well to allow adhesion and proliferation prior to treatment. Based on growth kinetics, this corresponded to approximately 5000 and 40000 cells per well at the times of peptide exposure and final analysis (72 hours), respectively.

The cells were then incubated with 2.5 μ M and 7.5 μ M of the CPPs under four different incubation and feeding conditions, summarised below:

Group	Condition
Group 1	5-hour peptide incubation → medium change → daily FBS supplementation
Group 2	5-hour peptide incubation → medium change → daily FBS supplementation → additional medium change at 48h
Group 3	72-hour peptide incubation → daily FBS supplementation only
Group 4	48-hour peptide incubation → daily FBS supplementation → medium change at 48h

All assays were performed for 72 hours after initial seeding. For each condition, blank controls (no cells) and untreated controls (no peptide) were included.

After incubation, the culture medium was replaced with 100 μ L of fresh medium containing 10% CCK8 reagent (v/v) and the sample was incubated for two hours at 37°C. Absorbance was then measured at 450 nm (signal) and 620 nm (background) using a microplate reader.

4.12. Cell Proliferation Assay General Procedure

The cell proliferation assay was performed using the colorimetric CCK-8 method, which measures cell viability and proliferation by quantifying cellular metabolic activity.

- **Cell culture and seeding.** A549 human lung carcinoma cells were cultured in DMEM medium supplemented with 10% foetal bovine serum. The cells were seeded in 96-well plates at densities ranging from 1875 (24, 48, and 72 hours) to 3750 (1 hour) cells per well depending on the time point. Each well contained 100 μ L of complete medium and the plates were incubated overnight at 37°C in a 5% CO₂ atmosphere to allow cell adhesion.
- **Treatment.** The cells were treated with either CF-CPP-PNA conjugates, either alone or in combination with selected additives, such as chloroquine (150-500 μ M) or calcium chloride (5mM). Treatments were carried out for 1, 24, 48 or 72 hours with daily FBS supplementation, and complete medium replacement after 48 hours.



- **Assay procedure.** After the specific incubation period, the cells were washed once with HBSS. A solution of CCK-8 reagent (10% v/v) in DMEM was then added to each well (giving a total volume of 100 μ L/well). The plates were then incubated for 2-3 hours at 37°C. Absorbance was subsequently measured at 450 nm (signal) and 620 nm (background) using a microplate reader.
- **Data analysis.** The absorbance values were blank corrected and normalised to untreated control (100% viability). Data were expressed as the mean $\pm$ SD of the replicates.

4.13. Internalization Studies Protocol

Internalization studies were performed to quantify and compare the cellular uptake of fluorescently labelled CPP and CPP-PNA conjugates under controlled temperature and enzymatic conditions.

- **Standard curve without cell.** Aliquots of the 1 mM peptide stock solution in water were diluted in lysis buffer to obtain the following concentrations: 500 pM, 250 pM, 125 pM, 62.5 pM, 31.25 pM, 15.6 pM, 7.8 pM, 3.9 pM, and 1.9 pM, plus a blank. Each dilution was transferred to a 96-well plate for fluorescence measurements.
- **Standard curve with cell.** CHO-K1 or A549 cells were detached using trypsin, counted and adjusted to a final concentration of 10^6 cells per tube, then transferred into a Falcon tube. Following centrifugation for 5 minutes at 1000 rpm, the supernatant discarded, and the cells were resuspended in 3 mL of HBSS. The suspension was then aliquoted into 300 μ L per tube (10 tubes total), followed by a second centrifugation step (5 minutes, 3000 rpm). The cells were then lysed with 200 μ L of NP-40 lysis buffer and 1M NaCl. Aliquots of the same peptide concentrations used for the standard curve without cells were then added to these lysates. After sonication for 15 minutes, the fluorescence signal was measured.
- **Cell culture and seeding.** Cells (CHO-K1 or A549, depending in the experiment) were seeded at a density of 500000 cells/well in 1mL of complete medium, to have 1 million cells per well for the experiment, in 12-well plates and incubated overnight at 37°C and 5% CO₂ to allow cell attachment.
- **Peptide/conjugate incubation.** Before adding the peptides, the cells were washed with HBSS, followed by serum-free medium (DMEM/F12 or DMEM). The desired concentration of peptides or conjugates was added in 0.5mL of serum free medium, and the cells were incubated for 1 hour at either 37°C or 4°C.



- **Enzymatic treatment to remove surface-bound conjugate.** After incubation, the cells were washed three times with HBSS to remove any unbound peptide. To eliminate surface-associated fluorescence conjugate, the cells were treated with trypsin/EDTA (37°C) or pronase (4°C) for 5 minutes. For conjugates containing a disulfide linkage, a brief reduction step with TCEP (2 Mm for 1.5 minutes at 37°C) was introduced prior to enzymatic treatment in order to cleave extracellular disulfides.

- **Condition:**
 - After washing, incubate the cells with 0.5 mL/well of Pronase solution (Pronase 0.05% in 50 mM TRIS buffer) at 4°C and trypsin at 37°C.
 - After the incubation, add 100 µL/well of a complete protease inhibitor cocktail (prepared by dissolving 1 tablet in 1.5 mL of H₂O) at 4°C and 100 µL/well of trypsin inhibitor (5 mg/mL) at room temperature, followed by 100 µL/well of BSA (1 mg/mL).
 - Transfer the treated cells into Eppendorf tubes by rinsing each well with 0.5 mL of 50 mM TRIS buffer and collecting the suspension.
 - Keep all tubes on ice during handling. Centrifuge the Eppendorf tubes at 3000 rpm for 5 minutes at 4°C.
 - Discard the supernatant and resuspend the cell tube in 1 mL of 50 mM TRIS buffer containing 0.1% BSA (1 mg/mL).
 - Centrifuge again under the same conditions (5 minutes, 3000 rpm, 4°C), discard the supernatant.
 - Finally, add 200 µL of lysis solution to each tube and sonicate the samples for 15 minutes.

- **Fluorescence Analysis.** Analyze the supernatant (96wells-plate, 100µL/well) from lysed cells using fluorimetry plate reader (excitation wavelength at 494 nm, emission wavelength at 520 nm).

4.14. Conformational Studies by CD Spectroscopy

CD spectra were recorded on a Jasco-1500 spectrophotometer using Spectra Manager (Jasco) software. CD spectra were recorded at 25°C using a 1 cm quartz cell, a 320–200 nm measurement range, 50 nm/min scanning speed, 1 nm bandwidth, 4s response time, 0.5 nm data pitch and are the results of 4 scans. The measurements are taken in duplicate.



4.15. UV-melting Analysis

UV measurements were carried out on a IMPLEN NP80 UV/Vis spectrophotometer by using 1 cm quartz cells of 1 mL internal volume. Single strand PNA and complementary RNA were dissolved in ultrapure water and were annealed by warming up at 95°C and slowly cooling down to 15°C in Phosphate buffer 10 mM, 10 mM NaCl, pH = 7.4. The thermal denaturation experiment was carried out using 2.5 μ M duplex in Phosphate buffer 10 mM, 10 mM NaCl, pH = 7.4 at a 0.5°C/min scan speed, recording at 260 nm every 5°C. The measurements are taken in duplicate.

References

1. Ille, A.M., H. Lamont, and M.B. Mathews, *The Central Dogma revisited: Insights from protein synthesis, CRISPR, and beyond*. Wiley Interdiscip Rev RNA, 2022. **13**(5): p. e1718.
2. Zhao, J.H. and H.S. Guo, *RNA silencing: From discovery and elucidation to application and perspectives*. J Integr Plant Biol, 2022. **64**(2): p. 476-498.
3. Bouton, L., et al., *Small molecules modulating RNA splicing: a review of targets and future perspectives*. RSC Med Chem, 2024. **15**(4): p. 1109-1126.
4. Bhaskaran, M. and M. Mohan, *MicroRNAs: history, biogenesis, and their evolving role in animal development and disease*. Vet Pathol, 2014. **51**(4): p. 759-74.
5. Liu, K. and W. Cai, *miRNAs: Biosynthesis, mechanism of action, and applications in biological systems*. Gene Reports, 2025. **39**: p. 102208.
6. O'Brien, J., et al., *Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation*. (1664-2392 (Print)).
7. Saliminejad, K., et al., *An overview of microRNAs: Biology, functions, therapeutics, and analysis methods*. J Cell Physiol, 2019. **234**(5): p. 5451-5465.
8. Kozomara, A., M. Birgaoanu, and S. Griffiths-Jones, *miRBase: from microRNA sequences to function*. Nucleic Acids Res, 2019. **47**(D1): p. D155-d162.
9. Desvignes, T., et al., *miRNA Nomenclature: A View Incorporating Genetic Origins, Biosynthetic Pathways, and Sequence Variants*. Trends Genet, 2015. **31**(11): p. 613-626.
10. Medley, J.C., G. Panzade, and A.Y. Zinovyeva, *microRNA strand selection: Unwinding the rules*. Wiley Interdiscip Rev RNA, 2021. **12**(3): p. e1627.
11. Liu, B., et al., *Interplay between miRNAs and host genes and their role in cancer*. Briefings in Functional Genomics, 2019. **18**(4): p. 255-266.
12. Ha, M. and V.N. Kim, *Regulation of microRNA biogenesis*. Nature Reviews Molecular Cell Biology, 2014. **15**(8): p. 509-524.
13. Liu, B., et al., *Interplay between miRNAs and host genes and their role in cancer*. Brief Funct Genomics, 2018. **18**(4): p. 255-266.
14. Hynes, C. and P.K. Kakumani, *Regulatory role of RNA-binding proteins in microRNA biogenesis*. 2024. **Volume 11 - 2024**.
15. Lee, Y., et al., *MicroRNA genes are transcribed by RNA polymerase II*. Embo j, 2004. **23**(20): p. 4051-60.
16. Guo, F., *Chapter Five - Drosha and DGCR8 in MicroRNA Biogenesis*, in *The Enzymes*, F. Guo and F. Tamanoi, Editors. 2012, Academic Press. p. 101-121.
17. Han, J., et al., *The Drosha-DGCR8 complex in primary microRNA processing*. Genes Dev, 2004. **18**(24): p. 3016-27.
18. Yi, R., et al., *Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs*. Genes Dev, 2003. **17**(24): p. 3011-6.

19. Ciechanowska, K., M. Pokornowska, and A. Kurzyńska-Kokorniak, *Genetic Insight into the Domain Structure and Functions of Dicer-Type Ribonucleases*. 2021. **22**(2): p. 616.
20. Nakanishi, K., *Anatomy of RISC: how do small RNAs and chaperones activate Argonaute proteins?* (1757-7012 (Electronic)).
21. Desvignes, T., et al., *miRNA Nomenclature: A View Incorporating Genetic Origins, Biosynthetic Pathways, and Sequencing Variants*. Trends in Genetics, 2015. **31**(11): p. 613-626.
22. Nakanishi, K., *Anatomy of four human Argonaute proteins*. Nucleic Acids Res, 2022. **50**(12): p. 6618-6638.
23. Jungers, C.F. and S. Djuranovic, *Modulation of miRISC-Mediated Gene Silencing in Eukaryotes*. Front Mol Biosci, 2022. **9**: p. 832916.
24. Li, T., et al., *Translation suppresses exogenous target RNA-mediated microRNA decay*. Nature Communications, 2025. **16**(1): p. 5257.
25. Shang, R., et al., *microRNAs in action: biogenesis, function and regulation*. Nature Reviews Genetics, 2023. **24**(12): p. 816-833.
26. Yu, Y., et al., *Plant microRNA maturation and function*. Nature Reviews Molecular Cell Biology, 2025.
27. Jung, E., et al., *Global analysis of AGO2-bound RNAs reveals that miRNAs induce cleavage of target RNAs with limited complementarity*. Biochim Biophys Acta Gene Regul Mech, 2017. **1860**(11): p. 1148-1158.
28. Buhagiar, A.F. and B. Kleaveland, *To kill a microRNA: emerging concepts in target-directed microRNA degradation*. Nucleic Acids Research, 2024. **52**(4): p. 1558-1574.
29. Hu, W. and J. Collier, *What comes first: translational repression or mRNA degradation? The deepening mystery of microRNA function*. Cell Res, 2012. **22**(9): p. 1322-4.
30. Carthew, R.W. and E.J. Sontheimer, *Origins and Mechanisms of miRNAs and siRNAs*. Cell, 2009. **136**(4): p. 642-55.
31. Sarkar, N. and A. Kumar, *Paradigm shift: microRNAs interact with target gene promoters to cause transcriptional gene activation or silencing*. Experimental Cell Research, 2025. **444**(2): p. 114372.
32. Nalbant, E. and Y.Z. Akkaya-Ulum, *Exploring regulatory mechanisms on miRNAs and their implications in inflammation-related diseases*. Clinical and Experimental Medicine, 2024. **24**(1): p. 142.
33. Oliveto, S., N. Manfrini, and S. Biffo, *The power of microRNA regulation—insights into immunity and metabolism*. 2025. **599**(13): p. 1821-1851.
34. Friedman, R.C., et al., *Most mammalian mRNAs are conserved targets of microRNAs*. (1088-9051 (Print)).
35. Levantini, E. and M. Rizzo, *miRNAs: From Master Regulators of Gene Expression to Biomarkers Involved in Intercellular Communication*. Biomedicines, 2024. **12**(4).

References

36. Zhang, J., et al., *miR-let-7a suppresses α -Synuclein-induced microglia inflammation through targeting STAT3 in Parkinson's disease*. Biochemical and Biophysical Research Communications, 2019. **519**(4): p. 740-746.
37. Liu, J., et al., *Let-7 miRNAs Modulate the Activation of NF- κ B by Targeting TNFAIP3 and Are Involved in the Pathogenesis of Lupus Nephritis*. PLOS ONE, 2015. **10**(6): p. e0121256.
38. Li, Z.H., et al., *Let-7f-5p suppresses Th17 differentiation via targeting STAT3 in multiple sclerosis*. (1945-4589 (Electronic)).
39. França, G.S., et al., *Cellular adaptation to cancer therapy along a resistance continuum*. Nature, 2024. **631**(8022): p. 876-883.
40. Li, Y., et al., *Gain-of-Function Mutations: An Emerging Advantage for Cancer Biology*. Trends Biochem Sci, 2019. **44**(8): p. 659-674.
41. Otmani, K. and P. Lewalle, *Tumor Suppressor miRNA in Cancer Cells and the Tumor Microenvironment: Mechanism of Dereglulation and Clinical Implications*. 2021, **11**.
42. Cimmino, A., et al., *miR-15 and miR-16 induce apoptosis by targeting BCL2*. Proc Natl Acad Sci U S A, 2005. **102**(39): p. 13944-9.
43. Messina, S., *The RAS oncogene in brain tumors and the involvement of let-7 microRNA*. Mol Biol Rep, 2024. **51**(1): p. 531.
44. Mott, J.L., et al., *mir-29 regulates Mcl-1 protein expression and apoptosis*. Oncogene, 2007. **26**(42): p. 6133-40.
45. Kilanowska, A.A.-O. and S. Studzińska, *In vivo and in vitro studies of antisense oligonucleotides - a review*. (2046-2069 (Electronic)).
46. Di Martino, M.T., et al., *miR-221/222 as biomarkers and targets for therapeutic intervention on cancer and other diseases: A systematic review*. Mol Ther Nucleic Acids, 2022. **27**: p. 1191-1224.
47. Hashemi, M., et al., *Pre-clinical and clinical importance of miR-21 in human cancers: Tumorigenesis, therapy response, delivery approaches and targeting agents*. Pharmacological Research, 2023. **187**: p. 106568.
48. Wang, S., et al., *MicroRNA-24 in Cancer: A Double Side Medal With Opposite Properties*. Front Oncol, 2020. **10**: p. 553714.
49. Frixia, T., S. Donzelli, and G. Blandino, *Oncogenic MicroRNAs: Key Players in Malignant Transformation*. Cancers (Basel), 2015. **7**(4): p. 2466-85.
50. Dellago, H., M.R. Bobbili, and J. Grillari, *MicroRNA-17-5p: At the Crossroads of Cancer and Aging - A Mini-Review*. Gerontology, 2017. **63**(1): p. 20-28.
51. Ardizzone, A., et al., *Role of miRNA-19a in Cancer Diagnosis and Poor Prognosis*. Int J Mol Sci, 2021. **22**(9).
52. Martino, M.T.D., P. Tagliaferri, and P. Tassone, *MicroRNA in cancer therapy: breakthroughs and challenges in early clinical applications*. Journal of Experimental & Clinical Cancer Research, 2025. **44**(1): p. 126.

53. Brinkmann, K., et al., *miR17~92 restrains pro-apoptotic BIM to ensure survival of haematopoietic stem and progenitor cells*. Cell Death Differ, 2020. **27**(5): p. 1475-1488.
54. Lu, Y., et al., *Transgenic over-expression of the microRNA miR-17-92 cluster promotes proliferation and inhibits differentiation of lung epithelial progenitor cells*. Dev Biol, 2007. **310**(2): p. 442-53.
55. Concepcion, C.P., C. Bonetti, and A. Ventura, *The microRNA-17-92 family of microRNA clusters in development and disease*. Cancer J, 2012. **18**(3): p. 262-7.
56. Costinean, S., et al., *Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in Eμ-miR155 transgenic mice*. 2006. **103**(18): p. 7024-7029.
57. Ju, J., *Challenges and opportunities in microRNA-based cancer therapeutics*. Cell Reports Medicine, 2025. **6**(4): p. 102057.
58. Abak, A., et al., *MicroRNA-221: biogenesis, function and signatures in human cancers*. Eur Rev Med Pharmacol Sci, 2018. **22**(10): p. 3094-3117.
59. Di Martino, M.T., et al., *miR-221/222 as biomarkers and targets for therapeutic intervention on cancer and other diseases: A systematic review*. Molecular Therapy Nucleic Acids, 2022. **27**: p. 1191-1224.
60. Huyan, T., et al., *miR-221-5p and miR-186-5p Are the Critical Bladder Cancer Derived Exosomal miRNAs in Natural Killer Cell Dysfunction*. Int J Mol Sci, 2022. **23**(23).
61. Tepebaşı, M.Y. and Ö. Öztürk, *miR-21, miR-221, and miR-222 upregulation in lung cancer promotes metastasis by reducing oxidative stress and apoptosis*. Rev Assoc Med Bras (1992), 2023. **69**(6): p. e20221688.
62. Braasch, D.A. and D.R. Corey, *Locked nucleic acid (LNA): fine-tuning the recognition of DNA and RNA*. Chemistry & Biology, 2001. **8**(1): p. 1-7.
63. Zamore, P.D., et al., *RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals*. (0092-8674 (Print)).
64. Pellestor, F. and P. Paulasova, *The peptide nucleic acids (PNAs), powerful tools for molecular genetics and cytogenetics*. (1018-4813 (Print)).
65. Singh, G. and V. Monga, *Peptide Nucleic Acids: Recent Developments in the Synthesis and Backbone Modifications*. Bioorganic Chemistry, 2023. **141**: p. 106860.
66. Sannigrahi, A., et al., *Peptide nucleic acids: Recent advancements and future opportunities in biomedical applications*. Bioorganic Chemistry, 2025. **155**: p. 108146.
67. De Costa, N.T. and J.M. Heemstra, *Evaluating the effect of ionic strength on duplex stability for PNA having negatively or positively charged side chains*. (1932-6203 (Electronic)).
68. Pfeffer, N.J., et al., *Strand-invasion of duplex DNA by peptide nucleic acid oligomers*. 1993. **90**(22): p. 10648-10652.
69. Wierzba, A.J., et al., *Unveiling the promise of peptide nucleic acids as functional linkers for an RNA imaging platform*. RSC Chem Biol, 2025. **6**(2): p. 249-262.
70. Wittung, P., et al., *DNA-like double helix formed by peptide nucleic acid*. Nature, 1994. **368**(6471): p. 561-563.

References

71. Collotta, D., et al., *Antisense oligonucleotides: a novel Frontier in pharmacological strategy*. (1663-9812 (Print)).
72. Troisi, R.A.-O., et al., *Steric hindrance and structural flexibility shape the functional properties of a guanine-rich oligonucleotide*. (1362-4962 (Electronic)).
73. Lima, W.F., V. Mohan, and S.T. Crooke, *The Influence of Antisense Oligonucleotide-induced RNA Structure on Escherichia coli RNase H1 Activity**. Journal of Biological Chemistry, 1997. **272**(29): p. 18191-18199.
74. Knudsen, H. and P.E. Nielsen, *Antisense Properties of Duplex- and Triplex-Forming PNAs*. Nucleic Acids Research, 1996. **24**(3): p. 494-500.
75. Seo, J., E.W. Ottesen, and R.N. Singh, *Antisense methods to modulate pre-mRNA splicing*. Methods Mol Biol, 2014. **1126**: p. 271-83.
76. Faria, M., et al., *Targeted inhibition of transcription elongation in cells mediated by triplex-forming oligonucleotides*. 2000. **97**(8): p. 3862-3867.
77. Hansen, M.E., T. Bentin, and P.E. Nielsen, *High-affinity triplex targeting of double stranded DNA using chemically modified peptide nucleic acid oligomers*. Nucleic Acids Research, 2009. **37**(13): p. 4498-4507.
78. Malik, S., et al., *Antitumor efficacy of a sequence-specific DNA-targeted γ PNA-based c-Myc inhibitor*. Cell Rep Med, 2024. **5**(1): p. 101354.
79. Kemme, C.A., et al., *Regulation of transcription factors via natural decoys in genomic DNA*. (2154-1272 (Electronic)).
80. Moggio, L., et al., *Alternate PNA-DNA chimeras (PNA-DNA)(n): synthesis, binding properties and biological activity*. (0006-3525 (Print)).
81. Quijano, E., et al., *Therapeutic Peptide Nucleic Acids: Principles, Limitations, and Opportunities*. Yale J Biol Med, 2017. **90**(4): p. 583-598.
82. Zhang, X., et al., *Polypeptide-protein conjugation: A new paradigm for therapeutic protein delivery*. Journal of Controlled Release, 2025. **384**: p. 113953.
83. Nsairat, H., et al., *Recent advances in using liposomes for delivery of nucleic acid-based therapeutics*. OpenNano, 2023. **11**: p. 100132.
84. Avitabile, C., et al., *Recent Cutting-Edge Technologies for the Delivery of Peptide Nucleic Acid*. Chemistry, 2025. **31**(34): p. e202500469.
85. Volpi, S., et al., *Multifunctional Delivery Systems for Peptide Nucleic Acids*. 2021. **14**(1): p. 14.
86. Du, J.-J., et al., *Applications of cell penetrating peptide-based drug delivery system in immunotherapy*. 2025. **Volume 16 - 2025**.
87. Gori, A., et al., *Cell Penetrating Peptides: Classification, Mechanisms, Methods of Study, and Applications*. ChemMedChem, 2023. **18**(17): p. e202300236.
88. Langel, Ü., *Cell-Penetrating Peptides and Transportan*. Pharmaceutics, 2021. **13**(7).
89. Nugrahadhi, P.P., et al., *Designing Formulation Strategies for Enhanced Stability of Therapeutic Peptides in Aqueous Solutions: A Review*. 2023. **15**(3): p. 935.

90. Cautain, B., et al., *Components and regulation of nuclear transport processes*. Febs j, 2015. **282**(3): p. 445-62.
91. Braun, K., et al., *A biological transporter for the delivery of peptide nucleic acids (PNAs) to the nuclear compartment of living cells*. J Mol Biol, 2002. **318**(2): p. 237-43.
92. Malinowska, A.A.-O., H.A.-O. Huynh, and S.A.-O. Bose, *Peptide-Oligonucleotide Conjugation: Chemistry and Therapeutic Applications*. (1467-3045 (Electronic)).
93. Lindgren, M., et al., *Cell-penetrating peptides*. Trends in Pharmacological Sciences, 2000. **21**(3): p. 99-103.
94. Frankel, A.D. and C.O. Pabo, *Cellular uptake of the tat protein from human immunodeficiency virus*. Cell, 1988. **55**(6): p. 1189-93.
95. Green, M. and P.M. Loewenstein, *Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein*. Cell, 1988. **55**(6): p. 1179-88.
96. Vivès, E., P. Brodin, and B. Lebleu, *A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus*. J Biol Chem, 1997. **272**(25): p. 16010-7.
97. Koren, E. and V.P. Torchilin, *Cell-penetrating peptides: breaking through to the other side*. Trends in Molecular Medicine, 2012. **18**(7): p. 385-393.
98. Xiao, W., et al., *Advance in peptide-based drug development: delivery platforms, therapeutics and vaccines*. Signal Transduction and Targeted Therapy, 2025. **10**(1): p. 74.
99. Derakhshankhah, H. and S. Jafari, *Cell penetrating peptides: A concise review with emphasis on biomedical applications*. Biomedicine & Pharmacotherapy, 2018. **108**: p. 1090-1096.
100. Bottens, R.A. and T. Yamada, *Cell-Penetrating Peptides (CPPs) as Therapeutic and Diagnostic Agents for Cancer*. LID - 10.3390/cancers14225546 [doi] LID - 5546. (2072-6694 (Print)).
101. Walrant, A., et al., *Membrane Crossing and Membranotropic Activity of Cell-Penetrating Peptides: Dangerous Liaisons?* Acc Chem Res, 2017. **50**(12): p. 2968-2975.
102. Xu, J., et al., *Cell-penetrating peptide: a means of breaking through the physiological barriers of different tissues and organs*. Journal of Controlled Release, 2019. **309**: p. 106-124.
103. Joliot, A., et al., *Antennapedia homeobox peptide regulates neural morphogenesis*. (0027-8424 (Print)).
104. Derossi, D., et al., *The third helix of the Antennapedia homeodomain translocates through biological membranes*. (0021-9258 (Print)).
105. Derossi, D., et al., *Cell internalization of the third helix of the Antennapedia homeodomain is receptor-independent*. J Biol Chem, 1996. **271**(30): p. 18188-93.
106. Czajlik, A., et al., *Investigation of penetratin peptides. Part 1. The environment dependent conformational properties of penetratin and two of its derivatives*. J Pept Sci, 2002. **8**(4): p. 151-71.
107. Sorkin, A. and M. von Zastrow, *Endocytosis and signalling: intertwining molecular networks*. Nature Reviews Molecular Cell Biology, 2009. **10**(9): p. 609-622.

References

108. Zahraoui, A., *Chapter Twenty-Four - MICAL-Like1 in Endosomal Signaling*, in *Methods in Enzymology*, P.M. Conn, Editor. 2014, Academic Press. p. 419-437.
 109. Depierre, M., C. Pompili, and F. Niedergang, *Phagocytosis at a glance*. Journal of Cell Science, 2025. **138**(12).
 110. Stillwell, W., *Chapter 19 - Membrane Transport*, in *An Introduction to Biological Membranes (Second Edition)*, W. Stillwell, Editor. 2016, Elsevier. p. 423-451.
 111. Mani, I. and V. Singh, *Chapter One - An overview of receptor endocytosis and signaling*, in *Progress in Molecular Biology and Translational Science*, I. Mani and V. Singh, Editors. 2023, Academic Press. p. 1-18.
 112. Schultz, S.G., *Membrane Transport, General Concepts*, in *Encyclopedia of Biological Chemistry (Second Edition)*, W.J. Lennarz and M.D. Lane, Editors. 2013, Academic Press: Waltham. p. 49-51.
 113. Smith, D., et al., *Passive Lipoidal Diffusion and Carrier-Mediated Cell Uptake Are Both Important Mechanisms of Membrane Permeation in Drug Disposition*. Molecular Pharmaceutics, 2014. **11**(6): p. 1727-1738.
 114. Yang, N.J. and M.J. Hinner, *Getting across the cell membrane: an overview for small molecules, peptides, and proteins*. Methods Mol Biol, 2015. **1266**: p. 29-53.
 115. Neverisky, D.L. and G.W. Abbott, *Ion channel-transporter interactions*. Crit Rev Biochem Mol Biol, 2015. **51**(4): p. 257-67.
 116. Stillwell, W., *Membrane Transport*.
 117. Chen, I. and F. Lui, *Physiology, Active Transport*, in *StatPearls*. 2025, StatPearls Publishing
- Copyright © 2025, StatPearls Publishing LLC.: Treasure Island (FL) ineligible companies. Disclosure: Forshing Lui declares no relevant financial relationships with ineligible companies.
118. Wang, H. and A. Shi, *Endosomes: An underappreciated player in cell signal transduction*. Fundamental Research, 2025.
 119. Patel, S.G., et al., *Cell-penetrating peptide sequence and modification dependent uptake and subcellular distribution of green florescent protein in different cell lines*. Scientific Reports, 2019. **9**(1): p. 6298.
 120. Ouyang, J., Y. Sheng, and W. Wang, *Recent Advances of Studies on Cell-Penetrating Peptides Based on Molecular Dynamics Simulations*. Cells, 2022. **11**(24).
 121. Kawamoto, S., et al., *Inverted micelle formation of cell-penetrating peptide studied by coarse-grained simulation: importance of attractive force between cell-penetrating peptides and lipid head group*. J Chem Phys, 2011. **134**(9): p. 095103.
 122. Pandidan, S. and A. Mechler, *Latest developments on the mechanism of action of membrane disrupting peptides*. Biophys Rep, 2021. **7**(3): p. 173-184.
 123. Ruseska, I. and A. Zimmer, *Internalization mechanisms of cell-penetrating peptides*. Beilstein Journal of Nanotechnology, 2020. **11**: p. 101-123.

124. Behzadipour, Y. and S. Hemmati, *Covalent conjugation and non-covalent complexation strategies for intracellular delivery of proteins using cell-penetrating peptides*. Biomedicine & Pharmacotherapy, 2024. **176**: p. 116910.
125. Porosk, L. and Ü. Langel, *Approaches for evaluation of novel CPP-based cargo delivery systems*. 2022. **Volume 13 - 2022**.
126. Temsamani, J. and P. Vidal, *The use of cell-penetrating peptides for drug delivery*. Drug Discov Today, 2004. **9**(23): p. 1012-9.
127. Saito, G., J.A. Swanson, and K.D. Lee, *Drug delivery strategy utilizing conjugation via reversible disulfide linkages: role and site of cellular reducing activities*. Adv Drug Deliv Rev, 2003. **55**(2): p. 199-215.
128. Vutti, S., et al., *Click Chemistry Mediated Functionalization of Vertical Nanowires for Biological Applications*. 2016. **22**(2): p. 496-500.
129. Genchi, G., *An overview on d-amino acids*. Amino Acids, 2017. **49**(9): p. 1521-1533.
130. Carson, L.M. and E.E. Watson, *Peptide Nucleic Acids: From Origami to Editing*. 2024. **89**(11): p. e202400305.
131. Merrifield, R.B., *Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide*. Journal of the American Chemical Society, 1963. **85**(14): p. 2149-2154.
132. Amblard, M., et al., *Methods and protocols of modern solid phase peptide synthesis*. Molecular Biotechnology, 2006. **33**(3): p. 239-254.
133. Jaradat, D.s.M.M., *Thirteen decades of peptide synthesis: key developments in solid phase peptide synthesis and amide bond formation utilized in peptide ligation*. Amino Acids, 2018. **50**(1): p. 39-68.
134. Vaino, A.R. and K.D. Janda, *Solid-phase organic synthesis: a critical understanding of the resin*. J Comb Chem, 2000. **2**(6): p. 579-96.
135. Kassem, T., et al., *To Rink or Not to Rink Amide Link, that is the Question to Address for More Economical and Environmentally Sound Solid-Phase Peptide Synthesis*. International Journal of Peptide Research and Therapeutics, 2009. **15**(3): p. 211-218.
136. Barlos, K., et al., *2-Chlorotrityl chloride resin. Studies on anchoring of Fmoc-amino acids and peptide cleavage*. Int J Pept Protein Res, 1991. **37**(6): p. 513-20.
137. Chang, C.D., et al., *Preparation and properties of Nalpha-9-fluorenylmethoxycarbonylamino acids bearing tert.-butyl side chain protection*. Int J Pept Protein Res, 1980. **15**(1): p. 59-66.
138. Fields, G.B. and R.L. Noble, *Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids*. Int J Pept Protein Res, 1990. **35**(3): p. 161-214.
139. Crisma, M., et al., *Peptide helices based on alpha-amino acids*. Biopolymers, 2006. **84**(1): p. 3-12.
140. Carpino, L.A., *1-Hydroxy-7-azabenzotriazole. An efficient peptide coupling additive*. Journal of the American Chemical Society, 1993. **115**(10): p. 4397-4398.
141. Bodanszky, M. and J.Z. Kwei, *SIDE REACTIONS IN PEPTIDE SYNTHESIS*. 1978. **12**(2): p. 69-74.

References

142. Manne, S.R., et al., *Minimizing side reactions during amide formation using DIC and oxymapure in solid-phase peptide synthesis*. Tetrahedron Letters, 2021. **85**: p. 153462.
143. Fields, G.B., *Introduction to peptide synthesis*. Curr Protoc Protein Sci, 2002. **Chapter 18**: p. 18.1.1-18.1.9.
144. King, D.S., C.G. Fields, and G.B. Fields, *A cleavage method which minimizes side reactions following Fmoc solid phase peptide synthesis*. Int J Pept Protein Res, 1990. **36**(3): p. 255-66.
145. Ralhan, K., V.G. KrishnaKumar, and S. Gupta, *Piperazine and DBU: a safer alternative for rapid and efficient Fmoc deprotection in solid phase peptide synthesis*. RSC Advances, 2015. **5**(126): p. 104417-104425.
146. Mthembu, S.N., et al., *Solid-Phase Synthesis of C-Terminus Cysteine Peptide Acids*. Organic Process Research & Development, 2022. **26**(12): p. 3323-3335.
147. Tornøe, C.W., C. Christensen, and M. Meldal, *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): p. 3057-64.
148. Gaetke, L.M. and C.K. Chow, *Copper toxicity, oxidative stress, and antioxidant nutrients*. Toxicology, 2003. **189**(1-2): p. 147-63.
149. Gierlich, J., et al., *Click chemistry as a reliable method for the high-density postsynthetic functionalization of alkyne-modified DNA*. Org Lett, 2006. **8**(17): p. 3639-42.
150. *Organometallic Chemistry: Volume 43*, ed. N.J. Patmore and P.I.P. Elliott. 2020: The Royal Society of Chemistry.
151. Nmagu, D., S.K. Singh, and K.H. Lee, *Chapter Thirteen - Creation of monoclonal antibody expressing CHO cell lines grown with sodium butyrate and characterization of resulting antibody glycosylation*, in *Methods in Enzymology*, W.B. O'Dell and Z. Kelman, Editors. 2021, Academic Press. p. 267-295.
152. Cooper, J.R., et al., *Long Term Culture of the A549 Cancer Cell Line Promotes Multilamellar Body Formation and Differentiation towards an Alveolar Type II Pneumocyte Phenotype*. PLoS One, 2016. **11**(10): p. e0164438.
153. Sundararajan, R., et al., *10 - Irreversible electroporation: a drug-free cancer treatment*, in *Electroporation-Based Therapies for Cancer*, R. Sundararajan, Editor. 2014, Woodhead Publishing. p. 219-243.
154. Wang, N., et al., *Negative Regulation of PTEN by MicroRNA-221 and Its Association with Drug Resistance and Cellular Senescence in Lung Cancer Cells*. 2018. **2018**(1): p. 7908950.
155. Adan, A., Y. Kiraz, and Y. Baran, *Cell Proliferation and Cytotoxicity Assays*. Curr Pharm Biotechnol, 2016. **17**(14): p. 1213-1221.
156. Erbacher, P., et al., *Putative role of chloroquine in gene transfer into a human hepatoma cell line by DNA/lactosylated polylysine complexes*. Exp Cell Res, 1996. **225**(1): p. 186-94.
157. Zaitsev, S., et al., *Histone H1-mediated transfection: role of calcium in the cellular uptake and intracellular fate of H1-DNA complexes*. Acta Histochemica, 2002. **104**(1): p. 85-92.

158. Shiraishi, T., S. Pankratova, and P.E. Nielsen, *Calcium ions effectively enhance the effect of antisense peptide nucleic acids conjugated to cationic tat and oligoarginine peptides*. Chem Biol, 2005. **12**(8): p. 923-9.
159. Cordier, C., et al., *Delivery of Antisense Peptide Nucleic Acids to Cells by Conjugation with Small Arginine-Rich Cell-Penetrating Peptide (R/W)9*. PLOS ONE, 2014. **9**(8): p. e104999.
160. Schneider, A.F.L., et al., *Cellular uptake of large biomolecules enabled by cell-surface-reactive cell-penetrating peptide additives*. Nature Chemistry, 2021. **13**(6): p. 530-539.
161. Illien, F., et al., *Quantitative fluorescence spectroscopy and flow cytometry analyses of cell-penetrating peptides internalization pathways: optimization, pitfalls, comparison with mass spectrometry quantification*. Sci Rep, 2016. **6**: p. 36938.
162. Le Jeune, M., et al., *Conjugation of Oligo-His Peptides to Magnetic γ -Fe₂O₃@SiO₂ Core-Shell Nanoparticles Promotes Their Access to the Cytosol*. ACS Applied Materials & Interfaces, 2022. **14**(13): p. 15021-15034.
163. Shakeel, S., S. Karim, and A. Ali, *Peptide nucleic acid (PNA) — a review*. 2006. **81**(6): p. 892-899.
164. Wienken, C.J., et al., *Thermophoretic melting curves quantify the conformation and stability of RNA and DNA*. Nucleic Acids Res, 2011. **39**(8): p. e52.

CHAPTER II

Structural and Functional Analysis of Vimentin-Derivative Peptide in DNA Binding



1. Introduction

This chapter investigates the interaction between Vimentin derived peptides and G-quadruplexes DNA structures through a combination of peptide synthesis, conformational and biophysical studies. Peptides corresponding to different regions of Vimentin were synthesised using Fmoc-based solid-phase peptide synthesis (SPPS), with specific optimisation steps.

Electrophoretic mobility shift assays, fluorescence melting experiments and surface plasmon resonance analyses revealed that one of the synthesised fragments binds to the G-quadruplex.

Interaction studies were conducted by Prof. Sissi research group at the Department of Pharmaceutical and Pharmacological Sciences at the University of Padova.

1.1. Structural Proteins and Cytoskeleton

The cytoskeleton represents one of the most sophisticated molecular architectures found in eukaryotic cells. Rather than being a static structure, it is a dynamic, plastic and multifunctional system capable of integrating mechanical, biochemical and topological signals.[1] It defines cellular morphology, directs intracellular trafficking, coordinates cell division and mediates interactions with the surrounding microenvironment. In the context of cancer biology, the cytoskeleton plays an even more critical role, as alterations to its organisation and composition directly contribute to neoplastic transformation, invasion and metastasis.[2]

The cytoskeleton is structurally composed of three main filament systems (Figure 2.1): actin microfilaments, microtubules, and intermediate filaments. These systems are interconnected and cooperate to provide the cell with the mechanical and functional properties required for survival and adaptation to changing conditions. The distinction between these filament types is not only based on size or protein composition, but also reflects fundamental differences in assembly dynamics, polarity, and function.[3]

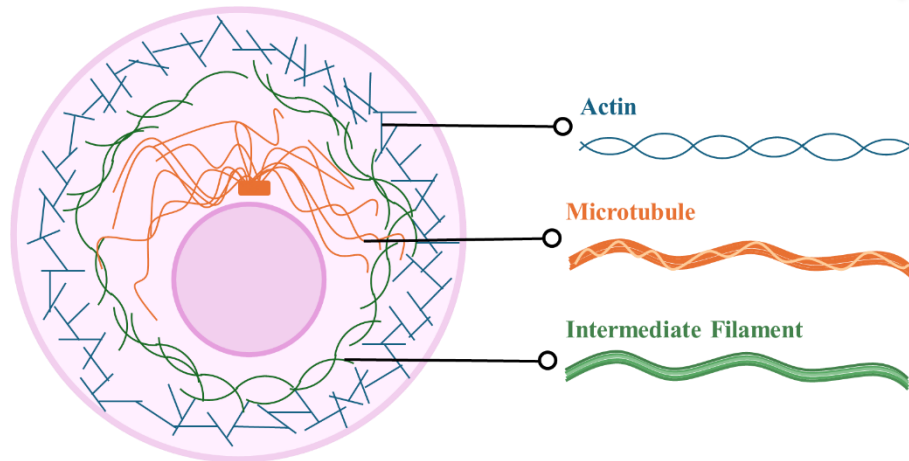


Figure 2.1. Schematic representation of cytoskeleton in cells.

- Actin microfilaments are composed of F-actin polymers formed by the polymerisation of G-actin monomers. They are particularly dense in the cortical region just beneath the plasma membrane, which allows for the rapid modulation of cell shape and the generation of forces necessary for migration and phagocytosis. This polymerisation process is regulated by protein complexes such as Arp2/3 and formins, which are in turn controlled by Rho-family small GTPases. In tumours, the dysregulation of this pathway can promote a highly motile and invasive phenotype.[4]
- Microtubules, which are formed by α - and β -tubulin heterodimers, constitute a radial network that emanates from the centrosome. Their intrinsic polarity, with dynamic '+' and '-' ends anchored at the centrosome, is crucial for directing intracellular traffic and ensuring proper chromosome segregation during mitosis. Microtubules are also a key target of chemotherapeutic drugs; destabilisers such as colchicine and vincristine, and stabilisers like taxanes, exploit microtubule dynamics to inhibit cell division. In cancer, alterations in microtubule stability and orientation can promote aneuploidy and metabolic adaption.[5]
- Intermediate filaments (IFs) are characterised by their stability and lack of intrinsic polarity. With a diameter of approximately 10-12 nm, they provide mechanical resistance and structural stability. Their composition is tissue-specific: keratins in epithelial cells, desmin in muscle cells, glial fibrillary acidic protein (GFAP) in glial cells and Vimentin in mesenchymal cells. Intermediate filaments not only mechanically reinforce cells, but also actively interact with the plasma membrane, actin filaments and microtubule, thereby participating in intracellular signalling and morphological adaptation.[6]

The concept of mechanotransduction is particularly relevant when it comes to understanding the impact of the cytoskeleton in tumour biology. Cells send and respond to the physical properties of their microenvironment, such as matrix stiffness, shear stress, compression or tension, by translating

Introduction



mechanical stimuli into biochemical signals that modulate cellular behaviour. Tumour microenvironments often display increased rigidity due to abnormal collagen and fibronectin deposition and altered matrix crosslinking. This prompts cytoskeletal reorganization that supports invasive phenotypes and activities transcriptional programmes associated with epithelial-mesenchymal transition (EMT).[7]

EMT is an example of how the cytoskeleton is extensively remodelled during tumour progression. Epithelial cells lose their apico-basal polarity and adherens junctions, instead acquiring an elongated, motile mesenchymal morphology.[7] At the cytoskeletal level, down regulation of keratin coincides with upregulation of Vimentin, which is the primary intermediate filament protein in these cells. This replacement is not merely a marker of cellular state, but fundamentally alters the cell mechanical and dynamic properties, enhancing deformability and enabling migration through confined spaces.[8]

In oncology, cytoskeleton acts as a regulatory hub integrating growth, adhesion, stress and differentiation signals. The interconnection among the three filamentous components coordinates force generation, migration directionality and resistance to mechanical stress. Vimentin acts as a mechanical and functional ‘hinge’ between the networks of actin and microtubules, stabilising protrusions such as lamellipodia and filopodia, and anchoring key organelles such as the nucleus and mitochondria.[9]

1.2. Intermediate Filaments and Vimentin

Intermediate filaments (IFs) provide structural integrity and elasticity, as well as the ability to absorb mechanical stress without breaking. Unlike actin filaments or microtubules, IFs are non-polar and do not serve as tracks for motor proteins or directly generate motile force. These properties make them particularly important in cells that undergo substantial mechanical deformation, such as fibroblasts, endothelial cells, and muscle cells.[6]

Biochemically, IFs are composed of modular proteins with a highly conserved α -helical rod domain, which is flanked by less structured N- and C-terminal head and tail domains. Filament assembly is hierarchical: parallel dimers form antiparallel tetramers, which then assemble into lateral protofilaments with a diameter of around 10-12nm. The antiparallel arrangement explains the lack of polarity, while the modular domains allow flexibility and regulated assembly and disassembly in response to cellular signals.[10]

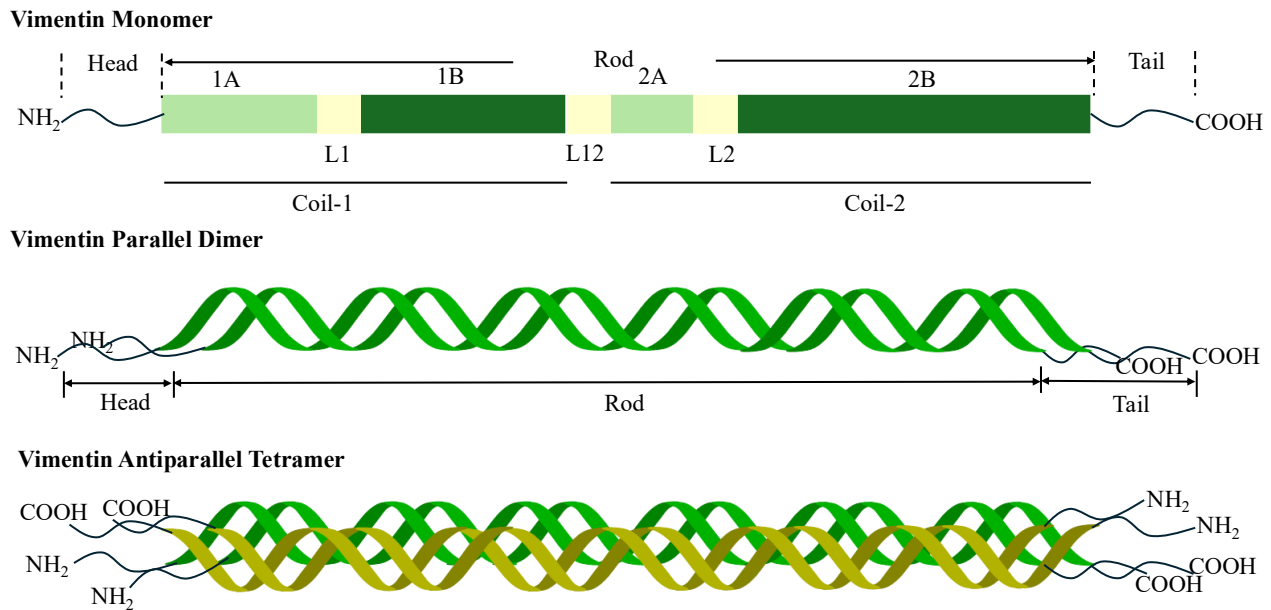


Figure 2.2. Schematic dynamic structure of Vimentin.

Vimentin, a type III filament, is the principal IF in mesenchymal cells including fibroblasts, endothelial cells, monocytes/macrophages, and various stem cell populations. Under physiological conditions, its expression is tightly regulated in accordance with the differentiation state and mechanical demands. In many epithelial cancers, however, Vimentin is aberrantly upregulated, marking cells undergoing EMT and correlating with enhanced invasiveness.[11]

From a structural standpoint (Figure 2.2), the monomeric unit of Vimentin consists of two intrinsically disordered domains (N- and C-terminal) flanking a central α -helical rod domain. The rod region is organised into two main coiled-coil segments (coil-1 and coil-2), which are further divided into subdomains A and B. Two flexible, non-helical linkers (L1A and L12) connect coil-1A to coil-1B and coil-1 to coil-2, respectively.[12]

In its soluble form, Vimentin exists as a tetramer. Assembly begins with the parallel alignment of two monomeric units to form a coiled-coil dimer. Subsequently, two dimers associate laterally in distinct orientations to generate tetrameric species. Cross-linking and structural studies have identified four principal dimer-dimer association mode [13, 14]:

- In the A11 arrangement, two dimers associate antiparallely, with the coil-1B regions aligned in register.
- In the A22 configuration, however, it is the coil-2 segments that are aligned.
- The A12 mode involves complete antiparallel overlap of the two dimers.
- The ACN configuration is the only parallel association in which the C- and N-terminal domains of the opposing dimers interact.

Introduction



The lateral association of these tetramers gives rise to unit-length filaments (ULFs), which then undergo longitudinal annealing to form mature intermediate filaments. Contrary to the traditional view of Vimentin as a static cytoplasmatic network, it is highly dynamic. Its filaments undergo constant cycles of polymerization and depolymerization, which are controlled by post-translational modifications such as the phosphorylation of N-terminal residues by kinases including Cdk1, PKA, PKC and Rho kinase. This dynamic turnover enables cells to rapidly reorganise their cytoskeleton in response to mitotic cues, migratory stimuli or extracellular mechanical forces. Vimentin filaments also act as mechanical bridges that link the actin and microtubule networks. They stabilise protrusions such as lamellipodia and filopodia, anchor organelles including the nucleus and mitochondria, and maintain cellular integrity under stress.[15]

Furthermore, extracellular stimuli such as growth factors chemokines, cytokines or mechanical cues can modulate Vimentin solubility, subcellular localisation and protein-binding capacity, thereby facilitating transitions between stationary and motile states.

Thus, Vimentin is not merely a marker of mesenchymal identity; it is also an active regulator of cell mechanics, organelle positioning and intracellular signalling, making it a key player in tumour progression.[16]

1.3. Vimentin in Cell Migration and EMT

Cell migration is a highly regulated process that is essential for embryogenesis, wound healing and immune surveillance. However, tumour cells exploit this process to invade surrounding tissues and disseminate to distant sites. Vimentin is a key mediator of this process and a canonical marker of EMT.[17]

EMT is a reversible differentiation programme that enables epithelial cells, which are normally polarised and adherent, to acquire a mesenchymal phenotype characterised by morphological plasticity, loss of polarity, and enhanced motility. EMT involves the downregulation of epithelial markers such as E-cadherin and the upregulation of mesenchymal markers including Vimentin, N-cadherin, and fibronectin.[18]

Vimentin actively remodels the cytoskeleton to facilitate directional migration. In cells undergoing EMT, Vimentin filaments reorganise from a compact perinuclear arrangement into a more extended, polarised network that is aligned with the migration axis. This structural reorganisation supports the formation of lamellipodia and filopodia, stabilises protrusions against mechanical stress and protect the nucleus when passing through confined extracellular matrix (ECM) spaces.[19]

The role of Vimentin in migration encompasses several interrelated functions:



1. **Mechanical resilience and deformability.** Vimentin filaments act as a cushion that absorbs mechanical stress, thereby protecting the nucleus and DNA from damage during migration. Cells lacking Vimentin exhibit increased nuclear fragility and impaired transmigration through tight pores.[20]
2. **Intracellular signalling.** Vimentin interacts with proteins that organise actin and components of focal adhesions, including FAK, paxillin, and integrins. Phosphorylation by kinases such as PAK1 or ROCK modulates adhesion assembly and disassembly, thereby influencing migration speed and persistence.[20]
3. **Organelle positioning and metabolic support.** Vimentin organises the distribution of mitochondria and the Golgi apparatus to sustain the supply of energy and the targeted secretion of proteases such as MMPs (matrix metalloproteinases), thereby facilitating ECM degradation and invasion.[20]

In tumours, high Vimentin levels are associated with aggressive phenotypes and poor prognosis in various cancers, including triple-negative breast cancer, non-small cell lung carcinoma, colorectal carcinoma. Vimentin also supports partial EMT states, in which cells co-express epithelial and mesenchymal markers while maintaining proliferative capacity and gaining motility and chemoresistance. The presence of surface-expressed Vimentin on migrating tumour cells provides additional therapeutic opportunities, enabling selective targeting antibodies or nanoparticle-based systems.[21]

1.4. Vimentin in Metastasis and Tumour Progression

Metastasis is the primary cause of cancer-related death, and Vimentin plays a central role in every stage of the metastatic cascade, including local invasion, intravasation, survival in circulation, extravasation and colonisation of secondary sites.

- **Local invasion.** Vimentin facilitates the formation of invasive protrusions and the targeted secretion of MMPs to degrade the ECM. Knockdown experiments confirm its role in 3D motility.[22]
- **Intravasation.** High Vimentin expression enhances endothelial transmigration by regulating integrins and adhesion molecules, such as ICAM-1 and VCAM-1.[23]
- **Circulating tumour cells (CTCs).** Vimentin networks protect cells against shear stress, anoikis and immune attacks, and may facilitate the formation of CTC clusters, thereby increasing survival and metastatic efficiency.[24]

Introduction



- **Metastatic colonisation.** Persistent Vimentin expression modulates interactions with fibroblasts, immune cells and the endothelium, thereby promoting angiogenesis (via VEGF and IL-8) and the establishment of a permissive microenvironment.[25]

Vimentin may also play a role in metastatic dormancy by supporting survival under conditions of low proliferation and high stress, which could contribute to late relapse. Clinically, elevated Vimentin levels predict poor outcomes, therapy resistance and an increased metastatic burden. The detection of circulating Vimentin-positive cells or soluble Vimentin offers non-invasive prognostic and diagnostic opportunities.[26]

1.5. Interplay between Vimentin and G-Quadruplexes Structures

Recent evidence suggests the existence of potential functional links between Vimentin and G-quadruplex (G4) nucleic acid structures, a non-canonical DNA structure.

Although Vimentin is predominantly localised in the cytoplasm, it has also been detected in the nucleus and nuclear matrix, particularly in poorly differentiated metastatic cancer cell.[27] At the nuclear level, the soluble tetrameric form of Vimentin has been found to interact with G-rich DNA sequences that can potentially fold into G4 structures.[27] In certain genomic regions, the high guanine content enables the formation of multiple adjacent G4 units that can interact to generate higher-order structural assemblies known as G4 repeats (G4rp). These complex, three-dimensional elements appear to act as active platforms for the selective recruitment of proteins involved in gene expression and cell cycle regulation.[28]

Recent studies have demonstrated that tetrameric Vimentin binds to G4rp with high affinity (with dissociation constants in the nanomolar range) and remarkable selectivity; it does not interact with single G4s or other DNA structural arrangements.[29] This binding is independent of the topology or the number of G4 units within the G4rp, as demonstrated using both telomeric and oncogenic promoter sequences. Bioinformatic analyses have also revealed that G4rp elements are enriched in the promoter regions of genes associated with EMT. This suggests that the Vimentin-G4rp complex may constitute a novel regulatory hub involved in transcriptional control during tumour invasion and metastasis.[29, 30]

Overall, these findings suggest a direct or indirect mechanism through which Vimentin, traditionally considered a cytoskeletal protein, could affect chromatin organisation and influence the formation or stabilisation of G4 structures. This interaction could serve as an integrative link between mechanical signalling and genomic regulation, revealing a new mechano-epigenetic axis with significant implications for tumour progression and as a promising target for anti-cancer intervention.[29]



1.6. G-Quadruplexes and Their Biological Role

Since 1953, the most familiar form of DNA has been the double helix, known as the B-form.[31] However, it is now recognised that DNA can adopt a variety of alternative conformations influenced by specific sequences, the cellular environment and interactions with different proteins. These ‘non-canonical’ secondary structures, including Z-DNA, triplexes, cruciform structures and G4s, play a crucial role in biological processes such as duplication, transcription, replication and cellular apoptosis.[32] The abundance of genomic data on DNA-protein and DNA-ligand interactions underlying both physiological and pathological cellular mechanisms has driven research into identifying novel molecular targets that can selectively interact with their biological partners. In this context, G4 – secondary structures formed by guanine-rich sequences – have attracted significant attention due to their potential presence in regulatory regions and their ability to selectively bind various types of ligands.[33] Computational studies across different organisms have revealed that sequences capable of forming quadruplexes (G4 motifs) are not randomly distributed across the genome but are closely associated with functional domains.[34] These sequences are frequently found in oncogene promoters, telomeric ends and the 5’ regions of certain mRNA molecules, and they often exhibit conserved base compositions across species. The non-random distribution and evolutionary conservation of their genomic positions support the hypothesis that these structures have regulatory functions within the cell.[35] Interest in quadruplexes has grown exponentially, with numerous studies clarifying the relationships between sequence, localisation, and biological function. Key functions of G4s include:

1. Regulation of transcription in numerous oncogene promoters;[36]
2. Control of post-transcriptional processes when present in mRNAs;[37]
3. Participation in telomere architecture in many organisms;[38]
4. Serving as molecular scaffolds for the design and synthesis of aptamers.[33]

Transcriptional regulation by G-quadruplexes depends on the guanine-rich sequences located in the region upstream of the transcription start sites of genes that are involved in cell growth and proliferation. Potential quadruplex-forming sequences have been identified in the promoters of human oncogenes such as c-MYC,[39] c-KIT,[40] BCL-2,[41] and K-RAS,[42] . The expression of these oncogenes is associated with numerous cancers. The equilibrium between quadruplex and single-stranded DNA determines transcriptional regulation: the formation of a quadruplex in these promoters silences the oncogene as steric hindrance prevents access by the transcription machinery.[43]

G4s can also regulate post-transcriptional processes at the RNA level. Some RNA quadruplexes have been found in the genomes of the human immunodeficiency virus (HIV) and herpes simplex virus,

Introduction



where they modulate dimerization and translation.[44] Others are present in the 5' UTR of oncogene mRNA, where they block ribosomal scanning and reduce protein synthesis.[45]

Finally, G-quadruplexes contribute to telomere structure: the focus of this aspect is explored in the following section.

1.6.1. G-Quadruplexes in Telomeres

The most pharmacologically interesting biological role of quadruplexes relates to telomere function. Telomeres are nucleoprotein complexes located at the ends of linear eukaryotic chromosomes. They protect DNA from 'erosion' during cell division, prevent chromosome ends from fusing during replication and safeguard nucleic acids from nuclease attacks.[38]

Human telomeric DNA consists of highly repeated sequences (TTAGGG) (Figure 2.3). In eukaryotic cells, most of the telomeric region consists of double-stranded DNA (approximately 5-15 kilobases), while the terminal ~200 nucleotides form a single-stranded 3' overhang containing multiple TTAGGG repeats.[46] In vertebrates, this DNA associates with *sic* proteins (POT1, TRF1, TRF2, TIN2, TPP1, and Rap1), which together form the Shelterin complex. This complex regulates the conformational equilibrium of the telomeric portion and indirectly controls telomerase activity.[47]

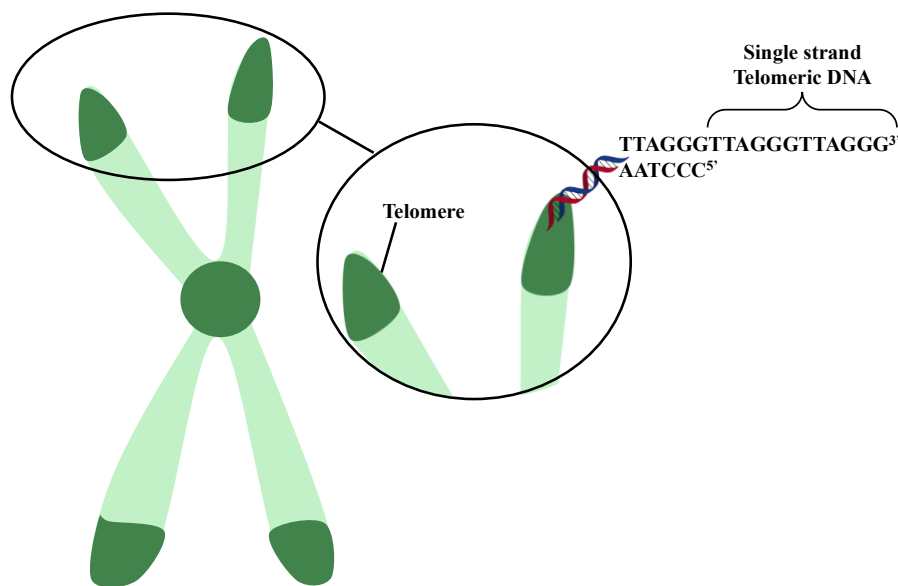


Figure 2.3. The human telomere is located at the end of the chromosome. It consists of a double-stranded DNA segment and a short single-stranded segment at the 3' end.

Telomeres are crucial for maintaining DNA length during the cell cycle, a process that is closely linked to senescence and apoptosis. In somatic cells, telomeres shorten with each replication due to DNA polymerase's inability to fully replicate the lagging strand. This shortening continues until a critical length is reached, triggering cellular senescence and apoptosis.[48]



In contrast, cancer cells overexpress the enzyme telomerase. This reverse transcriptase synthesises TTAGGG sequences at the 3' end of the free DNA strand. This shortening continues until a critical length is reached, triggering cellular senescence and apoptosis.[49]

In contrast, cancer cells overexpress the enzyme telomerase. This reverse transcriptase synthesises TTAGGG sequences at 3' end of the free DNA strand, thereby preventing normal shortening and granting 'cellular immortality' to tumour cells. Therefore, inhibiting telomerase could block this immortality mechanism, representing a promising anticancer strategy.[50] One approach is to stabilise telomeric DNA in a conformation that is inaccessible to the enzyme. Telomerase requires a linear, single-stranded 3' overhang to function. This region contains guanine-rich sequences that can fold into quadruplexes, creating a steric hindrance for the enzyme. Small molecules that stabilise these quadruplexes prevent telomerase access, thereby inhibiting its activity.[51]

Therefore, the interaction and stabilisation of telomeric G4s by potential drugs is a crucial aspect of new anti-cancer strategies. Rational drug design requires a thorough understanding of the binding properties of telomeric sequences that can form quadruplex structures.

1.6.2. Structural Features of G-Quadruplexes

The fundamental unit of G4s is the G-tetrad (also known as a G-quartet), which is a planar cyclic structure composed of four guanines connected via Hoogsteen hydrogen bonds (Figure 2.4). These differ from the more well-known Watson-Crick bonds. These four bases form the tetrad plane, and the carbonyl oxygens (O6) of each guanine create a central cavity that can accommodate monovalent cations, such as sodium and potassium, which contribute to the stability of the structure.[33]

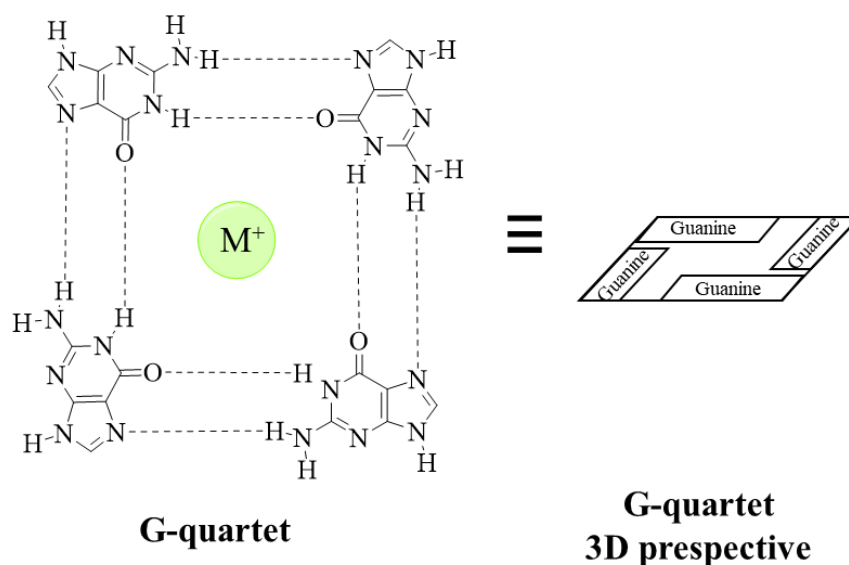


Figure 2.4. Schematic representation of a G-quartet rearrangement.

