Università di Napoli Federico II

Dipartimento di Farmacia

DOTTORATO DI RICERCA IN

SCIENZA DEL FARMACO

XXVIII CICLO

Design and Synthesis of Peptide and Peptidomimetics Involved in Antiretroviral and Anticancer Therapies

Coordinatore
Chiar.ma Prof. Maria Valeria D'Auria

Tutor
Chiar.mo Prof. Vincenzo Santagada

Dottoranda
Irene Saccone

Contents

Unit 1 - **Inhibitors of the transactivation mechanism of Tat-TAR and of the chaperonic protein NC**

Unit 2 - **Synthesis of a PEGylated ligation auxiliary and its application in the preparation of homogeneously glycosylated linker domain of IGF-BP6 protein**

“La mente è come un paracadute:

Funziona solo se si apre”

Albert Einstein

Alla mia famiglia:

Insieme di mattoni di unione, pazienza ed amore

RIASSUNTO

Il mio lavoro di dottorato è stato incentrato sulla progettazione e sintesi di peptidi e peptidomimetici coinvolti nella terapia antiretrovirale ed antitumorale.

La prima parte del mio lavoro, della durata di due anni e mezzo, è stata svolta presso il Dipartimento di Farmacia dell'Università Federico II di Napoli, presso il gruppo di ricerca guidato dal Prof. Vincenzo Santagada ed è stata focalizzata alla ricerca di nuovi potenti inibitori *in vitro* della proteina NC del virus dell'HIV-1.

La necessità di disporre di nuovi farmaci anti-HIV costituisce sicuramente un'esigenza di interesse globale. Gli effetti collaterali e lo sviluppo di fenomeni di resistenza hanno limitato molto l'impiego terapeutico delle classi di medicinali attualmente disponibili, per cui l'attenzione del mondo della ricerca è da tempo rivolta verso nuovi target per lo sviluppo di nuovi farmaci antiretrovirali.

La nostra attenzione si è rivolta, pertanto, ad un innovativo eccellente target molecolare, la proteina chaperonica nucleocapsidica (NC), costituita da 55 aminoacidi basici, caratterizzata dalla presenza di due domini "zinc-fingers" altamente conservati, che gioca un importante ruolo in diversi steps della replicazione del virus dell'HIV, tra cui il processo di trascrizione.

La proteina NC svolge un ruolo di mediatore nella reazione di *annealing* dell'acido nucleico TAR RNA con la sua copia di DNA cTAR attraverso la fusione

delle loro strutture secondarie, per la formazione di un ibrido più stabile di RNA/DNA.

La protein Tat (elemento transattivatore nel processo di trascrizione) aumenta la frequenza della trascrizione a seguito del legame alla regione TAR all' estremità 5' del trascritto nascente. In assenza di Tat, il processo di trascrizione inizia normalmente ma si riduce l' efficienza del processo, infatti la polimerasi non trascrive più di 100 nucleotidi.

Al fine di individuare nuovi potenti inibitori della proteina NC, durante il mio periodo di dottorato mi sono occupata della sintesi di tre nuove serie di molecole. In particolare, basandoci su precisi requisiti strutturali, consistenti nella presenza di sostituenti amminoacidici carichi positivamente nelle posizioni 2,6 dell'anello antrachinonico, mi sono occupata della sintesi in soluzione di tre nuove serie di composti denominate 2,6-AQ-Gly-X (composti **4a**-**4l**),2,6-AQ-L-Ala-X (composti **5a**-**5l**) e 2,6-AQ-D-Ala-X (**6a**-**6l**) caratterizzate dalla presenza di *linker* differenti tra lo *scaffold* antrachinonico e le catene laterali aminoacidiche.

Le nuove serie di composti sono state saggiate, al fine di verificare la loro capacità di inibire la formazione del complesso Tat-TAR e l'attività chaperonica della proteina NC, da parte del gruppo di ricerca della Prof.ssa Barbara Gatto presso il Dipartimento di Scienze Farmaceutiche dell'Università degli Studi di Padova, mediante studi di: *High Throughput Screening* (HTS), *Fluorescence Quenching Assay* (FQA); Saggio NAME (*Nucleocapsid Annealing-Mediated Electrophoreis*).

Inoltre, al fine di ampliare lo spettro di informazioni relative all'interazione tra i composti sintetizzati e gli acidi nucleici di interesse, sono stati eseguiti studi di ESI-MS presso l'RNA Institute – University of New York sotto la supervisione del Prof. Dan Fabris.

I composti sono stati ottenuti seguendo un'eccellente strategia sintetica che ha condotto ad alte rese e purezza elevata.

Confrontando le strutture chimiche delle due serie di derivati dei quali ad oggi si sono resi disponibili i risultati dell'attività biologica (serie 2,6-AQ-Gly-X, composti da **4a** a **4i** e serie 2,6-AQ-L-Ala-X, composti da **5a** a **5i**), caratterizzate dalla presenza di *linker* diversi, possiamo concludere che la presenza di un residuo di ornitina in catena laterale, conferisce caratteristiche favorevoli per potenziare l'attività inibitoria di entrambe le serie di composti antrachinonici testati.

Questo assunto è confermato dal fatto che i composti **4g** e **5g** presentano un'elevata attività sia nel saggio NAME che in quello di HTS. Dai risultati emerge altresì che il derivato **4h,** caratterizzato da una lisina come residuo amminoacidico terminale**,** rappresenti il composto più promettente della serie con la glicina come spacer.

La considerazione che la catena della lisina, più lunga di un unità metilenica rispetto a quella dell'ornitina, conferisca migliori proprietà di inibizione della proteina NC, suggerisce che la lunghezza della catena gioca sicuramente un ruolo fondamentale sull'attività di tali composti.

In conclusione, abbiamo dimostrato che questi nuovi derivati interagiscono preferenzialmente con le strutture non-canoniche di TAR e cTAR, stabilizzano le loro strutture dinamiche e interferiscono con l'attività della proteina NC. Il *lead compound* è rappresentato dal composto **4h**, che agisce come legante nei confronti sia di TAR che di cTAR, così come potente inibitore della proteina NC (IC_{50}= 2.12 μM).

La seconda parte del mio lavoro di dottorato è stata svolta presso il Dipartimento di Chimica Biologica dell'Università di Vienna, presso il gruppo di ricerca del Prof. Christian Becker, avendo come tutor la Dottoressa Claudia Bello ed è stata incentrata su nuove strategie sintetiche per la glicosilazione di un frammento della proteina IGFBP6.

In passato sono stati sviluppati molti metodi per ottenere *in vitro* peptidi e proteine selettivamente glicosilati con saccaridi complessi, basati su strategie di sintesi chimica o usando altri approcci basati sull´utilizzo di enzimi. Gli approcci enzimatici, tuttavia, hanno come svantaggio il fatto di sviluppare reazioni di glicosilazione non quantitative a causa del limitato accesso degli enzimi ai substrati, soprattutto quando la glicosilazione è condotta in fase solida, cioè con il substrato ancora legato alla resina. Tale limitazione conduce all´ottenimento di rese basse dei prodotti di glicosilazione.

Per superare tale problemi in un primo approccio è stato proposto di modificare i peptidi inserendo il PEG alla loro estremitá N-terminale. Il peptide glicosilato poteva essere recuperato dopo rimozione proteolitica del PEG alla fine della glicosilazione.

Successivamente, tale nuova strategia basata sull'utilizzo del PEG, è stata associata all'introduzione di un ausiliario rimuovibile mediante radiazioni UV, che facesse da tramite per la glicosilazione e ligazione di diversi peptidi.

A tal fine mi sono occupata di sintetizzare tale ausiliario, utilizzato precedentemente dal gruppo di ricerca per la glicosilazione di analoghi della mucina, in rese più elevate al fine di poter applicare tale strategia ad un nuovo target, la proteina IGFBP6.

L´insulin growth factor (IGF), é un elemento fondamentale per la crescita fisiologica ed é anche implicato in numerose patologie incluso il cancro; se ne classificano due distinti sottoptipi denominati 1 e 2.

La sua attività è modulata da una famiglia di proteine dotate di elevata affinità per l´IGF-1 e IGF-2: in particolare IGFBP6, target di nostro interesse, si distingue per la propria marcata affinità di legame per l'IGF-2 rispetto all'IGF-1.

Il principale ruolo della proteina IGFBP6 consiste nell'inibizione delle azioni dell'IGF-2, ma studi recenti hanno anche indicato come presenti anche delle attività indipendenti rispetto a tale target, che includono l'inibizione del processo di angiogenesi e l'induzione alla migrazione delle cellule tumorali. IGFBP6 possiede caratteristiche chimiche peculiari: è composto da tre domini strutturali ed inoltre sia l´estremità N-terminale che quella C-terminale si presentano ricche in cisteine.

L'estremità C-terminale possiede azioni e funzioni indipendenti dall'IGF. Tutte le IGFBP sono soggette ad un gran numero di modifiche post-traslazionali che influenzano le loro azioni mediate da un gran numero di proteasi.

Le IGFBP3 e IGFBP4 vengono N-glicosilate, mentre IGFBP 5 e 6 vengono O-glicosilate. Il processo di glicosilazione non agisce direttamente influenzando le azioni di tali proteine, ma diminuisce la loro percentuale di legame nei confronti dei glicosamminoglicani, inibendo il loro processo proteolitico.

IGFBP6 è coinvolta nella patogenesi di fenomeni tumorali: cancro al seno, alla prostata, al colon, al collo e alla testa del pancreas, alle ovaie e nel neuroblastoma rappresentano infatti le patologie nelle quali il ruolo di tale proteina è stato elucidato chiaramente.

Pertanto, la sua inibizione può rivelarsi un interessante target per la terapia del cancro, soprattutto per quanto riguarda bambini o giovani pazienti adolescenti, dove una specifica inibizione del recettore IGF-2 risulta essere vantaggiosa per bloccare gli alti livelli di IGF-2 autocrino, senza andare a perturbare la crescita fisiologica mediata dalla presenza di IGF-1.