Introduction



The guanine planes expose a large, electron rich surface area due to aromatics rings, enabling stacking interactions between tetrads – a feature common to all quadruplexes.[33]

Despite this common motif, G-quadruplexes exhibit high polymorphism resulting from variations in the number of tetrads, strand polarity, loop length, and loop position. Quadruplexes can generally be classified according to:[52]

1. Strand stoichiometry
2. Sugar-phosphate backbone orientation
3. Sugar conformation
4. Loop orientation.

Based on the number of strands, they are distinguished in (Figure 2.5):[53]

1. Monomolecular: a single strand folding onto itself
2. Bimolecular: two strands with intra- or intermolecular arrangements
3. Tetramolecular: association of four strands.

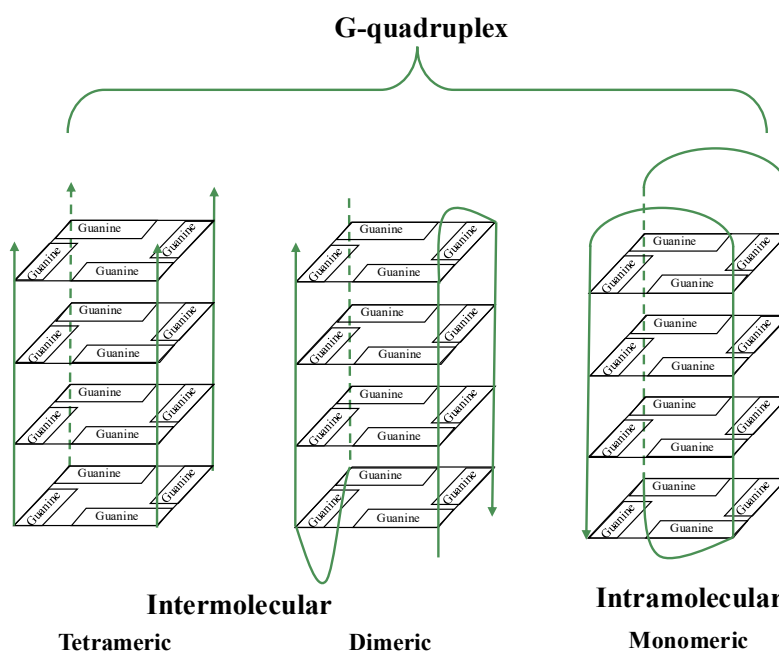


Figure 2.5. Representation of a tetrameric, dimeric and monomeric G-quadruplex.

The sugar-phosphate backbone can adopt parallel, antiparallel or '3+1' orientations (three strands in one direction and one in the opposite direction). These orientations are linked to the *syn/anti* guanine configuration within the tetrad (Figure 2.6).[54]

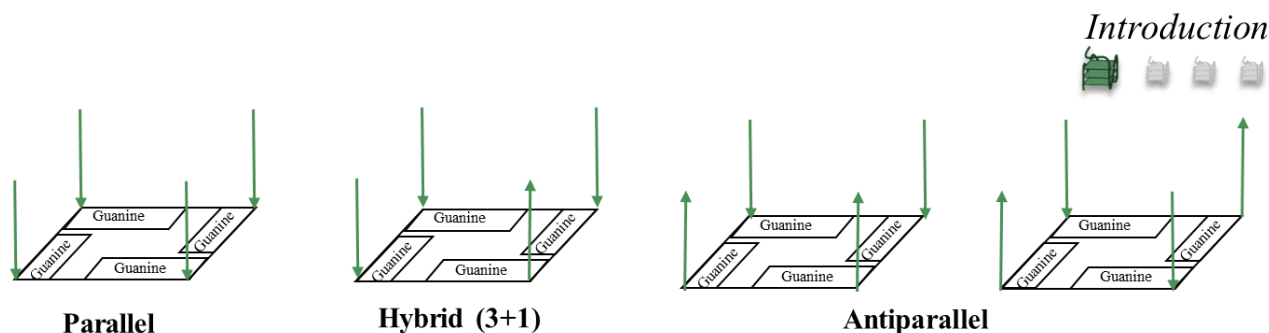


Figure 2.6. Representation of the parallel, hybrid and antiparallel conformation.

Loops, which connect bases within the same tetrad or between different tetrads, are essential for specific ligand recognition. They are classified as follows (Figure 2.7):[55]

1. Edgewise or lateral loops: connecting adjacent antiparallel strands
2. Diagonal loops: connecting opposite antiparallel strands
3. Propeller loops: which connect adjacent parallel strands

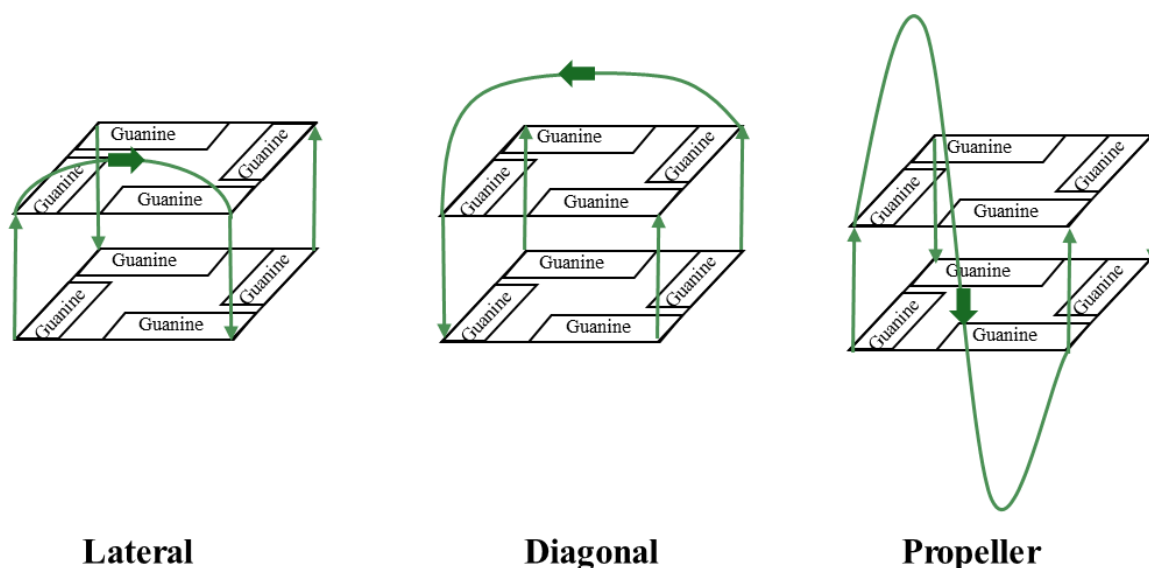


Figure 2.7. Representation of lateral, diagonal and propeller conformation.

Structural stability and conformation variability also depend on cations occupying the central cavity. Quadruplexes are more stable with Na^+ and K^+ , which are abundant in cellular environments, and are destabilised by Cs^+ and Li^+ , with the latter inhibiting quadruplex formation. Not only do sodium and potassium stabilise the structure, but they also influence the formation kinetics and the adopted conformation, which is particularly important for human telomeric sequences. Therefore, precise knowledge of telomeric structure is essential for investigating interactions with potential drugs.[56]

1.6.3. Structural Polymorphism of Human Telomeric G-Quadruplexes

The conformations of G4s can vary depending on the ions present in solution and the experimental conditions. High structural polymorphism has been observed in human telomeric sequences. The first type of quadruplex was characterised in solution in the presence of K^+ by analysing the sequence

Introduction



d[AGGG(TTAGGG)₃] (22 mer) using X-ray crystallography. This crystal structure exhibits a parallel conformation, in which the oligonucleotide strand folds in such a way that all segments of the sugar-phosphate backbone are oriented identically. The three G-tetrad planes, with all guanines in the anti-conformation, are connected by three symmetric propeller TTA loops. [57]

The same sequence, when studied in solution using NMR, revealed two conformations in equilibrium containing the three guanine planes. These are referred to as hybrid or 3+1 structures. These conformations, named hybrid 1 and hybrid 2, differ in the arrangement of the loops and the opposite strand.[58-60]

In 3+1 conformations, three strands are oriented in the same direction while the fourth is oriented in the opposite direction. In hybrid 1, the opposite strand is the third strand from the 5' end; in hybrid 2, it is the second strand. Additionally, the order of the loops from the 5' end is propeller-lateral-lateral in hybrid 1 and lateral-lateral-propeller in hybrid 2. The two forms coexist in solution, and their equilibrium is influenced by the presence of flanking bases at the 3' end, typically thymine. For example, the presence of thymine at this position appears to favour hybrid 2.[59, 60]

Successively, a third form has been observed under the same saline conditions: an antiparallel 'basket' conformation which differs in the number of guanine planes. The four strands have opposite polarity in pairs, and their folding resembles a basketball hoop, hence the name.[61]

This structural polymorphism has sparked debates about which conformation is biologically relevant. It should be noted that the reported structural studies were performed on telomeric sequences containing no more than four guanine tracts – sufficient to form a single quadruplex. However, the 3' end of the telomeric strand consists of around 200 nucleotides that can form consecutive quadruplexes and influence each other or the binding properties arising from these interactions.[62] Recent studies have shown that, when multiple G4 units occur in proximity, they can interact to form higher-order assemblies known as G4rp. In these structures, adjacent G4 modules are not independent, but rather exhibit inter-quadruplex crosstalk, resulting in unique three-dimensional arrangements with distinct biophysical properties. G4 repeats have been proposed to act as recognition platforms for specific proteins involved in chromatin organisation and gene regulation. At telomeres, G4rp formations could contribute to the structural compartmentalisation of the single-stranded overhang, thereby influencing telomere protection, accessibility and protein recruitment.[63]

The presence of non-isolated quadruplexes and their mutual influence are critical factors in designing selective telomeric ligands.



1.7. Aim of the chapter

The main objective of this chapter was to investigate the molecular determinants responsible for the interaction between Vimentin and G-quadruplex DNA structures, and to identify the protein fragment that could modulate or prevent this interaction.

To this end three peptide fragments derived from Vimentin were designed and synthesised, both with the free and acetylated N-terminus:

- 11-30 and 68-79, which belong to the N-terminal head domain
- 270-280, which is located within the central coiled-coil rod domain

These fragments were selected based on their structural and functional relevance to Vimentin assembly, as well as their potential involvement in nucleic acid recognition. The peptides were synthesised using optimised SPPS protocols and characterised through spectroscopic analyses. They were then evaluated for their conformational properties and binding behaviour towards G4-forming sequences.

The overall aim was to elucidate which Vimentin-derived regions contribute most to G4 binding and to assess their binding selectivity against duplex and single stranded DNA. The aim was also to provide a structural and biochemical framework for the rational design of peptide-based modulators capable of interfering with G4s.



2. Results and Discussion

In the previous chapter, particular attention was devoted to optimising solid-phase peptide synthesis (SPPS) for preparing peptides and peptide nucleic acid conjugates (Figure 2.8). The methodological refinements developed – such as selecting suitable resins, coupling reagents and optimising deprotection were essential for achieving high coupling efficiency minimising racemisation and improving product purity.

These optimised protocols provide a robust and versatile framework that can be readily adapted to synthesise a wide range of peptide sequences, including more complex or aggregation-prone systems. A similar synthetic strategy and methodological considerations were adopted as the starting point for the present work, which applies SPPS to prepare new peptide fragments derived from the Vimentin protein.

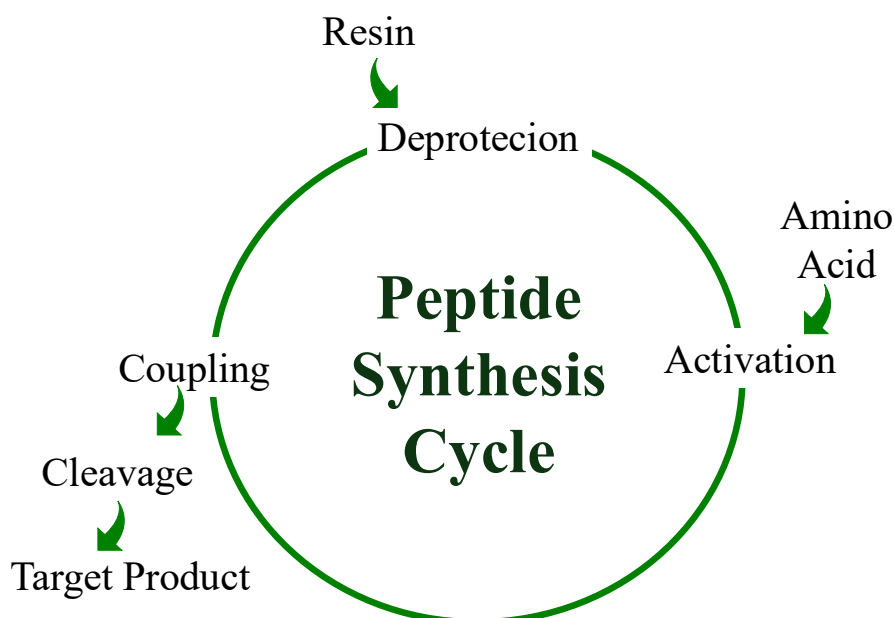


Figure 2.8. Schematic representation of the SPPS cycle.

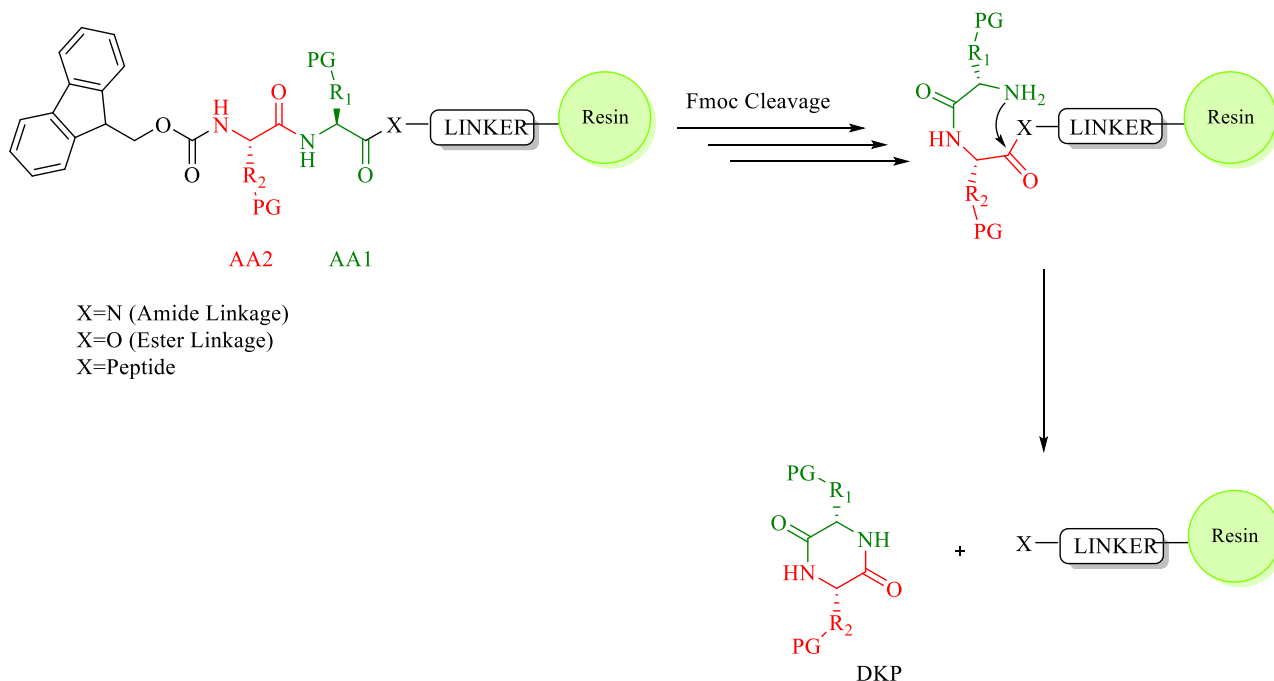
2.1. Diketopiperazine Formation and Strategies for Its Prevention

Of the various side reaction that occur during SPPS, the formation of diketopiperazines (DKPs) is one of the most detrimental (Scheme 2.1). These cyclic dipeptides originate from an intramolecular cyclisation between two amino acids attached to the resin. This leads to the premature cleavage of the dipeptide fragment and corresponding decrease in the yield of the desired peptide. [64, 65]

The classical mechanism of DKP formation begins with removal of the Fmoc protecting group from the second amino acid residue (AA2), generating a free nucleophilic α -amino group. This newly



deprotected amine can then attack the ester linkage intramolecularly, either connecting the first amino acid (AA1) to the linker or to the growing peptide chain. The resulting aminolysis reaction produces a stable six-membered 2,5-diketopiperazine ring, which simultaneously cleaves the dipeptide fragment from the resin and terminates further chain elongation.



Scheme 2.1. Diketopiperazine Formation.

Certain structural features and reaction conditions can make DKP formation particularly likely. The presence of glycine or proline as the second residue favours cyclisation due to their small or conformationally constrained side chains, which facilitate the intramolecular attack. Similarly, resins that provide high conformational flexibility or extended basic treatment during Fmoc deprotection can accelerate the process.

Several preventive measures can be considered and implemented where appropriate:

- Rapid and low-temperature Fmoc deprotection minimises exposure to basic conditions, which helps to suppress the nucleophilic attack responsible for DKP formation.
- Performing rapid washing after Fmoc removal reduces the time window in which cyclisation can occur
- In aggregation-prone or flexible sequences, temporarily introducing pseudoproline motifs effectively blocks the backbone conformation required for cyclisation.
- Maintaining good resin solvation and avoiding excessive mechanical agitation (e.g. vigorous vortexing or nitrogen bubbling) was also found to reduce unwanted side reactions

By carefully optimising the initial coupling conditions and adopting appropriate preventive strategies, the risk of DKP formation can be minimised effectively.



2.1.1. Synthesis of H-11-30-NH₂ Peptide

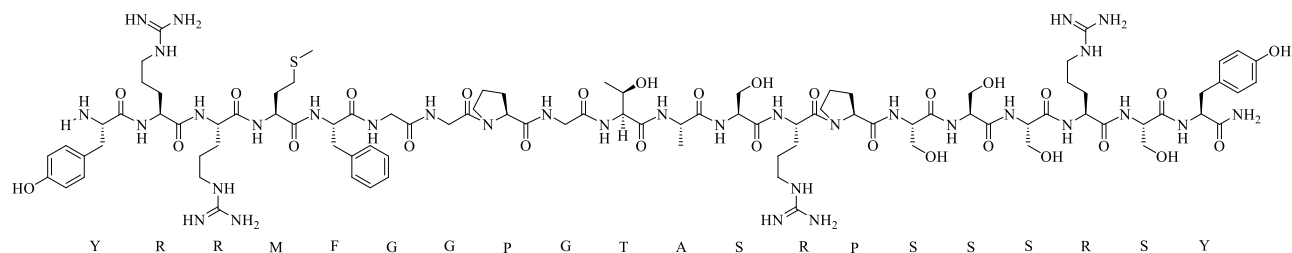


Figure 2.9. Structure of H-11-30-NH₂ peptide.

The H-11-30-NH₂ peptide (Figure 2.9), a fragment of Vimentin N-terminal region, was designed as a positive control for the internalization assays. It was synthesised manually using SPPS.

This Vimentin-derived fragment was particularly challenging to synthesise due to its sequence composition favouring the formation of DKP intermediates and truncated side products. The presence of residues such as Gly and Pro in the initial elongation steps can lead to intramolecular nucleophilic attack of the newly deprotected amine. This results in premature cleavage and loss of the growing peptide chain.

To overcome these issues, the synthesis was conducted manually to allow greater control over each coupling and deprotection step. This approach enabled the reaction times, agitation and washing steps to be finely adjusted, thereby minimising the accumulation of reactive intermediates responsible for side reactions. Additionally, double coupling cycles were systematically applied at positions that had previously shown incomplete conversion or deletion sequences, namely Tyr¹¹, Arg¹², Arg¹³, Pro¹⁸, Gly¹⁹, Thr²⁰, Ala²¹, Ser²², Arg²³, Tyr³⁰. This approach ensured complete coupling efficiency and aggregation.

During Fmoc deprotection, a 20% v/v solution of piperidine in DMF was used up to Gly¹⁹. This was followed by a modified protocol involving the introduction of 1M Oxyma Pure into the piperidine mixture for the remaining elongation steps. Adding Oxyma lowered the effective solution basicity, reducing the rate of DKP formation. Furthermore, the wash steps following Fmoc removal were conducted more quickly than in the standard protocol to minimise the deprotected peptide exposure to basic conditions and prevent intramolecular cyclisation. Each washing cycle comprised short DMF rinses followed by immediate neutralisation through coupling rather than the longer washing intervals employed in standard and automated procedures.

Particular attention was also given to the cleavage and side chain deprotection phase, as the presence of a methionine residue made the peptide highly susceptible to oxidation (Figure 2.10).

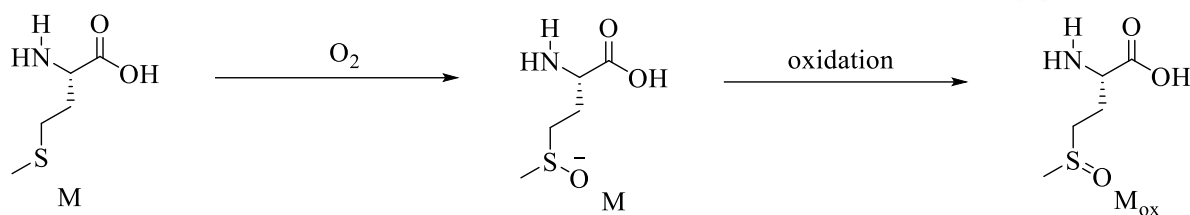


Figure 2.10. Methionine oxidation.

Initial cleavage trials using TFA/TIS/H₂O resulted in the partial oxidation of the methionine residue to the corresponding sulfoxide (Figure 2.11, A). When the scavenger system was changed to TFA/DODT/TIS/H₂O, incomplete removal of the protecting groups was observed, indicating insufficient scavenging activity (Figure 2.11, B). The optimal balance between complete deprotection and preservation of methionine was obtained using TFA/EDT/H₂O. In this mixture, EDT acts as both a strong nucleophile and a reducing agent. It efficiently traps carbocationic intermediates and prevents the oxidation of sulfur-containing residues (Figure 2.11, C).

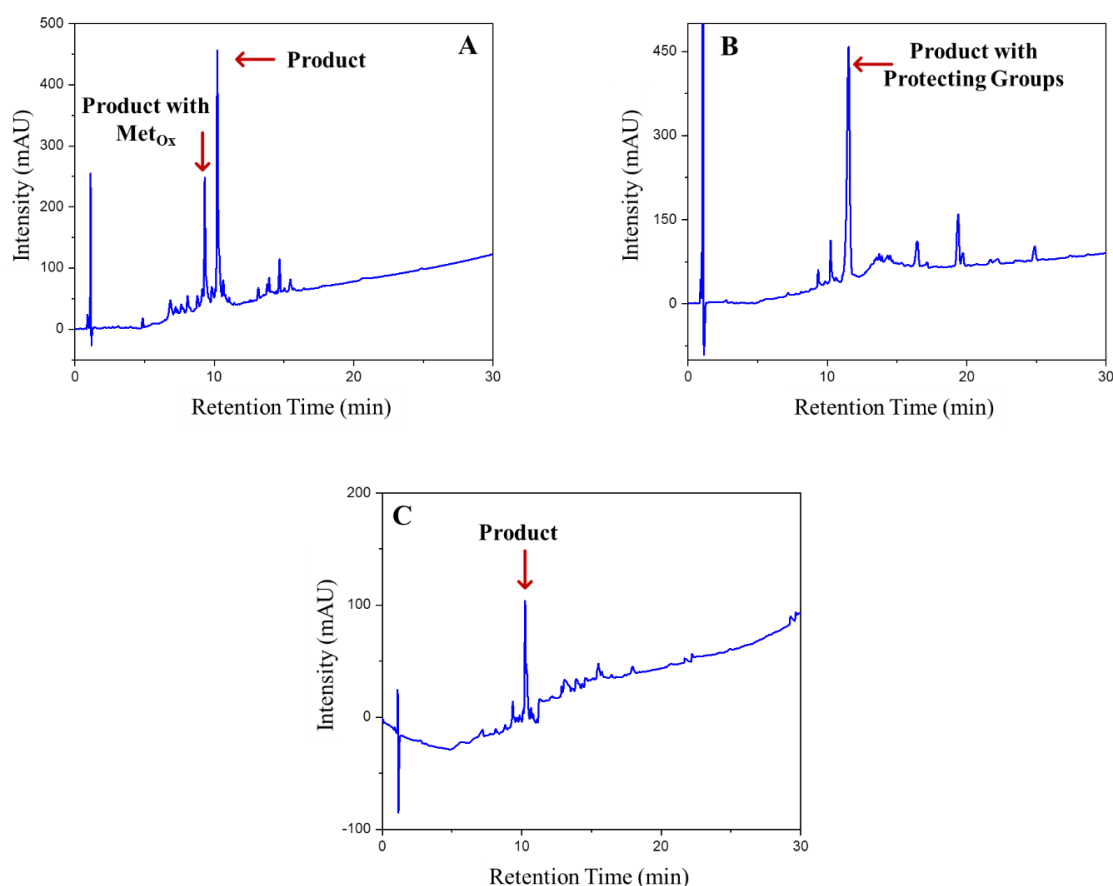


Figure 2.11. **A)** Analytical HPLC chromatogram of the H-11-30-NH₂ crude, after the cleavage in presence of TIS, recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes. The product peak is indicated by the arrow. **B)** Analytical HPLC chromatogram of the H-11-30-NH₂ crude, after the cleavage in presence of DODT, recorded as previously described for A). The product peak is indicated by the arrow. **C)** Analytical HPLC chromatogram of the H-11-30-NH₂ crude, after the cleavage in presence of EDT, recorded as previously described for A). The product peak is indicated by the arrow.

Results and Discussion



Following resin cleavage and deprotection, the crude product was purified by RP-HPLC using Luna C18 column (300 Å, 10 µm) and a binary solvent gradient ranging from 10% to 35% solvent over 25 minutes. Fractions corresponding to the major product were collected, combined and lyophilised. The identity and purity of the purified peptide were verified by analytical HPLC (Figure 2.12) and ESI-MS (Figure 2.13), confirming the successful synthesis of the desired sequence.

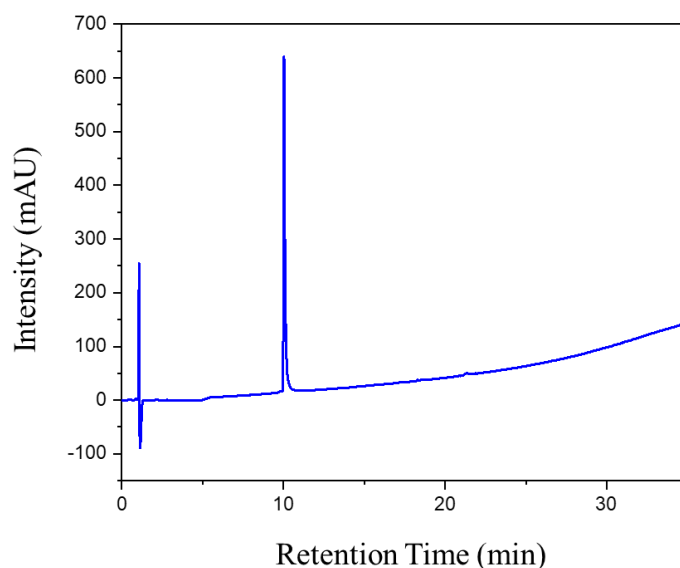


Figure 2.12. Analytical HPLC chromatogram of the H-11-30-NH₂ peptide recorded using a Kinetex C18 column, 1 mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

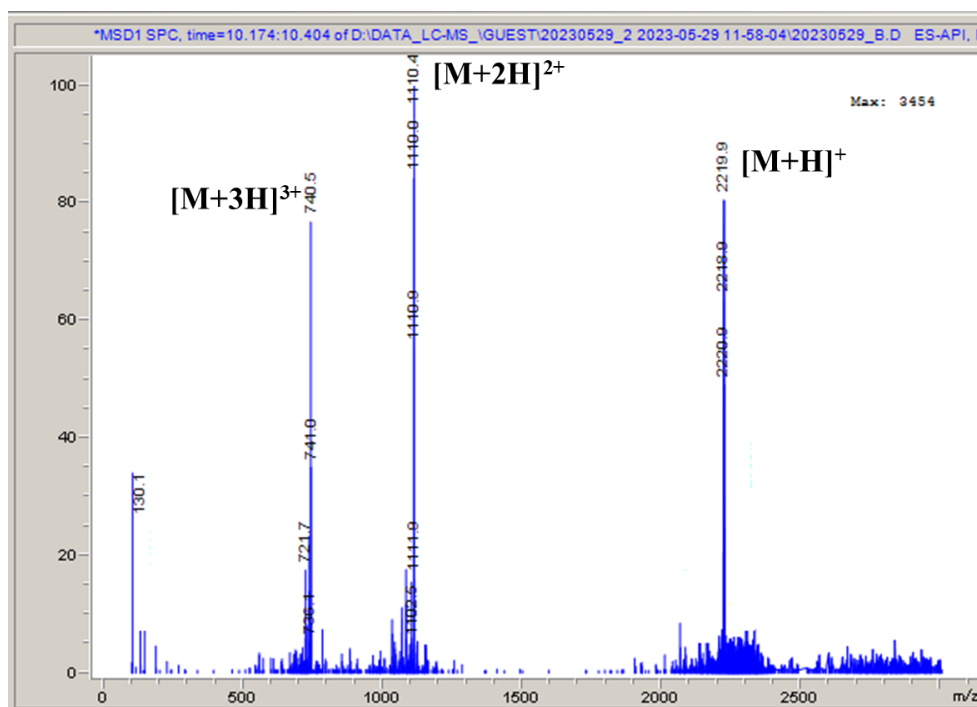


Figure 2.13. ESI-MS spectrum of H-11-30-NH₂. Molecular formula: C₉₄H₁₄₇N₃₃O₂₈S, [M+H]¹⁺ *Calculated*: 2219.1 Da; *Experimental mass*: [M+H]¹⁺: 2218.9 Da, [M+2H]²⁺: 1109.6 Da, [M+3H]³⁺: 740.2 Da.



2.1.2. Synthesis of Ac-11-30-NH₂ Peptide

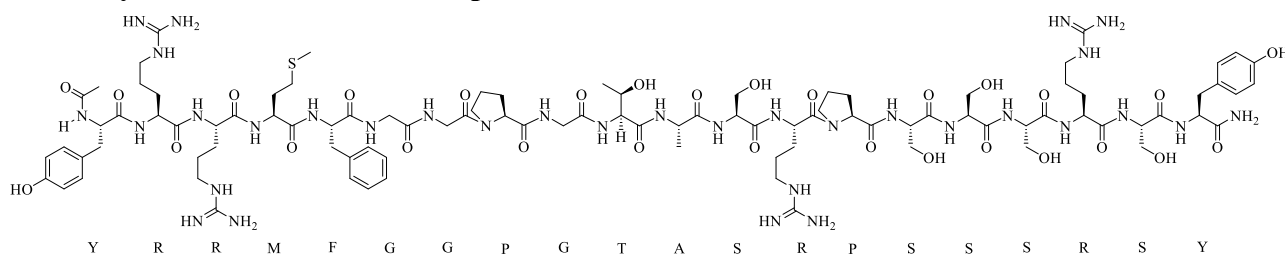


Figure 2.14. Structure of Ac-11-30-NH₂ peptide.

The Ac-11-30-NH₂ peptide (Figure 2.14) was synthesised using the same optimised manual SPPS protocol as for the non-acetylated analogue. Following completion of the chain assembly, N-terminal acetylation was performed in two consecutive one-hour coupling steps using a mixture of acetic anhydride, DIPEA and DMF.

The cleavage and deprotection steps were carried out under the same optimised conditions as for H-11-30-NH₂ to prevent oxidation of the methionine residue. The crude peptide was purified by RP-HPLC (Figure 2.15, Figure 2.16) using the same chromatographic protocol as described above.

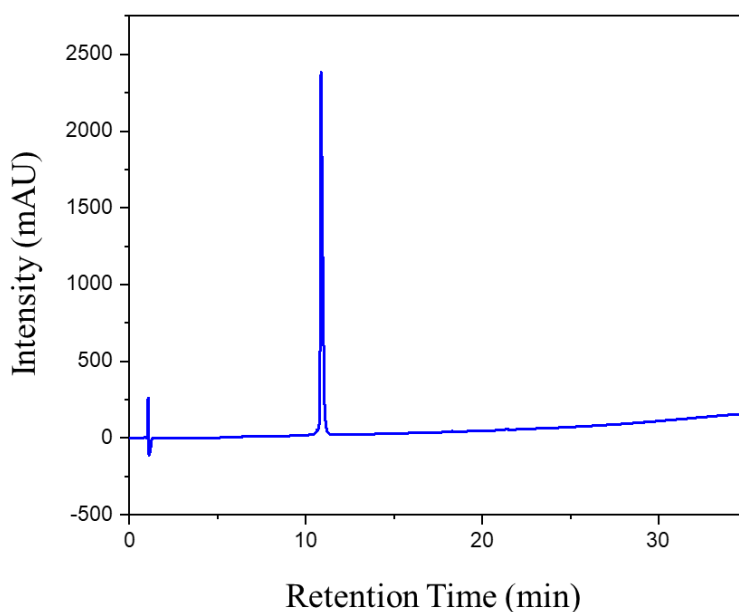


Figure 2.15. Analytical HPLC chromatogram of the Ac-11-30-NH₂ peptide recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

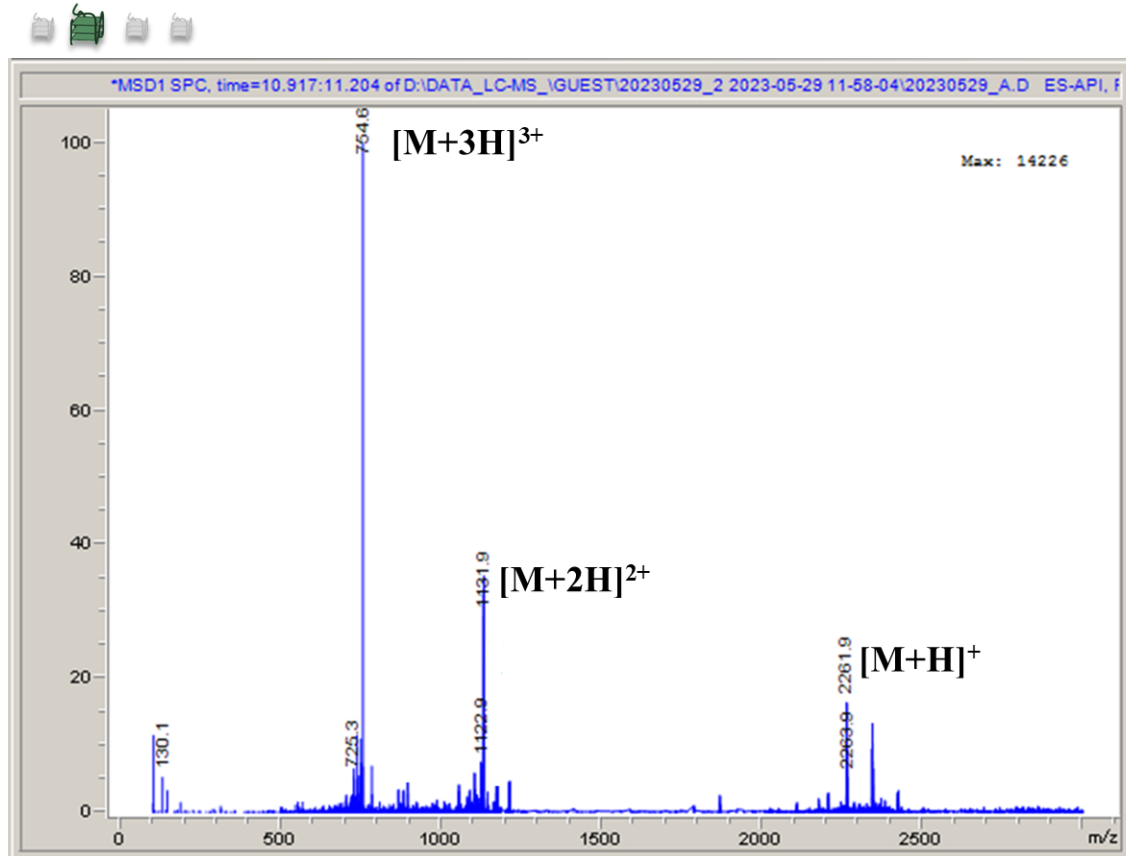


Figure 2.16. ESI-MS spectrum of Ac-11-30-NH₂. Molecular formula: C₉₆H₁₄₉N₃₃O₂₈S, [M+H]¹⁺ Calculated: 2261.5 Da; Experimental mass: [M+H]⁺:2260.9 Da, [M+2H]²⁺:1131.4 Da, [M+3H]³⁺:754.3 Da.

2.1.3. Synthesis of H-68-79-NH₂ Peptide

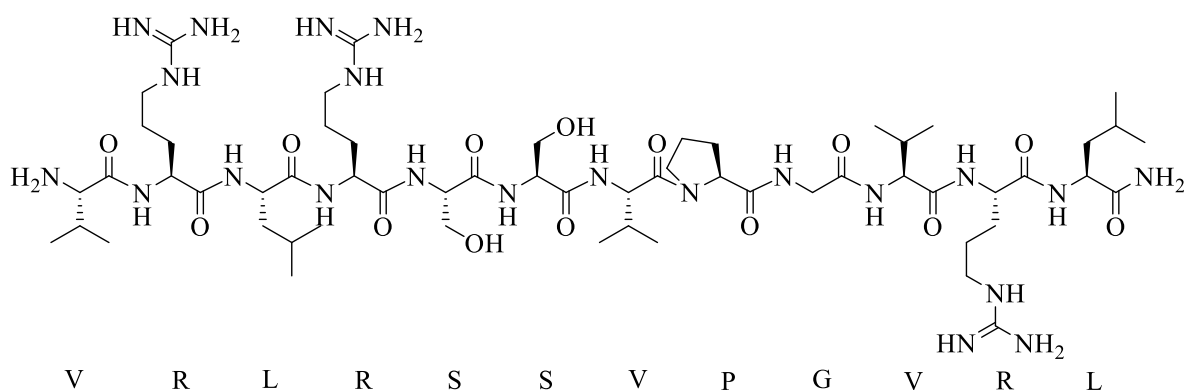


Figure 2.17. Structure of H-68-79-NH₂ peptide.

The H-68-79-NH₂ peptide (Figure 2.17), which was designed as a negative control of the N-terminal region for internalisation assays, was synthesised manually using SPPS.

The synthesis proceeded smoothly with no significant side reactions or deletion sequences occurring. The sequence was then purified using preparative reverse-phase HPLC with a Luna C18 column (300 Å, 10 µm) and a binary solvent gradient ranging from 10% to 40% solvent B over 40 minutes. Fractions corresponding to the main product were collected and lyophilised. The identity and purity of the final compound were confirmed by analytical HPLC (Figure 2.18) and ESI-MS (Figure 2.19).

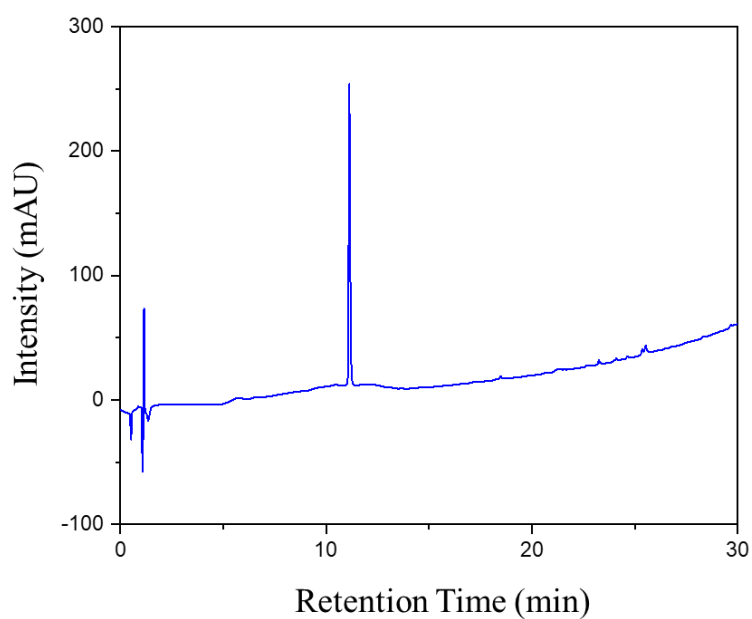


Figure 2.18. Analytical HPLC chromatogram of the H-68-79-NH₂ peptide recorded using a Kinetex C18 column, 1 mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

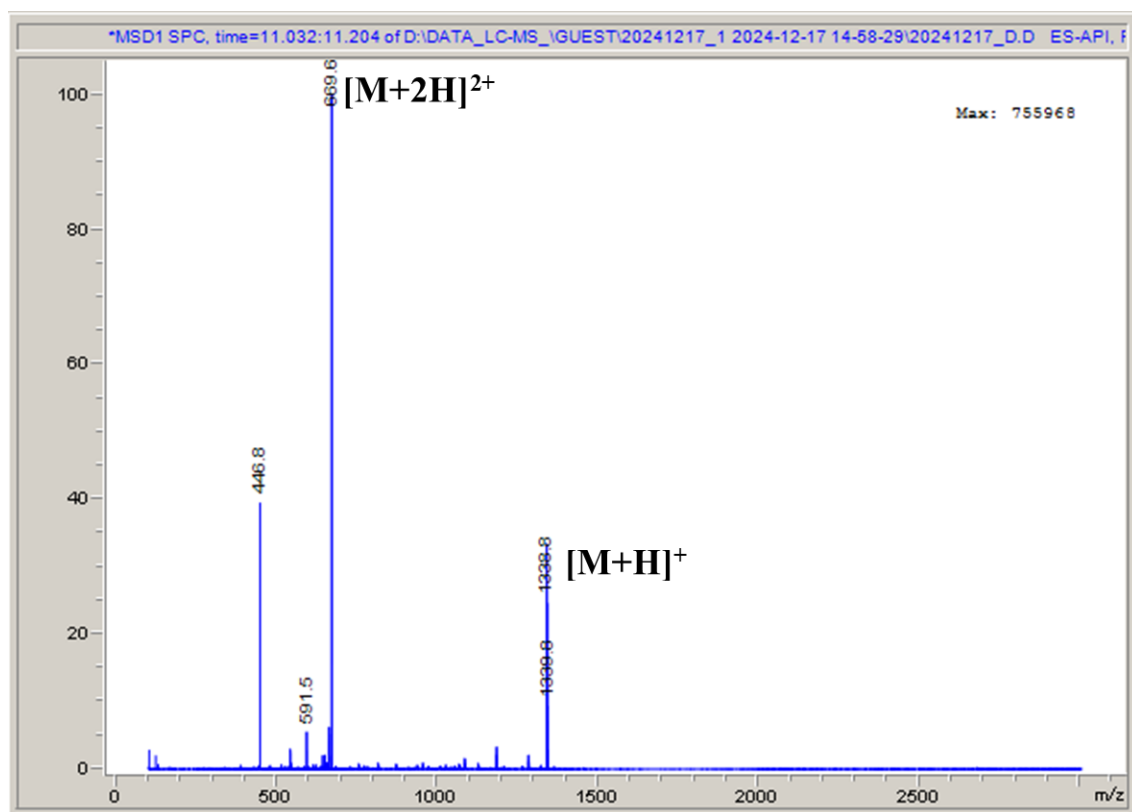


Figure 2.19. ESI-MS spectrum of H-68-79-NH₂. Molecular formula: C₅₈H₁₀₈N₂₂O₁₄, [M+H]¹⁺ Calculated: 1338.6 Da; Experimental mass: [M+H]⁺: 1337.8 Da, [M+2H]²⁺: 669.1 Da.



2.1.4. Synthesis of Ac-68-79-NH₂ Peptide

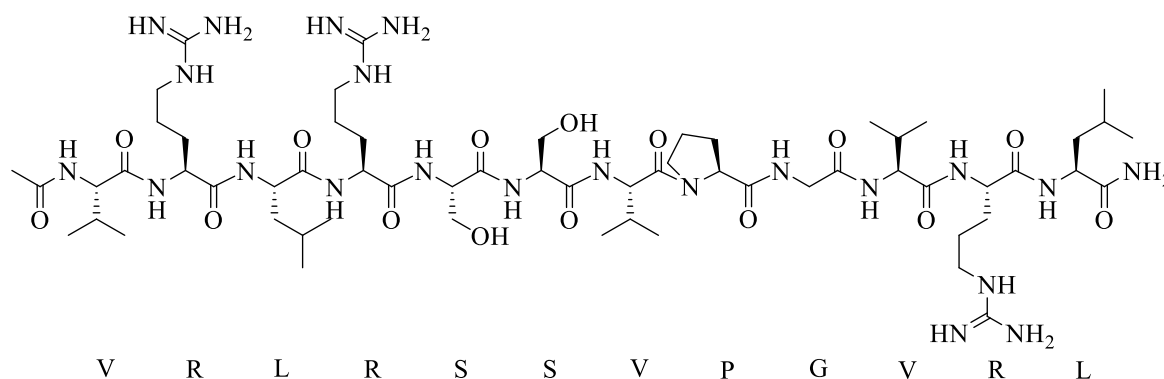


Figure 2.20. Structure of Ac-68-79-NH₂ peptide.

The Ac-68-79-NH₂ peptide (Figure 2.20) was synthesised manually using SPPS under optimised conditions previously established for the H-68-79-NH₂ sequence.

Once chain assembly was complete, N-terminal acetylation was performed in two consecutive one-hour coupling steps using a mixture of acetic anhydride, DIPEA and DMF. After final cleavage from the resin, the crude peptide was purified using the same chromatographic protocol as for the H-68-79-NH₂ peptide (Figure 2.21, Figure 2.22).

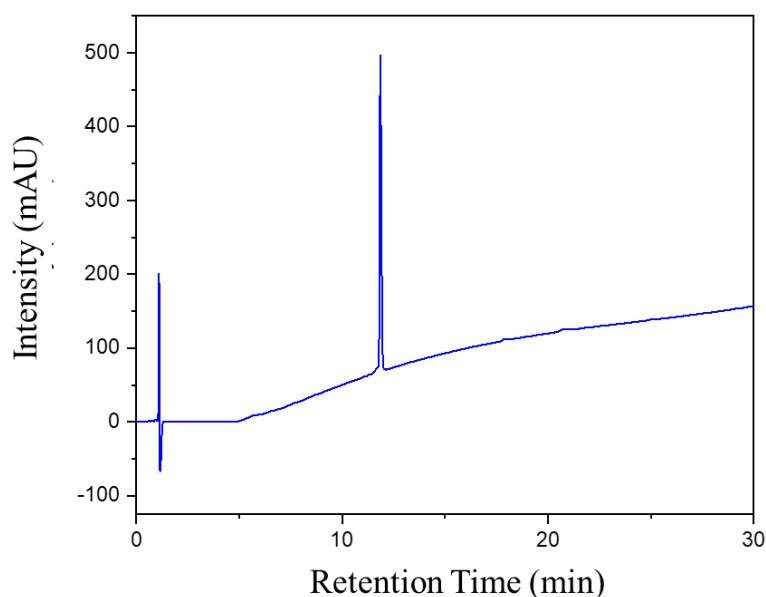


Figure 2.21. Analytical HPLC chromatogram of the Ac-68-79-NH₂ peptide recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

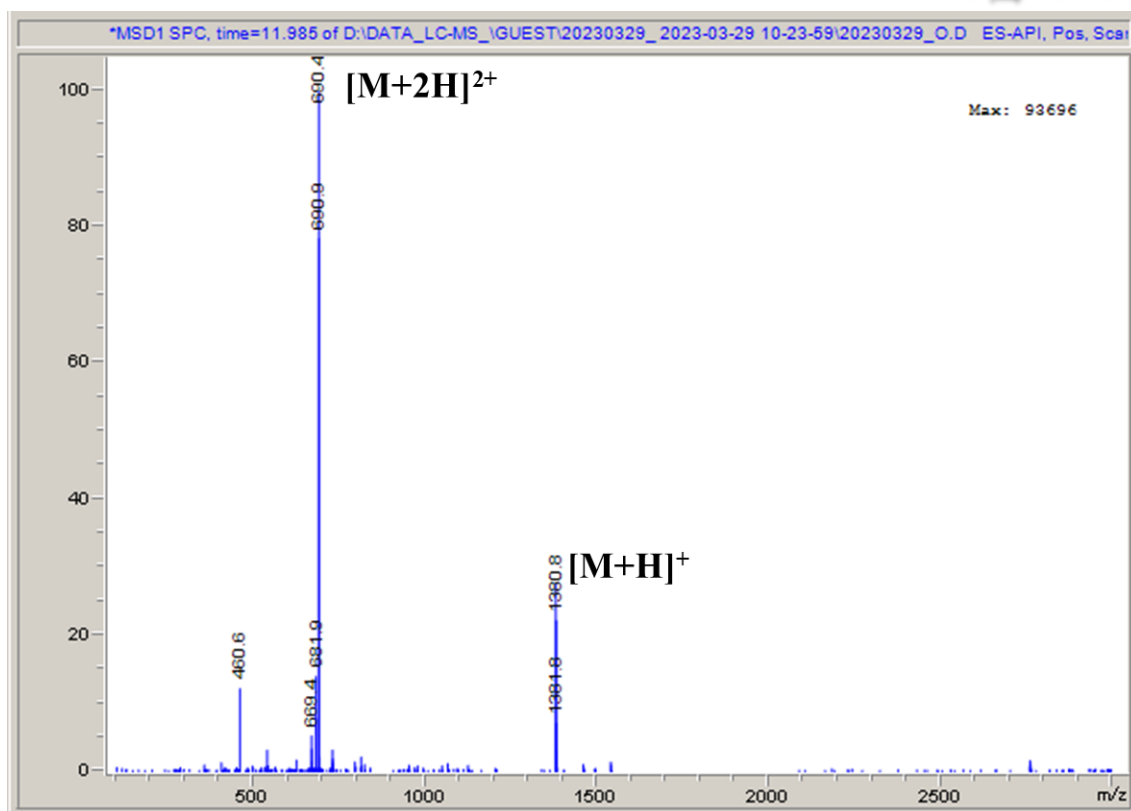


Figure 2.22. ESI-MS spectrum of Ac-68-79-NH₂. Molecular formula: C₉₆H₁₄₉N₃₃O₂₈S, [M+H]¹⁺ Calculated: 1379.8 Da; Experimental mass: [M+H]⁺:1379.8 Da, [M+2H]²⁺:690.0 Da.

2.1.5. Synthesis of H-270-280-NH₂ Peptide

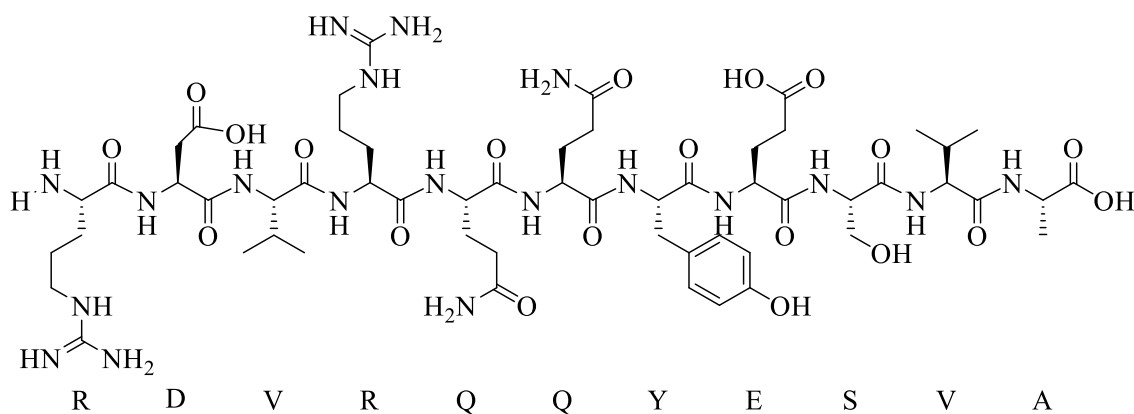


Figure 2.23. Structure of H-270-280-NH₂ peptide.

The H-270-280-NH₂ peptide (Figure 2.23) was designed as another positive control from the protein central domain. It was synthesised manually using SPPS.

To ensure complete coupling and prevent sequence truncation, double coupling steps were performed for the two consecutive Gln residues at positions 274 and 275. The synthesis proceeded smoothly with no evidence of significant side reactions or deletion sequences.

Results and Discussion



Following resin cleavage and deprotection, the crude product was purified by RP-HPLC using Luna C18 column (300 Å, 10 µm) and a binary solvent gradient ranging from 5% to 35% solvent over 30 minutes. Fractions corresponding to the major product were collected, combined and lyophilised. The identity and purity of the purified peptide were verified by analytical HPLC (Figure 2.24) and ESI-MS (Figure 2.25), confirming the successful synthesis of the desired sequence.

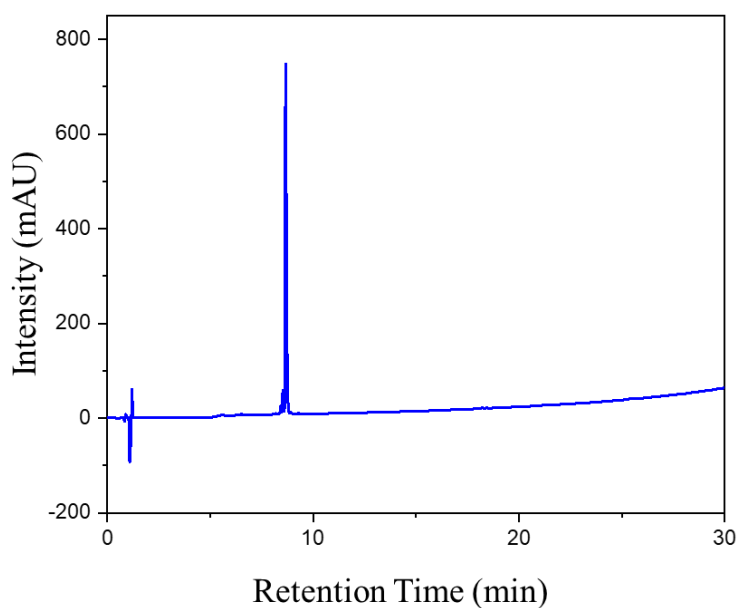


Figure 2.24. Analytical HPLC chromatogram of the H-270-280-NH₂ peptide recorded using a Kinetex C18 column, 1 mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

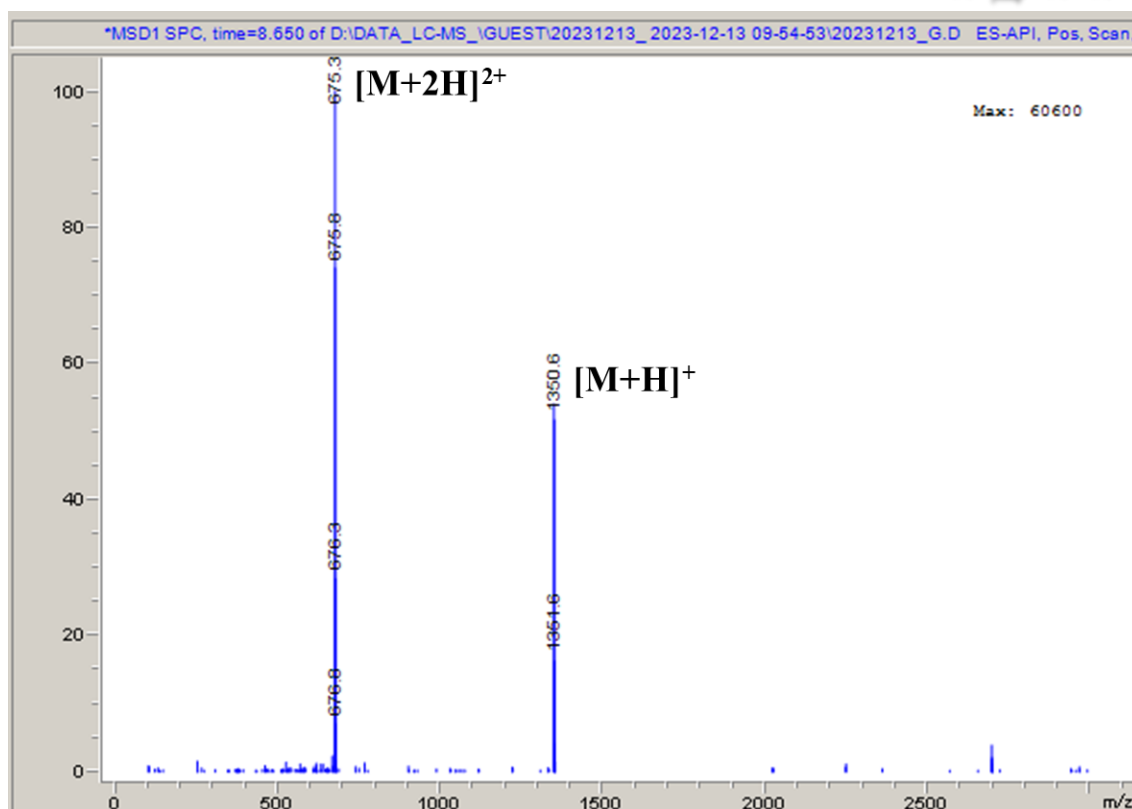


Figure 2.25. ESI-MS spectrum of H-270-280-NH₂. Molecular formula: C₉₆H₁₄₉N₃₃O₂₈S, [M+H]¹⁺ Calculated: 1349.5 Da; Experimental mass: [M+H]⁺:1349.6 Da, [M+2H]²⁺:675.3 Da.

2.1.6. Synthesis of Ac-270-280-NH₂ Peptide

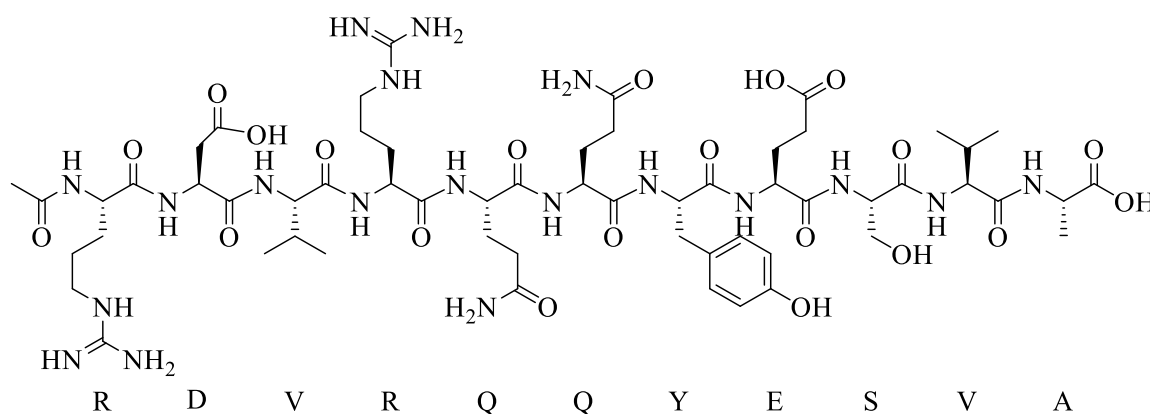


Figure 2.26. Structure of Ac-270-280-NH₂.

The Ac-270-280-NH₂ peptide (Figure 2.26) was synthesised manually using SPPS under the optimised conditions previously established for the H-270-280-NH₂ peptide. Following chain assembly, N-terminal acetylation was performed in two consecutive one-hour coupling steps using a mixture of acetic anhydride, DIPEA and DMF.

After resin cleavage and global deprotection, the crude peptide was purified using the same preparative RP-HPLC protocol as for H-270-280-NH₂ sequence (Figure 2.27, Figure 2.28).

Results and Discussion

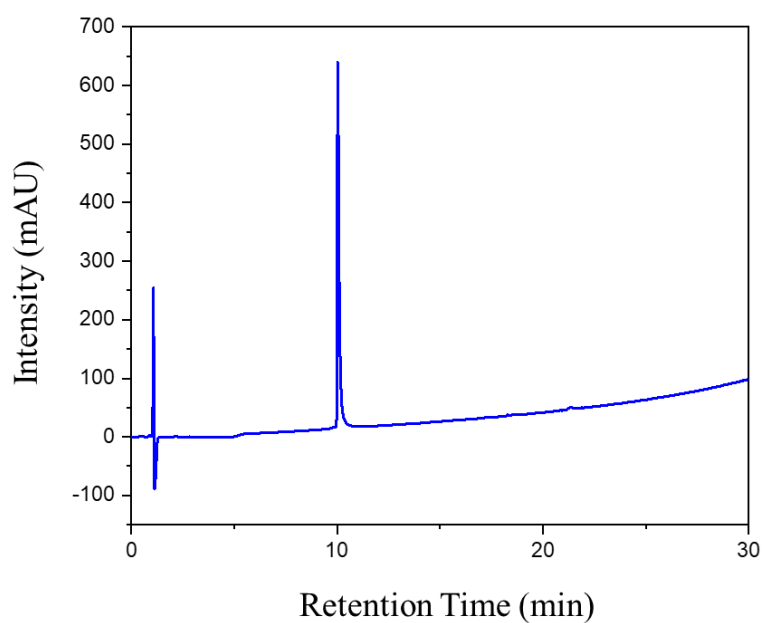


Figure 2.27. Analytical HPLC chromatogram of the Ac-270-280-NH₂ peptide recorded using a Kinetex C18 column, 1 mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

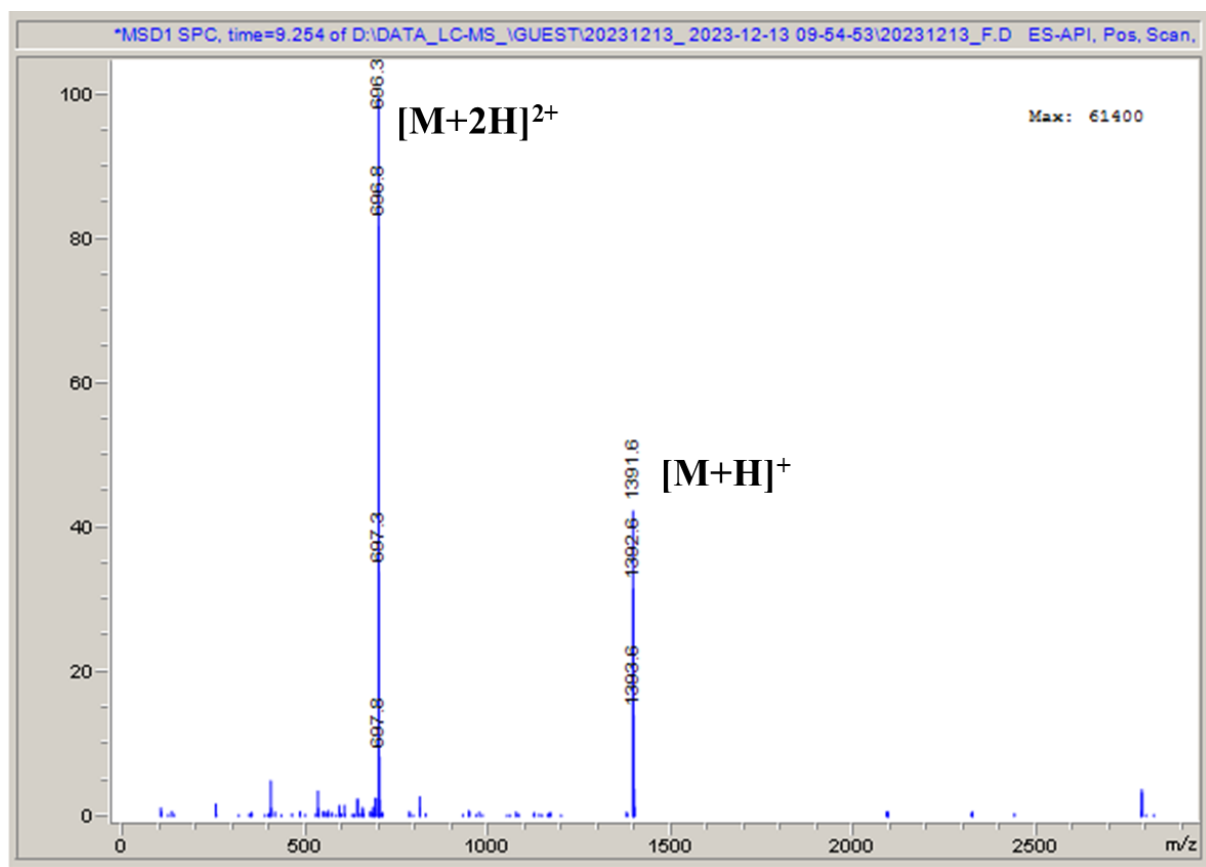


Figure 2.28. ESI-MS spectrum of Ac-270-280-NH₂. Molecular formula: C₅₈H₉₃N₁₉O₂₁, $[M+H]^+$ Calculated: Da; Experimental mass: $[M+H]^+$: 2260.9 Da, $[M+2H]^{2+}$: 1131.4 Da, $[M+3H]^{3+}$: 754.3 Da.