Mi sono quindi occupata della sintesi, mediante strategia sintetica in fase solida via Fmoc, di un frammento peptidico di 27 amminoacidi, parte della sequenza dell' insulin-like growth factor-binding protein6(IGFBP6):

127-ARPQDVNRRDQQRNPGTSTTPSQPNSA-153

A tale peptide è stato aggiunto in seguito un amminoacido modificato la Fmoc-Thr(GalNAc-Ac_3)-OH, al cui *coupling* è seguita una reazione di iodoacetilazione. L'ultima fase del progetto ha previsto l'esecuzione di tests preliminari per legare il peptide all'ausiliario. Attualmente il progetto viene continuato dal gruppo del Prof. Christian Becker e dalla Dottoressa Claudia Bello.

This PhD thesis is divided into two parts, the first part carried out at the Department of Pharmacy at the University Federico II of Naples, the second one at the Department of Biological Chemistry at the University of Vienna. Both of them are focused on design and synthesis of peptides and peptidomimetics involved in pharmacological therapies.

Abstract Unit 1

The development of new anti-HIV-1 drugs is a global need, due to the resistance and the side effects arising from the therapies currently in use, for this reason we focused on an innovative target, the nucleocapsidic chaperonic protein(NC) of HIV-1 virus, which is involved in many steps of the virus replication, expecially the trascription. Moreover it plays a role of mediator in the annealing reaction of TAR with cTAR.

2,6-Dipeptidyl-anthraquinones are a promising class of nucleic acid-binding compounds that act as NC inhibitors *in vitro*. We designed, synthesized and tested new series of 2,6-disubstituted-anthraquinones, which are able to bind viral nucleic acid substrates of NC. We demonstrated that these novel derivatives interact preferentially with non-canonical structures of TAR and cTAR, stabilize their dynamics structures, such as *bulge and loop*, and interfere with NC chaperone activity.

UNIT 1: Inhibitors of the transactivation mechanism of Tat-TAR and of the chaperonic protein NC.

Chapter 1

Introduction

1.1 Acquired immune deficiency syndrome (AIDS): human immunodeficiency virus type (HIV-1)

Acquired immunodeficiency syndrome (AIDS) of humans is caused by two *Lentiviruses*, human immunodeficiency viruses types 1 and 2 (HIV-1 and HIV-2) [1].

Both HIVs are the result of multiple cross-species transmissions of simian immunodeficiency viruses (SIVs) naturally infecting African primates. Most of these transfers resulted in viruses that spread in humans to only a limited extent. However, one transmission event, involving SIV from chimpanzees in southeastern Cameroon, gave rise to HIV-1 group M, the principal cause of the AIDS pandemic.

Acquired Immune Deficiency Syndrome (AIDS) was first recognized as a new disease in 1981 when increasing numbers of young homosexual men succumbed to unusual opportunistic infections and rare malignancies. A retrovirus, now termed human immunodeficiency virus type 1 (HIV-1), was subsequently identified as the causative agent of what has since become one of the most devastating infectious diseases to have emerged in recent history.

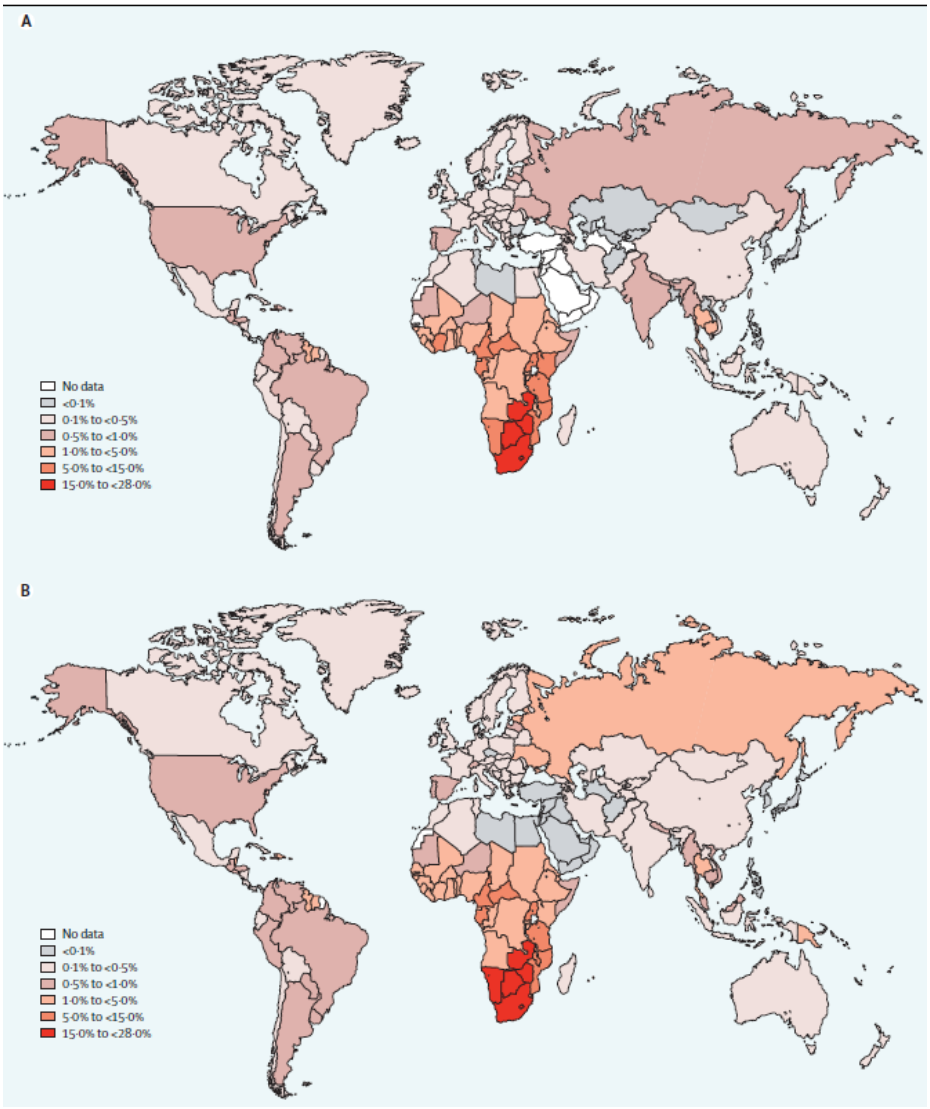


Figure 1: HIV prevalence in individuals aged 15-49 years.

HIV-1 spreads by sexual, percutaneous and perinatal routes, however, 80% of adults acquire HIV-1 following exposure at mucosal surfaces, and AIDS is thus primarily a sexually transmitted disease.

Since its first identification almost three decades ago, the pandemic form of HIV-1, also called the main (M) group, has infected at least 60 million people and caused more than 25 million deaths (Figure 1) [2].

Developing countries have experienced the greatest HIV/AIDS morbidity and mortality, with the highest prevalence rates recorded in young adults in sub-Saharan Africa. Although antiretroviral treatment has reduced the toll of AIDS- related deaths, access to therapy is not universal, and the prospects of curative treatments and an effective vaccine are uncertain. Thus, AIDS will continue to pose a significant public health threat for decades to come.

Ever since HIV-1 was first discovered, the reasons for its sudden emergence, epidemic spread and unique pathogenicity have been a subject of intense study. A first clue came in 1986 when a morphologically similar, but antigenically distinct, virus was found to cause AIDS in patients in western Africa [3].

1.2 The origin of the virus

HIV-1 has long been suspected to be of chimpanzee origin [4] however, until recently, the perceived lack of a chimpanzee reservoir left the source of HIV-1 open to question. These uncertainties have since been resolved by non-invasive testing of wild-living ape populations.

It is now well established that all naturally occurring SIVcpz (SIV chimpanzees) strains fall into two subspecies-specific lineages, termed SIVcpz*Ptt* and SIVcpz*Pts*, respectively, that are restricted to the home ranges of their respective hosts. Viruses from these two lineages are quite divergent, differing at about 30%-50% of sites in their Gag, Pol, and Env protein sequences [5].

Interestingly, population genetic studies have shown that central and eastern chimpanzees are barely differentiated, calling into question their status as separate subspecie. However, the fact that they harbor distinct SIVcpz lineages suggests that central and eastern chimpanzees have been effectively isolated for some time.

In addition, molecular epidemiological studies in southern Cameroon have shown that SIVcpz*Ptt* strains show phylogeographic clustering, with viruses from particular areas forming monophyletic lineages, and the discovery of SIVgor (SIV gorillas) has identified a second ape species as a potential reservoir for human infection [6].

Collectively, these findings have allowed the origins of HIV-1 to be unraveled [7].

HIV-1 is not just one virus, but comprises four distinct lineages, termed groups M, N, O, and P, each of which resulted from an independent cross-species transmission event. Group M was the first to be discovered and represents the pandemic form of HIV-1[8]; it has infected millions of people worldwide and has been found in virtually every country on the globe.

Group O was discovered in 1990 and is much less prevalent than group M. It represents less than 1% of global HIV-1 infections, and is largely restricted to Cameroon, Gabon, and neighboring countries.

Group N was identified in 1998, and is even less prevalent than group O; so far, only 13 cases of group N infection have been documented, all in individuals from Cameroon.

Finally, group P was discovered in 2009 in a Cameroonian woman living in France [9]. Despite extensive screening, group P has thus far only been identified in one other person, also from Cameroon [10]. Although members of all of these groups are capable of causing $CD4^+$ T-cell depletion and AIDS.

How humans acquired the precursors of HIV-1 groups M, N, O, and P is not known; however, based on the biology of these viruses, transmission must have occurred through cutaneous or mucous membrane exposure to infected ape blood and/or body fluids. Such exposures occur most commonly in the context of bushmeat hunting. Whatever the circumstances, it seems clear that human–ape encounters in west central Africa have resulted in four independent cross-species transmission events. Molecular clock analyses have dated the onset of the group M and O epidemics to the beginning of the twentieth century [10].

In contrast, groups N and P appear to have emerged more recently, although the sequence data for these rare groups are still too limited to draw definitive conclusions. Obviously they differ vastly in distribution within the human population.

1.3 HIV Replication

HIV can only replicate inside human cells and begins its cycle by entries those cell that present CD4 receptors on their surface. This receptor is usually found on immune cells such as T-cells on macrophages. The gp120 surface protein interacts with the CD4 marker resulting in attachment of the virus to the host cell. However, for entry, the virus also needs a co-receptor in addition to the CD4 receptor i.e. CCR5(R5) on macrophages and CXCR4(X4) on T cells [11].

The interaction between the CD4-gp 120 complex and co-receptor is required for changes in the cell membrane to facilitate viral entry.

The HIV-1 genome encodes nine open reading frames. Three of these encode the Gag, Pol, and Env polyproteins, which are subsequently proteolyzed into individual proteins common to all retroviruses. The four Gag proteins, MA (matrix), CA (capsid), NC (nucleocapsid), and p6, and the two Env proteins, SU (surface or gp120) and TM (transmembrane or gp41), are structural components that make up the core of the virion and outer membrane envelope. The three Pol proteins, PR (protease), RT (reverse transcriptase), and IN (integrase), provide essential enzymatic functions and are also encapsulated within the particle.

HIV-1 encodes six additional proteins, often called accessory proteins, three of which (Vif, Vpr and Nef) are found in the viral particle. Two other accessory proteins, Tat and Rev, provide essential gene regulatory functions, and the last protein, Vpu, indirectly assists in assembly of the virion [11].

The retroviral genome is encoded by an ~9-kb RNA, and two genomic-length RNA molecules are also packaged in the particle. Thus, in simplistic terms, HIV-1 may be considered as a molecular entity consisting of 15 proteins and one RNA.

The HIV-1 promoter is located in the 5′ LTR and contains a number of regulatory elements important for RNA polymerase II transcription. Sites for several cellular transcription factors are located upstream of the start site, including sites for NF-κB, Sp1, and TBP. These cellular factors help control the rate of transcription initiation from the integrated provirus and their abundance in different cell types or at different times likely determines whether a provirus is quiescent or actively replicating.

Despite the importance of these factors, transcription complexes initiated at the HIV-1 promoter are rather inefficient at elongation and require the viral protein Tat to enhance the processivity of transcribing polymerases. Under some conditions, Tat may also enhance the rate of transcription initiation.

Tat increases production of viral mRNAs ~100-fold and consequently is essential for viral replication. It is not yet clear which features of the HIV-1 promoter cause initiating transcription complexes to be poorly processive, but experiments in which the TATA box and downstream sequences have been interchanged with different promoters suggest an important role for these regions[12].

In the absence of Tat, polymerases generally do not transcribe beyond a few hundred nucleotides, though they do not appear to terminate at specific sites. It is not

yet clear how Tat causes transcribing polymerases to become sufficiently processive to completely transcribe the ~9-kb viral genome, but recent experiments suggest that Tat may assemble into transcription complexes and recruit or activate factors that phosphorylate the RNA polymerase II C-terminal domain (CTD), including the general transcription factor TFIIH and other novel kinases.

These findings support a model in which Tat enhances phosphorylation of the CTD, a process known to occur as RNA polymerase II converts from an initiating to an elongating enzyme.

Unlike typical transcriptional activators, Tat binds not to a DNA site but rather to an RNA hairpin known as TAR (trans-activating response element), located at the 5′ end of the nascent viral transcripts. An arginine-rich domain of Tat helps mediate binding to a three-nucleotide *bulge* region of TAR, with one arginine residue being primarily responsible for recognition. NMR studies of TAR complexed to arginine show a base-specific contact between the arginine side chain and a guanine in the RNA major groove. The complex is stabilized by additional contacts to the phosphate backbone and by formation of a U-A: U base triple between a *bulge* nucleotide and a base pair above the *bulge* (Figure 2)[12].

NMR studies of the full-length 86–amino acid Tat protein have suggested that a hydrophobic core region of about 10 amino acids adopts a defined structure but that the rest of the molecule, including the arginine-rich RNA-binding domain, is

relatively disordered. It seems likely that Tat requires interactions with cellular proteins in addition to TAR to adopt a stable structure.

Aside from proteins of the transcription apparatus, another protein is needed to bind to the *loop* of the TAR hairpin, apparently helping to stabilize the Tat-TAR interaction.

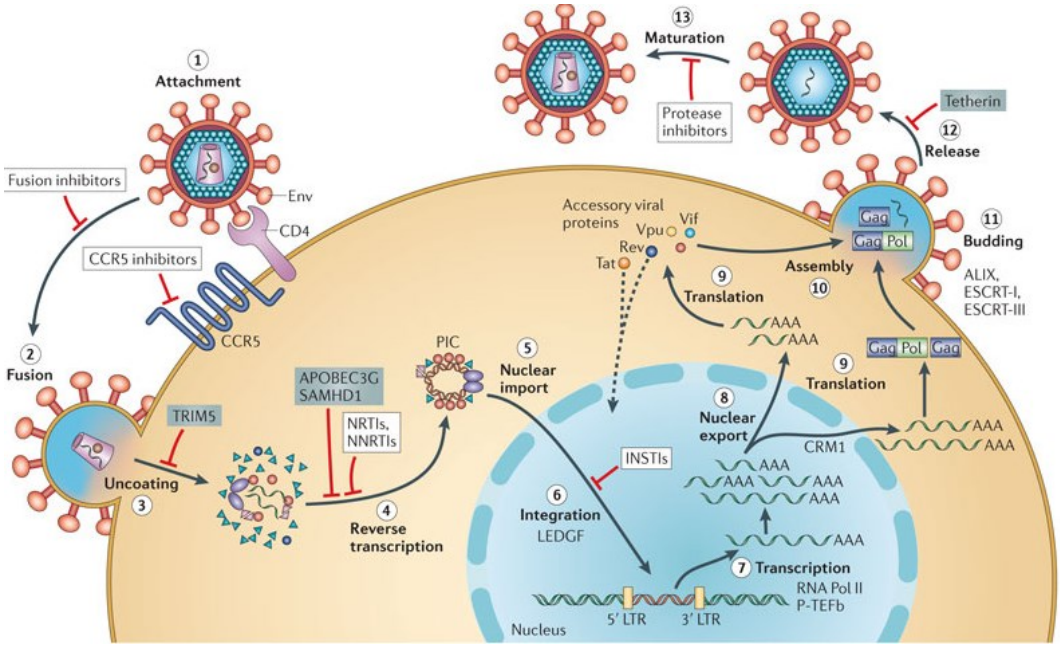


Figure 2: HIV replication cycle.

Reverse transcription is a critical step for retroviral replication. Reverse transcription is catalyzed by Reverse Transcriptase(RT) and consists of a complex series of events that culminate in the synthesis of a linear double-stranded DNA copy of the viral RNA genome. The conversion of a single-stranded RNA genome to a double-stranded DNA is achieved by employing two obligatory strand transfer steps, minus-strand transfer and plus-strand transfer. During the assembly of virus, both host tRNA and viral RNA are selectively packaged into the virions (Figure 3) [13].

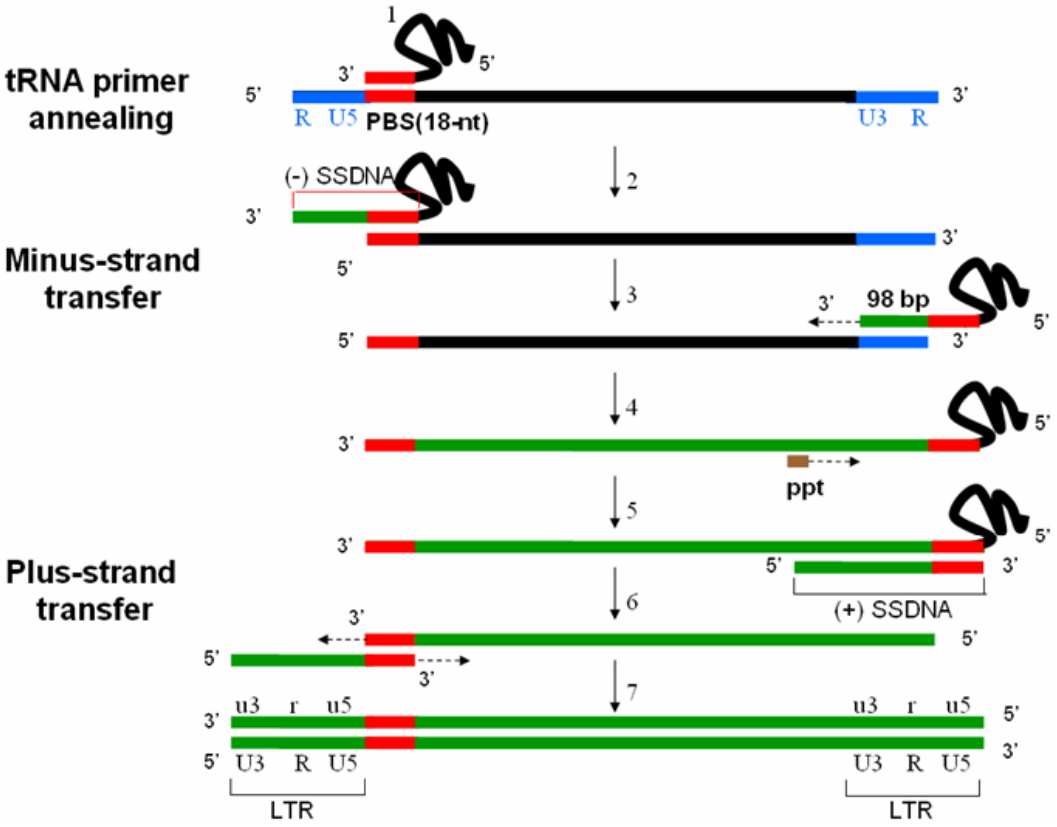


Figure 3: Mechanism of reverse transcription.

1.4 AIDS and HIV-1 Antiretroviral Drug Therapy

HIV infection is characterized by a gradual decrease in T cell levels (in particular T cells carrying the CD4 marker on their surface which include T helper cells) through three main mechanisms: direct viral killing of infected cells; increased rates of apoptosis in infected cells and killing of infected CD4+T cells by CD8 cytotoxic lymphocytes that recognize infected cells [14].

When $CD4^{+}$T cell number decline below a critical level, cell mediated immunity is lost. AIDS is the final stage of HIV infection and is used to describe a highly decimated $CD4^{+}$T cell count and the appearance of various opportunistic infections as a result of the progressive failure of the immune system.