2.2. Methionine Substitution with Norleucine for Enhanced Stability

To synthesise the next peptide variants, Met was systematically replaced with norleucine (Nle) to enhance the chemical stability of the peptides during synthesis and storage. This substitution is a well-established strategy in peptide chemistry, since Nle is isosteric analogue of methionine which preserves its hydrophobicity and steric profile, but lacks the sulfur atom, that is highly susceptible to oxidation (Figure 2.29). Met modification can alter the polarity and conformation of peptides, compromise their biological activity.

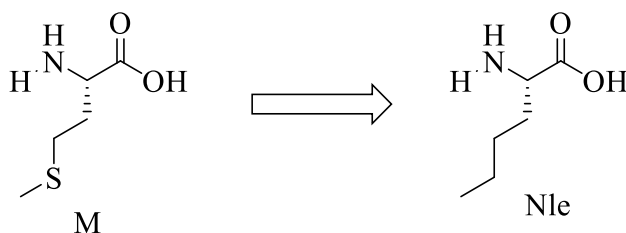


Figure 2.29. Comparison between methionine and Nle.

Replacing Met with Nle eliminates the risk of oxidation without affecting overall physicochemical properties of the peptide – in particular, its side chain volume remains essentially unchanged. This makes Nle an ideal conservative substitute for stabilising the peptide during acidic cleavage, lyophilisation and long-term storage.

To confirm that this substitution did not affect the peptide biological function, preliminary binding experiments were conducted with the G4.

2.2.1. Convergent Synthesis of H-[Nle¹⁴]-11-30-NH₂ peptide

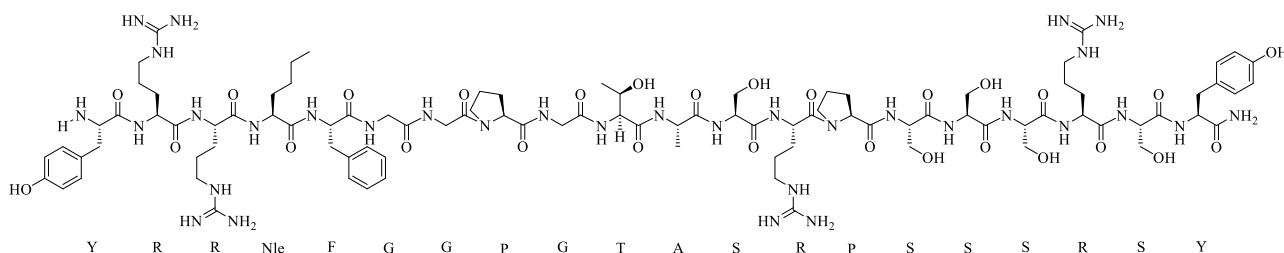


Figure 2.30. Structure of H-[Nle¹⁴]-11-30-NH₂ peptide.

This peptide was designed to verify the effect of the Met-to-Nle substitution on the structural and synthetic behaviour. It was obtained through the on-resin conjugation of two protected peptide fragments – Fmoc-[Nle¹⁴]-11-17-OH and H-18-30-NH₂ – using DIC/HOAt as the coupling agents (Figure 2.30). The condensation reaction was allowed to proceed for four hours under gentle agitation to ensure complete fragment coupling and minimise potential side reactions.

The synthesis proceeded smoothly with no significant side reactions or deletion sequences occurring. The sequence was then purified using preparative reverse-phase HPLC with a Luna C18 column (300

Results and Discussion



Å, 10 µm) and a binary solvent gradient ranging from 5% to 35% solvent B over 40 minutes. Fractions corresponding to the main product were collected and lyophilised. The identity and purity of the final compound were confirmed by analytical HPLC (Figure 2.31) and ESI-MS (Figure 2.32).

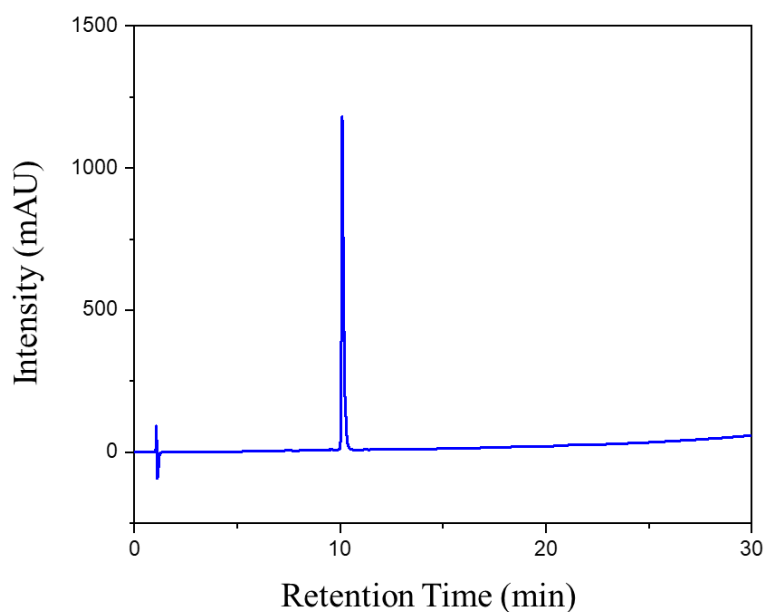


Figure 2.31. Analytical HPLC chromatogram of the H -[Nle¹⁴]-11-30-NH₂ peptide recorded using a Kinetex C18 column, 1 mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

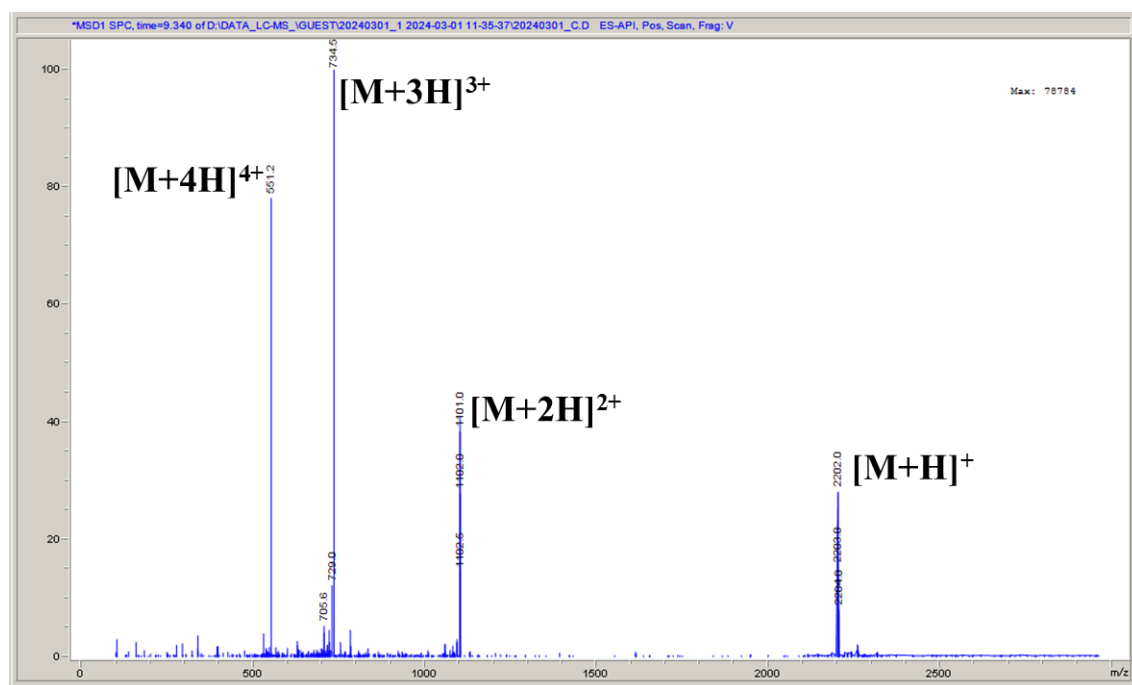


Figure 2.32. ESI-MS spectrum of H -[Nle¹⁴]-11-30-NH₂. Molecular formula: $C_{95}H_{149}N_{33}O_{28}$, $[M+H]^+$ Calculated: 2201.4 Da; Experimental mass: $[M+H]^+$: 2201.0 Da, $[M+2H]^{2+}$: 1100.5 Da, $[M+3H]^{3+}$: 734.2 Da, $[M+4H]^{4+}$: 551.0 Da.



2.3. Alanine Scanning

The alanine scanning (Ala scan) approach is one of the most powerful and widely used tools in peptide and protein chemistry for determining structure-activity relationships. This method involves systematically replacing selected amino acid residues within a peptide or protein sequence with alanine and then evaluating how each substitution affects the molecule biological or biophysical properties.

Alanine is the residue of choice for this type of study because it is small, uncharged and chemically inert. Its methyl side chain lacks the capacity for hydrogen bonding or electrostatic interaction, yet it preserves the overall peptide backbone conformation and maintains α -helical or β -structural propensities comparable to those of most natural residues. This enables to effectively ‘silence’ the side chain functionality of a specific residue – whether aromatic, charged, or polar – without introducing steric distortion or significant conformational rearrangements. Consequently, any observed difference in activity or binding affinity can be directly attributed to the loss of the original side chain functional contribution rather than to a global change in peptide folding.

Practically, the Ala scan technique involves synthesising a series of peptide analogues, each with a single Ala substitution in a different position. The functional or binding properties of these analogues are then compared with those of the wild-type sequence. Systematic analysis of how each substitution affects interaction with the target makes it possible to map the key residues involved in molecular recognition, quantify their relative importance and identify potential binding spots.

In this study, Ala scan was applied to the Vimentin-derived H-11-30-NH₂ peptide: six variants (Table 2.1) were synthesised – three involving aromatic residues and three involving the Arg residues – to explore how each side chain influences recognition of the G4 sequence.

Table 2.1. List of all peptide sequences synthesized for Ala scan.

Ala Scan	Peptide Sequence
1	H-ARRNleFGGPGTASRPSSRSY-NH ₂
2	H-YRRNleAGGPGTASRPSSRSY-NH ₂
3	H-YRRNleFGGPGTASRPSSRS Δ -NH ₂
4	H-Y Δ ANleFGGPGTAS Δ PSSS Δ SY-NH ₂
5	H-Y Δ ANleFGGPGTASRPSSRSY-NH ₂
6	H-YRRNleFGGPGTAS Δ PSSS Δ SY-NH ₂

This approach provides a quantitative and mechanistic understanding of how each side chains modulate the interaction, laying the groundwork for the rational design of improved peptide analogues with enhanced stability and binding selectively towards G4 structures.



2.4. Fragment-Based and Convergent Peptide Synthesis

As peptide length and sequence complexity increase, traditional stepwise SPPS is often limited by incomplete couplings, side reactions and cumulative yield losses. Long or aggregation prone sequences pose significant synthetic challenges due to steric hindrance and intra-chain hydrogen bonding during elongation. To overcome these difficulties and obtain high-purity material for subsequent structural and binding studies, a fragment based convergent synthesis approach was employed (Figure 2.33).

In this strategy, the target peptide is divided into shorter fragments that are synthesised independently using Fmoc-based SPPS, following optimised coupling and deprotection cycles. Unlike traditional linear synthesis, where the peptide is assembled continuously, this method allows the synthesis of shorter, more manageable fragments. These fragments are synthesised in parallel with the appropriate protective groups, which are attached to each amino acid to prevent undesired side reactions.

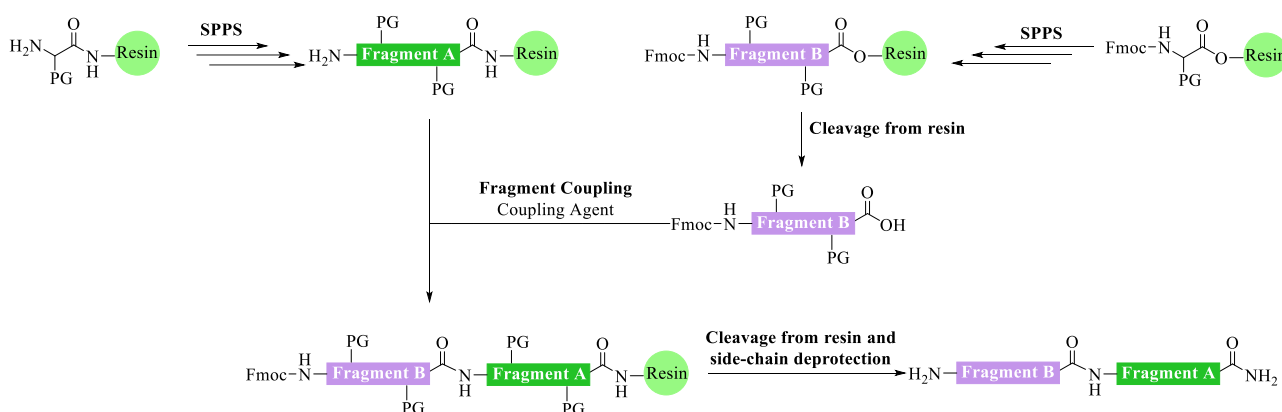


Figure 2.33. Schematic representation of a convergent synthesis strategy.

Once purified, the fragments are joined through solution-phase condensation or on-resin coupling, depending on their solubility and compatibility with the resin. High efficiency coupling reagents are generally used to promote fragment ligation while minimising racemisation. The reaction is conducted under mild conditions to preserve side chain protecting groups until the final global deprotection and cleavage step.

In this study, alanine-substituted Vimentin fragments were synthesised using an approach that combined Fmoc-based SPPS for fragments preparation and convergent fragment coupling for full-sequence assembly. This strategy provided multiple advantages:

- Improved overall yield and purity due to shorter synthesis times and minimised side reactions.
- Reduced aggregation, since difficult regions were handled as soluble intermediates.

Finally, this methodology was essential for producing enough purified peptide variants for interaction studies with G4s.



2.4.1. Synthesis of the Fmoc-18-30-NH₂ fragment

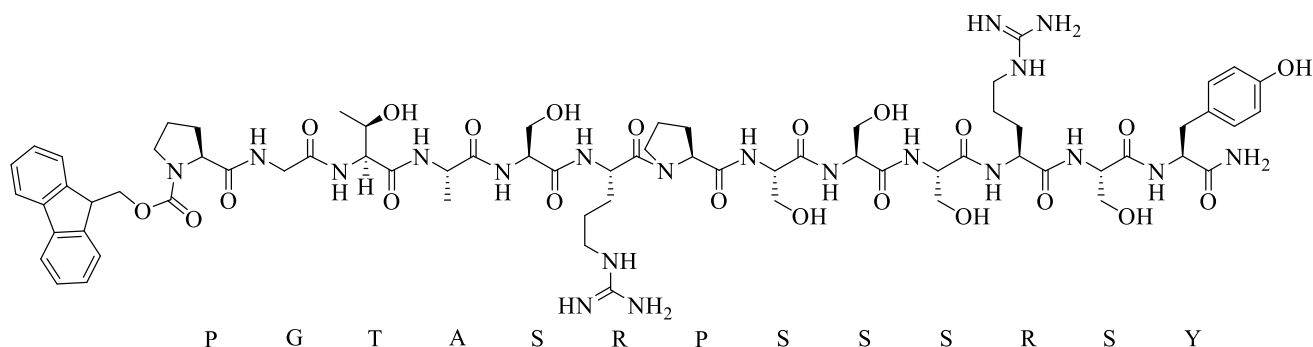


Figure 2.34. Structure of Fmoc-18-30-NH₂.

The Fmoc-18-30-NH₂ peptide (Figure 2.34) was synthesised manually using Fmoc-based SPPS. To ensure complete coupling efficiency and prevent sequence truncation, double coupling steps were systematically introduced for residues Pro¹⁸, Gly¹⁹, Thr²⁰, Ala²¹, Ser²², Arg²³ and Tyr³⁰. The synthesis proceeded smoothly with no evidence of significant side reactions or deletion sequences. Upon completion of the elongation steps, the peptide remained attached to the resin and was subsequently employed as a fragment in the convergent synthesis strategy. This enabled its coupling with the complementary peptide segment under optimised fragment condensation conditions (Figure 2.35, Figure 2.36).

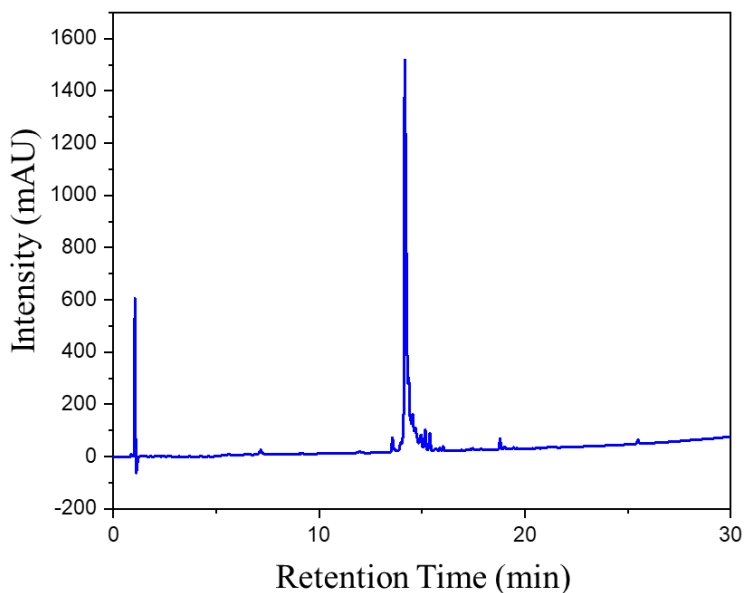


Figure 2.35. Analytical HPLC chromatogram of the Fmoc-18-30-NH₂ peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

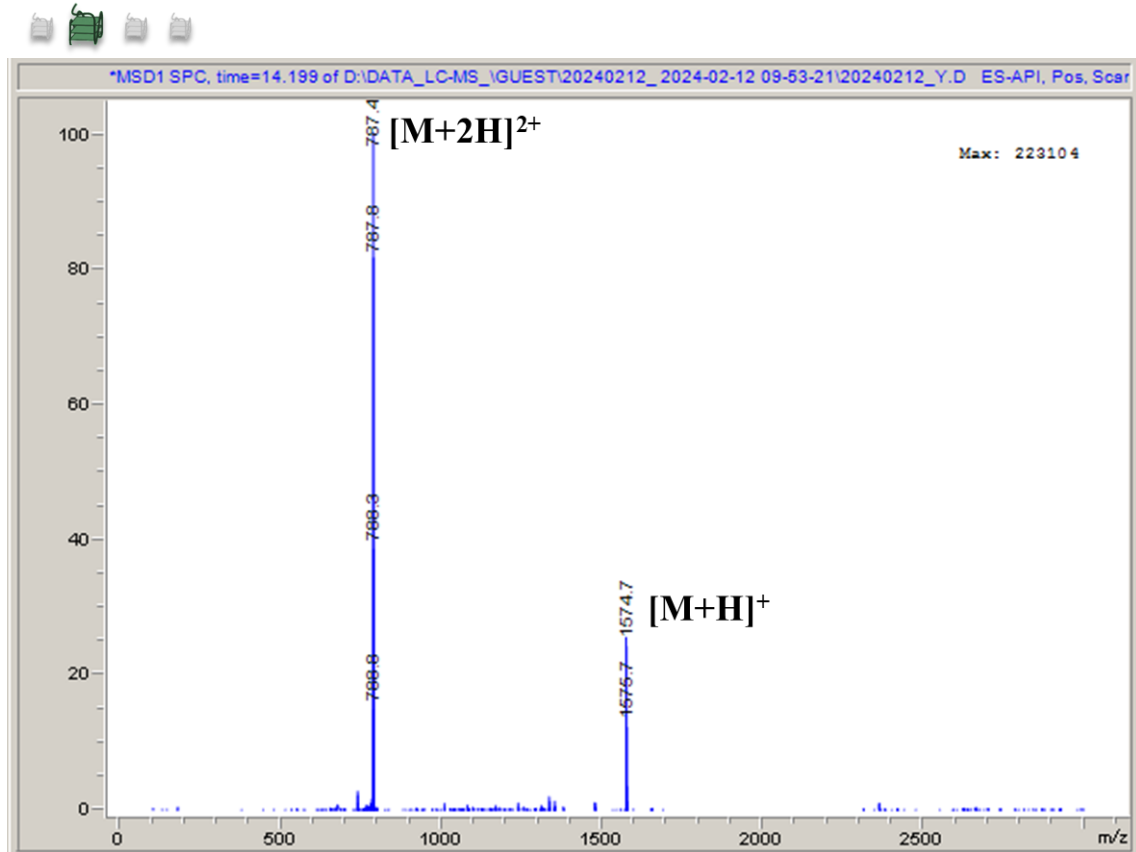


Figure 2.36. ESI-MS spectrum of Fmoc-18-30-NH₂. Molecular formula: C₇₀H₁₀₀N₂₀O₂₂, [M+H]¹⁺ Calculated: 1573.7 Da; Experimental mass: [M+H]¹⁺:1573.7 Da, [M+2H]²⁺: 787.0 Da.

2.4.2. Synthesis of the Fmoc-[Ala³⁰]-18-30-NH₂ fragment

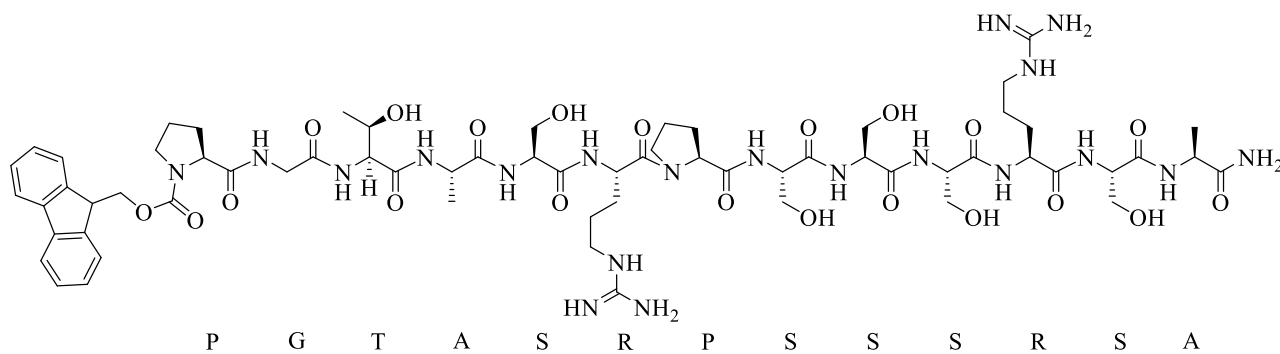


Figure 2.37. Structure of Fmoc-[Ala³⁰]-18-30-NH₂ fragment.

The Fmoc-[Ala³⁰]-18-30-NH₂ peptide (Figure 2.37) was synthesised manually using Fmoc-based SPPS. To ensure complete coupling efficiency and prevent sequence truncation, double coupling steps were systematically introduced for residues Pro¹⁸, Gly¹⁹, Thr²⁰, Ala²¹, Ser²², and Arg²³. The synthesis proceeded smoothly with no evidence of significant side reactions or deletion sequences. Upon completion of the elongation steps, the peptide remained attached to the resin and was subsequently employed as a fragment in the convergent synthesis strategy. This enabled its coupling with the complementary peptide segment under optimised fragment condensation conditions (Figure 2.38, Figure 2.39).

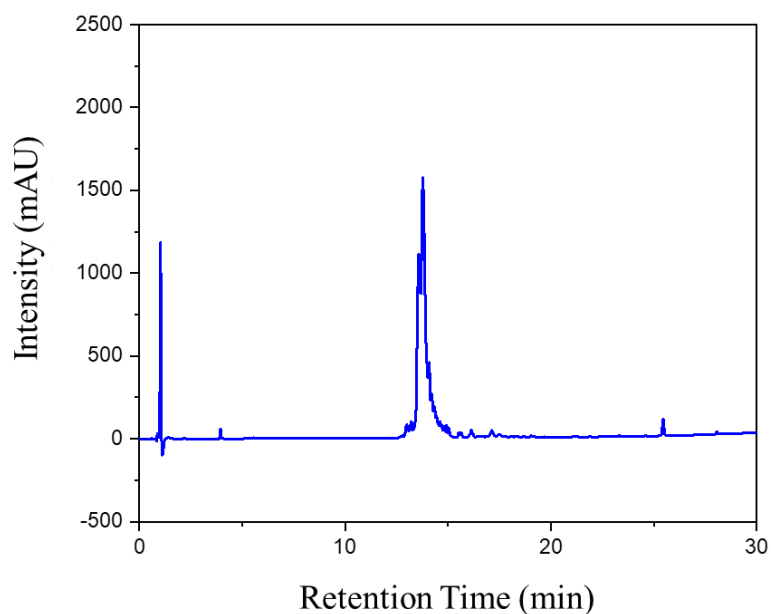


Figure 2.38. Analytical HPLC chromatogram of the Fmoc-[Ala³⁰]-18-30-NH₂ fragment crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

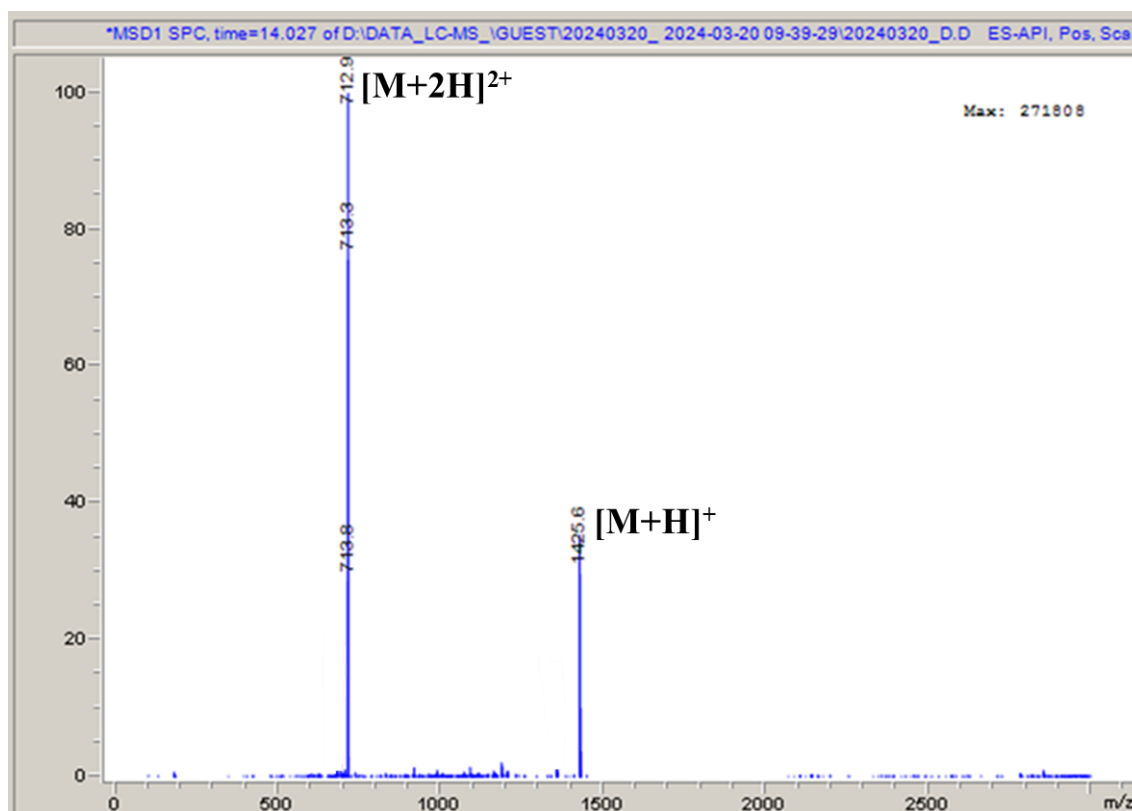


Figure 2.39. ESI-MS spectrum of Fmoc-[Ala³⁰]-18-30-NH₂ fragment. Molecular formula: C₆₄H₉₆N₂₀O₂₁, [M+H]¹⁺ Calculated: 1425.3 Da; Experimental mass: [M+H]⁺: 1424.6 Da, [M+2H]²⁺: 712.4 Da.



2.4.3. Synthesis of the Fmoc-[Ala^{23,28}]-18-30-NH₂ fragment

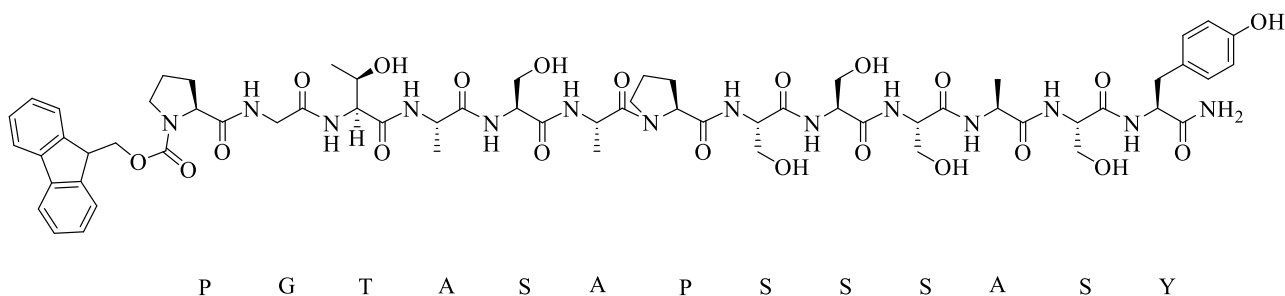


Figure 2.40. Structure of Fmoc-[Ala^{23,28}]-18-30-NH₂ fragment.

The Fmoc-[Ala^{23,28}]-18-30-NH₂ peptide (Figure 2.40) was synthesised manually using Fmoc-based SPPS. To ensure complete coupling efficiency and prevent sequence truncation, double coupling steps were systematically introduced for residues Pro¹⁸, Gly¹⁹, Thr²⁰, Ala²¹, Ser²², and Tyr³⁰. The synthesis proceeded smoothly with no evidence of significant side reactions or deletion sequences. Upon completion of the elongation steps, the peptide remained attached to the resin and was subsequently employed as a fragment in the convergent synthesis strategy. This enabled its coupling with the complementary peptide segment under optimised fragment condensation conditions (Figure 2.41, Figure 2.42).

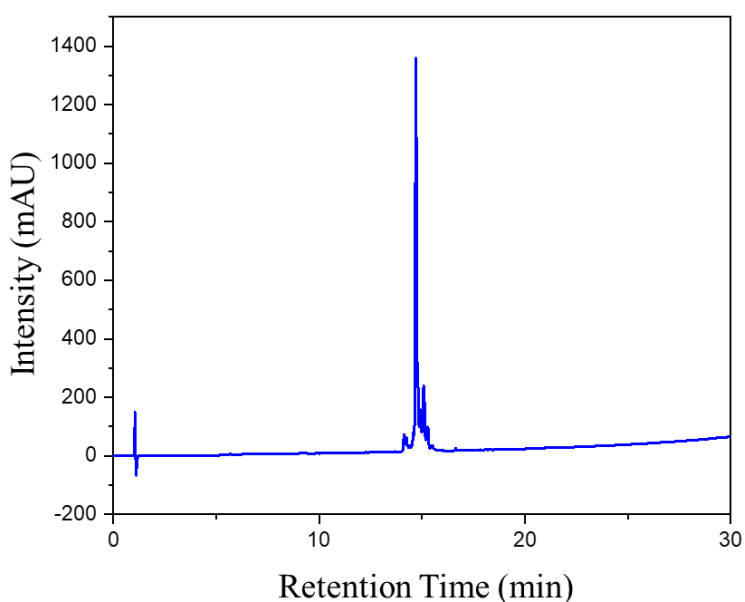


Figure 2.41. Analytical HPLC chromatogram of the Fmoc-[Ala^{23,28}]-18-30-NH₂ fragment recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

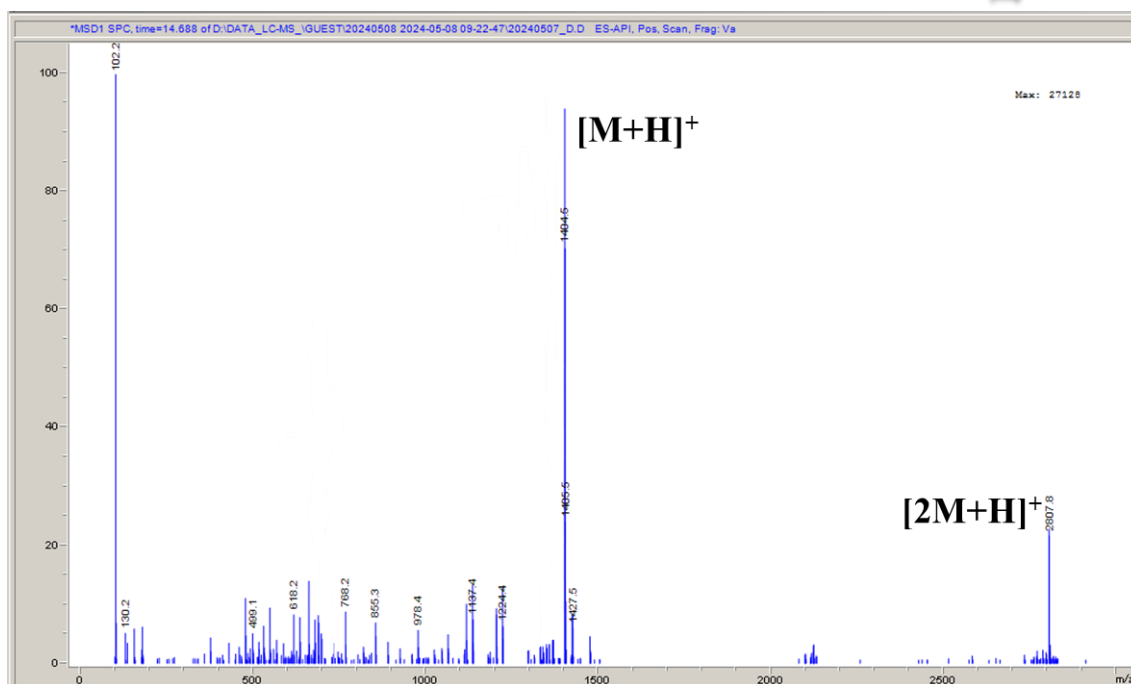


Figure 2.42. ESI-MS spectrum of Fmoc-[Ala^{23,28}]-18-30-NH₂ fragment. Molecular formula: C₆₄H₈₆N₁₄O₂₂, [M+H]¹⁺ Calculated: 1403.6 Da; Experimental mass: [M+2H]²⁺: 2201.0 Da, [2M+H]⁺: 2806.8 Da.

2.4.4. Synthesis of the Fmoc-[Nle¹⁴]-11-17-OH fragment

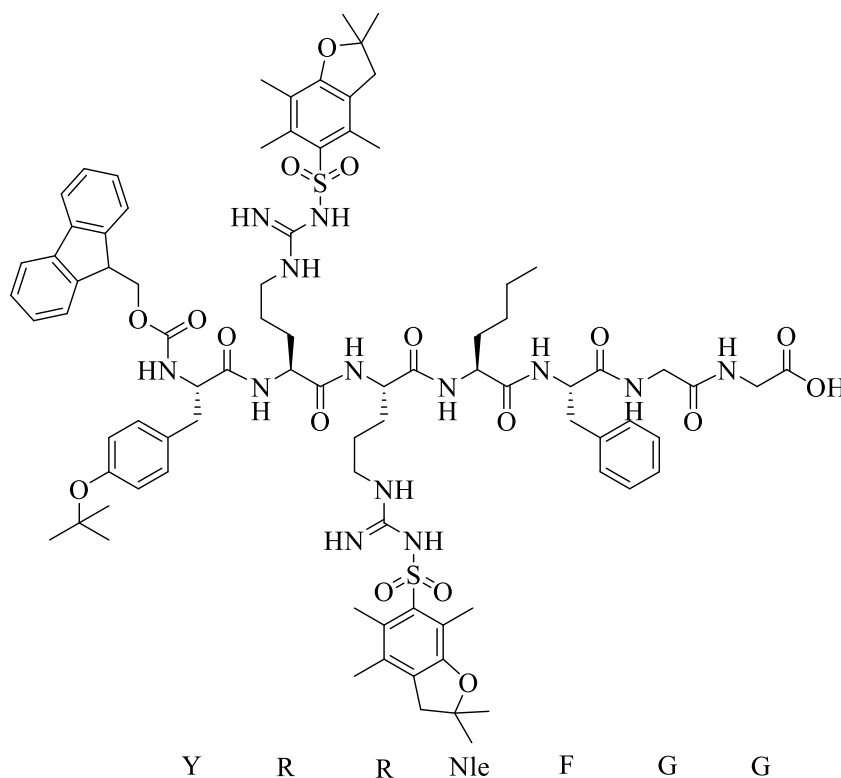


Figure 2.43. Structure of the Fmoc-[Nle¹⁴]-11-17-OH fragment.

The Fmoc-[Nle¹⁴]-11-17-OH peptide (Figure 2.43) was synthesised manually using Fmoc-based SPPS. To ensure complete coupling efficiency and minimise the occurrence of truncated sequences,

Results and Discussion



double coupling steps were performed for residues: Tyr¹¹, Arg¹², Arg¹³. The synthesis proceeded smoothly with no evidence of significant side reactions or deletion sequences.

Upon completion of the elongation steps, the peptide was cleaved from the resin using a 30% hexafluoroisopropanol (HFPI) solution in DCM. This mild cleavage procedure effectively released the protected peptide fragment while preserving the side chain protecting groups on all residues. This allowed the fragment to be used directly in the subsequent convergent coupling reaction without the need for further modification (Figure 2.44, Figure 2.45).

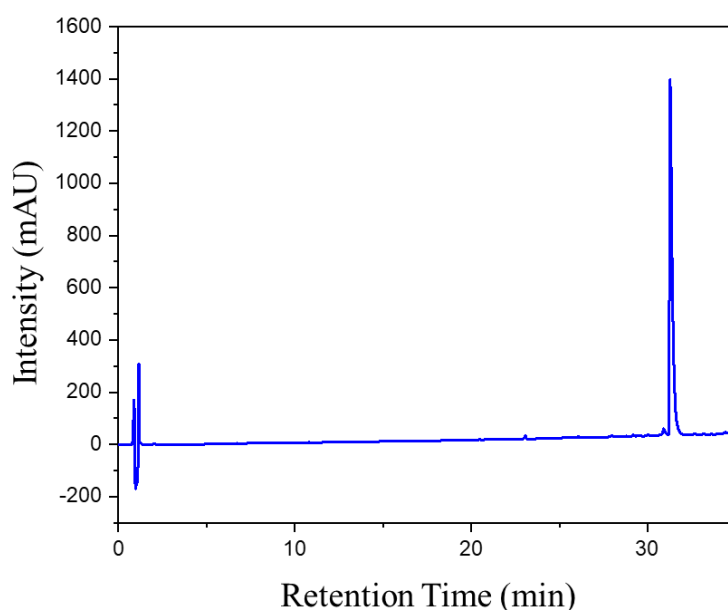


Figure 2.44. Analytical HPLC chromatogram of the Fmoc-[Nle¹⁴]-11-17-OH peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

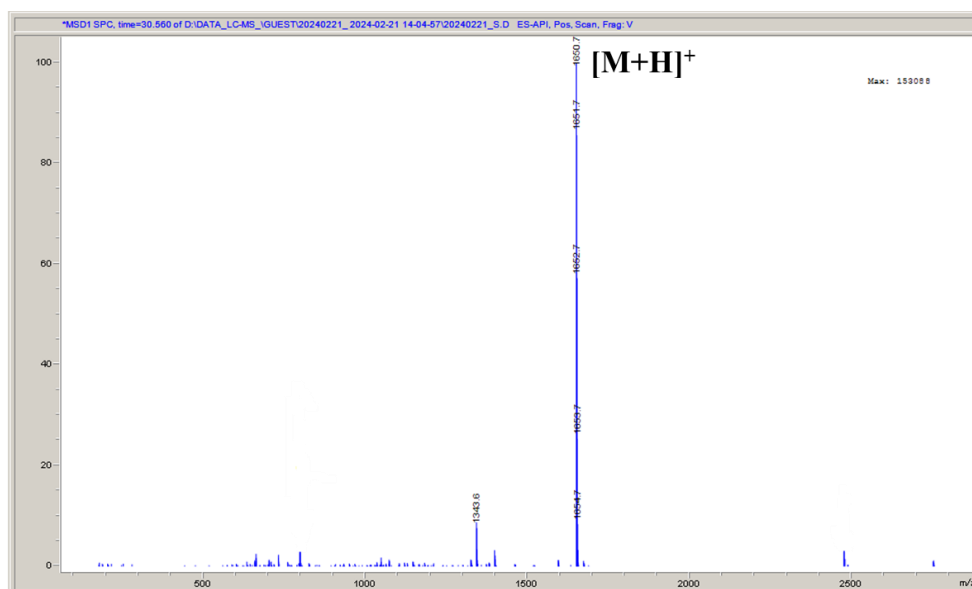


Figure 2.45. ESI-MS spectrum of Fmoc-[Nle¹⁴]-11-17-OH. Molecular formula: C₈₅H₁₁₁N₁₃O₁₇S₂, [M+H]¹⁺ Calculated: 1649.9 Da; Experimental mass: [M+H]⁺:1649.7 Da.



2.4.5. Synthesis of the Fmoc-[Ala¹¹, Nle¹⁴]-11-17-OH fragment

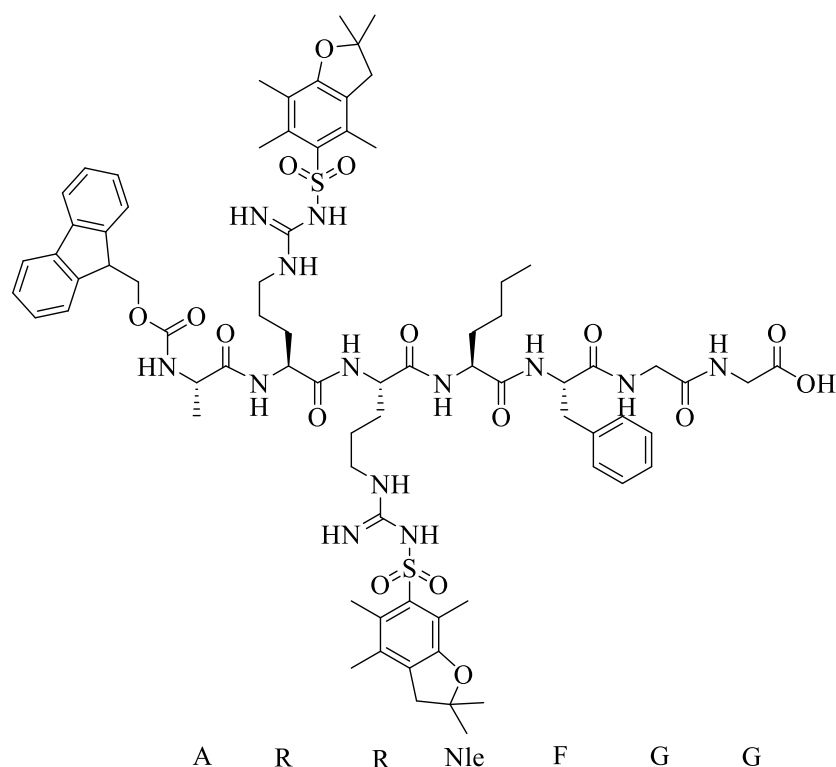


Figure 2.46. Structure of Fmoc-[Ala¹¹, Nle¹⁴]-11-17-OH peptide.

The Fmoc-[Ala¹¹, Nle¹⁴]-11-17-OH peptide (Figure 2.46) was synthesised manually using Fmoc-based SPPS. To ensure complete coupling efficiency and minimise the occurrence of truncated sequences, double coupling steps were performed for residues Arg¹² and Arg¹³. The synthesis proceeded smoothly with no evidence of significant side reactions or deletion sequences.

Upon completion of the elongation steps, the peptide was cleaved from the resin using a 30% HFPI solution in DCM. This mild cleavage procedure effectively released the protected peptide fragment while preserving the side chain protecting groups on all residues. This allowed the fragment to be used directly in the subsequent convergent coupling reaction without the need for further modification (Figure 2.47, Figure 2.48).

Results and Discussion

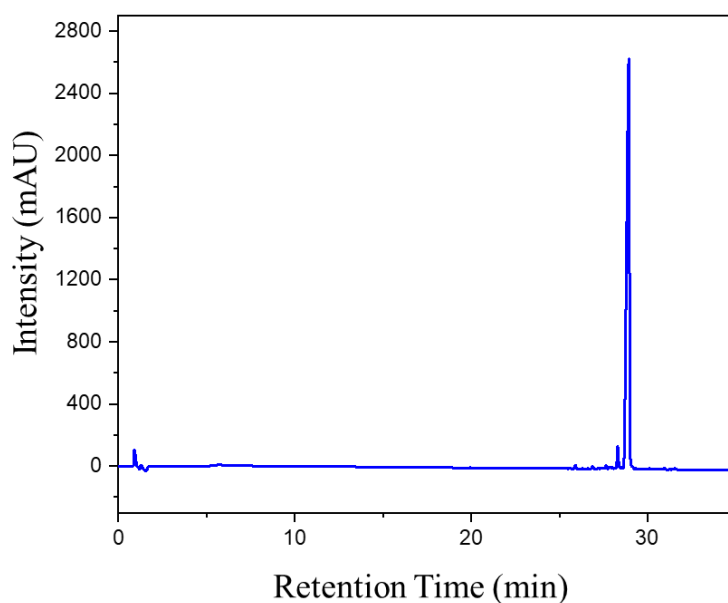


Figure 2.47. Analytical HPLC chromatogram of the Fmoc-[Ala^I, Nle^I]-11-17-OH peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

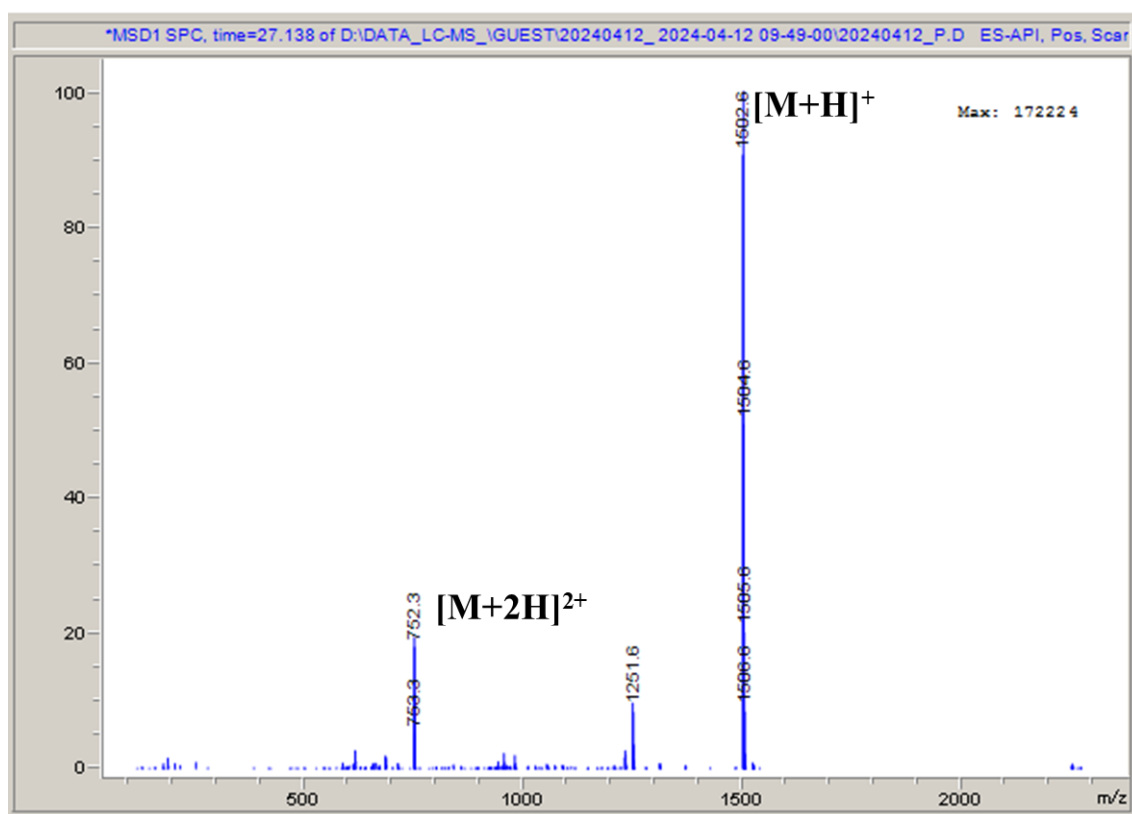


Figure 2.48. ESI-MS spectrum of Fmoc-[Ala^I, Nle^I]-11-17-OH. Molecular formula: C₇₅H₉₉N₁₃O₁₆S₂, [M+H]¹⁺ Calculated: 1501.7 Da; Experimental mass: [M+H]⁺: 1501.6 Da, [M+2H]²⁺: 751.8 Da.



2.4.6. Synthesis of the Fmoc-[Ala¹⁵, Nle¹⁴]-11-17-OH fragment

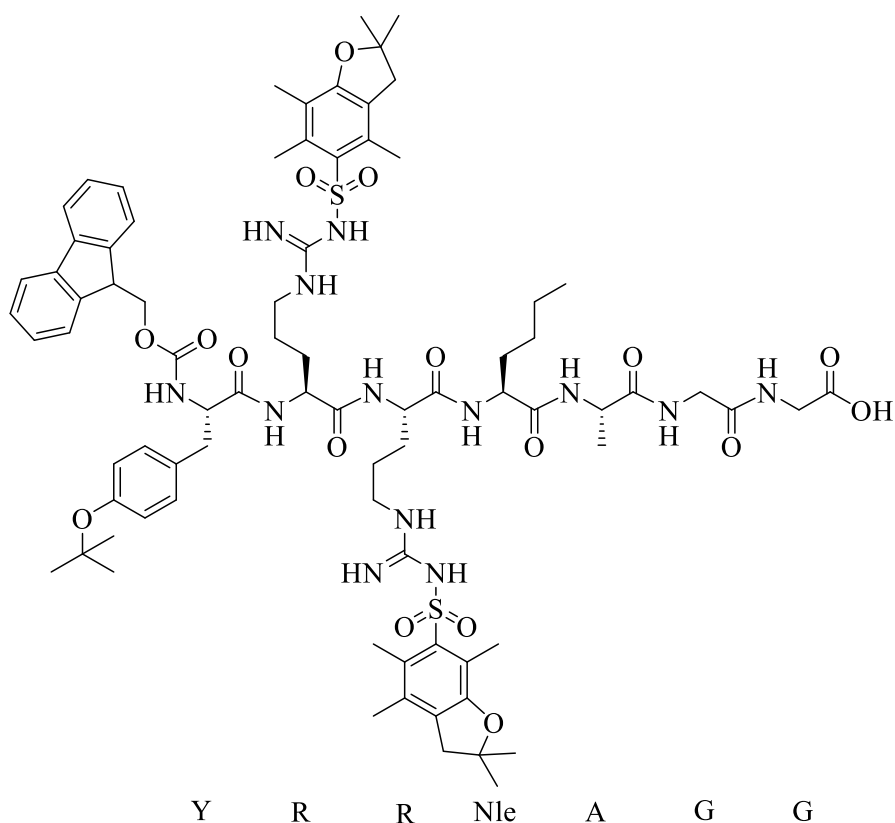


Figure 2.49. Structure of Fmoc-[Ala¹⁵, Nle¹⁴]-11-17-OH peptide.

The Fmoc-11-[Ala¹⁵, Nle¹⁴]-11-17-OH peptide (Figure 2.49) was synthesised manually using Fmoc-based SPPS. To ensure complete coupling efficiency and minimise the occurrence of truncated sequences, double coupling steps were performed for residues: Tyr¹¹, Ala¹², and Ala¹³. The synthesis proceeded smoothly with no evidence of significant side reactions or deletion sequences.

Upon completion of the elongation steps, the peptide was cleaved from the resin using a 30% HFPI solution in DCM. This mild cleavage procedure effectively released the protected peptide fragment while preserving the side chain protecting groups on all residues. This allowed the fragment to be used directly in the subsequent convergent coupling reaction without the need for further modification (Figure 2.50, Figure 2.51).

Results and Discussion

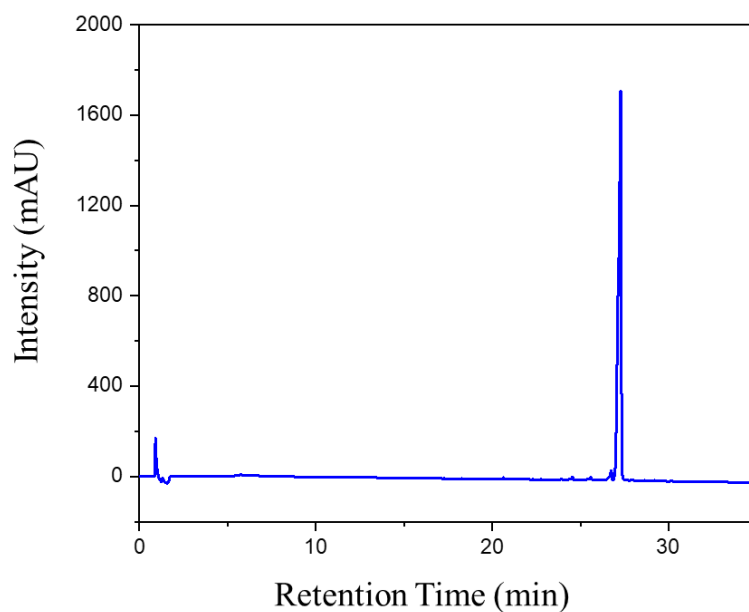


Figure 2.50. Analytical HPLC chromatogram of the Fmoc-[Ala¹⁵, Nle¹⁴]-11-17-OH peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

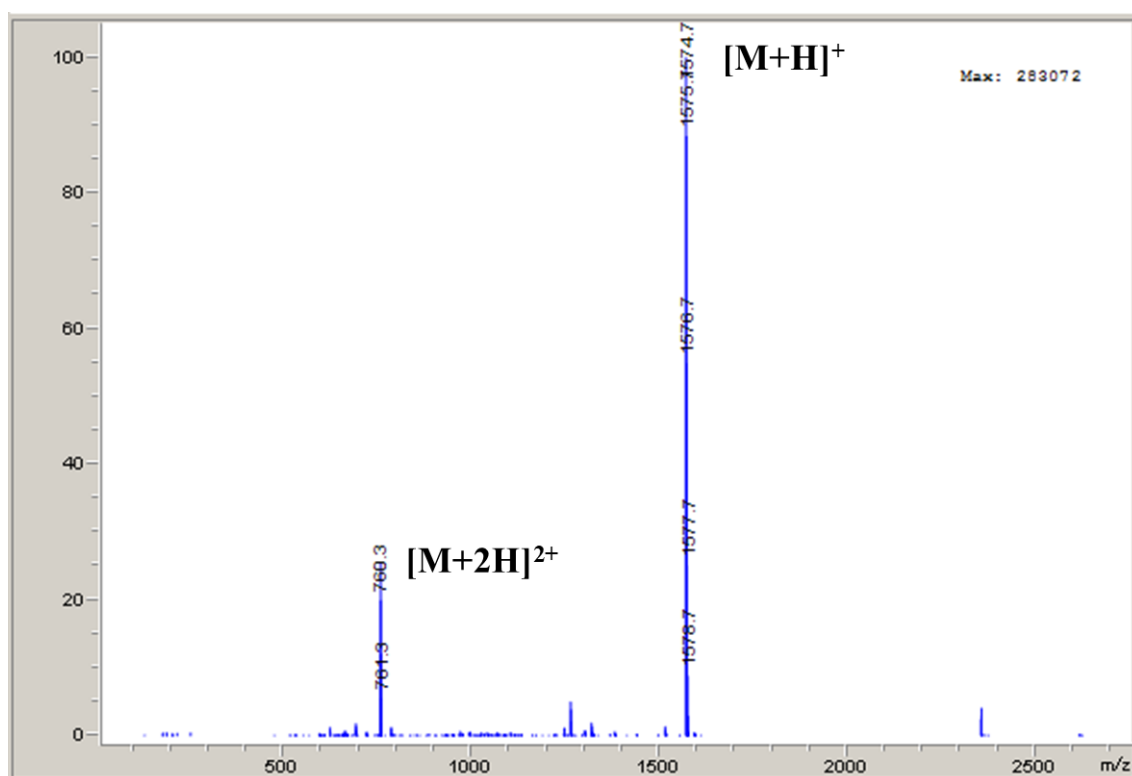


Figure 2.51. Figure 2.52. ESI-MS spectrum of Fmoc-[Ala¹⁵, Nle¹⁴]-11-17-OH. Molecular formula: C₇₉H₁₀₇N₁₃O₁₇S₂, [M+H]¹⁺ Calculated: 1554.9 Da; Experimental mass: [M+H]⁺:1574.7 Da, [M+2H]²⁺:1100.5 Da.



2.4.7. Synthesis of the Fmoc-[Ala^{12,13}, Nle¹⁴]-11-17-OH fragment

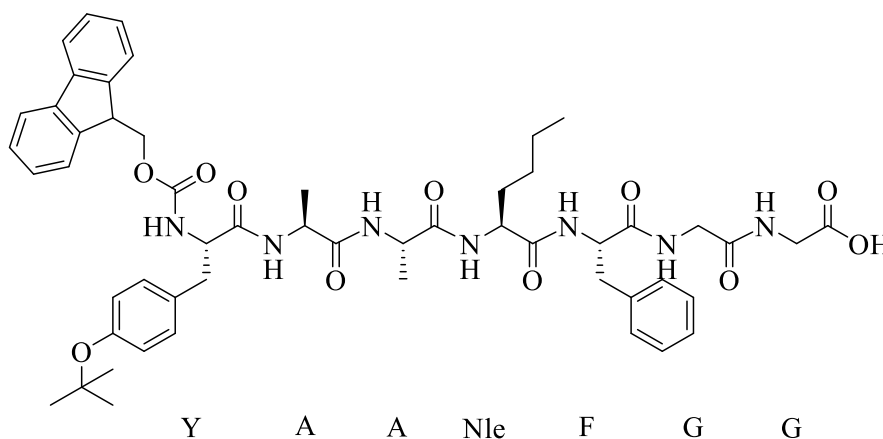


Figure 2.53. Structure of Fmoc-[Ala^{12,13}, Nle¹⁴]-11-17-OH peptide.

The Fmoc-[Ala^{12,13}, Nle¹⁴]-11-17-OH peptide (Figure 2.53) was synthesised manually using Fmoc-based SPPS. To ensure complete coupling efficiency and minimise the occurrence of truncated sequences, double coupling steps were performed for residues: Tyr¹¹, Ala¹², and Ala¹³. The synthesis proceeded smoothly with no evidence of significant side reactions or deletion sequences.

Upon completion of the elongation steps, the peptide was cleaved from the resin using a 30% HFPI solution in DCM. This mild cleavage procedure effectively released the protected peptide fragment while preserving the side chain protecting groups on all residues. This allowed the fragment to be used directly in the subsequent convergent coupling reaction without the need for further modification (Figure 2.54, Figure 2.55).

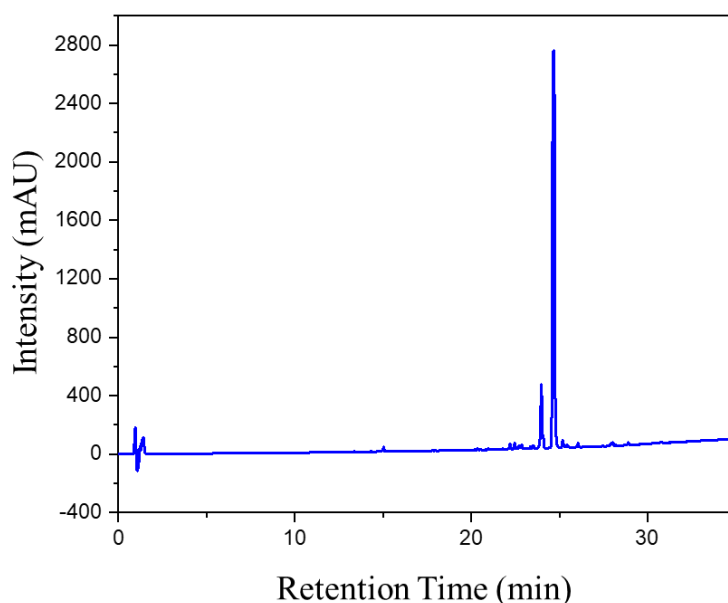


Figure 2.54. Analytical HPLC chromatogram of the Fmoc-[Ala^{12,13}, Nle¹⁴]-11-17-OH peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

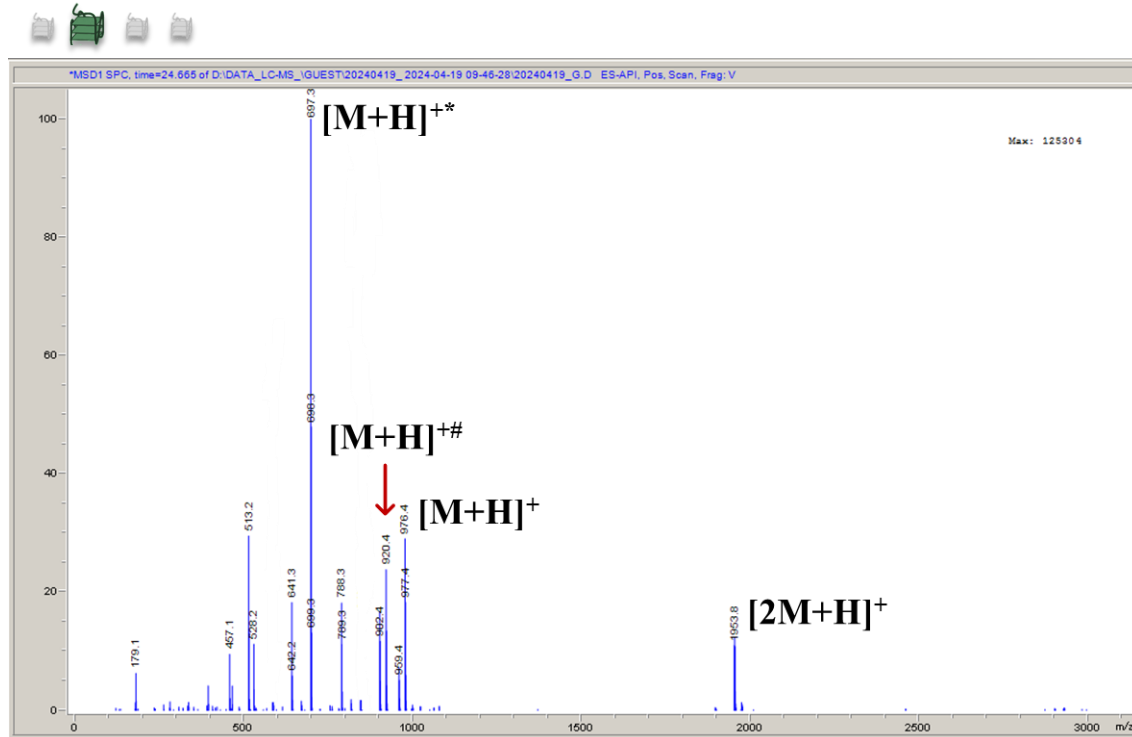


Figure 2.55. ESI-MS spectrum of Fmoc-[Ala^{12,13}, Nle¹⁴]-11-17-OH. In particular, the $[M+H]^+*$ signal represents the mass of the peptide after loss of the Boc protecting group, while $[M+H]^{+\#}$ signal represents the mass of the peptide after loss of both Fmoc and Boc protecting groups. Molecular formula: C₅₃H₆₅N₇O₁₁, $[M+H]^+$ Calculated: 976.5 Da; Experimental mass: $[2M+H]^+$: 1953.8 Da, $[M+H]^+$: 976.4 Da.