Any secondary condition, symptom, or other disorders caused by an AIDS weakened immune system is a complication of AIDS. The opportunistic infections associated with AIDS are not restricted to the immune system. Some of the more common are caused by: Pneumocystis pneumonia, Cytomegalovirus, Candidiasis, Mycobacterium tuberculosis and Herpes simplex. AIDS causes also common diseases affecting the gastrointestinal tract, neurological and psychiatric effects, increasing incidence of tumors and cancers (Kaposi's sarcoma, Burkitt's lymphoma, and cervical cancer in HIV infected women due to the human papillomavirus) [15].

Currently there are different classes of anti-HIV drugs that however have several limitations especially the resistance of the virus.

HIV-1 reverse transcriptase (RT) contributes to the development of resistance to all anti-AIDS drugs by introducing mutations into the viral genome. Until today, the most significant advance in the medical management of HIV-1 infection has been the treatment of patients with antiviral drugs, which can suppress HIV-1 replication to undetectable levels.

The discovery of HIV-1 as the causative agent of AIDS together with an ever-increasing understanding of the virus replication cycle have been instrumental in this effort by providing researchers with the knowledge and tools required to prosecute drug discovery efforts focused on targeted inhibition with specific pharmacological agents [15]. To date, an arsenal of 24 Food and Drug Administration (FDA)-approved drugs are available for treatment of HIV-1 infections. These drugs are distributed into six distinct classes based on their molecular mechanism and resistance profiles: (1) nucleoside-analog reverse transcriptase inhibitors (NNRTIs), (2) non–nucleoside reverse transcriptase inhibitors (NNRTIs), (3) integrase inhibitors, (4) protease inhibitors (PIs), (5) fusion inhibitors, and (6) coreceptor antagonists [16].

HIV-1 entry exploits several host proteins for a set of intricate events leading to membrane fusion and virus core release into the cytoplasm.

The treatment of HIV-1 infection was revolutionized in the mid-1990s by the development of inhibitors of the reverse transcriptase and protease [17] two of three

essential enzymes of HIV-1, and the introduction of drug regimens that combined these agents to enhance the overall efficacy and durability of therapy.

Since the first HIV-1 specific antiviral drugs were given as monotherapy in the early 1990s, the standard of HIV-1 care evolved to include the administration of a cocktail or combination of antiretroviral agents (ARVs). The advent of combination therapy, also known as HAART, for the treatment of HIV-1 infection was seminal in reducing the morbidity and mortality associated with HIV-1 infection and AIDS [18].

Combined antiretroviral therapy dramatically suppresses viral replication and reduces the plasma HIV-1 viral load (vLoad) to below the limits of detection of the most sensitive clinical assays (<50 RNA copies/mL) resulting in a significant reconstitution of the immune system [19] as measured by an increase in circulating $CD4^+$ T-lymphocytes. Importantly, combination therapy using three antiretroviral agents directed against at least two distinct molecular targets is the underlying basis for forestalling the evolution drug resistance [20].

1) NRTIs (Nucleoside/Nucleotide reverse transcriptase inhibitors**).**

NRTIs were the first class of drugs to be approved by the FDA. NRTIs are administered as prodrugs, which require host cell entry and phosphorylation by cellular kinases before enacting an antiviral effect. Lack of a 3′-hydroxyl group at the sugar (2′-deoxyribosyl) moiety of the NRTIs prevents the formation of a 3′-5′-phosphodiester bond between the NRTIs and incoming 5′-nucleoside triphosphates, resulting in termination of the growing viral DNA chain. Chain termination can occur

during RNA-dependent DNA or DNA-dependent DNA syntheses, inhibiting production of either the (−) or (+) strands of the HIV-1 proviral DNA.

Currently, there are eight FDA-approved NRTIs: abacavir (ABC), didanosine (ddI), emtricitabine (FTC), lamivudine (3TC), stavudine (d4T), zalcitabine (ddC), zidovudine (AZT) and tenofovir disopropyl fumarate (TDF) (Figure 4) [21].

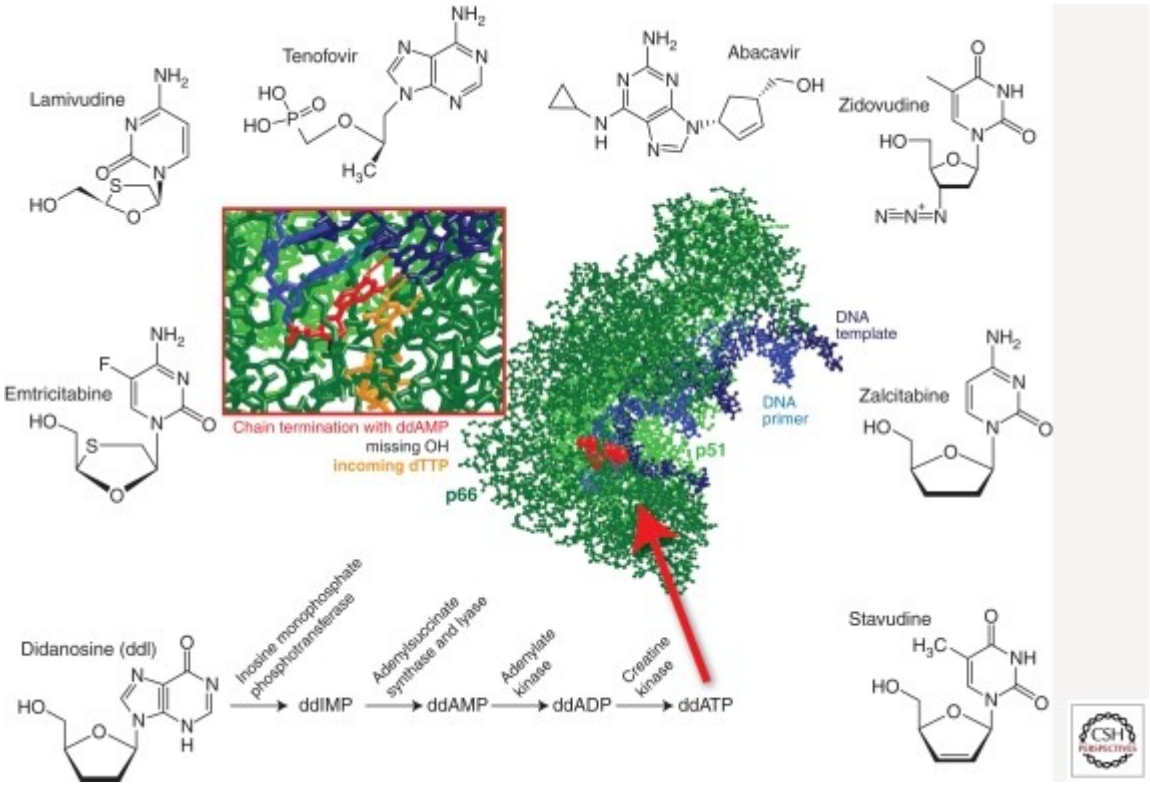


Figure 4: Nucleos(t)ide reverse transcriptase inhibitors and X-ray crystal structure of HIV-1 RT in complex with DNA primer/template chain terminated with ddAMP and with an incoming dTTP.

Resistance to NRTIs is mediated by two mechanisms:

1. ATP-dependent pyrophosphorolysis, which is the removal of NRTIs from the 3′ end of the nascent chain, and reversal of chain termination;
2. increased discrimination between the native deoxyribonucleotide substrate and the inhibitor.

2)Non-Nucleoside Reverse Transcriptase Inhibitors

NNRTIs inhibit HIV-1 RT by binding and inducing the formation of a hydrophobic pocket proximal to, but not overlapping, the active site.

The binding of NNRTIs changes the spatial conformation of the substrate-binding site and reduces polymerase activity. The NNRTI-binding pocket only exists in the presence of NNRTIs and consists of hydrophobic residues (Y181, Y188, F227, W229, and Y232), and hydrophilic residues such as K101, K103, S105, D192, and E224 of the p66 subunit and E138 of the p51 subunit.

Unlike NRTIs, these non/uncompetitive inhibitors do not inhibit the RT of other *Lentiviruses* such as HIV-2 and Simian Immunodeficiency Virus (SIV).

Currently, there are four approved NNRTIs: etravirine, delavirdine, efavirenz, and nevirapine, and several in development, including rilpivirine in Phase III (Figure 5) [22].

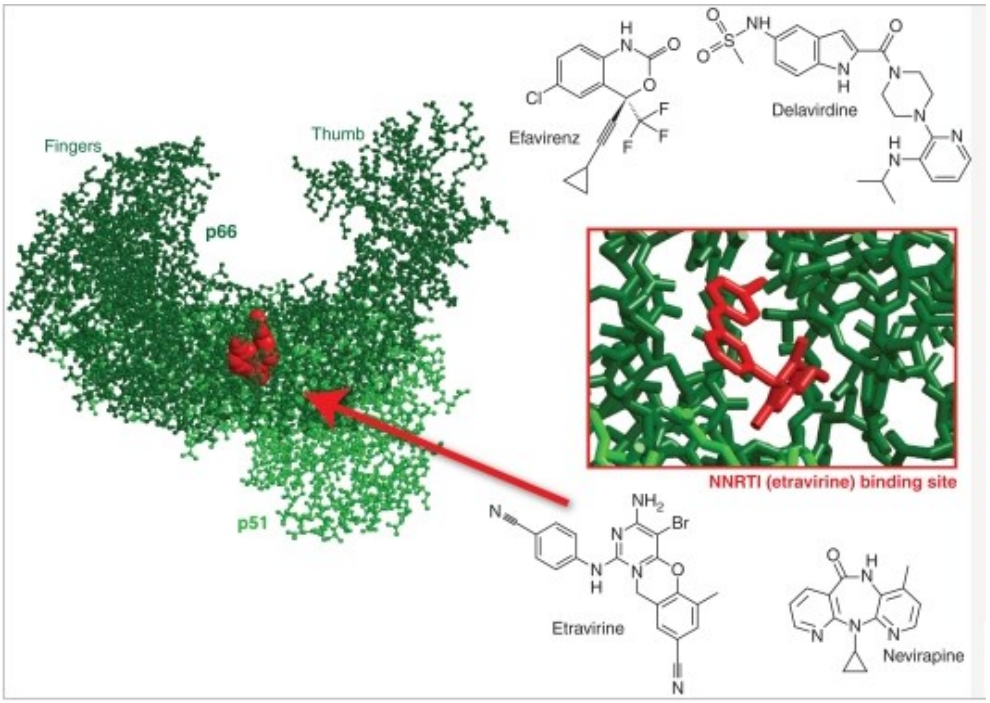


Figure 5: Non–nucleoside RT inhibitors and the X-ray crystal structure of HIV-1 RT complexed with etravirine.

NNRTI resistance generally results from amino acid substitutions such as L100, K101, K103, E138, V179, Y181, and Y188 in the NNRTI-binding pocket of RT.

3) Integrase Inhibitors (INIs or InSTIs)

Integrase was the most recent HIV-1 enzyme to be successfully targeted for drug development . Raltegravir (RAL, MK-0518) was FDA approved in 2007, and other integrase inhibitors, including Elvitegravir (EVG, GS-9137) are progressing through clinical development [23].

Integrase catalyzes 3′ end processing and viral DNA and strand transfer. All integrase inhibitors in development target the strand transfer reaction and are thus referred to as either INIs or more specifically, integrase strand transfer inhibitors (InSTIs).

The selective effect on strand transfer is a result of a now well-defined mechanism of action in which the inhibitor binds only to the specific complex between integrase and the viral DNA and interacts with the two essential magnesium metal ion cofactors in the integrase active site and also the DNA.

Therefore, all InSTIs are comprised of two essential components: a metal-binding pharmacophore, which sequesters the active site magnesiums, and a hydrophobic group, which interacts with the viral DNA as well as the enzyme in the complex.

InSTIs are therefore the only class that interacts with two essential elements of the virus, the integrase enzyme as well as the viral DNA, which is the substrate for integration.

The recent co-crystallization of the foamy virus integrase DNA complex or intasome with both RAL and EVG corroborates the biochemical mechanism and provides a structural basis for understanding the unique breadth of antiviral activity that has been observed for InSTIs across all HIV-1 subtypes as well as other retroviruses, such as HIV-2 and XMRV (Figure 6) [24].

In the cocrystal structure, the general architecture and aminoacids within the active site of the foamy virus intasome are highly conserved with other retroviral integrases, as are the immediate surrounding interactions with InSTIs. The common mechanism of action and conserved binding mode for InSTIs also has important implications for understanding resistance to the class.

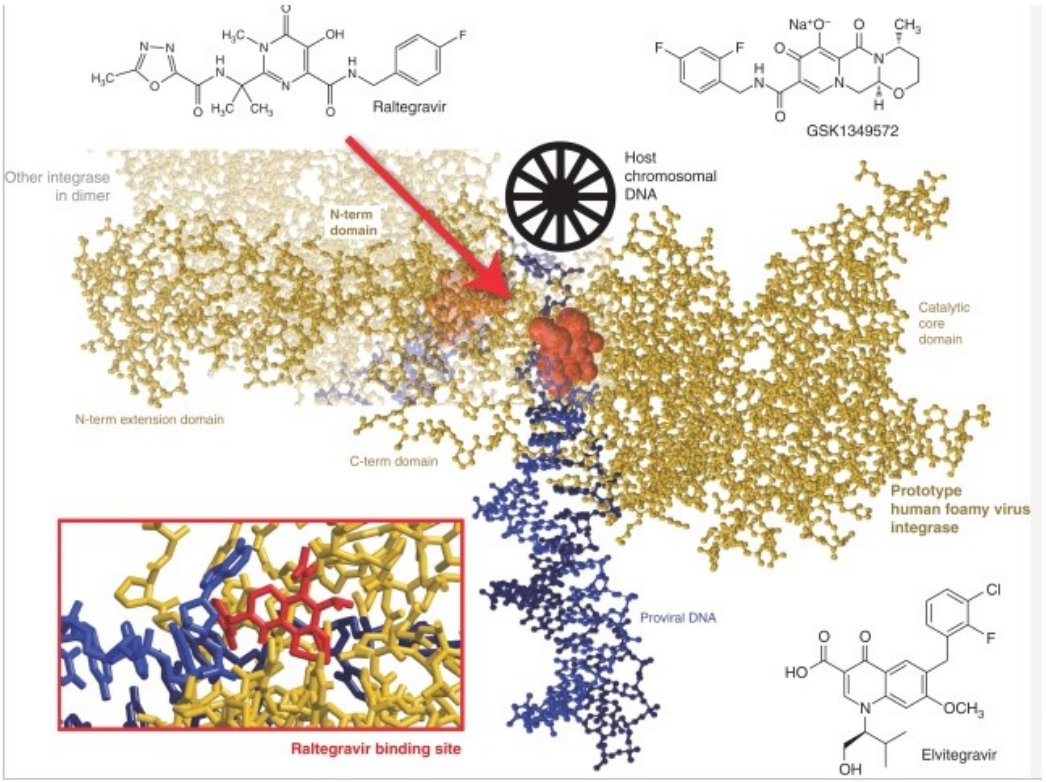


Figure 6: Integrase strand transfer inhibitors and the crystal structure of prototype human foamy virus integrase (as a model of HIV-1 IN) complexed to dsDNA and Raltegravir.

4) Protease Inhibitors (PI)

The HIV-1 protease is the enzyme responsible for the cleavage of the viral gag and gag-pol polyprotein precursors during virion maturation.

Ten PIs are currently approved (Figure 7) [25]: amprenavir (APV), atazanavir (ATZ), darunavir (TMC114), fosamprenavir, indinavir (IDV), lopinavir (LPV), nelfinavir (NFV), ritonavir (RTV), saquinavir (SQV), and tipranavir.

Because of its vital role in the life cycle of HIV-1 and relatively small size (11 kDa), it was initially expected that resistance to protease inhibitors would be rare.

However, the protease gene has great plasticity, with polymorphisms observed in 49 of the 99 codons, and more than 20 substitutions known to be associated with resistance. The emergence of protease inhibitor resistance likely requires the stepwise accumulation of primary and compensatory mutations and each PI usually selects for certain signature primary mutations and a characteristic pattern of compensatory mutations [25].

Unlike NNRTIs, primary drug-resistant substitutions are rarely observed in the viral populations in protease inhibitor-naive individuals.

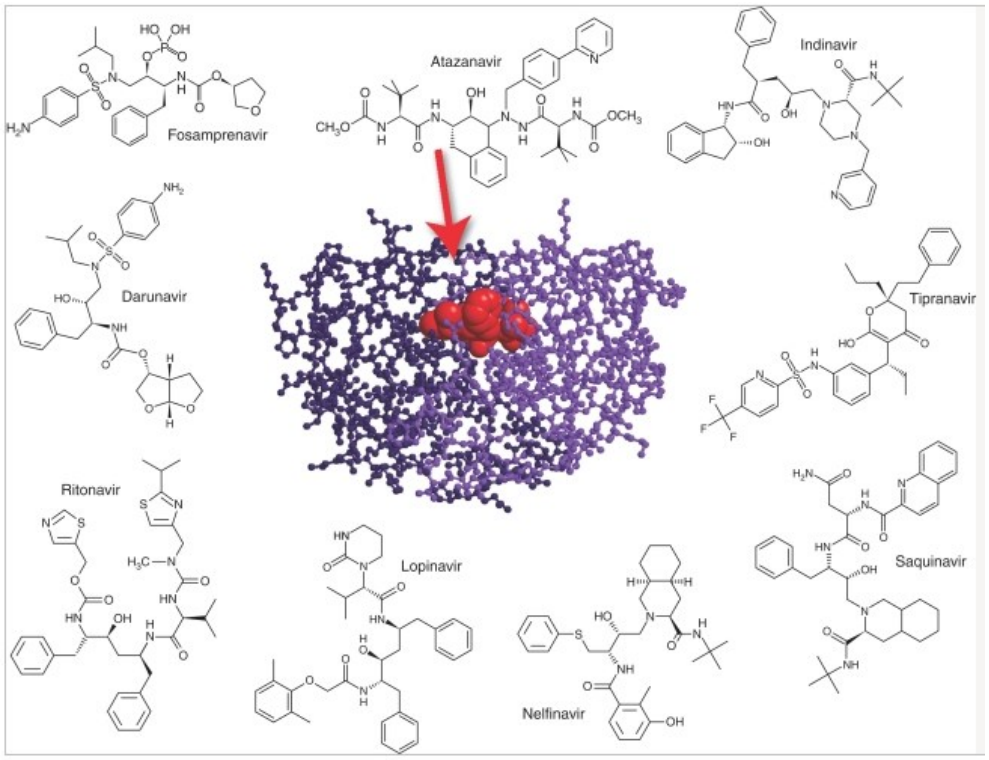


Figure 7: Protease inhibitors and the crystal structure of HIV-1 protease complexed with atazanavir.

5) Fusion Inhibitors

The crystal structure of the gp41 ectodomain and of the ectodomain partnered with an inhibitory peptide (C34) revealed that the fusion-active conformation of gp41 was a six-helix bundle in which three N helices form an interior, trimeric *coiled coil* onto which three antiparallel C helices pack.

Peptide fusion inhibitors were designed based on the discovery that two homologous domains in the viral gp41 protein must interact with each other to promote fusion, and that mimicry of one of these domains by a heterologous protein can bind and disrupt the intramolecular interactions of the virus protein.

Alpha-helical peptides homologous to the leucine *zipper* domain of gp41 had significant antiviral activity against HIV-1, and this activity depended on their ordered solution structure. Rational design of helical inhibitors ultimately produced a molecule (T-20, enfuvirtide) with potent antiviral activity *in vivo* (Figure 8) [26].