2.4.8. Convergent Synthesis of H-[Ala¹¹, Nle¹⁴]-11-30-NH₂ peptide

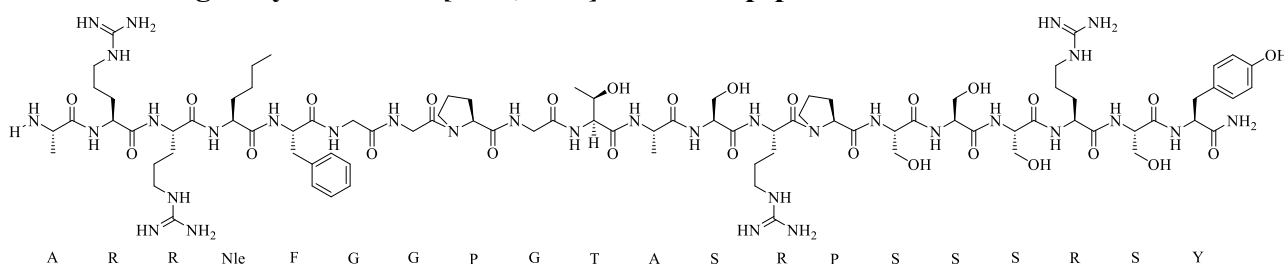


Figure 2.56. Structure of H-[Ala¹¹, Nle¹⁴]-11-30-NH₂ peptide.

The peptide H-[Ala¹¹, Nle¹⁴]-11-30-NH₂ (Figure 2.56) corresponds to the variant in which Ala was introduced in place of Tyr¹¹. It was obtained through the on-resin conjugation of two protected peptide fragments – Fmoc-[Ala¹¹, Nle¹⁴]-11-17-OH and Fmoc-18-30-NH₂ – using DIC/HOAt as the coupling agents. The condensation reaction was allowed to proceed for four hours under gentle agitation to ensure complete fragment coupling and minimise potential side reactions.

The synthesis proceeded smoothly with no significant side reactions or deletion sequences occurring. The peptide was then purified using preparative reverse-phase HPLC with a Luna C18 column (300 Å, 10 µm) and a binary solvent gradient ranging from 5% to 35% solvent B over 35 minutes. Fractions corresponding to the main product were collected and lyophilised. The identity and purity of the final compound were confirmed by analytical HPLC (Figure 2.57) and ESI-MS (Figure 2.58).

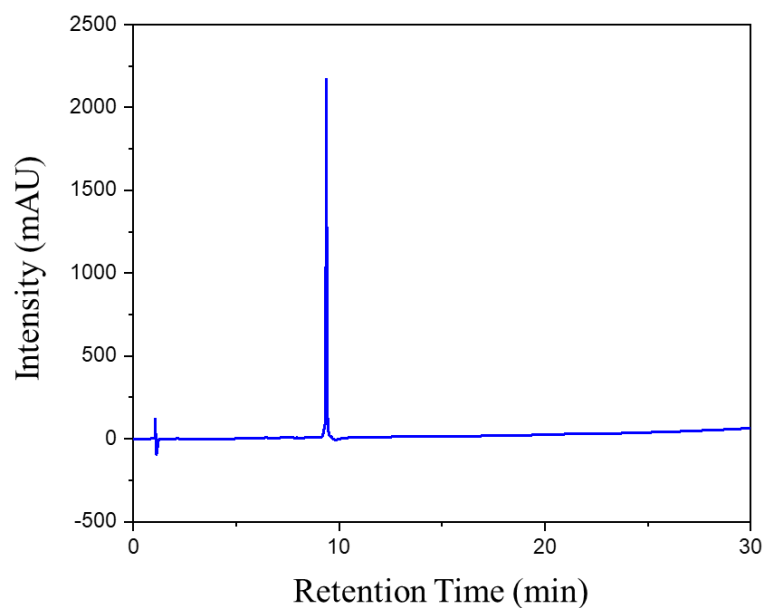


Figure 2.57. Analytical HPLC chromatogram of the H -[Ala¹¹, Nle¹⁴]-11-30-NH₂ peptide recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

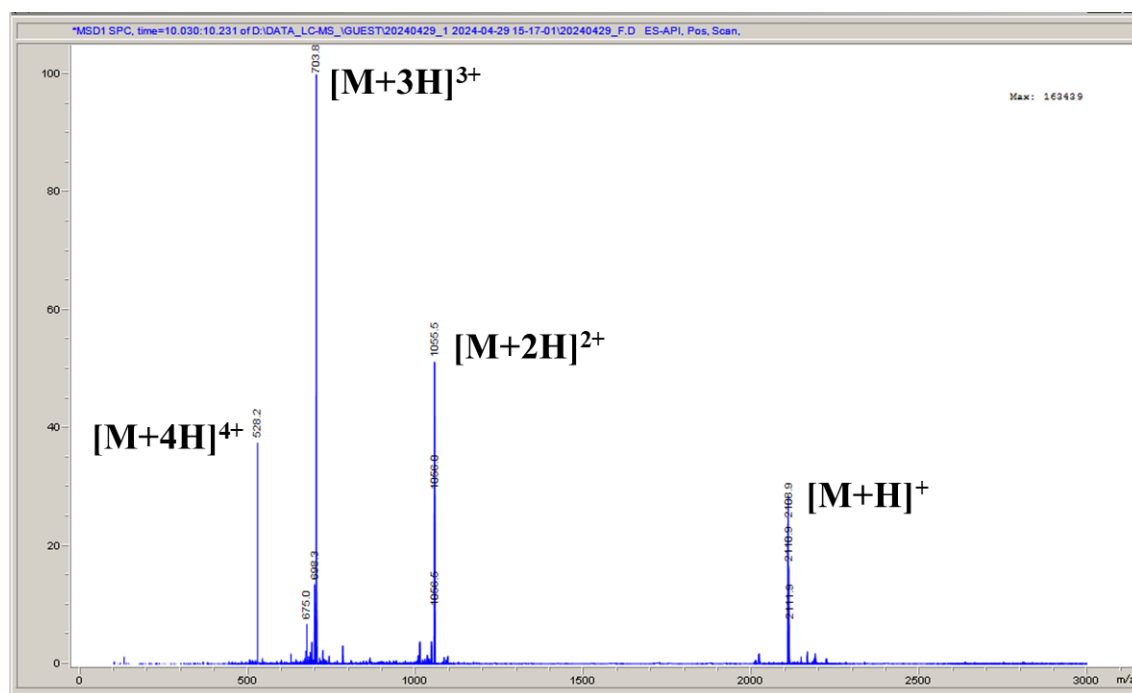


Figure 2.58. ESI-MS spectrum of H -[Ala¹¹, Nle¹⁴]-11-30-NH₂. Molecular formula: $C_{89}H_{145}N_{33}O_{27}$, $[M+H]^+$ Calculated: 2108.1 Da; Experimental mass: $[M+H]^+$: 2108.9 Da, $[M+2H]^2+$: 1055.0 Da, $[M+3H]^3+$: 703.5 Da, $[M+4H]^4+$: 528.0 Da.

Chemical structure of the 100-residue protein, showing the sequence: Y R R Nle A G G P G T A S R P S S S R S Y. The structure includes various side chains, including a hydroxyl group (HO), a guanidino group (HN=NH), and a hydroxyl group (OH).

Figure 2.59. Structure of H-[Ala¹⁵, Nle¹⁴]-11-30-NH₂ peptide.

The synthesis proceeded smoothly with no significant side reactions or deletion sequences occurring. The sequence was then purified using preparative reverse-phase HPLC with a Luna C18 column (300 Å, 10 µm) and a binary solvent gradient ranging from 5% to 35% solvent B over 35 minutes. Fractions corresponding to the main product were collected and lyophilised. The identity and purity of the final compound were confirmed by analytical HPLC (Figure 2.60) and ESI-MS (Figure 2.61).

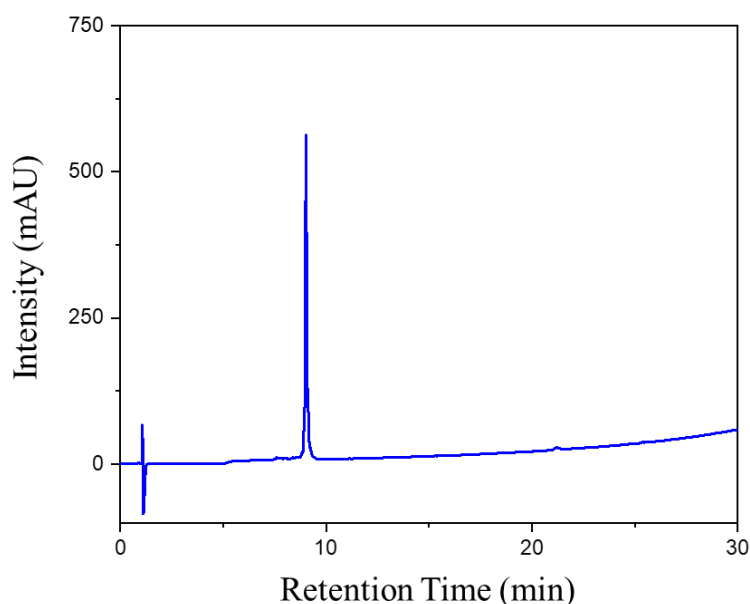


Figure 2.60. Analytical HPLC chromatogram of the H-[Ala¹⁵, Nle¹⁴]-11-30-NH₂ peptide recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

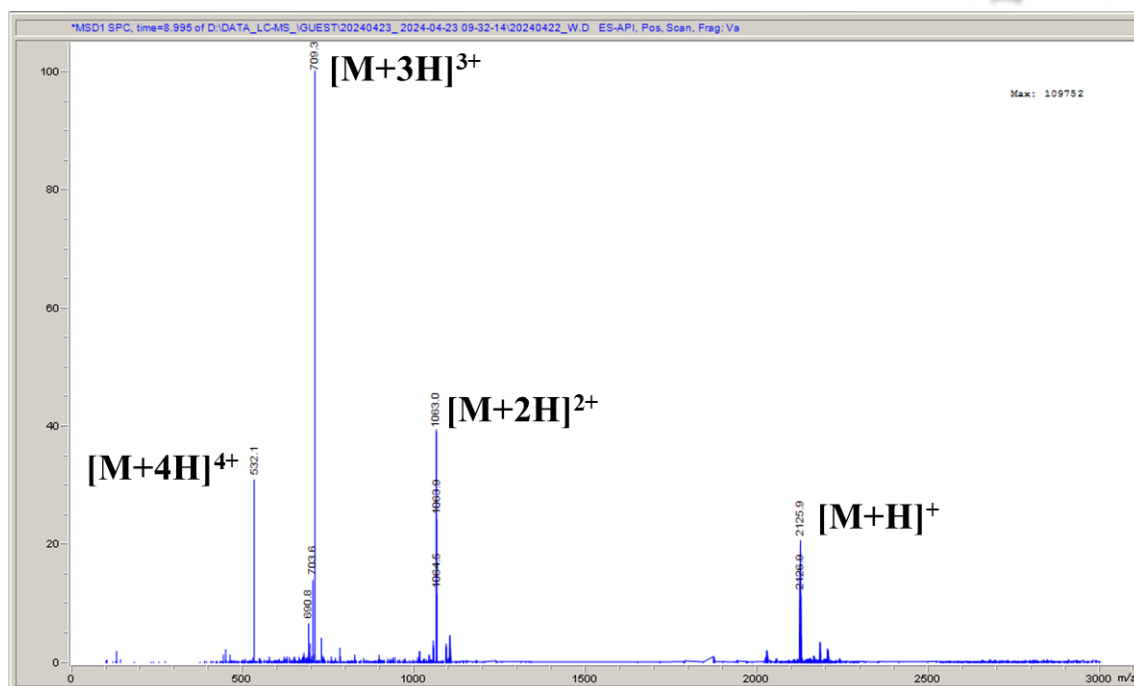


Figure 2.61. ESI-MS spectrum of H-[Ala¹⁵, Nle¹⁴]-11-30-NH₂ peptide. Molecular formula: C₈₉H₁₄₅N₃₃O₂₈, [M+H]¹⁺ Calculated: 2125.2 Da; Experimental mass: [M+H]¹⁺: 2124.9 Da, [M+2H]²⁺: 1062.5 Da, [M+3H]³⁺: 709.0 Da, [M+4H]⁴⁺: 531.9 Da.

2.4.10. Convergent Synthesis of H-[Ala³⁰, Nle¹⁴]-11-30-NH₂ peptide

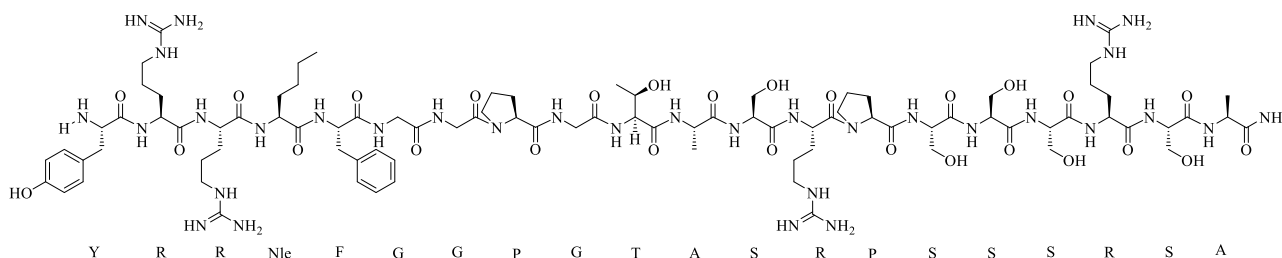


Figure 2.62. Structure of H-[Ala³⁰, Nle¹⁴]-11-30-NH₂ peptide.

The peptide H-[Ala³⁰, Nle¹⁴]-11-30-NH₂ (Figure 2.62) correspond to the variant in which Ala was introduced in place of Tyr³⁰ residue. It was obtained through the on-resin conjugation of two protected peptide fragments – Fmoc-[Nle¹⁴]-11-17-OH and H-[Ala³⁰]-18-30-NH₂ – using DIC/HOAt as the coupling agents. The condensation reaction was allowed to proceed for four hours under gentle agitation to ensure complete fragment coupling and minimise potential side reactions.

The synthesis proceeded smoothly with no significant side reactions or deletion sequences occurring. The sequence was then purified using preparative reverse-phase HPLC with a Luna C18 column (300 Å, 10 µm) and a binary solvent gradient ranging from 5% to 35% solvent B over 35 minutes. Fractions corresponding to the main product were collected and lyophilised. The identity and purity of the final compound were confirmed by analytical HPLC (Figure 2.63) and ESI-MS (Figure 2.64).

Results and Discussion

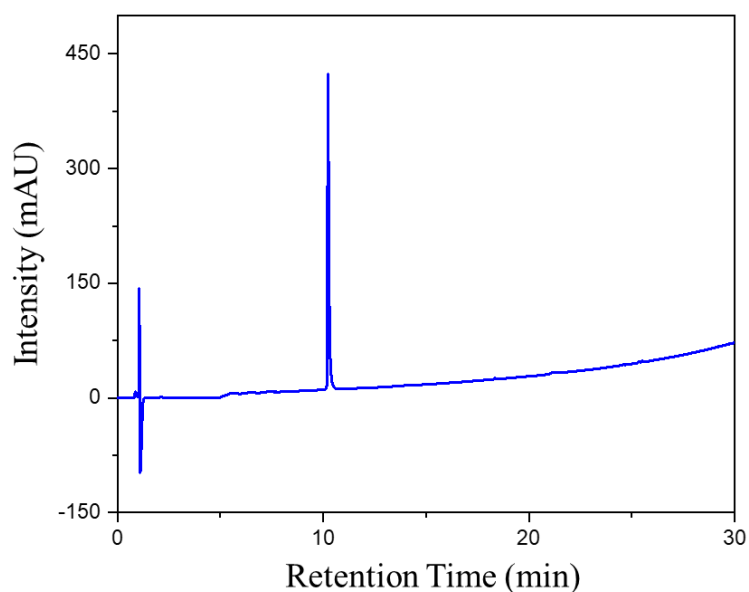


Figure 2.63. Analytical HPLC chromatogram of the H-[Ala³⁰, Nle¹⁴]-11-30-NH₂ peptide recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

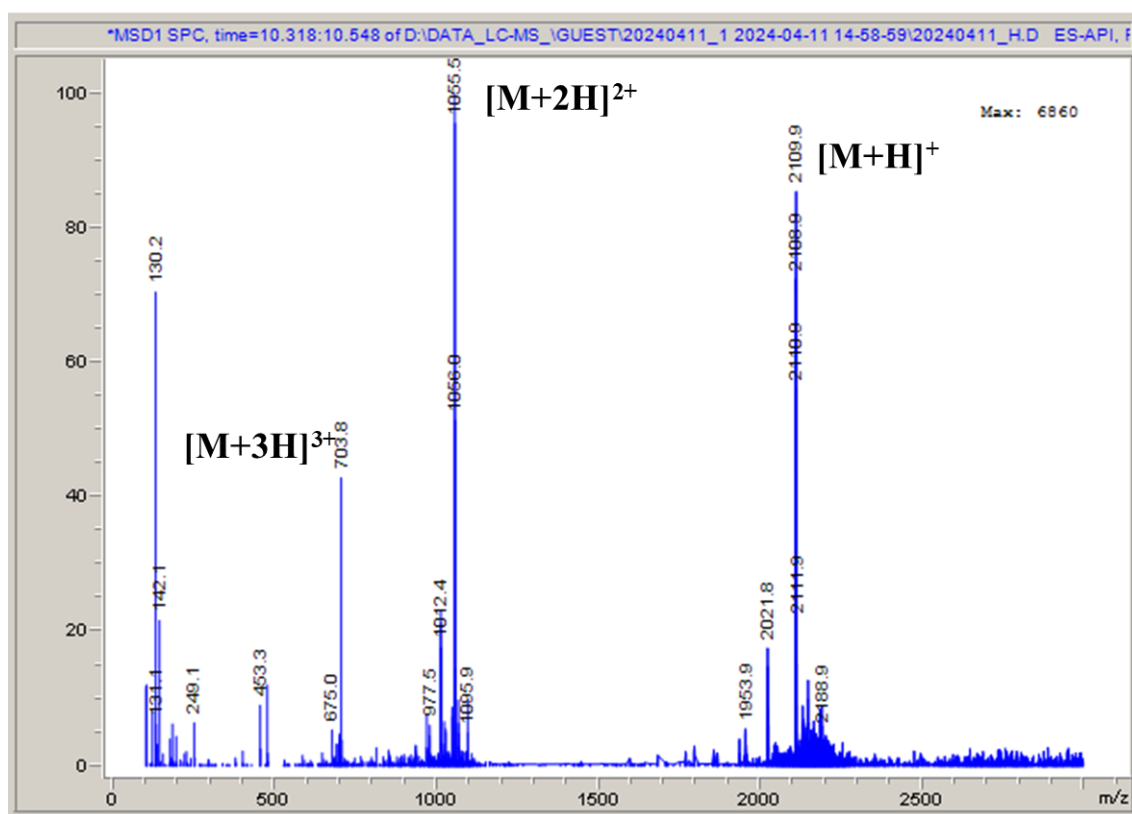


Figure 2.64. ESI-MS spectrum of H-[Ala³⁰, Nle¹⁴]-11-30-NH₂ peptide. Molecular formula: C₈₉H₁₄₅N₃₃O₂₇, [M+H]¹⁺ Calculated: 2109.3 Da; Experimental mass: [M+H]¹⁺:2108.9 Da, [M+2H]²⁺:1055.0 Da, [M+3H]³⁺:703.5 Da.



2.4.11. Convergent Synthesis of H-[Ala^{12,13,23,28}, Nle¹⁴]-11-30-NH₂ peptide

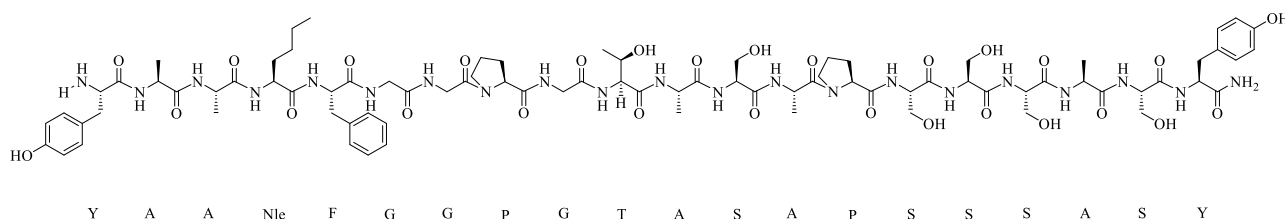


Figure 2.65. Structure of H-[Ala^{12,13,23,28}, Nle¹⁴]-11-30-NH₂ peptide.

The peptide H-[Ala^{12,13,23,28}, Nle¹⁴]-11-30-NH₂ (Figure 2.65) corresponds to the variant in which Ala was introduced in place of Arg¹², Arg¹³, Arg²³ and Arg²⁸ residues. It was obtained through the on-resin conjugation of two protected peptide fragments – Fmoc-[Ala^{12,13}, Nle¹⁴]-11-17-OH and H-[Ala^{23,28}]-18-30-NH₂ – using DIC/HOAt as the coupling agents. The condensation reaction was allowed to proceed for four hours under gentle agitation to ensure complete fragment coupling and minimise potential side reactions.

The synthesis proceeded smoothly with no significant side reactions or deletion sequences occurring. The sequence was then purified using preparative reverse-phase HPLC with a Luna C18 column (300 Å, 10 µm) and a binary solvent gradient ranging from 5% to 35% solvent B over 35 minutes. Fractions corresponding to the main product were collected and lyophilised. The identity and purity of the final compound were confirmed by analytical HPLC (Figure 2.66) and ESI-MS (Figure 2.67).

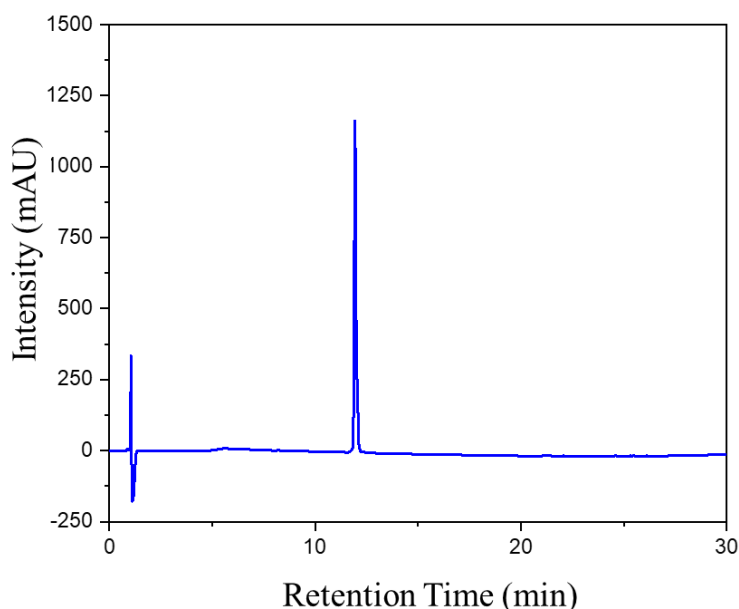


Figure 2.66. Analytical HPLC chromatogram of the H-[Ala^{12,13,23,28}, Nle¹⁴]-11-30-NH₂ peptide recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

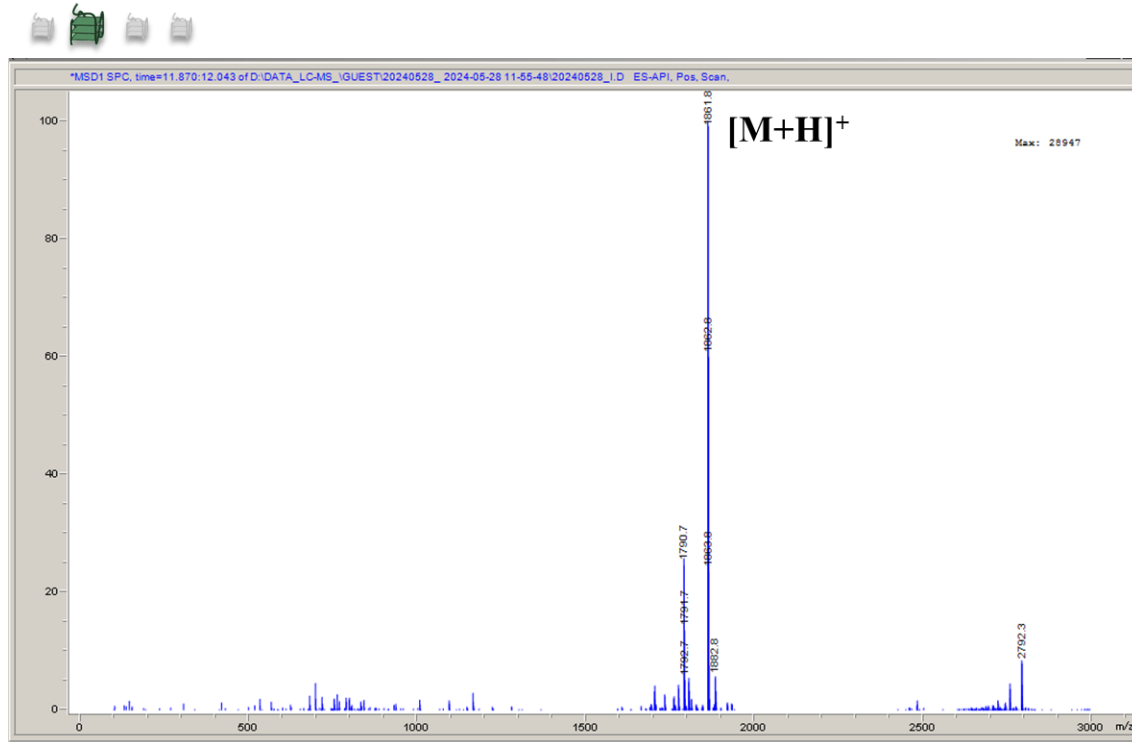


Figure 2.67. ESI-MS spectrum of H-[Ala^{12,13,23,28}, Nle¹⁴]-11-30-NH₂ peptide. Molecular formula: C₈₃H₁₂₁N₂₁O₂₈, [M+H]¹⁺ Calculated: 1861.0 Da; Experimental mass: [M+H]⁺:1860.8.

2.4.12. Convergent Synthesis of H-[Ala^{12,13}, Nle¹⁴]-11-30-NH₂ peptide

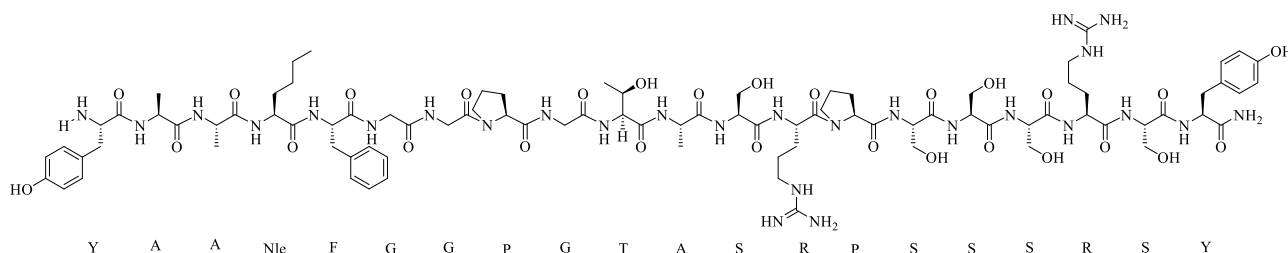


Figure 2.68. Structure of H-[Ala^{12,13}, Nle¹⁴]-11-30-NH₂ peptide.

The peptide H-[Ala^{12,13}, Nle¹⁴]-11-30-NH₂ (Figure 2.68) corresponds to the variant in which Ala was introduced in place of Arg¹² and Arg¹³ residues. It was obtained through the on-resin conjugation of two protected peptide fragments – Fmoc-[Ala^{12,13}, Nle¹⁴]-11-17-OH and H-18-30-NH₂ – using DIC/HOAt as the coupling agents. The condensation reaction was allowed to proceed for four hours under gentle agitation to ensure complete fragment coupling and minimise potential side reactions. The synthesis proceeded smoothly with no significant side reactions or deletion sequences occurring. The sequence was then purified using preparative reverse-phase HPLC with a Luna C18 column (300 Å, 10 µm) and a binary solvent gradient ranging from 5% to 35% solvent B over 35 minutes. Fractions corresponding to the main product were collected and lyophilised. The identity and purity of the final compound were confirmed by analytical HPLC (Figure 2.69) and ESI-MS (Figure 2.70).

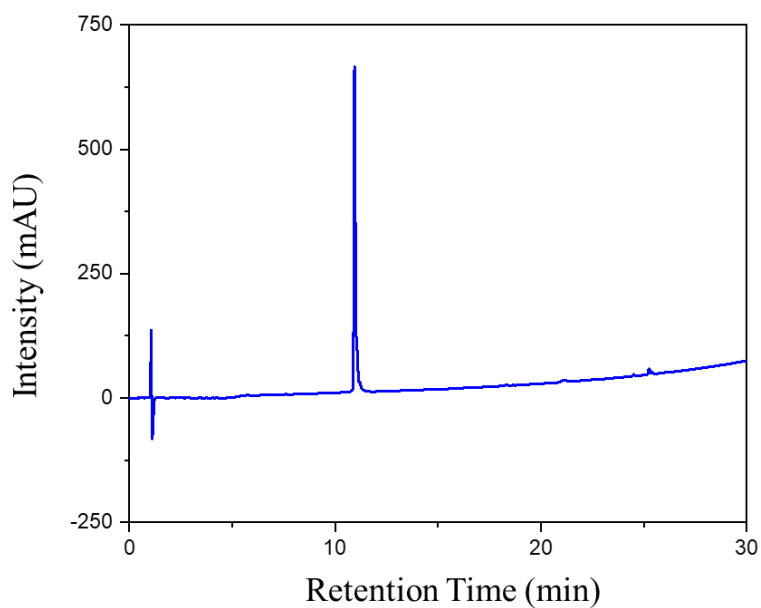


Figure 2.69. Analytical HPLC chromatogram of the H -[Ala^{12,13}, Nle¹⁴]-11-30-NH₂ peptide recorded using a Kinetex C18 column, 1 mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

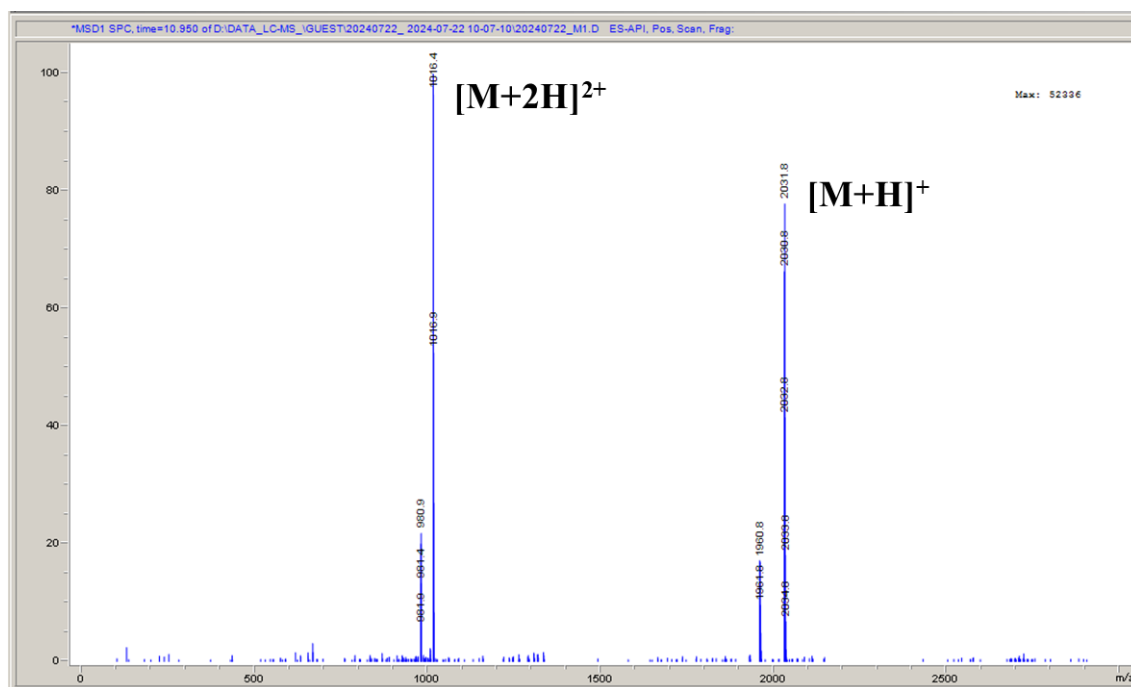


Figure 2.70. ESI-MS spectrum of H -[Ala^{12,13}, Nle¹⁴]-11-30-NH₂ peptide. Molecular formula: $C_{89}H_{135}N_{27}O_{28}$, $[M+H]^+$ Calculated: 2031.2 Da; Experimental mass: $[M+H]^+$: 2030.8 Da, $[M+2H]^{2+}$: 1015.9 Da.



2.4.13. Convergent Synthesis of H-[Ala^{23,28}, Nle¹⁴]-11-30-NH₂ peptide

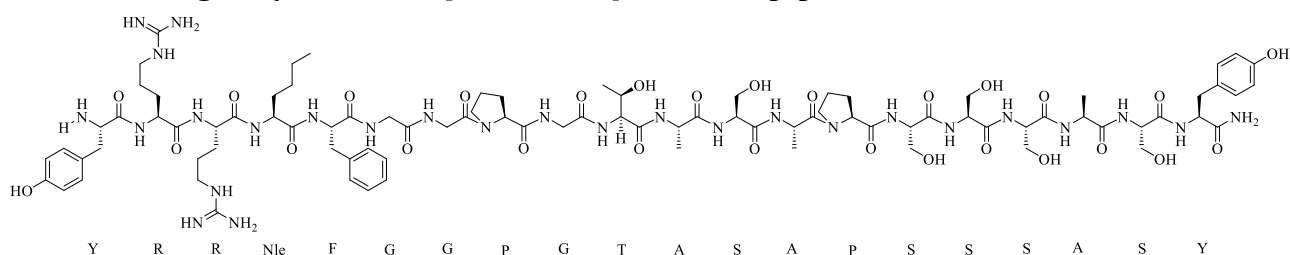


Figure 2.71. Structure of H-[Ala^{23,28}, Nle¹⁴]-11-30-NH₂ peptide.

The peptide H-[Ala^{23,28}, Nle¹⁴]-11-30-NH₂ (Figure 2.71) corresponds to the variant in which Ala was introduced in place of Arg²³ and Arg²⁸ residues. It was obtained through the on-resin conjugation of two protected peptide fragments – Fmoc-[Nle]¹⁴-11-17-OH and H-[Ala^{23,28}]-18-30-NH₂ – using DIC/HOAt as the coupling agents. The condensation reaction was allowed to proceed for four hours under gentle agitation to ensure complete fragment coupling and minimise potential side reactions. The synthesis proceeded smoothly with no significant side reactions or deletion sequences occurring. The peptide was then purified using preparative reverse-phase HPLC with a Luna C18 column (300 Å, 10 µm) and a binary solvent gradient ranging from 5% to 35% solvent B over 35 minutes. Fractions corresponding to the main product were collected and lyophilised. The identity and purity of the final compound were confirmed by analytical HPLC (Figure 2.72) and ESI-MS (Figure 2.73).

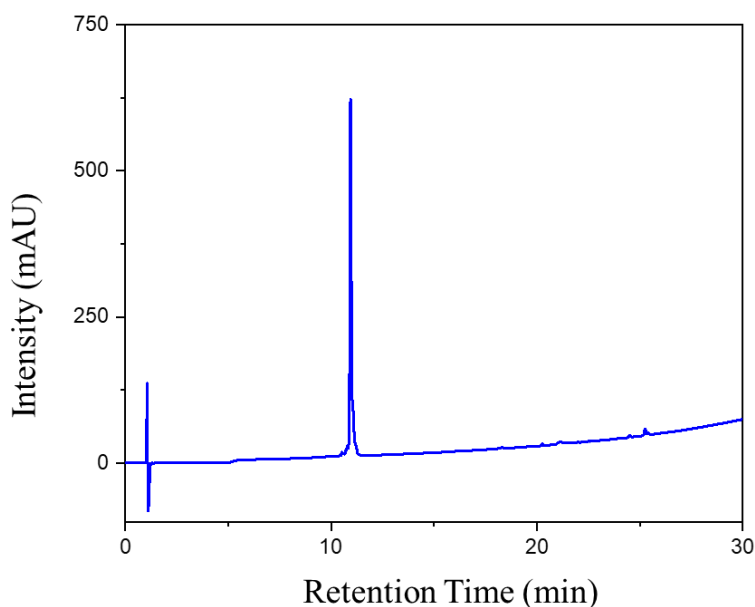


Figure 2.72. Analytical HPLC chromatogram of the H-[Ala^{23,28}, Nle¹⁴]-11-30-NH₂ peptide recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

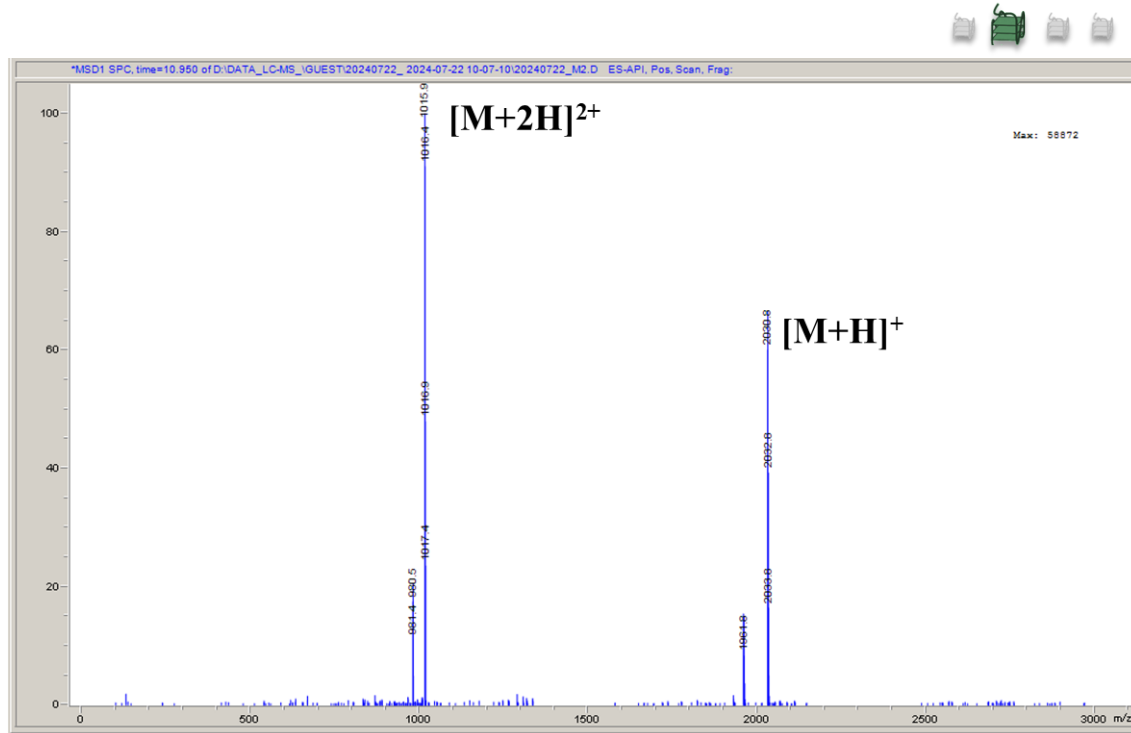


Figure 2.73. ESI-MS spectrum of H-[Ala^{23,28}, Nle¹⁴]-11-30-NH₂ peptide. Molecular formula: C₈₉H₁₃₅N₂₇O₂₈, $[M+H]^+$ Calculated: 2031.2 Da; Experimental mass: $[M+H]^+$: 2030.8 Da, $[M+2H]^{2+}$: 1015.9 Da.



2.5. Circular Dichroism Characterization of Vimentin-derived Peptides

The secondary structure tendencies of the synthesised Vimentin-derived peptides (H-11-30-NH₂, Ac-11-30-NH₂, H-68-79-NH₂, Ac-68-79-NH₂, H-270-280-NH₂, Ac-270-280-NH₂) were investigated using CD spectroscopy under different solvent conditions. CD spectra were collected in TRIS buffer, in a 1:1 TRIS:TFE mixture and in pure TFE. Increasing the TFE content gradually provides a means of evaluating the ability of the peptides to adopt ordered conformations. TFE allows to stabilise helical conformations although structures containing turns and β -sheets are also possible in its presence. Plausible mechanisms by which TFE affects peptide structure include enhancing internal hydrogen bonding, disrupting the structure of water, reducing the hydrophobic effect, penetrating cosolvent molecules into the peptide core, and preferentially solvating certain peptide chain groups.[66]

In aqueous buffer, all peptides exhibited a single strong negative band near 200 nm. This indicates their limited ability to adopt ordered secondary conformations under these conditions. However, when transferred to a membrane-mimetic environment, subtle but significant differences among the peptide regions became evident, reflecting their tendencies to form specific structural elements.

2.5.1. Peptide 11-30 (H-11-30-NH₂ and Ac-11-30-NH₂)

Both peptides, which correspond to the 11-30 region of Vimentin, exhibited a significant increase in ellipticity, as well as the formation of two negative bands at approximately 208 and 222 nm, when analysed in a TFE environment (Figure 2.74). This indicates the partial induction of an ordered secondary structure. In the TRIS:TFE mixture, the spectra displayed broad minima with feature indicative of mixed conformations. In pure TFE, this effect became more pronounced. The acetylated variant exhibited slightly increased overall ellipticity, indicating that N-terminal capping reduces electrostatic repulsion and stabilises local hydrogen bonding.

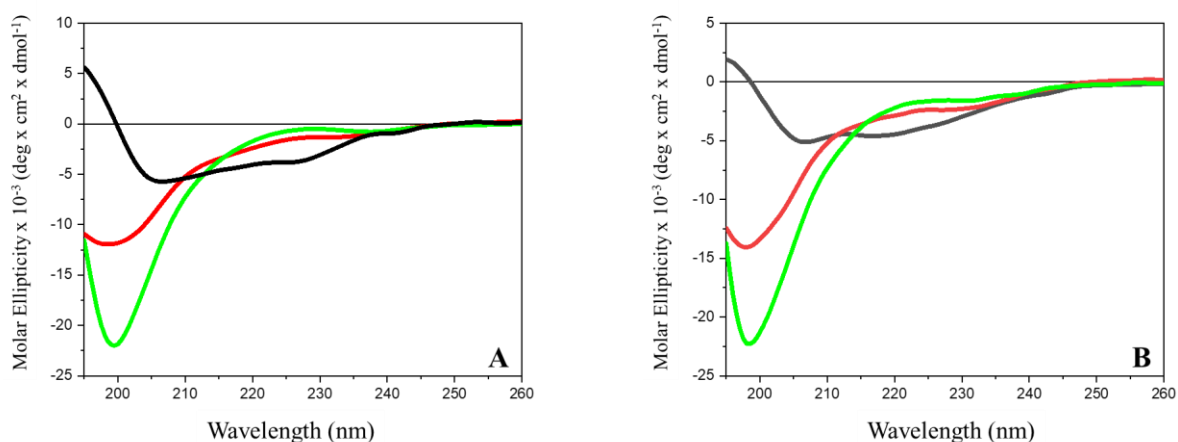


Figure 2.74. CD spectra of **A)** Ac-11-30-NH₂ and **B)** H-11-30-NH₂ show the spectrum of the peptide in TRIS buffer, 10mM, pH=8.4 (in green), in TRIS:TFE=50:50 mixture (in red) and in TFE (in black).



To further investigate this aggregation phenomenon, additional CD spectra of Ac-11-30-NH₂ were recorded at various peptide concentrations in TFE (Figure 2.75). While the overall spectral shape remained similar, a marked dependence on concentration was observed. At higher concentrations the spectrum displayed two strong negative maxima with a $\theta_{222}/\theta_{208}$ ratio greater than 1, indicating that the peptide aggregates in the hydrophobic medium.[67, 68] Upon dilution, the overall ellipticity decreased and the 222 nm increased in relative intensity. This shift indicates partial disaggregation and the recovery of more intramolecularly folded species.

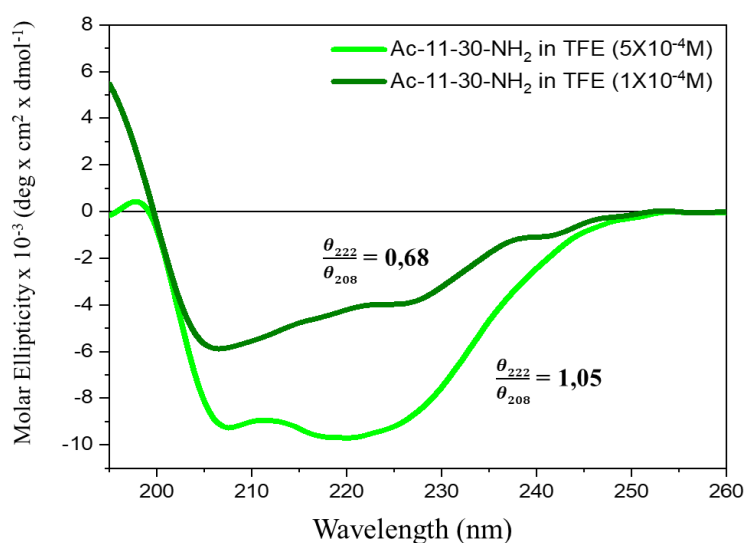


Figure 2.75. CD spectra of Ac-11-30-NH₂ at different concentration in TFE solvent.

These results demonstrate that the structured features observed at high concentrations mainly arise from self-association phenomena.

2.5.2. Peptides 68-79 (H-68-79-NH₂ and Ac-68-79-NH₂)

The 68-79 region exhibit the most pronounced conformational transition among the analysed fragments (Figure 2.76). In TRIS buffer, both variants displayed a broad negative band centred around 200 nm, indicative of a predominantly disordered structure. As the TFE content increased, the spectra evolved to show two distinct negative bands at around 208 and 222 nm, which is consistent with the partial induction of α -helical or turn-like conformations. While the characteristic positive band at around 190 nm, which is typical of a fully formed α -helix, was not observed, the shift in ellipticity and the emergence of the double minimum clearly demonstrate a tendency towards local secondary structure organisation rather than complete random coil behaviour.

As expected, this effect was more evident in pure TFE, particularly for the acetylated peptide, which exhibited a more intense and better-defined band at 222 nm. The improved stability of this variant is

Results and Discussion



likely due to N-terminal acetylation, which reduces electrostatic repulsion and promotes intramolecular hydrogen bonding.

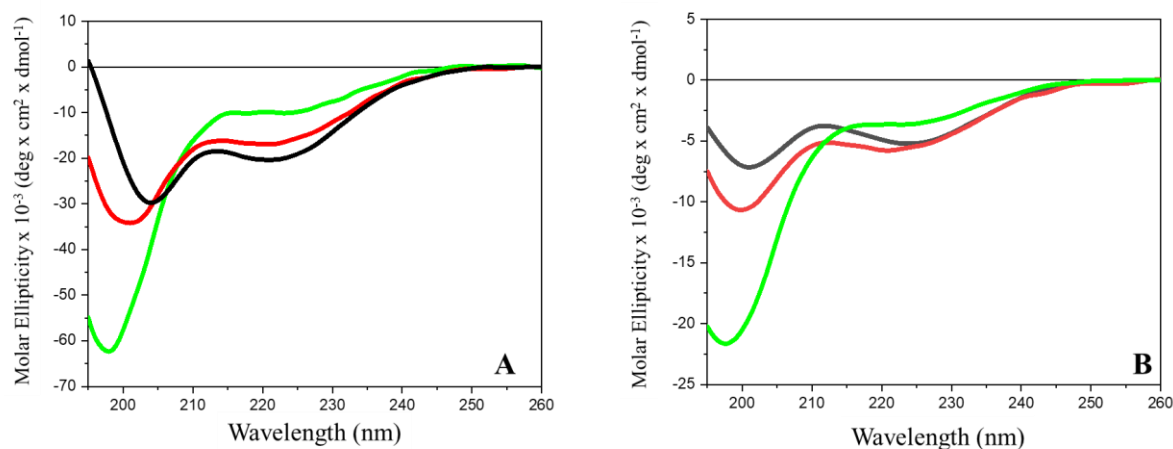


Figure 2.76. CD spectra of **A)** Ac-68-79-NH₂ and **B)** H-68-79-NH₂ show the spectrum of the peptide in TRIS buffer; 10mM, pH=8.4 (in green), in TRIS:TFE=50:50 mixture (in red) and in TFE (in black).

2.5.3. Peptides 270-280 (H-270-280-NH₂ and Ac-270-280-NH₂)

CD spectroscopy revealed that the conformational preferences of the 270-280 Vimentin fragments are influenced by the N-terminus modification (Figure 2.76).

For the non-acetylated peptide, an increasing TFE concentration resulted in the gradual emergence of two negative minima at 208 and 222 nm, which are characteristic of α -helical folding. This partial helical stabilization suggests that free N-terminal amine may contribute to intramolecular hydrogen bonding through favourable electrostatic interactions.

In contrast, the acetylated form displayed different spectral evolution. Upon TFE addition, a broad negative maximum emerged near 215 nm and a positive band appear around 195 nm, consistent with the formation of β -sheet or β -turn-like arrangements. This behaviour likely reflects the propensity of the acetylated variant to undergo intermolecular association and β -aggregation, since the neutralisation of the terminal charge promotes backbone hydrogen bonding between peptides chains.

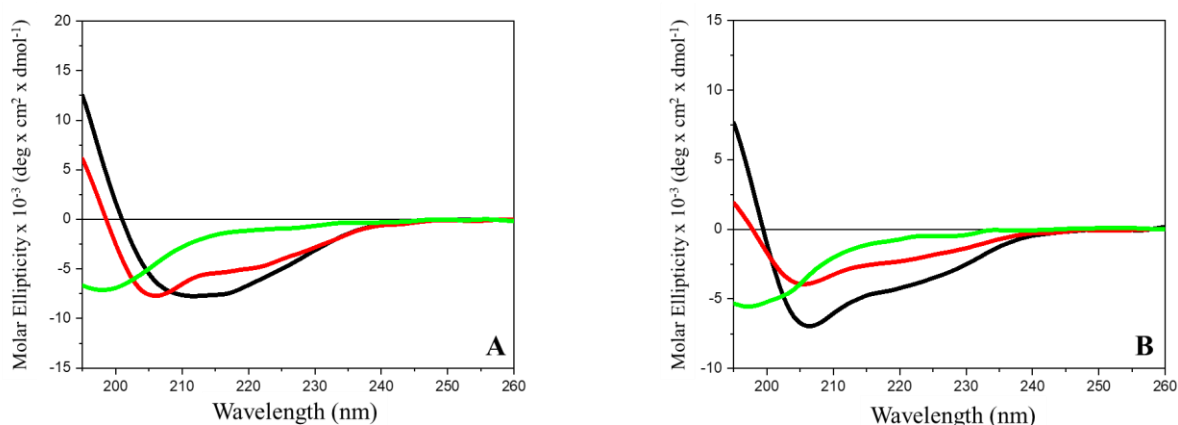


Figure 2.77. CD spectra of **A)** Ac-270-280-NH₂ and **B)** H-270-280-NH₂ show the spectrum of the peptide in TRIS buffer, 10mM, pH=8.4 (in green), in TRIS:TFE=50:50 mixture (in red) and in TFE (in black).

In aqueous buffer solution, all of the Vimentin-derived peptide fragments exhibited a single, broad negative band centred around 200 nm. This indicates an intrinsically disordered conformation. This behaviour reflects the limited propensity of these short regions to adopt stable secondary structures under physiologically aqueous conditions. Although subtle differences among the sequences became apparent when the peptides were analysed in membrane-mimicking environments, such as TFE, where certain fragments showed enhanced α -helical or β -type structural tendencies, these conformational preferences only emerged under strongly stabilising conditions.

For the purpose of the interaction studies presented in the following sections, it is important to note that under aqueous buffer conditions, all peptides remain predominantly disordered. This ensures that the structural state characterised here corresponds to the conformational ensemble in which peptide-DNA and peptide-G4 interactions were evaluated.

2.6. NMR Spectroscopy Characterization of Vimentin Fragment

2D NMR spectroscopy is a powerful analytical technique for investigating the structural and conformational properties of biomolecules. In this study, 2D NMR experiments -Total Correlation Spectroscopy (TOCSY) and Nuclear Overhauser Effect Spectroscopy (NOESY)- were employed to elucidate both-through bond and through-space connectivities within the synthesised peptides.

TOCSY experiments provide information on scalar (through-bond) spin-spin couplings, enabling complete spin systems corresponding to individual amino acid residues to be identified. In a TOCSY spectrum, each amino acid residue appears as a distinct spin system, isolated from neighbouring residues by the intervening carbonyl group. The spectrum displays cross-peaks between not only NH and CH α protons, but also all protons linked by a continuous chain of scalar couplings within the same residue.[69]

Results and Discussion



However, protons from N-terminal amine often undergo rapid solvent exchange and may not be visible in this region.

In contrast, NOESY experiments reveal dipolar interactions through space between spatially close protons, typically within 5 Å, based on the nuclear Overhauser effect (NOE), which reflects the transfer of nuclear spin polarization between neighbouring nuclei. The presence and intensity of NOE-cross peaks provide insights into the three-dimensional folding of the peptide and can be used to identify secondary structural motifs, such as turns, helices or β -contacts. The NH-CH α region (7-9 ppm for NH and 3-5 ppm for CH α) is particularly informative and serves as a fingerprint region for peptide sequence analysis. Unlike TOCSY, NOESY cross-peaks depend not on scalar couplings, but on the proximity of protons in space. This makes NOESY an essential tool for conformational analysis in solution.[70]

Together, TOCSY and NOESY experiments offer complementary information: the former defines through-bond connectivity and establishes residue-specific spin systems, while the latter highlights spatial correlations that indicate the folding and compactness of the peptide.

In this study, NMR spectroscopy was used to characterise the acetylated peptide of the vimentin N-terminal region. These fragments are the most interesting for internalisation studies, with one acting as a positive control and the other as a negative control.

2.6.1. NMR Analysis of Ac-11-30-NH₂ Peptide

The ¹H-NMR spectrum of Ac-11-30-NH₂ peptide shows features that are consistent with partial aggregation in TFE solution. In the amide region (7.5-8.5 ppm), broad, poorly resolved resonances reflecting conformational heterogeneity and intermolecular self-association are observed. These are both typical of peptides with a tendency to aggregate.

Despite the intrinsic broadening, two-dimensional NMR experiments (TOCSY and NOESY) enabled individual spin systems to be identified and assigned, providing an insight into the peptide conformational preferences. In TFE, the spectra exhibited the characteristic sequential NH-NH and NH-CH α cross-peaks that are diagnostic of an α -helical structures. These findings confirm and complement the CD spectroscopy, demonstrating that the Ac-11-30-NH₂ peptide possesses an intrinsic α -helical propensity in helix-promoting conditions while retaining a certain degree of aggregation.

The one- and two-dimensional NMR spectra supporting these observations are reported in Appendix.



2.6.2. NMR Analysis of Ac-68-79-NH₂ Peptide

The ¹H-NMR spectrum of Ac-68-79-NH₂ peptide shows clearly defined resonances, suggesting that the peptide primarily adopts a flexible structure in TFE. Compared to the Ac-11-30-NH₂ peptide, the amide proton region (7.0-8.5 ppm), exhibits sharper and more distinct peaks, suggesting reduced aggregation and a higher degree conformational homogeneity.

Two-dimensional NMR experiments (TOCSY and NOESY) enabled sequential assignment to be made and revealed short- to medium- range NOE connectivities, which are consistent with the presence of a α -helical elements. These spectral features corroborate the CD results, which indicated a helical propensity for this peptide.

The one- and two-dimensional NMR spectra supporting these observations are reported in Appendix.

2.7. Interaction Studies

The interaction studies were carried out in collaboration with Prof. Sissi research group at the Department of Pharmaceutical and Pharmacological Sciences at the University of Padova at the University of Padova. Following an initial comparative analysis of all synthesised peptides, the focus shifted towards the non-acetylated variants, as these exhibited similar overall behaviour to their acetylated counterparts, albeit with slightly stronger interactions, when present. This difference is likely due to the presence of an additional positive charge at the free N-terminus, which could increase electrostatic attraction towards the negatively charged G-quadruplex DNA backbone.

2.7.1. Electrophoretic Mobility Shift Assay

Electrophoretic Mobility Shift Assay (EMSA) is a well-established technique for studying protein-nucleic acid interactions, for example. It enables specific binding events to be identified based on the decreased electrophoretic mobility of a nucleic acid when it is complexed with a protein. The assay is based on the fact that a DNA-protein complex migrates more slowly through a non-denaturing polyacrylamide gel than free DNA due to its increased molecular weight and altered charge-to-mass ratio.[71]

In this study, EMSA experiments were performed to investigate the interaction between Vimentin and the G-quadruplex-forming sequence 2TEL, and to evaluate the impact of synthesised Vimentin-derived peptides in this interaction. The assays were performed in 5 mM TRIS buffer (pH= 8.4) containing 150 mM KCl, which stabilises G4s, using 0.5 μ M 2TEL and 8 μ M Vimentin. Increasing concentrations of peptides (5-180 μ M) were tested to evaluate their ability to modulate the formation of the Vimentin-G4 complex. The experimental setup involved pre-incubating the DNA with KCl and peptides, followed by adding Vimentin and a second incubation step. This ensured that the DNA could

Results and Discussion



interact with either the peptide or Vimentin to the formation of a complex. The gels were analysed using both SYBER GREEN II staining and Coomassie Brilliant Blue staining to detect protein-containing complexes.

2.7.1.1. Interaction of Vimentin-Derived Peptides with the Vimentin-G4 Complex

The EMSA results revealed that the three Vimentin-derived peptides (H-11-30-NH₂, H-68-79-NH₂, and H-270-280-NH₂) exert distinct effects on the interaction between Vimentin and the 2TEL G4 sequence (Figure 2.78).

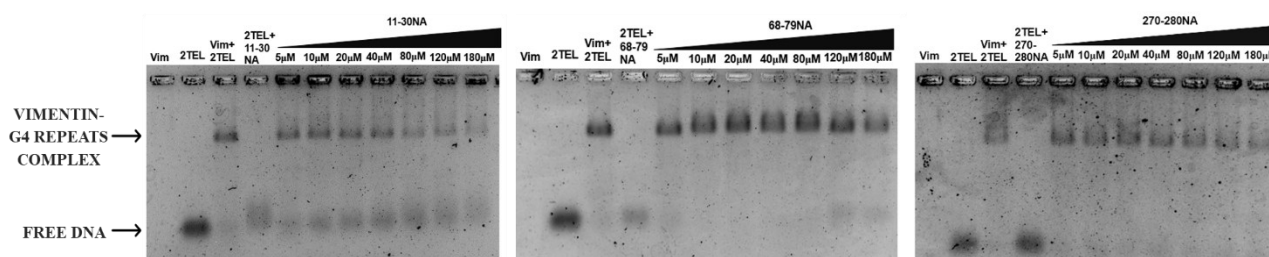


Figure 2.78. Electrophoretic Mobility Shift Assay of Vimentin-derived peptides with the G4 repeat-forming sequence 2TEL.

In the control lanes containing only 2TEL, a single lower band corresponding to free DNA was observed. However, the presence of Vimentin induced a pronounced shift, which is consistent with the formation of the Vimentin-G4 complex.

When increasing concentrations of peptides were added, differential modulation of this complex was detected.

- The H-11-30-NH₂ peptide produced a clear, concentration dependent decrease in the intensity of the shifted band that corresponds to the Vimentin-G4 complex. This suggests competitive or allosteric interference with Vimentin G4-binding region. The peptide likely interacts with Vimentin itself – possibly with its head domain, from which it is derived – thereby hindering its ability to bind G4 DNA. The ethidium bromide panel confirms a gradual restoration of the free DNA band, which is consistent with the displacement of the G4 from the Vimentin complex.
- The H-68-79-NH₂ peptide did not induce disruption of the complex. This may be due to a lack of affinity for DNA, or to its lesser involvement in G4 recognition.
- The H-270-280-NH₂ peptide had a little effect, even at the highest concentrations tested. The Vimentin-G4 complex remained largely intact and there was no significant reappearance of the free DNA band. This suggests that the central domain region represented by this peptide has a weak interaction to compete effectively under these conditions.

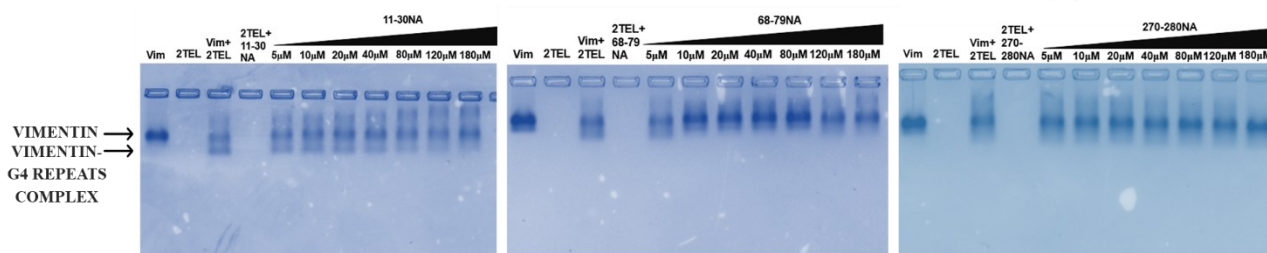


Figure 2.79. Electrophoretic Mobility Shift Assay of Vimentin-derived peptides with the G4 repeat-forming sequence 2TEL.

Coomassie-stained gels further confirm these observations: the bands corresponding to free Vimentin and Vimentin-G4 complex remain largely unchanged in the presence of H-68-79-NH₂ peptide. In contrast, in the H-11-30-NH₂ fragment, the complex band progressively weakens in line with peptide-induced displacement (Figure 2.79).

Overall, these results suggest that the N-terminal region of Vimentin (residues 11-30) is crucial for mediating G4 binding and that the synthetic peptide derived from this region can disrupt the Vimentin-G4 interaction effectively and in a dose-dependent manner. Conversely, the central and the other N-terminal fragments appear to play secondary roles rather than contributing to DNA recognition.

2.7.2. Fluorescence Melting Assay

In this work, fluorescence melting (FM) analysis is a technique used to evaluate the thermal stability of nucleic acid structures and how this is affected by ligand binding (Figure 2.80). In this method, a fluorophore-quencher-labelled oligonucleotide detects conformational changes by measuring variations in fluorescence intensity as the temperature increases. When the structure is folded, such as in the case of a G4, the proximity of the fluorophore and quencher reduces fluorescence emission. Upon thermal denaturation and unfolding, the two groups are separated, resulting in a measurable increase in fluorescence. The melting temperature (T_m), defined as the midpoint of this transition, reflects the stability of the folded structure.[72]

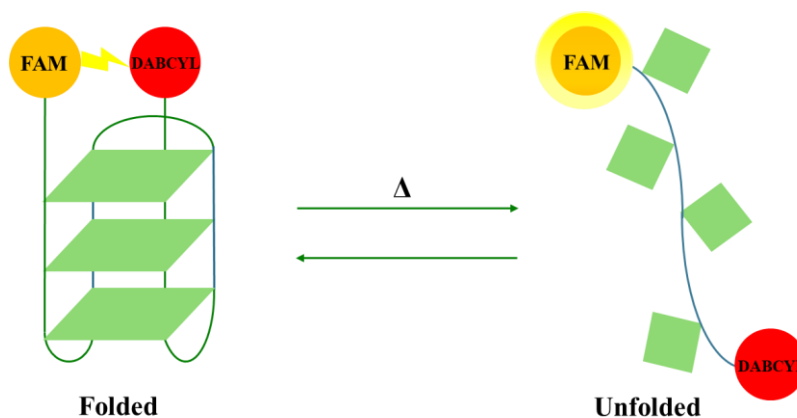


Figure 2.80. Schematic representation of the melting representation assay.

Results and Discussion



In this study, fluorescence melting experiments were employed to investigate the interaction Vimentin-derived peptides with G4 DNA (Tel23 FD) and duplex DNA. Changes in the melting temperatures (ΔT_m) were monitored when increasing concentrations of peptide (0-80 μM) were added to a solution of 50 mM KCl and 10 mM Li_3PO_4 buffer at pH= 7.4. Control experiments using duplex DNA and scrambled single-stranded sequences were also performed to assess selectivity. The observed T_m shifts provided direct stabilisation or destabilisation of nucleic acid structures.

2.7.2.1. Peptide-G4 and DNA Interaction Analysis

Fluorescence melting experiments revealed interactions between Vimentin-derived peptides and G4. H-11-30- NH_2 peptide exhibited the most pronounced effect (Figure 2.81).

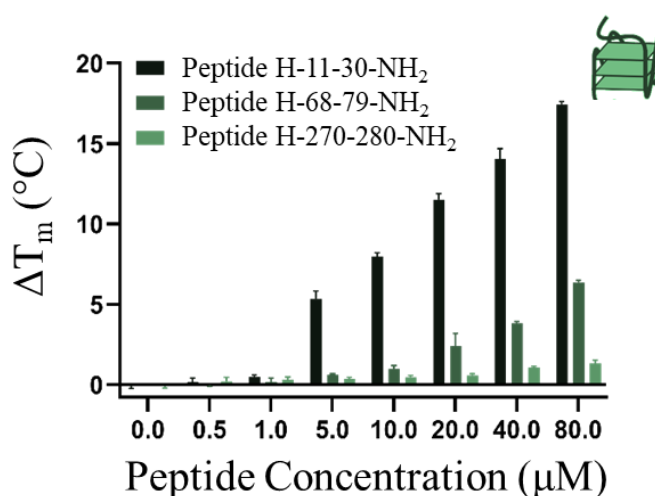


Figure 2.81. Fluorescence melting analysis of the interaction between Vimentin-derived peptides and the G4-forming sequence Tel23.

When titrated with the Tel23 G4 sequence, this peptide induced a concentration-dependent increase in ΔT_m , reaching almost $+18^{\circ}\text{C}$ at 80 μM . This significant thermal stabilisation indicates a direct interaction that reinforces the G4 secondary structure, either by stabilising the Hoogsten hydrogen bonds within the guanine tetrads, or by reducing the flexibility of the loop regions.

In contrast, H-68-79- NH_2 and H-270-280- NH_2 peptides produced only minimal increases in ΔT_m , suggesting a much weaker or transient interaction with the quadruplex scaffold.

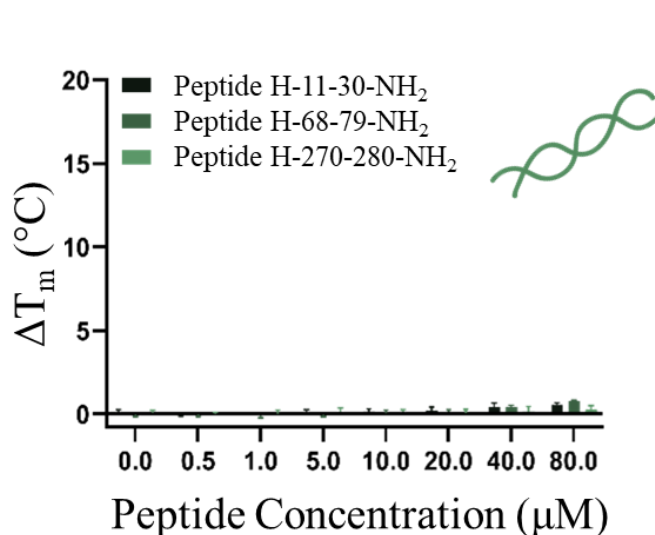


Figure 2.82. Fluorescence melting analysis of the interaction between Vimentin-derived peptides and duplex DNA.

When the same peptides were tested on duplex DNA, none of them caused significant changes to the melting temperature (Figure 2.82). This confirms that the observed effects are not due to non-specific electrostatic interactions with the DNA backbone but instead depend on the unique structural features of the G4 topology. H-11-30-NH₂ peptide preference for G4 over duplex DNA implies that it recognises specific conformational or electrostatic motifs associated with guanine quartets, such as planar π -stacked surfaces or the localised negative potential at the tetrad-loop interfaces.

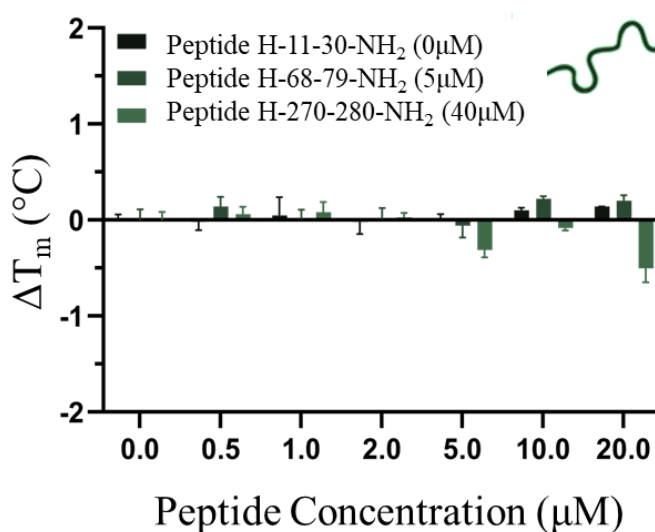


Figure 2.83. Fluorescence melting analysis of the interaction between H-11-30-NH₂ peptide and ssDNA.

To further evaluate selectivity, competition assays were conducted by introducing escalating concentrations of single-stranded DNA (ssDNA) in presence of peptide H-11-30-NH₂ (Figure 2.83). The addition of ssDNA did not significantly affect the melting temperature of the G4 in the presence of peptide H-11-30-NH₂, even at the highest tested ssDNA concentration (20 μM). This confirms that peptide H-11-30-NH₂ binds preferentially to the folded G4 structure rather than to unstructured DNA, demonstrating clear conformational selectivity.

Results and Discussion



From a structural standpoint, this selectivity may be due to the presence of aromatic and cationic residues within the peptide sequence. The 11-30 region of Vimentin is rich in arginine and aromatic residues, which can engage in π - π stacking interactions with the guanine bases, as well as forming electrostatic contacts with the phosphate backbone. The observed dose-dependent stabilisation likely reflects multiple binding events or cooperative interactions along quadruplex surface.

In contrast, H-68-79-NH₂ and H-270-280-NH₂ peptides, which lack a comparable distribution of cationic and aromatic residues, appear less capable of forming such stabilising contacts. Their modest or negligible ΔT_m shifts suggest that these regions may not directly participate in G4 recognition in the full-length Vimentin protein but could instead contribute to secondary contacts or structural support in the overall binding process.

Taken together, the fluorescence melting and competition data highlight that the Vimentin H-11-30-NH₂ peptide fragment acts as a G4 stabilising element with high structural and topological specificity. Combined with EMSA data showing its ability to disrupt the Vimentin-G4 complex, these findings strongly support a model in which this N-terminal region is the key recognition motif responsible for Vimentin selective interaction with G4 DNA.

2.7.2.2. Effect of Nle substitution

These fluorescence melting data clearly demonstrate that replacing Met¹⁴ with Nle does not alter the G4-binding profile of the Vimentin H-11-30-NH₂ peptide (Figure 2.84).

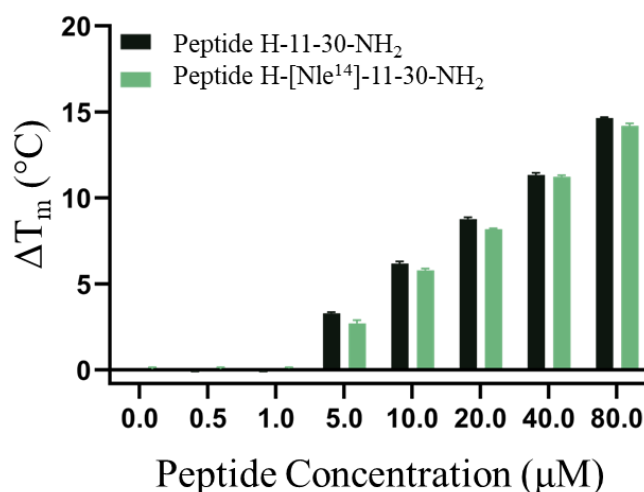


Figure 2.84. Fluorescence melting analysis of the interaction between H-11-30-NH₂ peptide and its Nle substituted analogue H-[Nle¹⁴]-11-30-NH₂ peptide with the G4-forming sequence Tel23.

With the G4 conformation, there is a clear, dose-dependent increase in the Tel23 FD melting temperature. The ΔT_m values increase almost in parallel up to 80 μ M, and the small differences observed at intermediate concentrations remain within the limits of experimental variability. This indicates that the side chain at position 14 is not a key determinant of G4 recognition or stabilisation.



Removing the sulphur atom and replacing Met with isosteric, non-oxidable Nle leaves the interaction that matter intact.

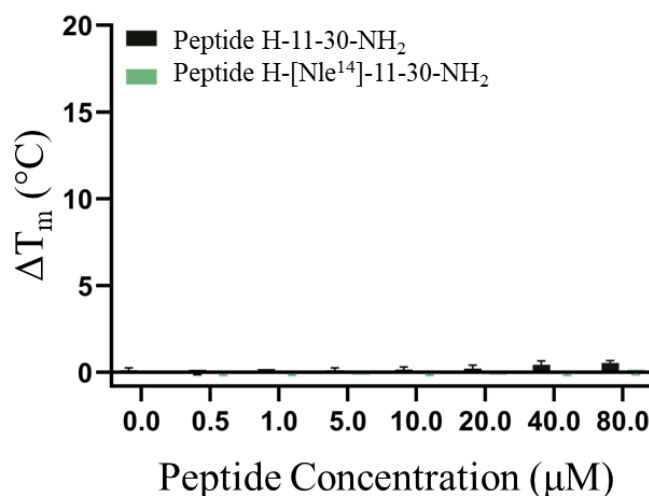


Figure 2.85. Fluorescence melting analysis of the interaction between H-11-30-NH₂ peptide and its Nle substituted analogue H-[Nle¹⁴]-11-30-NH₂ peptide with the duplex DNA.

The situation with the duplex DNA is the same as in previous melting experiments: both peptides give a similar value of ΔT_m over the whole concentration range. Therefore, the Met→Nle substitution does not introduce any loss of selectivity – the peptide remains G4-preferring and does not bind to duplex DNA (Figure 2.85).

These data validate the Met→Nle replacement as a synthetic/stability modification: it prevents possible Met oxidation during synthesis, work-up, or storage, yet does not alter the functionality of the peptide. Therefore, the Nle containing version can be used in the rest of the biophysical studies allowing a direct comparison with the native sequence.

2.7.2.3. Fluorescence Melting Analysis of Ala Scan Peptides

In order to analyse the molecular determinants underlying the interaction between the Vimentin-derived peptide H-11-30-NH₂ and G4, a systematic alanine scanning approach was employed (Figure 2.86). Substitutions were made to either the aromatic residues (Tyr¹¹, Phe¹⁵, Tyr³⁰) or the positively charged arginines (Arg¹², Arg¹³, Arg²³, Arg²⁸). This allowed to evaluate their respective contributions to G4 recognition and stabilisation.

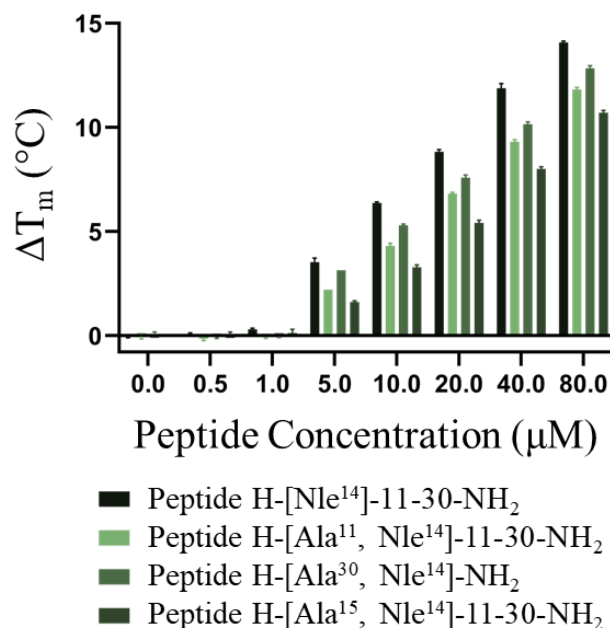


Figure 2.86. Fluorescence melting analysis of the interaction between *H*-[Nle¹⁴]-11-30-NH₂ peptide and its aromatic substituted analogue peptides with G4 forming sequence Tel23.

The fluorescence melting data revealed that replacing aromatic residues only moderately affected G4 stabilisation. All aromatic-substituted peptides exhibited a concentration-dependent increase in ΔT_m , albeit with slightly reduced efficiency compared to the wild-type peptide. The peptide with the Phe¹⁵ substitution plays a stronger role in G4 stabilisation than Tyr¹¹ or Tyr³⁰. This likely reflects the Phe¹⁵ ability to form a stable stacking contact with the tetrads. The Tyr residues, located closer to the termini, may engage more transiently through hydrogen-bond-assisted positioning thanks to their hydroxyl groups rather than direct stacking. Therefore, Tyr probably contribute to orientation, whereas Phe enhances the strength of binding and the overall stabilisation of the G4-peptide complex. Nevertheless, the persistence of significant stabilisation across all peptides indicates that aromatic contacts play a supportive rather than a determining role in complex formation.

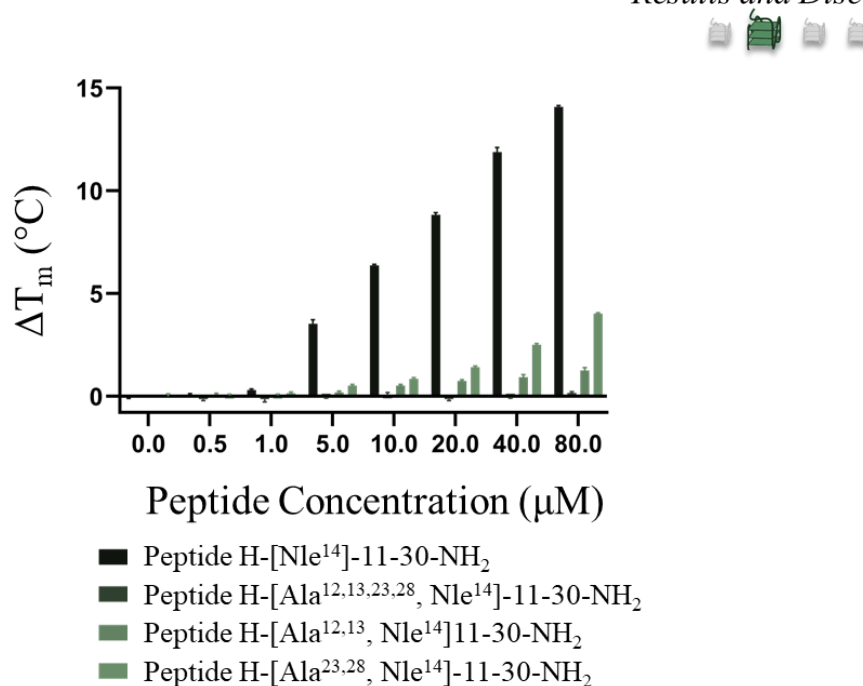


Figure 2.87. Fluorescence melting analysis of the interaction between H-[Nle]¹⁴-11-30-NH₂ peptide and its arginine substituted analogue peptides with G4 forming sequence Tel23.

In contrast, substitution of the arginine residues caused drastic loss of G4 stabilisation (Figure 2.87). The peptide in which all four arginines were replaced by alanines, exhibited an almost complete absence of an increase in ΔT_m , while the double substitution of arginines displayed only residual activity at the highest concentrations. These results demonstrate that electrostatic interactions mediated by arginine side chains are the main driving force for peptide binding to G4 DNA. The guanidinium groups likely engage with the negatively charged grooves and phosphate backbone of the folded G4 structure, providing both affinity and structural stabilisation.

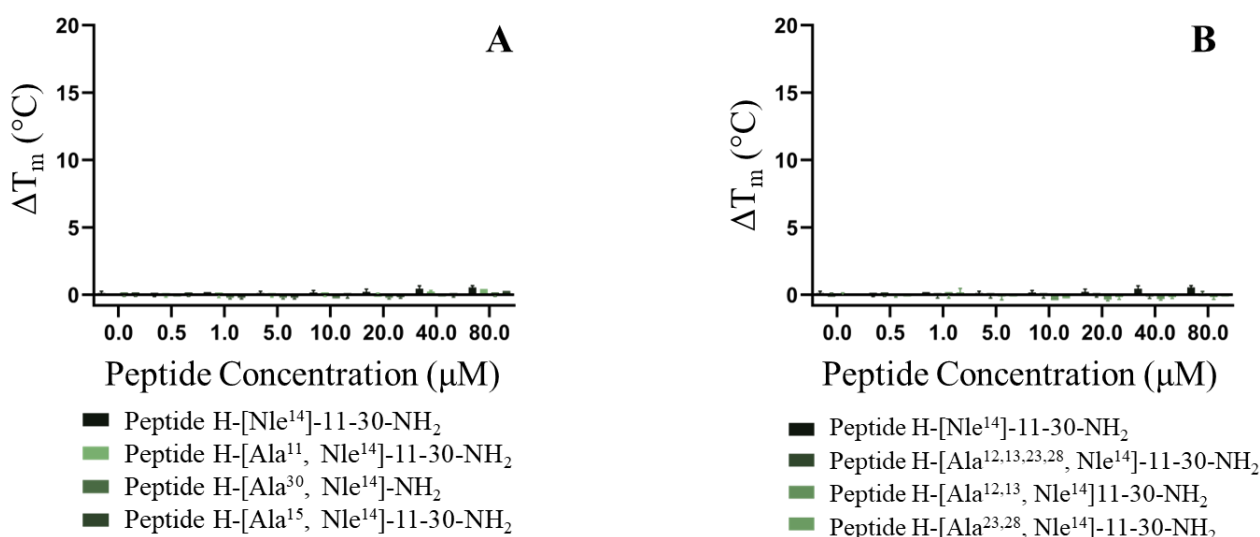


Figure 2.88. Fluorescence melting analysis of the interaction between H-[Nle]¹⁴-11-30-NH₂ peptide and its A) aromatic and B) arginine substituted analogue peptides with G4 forming sequence Tel23.

Results and Discussion



Crucially, none of these peptides exhibited a measurable interaction with duplex DNA. This confirms that the observed stabilization originates from the specific recognition of the G4 topology, rather than from non-specific charge-charge attraction (Figure 2.88).

Together, these findings establish that the binding of the Vimentin H-11-30-NH₂ peptide to G4 is primarily charge-driven, with aromatic residues contributing to the interaction secondarily. This electrostatic mechanism, coupled with shape complementarity to the G4 structure, explains the peptide ability to recognise G-quadruplex motifs selectively and functionally.

2.7.3. Surface Plasmon Resonance Analysis

Surface Plasmon Resonance (SPR) is a label-free, real time biophysical technique that enables the quantitative characterisation of biomolecular interactions (Figure 2.89). It provides kinetic and thermodynamic parameters, such as the association (k_a) and dissociation (k_d) rate constants and the equilibrium dissociation constant (K_D).

In SPR, one binding partner (the ligand) is immobilised on a sensor chip – typically through biotin-streptavidin coupling – while the other (the analyte) is passed over the surface at increasing concentrations. Binding events induce changes in the refractive index near the sensor surface, which are detected as shifts in the SPR signal. The association and dissociation phases are recorded in real time, enabling the determination of the kinetic and equilibrium binding parameters.[73]

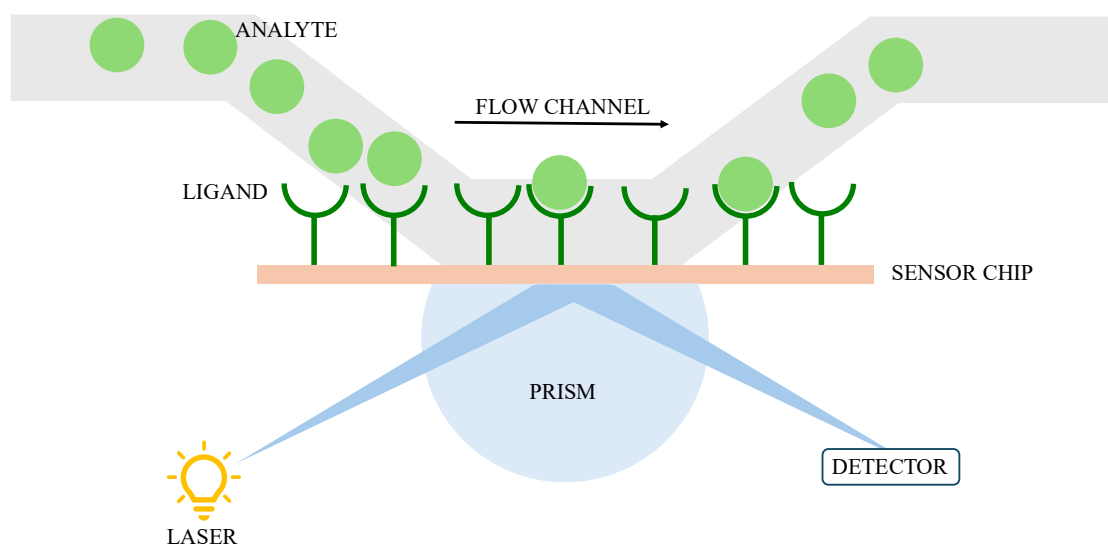


Figure 2.89. Schematic representation of SPR assay.

In the present study, SPR was employed to examine the interaction between Vimentin-derived peptides and various DNA substrates, including a G4-forming sequence, a G4 repeat and single-stranded DNA (ssDNA). DNA oligonucleotides were biotinylated and immobilised on a streptavidin-coated chip at a concentration of 50 nM. Meanwhile, peptide solutions at increasing concentrations (5-50 μ M) were injected in TRIS-KCl buffer (pH= 8.4) containing 0.025% Tween-20 to minimise



non-specific binding. The dissociation phase was monitored using a 10 mM glycine buffer solution at pH= 2.5 to ensure complete regeneration of the chip surface between runs.

2.7.3.1. Binding of Vimentin-derived peptides to G4 DNA structures

SPR analysis revealed a clear and selective binding profile for H-11-30-NH₂ peptide, whereas the other two Vimentin-derived fragments (H-68-79-NH₂ and H-270-280-NH₂ peptide) showed no measurable affinity under the tested conditions (Figure 2.90).

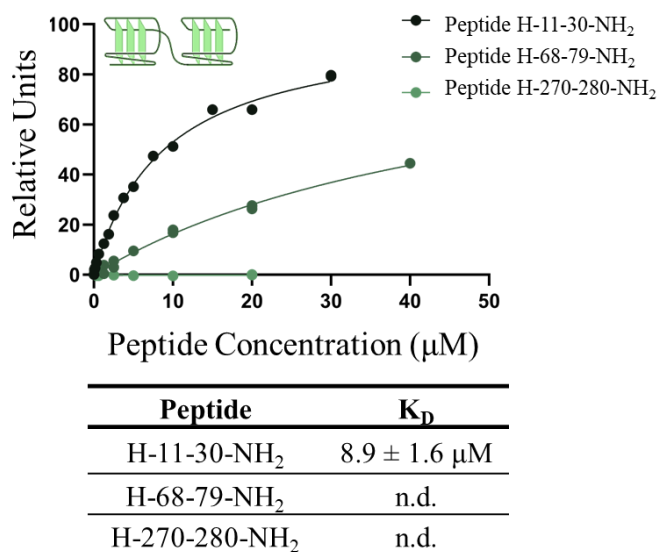


Figure 2.90. SPR analysis of Vimentin-derived peptides binding to the G4-repeats forming sequence.

The curves showed that the SPR response for H-11-30-NH₂ peptide increased in a concentration-dependent manner, followed by reversible dissociation upon buffer washing. This is indicative of a specific, non-covalent interaction with the G4 substrates (Figure 2.91).

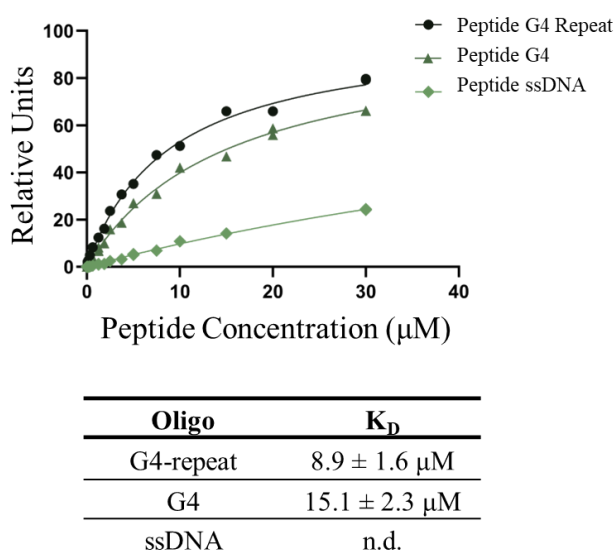


Figure 2.91. SPR analysis of H-11-30-NH₂ peptide binding to different DNA structure.

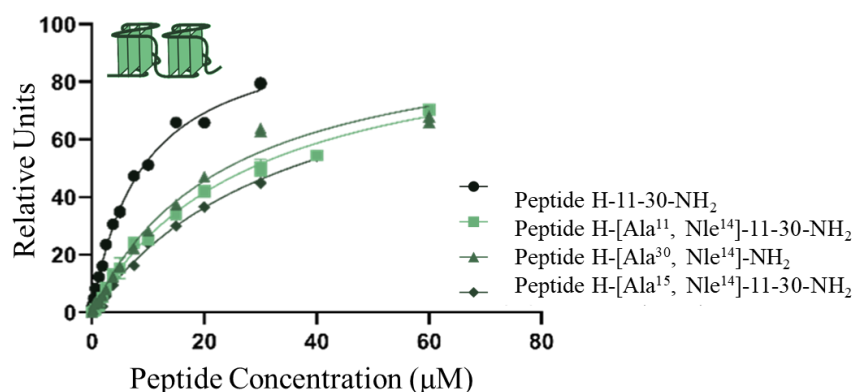
Results and Discussion



Quantitative fitting of the steady-state binding curves yielded K_D values of $8.9 \pm 1.6 \mu\text{M}$ for the G4 repeat and $15.1 \pm 2.3 \mu\text{M}$ for the single G4 structure. These values confirm moderate to high affinity, suggesting that the peptide interacts efficiently with both individual G4 units and G4 multimeric arrays. In contrast, no detectable binding was observed towards ssDNA, further supporting the notion that the peptide recognises G4-specific structural motifs rather than generic nucleic acid backbones. The stronger affinity for the G4 repeat compared to the single G4 structure suggests a multivalent binding mode, possibly involving cooperative interactions between peptide molecules and adjacent G4 units. This is consistent with EMSA and fluorescence melting results, in which peptide H-11-30-NH₂ stabilised G4 structures and competed with full-length Vimentin for G4 binding. Therefore, these SPR findings corroborating that the N-terminal segment Vimentin (residues 11-30) constitutes the principal G4-recognition motif capable of specifically and moderately strong binding to telomeric G4 structures.

2.7.3.2. Aromatic-Substituted Peptides SPR Analysis

SPR experiments comparing the wild type H-11-30-NH₂ peptide with the three aromatic-substituted peptides provide key mechanistic insight (Figure 2.92).



Peptide	K_D
Peptide H-11-30-NH ₂	$8.9 \pm 1.6 \mu\text{M}$
Peptide H-[Ala ¹¹ , Nle ¹⁴]-11-30-NH ₂	$28.1 \pm 3.9 \mu\text{M}$
Peptide H-[Ala ³⁰ , Nle ¹⁴]-NH ₂	$23.8 \pm 5.0 \mu\text{M}$
Peptide H-[Ala ¹⁵ , Nle ¹⁴]-11-30-NH ₂	$35.0 \pm 7.7 \mu\text{M}$

Figure 2.92. Fluorescence melting analysis of the interaction between H-11-30-NH₂ peptide and its aromatic substituted analogue peptides with G4-repeat forming sequence.

Although all three substitutions reduce binding affinity: replacing Phe¹⁵ results in the greatest reduction in affinity (~4 fold), followed by the Tyr¹¹ (~3.2 fold) and Tyr³⁰ (~2.7 fold) substitutions. This suggests that Phe¹⁵ plays the most critical role in the peptide-G4 interaction and is likely to engage in direct π - π stacking with one of the terminal guanine tetrads of the G4. The fact that the substitutions of Tyr¹¹ and Tyr³⁰ also impair binding, albeit to a lesser extent, implies that these residues



contribute secondary stabilising interactions, perhaps via hydrogen bonding, but are not primary stacking anchors. Thus, the aromatic network appears to be hierarchical: Phe¹⁵ acts as a central stacking anchor, while the Tyr residues fine-tune the interaction geometry and help to orient the peptide on the G4 surface.

In summary the aromatic residues may enhance stability and specificity, with Phe¹⁵ acting as the key stacking element.

Conclusion



3. Conclusion

In this chapter, structural and functional properties of selected Vimentin-derived peptides were explored. These peptides were designed to investigate and disrupt the interaction between Vimentin and G4 DNA structures. A combination of spectroscopic, biophysical and binding analyses provided an integrated view of how specific regions of Vimentin contribute to G4 recognition and stabilisation. CD and NMR spectroscopy revealed that the synthesised peptides are mainly disordered in aqueous solution but can acquire a secondary structure under helical-promoting conditions.

Functional assays demonstrated that the 11-30 fragment is the key region mediating Vimentin's interaction with G4-forming DNA sequences. EMSA experiments showed that increasing peptide concentrations can disrupt the Vimentin-G4 complex. Fluorescence melting and SPR studies confirmed the direct binding of the 11-30 fragment to the G4 repeats with micromolar affinity, but not to duplex or single stranded DNA. Competition experiments further indicated that this interaction is highly selective towards the G4 topology.

Ala scan analyses provided insight into the molecular determinants of binding. Substituting aromatic residues reduced G4 affinity moderately, suggesting that these residues play a role in structural stabilisation or π -stacking interactions. In contrast, substitution of the positively charged Arg residues resulted in a significant loss of binding and G4 stabilisation ability, emphasising the significant contribution of electrostatic interactions to G4 recognition. Taken together, these findings suggest that the H-11-30-NH₂ peptide primarily binds to G4 structures through charge complementarity, which is potentially reinforced by local conformational ordering within the peptide.

Although this study focuses on isolated Vimentin-derived peptide fragments, their functional interpretation becomes more relevant when considered in the context of the full-length Vimentin structure. Vimentin consists of a central α -helical coiled-coil rod domain, which is flanked by intrinsically disordered N- and C-terminal domains. Predictive structural models suggests that the rod domain adopts a stable dimeric coiled-coil structure that is responsible for filament assembly. Meanwhile, the terminal regions are predicted to retain a high degree of conformational flexibility and solvent exposure. Several of the peptide fragments identified as G4-binding competent correspond to regions predicted to be structurally accessible or dynamically disordered in the intact protein. This observation supports the hypothesis that G4 recognition may occur through flexible, surface-exposed segments capable of adapting to non-canonical DNA topologies. Rather than acting as isolated binding motifs, these regions may function as dynamic nucleic acid-interacting modules within the full-length protein, whose accessibility could be modulated by the state of filament assembly, post-translational modifications, or the local cellular context. Therefore, integrating the

Conclusion



peptide-level mapping with structural models of Vimentin strengthens the biological interpretation of the data and suggests that G4 interaction may represent an additional regulatory dimension of Vimentin, beyond its canonical cytoskeletal role

Overall, these results identify the Vimentin 11-30 region as a minimal G4-interacting motif capable of selectively recognising and stabilising G4 DNA, while maintaining the physicochemical features necessary for molecular recognition.



4. Experimental Part

4.1. General Procedure for Solid-Phase Synthesis of Peptides

Manual approach to solid-phase synthesis was employed to prepare the peptide sequences described in this thesis.

6. **Resin swelling.** The loaded resin was placed in the syringe reactor and swollen by the addition of DMF, followed by 40 minutes of mechanical agitation. The solvent is then removed by filtration directly through the syringe.
7. **Fmoc Deprotection.** The Fmoc protecting group is removed by adding either a 20% piperidine or 20% piperidine in DMF in the presence of Oxyma Pure 1M. The mixture is mechanically agitated for 5 minutes, followed by a second addition of the same solution and agitation for an additional 15 minutes.
8. **Washing step.** After each operation, the solution is removed by filtration, and the resin is washed with five portions of DMF. Each wash consists of adding DMF, agitating for a few minutes and filtering off the solvent.
9. **Coupling reaction.** The coupling mixture is prepared using 3 equivalents of Fmoc-AminoAcid-OH, together with an equimolar amount of Oxyma Pure in the presence of DIC in anhydrous DMF. This solution is then transferred to the reactor containing the resin. The mixture is mechanically agitated for 50 minutes. Then, the resin is washed with five portions of DMF (for double coupling, after filtering the reaction mixture, a second portion of the activated amino acid or nucleobase is added, and the resin is agitated before washing).
10. **Chain Elongation.** Steps 2 to 4 are then repeated for the incorporation of subsequent monomers until the desired sequence is complete.

The peptide chain can be acetylated to complete the sequence.

4.2. Acetylation

Once the sequence has been synthesised, the peptide can be acetylated at the N-terminus by adding an acetyl group. The reaction solution contains 6 eq of Ac_2O and 1.5 eq of DIPEA. For example, for a 10mL solution: 0.475 mL of Ac_2O and 0.225 mL of DIPEA, which is then brought up to volume with DMF. 2 mL of the solution is transferred to a syringe reactor, stirred for 60 minutes, and this process is repeated twice after emptying by filtration. Finally, five washes are performed with 2 mL of DMF each, followed by five final washes with DCM.



4.3. Convergent synthesis of peptide fragments

Convergent synthesis involves the following steps:

1. **Resin swelling.** The resin loaded into the syringe is swollen by adding 3 mL of DMF and mechanically shaking the syringe for 40 minutes. The solvent is then removed by filtration using a nitrogen flow directly into the syringe.
2. **Coupling reaction.** The coupling reaction involves preparing the coupling mixture with 2 eqs of Peptide fragment, protected at its N-terminus and side chains, and 2 eqs of HOAt activator in 2 mL of anhydrous DMF. This mixture is then transferred to a reactor containing the resin, to which 2 eqs of DIC are added to promote the reaction. The resin is then mechanically stirred for four hours and heated in a sand bath at 40°C. Two hours after the reaction begins, a further 2 eqs of DIC are added. After 4 hours, the resin is washed with five 2 mL aliquots of DMF.
3. **Fmoc deprotection.** The Fmoc protecting group is removed by treating the resin with 2 mL of 20% piperidine solution in DMF, agitated mechanically for 5 minutes. A second 2 mL portion of the same solution is then added, and agitation continues for an additional 15 minutes.
4. **Washing.** At the end of each step, the solution is filtered from the syringe, and the resin is washed with five portions of DMF.

4.4. Cleavage from the resin

At the end of the synthesis process, the peptide is cleaved from the resin. The cleavage procedure varies depending on the types of amino acids present in the sequence.

A standard cleavage cocktail consisting of TFA/TIS/H₂O (95:2.5:2.5, v/v/v) is typically employed.

However, if the sequence contains one or more methionine residues, a different mixture is used to prevent oxidation and protect sensitive chains: TFA/H₂O/DODT/TIS or TFA/H₂O/EDT/TIS (94:2.5:2.5:1, v/v/v/v). The cleavage reaction is carried out under gentle agitation for 2-3 hours.

To retain side chain protecting groups after peptide-resin cleavage, a milder mixture of 30% 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP) in DCM is used instead. The cleavage reaction is carried out twice for 1 hour under gentle agitation.

Once complete, the reaction mixture is filtered to collect the filtrate containing the crude peptide. Most of the solvent is then removed under a nitrogen stream and the peptide is precipitated by adding cold Et₂O.

The resulting suspension is then centrifuged three times at 6500 rpm for six minutes. The supernatant is removed after each cycle, and a fresh aliquot of cold diethyl ether is added to wash the precipitate.

Experimental Part



The procedure for ‘Coupling of the first amino acid onto 2-chlorotrityl chloride resin’, ‘determination of resin loading after Fmoc deprotection by UV-Vis spectroscopy and ‘analytical cleavage’ are reported in the Experimental Section of Chapter I.

4.5. Synthesis

4.5.1. Synthesis of H-11-30-NH₂ peptide

Sequence:

H-Tyr-Arg-Arg-Met-Phe-Gly-Gly-Pro-Gly-Thr-Ala-Ser-Arg-Pro-Ser-Ser-Ser-Arg-Ser-Tyr-NH₂

Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)	
Rink Amide		0.65		0.125		0.08	
Step	Residue	Exc. (eq)	Fmoc-AA-OH Coupling (mg)	Activating Agent 1 (g)		Activating Agent 2 (μL)	
1	Y (Tyr)	3	0.2241	Oxyma	0.0693	DIC	75.5
2	S (Ser)	3	0.1870	Oxyma	0.0693	DIC	75.5
3	R (Arg)	3	0.3163	Oxyma	0.0693	DIC	75.5
4	S (Ser)	3	0.1870	Oxyma	0.0693	DIC	75.5
5	S (Ser)	3	0.1870	Oxyma	0.0693	DIC	75.5
6	S (Ser)	3	0.1870	Oxyma	0.0693	DIC	75.5
7	P (Pro)	3	0.1645	Oxyma	0.0693	DIC	75.5
8	R (Arg)	3	0.3163	Oxyma	0.0693	DIC	75.5
9	S (Ser)	3	0.1870	Oxyma	0.0693	DIC	75.5
10	A (Ala)	3	0.1518	Oxyma	0.0693	DIC	75.5
11	T (Thr)	3	0.3163	Oxyma	0.0693	DIC	75.5
12	G (Gly)	3	0.1450	Oxyma	0.0693	DIC	75.5
13	P (Pro)	3	0.1645	Oxyma	0.0693	DIC	75.5
14	G (Gly)	3	0.1450	Oxyma	0.0693	DIC	75.5
15	G (Gly)	3	0.1450	Oxyma	0.0693	DIC	75.5
16	F (Phe)	3	0.1811	Oxyma	0.0693	DIC	75.5
17	M (Met)	3	0.1889	Oxyma	0.0693	DIC	75.5
18	R (Arg)	3	0.3163	Oxyma	0.0693	DIC	75.5
19	R (Arg)	3	0.3163	Oxyma	0.0693	DIC	75.5
20	Y (Tyr)	3	0.2241	Oxyma	0.0693	DIC	75.5



The peptide synthesis was carried out manually.

Double coupling was performed for the following amino acids: Tyr¹¹, Arg¹², Arg¹³, Pro¹⁸, Gly¹⁹, Thr²⁰, Ala²¹, Ser²², Arg²³, Tyr³⁰.

Fmoc deprotection was performed using 20% piperidine solution in DMF up to the Thr in position 20 residue and subsequently using a 20% piperidine solution in DMF containing Oxyma Pure 1M.

After the coupling of the Ser at position 25, Thr at position 20 and Gly at position 16, the sequence was verified by analytical cleavage.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5µm, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using Luna C18, 300Å, 10µm column. Increasing amounts of the crude product were injected progressively over the course of eight consecutive runs. The elution was conducted with a gradient of 5-35% B over 30 minutes.

The fractions displaying the highest purity were pooled, acetonitrile was removed under reduced pressure, and the peptide was subsequently lyophilised.

Final yield after purification: 4%

Purity (HPLC): 98%

Retention Time: 10.2 min

Molecular Formula: C₉₄H₁₄₇N₃₃O₂₈S, [M+H]⁺_{Calculated}=2219.1

MS: [M+H]⁺=2218.9; [M+2H]²⁺=1109.6; [M+3H]³⁺=740.2

Experimental Part



4.5.2. Synthesis of Ac-11-30-NH₂ peptide

Sequence:

Ac-Tyr-Arg-Arg-Met-Phe-Gly-Gly-Pro-Gly-Thr-Ala-Ser-Arg-Pro-Ser-Ser-Ser-Arg-Ser-Tyr-NH₂

Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)	
Rink Amide		0.65		0.125		0.08	
Step	Residue	Exc. (eq)	Fmoc-AA-OH Coupling (mg)	Activating Agent 1 (g)		Activating Agent 2 (μL)	
1	Y (Tyr)	3	0.2241	Oxyma	0.0693	DIC	75.5
2	S (Ser)	3	0.1870	Oxyma	0.0693	DIC	75.5
3	R (Arg)	3	0.3163	Oxyma	0.0693	DIC	75.5
4	S (Ser)	3	0.1870	Oxyma	0.0693	DIC	75.5
5	S (Ser)	3	0.1870	Oxyma	0.0693	DIC	75.5
6	S (Ser)	3	0.1870	Oxyma	0.0693	DIC	75.5
7	P (Pro)	3	0.1645	Oxyma	0.0693	DIC	75.5
8	R (Arg)	3	0.3163	Oxyma	0.0693	DIC	75.5
9	S (Ser)	3	0.1870	Oxyma	0.0693	DIC	75.5
10	A (Ala)	3	0.1518	Oxyma	0.0693	DIC	75.5
11	T (Thr)	3	0.3163	Oxyma	0.0693	DIC	75.5
12	G (Gly)	3	0.1450	Oxyma	0.0693	DIC	75.5
13	P (Pro)	3	0.1645	Oxyma	0.0693	DIC	75.5
14	G (Gly)	3	0.1450	Oxyma	0.0693	DIC	75.5
15	G (Gly)	3	0.1450	Oxyma	0.0693	DIC	75.5
16	F (Phe)	3	0.1811	Oxyma	0.0693	DIC	75.5
17	M (Met)	3	0.1889	Oxyma	0.0693	DIC	75.5
18	R (Arg)	3	0.3163	Oxyma	0.0693	DIC	75.5
19	R (Arg)	3	0.3163	Oxyma	0.0693	DIC	75.5
20	Y (Tyr)	3	0.2241	Oxyma	0.0693	DIC	75.5

The peptide synthesis was carried out manually.

Double coupling was performed for the following amino acids: Tyr¹¹, Arg¹², Arg¹³, Pro¹⁸, Gly¹⁹, Thr²⁰, Ala²¹, Ser²², Arg²³, Tyr³⁰.

Fmoc deprotection was performed using 20% piperidine solution in DMF up to the Thr at position 20 residue and subsequently using a 20% piperidine solution in DMF containing Oxyma Pure 1M.



After the coupling of the Ser at position 25, Thr at position 20 and Gly at position 16, the sequence was verified by analytical cleavage.

The standard procedure using acetic anhydride was followed for the acetylation step.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5 μ m, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using Luna C18, 300Å, 10 μ m column. Increasing amounts of the crude product were injected progressively over the course of eight consecutive runs.

The elution was conducted with a gradient of 5-35% B over 30 minutes.

The fractions displaying the highest purity were pooled, acetonitrile was removed under reduced pressure, and the peptide was subsequently lyophilised.

Final yield after purification: 6%

Purity (HPLC): 95%

Retention Time: 10.9 min

Molecular Formula: C₉₆H₁₄₉N₃₃O₂₉S, [M+H]⁺_{Calculated}=2261.5

MS: [M+H]⁺=2260.9; [M+2H]²⁺=1131.4; [M+3H]³⁺=754.3

Experimental Part



4.5.3. Synthesis of H-68-79-NH₂ peptide

Sequence:

H-Val-Arg-Leu-Arg-Ser-Ser-Val-Pro-Gly-Val-Arg-Leu-NH₂

Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)	
Rink Amide		0.65		0.250		0.1625	
Step	Residue	Exc. (eq)	Fmoc-AA-OH Coupling (mg)	Activating Agent 1 (g)		Activating Agent 2 (μL)	
1	L (Leu)	3	0.1723	Oxyma	0.0693	DIC	75.5
2	R (Arg)	3	0.3163	Oxyma	0.0693	DIC	75.5
3	V (Val)	3	0.1655	Oxyma	0.0693	DIC	75.5
4	G (Gly)	3	0.1450	Oxyma	0.0693	DIC	75.5
5	P (Pro)	3	0.1645	Oxyma	0.0693	DIC	75.5
6	V (Val)	3	0.1655	Oxyma	0.0693	DIC	75.5
7	S (Ser)	3	0.1870	Oxyma	0.0693	DIC	75.5
8	S (Ser)	3	0.1870	Oxyma	0.0693	DIC	75.5
9	R (Arg)	3	0.3163	Oxyma	0.0693	DIC	75.5
10	L (Leu)	3	0.1723	Oxyma	0.0693	DIC	75.5
11	R (Arg)	3	0.3163	Oxyma	0.0693	DIC	75.5
12	V (Val)	3	0.1655	Oxyma	0.0693	DIC	75.5

The peptide synthesis was carried out manually.

Fmoc deprotection was performed using 20% piperidine solution in DMF.

After the coupling of the Ser at position 72 and again at the end of the synthesis, the sequence was verified by analytical cleavage.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5μm, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using Luna C18, 300Å, 10μm column. Increasing amounts of the crude product were injected progressively over the course of five consecutive runs. The elution was conducted with a gradient of 10-40% B over 40 minutes.

The fractions displaying the highest purity were pooled, the acetonitrile was removed under reduced pressure, and the peptide was subsequently lyophilised.

Final yield after purification: 28%



Purity (HPLC): 97%

Retention Time: 10.9 min

Molecular Formula: $C_{58}H_{108}N_{22}O_{14}$, $[M+H]^+$ Calculated=1338.6

MS: $[M+H]^+$ =1338.8; $[M+2H]^{2+}$ =669.4

4.5.4. Synthesis of Ac-68-79-NH₂ peptide

Sequence:

Ac-Val-Arg-Leu-Arg-Ser-Ser-Val-Pro-Gly-Val-Arg-Leu-NH₂

Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)	
Rink Amide		0.65		0.250		0.1625	
Step	Residue	Exc. (eq)	Fmoc-AA-OH Coupling (mg)	Activating Agent 1 (g)	Activating Agent 2 (μL)		
1	L (Leu)	3	0.1723	Oxyma 0.0693	DIC	75.5	
2	R (Arg)	3	0.3163	Oxyma 0.0693	DIC	75.5	
3	V (Val)	3	0.1655	Oxyma 0.0693	DIC	75.5	
4	G (Gly)	3	0.1450	Oxyma 0.0693	DIC	75.5	
5	P (Pro)	3	0.1645	Oxyma 0.0693	DIC	75.5	
6	V (Val)	3	0.1655	Oxyma 0.0693	DIC	75.5	
7	S (Ser)	3	0.1870	Oxyma 0.0693	DIC	75.5	
8	S (Ser)	3	0.1870	Oxyma 0.0693	DIC	75.5	
9	R (Arg)	3	0.3163	Oxyma 0.0693	DIC	75.5	
10	L (Leu)	3	0.1723	Oxyma 0.0693	DIC	75.5	
11	R (Arg)	3	0.3163	Oxyma 0.0693	DIC	75.5	
12	V (Val)	3	0.1655	Oxyma 0.0693	DIC	75.5	

The peptide synthesis was carried out manually.

Fmoc deprotection was performed using 20% piperidine solution in DMF.

After the coupling of the Ser at position 72 and again at the end of the synthesis, the sequence was verified by analytical cleavage.

The standard procedure using acetic anhydride was followed for the acetylation step.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5μm, 100Å, 4.6x100 mm column.

Experimental Part



Purification was performed by preparative HPLC using Luna C18, 300Å, 10µm column. Increasing amounts of the crude product were injected progressively over the course of five consecutive runs. The elution was conducted with a gradient of 10-40% B over 40 minutes.

The fractions displaying the highest purity were pooled, the acetonitrile was removed under reduced pressure, and the peptide was subsequently lyophilised.

Final yield after purification: 10%

Purity (HPLC): 99%

Retention Time: 11.9 min

Molecular Formula: C₆₀H₁₁₀N₂₂O₁₅, [M+H]⁺_{Calculated}=1379.8

MS: [M+H]⁺=1379.8; [M+2H]²⁺=690.0

4.5.5. Synthesis of H-270-280-NH₂ peptide

Sequence:

H-Arg-Asp-Val-Arg-Gln-Gln-Tyr-Glu-Ser-Val-Ala-NH₂

Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)	
Rink Amide		0.65		0.200		0.13	
Step	Residue	Exc. (eq)	Fmoc-AA-OH Coupling (mg)	Activating Agent 1 (g)	Activating Agent 2 (g)	Activating Agent 1 (g)	Activating Agent 2 (g)
1	A (Ala)	3	0.1214	Oxyma	0.0831	DIC	60.4
2	V (Val)	3	0.1324	Oxyma	0.0831	DIC	60.4
3	S (Ser)	3	0.1495	Oxyma	0.0831	DIC	60.4
4	E (Glu)	3	0.1660	Oxyma	0.0831	DIC	60.4
5	Y (Tyr)	3	0.1792	Oxyma	0.0831	DIC	60.4
6	Q (Gln)	3	0.2382	Oxyma	0.0831	DIC	60.4
7	Q (Gln)	3	0.2382	Oxyma	0.0831	DIC	60.4
8	R (Arg)	3	0.2531	Oxyma	0.0831	DIC	60.4
9	V (Val)	3	0.1324	Oxyma	0.0831	DIC	60.4
10	D (Asp)	3	0.1605	Oxyma	0.0831	DIC	60.4
11	R (Arg)	3	0.2531	Oxyma	0.0831	DIC	60.4

The peptide synthesis was carried out manually.

Double coupling was performed for the following amino acids: Gln²⁷⁴, Gln²⁷⁵

Fmoc deprotection was performed using 20% piperidine solution in DMF.



The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5 μ m, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using Luna C18, 300Å, 10 μ m column. Increasing amounts of the crude product were injected progressively over the course of six consecutive runs. The elution was conducted with a gradient of 5-35% B over 30 minutes.

The fractions displaying the highest purity were pooled, the acetonitrile was removed under reduced pressure, and the peptide was subsequently lyophilised.

Final yield after purification: 11%

Purity (HPLC): 98%

Retention Time: 8.11 min

Molecular Formula: C₅₆H₉₁N₁₉O₂₀, [M+H]⁺Calculated=1349.5

MS: [M+H]⁺=1349.6; [M+2H]²⁺=675.3

4.5.6. Synthesis of Ac-270-280-NH₂ peptide

Sequence:

Ac-Arg-Asp-Val-Arg-Gln-Gln-Tyr-Glu-Ser-Val-Ala-NH₂

Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)	
Rink Amide		0.65		0.200		0.13	
Step	Residue	Exc. (eq)	Fmoc-AA-OH Coupling (mg)	Activating Agent 1 (g)	Activating Agent 2 (μL)		
1	A (Ala)	3	0.1214	Oxyma 0.0831	DIC	60.4	
2	V (Val)	3	0.1324	Oxyma 0.0831	DIC	60.4	
3	S (Ser)	3	0.1495	Oxyma 0.0831	DIC	60.4	
4	E (Glu)	3	0.1660	Oxyma 0.0831	DIC	60.4	
5	Y (Tyr)	3	0.1792	Oxyma 0.0831	DIC	60.4	
6	Q (Gln)	3	0.2382	Oxyma 0.0831	DIC	60.4	
7	Q (Gln)	3	0.2382	Oxyma 0.0831	DIC	60.4	
8	R (Arg)	3	0.2531	Oxyma 0.0831	DIC	60.4	
9	V (Val)	3	0.1324	Oxyma 0.0831	DIC	60.4	
10	D (Asp)	3	0.1605	Oxyma 0.0831	DIC	60.4	
11	R (Arg)	3	0.2531	Oxyma 0.0831	DIC	60.4	

Experimental Part



The peptide synthesis was carried out manually.

Double coupling was performed for the following amino acids: Gln²⁷⁴, Gln²⁷⁵

Fmoc deprotection was performed using 20% piperidine solution in DMF.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5µm, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using Luna C18, 300Å, 10µm column. Increasing amounts of the crude product were injected progressively over the course of six consecutive runs.

The elution was conducted with a gradient of 5-35% B over 30 minutes.

The fractions displaying the highest purity were pooled, the acetonitrile was removed under reduced pressure, and the peptide was subsequently lyophilised.

Final yield after purification: 10%

Purity (HPLC): 97%

Retention Time: 9.35 min

Molecular Formula: C₅₈H₉₃N₁₉O₂₁, [M+H]⁺_{Calculated}=1390.7

MS: [M+H]⁺=1390.6; [M+2H]²⁺=695.8



4.5.7. Synthesis of Fmoc-18-30-NH₂ peptide

Sequence:

Fmoc-Pro-Gly-Thr-Ala-Ser-Arg-Pro-Ser-Ser-Ser-Arg-Ser-Tyr- NH₂

Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)	
Rink Amide		0.65		0.250		0.1625	
Step	Residue	Exc. (eq)	Fmoc-AA-OH Coupling (mg)	Activating Agent 1 (g)		Activating Agent 2 (μL)	
1	Y (Tyr)	3	224.06	Oxyma	36.9	DIC	40.3
2	S (Ser)	3	186.91	Oxyma	36.9	DIC	40.3
3	R (Arg)	3	316.29	Oxyma	36.9	DIC	40.3
4	S (Ser)	3	186.91	Oxyma	36.9	DIC	40.3
5	S (Ser)	3	186.91	Oxyma	36.9	DIC	40.3
6	S (Ser)	3	186.91	Oxyma	36.9	DIC	40.3
7	P (Pro)	3	164.48	Oxyma	36.9	DIC	40.3
8	R (Arg)	3	316.29	Oxyma	36.9	DIC	40.3
9	S (Ser)	3	186.91	Oxyma	36.9	DIC	40.3
10	A (Ala)	3	151.76	Oxyma	36.9	DIC	40.3
11	T (Thr)	3	193.78	Oxyma	36.9	DIC	40.3
12	G (Gly)	3	144.93	Oxyma	36.9	DIC	40.3
13	P (Pro)	3	164.48	Oxyma	36.9	DIC	40.3

The peptide synthesis was carried out manually.

Double coupling was performed for the following amino acids: Pro¹⁸, Gly¹⁹, Thr²⁰, Ala²¹, Ser²², Arg²³, Tyr³⁰.

Fmoc deprotection was performed using 20% piperidine solution in DMF.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5μm, 100Å, 4.6x100 mm column.

The peptide remained on the resin with the final Fmoc group still attached.

Purity (HPLC): 88%

Retention Time: 14.16 min

Molecular Formula: C₇₀H₁₀₀N₂₀O₂₂, [M+H]⁺ Calculated=1573.7

MS: [M+H]⁺=1573.7; [M+2H]²⁺=787.0

Experimental Part



4.5.8. Synthesis of Fmoc-[Ala³⁰]-18-30-NH₂ peptide

Sequence:

Fmoc-Pro-Gly-Thr-Ala-Ser-Arg-Pro-Ser-Ser-Ser-Arg-Ser-Ala- NH₂

Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)	
Rink Amide		0.65		0.250		0.1625	
Step	Residue	Exc. (eq)	Fmoc-AA-OH Coupling (mg)	Activating Agent 1 (g)		Activating Agent 2 (μL)	
1	Y (Tyr)	3	224.06	Oxyma	36.9	DIC	40.3
2	S (Ser)	3	186.91	Oxyma	36.9	DIC	40.3
3	R (Arg)	3	316.29	Oxyma	36.9	DIC	40.3
4	S (Ser)	3	186.91	Oxyma	36.9	DIC	40.3
5	S (Ser)	3	186.91	Oxyma	36.9	DIC	40.3
6	S (Ser)	3	186.91	Oxyma	36.9	DIC	40.3
7	P (Pro)	3	164.48	Oxyma	36.9	DIC	40.3
8	R (Arg)	3	316.29	Oxyma	36.9	DIC	40.3
9	S (Ser)	3	186.91	Oxyma	36.9	DIC	40.3
10	A (Ala)	3	151.76	Oxyma	36.9	DIC	40.3
11	T (Thr)	3	193.78	Oxyma	36.9	DIC	40.3
12	G (Gly)	3	144.93	Oxyma	36.9	DIC	40.3
13	P (Pro)	3	164.48	Oxyma	36.9	DIC	40.3

The peptide synthesis was synthesised manually.

Double coupling was performed for the following amino acids: Pro¹⁸, Gly¹⁹, Thr²⁰, Ala²¹, Ser²², Arg²³.

Fmoc deprotection was performed using 20% piperidine solution in DMF.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5μm, 100Å, 4.6x100 mm column.

The peptide remained on the resin maintaining the N-terminus Fmoc protection.

Purity (HPLC): 85%

Retention Time: 13.77 min

Molecular Formula: C₆₄H₉₆N₂₀O₂₁, [M+H]⁺Calculated=1425.3

MS: [M+H]⁺=1425.6; [M+2H]²⁺=712.4



4.5.9. Synthesis of Fmoc-[Ala^{23,28}]-18-30-NH₂ peptide

Sequence:

Fmoc-Pro-Gly-Thr-Ala-Ser-Ala-Pro-Ser-Ser-Ser-Ala-Ser-Tyr- NH₂

Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)	
Rink Amide		0.65		0.250		0.1625	
Step	Residue	Exc. (eq)	Fmoc-AA-OH Coupling (mg)	Activating Agent 1 (g)		Activating Agent 2 (μL)	
1	Y (Tyr)	3	224.06	Oxyma	36.9	DIC	40.3
2	S (Ser)	3	186.91	Oxyma	36.9	DIC	40.3
3	R (Arg)	3	316.29	Oxyma	36.9	DIC	40.3
4	S (Ser)	3	186.91	Oxyma	36.9	DIC	40.3
5	S (Ser)	3	186.91	Oxyma	36.9	DIC	40.3
6	S (Ser)	3	186.91	Oxyma	36.9	DIC	40.3
7	P (Pro)	3	164.48	Oxyma	36.9	DIC	40.3
8	R (Arg)	3	316.29	Oxyma	36.9	DIC	40.3
9	S (Ser)	3	186.91	Oxyma	36.9	DIC	40.3
10	A (Ala)	3	151.76	Oxyma	36.9	DIC	40.3
11	T (Thr)	3	193.78	Oxyma	36.9	DIC	40.3
12	G (Gly)	3	144.93	Oxyma	36.9	DIC	40.3
13	P (Pro)	3	164.48	Oxyma	36.9	DIC	40.3

The peptide synthesis was synthesised manually.

Double coupling was performed for the following amino acids: Pro¹⁸, Gly¹⁹, Thr²⁰, Ala²¹, Ser²², Tyr³⁰.

Fmoc deprotection was performed using 20% piperidine solution in DMF.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5μm, 100Å, 4.6x100 mm column.

The peptide remained on the resin with the final Fmoc group still attached.

Purity (HPLC): 91%

Retention Time: 14.71 min

Molecular Formula: C₆₄H₈₆N₁₄O₂₂, [M+H]⁺ Calculated=1403.6

[M+H]⁺=1403.6

Experimental Part



4.5.10. Synthesis of Fmoc-[Nle¹⁴]-11-17-OH peptide

Sequence:

Fmoc-Tyr-Arg-Arg-Nle-Phe-Gly-Gly-OH

Resin		Resin Loading (mmol/g)	Resin (g)	Scale (mmol)			
Gly-CTC-OH		1.1	0.250	0.275			
Step	Residue	Exc. (eq)	Fmoc-AA-OH Coupling (mg)	Activating Agent 1 (g)		Activating Agent 2 (μL)	
1	G (Gly)	3	245.47	Oxyma	0.1172	DIC	127.7
2	F (Phe)	3	319.61	Oxyma	0.1172	DIC	127.7
3	nL (Nle)	3	291.56	Oxyma	0.1172	DIC	127.7
4	R (Arg)	3	535.26	Oxyma	0.1172	DIC	127.7
5	R (Arg)	3	535.26	Oxyma	0.1172	DIC	127.7
6	Y (Tyr)	3	379.17	Oxyma	0.1172	DIC	127.7

The peptide synthesis was carried out manually.

Double coupling was performed for the following amino acids: Tyr¹¹, Arg¹², Arg¹³.

Fmoc deprotection was performed using 20% piperidine solution in DMF. The peptide was left with the final group still attached.

At the end of the synthesis, the sequence was cleaved from the resin using a solution of 30% HIFP in DCM.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5μm, 100Å, 4.6x100 mm column.

Final yield: 41%

Purity (HPLC): 97%

Retention Time: 31.27 min

Molecular Formula: C₈₅H₁₁₁N₁₃O₁₇S₂, [M+H]⁺Calculated=1649.9

MS: [M+H]⁺=1649.7



4.5.11. Synthesis of Fmoc-[Ala¹¹, Nle¹⁴]-11-17-OH peptide

Sequence:

Fmoc-Ala-Arg-Arg-Nle-Phe-Gly-Gly-OH

Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)	
Gly-CTC-OH		1.1		0.250		0.275	
Step	Residue	Exc. (eq)	Fmoc-AA-OH Coupling (mg)	Activating Agent 1 (g)		Activating Agent 2 (μL)	
1	G (Gly)	3	245.47	Oxyma	0.1172	DIC	127.7
2	F (Phe)	3	319.61	Oxyma	0.1172	DIC	127.7
3	nL (Nle)	3	291.56	Oxyma	0.1172	DIC	127.7
4	R (Arg)	3	535.26	Oxyma	0.1172	DIC	127.7
5	R (Arg)	3	535.26	Oxyma	0.1172	DIC	127.7
6	A (Ala)	3	256.82	Oxyma	0.1172	DIC	127.7

The peptide synthesis was carried out manually.