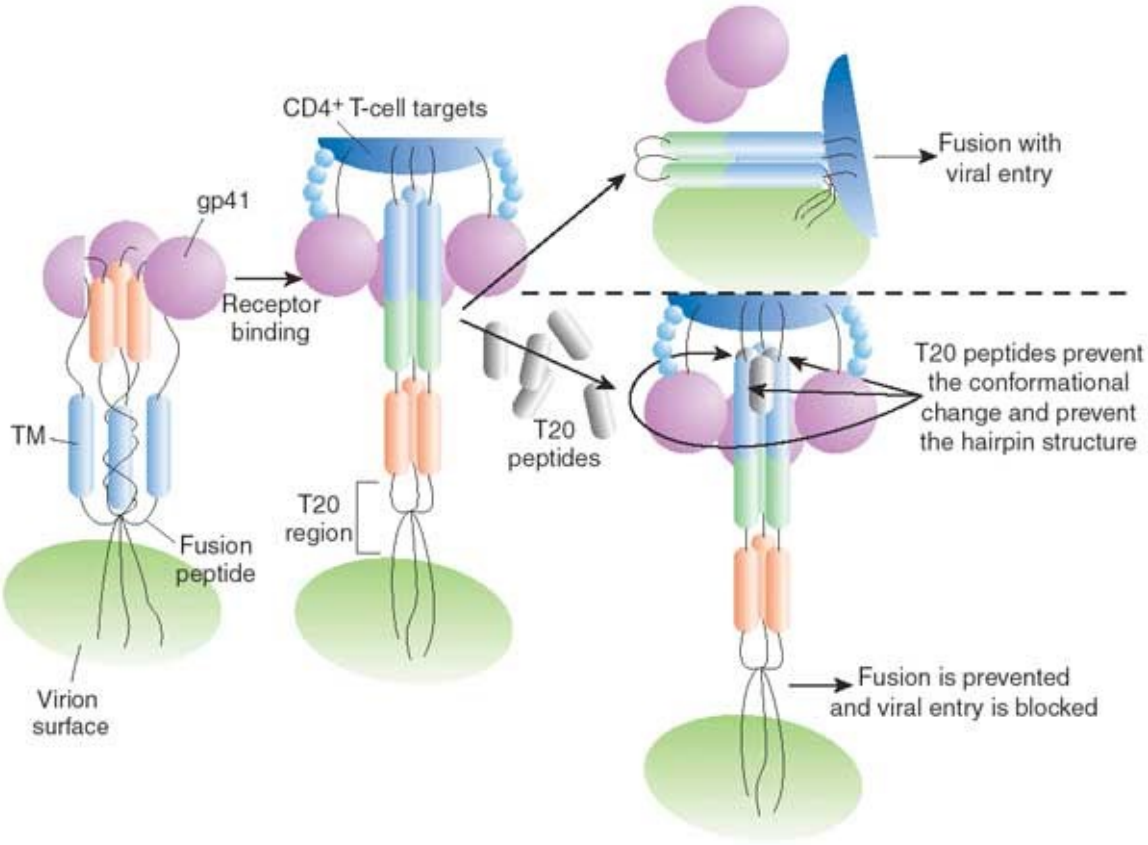


Figure 8: Fusion inhibitors - mechanism of action. Enfuvirtide binds to the transmembrane domain and prevents fusion to the host cell and viral entry.

Resistance to early alpha-helical inhibitors was shown to be mediated by mutations in the amino-terminal region of gp41, which provide further evidence for binding of these peptides to the virus.

6) Small-Molecule CCR5 Antagonists

Small-molecule CCR5 antagonists bind to hydrophobic pockets within the transmembrane helices of CCR5. This site does not overlap the binding sites of either CCR5 agonists or HIV-1 envelope. Instead, drug binding induces and stabilizes a receptor conformation that is not recognized by either. Thus, these molecules are considered allosteric inhibitors. Ideally, a small-molecule inhibitor of CCR5 would block binding by HIV-1 envelope, but continue to bind native chemokines and allow signal transduction [27].

Most small-molecule inhibitors, however, are pure antagonists of the receptor.

Oral administration of small-molecule antagonists has been shown to inhibit viral replication in macaque models and to prevent vaginal transmission. Thus far, three antagonists (VCV, MVC, and Aplaviroc) have been shown to inhibit virus replication in humans [28].

The compound MVC was approved for therapeutic use by the FDA in 2007. Small-molecule CCR5 inhibitors have been used to select for drug resistance in peripheral blood mononuclear cell cultures (PBMC), which express CCR5 and CXCR4, as well as a variety of other chemokine receptors that could potentially substitute for HIV-1 coreceptors.

1.5 New target for inhibitors of HIV-1 replication: NC (Nucleocapsid Protein)

Despite the success of protease and reverse transcriptase inhibitors, and due to the toxicity of the anti-HIV drugs and their adverse reactions (responsible for failures of many treatments), new drugs to suppress HIV-1 replication are still needed [29].

Several other early events in the viral life cycle (stages before the viral genome is inserted into host cell DNA) are susceptible to drugs, including virus attachment to target cells, membrane fusion and post-entry events such as integration, accessory-gene function and assembly of viral particles.

New molecular targets that can be exploited for the development of new classes of HIV drugs are hence actively sought. Moreover regarding HIV-1, viral proteins involved in the regulation of numerous viruses recognize nucleic acids.

One of these unexplored targets can be precisely represented by the HIV-1 nucleocapsid NCp(NC), which is required at several steps of HIV replication: during retro-transcription it favors the strand transfer steps through its chaperone activities, allowing the helix destabilization and the annealing of viral nucleic acids.

In the absence of NC, the synthesis of DNA catalyzed by Reverse Transcriptase (RT) is in fact incomplete and the enzyme stops at stable structured sites in the template strand [30].

The sequence TAR(Transactivation Responsive Element) is a stable *bulge–loop* RNA in the R region of the HIV genome required for efficient viral

transcription. TAR RNA binds HIV-1 Tat (transactivator of transcription), and TAR–Tat complex inhibition has been thoroughly explored.

Interestingly, TAR is not only necessary for Tat-mediated functions, but it also plays a critical role in the minus strand transfer during reverse transcription: the TAR RNA located at the 50-end of the viral genome must be paired with its complementary DNA sequence, cTAR, synthesized by RT in the minus-strong-stop DNA. The annealing of TAR/cTAR, not favored due to the very stable secondary structures of TAR and cTAR, is catalysed by NC: melting of secondary structures of viral RNA and DNA makes them available for re-pairing in alternative combinations during the strand transfer to form more stable annealed hybrids (Figure 9) [31].

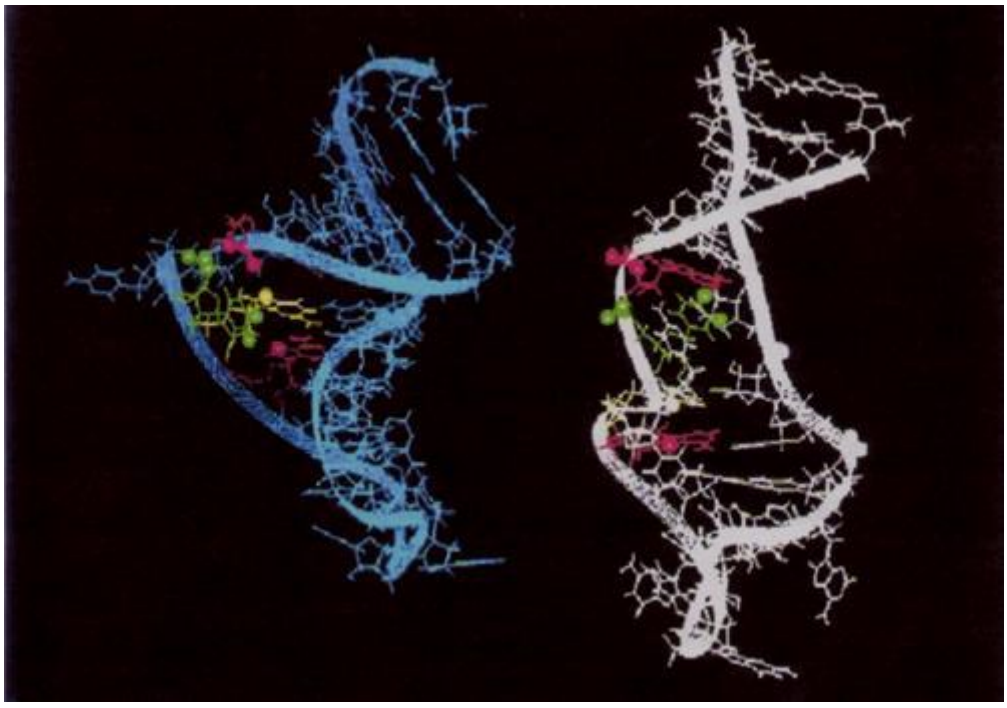


Figure 9: Representation of the major groove of free TAR (on the left) and of the Tat-bound TAR (on the right). A22 and A27 are represented in red; U23 in green and G26 in yellow .

So the HIV-1 nucleocapsid protein (NC) represents a key protein: structurally is a small 55-amino acids basic protein with two CysCysHisCys zinc-binding domains, that strongly bind Zn^{2+} and specifically recognizes the Ψ-site of the viral RNA.

It plays a number of crucial roles in the viral lifecycle, including reverse transcription and RNA encapsidation.

NC destabilizes nucleic acid structures and promotes the formation of annealed substrates for HIV-1 reverse transcription elongation. Short helical nucleic acid segments bordered by *bulges and loops*, such as the Trans-Activation Response element (TAR) of HIV-1 and its complementary sequence (cTAR), are nucleation elements for helix destabilization by NC and also preferred recognition sites for threading intercalators.

Its activity is mainly related to its 11amminoacids located in the N-terminus domain, instead destabilizing power is due to its zinc fingers domain (Figure 10) [32].

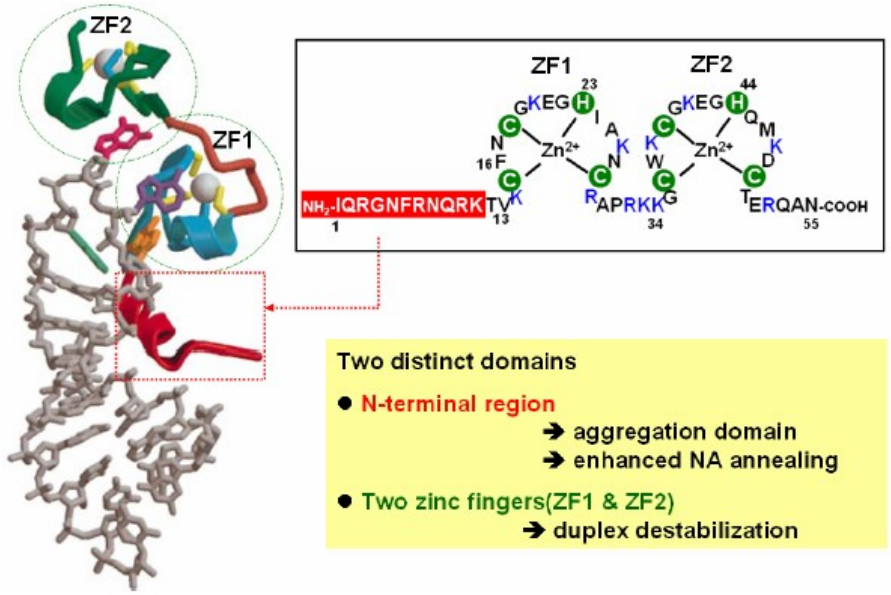


Figure 10: Three-dimensional NMR structure of a NC/RNA hairpin complex.