Double coupling was performed for the following amino acids: Arg¹², Arg¹³.

Fmoc deprotection was performed using 20% piperidine solution in DMF. The peptide was left with the final group still attached.

At the end of the synthesis, the sequence was cleaved from the resin using a solution of 30% HIFP in DCM.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5μm, 100Å, 4.6x100 mm column.

Final yield: 38%

Purity (HPLC): 98%

Retention Time: 28.93 min

Molecular Formula: C₇₅H₉₉N₁₃O₁₆S₂, [M+H]⁺_{Calculated}=1501.7

MS: [M+H]⁺=1501.6; [M+2H]²⁺=751.8

Experimental Part



4.5.12. Synthesis of Fmoc-[Ala¹⁵, Nle¹⁴]-11-17-OH peptide

Sequence:

Fmoc-Tyr-Arg-Arg-Nle-Ala-Gly-Gly-OH							
Resin		Resin Loading (mmol/g)		Resin (g)		Scale (mmol)	
Gly-CTC-OH		1.1		0.250		0.275	
Step	Residue	Exc. (eq)	Fmoc-AA-OH Coupling (mg)	Activating Agent 1 (g)		Activating Agent 2 (μL)	
1	G (Gly)	3	245.47	Oxyma	0.1172	DIC	127.7
2	A (Ala)	3	256.82	Oxyma	0.1172	DIC	127.7
3	nL (Nle)	3	291.56	Oxyma	0.1172	DIC	127.7
4	R (Arg)	3	535.26	Oxyma	0.1172	DIC	127.7
5	R (Arg)	3	535.26	Oxyma	0.1172	DIC	127.7
6	Y (Tyr)	3	379.17	Oxyma	0.1172	DIC	127.7

The peptide synthesis was carried out manually.

Double coupling was performed for the following amino acids: Tyr¹¹, Arg¹², Arg¹³.

Fmoc deprotection was performed using 20% piperidine solution in DMF. The peptide was left with the final group still attached.

At the end of the synthesis, the sequence was cleaved from the resin using a solution of 30% HIFP in DCM.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5μm, 100Å, 4.6x100 mm column.

Final yield: 47%

Purity (HPLC): 98%

Retention Time: 27.28 min

Molecular Formula: C₇₉H₁₀₇N₁₃O₁₇S₂, [M+H]⁺_{Calculated}=1574.9

MS: [M+H]⁺=1574.7; [M+2H]²⁺=761.3



4.5.13. Synthesis of Fmoc-[Ala^{12,13}, Nle¹⁴]-11-17-OH peptide

Sequence:

Fmoc-Tyr-Ala-Ala-Nle-Ala-Gly-Gly-OH

Resin		Resin Loading (mmol/g)	Resin (g)	Scale (mmol)			
Gly-CTC-OH		1.1	0.250	0.275			
Step	Residue	Exc. (eq)	Fmoc-AA-OH Coupling (mg)	Activating Agent 1 (g)	Activating Agent 2 (μL)		
1	G (Gly)	3	245.47	Oxyma 0.1172	DIC	127.7	
2	F (Phe)	3	319.61	Oxyma 0.1172	DIC	127.7	
3	nL (Nle)	3	291.56	Oxyma 0.1172	DIC	127.7	
4	A (Ala)	3	256.82	Oxyma 0.1172	DIC	127.7	
5	A (Ala)	3	256.82	Oxyma 0.1172	DIC	127.7	
6	Y (Tyr)	3	379.17	Oxyma 0.1172	DIC	127.7	

The peptide synthesis was carried out manually.

Double coupling was performed for the following amino acids: Tyr¹¹, Ala¹², Ala¹³.

Fmoc deprotection was performed using 20% piperidine solution in DMF. The peptide remained on the resin maintaining the final N-terminus Fmoc group protection. At the end of the synthesis, the sequence was cleaved from the resin using a solution of 30% HIFP in DCM.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5μm, 100Å, 4.6x100 mm column.

Final yield: 53%

Purity (HPLC): 93%

Retention Time: 24.64 min

Molecular Formula: C₅₃H₆₅N₇O₁₁, [M+H]⁺ Calculated=976.5

MS: [M+H]⁺=976.4

4.5.14. Synthesis of H-[Nle¹⁴]-11-30-NH₂ peptide

Sequence:

H-Tyr-Arg-Arg-Nle-Phe-Gly-Gly-Pro-Gly-Thr-Ala-Ser-Arg-Pro-Ser-Ser-Ser-Arg-Ser-Tyr-NH₂

Fmoc deprotection was performed using 20% piperidine solution in DMF.

The peptide was obtained through a coupling reaction involving the protected fragment Fmoc-[Nle¹⁴]-11-17-OH and the fragment H-18-30-NH₂ on the resin. This reaction was carried out over a period

Experimental Part



of four hours in the presence of 2eq of HOAt and an equimolar amount of DIC. After two hours, an additional two equivalents of DIC were added.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5 μ m, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using Luna C18, 300Å, 10 μ m column. Increasing amounts of the crude product were injected progressively over the course of three consecutive runs. The elution was conducted with a gradient of 5-35% B over 30 minutes.

The fractions displaying the highest purity were pooled, acetonitrile was removed under reduced pressure, and the peptide was subsequently lyophilised.

Final yield after purification: 7%

Purity (HPLC): 99%

Retention Time: 10.1 min

Molecular Formula: C₉₅H₁₄₉N₃₃O₂₈, [M+H]⁺_{Calculated}=2201.4

MS: [M+H]⁺= 2201.0; [M+2H]²⁺=1100.5; [M+3H]³⁺=734.2; [M+4H]⁴⁺=551.0

4.5.15. Synthesis of H-[Ala¹¹, Nle¹⁴]-11-30-NH₂ peptide

Sequence:

H-Ala-Arg-Arg-Nle-Phe-Gly-Gly-Pro-Gly-Thr-Ala-Ser-Arg-Pro-Ser-Ser-Ser-Arg-Ser-Tyr-NH₂

Fmoc deprotection was performed using 20% piperidine solution in DMF.

The peptide was obtained through a coupling reaction involving the protected fragment Fmoc-[Ala¹¹, Nle¹⁴]-11-17-OH and the fragment Fmoc-18-30-NH₂ on the resin. This reaction was carried out over a period of four hours in the presence of 2eq of HOAt and an equimolar amount of DIC. After two hours, an additional two equivalents of DIC were added.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5 μ m, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using Luna C18, 300Å, 10 μ m column. Increasing amounts of the crude product were injected progressively over the course of three consecutive runs. The elution was conducted with a gradient of 5-35% B over 30 minutes.

The fractions displaying the highest purity were pooled, acetonitrile was removed under reduced pressure, and the peptide was subsequently lyophilised.

Final yield after purification: 17%

Purity (HPLC): 99%



Retention Time: 9.38 min

Molecular Formula: $C_{89}H_{145}N_{33}O_{27}$, $[M+H]^+_{\text{Calculated}}=2108.1$

MS: $[M+H]^+=2108.9$; $[M+2H]^{2+}=1055.0$; $[M+3H]^{3+}=703.5$; $[M+4H]^{4+}=528.0$

4.5.16. Synthesis of H-[Ala¹⁵, Nle¹⁴]-11-30-NH₂ peptide

Sequence:

H-Tyr-Arg-Arg-Nle-Ala-Gly-Gly-Pro-Gly-Thr-Ala-Ser-Arg-Pro-Ser-Ser-Ser-Arg-Ser-Tyr-NH₂

Fmoc deprotection was performed using 20% piperidine solution in DMF.

The peptide was obtained through a coupling reaction involving the protected fragment Fmoc-[Ala¹⁵, Nle¹⁴]-11-17-OH and the fragment H-18-30-NH₂ on the resin. This reaction was carried out over a period of four hours in the presence of 2eq of HOAt and an equimolar amount of DIC. After two hours, an additional two equivalents of DIC were added.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5μm, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using Luna C18, 300Å, 10μm column. Increasing amounts of the crude product were injected progressively over the course of three consecutive runs. The elution was conducted with a gradient of 5-35% B over 30 minutes.

The fractions displaying the highest purity were pooled, acetonitrile was removed under reduced pressure, and the peptide was subsequently lyophilised.

Final yield after purification: 7%

Purity (HPLC): 99%

Retention Time: 9.03 min

Molecular Formula: $C_{89}H_{145}N_{33}O_{28}$, $[M+H]^+_{\text{Calculated}}=2125.2$

MS: $[M+H]^+=2124.9$; $[M+2H]^{2+}=1062.5$; $[M+3H]^{3+}=709.0$

4.5.17. Synthesis of H-[Ala³⁰, Nle¹⁴]-11-30-NH₂ peptide

Sequence:

H-Tyr-Arg-Arg-Nle-Phe-Gly-Gly-Pro-Gly-Thr-Ala-Ser-Arg-Pro-Ser-Ser-Ser-Arg-Ser-Ala-NH₂

Fmoc deprotection was performed using 20% piperidine solution in DMF.

The peptide was obtained through a coupling reaction involving the protected fragment Fmoc-[Nle¹⁴]-11-17-OH and the fragment H-[Ala³⁰]-18-30-NH₂ on the resin. This reaction was carried out over a period of four hours in the presence of 2eq of HOAt and an equimolar amount of DIC. After two hours, an additional two equivalents of DIC were added.

Experimental Part



The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5 μ m, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using Luna C18, 300Å, 10 μ m column. Increasing amounts of the crude product were injected progressively over the course of three consecutive runs. The elution was conducted with a gradient of 5-35% B over 30 minutes.

The fractions displaying the highest purity were pooled, acetonitrile was removed under reduced pressure, and the peptide was subsequently lyophilised.

Final yield after purification: 10%

Purity (HPLC): 98%

Retention Time: 10.22 min

Molecular Formula: C₈₉H₁₄₅N₃₃O₂₇, [M+H]⁺_{Calculated}=2109.3

MS: [M+H]⁺= 2108.9; [M+2H]²⁺=1055.0; [M+3H]³⁺= 703.5

4.5.18. Synthesis of H-[Ala^{12,13,23,28}, Nle¹⁴]-11-30-NH₂ peptide

Sequence:

H-Tyr-Ala-Ala-Nle-Phe-Gly-Gly-Pro-Gly-Thr-Ala-Ser-Ala-Pro-Ser-Ser-Ser-Ala-Ser-Tyr-NH₂

Fmoc deprotection was performed using 20% piperidine solution in DMF.

The peptide was obtained through a coupling reaction involving the protected fragment Fmoc-[Ala^{12,13}, Nle¹⁴]-11-17-OH and the fragment H-[Ala²³, Ala²⁸]-18--30-NH₂ on the resin. This reaction was carried out over a period of four hours in the presence of 2eq of HOAt and an equimolar amount of DIC. After two hours, an additional two equivalents of DIC were added.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5 μ m, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using Luna C18, 300Å, 10 μ m column. Increasing amounts of the crude product were injected progressively over the course of three consecutive runs. The elution was conducted with a gradient of 5-35% B over 30 minutes.

The fractions displaying the highest purity were pooled, acetonitrile was removed under reduced pressure, and the peptide was subsequently lyophilised.

Final yield after purification: 9%

Purity (HPLC): 99%

Retention Time: 11.94 min

Molecular Formula: C₈₃H₁₂₁N₂₁O₂₈, [M+H]⁺_{Calculated}=1861.0



MS: $[M+H]^+ = 1860.8$

4.5.19. Synthesis of H-[Ala^{12,13}, Nle¹⁴]-11-30-NH₂ peptide

Sequence:

H-Tyr-Ala-Ala-Nle-Phe-Gly-Gly-Pro-Gly-Thr-Ala-Ser-Arg-Pro-Ser-Ser-Ser-Arg-Ser-Tyr-NH₂

Fmoc deprotection was performed using 20% piperidine solution in DMF.

The peptide was obtained through a coupling reaction involving the protected fragment Fmoc--[Ala^{12,13}, Nle¹⁴]-11-17-OH and the fragment H-18-30-NH₂ on the resin. This reaction was carried out over a period of four hours in the presence of 2eq of HOAt and an equimolar amount of DIC. After two hours, an additional two equivalents of DIC were added.

The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5μm, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using Luna C18, 300Å, 10μm column. Increasing amounts of the crude product were injected progressively over the course of three consecutive runs. The elution was conducted with a gradient of 5-35% B over 30 minutes.

The fractions displaying the highest purity were pooled, acetonitrile was removed under reduced pressure, and the peptide was subsequently lyophilised.

Final yield after purification: 19%

Purity (HPLC): 98%

Retention Time: 10.94 min

Molecular Formula: C₈₉H₁₃₅N₂₇O₂₈, $[M+H]^+_{\text{Calculated}} = 2031.2$

MS: $[M+H]^+ = 2030.8$; $[M+2H]^{2+} = 1015.9$

4.5.20. Synthesis of H-[Ala^{23,28}, Nle¹⁴]-11-30-NH₂ peptide

Sequence:

H-Tyr-Arg-Arg-Nle-Phe-Gly-Gly-Pro-Gly-Thr-Ala-Ser-Ala-Pro-Ser-Ser-Ser-Ala-Ser-Tyr-NH₂

Fmoc deprotection was performed using 20% piperidine solution in DMF.

The peptide was obtained through a coupling reaction involving the protected fragment Fmoc-11-[Nle¹⁴]-11-17-OH and the fragment H-[Ala^{23,28}]-18-30-NH₂ on the resin. This reaction was carried out over a period of four hours in the presence of 2eq of HOAt and an equimolar amount of DIC. After two hours, an additional two equivalents of DIC were added.

Experimental Part



The presence of the desired peptide in the crude mixture was confirmed by HPLC-MS analysis under the following conditions: 5-95% B in 30 min on a Phenomenex Kinetex C18, 3.5 μ m, 100Å, 4.6x100 mm column.

Purification was performed by preparative HPLC using Luna C18, 300Å, 10 μ m column. Increasing amounts of the crude product were injected progressively over the course of three consecutive runs. The elution was conducted with a gradient of 5-35% B over 30 minutes.

The fractions displaying the highest purity were pooled, acetonitrile was removed under reduced pressure, and the peptide was subsequently lyophilised.

Final yield after purification: 10%

Purity (HPLC): 99%

Retention Time: 10.94min

Molecular Formula: C₈₉H₁₃₅N₂₇O₂₈, [M+H]⁺_{Calculated}=2031.2

MS: [M+H]⁺=2030.8; [M+2H]²⁺=1015.9

4.6. Conformational Studies by CD Spectroscopy

CD spectra were recorded on a Jasco-1500 spectrophotometer using Spectra Manager (Jasco) software. CD spectra were recorded at 25°C using a 0.1 cm quartz cell, a 260–190 nm measurement range, 50 nm/min scanning speed, 1 nm bandwidth, 4s response time, 0.5 nm data pitch and are the results of 4 scans. The measurements are taken in duplicate.

4.7. NMR Spectroscopy

NMR Spectroscopy 2D ¹H NMR (TOCSY and NOESY) spectra were recorded on a Bruker Avance DMX 600 and 400 MHz. The solvent used for the analysis is TFE, d₂. The spectra were processed on Sparky software.[74]

4.8. Electrophoretic Mobility Shift Assay

The reaction mixtures (20 μ) contained 5 mM TRIS-HCl (pH 8.4), 150 mM KCl, 8 μ M Vimentin and an increasing concentration of peptides, as indicated. The mixtures were incubated for 30 minutes. Following this, 4 μ L of a dye mixture (SYBR Green I or Coomassie Brilliant Blue) was added and the samples were immediately loaded onto a 1.5% Agarose gel in 0.5x TBE buffer. The gel was then electrophoresed in one dimension at 70V and 25°C (without cooling) for 45 minutes.



4.9. Fluorescence Thermal Assay

The thermal stability of fluorescently labelled DNA was determined in a Roche LightCycler, using an excitation source at 465 nm and recording the fluorescence emission at 510 nm. As intramolecular G4 forming model we used TEL23_FD; for double-stranded model DNA (dsDNA), the two complementary strands Scr22_F and Coscr22_D which were previously annealed at an equimolar concentration. Both models were designed to locate a fluorophore (6-FAM) and a quencher (Dabcyl) nearby when folded into the expected secondary structure thus resulting in a quenching of the FAM fluorescence signal. Reaction mixtures (20 μ l final volume) were assembled on 96-well plates. They contained 0.5 μ M DNA in 10 mM $\text{Li}_3(\text{PO})_4$, 50 mM KCl, pH 7.5 and increasing concentrations of the tested peptides (0-80 μ M). For competition assays, 0.5 μ M TEL23_FD with 0, 5 or 40 μ M peptide H-11-30-NH₂ were mixed with increasing concentrations of Scr22 in the above buffer. TEL23_FD samples were preliminarily heated and cooled in the 25-95 $^{\circ}\text{C}$ range at a 0.1 $^{\circ}\text{C}/\text{s}$ rate.

Data were derived from heating and annealing steps run at a 1 $^{\circ}\text{C}/\text{min}$ rate. Melting temperatures (T_m) were determined from the first derivatives of the slower melting profiles using the Roche LightCycler software. Each curve was repeated three times. For each oligonucleotide, ΔT_m was calculated as the variation of the T_m induced by the presence of the peptide.

4.10. Surface Plasmon Resonance

Surface Plasmon Resonance (SPR) measurements were performed on a Biacore X100. A streptavidin-coated sensor chip was first conditioned by flushing for 1 min a solution of 1 M NaCl, 50 mM NaOH in 50% isopropanol, and finally extensively washed with running buffer (5 mM Tris, 50 mM KCl, pH 8.4 with 0.025% Tween20). Previously annealed 5'-biotinylated oligonucleotides were immobilized on one cell of the chip surface by flowing a 50 nM DNA solution at a 10 μ l/min flow rate until the required Response Units (RU) were reached (RU 200 for 2TEL_BIOT and Scr44_BIOT; RU 400 for TEL_BIOT). A second cell was left blank as control. Sensorgrams were acquired using serial dilutions of peptide (0-40 μ M), which were injected at a 10 μ l/min flow rate for 100 s. After each run, a 30 s regeneration step was performed with 10 mM glycine buffer, pH 2.5, followed by a 100 s stabilization period with running buffer. Before use, all buffers and solutions were previously filtered on a 0.22 μ m (MCE Membrane Millipore TM). The RU detected at the equilibrium, were converted into relative RU (rRU) based on the maximal value predicted by the experimental RU variation and plotted as a function of peptide concentrations. These data were fitted to derive the peptide concentration required to produce 50% of the rRU variation (EC 50). Each data point was repeated at least in triplicate.

References

1. Hohmann, T. and F. Dehghani, *The Cytoskeleton-A Complex Interacting Meshwork*. Cells, 2019. **8**(4).
2. Gimeno-Agud, H., Y. Díaz-Osorio, and A. Oyarzábal, *Biological Basis of Cell Trafficking: A General Overview*. J Inherit Metab Dis, 2025. **48**(1): p. e12839.
3. Sachana, M., J. Flaskos, and A.J. Hargreaves, *Chapter 19 - In vitro biomarkers of developmental neurotoxicity*, in *Reproductive and Developmental Toxicology*, R.C. Gupta, Editor. 2011, Academic Press: San Diego. p. 227-252.
4. Slack, J.M.W., *Chapter 7 - Molecular Biology of the Cell*, in *Principles of Tissue Engineering (Fourth Edition)*, R. Lanza, R. Langer, and J. Vacanti, Editors. 2014, Academic Press: Boston. p. 127-145.
5. Logan, C.M. and A.S. Menko, *Microtubules: Evolving roles and critical cellular interactions*. Exp Biol Med (Maywood), 2019. **244**(15): p. 1240-1254.
6. Coelho-Rato, L.S., et al., *Intermediate filaments at a glance*. Journal of Cell Science, 2024. **137**(16).
7. Di, X., et al., *Cellular mechanotransduction in health and diseases: from molecular mechanism to therapeutic targets*. Signal Transduction and Targeted Therapy, 2023. **8**(1): p. 282.
8. Sivagurunathan, S., et al., *Expression of vimentin alters cell mechanics, cell-cell adhesion, and gene expression profiles suggesting the induction of a hybrid EMT in human mammary epithelial cells*. 2022. **Volume 10 - 2022**.
9. Moujaber, O. and U. Stochaj, *The Cytoskeleton as Regulator of Cell Signaling Pathways*. Trends in Biochemical Sciences, 2020. **45**(2): p. 96-107.
10. Eldirany, S.A., et al., *Recent insight into intermediate filament structure*. Current Opinion in Cell Biology, 2021. **68**: p. 132-143.
11. Huang, X., et al., *The unconventional role of vimentin intermediate filaments*. Current Opinion in Cell Biology, 2025. **93**: p. 102483.
12. Chernyatina, A.A., D. Guzenko, and S.V. Strelkov, *Intermediate filament structure: the bottom-up approach*. Curr Opin Cell Biol, 2015. **32**: p. 65-72.
13. Steinert, P.M., *Structure, function, and dynamics of keratin intermediate filaments*. J Invest Dermatol, 1993. **100**(6): p. 729-34.

14. Steinert, P.M., L.N. Marekov, and D.A. Parry, *Diversity of intermediate filament structure. Evidence that the alignment of coiled-coil molecules in vimentin is different from that in keratin intermediate filaments*. J Biol Chem, 1993. **268**(33): p. 24916-25.
15. Patteson, A.E., et al., *The vimentin cytoskeleton: when polymer physics meets cell biology*. Phys Biol, 2020. **18**(1): p. 011001.
16. Turner, M.D., et al., *Cytokines and chemokines: At the crossroads of cell signalling and inflammatory disease*. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 2014. **1843**(11): p. 2563-2582.
17. Alonso-Matilla, R., P.P. Provenzano, and D.J. Odde, *Physical principles and mechanisms of cell migration*. npj Biological Physics and Mechanics, 2025. **2**(1): p. 2.
18. Chen, T., et al., *Epithelial-mesenchymal transition (EMT): A biological process in the development, stem cell differentiation, and tumorigenesis*. J Cell Physiol, 2017. **232**(12): p. 3261-3272.
19. Guo, M., et al., *Vimentin intermediate filaments as structural and mechanical coordinators of mesenchymal cells*. Nature Cell Biology, 2025. **27**(8): p. 1210-1218.
20. Yuan, Z., P.A. Janmey, and C.A. McCulloch, *Structure and function of vimentin in the generation and secretion of extracellular vimentin in response to inflammation*. Cell Communication and Signaling, 2025. **23**(1): p. 187.
21. Satelli, A. and S. Li, *Vimentin in cancer and its potential as a molecular target for cancer therapy*. Cell Mol Life Sci, 2011. **68**(18): p. 3033-46.
22. Kwak, H.-I., et al., *Calpain-mediated vimentin cleavage occurs upstream of MT1-MMP membrane translocation to facilitate endothelial sprout initiation*. Angiogenesis, 2012. **15**(2): p. 287-303.
23. Nieminen, M., et al., *Vimentin function in lymphocyte adhesion and transcellular migration*. Nature Cell Biology, 2006. **8**(2): p. 156-162.
24. Jang, J., et al., *Vimentin-mediated buffering of internal integrin β 1 pool increases survival of cells from anoikis*. BMC Biology, 2024. **22**(1): p. 139.
25. van Beijnum, J.R., et al., *Extracellular vimentin mimics VEGF and is a target for anti-angiogenic immunotherapy*. Nature Communications, 2022. **13**(1): p. 2842.
26. Tabatabaee, A., et al., *Targeting vimentin: a multifaceted approach to combatting cancer metastasis and drug resistance*. Cancer and Metastasis Reviews, 2024. **43**(1): p. 363-377.
27. Tolstonog, G.V., et al., *Isolation of SDS-stable complexes of the intermediate filament protein vimentin with repetitive, mobile, nuclear matrix attachment region, and mitochondrial DNA*

References

- sequence elements from cultured mouse and human fibroblasts. DNA Cell Biol, 2001. 20(9): p. 531-54.*
28. Largy, E., J.L. Mergny, and V. Gabelica, *Role of Alkali Metal Ions in G-Quadruplex Nucleic Acid Structure and Stability. Met Ions Life Sci, 2016. 16: p. 203-58.*
29. Ceschi, S., et al., *Vimentin binds to G-quadruplex repeats found at telomeres and gene promoters. Nucleic Acids Res, 2022. 50(3): p. 1370-1381.*
30. Cozzaglio, M., et al., *G-quadruplexes formation within the promoter of TEAD4 oncogene and their interaction with Vimentin. Front Chem, 2022. 10: p. 1008075.*
31. Watson, J.D. and F.H. Crick, *Genetical implications of the structure of deoxyribonucleic acid. Nature, 1953. 171(4361): p. 964-7.*
32. Yang, D. and K. Okamoto, *Structural insights into G-quadruplexes: towards new anticancer drugs. Future Med Chem, 2010. 2(4): p. 619-46.*
33. Spiegel, J., S. Adhikari, and S. Balasubramanian, *The Structure and Function of DNA G-Quadruplexes. Trends in Chemistry, 2020. 2(2): p. 123-136.*
34. Wu, F., et al., *Genome-wide analysis of DNA G-quadruplex motifs across 37 species provides insights into G4 evolution. Communications Biology, 2021. 4(1): p. 98.*
35. Figueiredo, J., J.-L. Mergny, and C. Cruz, *G-quadruplex ligands in cancer therapy: Progress, challenges, and clinical perspectives. Life Sciences, 2024. 340: p. 122481.*
36. De-Paula, R.B., et al., *Enriched G4 forming repeats in the human genome are associated with robust well-coordinated transcription and reduced cancer transcriptome variation. J Biol Chem, 2024. 300(11): p. 107822.*
37. Kim, N., *The Interplay between G-quadruplex and Transcription. Curr Med Chem, 2019. 26(16): p. 2898-2917.*
38. Xu, Y. and M. Komiyama, *G-Quadruplexes in Human Telomere: Structures, Properties, and Applications. 2024. 29(1): p. 174.*
39. Bahls, B., et al., *G-Quadruplexes in c-MYC Promoter as Targets for Cancer Therapy. Biomedicines, 2023. 11(3).*
40. Saad, M., et al., *G-quadruplex forming sequences in the genes coding for cytochrome P450 enzymes and their potential roles in drug metabolism. Biochimie, 2023. 214: p. 45-56.*
41. Aguilar-Santelises, M., et al., *Bcl-2, Bax and p53 expression in B-CLL in relation to in vitro survival and clinical progression. Int J Cancer, 1996. 69(2): p. 114-9.*

42. Monsen, R.C., et al., *Long promoter sequences form higher-order G-quadruplexes: an integrative structural biology study of c-Myc, k-Ras and c-Kit promoter sequences*. Nucleic Acids Res, 2022. **50**(7): p. 4127-4147.
43. Chalikian, T.V., L. Liu, and J.R.B. Macgregor, *Duplex-tetraplex equilibria in guanine- and cytosine-rich DNA*. Biophysical Chemistry, 2020. **267**: p. 106473.
44. Ruzi, Z., et al., *RNA G-quadruplexes: emerging regulators of gene expression and therapeutic targets*. Functional & Integrative Genomics, 2025. **25**(1): p. 143.
45. Bugaut, A. and S. Balasubramanian, *5'-UTR RNA G-quadruplexes: translation regulation and targeting*. Nucleic Acids Res, 2012. **40**(11): p. 4727-41.
46. Cohen, S.B. and T.M. Bryan, *Telomeres: Structure and Function*, in *Handbook of Chemical Biology of Nucleic Acids*, N. Sugimoto, Editor. 2022, Springer Nature Singapore: Singapore. p. 1-17.
47. de Lange, T., *Shelterin: the protein complex that shapes and safeguards human telomeres*. Genes Dev, 2005. **19**(18): p. 2100-10.
48. Victorelli, S. and J.F. Passos, *Telomeres and Cell Senescence - Size Matters Not*. EBioMedicine, 2017. **21**: p. 14-20.
49. Fan, H.C., et al., *Telomeres and Cancer*. Life (Basel), 2021. **11**(12).
50. Boccardi, V. and L. Marano, *Aging, Cancer, and Inflammation: The Telomerase Connection*. 2024. **25**(15): p. 8542.
51. Webb, C.J., Y. Wu, and V.A. Zakian, *DNA repair at telomeres: keeping the ends intact*. Cold Spring Harb Perspect Biol, 2013. **5**(6).
52. Dell'Oca, M.C., et al., *Spotlight on G-Quadruplexes: From Structure and Modulation to Physiological and Pathological Roles*. Int J Mol Sci, 2024. **25**(6).
53. Kharel, P. and P. Ivanov, *Structure and Functions of RNA G-quadruplexes*, in *RNA Structure and Function*, J. Barciszewski, Editor. 2023, Springer International Publishing: Cham. p. 183-203.
54. Obara, P., P. Wolski, and T. Pańczyk, *Insights into the Molecular Structure, Stability, and Biological Significance of Non-Canonical DNA Forms, with a Focus on G-Quadruplexes and i-Motifs*. Molecules, 2024. **29**(19).
55. Chen, Y. and D. Yang, *Sequence, stability, and structure of G-quadruplexes and their interactions with drugs*. Curr Protoc Nucleic Acid Chem, 2012. **Chapter 17**: p. Unit17.5.

References

56. Nishio, M., K. Tsukakoshi, and K. Ikebukuro, *G-quadruplex: Flexible conformational changes by cations, pH, crowding and its applications to biosensing*. Biosensors and Bioelectronics, 2021. **178**: p. 113030.
57. Parkinson, G.N., M.P. Lee, and S. Neidle, *Crystal structure of parallel quadruplexes from human telomeric DNA*. Nature, 2002. **417**(6891): p. 876-80.
58. Luu, K.N., et al., *Structure of the human telomere in K⁺ solution: an intramolecular (3 + 1) G-quadruplex scaffold*. J Am Chem Soc, 2006. **128**(30): p. 9963-70.
59. Dai, J., et al., *Structure of the Hybrid-2 type intramolecular human telomeric G-quadruplex in K⁺ solution: insights into structure polymorphism of the human telomeric sequence*. Nucleic Acids Res, 2007. **35**(15): p. 4927-40.
60. Ambrus, A., et al., *Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution*. Nucleic Acids Res, 2006. **34**(9): p. 2723-35.
61. Hänsel, R., et al., *High-Resolution Insight into G-Overhang Architecture*. Journal of the American Chemical Society, 2013. **135**(7): p. 2816-2824.
62. Neidle, S. and G.N. Parkinson, *The structure of telomeric DNA*. Current Opinion in Structural Biology, 2003. **13**(3): p. 275-283.
63. Monsen, R.C., J.O. Trent, and J.B. Chaires, *G-quadruplex DNA: A Longer Story*. Accounts of Chemical Research, 2022. **55**(22): p. 3242-3252.
64. Yang, Y., *Chapter 6 - Intramolecular Cyclization Side Reactions*, in *Side Reactions in Peptide Synthesis*, Y. Yang, Editor. 2016, Academic Press: Oxford. p. 119-161.
65. Yang, Y. and L. Hansen, *Optimized Fmoc-Removal Strategy to Suppress the Traceless and Conventional Diketopiperazine Formation in Solid-Phase Peptide Synthesis*. ACS Omega, 2022. **7**(14): p. 12015-12020.
66. Buck, M., *Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins*. Q Rev Biophys, 1998. **31**(3): p. 297-355.
67. Batkishig, D., et al., *A crucial residue in the hydrophobic core of the solenoid structure of leucine rich repeats*. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 2021. **1869**(6): p. 140631.
68. Biondi, B., et al., *Structural and biological characterization of shortened derivatives of the cathelicidin PMAP-36*. Sci Rep, 2023. **13**(1): p. 15132.
69. Wüthrich, K. *NMR of proteins and nucleic acids*. 1988.
70. Keeler, J., *Understanding NMR Spectroscopy*. 2013.

71. Seo, M., L. Lei, and M. Egli, *Label-Free Electrophoretic Mobility Shift Assay (EMSA) for Measuring Dissociation Constants of Protein-RNA Complexes*. *Curr Protoc Nucleic Acid Chem*, 2019. **76**(1): p. e70.
72. Yang, R., et al., *Fluorescence labeling-based differential scanning fluorimetry, an effective method for protein thermal stability and protein-compound binding analysis*. *International Journal of Biological Macromolecules*, 2024. **281**: p. 136043.
73. Butt, M.A., *Surface Plasmon Resonance-Based Biodetection Systems: Principles, Progress and Applications—A Comprehensive Review*. 2025. **15**(1): p. 35.
74. Lee, W., M. Tonelli, and J.L. Markley, *NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy*. *Bioinformatics*, 2015. **31**(8): p. 1325-7.

CHAPTER III

Structural Characterisation of G4-Ligand Interactions Using Synchrotron Radiation Circular Dichroism



1. Introduction

This chapter examines the interaction between G-quadruplex DNA structures and two complementary classes of ligand: a library of tetrazole-based small molecules and the G4-binding peptide Rhau25. The small-molecule ligands were provided by Dr Pietro Spanu's research group at the Institute of Biomolecular Chemistry in Sassari. The Rhau25 peptide was provided by Dr Paolo Ruzza's research group at the Institute of Biomolecular Chemistry in Padova. The binding properties of these ligands were systematically evaluated using three representative G4-forming sequences (Htelo1, G3T3, and T95-2T) in both sodium- and potassium-rich environments, through high-throughput synchrotron radiation circular dichroism experiments conducted at the B23 beamline (Diamond Light Source, UK).

1.1. Circular Dichroism: Physical Principles and Theoretical Foundations

Circular Dichroism (CD) is a spectroscopic technique that measures the differential absorption of left- and right-handed circularly polarised light (CPL) by optically active molecules (Figure 3.1). CPL can be generated by passing non-polarised light (NPL) through a linear polariser to produce linearly polarised light (LPL), which then passes through a quarter-wave plate to become circularly polarised. In CPL, the electric field vector rotates around the propagation axis with constant amplitude.[1]

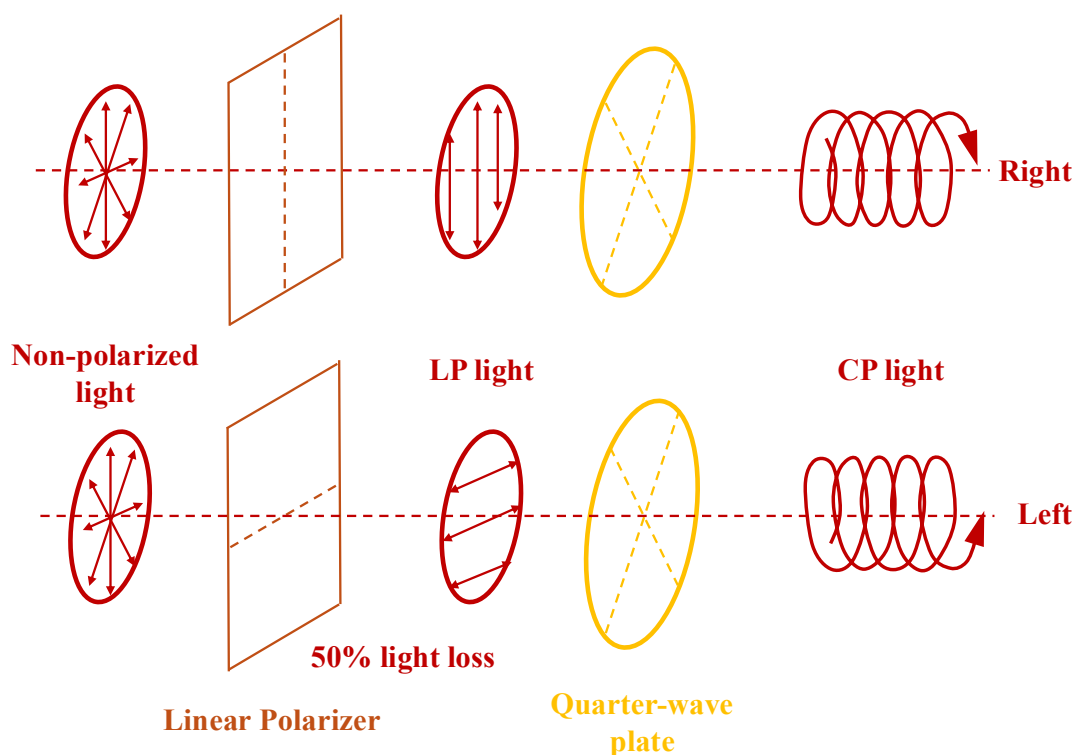


Figure 3.1. Schematic diagram of the physical method for obtaining CPL.



Both LPL and CPL are chiral entities because they consist of two non-superimposable mirror images of equal intensity. CD is closely related to optical rotation, whereby the plane of polarisation rotates when CPL passes through a chiral medium. In such a medium the two circular components are absorbed differently and, upon recombination, produce elliptically polarised light. The difference in absorption is described by the Lambert Beer law:[2]

$$\Delta A = A_L - A_R = \Delta \epsilon * l * c$$

where ΔA is the differential absorption signal, A_L and A_R are the absorption of left- and right-handed circularly polarised light, respectively, $\Delta \epsilon$ is the difference in the molar extinction coefficients, l is the optical path length, and c is the concentration. However, CD spectra are typically reported in terms of ellipticity (θ), which is related to ΔA by the following equation:[2]

$$\theta = 33\Delta A$$

The CD spectrum, which is obtained by plotting ΔA , θ or $\Delta \epsilon$ as a function of wavelength, is recorded using an instrument similar to a UV-Vis spectrophotometer. This instrument has an additional light polariser and a photoelastic modulator (PEM), which are placed between the monochromator and the sample.

The PEM alternately produces left- and right- handed circularly polarised light. Measurements are usually performed under a nitrogen atmosphere to prevent interference from oxygen (Figure 3.2).[2]

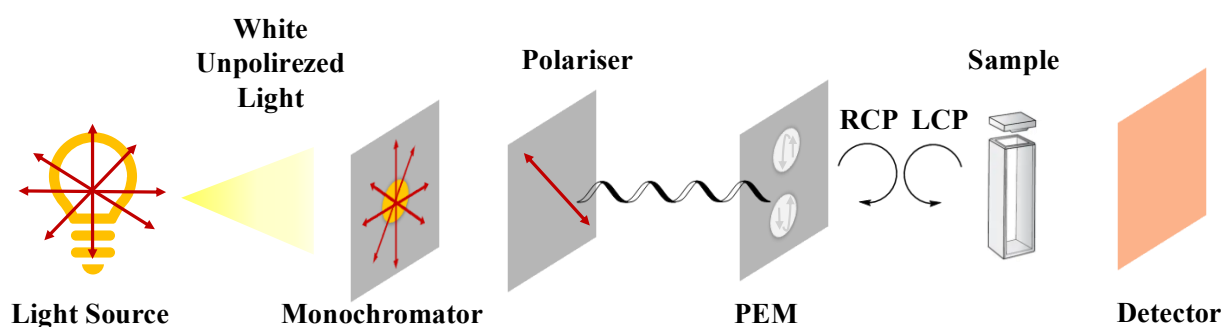


Figure 3.2. CD instrumentation scheme.



1.2. Applications of CD in Secondary Structure Analysis

1.2.1. Peptides and Proteins

For peptides and proteins, CD spectra are normalised to remove dependence on concentration and path length using the following equation:[3]

$$[\theta]_{MRW} = \frac{\theta * M}{c * l * n_{(r-1)}}$$

where $[\theta]_{MRW}$ is the mean residue ellipticity, M is the molecular weight, c is the concentration, l is the path length, and $n_{(r-1)}$ corresponds to the number of amide bonds, which is equal to the number of residues minus one for linear peptides.

CD spectroscopy is a powerful and rapid technique for investigating the secondary structure of biomolecules and their conformational changes as a function of temperature or in response to chemical agents. In order for a molecule to be CD-active, it must contain an intrinsic chromophore. Peptides and proteins can be chiral molecules by nature, or they can exhibit induced chirality when an achiral chromophore interacts with a nearby chiral centre. The most relevant chromophore in this context is the amide group, which is the most abundant in the peptide backbone. The optical activity of the amide group arises from two sources: the intrinsic chirality of the α -carbon (C_α) of each amino acid residue – except for glycine, which lacks chirality – and its insertion within an overall chiral structure, such as α -helices or β -sheets. Two spectral regions are typically used in peptide studies:[4]

- The far-UV (190-250 nm) is sensitive to the peptide backbone and provides information about secondary structure
- The near-UV region (250-350 nm) is sensitive to aromatic amino acids and disulfide bonds, reflecting tertiary structure

In the far-UV region range, the amide group acts as the main chromophore, giving rise to two electronic transitions:[4]

- The $n \rightarrow \pi^*$ transition, which appears as a negative band around 220 nm and is sensitive to solvent polarity
- The $\pi \rightarrow \pi^*$ transition, which appears as a positive band near 190 nm

Different secondary structures exhibit characteristic CD signatures (Figure 3.3):

- α -helices show two negative bands at 208 and 222 nm, as well as a positive band near 190 nm. The band at 208 and 190 nm arises from the splitting of the $\pi \rightarrow \pi^*$ transition into components polarised parallel ($\pi_{//}$) and perpendicular ($\pi_{\perp}$) to the helix axis, a process known as exciton splitting.
- β -sheets display a broad negative band in the 215-255 nm range and a positive band around 195 nm.



- random coils exhibit a single negative band near 200 nm.

In the near-UV region, signals originate from aromatic side chains (tryptophan, tyrosine, and phenylalanine) and disulfide bonds. While aromatic rings are intrinsically achiral, they become optically active in a chiral environment, reflecting tertiary packing and side-chain orientation. Due to the relatively low molar absorptivity of these chromophores ($\epsilon_{Trp,280} = 5500 \text{ M}^{-1}\text{cm}^{-1}$, $\epsilon_{Tyr,280} = 1490 \text{ M}^{-1}\text{cm}^{-1}$), higher sample concentration and longer optical paths are required.[5]

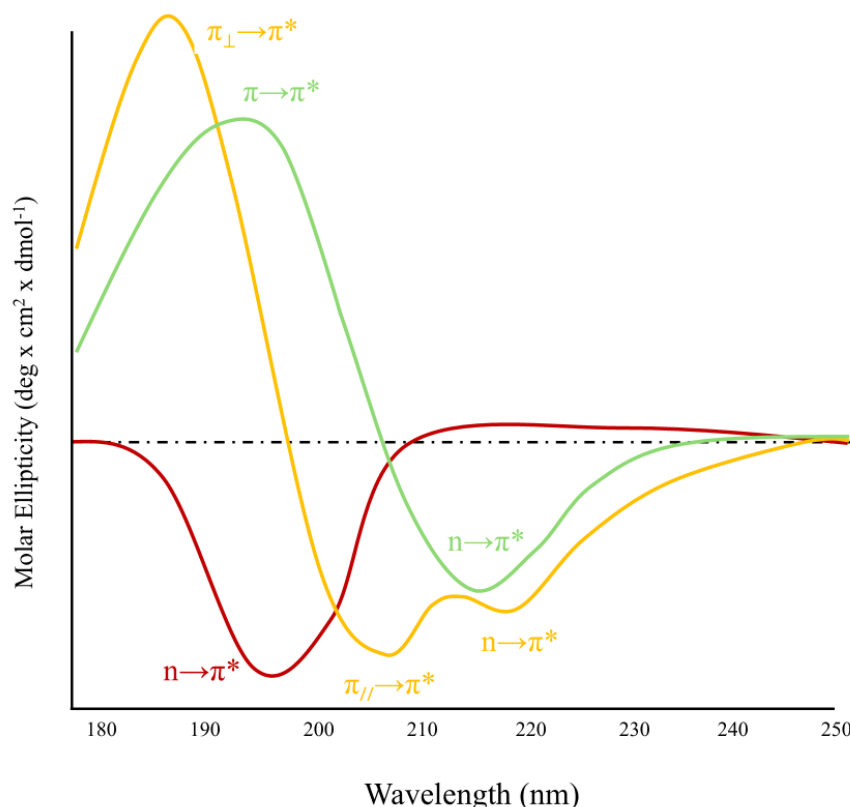


Figure 3.3. CD spectrum for an α -helix (yellow), β -sheet (green), and random coil (red).

The CD spectrum of a protein can be considered as the sum of the contributions from each secondary structure:[5]

$$[\theta_{MRE}] = f_{\alpha}[\theta]_{\alpha} + f_{\beta}[\theta]_{\beta} + f_T[\theta]_T + f_O[\theta]_O$$

Where $[\theta_{MRE}]$ is the measured mean residue ellipticity; $[\theta]_{\alpha}$, $[\theta]_{\beta}$, $[\theta]_T$, $[\theta]_O$ correspond to the reference ellipticities of the α -helix, β -sheet, and turn structures, as well as other structures (often random coil). The coefficients f represent the fractional content of each secondary structure.



1.2.2. Nucleic Acids

CD spectroscopy is equally valuable for investigating the structure and movement of nucleic acids. The optical activity of DNA and RNA derives from their chiral sugar-phosphate backbone and π -stacked nucleobases, which function as chromophores in the UV region of the spectrum.[6]

The CD spectra of nucleic acids are primarily governed by $\pi \rightarrow \pi^*$ electronic transitions in the aromatic bases (adenine, guanine, cytosine, thymine and uracil), which typically occur in the 220-300 nm range. These transitions are highly sensitive to base stacking, helical geometry and sequence context. Weaker $n \rightarrow \pi^*$ transitions may appear at longer wavelengths, particularly in regions rich in guanine and cytosine, but the $\pi \rightarrow \pi^*$ transitions dominate.[7]

Since isolated bases are planar and achiral, CD activity arises from exciton coupling between neighbouring chromophores within the helical framework of the backbone. Consequently, CD spectra provide direct insights into secondary structure, helix-coil transitions, and higher-order folding. Distinct DNA and RNA conformations exhibit characteristic CD profiles (Figure 3.4):[7]

- B-DNA (the physiological form) exhibits a positive band near 275-280 nm (base stacking) and a negative band around 245 nm (right-handed helix);
- A-DNA (dehydrated or RNA-like double helix) exhibits a strong positive band at ~260 nm and a negative band near 210 nm;
- Z-DNA (left-handed helix) exhibits a negative band around 295 nm and a positive band near 260 nm, which is indicative of reversed helicity;
- Single-stranded DNA/RNA exhibit weak and variable spectra due to the lack of ordered helicity and sequence-dependent stacking.

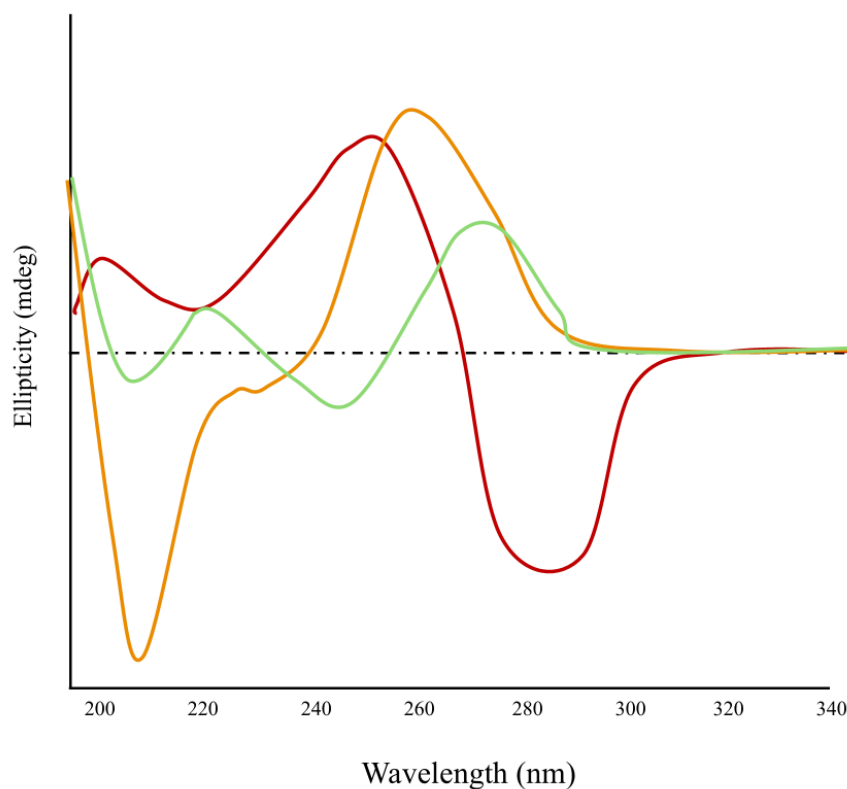


Figure 3.4. CD spectrum for an A-DNA (yellow), B-DNA (green), and Z-DNA (red).

CD spectra of nucleic acids are typically recorded between 320 and 220 nm, where $\pi \rightarrow \pi^*$ transitions related to the bases dominate.

1.2.3. G-Quadruplexes

The CD spectra of G4s are particularly sensitive to strand orientation and loop configuration, providing fingerprints for topology identification (Figure 3.5):[8]

- Parallel G4s present a strong positive band at ~260 nm and a negative band near 240 nm (all guanines in anti-conformation);
- Antiparallel G4s have a positive band around 295 nm and a negative one near 260 nm (guanines in syn-conformation);
- Hybrid (3+1) G4s have two positive bands at ~295 and ~265 nm, with a negative band near 240 nm, reflecting mixed orientations.

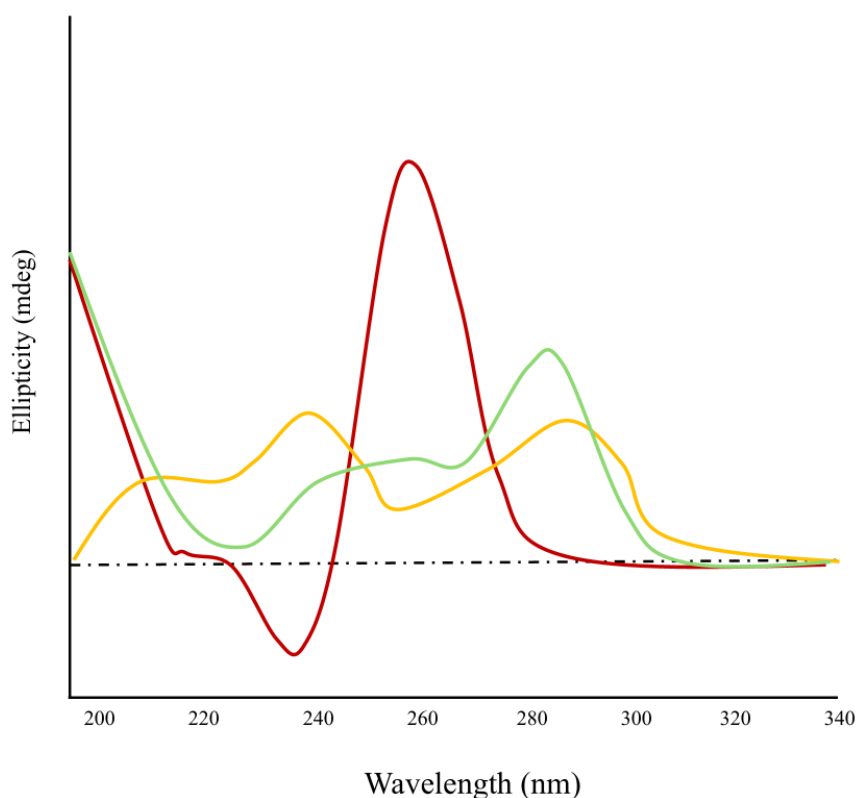


Figure 3.5. CD spectrum for an antiparallel (yellow), hybrid (green), and parallel (red) G-quadruplex.

These characteristic spectra result from $\pi \rightarrow \pi^*$ exciton coupling between stacked guanine bases. CD spectroscopy is an essential method for identifying G4 topology and monitoring structural changes upon variation of the type of cation, temperature, ligand binding, or protein interaction. It also provides insights into the folding equilibrium and stability of G4 motifs located in telomeres and gene promoters.[8]

1.3. Synchrotron Radiation: Principles and Properties

A synchrotron light source is a facility designed to produce extremely bright electromagnetic radiation (from infrared and ultraviolet to X-rays) by accelerating relativistic electrons and then forcing them to change direction using bending magnets or insertion devices, such as undulators or wigglers. As electrons travel near the speed of light, any deviation in their trajectory causes them to emit radiation. Relativistic and geometrical effects then compress this radiation into narrow cones, thereby increasing its directionality and brightness.[9]

Some of the key physico-chemical properties of synchrotron radiation include:[10]

- **High brilliance/intensity.** Synchrotron beams contain orders of magnitude more photons per unit of time, area, solid angle, and bandwidth than typical laboratory sources.



- **High polarisation.** Light can be polarised in a linear, circular or elliptical manner. This is particularly useful for techniques that depend on polarisation, such as circular dichroism and magnetic dichroism.
- **Wide tunability.** Monochromators can be used to adjust the settings of the insertion device, allowing a wide range of photon energies (wavelengths) to be selected, from far-UV or even vacuum-UV (below 190 nm) to hard X-rays.
- **Short pulses/time structure.** Synchrotron beams are often pulsed, enabling time-resolved studies, and the radiation is well collimated with low emittance.

These properties make synchrotron radiation especially valuable in structural biology, chemistry, materials science and physics. In the context of chiral biomolecules, the very high photon flux in the UV and vacuum-UV regions, and the ability to work with small sample volumes, are major advantages.[11]

1.4. HT-SRCD/Beamline B23 at Diamond Light Source

Beamline B23 at the Diamond Light Source in the UK is one of the world's leading facilities for synchrotron radiation circular dichroism (SRCD). It is equipped for high-throughput and advanced sample environments.

1.4.1. Conventional CD vs HT-SRCD

Conventional CD spectroscopy typically employs a xenon lamp and a single cuvette, requiring acquisition times of several minutes per sample.[12] This setup restricts both throughput and experimental flexibility, particularly when multiple conditions or ligands need to be tested. By contrast, the High-Throughput Synchrotron Radiation Circular Dichroism (HT-SRCD) system at B23 beamline of Diamond Light Source combines the exceptional properties of synchrotron radiation with advanced robotic automation to significantly enhance efficiency and analytical capabilities.[13]

Thanks to the intense, highly collimated and polarised synchrotron beam, measurements times are reduced to just a few seconds per sample and sample volumes are minimised to below 5 μ L. Automated sample changers and multi-position holders enable the simultaneous analysis of multiple experimental conditions, including ligand titrations, pH variations, temperature gradients and different buffer compositions, thereby providing an unparalleled capacity for rapid molecular screening.[14] These features establish HT-SRCD as an exceptionally powerful technique for structural biology and drug discovery, where high throughput and minimal sample consumption are essential.

Introduction



Compared to standard benchtop CD spectrometers, the B23 synchrotron beamline offers several remarkable advantages that significantly expand the reach and quality of CD spectroscopy experiments. The photon flux generated by synchrotron source is several orders of magnitude higher than that of conventional xenon lamps. This results in superior signal-to-noise ratios and faster acquisition times. It also provides access to the vacuum-UV region (below 190 nm), a spectral window that remains largely inaccessible to commercial instruments due to limitations in lamp technology and optical materials.[15]

Furthermore, the HT-SRCD system can operate with multi-cell sample plates or automated multi-position turrets. This minimises manual handling and reduces the risk of cross-contamination, while enabling the rapid screening of large numbers of conditions, within a single experimental session. In this setup, a vertical sample chamber has been designed specifically to accommodate multi-well plates for high-throughput analysis.[16] These plates are mounted on an X-Y motorised stage, enabling precise and automated positioning during measurements. Within this chamber, monochromatic light is directed upwards by a plane mirror, which reflects the incident beam towards the photo-elastic modulator that generates circularly polarised light before it passes through the sample (Figure 3.6, A). Using multi-well plates enables the simultaneous collection of multiple spectra, significantly accelerating data acquisition. This automated process is achieved through dedicated scripting that instruct the system to sequentially record spectra from each well, systematically progressing through the plate until all the positions have been analysed (Figure 3.6, B). This combination of automation, precision and synchrotron brilliance delivers unmatched throughput and data quality in chiroptical spectroscopy.[16]

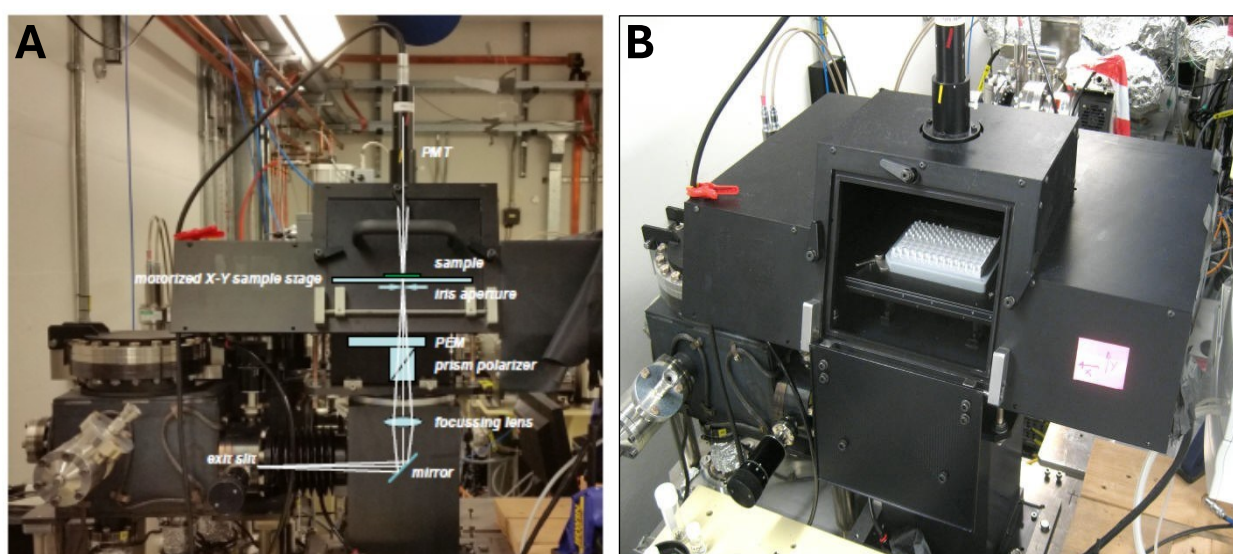


Figure 3.6. *A) Cartoon of the optical elements (exit slit, 45° plane mirror, focusing lens, polarizer prism, PEM, iris, X–Y motorized stage, horizontal sample and PMT detector). The sample is placed on the motorized X–Y stage. B) Picture of the 96-well plate held in the X–Y motorized stage inside the vertical sample chamber unit.[16]*



1.4.2. Technical Design and Optical Configuration

The B23 beamline is dedicated exclusively to UV-visible circular dichroism measurements of chiral systems, including proteins, nucleic acids, small ligands and polymers in both solution and thin film form. The beamline comprises two independently operated end stations:[17, 18]

- Module A is optimised for solution-state measurements in the vacuum-UV to near-UV range (125-500 nm). it is primarily dedicated to the study of chiral molecules in solution, including proteins, nucleic acids, peptides and small ligand. It is equipped with a variety of sample environments, including multi-position turrets for multiple cuvettes, temperature-controlled holders and adaptable interfaces for microfluidic chips. These features enable precise control over experimental parameters such as temperature, buffer composition and path length. this allows for the detailed exploration of structural stability and conformational changes in biomolecules under a wide variety of conditions.
- Module B extends into the visible region (190-650 nm) and can accommodate solid-state or anisotropic samples. It incorporates a Mueller Matrix Polarimeter (MMP), an instrument capable of performing full polarisation analysis. This system enables the separation of different optical contributions, such as linear dichroism, circular birefringence and optical rotation, which may otherwise overlap in conventional CD measurements. The MMP therefore allows for a more accurate interpretation of chiroptical data, particularly in anisotropic samples, such as thin films, oriented fibres, and chiral polymeric materials. This capability is also important for identifying and correcting for potential artefacts, ensuring that the recorded CD spectra truly reflect molecular chirality rather than macroscopic optical effects.

Together, they cover an exceptionally broad spectral range corresponding to photon energies of approximately 2^{-10} eV.[17]

The highly collimated and precisely focused synchrotron beam at B23 has a cross-sectional area of approximately $0.5\text{-}2\times 4\text{ mm}^2$, enabling accurate measurements on microvolume samples, capillaries, microfluidic devices or thin films. The optical setup incorporates subtractive (double) grating monochromators that effectively eliminate stray light, thereby extending the usable spectral range into the vacuum UV region. Detection is achieved using a high-sensitivity photomultiplier tube (PMT) optimised for a low noise level and a high signal-to-noise ratio across the entire wavelength range.[18]

These design features enable unparalleled spectral quality, sensitivity and reproducibility while accommodating a wide variety of sample formats and experimental conditions. The B23 beamline is



therefore a cutting-edge platform for exploring the conformational dynamics, stability, and interactions of chiral biomolecules with unmatched throughput and precision.

1.4.3. Applications

The B23 beamline has been successfully used for a variety of structural and biophysical studies. Typical applications include the study of protein-ligands interactions, conformational transitions, and protein folding and unfolding, as well as photostability assays.[19, 20] In nucleic acid research, the facility enables the detailed examination of the formation of secondary structures, such as G4s and hairpin loops, while also supporting the analysis of carbohydrates, chiral polymers and other optically active materials.[21]

Due to its high photon flux, B23 can also be used to perform UV-induced denaturation studies. While repeated scans or prolonged exposure to intense radiation can lead to photodegradation, this phenomenon can be deliberately exploited to probe the photostability and refolding capacity of biomolecules. The beamline offers dedicated controls for such experiments that regulate scan speed, slit width, and sample rotation, minimising unwanted damage while enabling reproducible monitoring of light-induced effects.[22]

Beyond solution studies, module B and the MMP open up new possibilities in materials science, particularly regarding the characterisation of chiral thin films, organic optoelectronic materials and supramolecular assemblies. The high degree of polarisation control enables mapping of local chiroptical properties with high spatial resolution, providing insights into molecular orientation and ordering at surfaces.[23]

1.4.4. Limitations

Despite its exceptional capabilities, SRCD spectroscopy at B23 presents certain practical challenges that must be considered during the design of experiments. For example, measurements in the vacuum-UV and far-UV regions require high-purity optical materials, such as quartz, to minimise absorption by the sample holder. Furthermore, as oxygen and water strongly absorb in this spectral range, experiments are usually carried out in a nitrogen-purged atmosphere to ensure spectral clarity and reproducibility. [17]

While the high brilliance of the synchrotron beam is advantageous, it also increases the risk of photodamage to sensitive biomolecules. To mitigate this, parameters such as scan speed, slit width and exposure time must be carefully optimised.