The N-terminal 12 residues of NC (shown in red in the Figure 9) bind to the major groove to the RNA stem (shown in gray) through the electrostatic interactions.

The function of NC relies on two interactions, electrostatic and π- π stacking interactions, with nucleic acids, NC protein appears to have two distinct domains implied in binding with nucleic acids.

The N-terminal basic helix of NC interacts nonspecifically with the major groove of the RNA stem by electrostatic interactions [33] and is responsible for the aggregation of NC/nucleic acid complexes.

The zinc finger domain interacts with specific base sequences in particular with preference for TG or UG rich regions [34]. The binding of zinc fingers to nucleic acids is responsible for destabilizing their secondary structure.

NMR further showed that phenylalanine in the first zinc finger (ZF1) and tryptophan in the second zinc finger (ZF2) engaged in aromatic π- π stacking interaction with the G residues in the single-stranded GGAG tetraloop as shown in Figure 9.

NC was observed to have highly specific interactions with bases in the nucleic acids loop since binds specifically to exposed (unpaired) nucleobases of RNA or DNA.

NC is crucial for many processes of viral life cycle in which interactions between nucleic acids and NC are the key. For example, NC is involved in the recognition and packaging for RNA genomes, in virus assembly, in selectively packaging tRNA primer and in many annealing steps during the reverse transcription [35].

As a nucleic acid chaperone protein, NC catalyzes nucleic acids conformational rearrangement that leads to the most thermodinamically stable structure. The chaperone activity of NC is believed to be derived from two main consequences of the nucleic-acid/NC interactions in this system.

Fist, NC lowers the energy cost of bringing two complementary strands of nucleic acids together by screening the negative charges of two oligonucleotides and perhaps through the binding of the N-terminal basic helix to the double-stranded stem of the hairpins.

The NC chaperone activity allows the protein to be a cofactor of RT and promote several annealing reactions during the HIV-1 reverse transcription process: NC makes it possible for RT to copy the *stem-loop* sequences in RNA and DNA templates by melting the structured sequences.

In additions, NC has been shown to facilitate the annealing of cellular tRNA onto the PBS of the viral RNA, the annealing of complementary PBS sequences in plus-strand transfer and the annealing of the transactivation response (TAR) region of the viral RNA to the complementary sequences (cTar) during minus-strand transfer [35].

NC inhibition represents therefore a new molecular approach in antiretroviral therapy: drugs targeting NC are currently not available, although different and unrelated classes of molecules such as peptidomimetics [36], RNA aptamers [37], zinc ejector agents [38], nucleomimetics and nucleic acid [39] binding drugs have been proposed.

Regard peptidomimetics: an alternative strategy to inhibit NC functions is to design peptides that can directly compete with NC for the binding to its RNA and DNA substrates i.e. cyclic peptides mimicking the spatial orientation of Phe16 and Trp37 residues in the hydrophobic plateau and containing two basic residues to mimic Arg26 and Arg32.

Anyway, the activity of the leading compound was not so satisfactory. Attempts are actually performed to better mimic the NC hydrophobic plateau, in order to

increase the activity of these peptidomimetics *i.e.* with small peptides able to specifically recognize NC targets that are composed of a cluster of Trp residues surrounded by basic residues.

These peptides were able to compete with NC target sequences binding mainly through stacking interactions between the Trp residues and the oligonucleotide bases and to stabilize the cTAR secondary structure inhibiting the NC-directed melting of the cTAR sequence.

The second class is represented by RNA aptamers which were selected using the SELEX method, obtaining aptamers of about 40 nucleotides in length [37]. Their secondary structures correspond to *stem-loops* in which the stems are rich in GC base pairs and the loops contain G and U residues.

Zinc ejector agents are oxidizing agents that are able to attack the sulfhydryl group of the Cys residue and eject the zinc, rendering the zinc fingers and by extension, the whole protein, inactive. Many classes of Zinc ejector agents and zinc ligands have been identified: NOBA (3-nitrosobenzamide), DIBA (disulfide benzamide).However, these compounds elicit serious side effects since they target cellular proteins binding zinc and this impairs their use in therapy [38].

Non zinc ejecting NC binders were also identified and these represent antagonists of the strong binding of NC to oligonucleotides. Many of them contain a xanthenyl ring structure while other had a related fluorescein or gallein-like structure (containing a xanthene ring substituted by two hydroxyls in positions 4' and 5', lead compound tetrachlorogallein). Hydroxyls groups likely play a critical role since they were found to be critical for binding to NC and are the only feature that the compounds share. Probably hydroxyl groups bind to the amide nitrogen of Gly35 with additional contacts at the carbonyl oxygen of Gly40 and Lys33 of NC and it is thought that these compounds might interfere with NC's ability to stack its aromatic residues with the bases of its nucleic acids target. Compounds able to inhibit NC activity by binding do not show any structure-activity correlation [40].

The last class is represented by intercalators: actinomycin D (widely used anti-cancer drug that inhibits the DNA-dependent DNA/RNA synthesis through DNA intercalation and binding to the minor groove) was the most active, but the main problem is that it interacts with a large range of nucleic acids, giving a poor specificity. As a consequence, a clinical use of such a molecule will likely generate critical side effects and the emergence of resistant strains[41].

Is now clear that the nucleocapsid protein of HIV-1 (NC) , is critical for proper viral packaging and replication, owing to its highly conserved nature, for these reasons it represents an ideal drug target [42].

Chapter 2

Aim of the research

As described above, NC mediates the annealing of TAR with cTAR by melting their secondary structures and making extensive complementary sequences available to anneal into a more stable RNA/DNA hybrid , as represented in Figure 11.

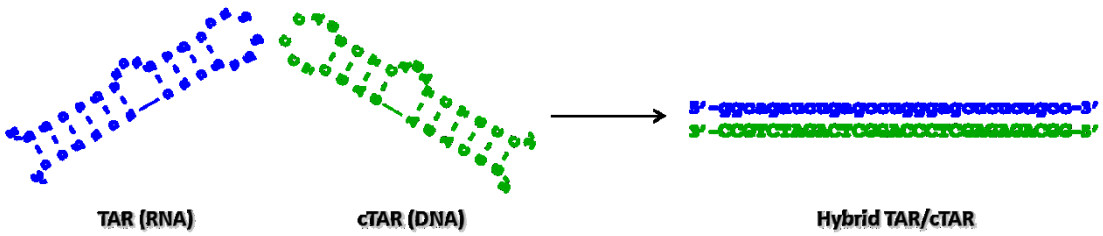


Figure 11: Sequence and secondary structure of constructs replicating TAR RNA and cTAR DNA, which were employed in the assays, performed for our compounds. Heat refolding or NC-mediated annealing can lead to the formation of the Hybrid TAR/cTAR.

Based on this and analogous functions in the viral lifecycle, it has become more clear that NC could represent an excellent molecular target for antiretroviral therapy, pursued by different approaches in recent years [42].

As starting point for the present PhD project, we decided to explore more in details the possibility of design and synthesize novel anthraquinonic derivatives able to bind the TAR RNA.

In particular the research group to which I belong has recently published that threading intercalators characterized by a polyaromatic nucleus di-substituted at opposite positions with charged amino acid side chains are preferential binders of dynamic RNA regions like the *bulge and loop* structure of TAR [43].

Similarly, molecules comprising an aromatic nucleus conjugated through a linker with cationic chains ("tripartite motif") were demonstrated to be able to inhibit TAR recognition by the Tat protein. This is particularly interesting given the parallels between Tat and NC as RNA chaperones and TAR binders by NC.

For this reason, searching for new NC inhibitors, the research group above mentioned identified a class of compounds consisting of intercalating anthraquinones that are known for their nucleic acid-binding properties. They demonstrated that these threading intercalators were able to inhibit NC *in vitro* by stabilizing dynamic nucleic acid structures and counteracting NC-induced destabilization. In particular, they shown that 2,6-derivatives bearing a β-Ala linker followed by different charged amino acids were able to bind preferentially to typical NC substrates, which resulted in their stabilization [44].

With the aim of exploring the Structure-Activity-Relationship (SAR) of these putative NC inhibitors, we have synthesized novel series of 2,6-dipeptidyl anthraquinones, which display different linkers between the anthraquinonic scaffold and the cationic side chains.

Named 2,6-AQ-Gly-X, 2,6-AQ-L-Ala-X, and 2,6-AQ-D-Ala-X, these three new series comprise respectively the compounds **4a-4l, 5a-5l** and **6a-6l** that are shown in Figure 12.