Finally achieving accurate measurements in the low-wavelength region requires stringent control of spectral purity and stray light. B23 optical design, which incorporates double-grating



monochromators and high-quality reflective optics, minimises such artefacts, enabling the collection of reliable data even in the challenging vacuum-UV regime. [18]

1.5. Aim of the Chapter

This chapter aims to provide a comprehensive analysis of the interaction between newly synthesised tetrazole-based small molecules and distinct G4 DNA structures. This will be achieved by using high-throughput synchrotron radiation circular dichroism as a powerful, label-free screening platform. By comparing ligand-binding in sodium- and potassium-rich environments, the study will seek to elucidate the influence of ionic conditions on G4 conformational landscapes and ligand accessibility. Through this systematic evaluation, the chapter aims to identify compounds that exhibit selective affinity towards particular G4 topologies, understand their structural impact on G4 folding and explore their potential as modulators of biologically relevant quadruplexes.[24]



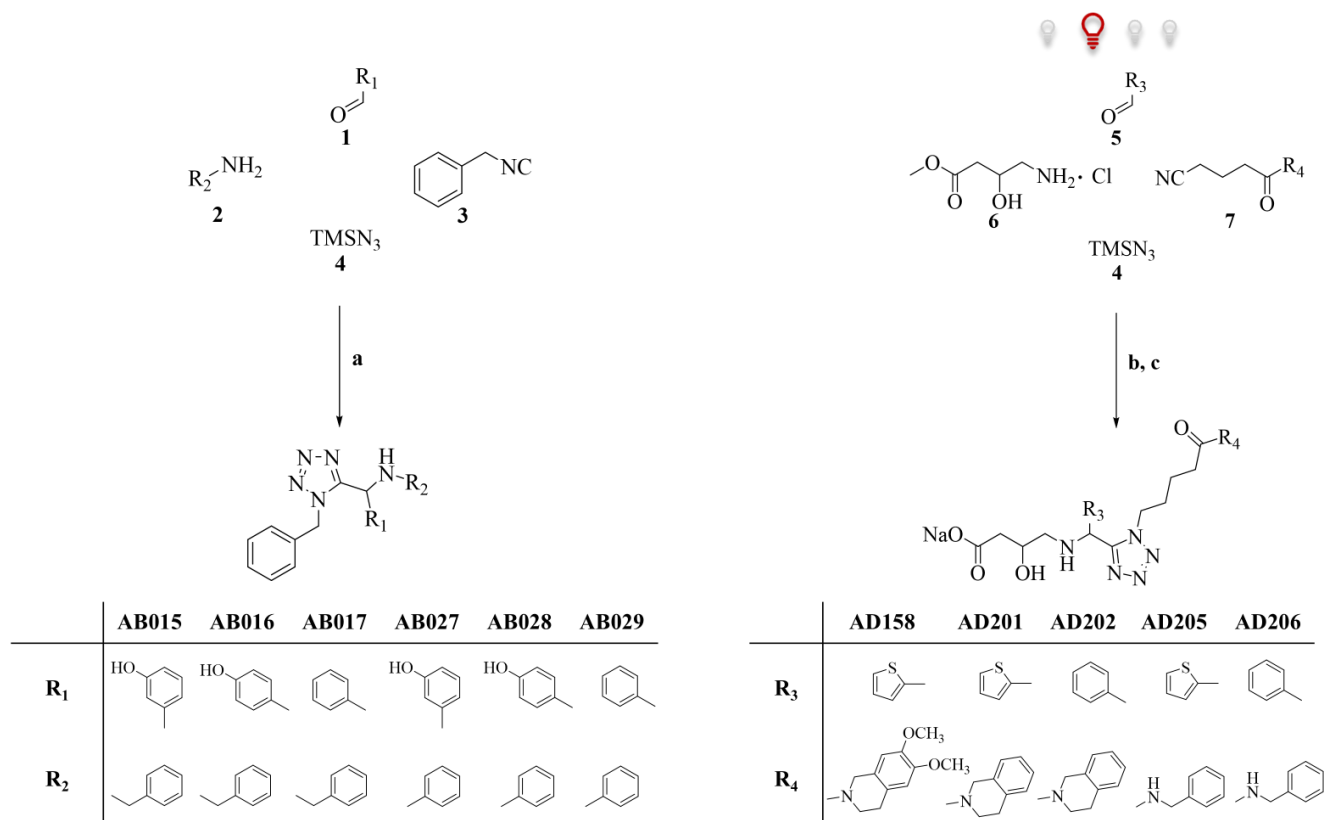
2. Results and Discussion

2.1. Ligand Design and Synthesis

The design of small molecules library was guided by the fundamental structural features required for the efficient recognition of G4s (Scheme 3.1). G4 ligands typically rely on extended aromatic scaffolds that can engage in π - π stacking interactions with planar G-tetrads, as well as functional groups that can modulate selectivity towards specific G4 topologies. Based on this, a series of tetrazole-based derivatives were selected as a molecular framework of interest. Tetrazoles are attractive candidates for exploring G4 binding preferences due to their aromaticity, synthetic accessibility, and versatile functionalisation patterns.

To generate structural diversity, a small library of tetrazole analogues was prepared using the Ugi-tetrazole four-component reaction: a one pot, multicomponent process offering rapid access to 1,5-disubstituted α -amino tetrazoles. Compounds AB015, AB016, AB017, AB027, AB028, and AB029 were synthesised in good yields by reacting commercially available benzaldehydes, primary amines, benzyl isocyanide and trimethylsilyl azide (TMSN₃) in methanol at room temperature.

A second subset of tetrazoles – AD158, AD201, AD202, AD205, and AD206 – was obtained under similar conditions, this time using substituted aromatic aldehydes and methyl-4-amino-3-hydroxybutanoate hydrochloride as the amine component, together with previously isocyanide derivatives. After the tetrazole products were formed, the corresponding esters underwent basic hydrolysis to yield the final carboxylate-containing ligands.



Scheme 3.1. Scheme of 1,5-disubstituted α -amino tetrazole synthesis. Reagents and conditions: **A**) CH_3OH , r.t., 1–5 days (63–80% yields); **B**) CH_3OH , Na_2SO_4 , Et_3N , r.t., 5 days (27–45% yields); **C**) NaOH , MeOH , r.t., 5 h (64–97% yields).

Overall, this synthetic strategy enabled the rapid generation of a structurally diverse panel of tetrazole derivatives suitable for probing G4 binding and selectivity.

2.2. HT-SRCD Analysis of G4-Ligand Interaction

A comprehensive screening was carried out using high-throughput synchrotron radiation circular dichroism to investigate how the newly synthesised tetrazole derivatives interact with different G4 topologies. A key advantage of HT-SRCD is that it provides direct structural information without the need for fluorophores, quenchers or surface immobilisation. This is particularly valuable in G4 research, where chemical modifications often alter folding equilibria or interfere with ligand binding. Unlike fluorescence-based melting assays (FRET), electrophoretic mobility shift assays (EMSA) or surface plasmon resonance (SPR) measurements, each of which requires dyes, immobilisation surfaces or tethering, the HT-SRCD approach measures native G4 structures in solution and monitors ligand-induced conformational changes with high spectral resolution. As the sample remain completely unmodified, the method captures authentic folding behaviour and true binding-induced perturbations, enabling reliable comparisons to be made across different G4 sequences and ionic environments. The high-throughput, plate-based format also allows for the rapid screening of large ligand libraries while minimising sample consumption. Using the B23 beamline at Diamond Light



Source, which is equipped with custom 96-well CD plates, enabled the reliable and reproducible acquisition of CD spectra, despite the reduced optical path lengths typical of high-throughput formats. The suitability of this method for studying G4-ligand interactions had already been demonstrated by Cuny *et al.* using pyrodoctatin (PDS) titrations into G4 DNA.[25]

2.2.1. Experimental Setup and G4 reference topologies

The three G4-forming sequences (Table 3.1) analysed represent complementary models of biologically relevant quadruplexes:

- HTelo1 corresponds to the canonical repeat found in human telomeric DNA, which naturally forms G4 hybrid structures under physiological ionic conditions. These telomeric quadruplexes play a key role in regulating telomere length, protecting chromosome ends and modulating telomerase activity, which makes them prime targets for cancer treatment.[26]
- T95-2T is a shortened sequence that adopts a highly stable parallel topology, representing promoter-type quadruplexes found in oncogenes such as c-MYC and KRAS. These promoter G4s are structurally compact and often rigid, playing regulatory roles in transcriptional activation and the response to replication stress.[27]
- G3T3 is a synthetic sequence model designed to mimic quadruplex motifs found in gene regulatory regions. Its marked structural plasticity – switching between antiparallel and hybrid topologies depending on ionic conditions – makes it a good benchmark for probing ligand-induced structural rearrangements.[28]

Combining these three systems provides a comprehensive overview of ligand behaviour across telomeric, promoter-like regulatory G4 landscapes.

Table 3.1. G-quadruplex sequence.

Strand	Sequence
HTelo1	5'-TAGGGTTAGGGTTAGGGTTAGGG-3'
T95-2T	5'-TTGGGTGGGTGGGTGGGT-3'
G3T3	5'-GGGTTTGGGTTTGGGTTTGGG-3'

These sequences were analysed using a custom 96-well plate comprising eight detachable strings of twelve 1 cm pathlength cells. Each cell can hold 350 μ L of sample and has an optical window of 6.6 mm. As the loading chimney reduce the illuminated area around 2.5 mm in diameter, a highly collimated, narrow beam is required, which can only be achieved with the synchrotron source at the B23 beamline (Diamond Light Source) and not with standard laboratory CD instruments.



Ligand stock solutions were prepared in methanol and diluted into a buffer containing either Na^+ or K^+ ions, ensuring that the final methanol concentration remained below 0.5%. Control experiments confirmed that this small amount of methanol had no detectable effect on the CD spectra of the G4s. The reference G4 topologies in the two ionic environments were assigned by comparing the resulting spectra with established signatures.

In 70 mM K^+ , both HTelo1 and G3T3 displayed features characteristic of hybrid G4 topologies. However, in Na^+ , HTelo1 maintained a hybrid fold, whereas G3T3 shifted to an antiparallel structure, as evidenced by the disappearance of the shoulder at 270 nm. Remarkably, T95-2T retained its parallel G4 signature – a strong positive band at 266 nm and a negative band at ~ 245 nm – independently of whether Na^+ or K^+ was present (Figure 3.7).

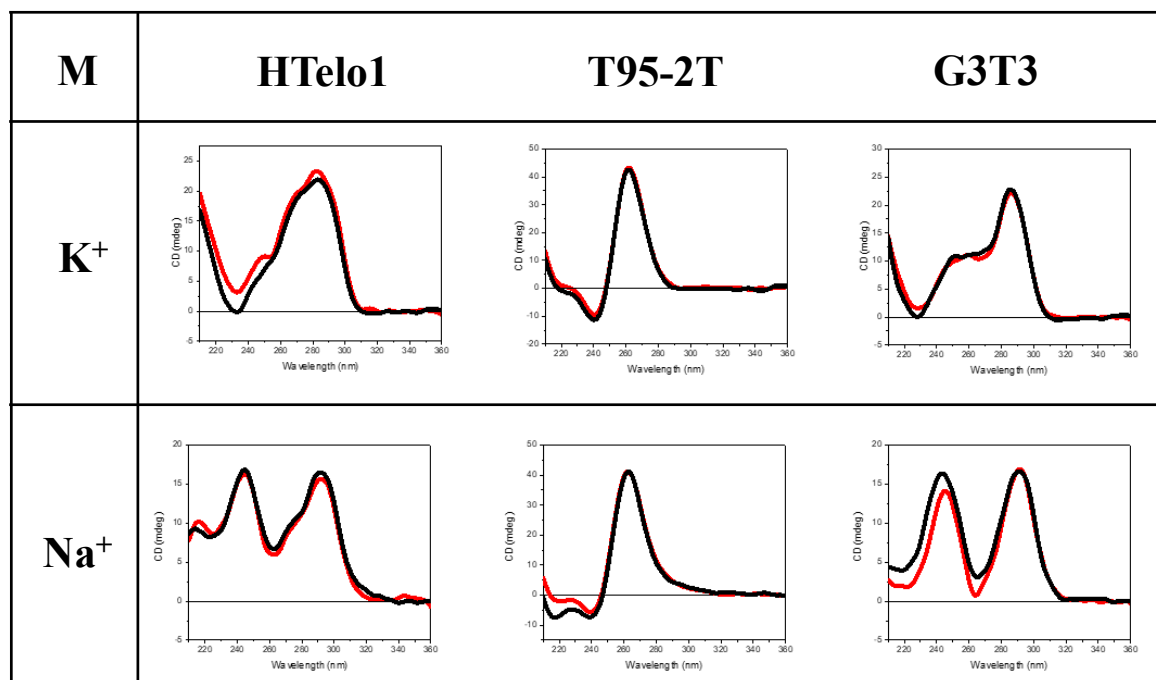


Figure 3.7. SRCD spectra in the 210–360 nm range of the three different G4 strands alone (black lines) or in the presence of <0.5% MeOH (red lines) in 10 mM Tris-HCl buffer, pH 7.4, in the presence of 70 mM K^+ or Na^+ ions, expressed as ellipticity (mdeg), were acquired using a 96-well plate at module A spectropolarimeter.

2.2.2. Ligand Screening by HT-SRCD

The interaction between the tetrazole derivatives and the three G4s was evaluated by comparing the CD spectrum of each G4-ligand mixture with the theoretical sum of the spectra of the individual components. Any differences between the two were interpreted as ligand-induced perturbations. These variations depended strongly on the G4 sequence and the identity of the monovalent cation.

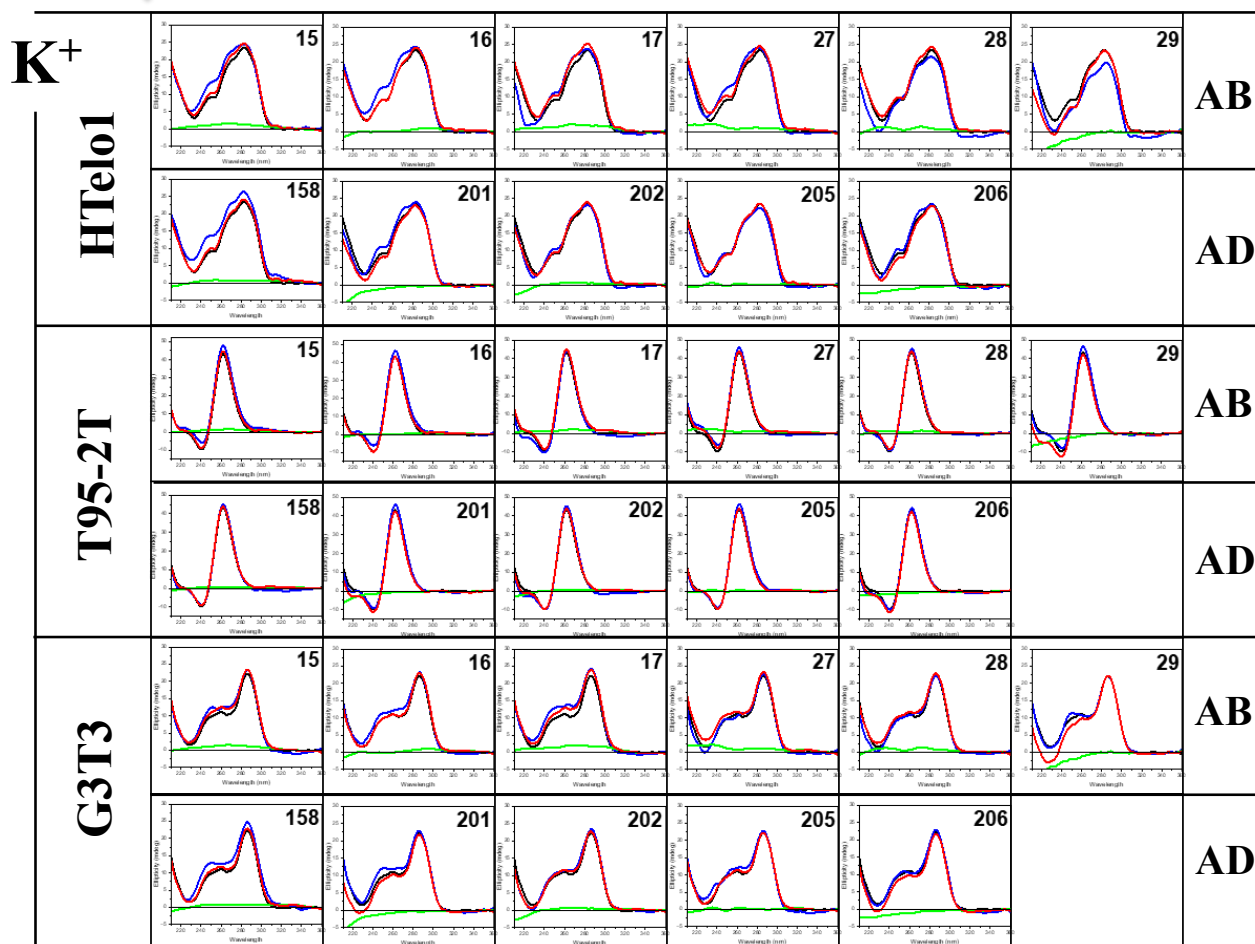


Figure 3.8. Summary of the SRCD spectra in the 210–360 nm range of G4 sequences (HTelo1, T95-2T, G3T3) in the presence of different ligands (from left to right: AB series = AB015, AB016, AB017, AB027, AB028, AB029; AD series = AD158, AD201, AD202, AD205, AD206) in 10 mM Tris-HCl buffer, pH 7.4, in the presence of 70 mM K⁺ acquired using a 96-well plate at module A. Each box contains the spectrum of the G4 alone (black), the measured (blue) and the calculated (red) spectra of the ligand–G4 complex and the spectrum of the ligand alone (green line) expressed as ellipticity (mdeg).

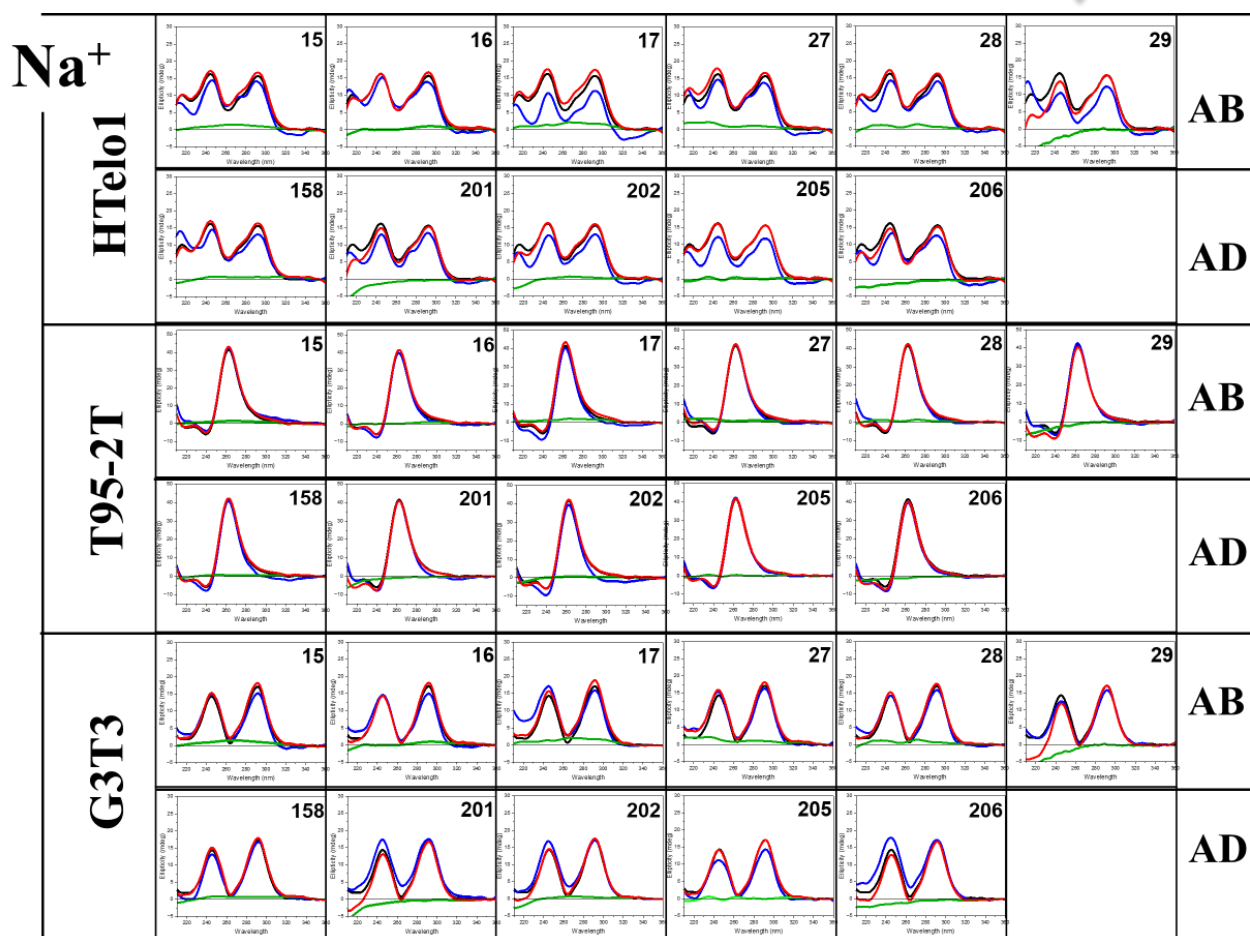


Figure 3.9. Summary of the SRCD spectra in the 210–360 nm range of G4 sequences (HTelo1, T95-2T, G3T3) in the presence of different ligands (from left to right: AB series = AB015, AB016, AB017, AB027, AB028, AB029; AD series = AD158, AD201, AD202, AD205, AD206) in 10 mM Tris-HCl buffer, pH 7.4, in the presence of 70 mM Na^+ acquired using a 96-well plate at module A. Each box contains the spectrum of the G4 alone (black), the measured (blue) and the calculated (red) spectra of the ligand–G4 complex and the spectrum of the ligand alone (green line) expressed as ellipticity (mdeg).

Across all experiments, T95-2T exhibited minimal spectral variation, suggesting that none of the compounds significantly interacted with this parallel G4 form under either Na^+ or K^+ conditions (Figure 3.8, Figure 3.9). This insensitivity suggests that the compact parallel arrangement of T95-2T offers accessible binding interfaces.

Consistently, none of the examined ligands induced any detectable topology switching in the studied G4 sequences; all perturbations were consistent with subtle local rearrangements rather than global structural transitions.

To compare the effects of the ligands more systematically, the percentage variation between the experimental CD spectrum of each G4-ligand mixture and the calculated sum of the individual spectra was evaluated. For HTelo1 and G3T3 in K^+ , the analysis was performed at the longer-wavelength positive band (~ 290 nm), which is characteristic of hybrid G4s. The results showed clear structure- and ion- dependent behaviour.

Results and Discussion

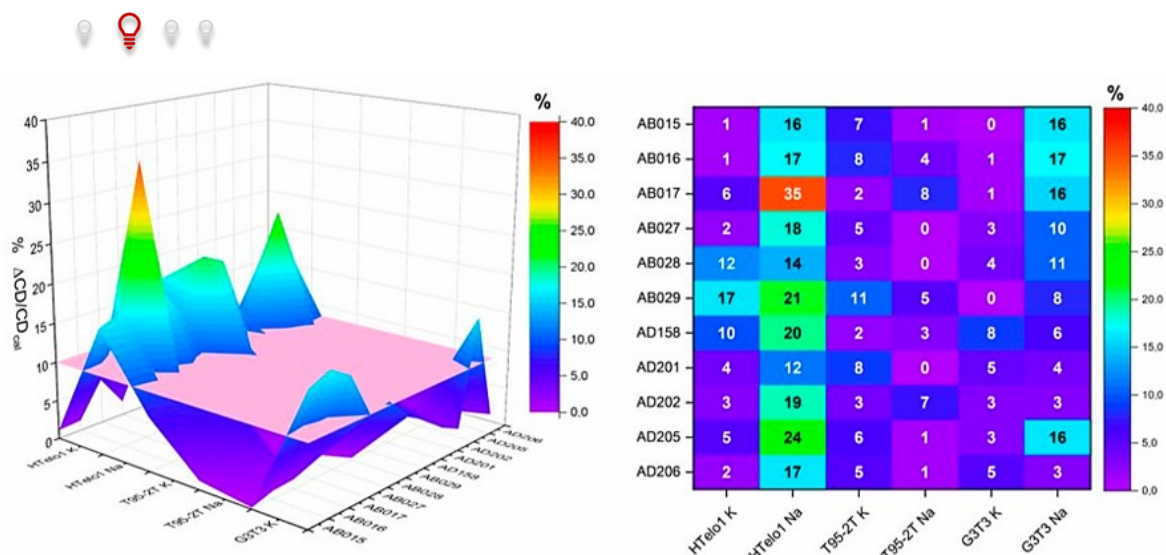


Figure 3.10. Summary of the percentage (%) change between the measured and calculated dichroic spectrum for each G4-ligand complex in the presence of potassium or sodium ions.

In Na^+ , all ligands produced $>10\%$ CD variation for HTelo1, indicating consistent ligand engagement. For G3T3, however, only AB015, AB016, AB017 and AD205 exceeded the 10% threshold, while AB027 and AB028 produced perturbations that were almost within the threshold. For T95-2T, no ligand generated variations above 10% in either ionic condition.

These observations (Figure 3.10), reveal that Na^+ facilitates ligand binding, particularly for HTelo1 and, to a lesser extent, G3T3. Sodium-stabilised G4s are known to adopt less compact structures, providing more accessible loops and grooves for ligand interaction.

In contrast to Na^+ , K^+ produced a markedly different interaction landscape. Most compounds failed to generate significant CD perturbations across all G4s tested, reflecting the higher stability and compactness of K^+ stabilised G4 topologies.

Only AB028 and AB029 induced measurable changes ($>10\%$) in HTelo1, with AB029 additionally affecting T95-2T – an effect not observed with Na^+ . For G3T3 in K^+ , no small molecule produced notable CD variation.

The different behaviour of Na^+ and K^+ highlights the sensitivity of ligand engagement to the conformational plasticity of the target G4 (Figure 3.11). Melting curve analyses further confirm the stabilising influence of K^+ . T_m values were consistently higher in potassium, with increases ranging from 10°C (HTelo1) to 32°C (T95-2T).

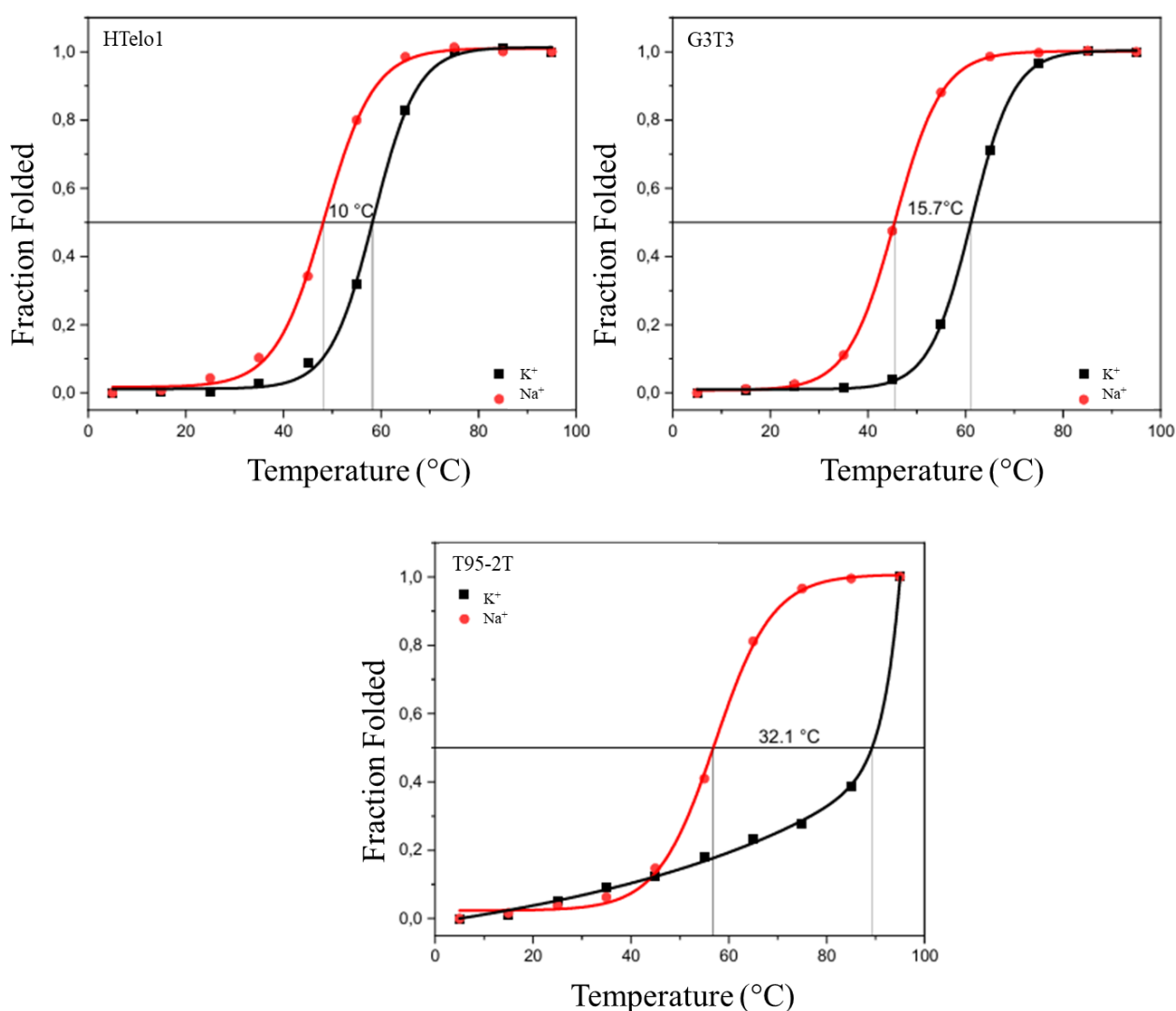


Figure 3.11. SRCD melting curves for HTelo 1, G3T3, and T95-2T G4s in 10 mM phosphate buffer, pH 7.4, in the presence of 70 mM potassium or sodium ions (indicated). The ellipticity was monitored at the maximum wavelength of each G4 spectrum.

2.2.3. Benchmarking with Rhau25 Peptide

To validate the assay and contextualise the performance of the tetrazole library, ligand screening was also performed with the well-characterised Rhau25 peptide, which has been shown to bind G4 structures (Figure 3.12).[21]

Rhau25 originates from the N-terminal recognition of the DHX36 helicase (also known as Rhau or G4-resolving helicase 1), which is one of the most specific and potent G4-binding proteins found in human cells. DHX36 plays a crucial role in the resolution of parallel and hybrid G4 structures during transcription, replication, and transition.[29] The N-terminal fragment of DHX36 retains a high affinity for G4s, which is driven by a combination of electrostatic contacts and stacking interactions at the 5' tetrad interface.[29] Therefore, Rhau25 peptide serve as an established positive control for

Results and Discussion



the recognition *in vitro*, providing a biologically meaningful benchmark for evaluating the binding behaviour and structural impact of synthetic small molecules. Including it in HT-SRCD screening enables direct comparison between peptide- and small-molecule-induced perturbations, helping to validate the method ability to resolve subtle conformational changes.

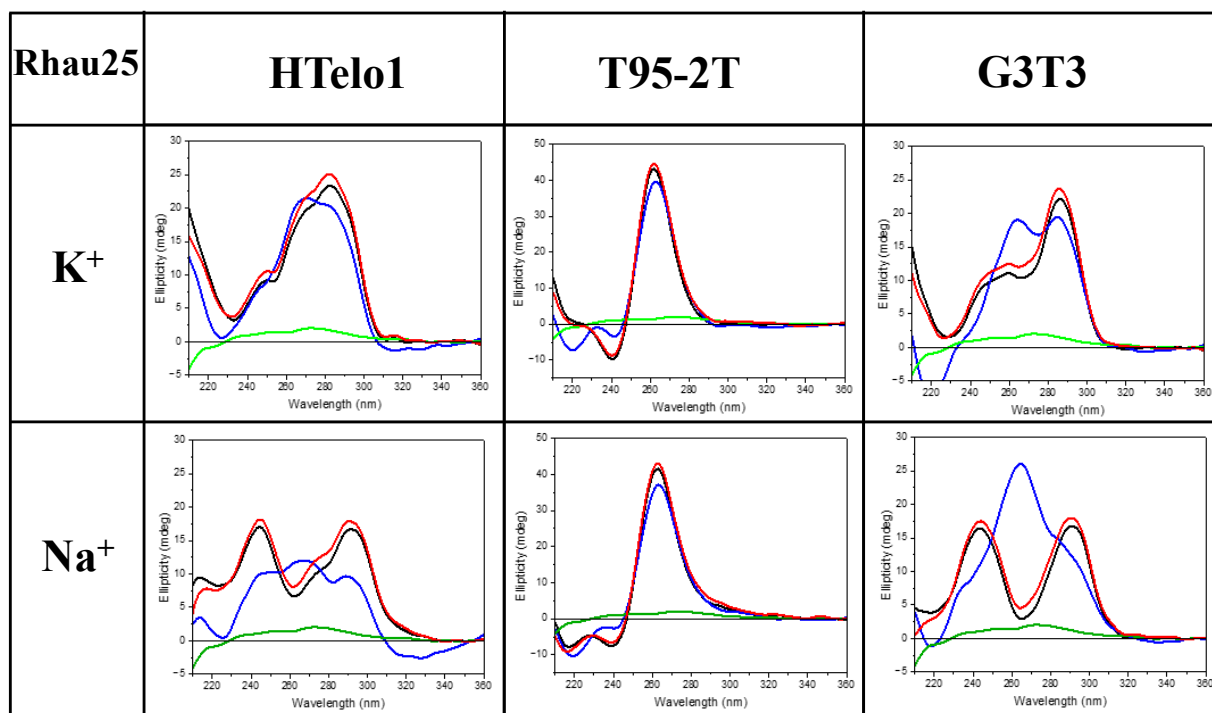


Figure 3.12. HT-SRCD spectra of G4 sequences in the presence of Rhau25 peptide in 10 mM Tris-HCl buffer, pH 7.4, in the presence of 70 mM K^+ or Na^+ ions, acquired using 2 string plates using B23 module A. Each box contains the spectrum of the G4 alone (black), the measured (blue) and the calculated (red) spectra of the ligand–G4 complex and the spectrum of the ligand alone (green line).

In the presence of K^+ , Rhau25 produced only modest spectral changes in HTelo1 and T95-2T, but induced a reorganisation of the CD shoulder of G3T3. This gave rise to a distinct positive band near 250 nm, indicating interaction without major topology remodelling.

In Na^+ , however, Rhau25 displayed a markedly stronger influence. For both HTelo1 and G3T3, it produced a pronounced positive band at approximately 270 nm, consistent with the induction of a partial parallel-like topology superimposed on the existing hybrid or antiparallel structures.

For T95-2T, the parallel topology persisted in both environments, with only a modest decrease in band intensity ($\sim 15\%$).

These results support the hypothesis that sodium promotes structural flexibility, enabling ligands, whether small molecules or peptides, to modulate G4 conformations more effectively.



3. Conclusion

This study highlights the pivotal role of G4 conformations and, consequently, ligand accessibility through a systematic evaluation of ligand binding across distinct G4-forming sequences and ionic environments. The data clearly demonstrate that sodium ions increase the conformational flexibility of HTelo1, G3T3, and T95-2T, favouring more open and dynamic G4 structures. This flexibility enables ligand engagement, rendering Na⁺-rich conditions particularly favourable for molecules that depend on groove or loop interactions rather than classical end-stacking. This finding provides a valuable framework for designing ligands that operate effectively in biological contexts where sodium predominates, potentially improving their therapeutic relevance.

Of the tetrazole-based small molecules investigated, several compounds were found to selectively alter the CD signature of HTelo1 in a NaCl buffer solution, while having negligible effects under KCl conditions. This sodium-dependent selectivity suggests that members of this chemical class could be designed to recognise G4 topologies with more relaxed conformations, such as those found in regulatory genomic regions, more effectively. This could be exploited to enhance target specificity and reduce undesirable interactions with more compact, potassium-stabilised G4 structures.

By contrast, the Rhau25 peptide exhibited a greater ability to disrupt G4 conformations, particularly in presence of Na⁺, where it prompted the formation of CD bands indicative of partial parallel rearrangements in both HTelo1 and G3T3. These results reaffirm the peptide versatility as a G4 binder and demonstrate how HT-SRCD can accurately detect ligand-induced conformational remodelling – a key aspect in understanding G4-protein recognition and the inhibition of G4-resolving enzymes.

Overall, the impact of this work is twofold. Firstly, it advances the understanding of how ionic conditions dictate G4 structural dynamics and ligand-binding behaviour – an essential consideration for the rational design of G4-targeted therapeutic agents, particularly in fields such as oncology and antiviral research. Secondly, it demonstrates the robustness and scalability of the B23 high-throughput SRCD platform, establishing it as an exceptionally powerful tool for the rapid screening of large ligand libraries. Its ability to provide high-quality structural fingerprints across numerous samples significantly accelerates the early stages of drug discovery.

In conclusion, the findings presented here highlight the necessity of integrating ionic context into the development of selective and efficacious G4 ligands. Future investigations should aim to further analyse the structural dynamics underlying ligand-induced rearrangements and optimise ligand scaffolds for improved binding profiles under physiologically relevant conditions. The successful application of HT-SRCD in this study validates its utility and underscores its potential to drive innovation at the intersection of structural biology, chemical biology, and pharmacological research.



4. Experimental Part

4.1. General Procedure for the Synthesis of Rhau25 peptide

The Rhau25 peptide (Ac-SMHPGHLKGREIGMWYAKKQGQKNK-NH₂) was synthesised using automated solid-phase peptide synthesis (SPPS) and standard Fmoc/HBTU chemistry. This process was performed using a Biotage® Syro Wave™ synthesizer (Biotage AB, Uppsala, Sweden) and Syro XP software, enabling the controlled and efficient sequential coupling of protected amino acids.

Following chain elongation, N-terminal acetylation was performed by treating the resin-bound peptide with acetic anhydride. Final cleavage from the Rink amide resin was achieved using a TFA cleavage containing TIS and water as scavengers to prevent unwanted side reactions. The crude peptide was then precipitated by the addition of cold diethyl ether, collected and dried.

Purification was performed using preparative reverse-phase HPLC on a Dionex Vydac C18 column (300 Å, 10 µm, 22 x 250 mm), with a Shimadzu preparative system equipped with LC-8A pumps, an SLC-8A controller, an SPD-6A UV detector and an ERC-3562 ERMA degasser.

Analytical characterisation of the purified material was performed using an Agilent 1260 Infinity II HPLC system coupled to an Agilent 6130 quadrupole LC-MS. The observed molecular ion peak was consistent with the calculated molecular mass of 2951.41 Da.

4.2. Preparation of G4-Forming Sequences

Synthetic DNA oligonucleotides were purchased from Eurogentec (Seraing, Belgium) and dissolved in ultrapure water. The concentration of each stock solution was determined spectrophotometrically at 90°C using the following extinction coefficients at 260 nm: $\epsilon = 236.500 \text{ M}^{-1} \text{ cm}^{-1}$ for Htelo1; $\epsilon = 197.300 \text{ M}^{-1} \text{ cm}^{-1}$ for G3T3; and $\epsilon = 172.600 \text{ M}^{-1} \text{ cm}^{-1}$ for T95-2T.

The oligonucleotides were then diluted to a final concentration of 4.5 µM in 10 mM TRIS-HCl buffer solution (pH 7.4), supplemented with either 70 mM KCl or 70 mM NaCl depending on the experimental conditions. To promote proper G4 folding, each solution was annealed by heating to 90°C for 10 minutes, followed by slow cooling to room temperature.

4.3. Preparation of small-molecule ligands

Each small-molecule ligand was initially dissolved in methanol to prepare 1mg/mL stock solution. Immediately before use, the ligands were diluted with 10mM TRIS-HCl buffer solution (pH 7.4) containing either 70 mM Na⁺ or 70 mM K⁺, in order to obtain required final working concentration of 0.155 mM.



In all assays, the final methanol content was kept below 0.5% (v/v) to avoid disturbing the secondary structure of the G4 sequences.

4.4. High-Throughput Synchrotron Radiation Circular Dichroism

HT-SRCD measurements were performed at module A of the B23 beamline at the Diamond Light Source in the UK, using a 96-well microplate setup designed specifically for high-throughput circular dichroism analysis. The microplate was mounted in a vertical sample holder with a motorised X-Y stage to enable the automated, precise positioning of each well in the synchrotron beam.

For each measurement, 360 μL of ligand solution to achieve a final G4:ligand molar ratio of 1:2. CD spectra were acquired at 25°C over the 210-360 nm range. Each well was scanned once at a scan rate of 39 nm/min, with a data pitch of 1 nm, a bandwidth of 1 nm, and an integration time of 2 s per point. Control spectra containing G4 and ligand alone were collected under identical conditions.

SRCD melting experiments were conducted using 0.05 cm pathlength quartz cuvettes (Hellma Analytics) at temperatures ranging from 5°C to 95°C in 10°C increments. The cuvettes were allowed to equilibrate for 120 s at each temperature. Spectra were recorded from 200 nm to 360 nm. the DNA strand concentration was kept at 10 μM in a 10 mM TRIS-HCl buffer solution (pH 7.4), with either 70 mM Na^+ or K^+ ions present.

All collected spectra were processed and analysed using OlisWorks, CDApps,[30] and OriginPro 2025 (OriginLab).

References

1. Woody, R.W., *Circular dichroism*. Methods Enzymol, 1995. **246**: p. 34-71.
2. Zhang, D.W., M. Li, and C.F. Chen, *Recent advances in circularly polarized electroluminescence based on organic light-emitting diodes*. Chem Soc Rev, 2020. **49**(5): p. 1331-1343.
3. Moffitt, W., *THE OPTICAL ROTATORY DISPERSION OF SIMPLE POLYPEPTIDES. II*. 1956. **42**(10): p. 736-746.
4. Rogers, D.M., et al., *Electronic Circular Dichroism Spectroscopy of Proteins*. Chem, 2019. **5**(11): p. 2751-2774.
5. Pace, C.N., et al., *How to measure and predict the molar absorption coefficient of a protein*. Protein Sci, 1995. **4**(11): p. 2411-23.
6. Martin, S.R. and M.J. Schilstra, *Circular Dichroism and Its Application to the Study of Biomolecules*, in *Methods in Cell Biology*. 2008, Academic Press. p. 263-293.
7. Kypr, J., et al., *Circular dichroism and conformational polymorphism of DNA*. Nucleic Acids Research, 2009. **37**(6): p. 1713-1725.
8. del Villar-Guerra, R., R.D. Gray, and J.B. Chaires, *Characterization of Quadruplex DNA Structure by Circular Dichroism*. 2017. **68**(1): p. 17.8.1-17.8.16.
9. Munro, I.H. and G.V. Marr, *CHAPTER 1 - SYNCHROTRON RADIATION SOURCES*, in *Handbook on Synchrotron Radiation*, G.V. Marr, Editor. 1987, Elsevier: Amsterdam. p. 1-20.
10. Mills, D.M., et al., *Report of the Working Group on Synchrotron Radiation Nomenclature - brightness, spectral brightness or brilliance?* J Synchrotron Radiat, 2005. **12**(Pt 3): p. 385.
11. Wakatsuki, S., *Structural Biology Applications of Synchrotron Radiation and X-Ray Free-Electron Lasers*, in *Synchrotron Light Sources and Free-Electron Lasers: Accelerator Physics, Instrumentation and Science Applications*, E.J. Jaeschke, et al., Editors. 2020, Springer International Publishing: Cham. p. 1673-1716.
12. Rodger, A., M.A. Ismail, and M. Gore, *Introduction to circular dichroism*, in *Spectrophotometry and Spectrofluorimetry: A Practical Approach*. 2000, Oxford University Press. p. 0.
13. Hussain, R., et al., *Circular Dichroism and Synchrotron Radiation Circular Dichroism Applications to Biomaterials*, in *Radiation in Bioanalysis: Spectroscopic Techniques and Theoretical Methods*, A.S. Pereira, P. Tavares, and P. Limão-Vieira, Editors. 2019, Springer International Publishing: Cham. p. 147-172.

14. Siligardi, G. and R. Hussain, *CD Spectroscopy: An Essential Tool for Quality Control of Protein Folding*, in *Structural Proteomics: High-Throughput Methods*, R.J. Owens, Editor. 2015, Springer New York: New York, NY. p. 255-276.
15. Hussain, R. and G. Siligardi, *Characterisation of Conformational and Ligand Binding Properties of Membrane Proteins Using Synchrotron Radiation Circular Dichroism (SRCD)*, in *The Next Generation in Membrane Protein Structure Determination*, I. Moraes, Editor. 2016, Springer International Publishing: Cham. p. 43-59.
16. Hussain, R., et al., *High-throughput SRCD using multi-well plates and its applications*. Scientific Reports, 2016. **6**(1): p. 38028.
17. Hussain, R., et al., *DESIGN OF B23 CIRCULAR DICHROISM BEAMLINE AT DIAMOND LIGHT SOURCE*. 2008. **01**(02): p. 265-270.
18. Siligardi, G., et al., *B23 circular dichroism beamline at the Diamond Light Source*. Diamond Light Source Proceedings, 2010. **1**(SRMS-7): p. e104.
19. Kwan, T.O.C., et al., *Measuring Protein Aggregation and Stability Using High-Throughput Biophysical Approaches*. 2022, **9**.
20. Townsend, D.J., et al., *Circular Dichroism Spectroscopy Identifies the β -Adrenoceptor Agonist Salbutamol As a Direct Inhibitor of Tau Filament Formation in Vitro*. ACS Chemical Neuroscience, 2020. **11**(14): p. 2104-2116.
21. Honisch, C., et al., *Interaction of a Short Peptide with G-Quadruplex-Forming Sequences: An SRCD and CD Study*. 2021. **13**(8): p. 1104.
22. Honisch, C., et al., *Enhanced Protein Photo-Stability Analysis Using SRCD in the Presence of Phospholipid SUVs*. 2025. **31**(37): p. e202500792.
23. Hussain, R., T. Javorfi, and G. Siligardi, *CD Imaging at High Spatial Resolution at Diamond B23 Beamline: Evolution and Applications*. Frontiers in Chemistry, 2021. **9**: p. 616928.
24. Rotondo, M., et al., *Rapid G4 Ligand Screening Through Spectral Changes Using HT-SRCD with Minimal Material*. 2025. **30**(16): p. 3322.
25. Cuny, Z.G., *Characterization of Pyridostatin and its Interactions with c-MYC G-Quadruplexes*. . Bachelor Thesis, Mississippi State University, Starkville, MS, USA, 2017.
26. Xu, Y. and M. Komiyama, *G-Quadruplexes in Human Telomere: Structures, Properties, and Applications*. Molecules, 2023. **29**(1).
27. Barr, J., et al., *From loops to caps: discriminating peptide binding to distinct G-quadruplex tetrads using 5-furyl-2'-deoxyuridine fluorescent probes*. RSC Chem Biol, 2025.

References

28. Fujimoto, T., et al., *Thermal stability and hydration state of DNA G-quadruplex regulated by loop regions*. Nucleic Acids Symp Ser (Oxf), 2009(53): p. 237-8.
29. Antcliff, A., L.D. McCullough, and A.S. Tsvetkov, *G-Quadruplexes and the DNA/RNA helicase DHX36 in health, disease, and aging*. Aging (Albany NY), 2021. **13**(23): p. 25578-25587.
30. Hussain, R., et al., *CDApps: integrated software for experimental planning and data processing at beamline B23, Diamond Light Source*. Journal of Synchrotron Radiation, 2015. **22**(2): p. 465-468.

Conclusion

This thesis presents a multidisciplinary study of peptide- and small-molecule-based systems designed to recognise and modulate nucleic acid structures. It integrates synthetic chemistry, structural biophysics and cellular studies. The three research chapters describe the design, synthesis and evaluation of distinct classes of molecules – CPP-PNA conjugates, Vimentin-derived peptide fragments and tetrazole-based small ligands – for their ability to interact with biologically relevant nucleic acid targets. Despite the different molecular frameworks and experimental approaches employed, the overall findings converge towards a common theme: the rational modulation of regulatory nucleic acid networks that sustain tumour progression and therapeutic resistance. Cancer development is driven by the dysregulation of gene expression, structural plasticity, and signalling integration at multiple levels, and the projects described here address these levels of complexity. Rather than being independent investigations, targeting miRNAs with PNA-based conjugates, characterising Vimentin-G-Quadruplex interactions and screening small-molecule G4 ligands collectively outline a coherent framework centred on perturbed nucleic acid architectures underlying key cancer hallmarks

The first part of this thesis describes the synthesis of CPP-PNA conjugates that were engineered to silence the oncogenic microRNA, miR-221-5p. These conjugates were obtained using CuAAC and disulfide-based coupling strategies, with the aim of evaluating their cellular internalisation and antisense potential. The synthesis established robust methodologies for assembling the conjugates, and spectroscopic analyses confirmed that the PNAs could still hybridise with the complementary miRNA target. However, internalisation studies revealed critical physicochemical limitations: the conjugates aggregated significantly in aqueous media, resulting in variable uptake efficiencies. These issues highlight that even well-designed nucleic acid analogues can exhibit unpredictable behaviour at the bio-interface and the importance of formulation strategies and sequence-design optimisation for improving biological reproducibility. These findings emphasise key design considerations for antisense agents intended for use in cancer treatment. The methodological advances reported here, including optimised coupling procedures and comparative CPP evaluation, lay valuable groundwork for future development, particularly the assessment of CF-P14-(Triazole)-PNA-NH₂ as a next-generation delivery candidate. At the post-transcriptional level, the development of CPP-PNA conjugates that target miR-221-5p interferes directly with the regulatory RNA circuits that control cell proliferation and apoptosis. It is known that miR-221 represses tumour suppressor pathways and promotes progression through the G₁/S restriction point, which is a critical checkpoint that determines whether cell will commit to DNA replication or remain quiescent. The strategy explored in Chapter I attempts to neutralise this oncogenic microRNA in order to re-establish regulatory balance at the G₁/S

Conclusion

transition. This counteracts sustained proliferative signalling and resistance to growth-inhibitory cues. As dysregulation of this checkpoint involves alterations to the Rb/E2F axis and impaired p53-mediated surveillance, modulation of microRNA networks is a promising upstream approach to influence these key cell cycle regulators and potentially reinstate apoptotic competence

The second part of the thesis builds on this regulatory perspective by examining Vimentin, which is a hallmark of epithelial-mesenchymal transition and metastatic dissemination. EMT is intrinsically linked to invasion, metastatic potential, and therapy resistance; however, its molecular implications extended beyond cytoskeletal remodelling. The observation that Vimentin-derived peptides can interact with G-quadruplexes DNA structures suggests an intersection between structural plasticity and genome regulation. G4s are enriched in promoter regions of oncogenes and telomeric sequences, where they influence transcription programmes, replicative potential, and genome stability. Using an integrating approach combining CD and NMR spectroscopies, it was shown that each peptide fragment possesses subtle but distinct structural behaviours. While all sequences remained largely disordered in aqueous buffer, consistent with the intrinsically disordered nature of the corresponding protein regions, membrane-mimetic conditions such as TFE induced partially folded states, including helical or β -like conformations depending on the peptide. Alanine-scanning peptides provides insight into the interaction interface, revealing that the substitution of key aromatic or positively charged residues markedly weakened G4 recognition, confirming the pivotal role in stabilising peptide-DNA interactions. The ability of Vimentin Fragments to recognise these structures lends weight to the hypothesis that proteins traditionally associated with cellular architecture may also play a role in modulating nucleic acid topology. This would indirectly affect the transcriptional control of genes that govern proliferation and checkpoint transitions. In this context, cytoskeletal reprogramming and gene expression control appear to be interconnected processes that contribute to the acquisition of invasive behaviour and adaptation to oncogenic stress.

The third part of the study involved expanding the investigation of G4 recognition to include small-molecule tetrazole ligands and the G4-binding peptide Rhau25, using a technique called high throughput synchrotron radiation circular dichroism. This highly sensitive, label-free technique revealed that ligand-G4 interactions are highly significantly influenced by the ionic environment. Under sodium-rich conditions, where G4 structures adopt more open and flexible topologies, the tetrazole derivatives displayed greater accessibility to the G-tetrads. This led to measurable perturbations of the CD signatures. Several ligands exhibited selective recognition of HTelo1 and G3T3, suggesting topology- and sequence-dependent binding. In contrast, potassium stabilised G4s, which are more compact and conformationally rigid, exhibited reduced responsiveness to the addition of ligands. Unlike the small molecules, Rhau25 displayed distinctive behaviour: it was able to

remodel G4 conformations, promoting more parallel-like arrangements, particularly under sodium conditions. These findings highlight the potential of these ligands as scaffolds for the development of anticancer drugs and emphasise the value of HT-SRCD as a platform for ligand discovery and screening.

Taken together, the three research lines converge on a layered model of tumour biology intervention. This thesis approaches cancer as a regulatory network operating across RNA, DNA and protein interfaces by modulating oncogenic microRNAs, influencing protein-G4 interactions and directly targeting DNA secondary structures. This multiscale strategy reflects the intrinsic complexity of tumour systems, where sustained proliferation, evasion of apoptosis, acquisition of replicative immortality, genomic instability and metastatic competence are interconnected outcomes of deregulated nucleic acid control, rather than isolated phenomena.

Overall, this work expands our understanding of nucleic acid recognition and provides the conceptual and methodological foundations necessary for the development of next-generation anti-cancer agents that target microRNAs, G-quadruplexes and protein-DNA interfaces.

Materials and Methods

1. Materials

All reagents and solvents used in the processes of synthesis and purification of peptides have been purchased by the following chemical suppliers:

- Sigma Aldrich s.r.l. (Milan, IT): diethyl ether (Et₂O), piperidine, dimethylformamide triisopropylsilane (TIPS), diisopropylethylamine (DIPEA), acetic anhydride, dimethylformamide, (DMF), dichloromethane (DCM)
- Merck (Darmstadt, Germany): trifluoroacetic acid (TFA), 5-hexynoic acid
- Eurogentech (Seraing, Belgium): all the synthetic oligonucleotides
- PolyOrg. Inc. (Massachusetts, USA): Fmoc-AEG(nucleobase)-OH
- CBL (Patras, Greece): Rink Amide MBHA resin, 2-CTC resin were supplied
- Iris Biotech (Marktredwitz, Germany): DIC, Oxyma Pure, HBTU, HOBt·H₂O, HATU, Fmoc-protected amino acids
- Carlo Erba (Milan, IT): acetonitrile

HPLC grade water was obtained by filtration by an Elga Veolia PURELAB system (London, UK). The scales used were a Mettler AE 240 scale (Columbus, Ohio, USA) and a Mettler ΔT21 Comparator scale.

2. Instruments

- HPLC-MS analysis was performed on a Phenomenex Kinetex XB-C18 column (4.6 mm × 100 mm, 3.5 μm, 100 Å) with an Agilent Technologies 1260 Infinity II HPLC system and a 6130 quadrupole LC/MS instrument. Eluent A was 0.1% trifluoroacetic acid in water, while eluent B was 0.1% trifluoroacetic acid in acetonitrile.
- Preparative HPLC was performed on an ÄKTA pure system. The column for the separation was a Luna C18, 300Å, 10μm or a DAISOGEL SP120 C18, 300Å, 10μm column. Eluent A was 100% water + 0.05% TFA, eluent B: 10% water / 90% ACN + 0.05% TFA.
- Low-pressure flash chromatography was performed with chromatograph (Biotage Isolera Prime) equipped with a UV detector set at 216 and 280 nm. The sample solution was first absorbed onto a pre-column and then loaded onto the chromatography column. The column used was a C18SNAP 12g with a volume of 15 mL and a flow rate of 12 mL/min. Eluent A was 100% water + 0.05% TFA, eluent B: 10% water / 90% ACN + 0.05% TFA.
- CD spectra were recorded on a Jasco-1500 spectrophotometer using Spectra Manager (Jasco) software. CD spectra were recorded at 25°C using a 1 cm or a 0.1 cm quartz cell (Hellma Analytics), a 320–200 nm or 260–190 nm measurement range, 50 nm/min scanning speed, 1 nm bandwidth, 4s response time, 0.5 nm data pitch and are the results of 4 scans.

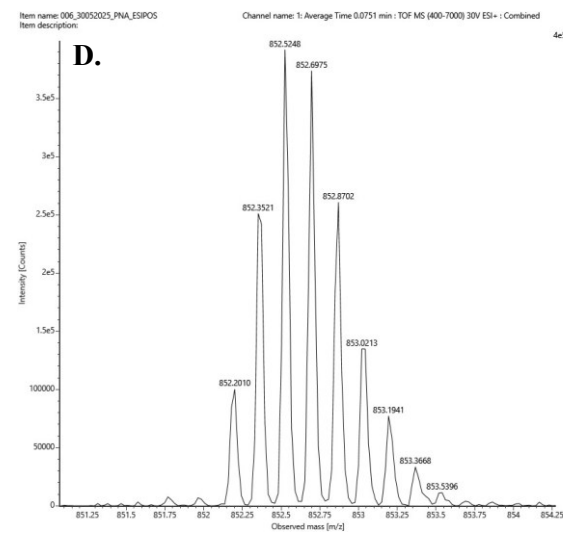
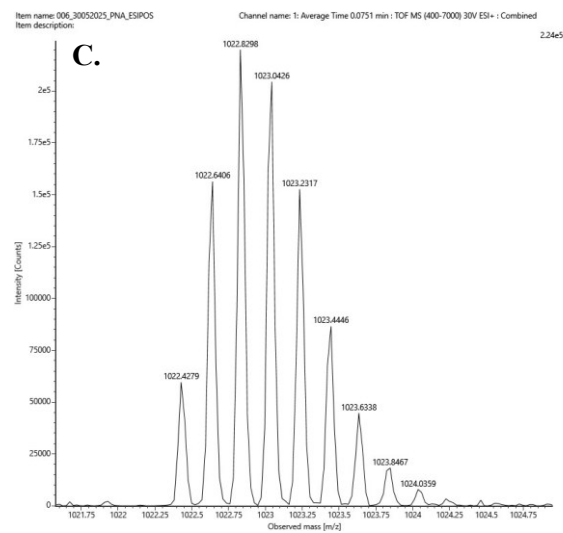
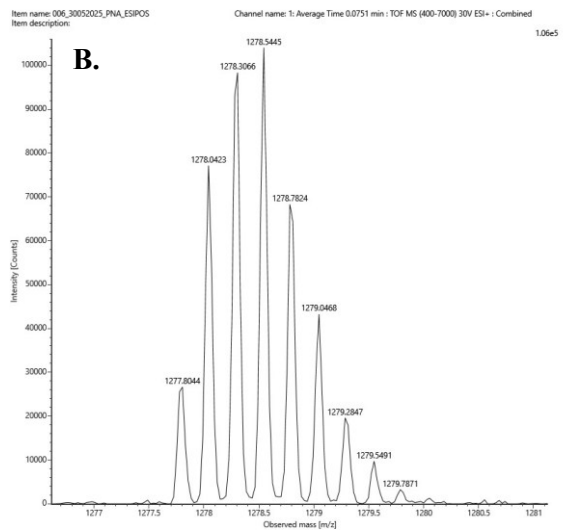
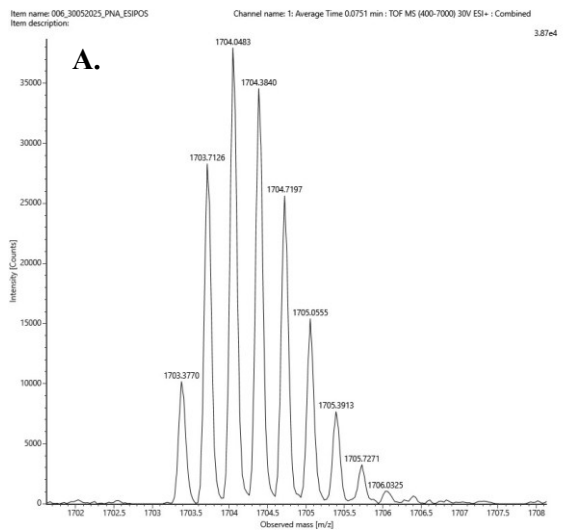
Materials and Methods

- UV-Vis measurements were carried out on a IMPLIN NP80 UV/Vis spectrophotometer by using 1 cm quartz cells of 1 mL internal volume (Hellma Analytics), at a 0.5°C/min scan speed, recording at 260 nm every 5°C, from 15°C to 95°C.
- NMR Spectroscopy 2D ^1H NMR (TOCSY and NOESY) spectra were recorded on a Bruker Avance DMX 600 and 400 MHz. The solvent used for the analysis is TFE, d2.

Appendix

Chapter I

H-AntimiRNA-221-5p



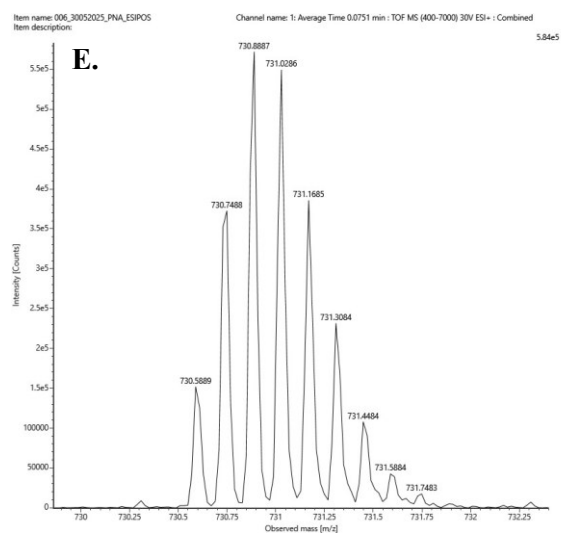


Figure 1A.1. MS Spectrometry of anti-miRNA-221-5p **A)** $[M+3H]^{3+}$; **B)** $[M+4H]^{4+}$; **C)** $[M+5H]^{5+}$; **D)** $[M+6H]^{6+}$; **E)** $[M+7H]^{7+}$.

Alkyne-AntimiRNA-221-5p

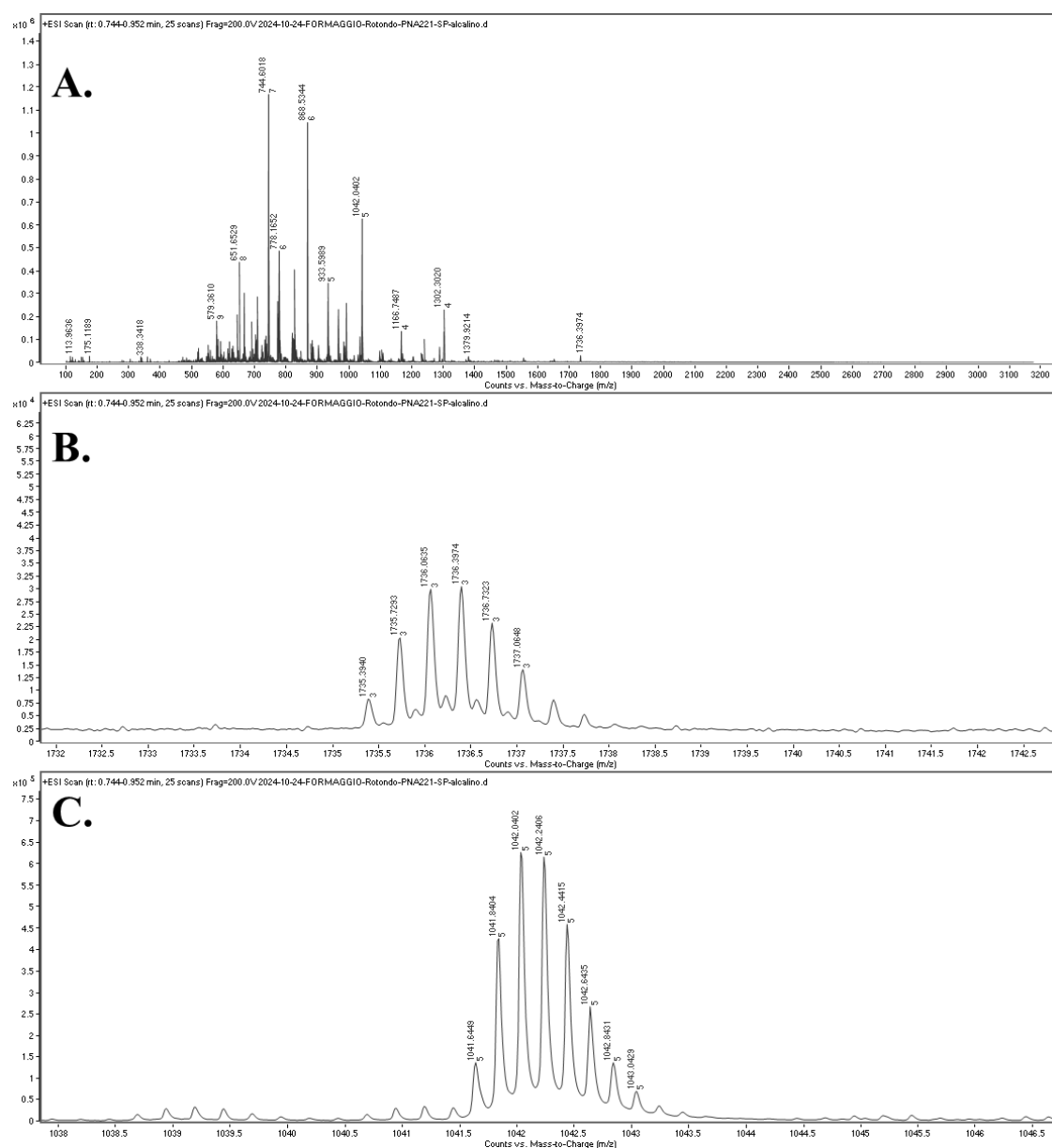


Figure 1A.2. A) MS Spectrometry of anti-miRNA-221-5p B) $[M+5H]^5+$; C) $[M+6H]^6+$.

H-Scramble-KK-NH₂ Sequence

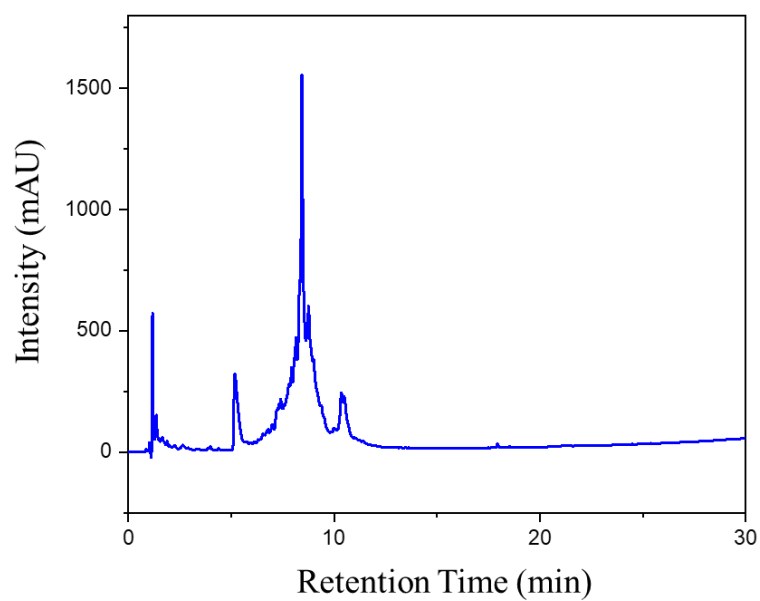


Figure 1A.3. Analytical HPLC chromatogram of the H-Scramble-KK-NH₂ PNA crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

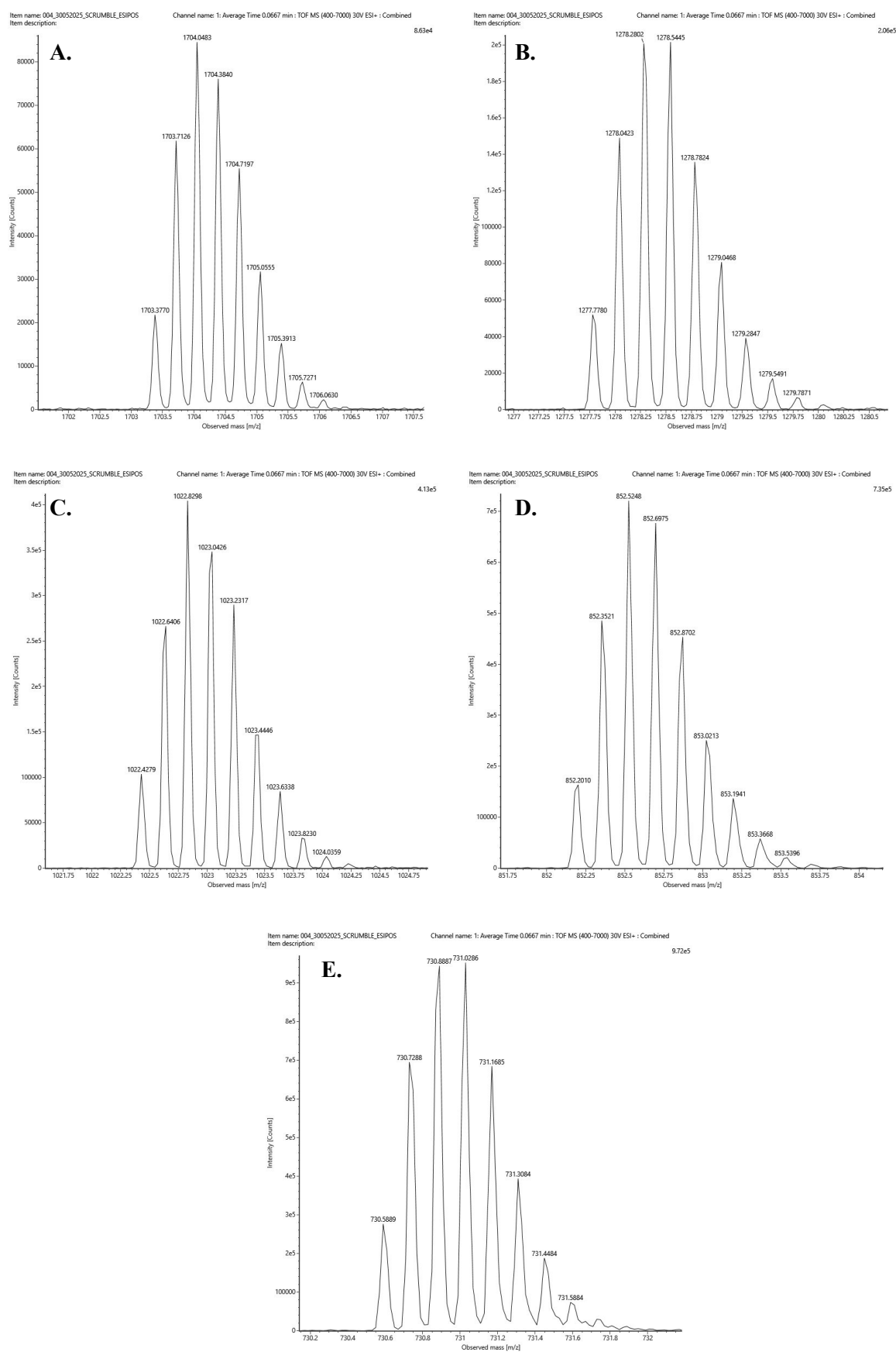


Figure 1A.4. MS Spectrometry of H-Scramble-KK-NH₂ PNA **A)** $[M+3H]^3+$; **B)** $[M+4H]^4+$; **C)** $[M+5H]^5+$; **D)** $[M+6H]^6+$; **E)** $[M+7H]^7+$.

CF-Tat-N₃-NH₂ Sequence

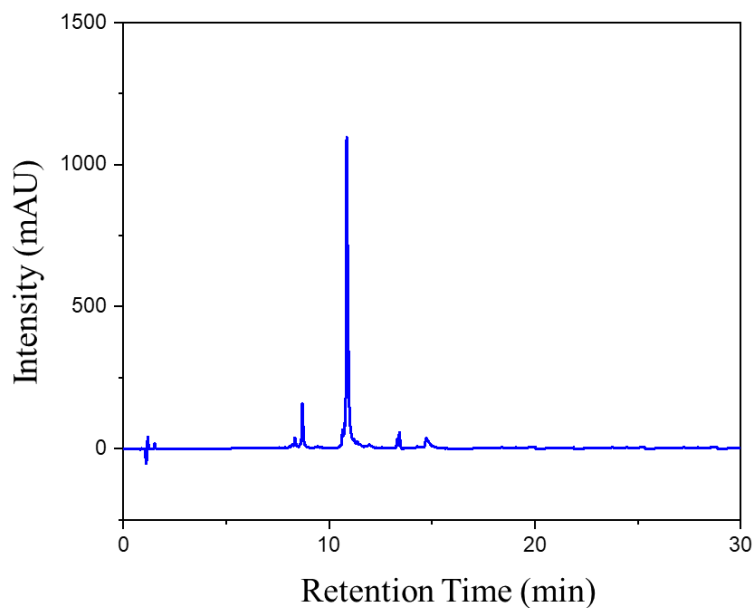


Figure 1A.5. Analytical HPLC chromatogram of the CF-Tat-N₃-NH₂ crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 45% over 30 minutes.

CF-PNT-N₃-NH₂

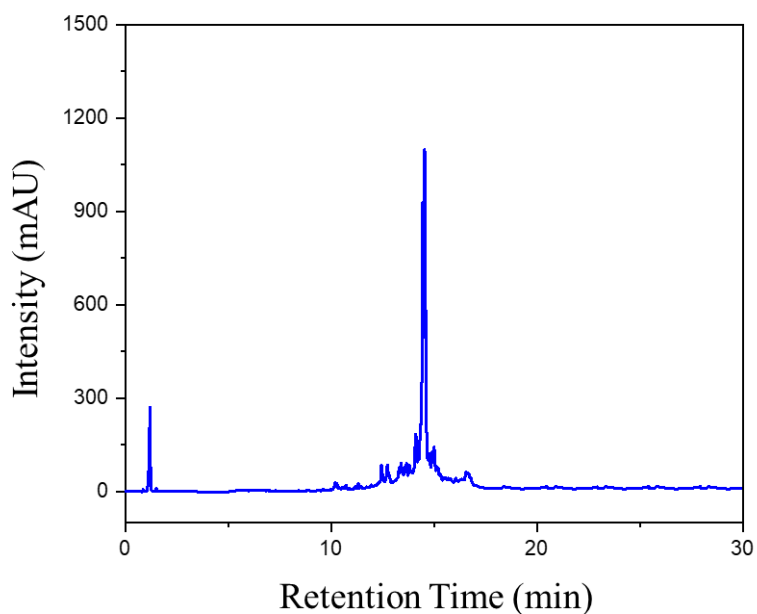


Figure 1A.6. Analytical HPLC chromatogram of CF-PNT-N₃-NH₂ recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 45% over 30 minutes.

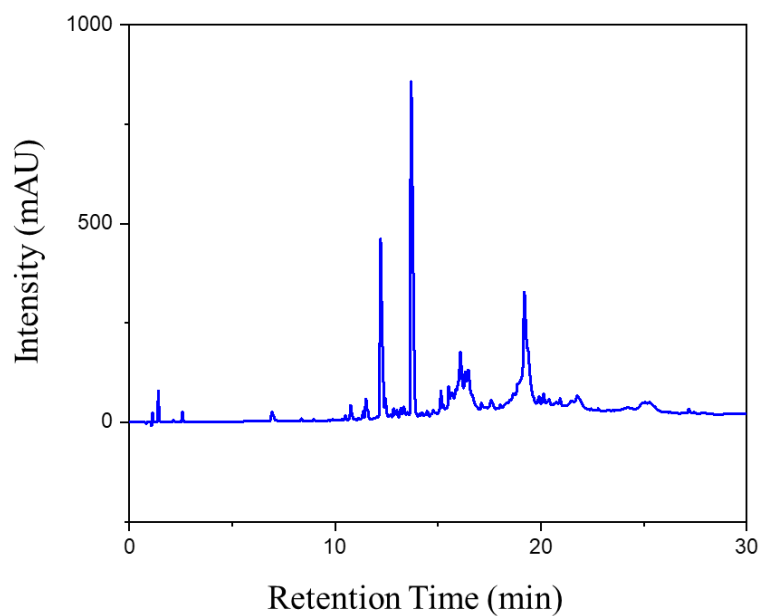
CF-P14-N₃-NH₂

Figure 1A.7. Analytical HPLC chromatogram of the CF-P14-N₃-NH₂ crude recorded using a Kinetex C18 column, 1 mL/min flow rate, with detection at 220 nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 45% over 30 minutes.

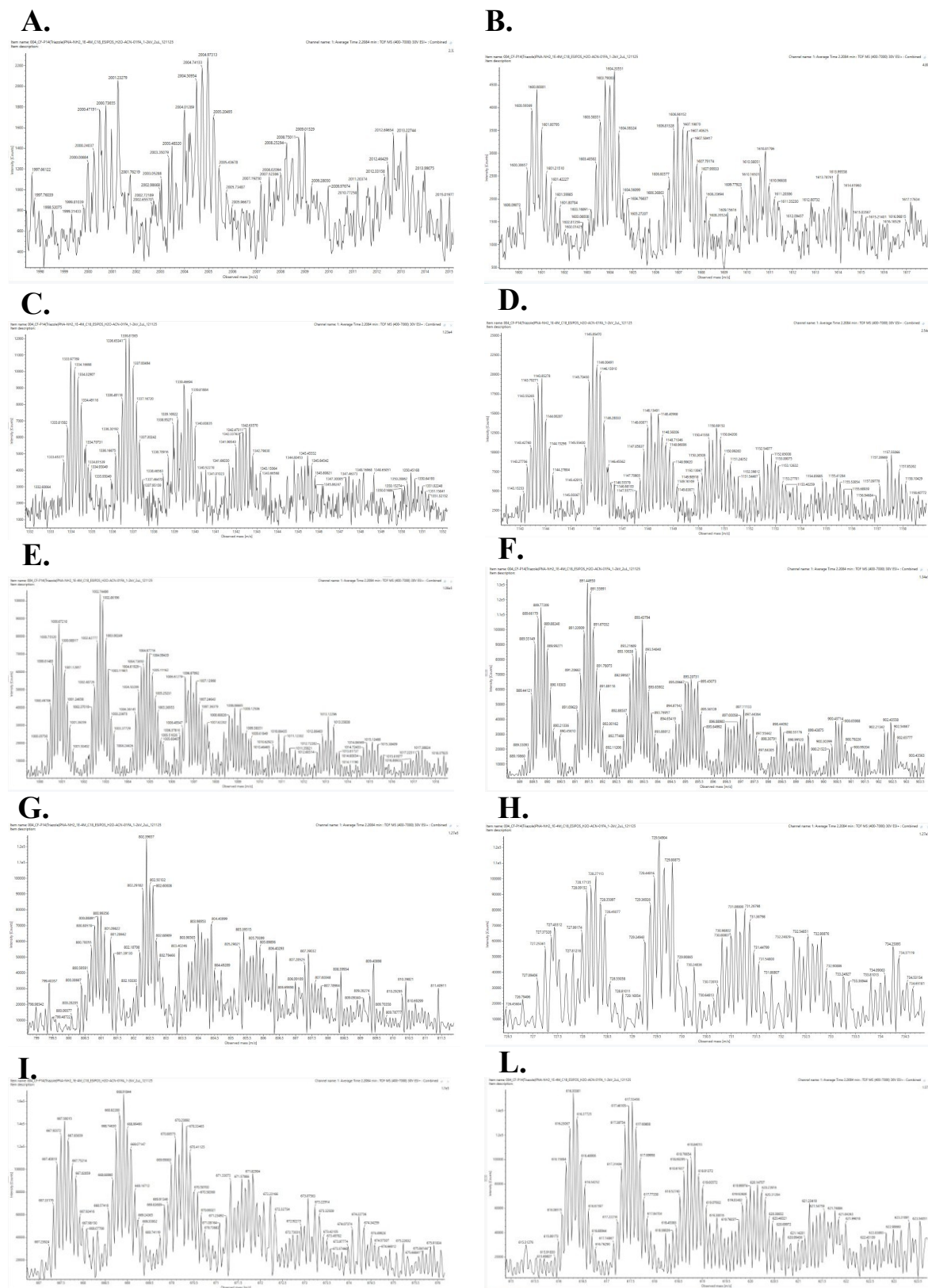
CF-P14-(Triazole)-PNA-NH₂

Figure 1A.8. MS Spectrometry of H-Scramble-KK-NH₂ PNA A) $[M+4H]^{4+}$; B) $[M+5H]^{5+}$; C) $[M+6H]^{6+}$; D) $[M+7H]^{7+}$; E) $[M+8H]^{8+}$; F) $[M+9H]^{9+}$; G) $[M+10H]^{10+}$; H) $[M+11H]^{11+}$; I) $[M+12H]^{12+}$; L) $[M+13H]^{13+}$.

Cell Proliferation Assay

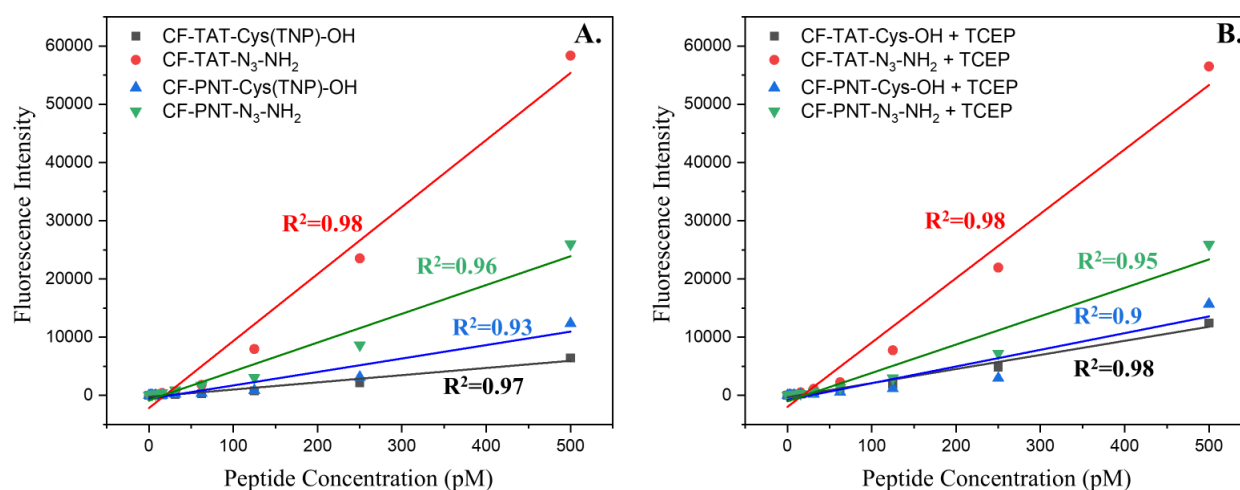
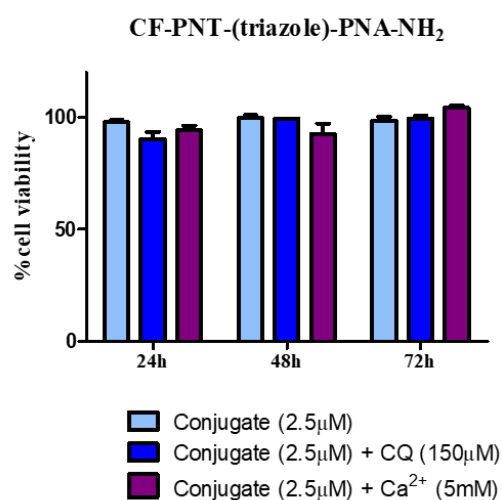


Figure 1A.9. Fluorescence assay of CF-CPPs in TRIS buffer in absence (left) and presence (right) of TCEP.

A.



B.

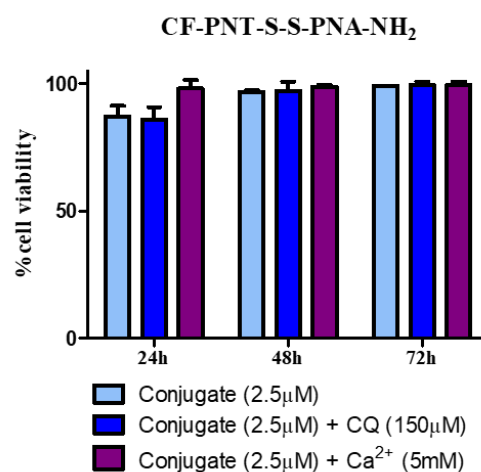


Figure 1A.10. A) CF-PNT(triazole)PNA-NH₂ and B) CF-PNT-S-S-PNA-NH₂ were tested at a final concentration of 2.5μM, either alone or in combination with chloroquine (CQ, μM) or calcium ions (Ca²⁺, 5 mM). Cell viability was evaluated after 24h, 48h and 72h using the CCK-8 assay. The percentage of viable cells is expressed relative to the corresponding control, consisting of cells incubated for the same duration under identical conditions (complete culture medium and serum), but without peptide or additives.

Internalization Assay

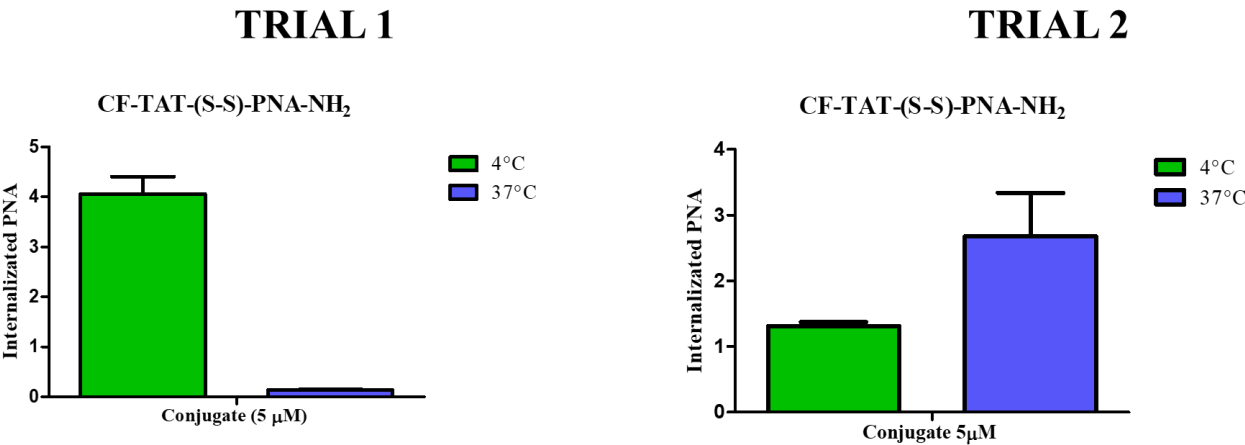


Figure 1A.11. Internalization of CF-Tat-S-S-PNA-NH₂ alone and in the presence of CQ (150 μ M) in A549 cells.

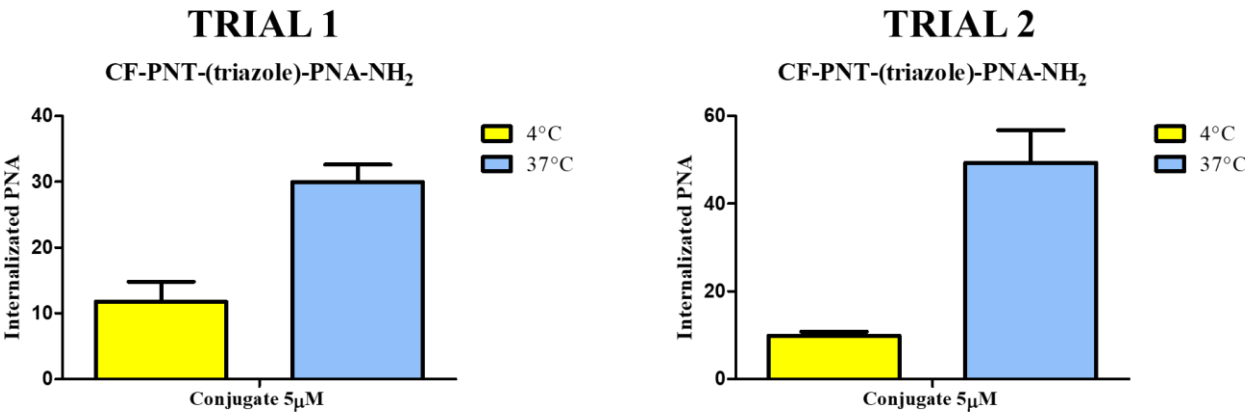


Figure 1A.12. Internalization of CF-PNT(triazole)PNA-NH₂ alone and in the presence of CQ (150 μ M) in A549 cells.

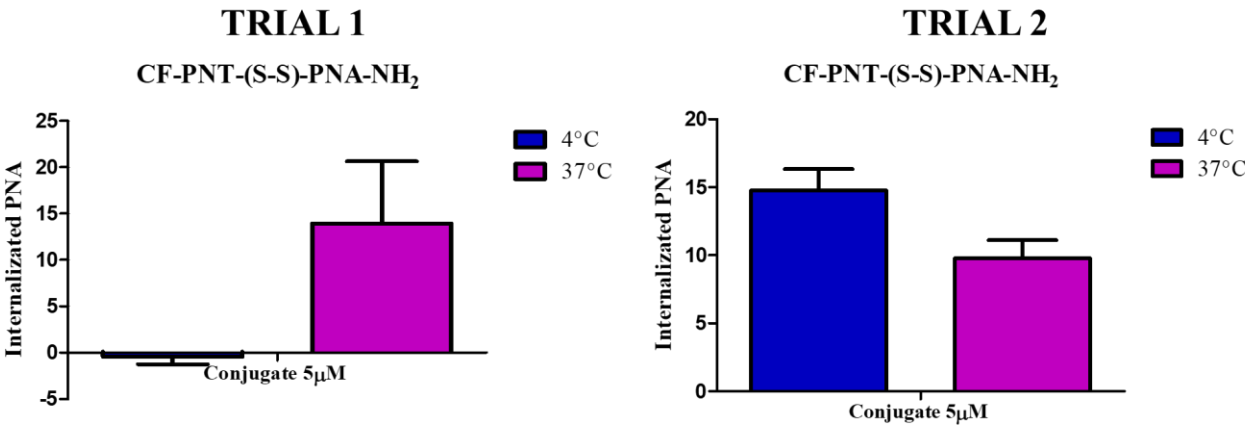


Figure 1A.13. Internalization of CF-PNT-S-S-PNA-NH₂ alone and in the presence of CQ (150 μ M) in A549 cells.

Chapter II

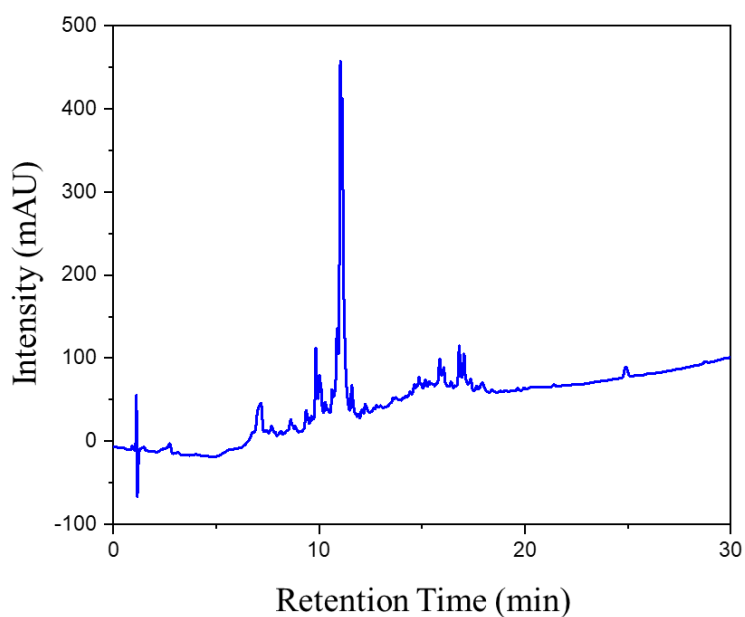
Ac-11-30-NH₂

Figure 2A.1. Analytical HPLC chromatogram of the Ac-11-30-NH₂ peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

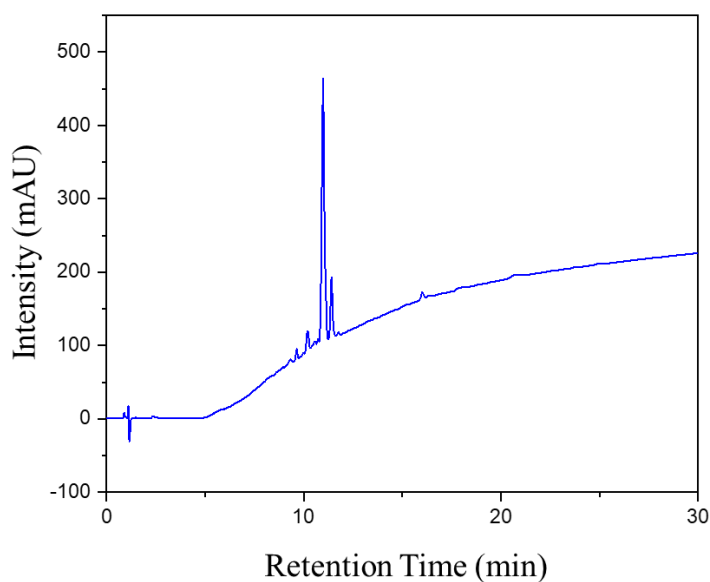
H-68-79-NH₂

Figure 2A.2. Analytical HPLC chromatogram of the H-68-79-NH₂ peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

Ac-68-79-NH₂

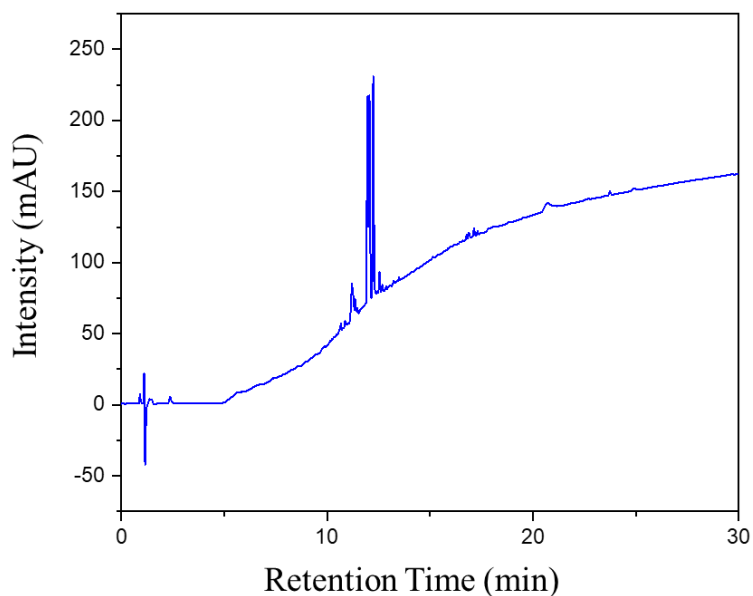


Figure 2A.3. Analytical HPLC chromatogram of the Ac-68-79-NH₂ peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

H-270-280-NH₂

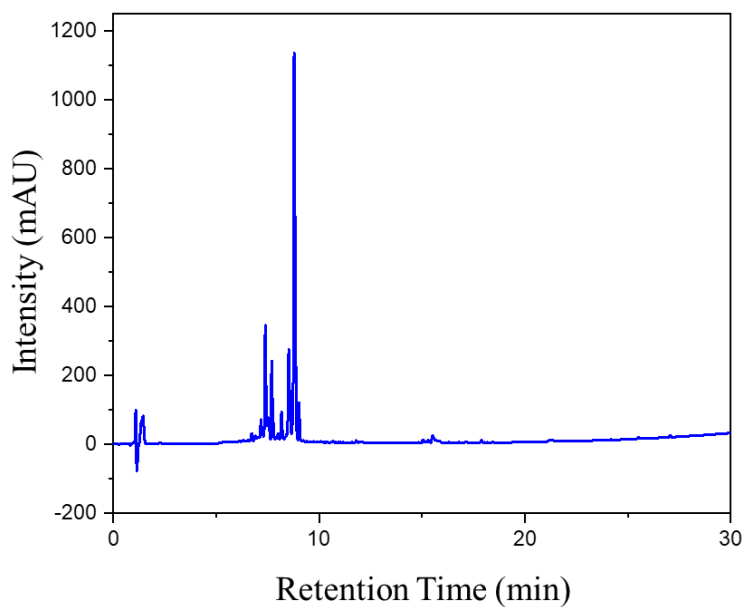


Figure 2A.4. Analytical HPLC chromatogram of the H-270-280-NH₂ peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

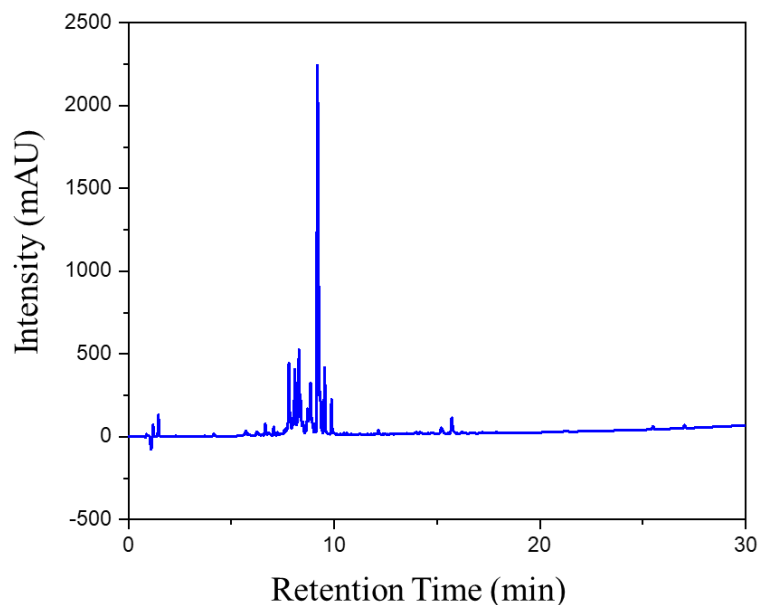
Ac-270-280-NH₂

Figure 2A.5. Analytical HPLC chromatogram of the Ac-270-280-NH₂ peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

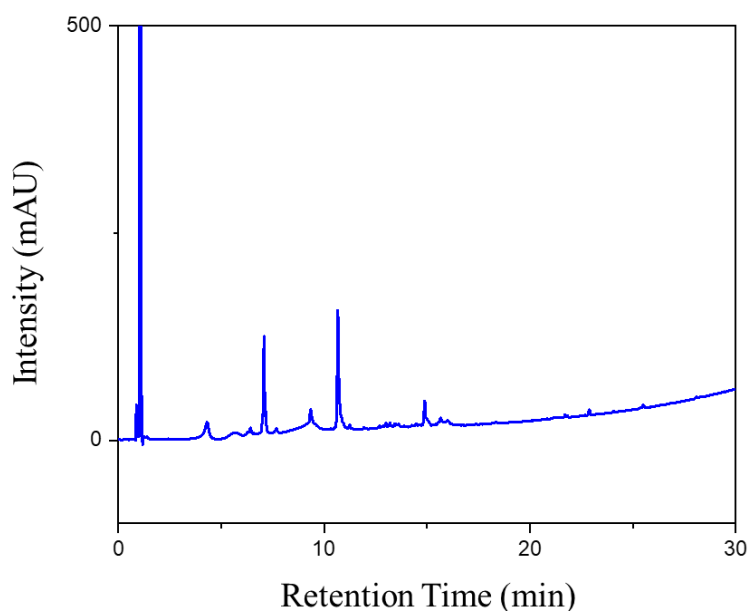
H-[Nle¹⁴]-11-30-NH₂ peptide

Figure 2A.6. Analytical HPLC chromatogram of the H-[Nle¹⁴]-11-30-NH₂ peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

H-[Ala¹¹, Nle¹⁴]-11-30-NH₂ peptide

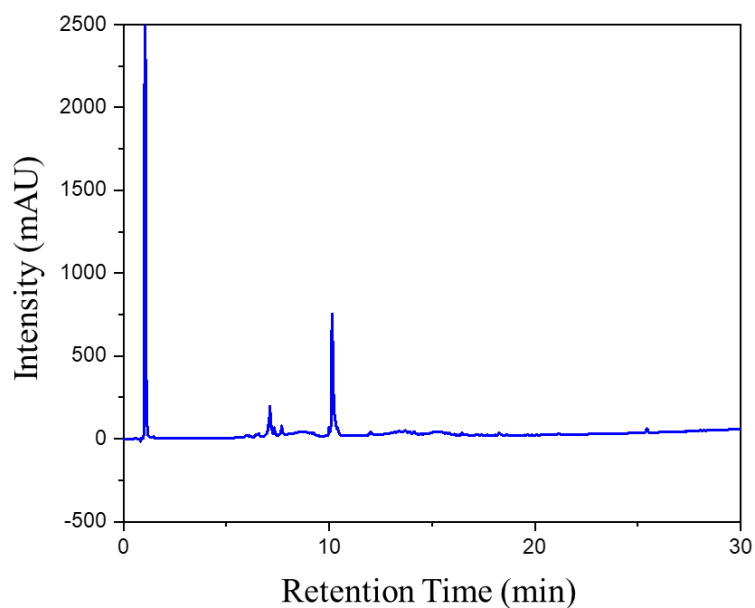


Figure 2A.7. Analytical HPLC chromatogram of the H-[Ala¹¹, Nle¹⁴]-11-30-NH₂ peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

H-[Ala¹⁵, Nle¹⁴]-11-30-NH₂

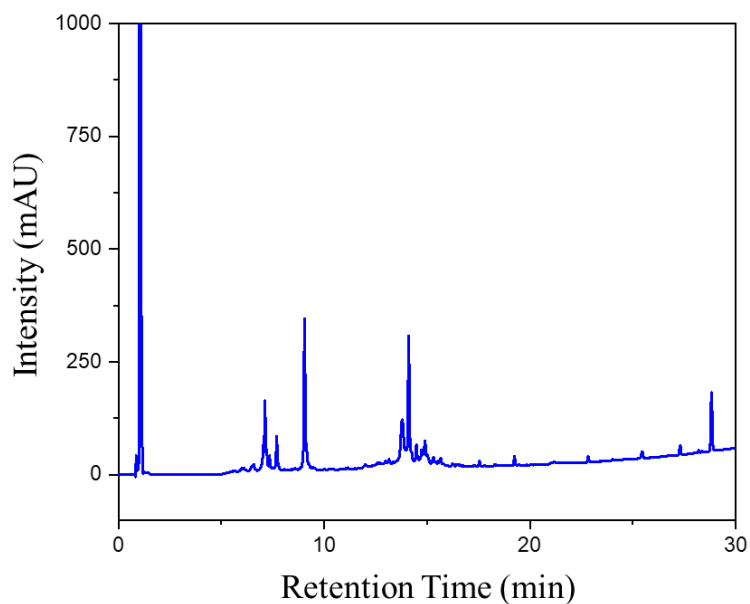


Figure 2A.8. Analytical HPLC chromatogram of the H-[Ala¹⁵, Nle¹⁴]-11-30-NH₂ peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

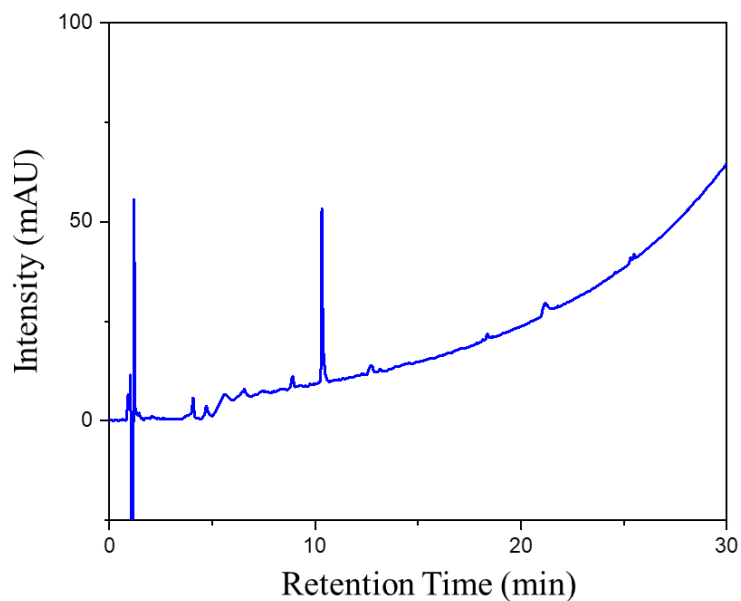
H-[Ala³⁰, Nle¹⁴]-11-30-NH₂

Figure 2A.9. Analytical HPLC chromatogram of the H-[Ala³⁰, Nle¹⁴]-11-30-NH₂ peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

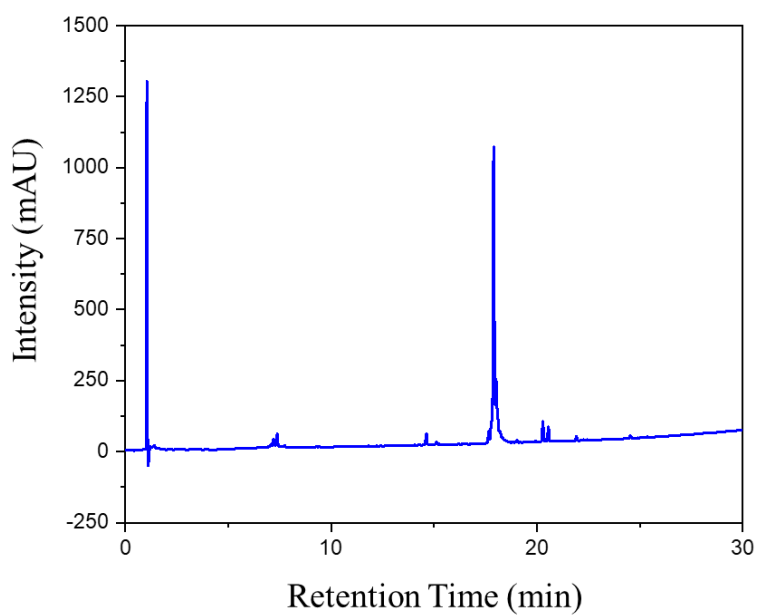
Fmoc-[Ala^{12,13,23,28}, Nle¹⁴]-11-30-NH₂

Figure 2A.10. Analytical HPLC chromatogram of the Fmoc-[Ala^{12,13,23,28}, Nle¹⁴]-11-30-NH₂ peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

Fmoc-[Ala^{12,13}, Nle¹⁴]-11-30-NH₂

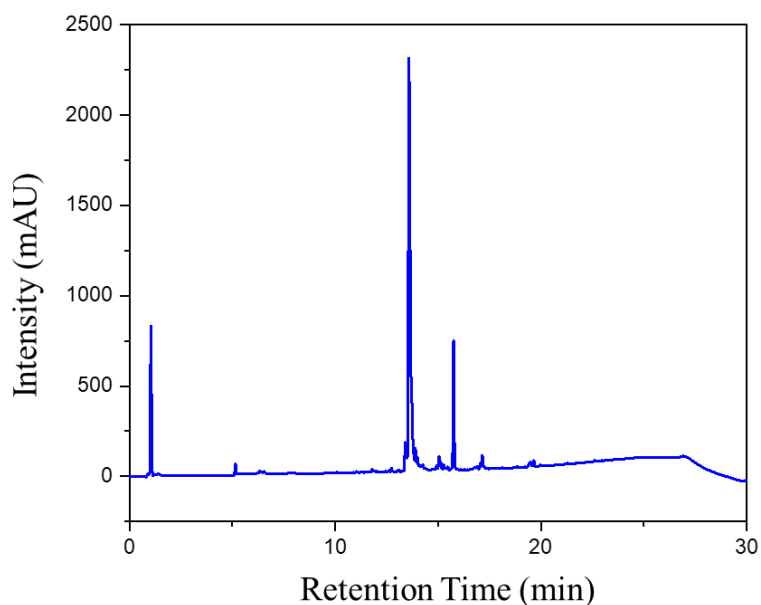


Figure 2A.11. Analytical HPLC chromatogram of the Fmoc-[Ala^{12,13}, Nle¹⁴]-11-30-NH₂ peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

H-[Ala^{23,28}, Nle¹⁴]-11-30-NH₂

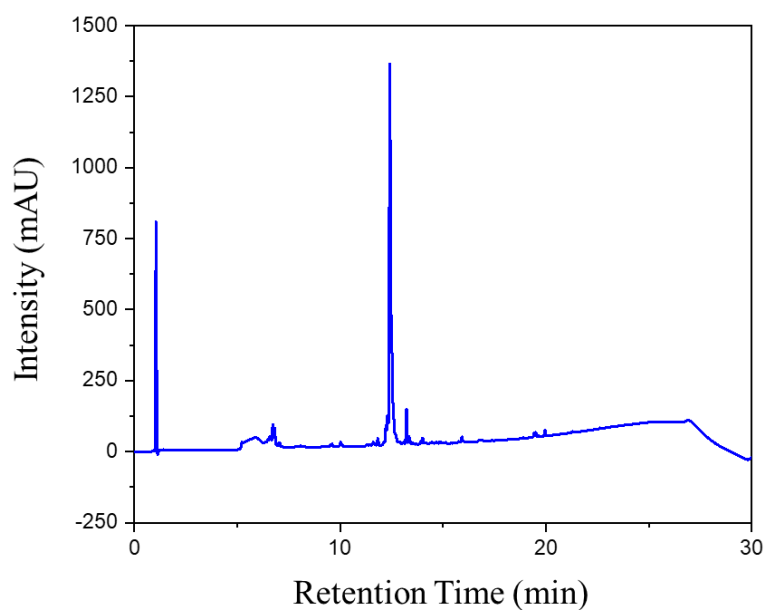


Figure 2A. 12. Analytical HPLC chromatogram of the H-[Ala^{23,28}, Nle¹⁴]-11-30-NH₂ peptide crude recorded using a Kinetex C18 column, 1mL/min flow rate, with detection at 220nm. Elution was initially performed at 5% of solvent B, followed by a linear gradient from 5% to 95% over 30 minutes.

NMR Spectra

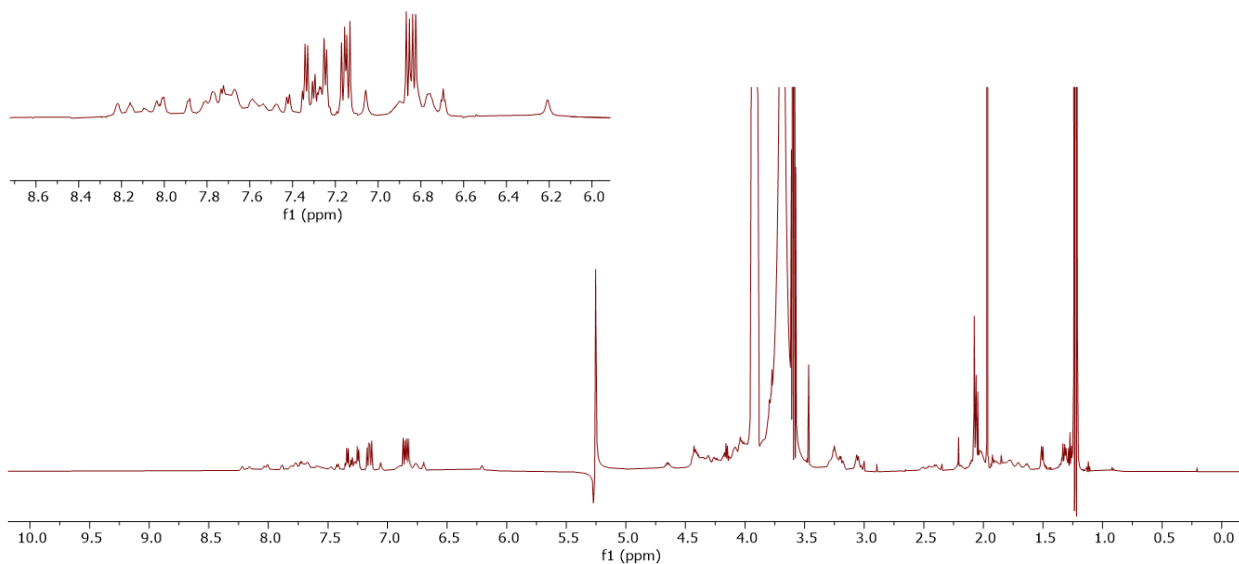
Ac-11-30-NH₂

Figure 2A.13. ¹H-NMR of Ac-11-30-NH₂ peptide in d₂-TFE.

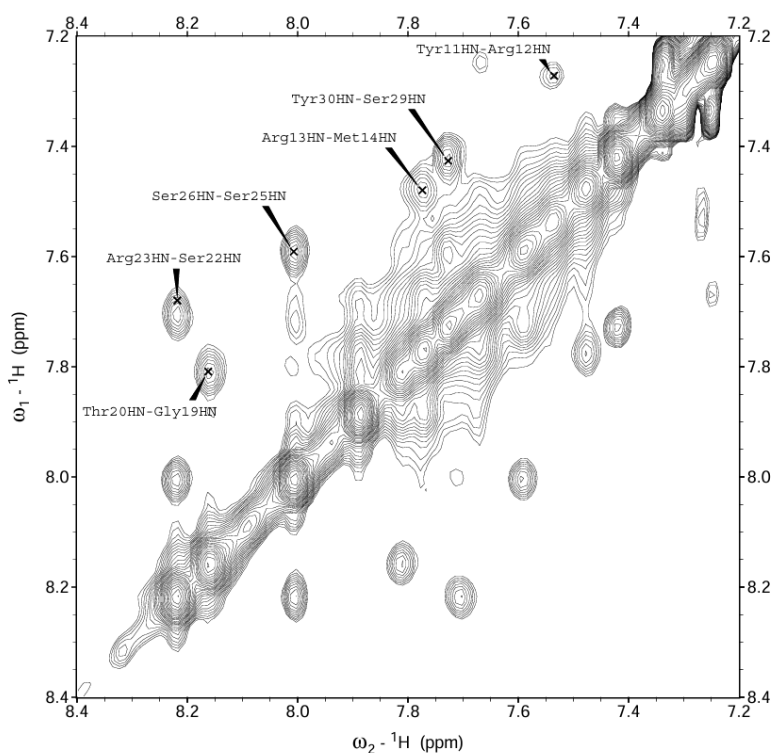


Figure 2A.14. NOESY spectrum of Ac-11-30-NH₂ peptide in TFE, d₂ showing the connectivities of the amino acids.

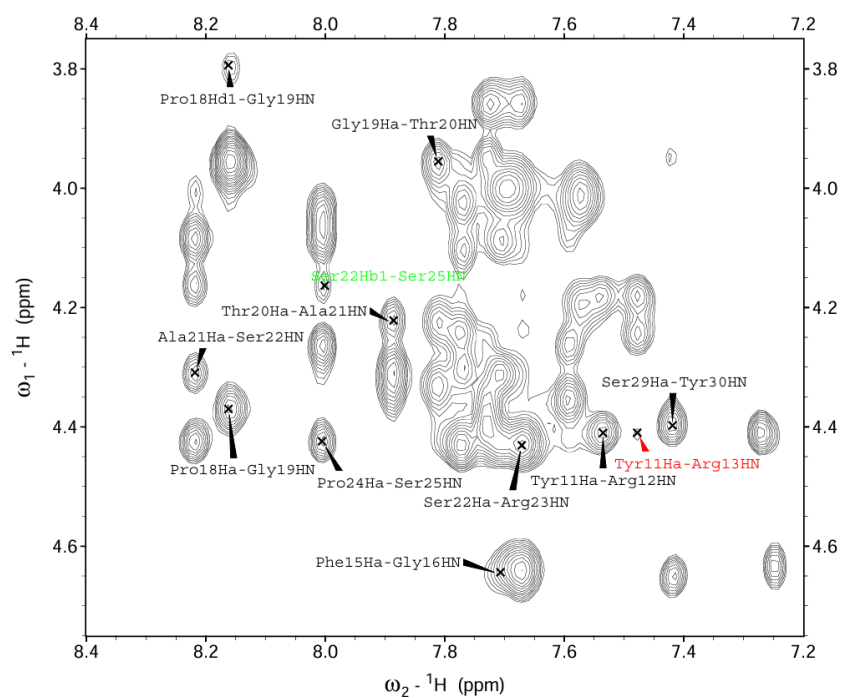


Figure 2A.15. NOESY spectrum of Ac-11-30-NH₂ peptide in TFE, d₂ showing the connectivities of the amino acids.

Table 2A.1. Proton resonances for Ac-11-30-NH₂ peptide.

H	NH	H α	H β	Other			
Ac	-	-	-	2.060			
NH ₂	7.056	-	-	-			
Tyr11	-	7.272	-	α 4.410	β 3.061		
Arg12	-	7.536	-	α 4.181	β 1.898	δ 1.783	γ 1.635
Arg13	-	7.478	-	α 4.245			
Met14	-	7.775	-	α 4.431	β 2.015	γ 2.512, 2.459	
Phe15	-	7.641	-	α 4.645	β 3.238	δ 3.031	
Gly16	-	7.707	-	α 4			
Gly17	-	7.571	4.0	α 4	β 4.0		
Pro18	-	-	-	α 4.372	β 2.396	δ 2.031	γ 3.657, 2.192, 2.097
Gly19	-	8.161	-	α 3.955			
Thr20	-	7.812	-	α 4.223	γ 1.310		
Ala21	-	7.887	-	α 4.309	β 1.506		
Ser22	-	8.218	-	α 4.428	β 4.164	δ 4.083	
Arg23	-	7.671	-	α 4.643	β 2.033	δ 3.238	γ 3.043, 1.829
Pro24	-	-	-	α 4.433	β 2.418	δ 4.030	γ 3.859, 2.225, 2.083
Ser25	-	8.004	-	α 4.262	β 4.083		
Ser26	-	7.592	-	α 4.355			
Ser27	-	7.769	-	α 4.443	β 4.106		
Arg28	-	7.595	-	-	β 2.016	γ 1.777	
Ser29	-	7.727	-	α 4.398	β 3.947		
Tyr30	-	7.420	-	α 4.651	β 3.185		

Ac-68-79-NH₂

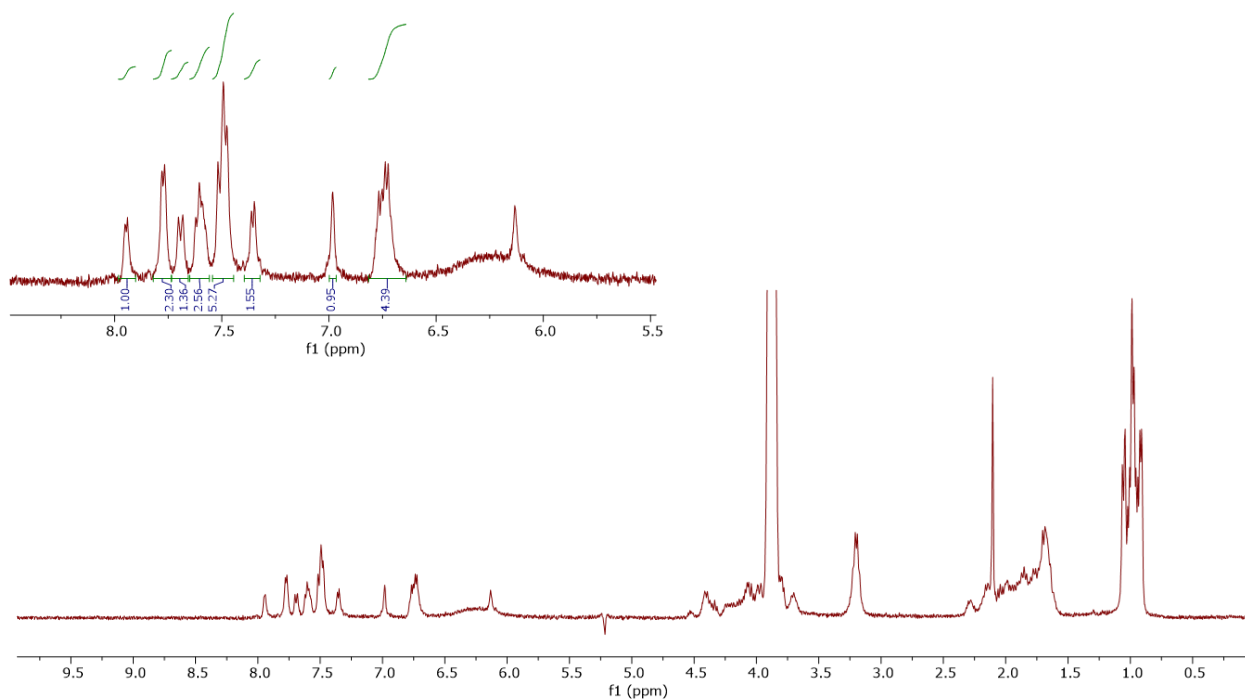


Figure 2A.16. ¹H-NMR of Ac-68-79-NH₂ peptide in d₂-TFE.

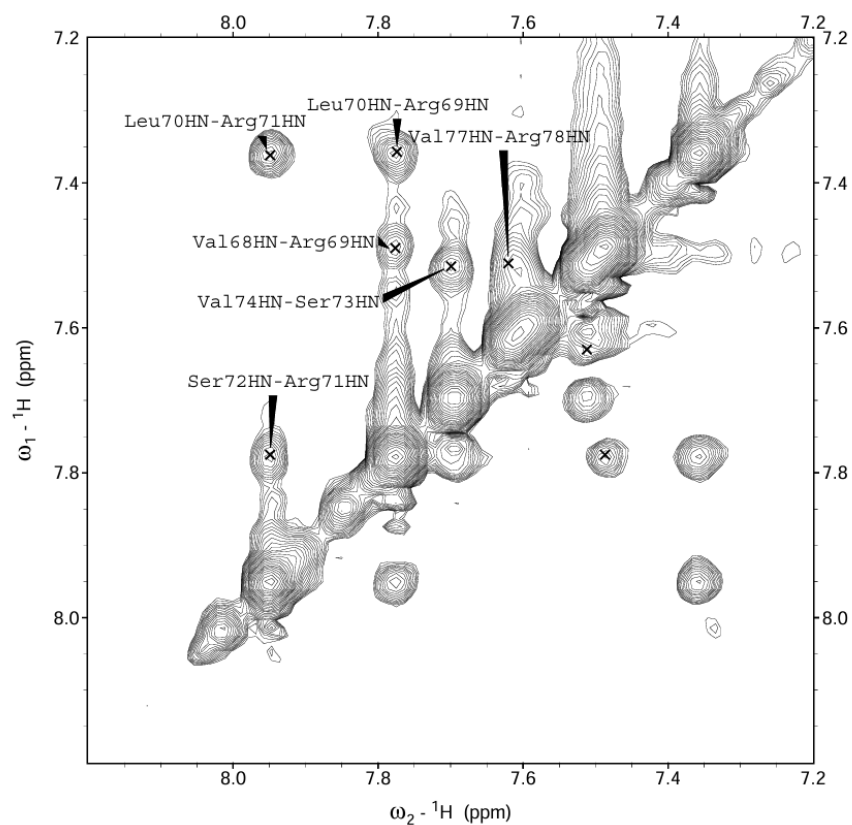


Figure 2A.17. NOESY spectrum of Ac-11-30-NH₂ peptide in TFE, d₂ showing the connectivities of the amino acids.

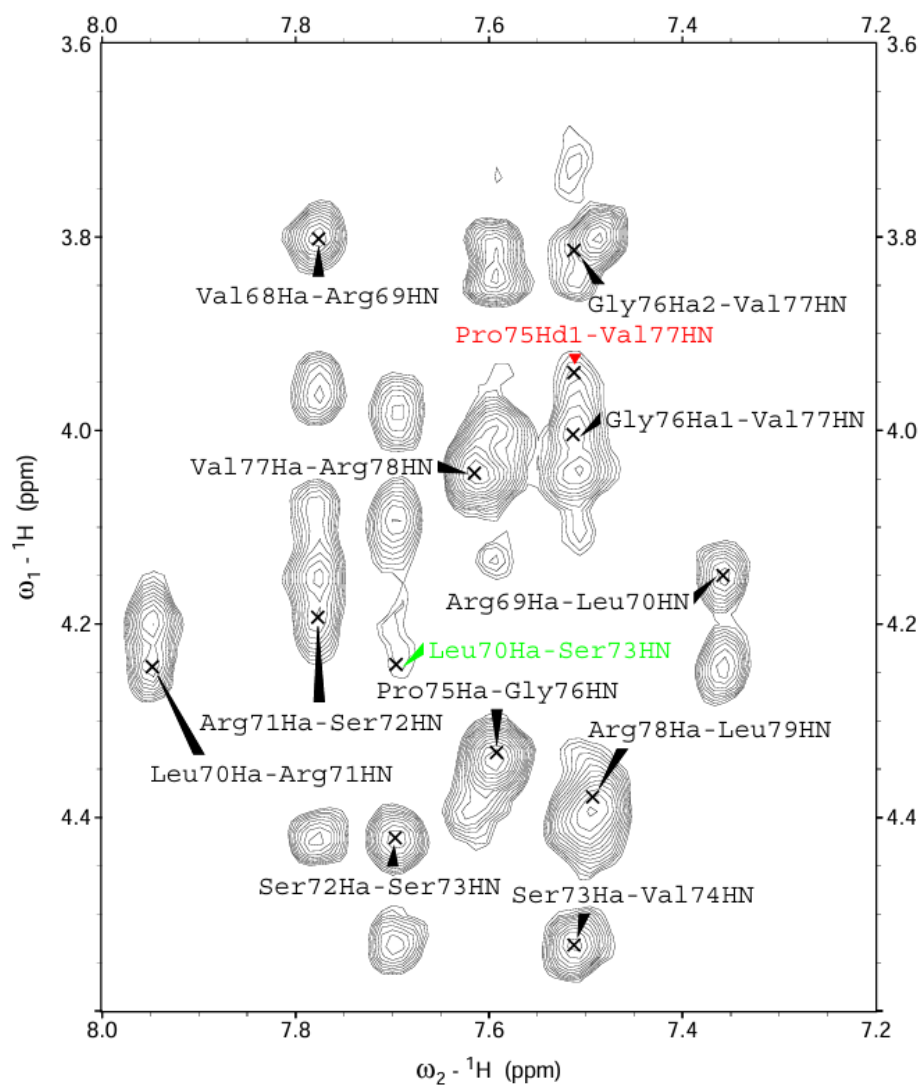


Figure 2A.18. NOESY spectrum of Ac-11-30-NH₂ peptide in TFE, d₂ showing the connectivities of the amino acids.

Appendix

Table 2A.2. Proton resonances for Ac-68-79-NH₂ peptide.

H	NH	H α	H β	Other
Ac	-	-	-	2.1, 3
Val68	7.486	3.802	2.050	γ 1.697, 1.621
Arg69	7.776	4.150	1.864	γ 1.740, 1.675
Leu70	7.359	4.243	1.771	δ 0.9890, 0.9200 γ 1.674
Arg71	7.949	4.193	-	β 1.957, 1.911 γ 1.776
Ser72	7.777	4.421	-	
Ser73	7.696	4.532	-	
Val74	7.513	-	-	
Pro75	-	4.333	-	β 2.298, 2.002 δ 3.939, 3.710 γ 2.146
Gly76	7.592	-	-	α 4.004, 3.814
Val77	7.513	4.044	2.164	γ 1.047, 0.9880
Arg78	7.616	4.379	-	β 1.909, 1.836 δ 0.9840, 0.9190 γ 1.047, 0.9881
Leu79	7.489	-	-	