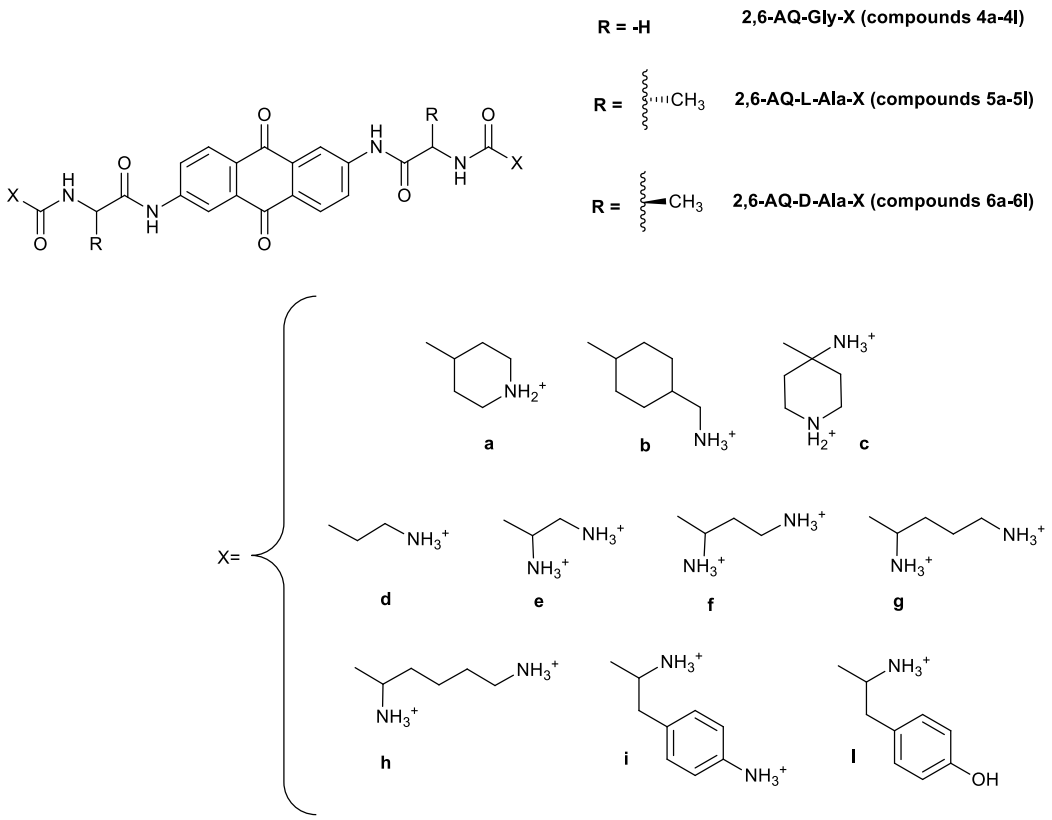


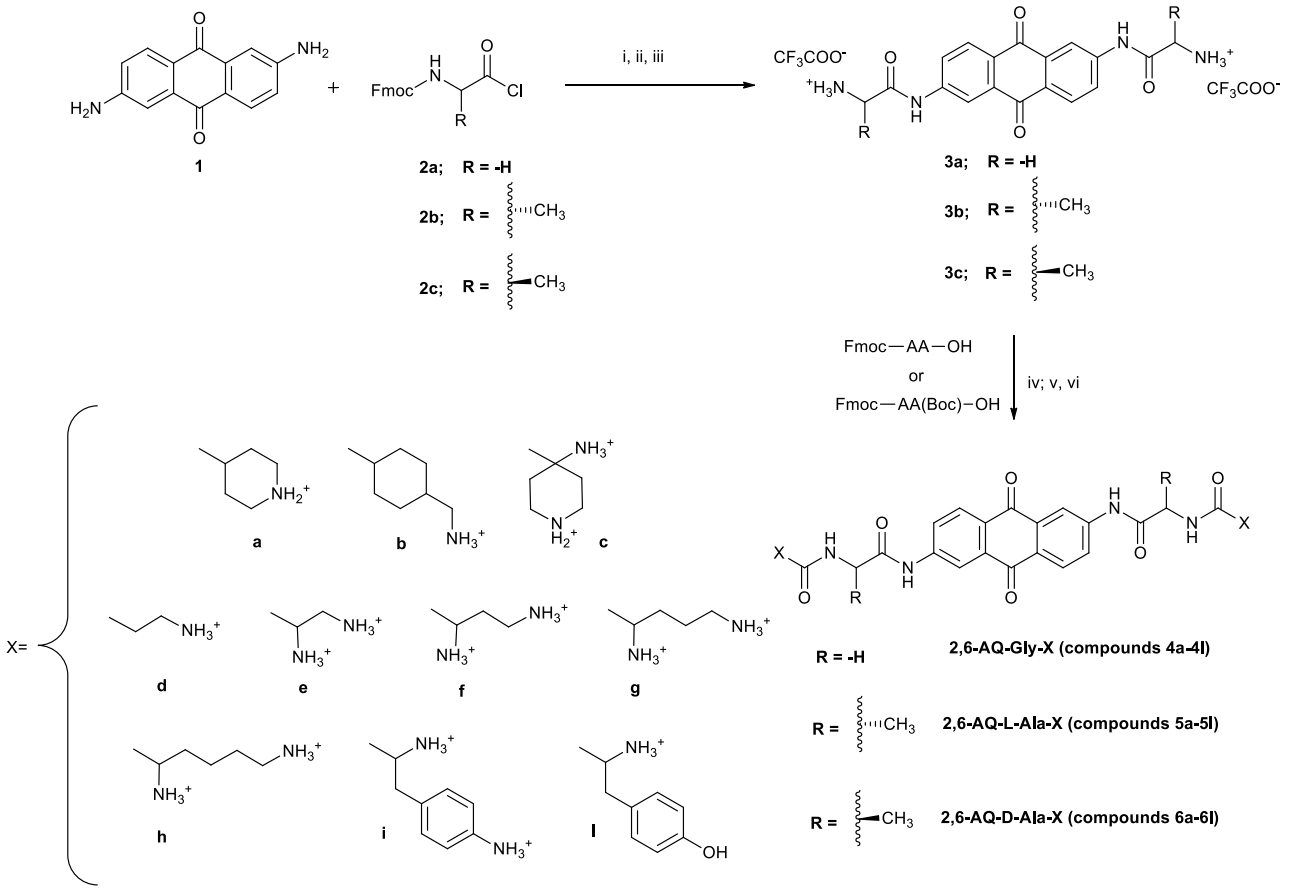
Figure 12: New series of compounds, named 2,6-AQ-Gly-X, 2,6-AQ-L-Ala-X and 2,6-AQ-D-Ala-X, respectively **4a-4l**, **5a-5l** and **6a-6l.**

Chapter 3

Synthetic Strategy

3.1 Synthesis of 2,6-disubstituted-anthraquinone derivatives

Compounds **4a-4l, 5a-5l** and **6a-6l** were obtained by standard solution synthesis according to the procedure reported previously for 2,6-disubstituted anthraquinones [43], which was appropriately modified to obtain the synthetic strategy summarized in Scheme 1.



Scheme 1. General synthetic pathway for the preparation of the anthraquinone derivatives included in the study. Reagents and conditions: (i) Pyridine, DMF; (ii) 33% Diethylamine, THF; (iii) TFA/H_2O; (iv) HBTU, DMAP, DMF; (v) 33% Diethylamine, THF; (vi) TFA/H_2O.

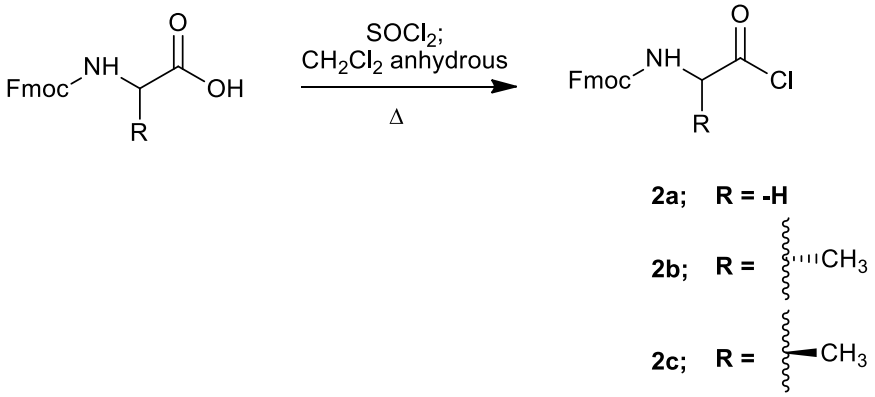
The general strategy is based on the condensation of 2,6-diaminoanthraquinone nucleus **1** with (9H-fluoren-9-yl)methyl-2-chloro-2-oxoethylcarbamate (Fmoc-Gly-Cl) **2a**, with (9H-fluoren-9-yl)methyl-1-chloro-1-oxopropan-2-yl carbamate (Fmoc-L-Ala-Cl) **2b** or with **2c** [(9H-fluoren-9-yl)metil-1-chloro-1-oxopropan-2-yl carbamate (Fmoc-D-Ala-Cl)] in presence of pyridine in DMF, stirring at room temperature for two hours.

Reaction mixture was dried and then treated with 33% diethylamine in THF in order to remove Fmoc- protecting group and successively reacted with trifluoroacetic acid in H_2O (9:1; v/v) giving intermediates **3a**, **3b** and **3c**, respectively.

Intermediates **3a**, **3b** and **3c** were successively coupled with the selected protected amino acid in presence of HBTU and DMAP in DMF. Fmoc- and Boc-protecting groups were removed with 33% diethylamine in THF and trifluoroacetic acid in H_2O (9:1; v/v), respectively, giving the desired compounds **4a-4l, 5a-5l** and **6a-6l**.

3.2 Synthesis of intermediates 2a, 2b and 2c

Fmoc-Gly-Cl (**2a**), Fmoc-L-Ala-Cl (**2b**) and Fmoc-D-Ala-Cl (**2c**) were prepared stirring at reflux for 30 minutes the corresponding commercially available Fmoc-amino acids treated with an excess of thionyl chloride in anhydrous CH_2Cl_2 (Scheme 2).



Scheme 2: Preparation of Fmoc-Gly-Cl (**2a**), Fmoc-L-Ala-Cl (**2b**) and Fmoc-D- Ala-Cl (**2c**).

Final compounds were purified by preparative RP-HPLC and were fully characterized by mass spectrometry, ^{1}H-NMR and ^{13}C-NMR. NMR and MS data for all final compounds were consistent with proposed structures.

Chapter 4

Experimental Section

4.1 Materials and Methods.

Protected α-amino acids were purchased from Bachem (Bubendorf, Switzerland) and Merck-Novabiochem (Darmstadt, Germany). The following protecting groups were employed for the selected amino acids: Fmoc for α-aminogroups and for the side-chains of 4-(amino)-piperidine-4-carboxylic acid, piperidine-4-carboxylic acid, and 4-(aminomethyl)-cyclohexanecarboxylic acid; Boc for the side chains of Lys, Orn, diaminobutyric, diaminopropionic, and 4-amino-phenylalanine residues.

Coupling and cleavage reagents, as well as solvents and other chemicals, were purchased from Sigma–Aldrich.

RP-HPLC preparative purification was routinely performed on a Shimadzu prominence LC-20 AP system equipped with a multiwavelength prominence SPD-20A UV-VIS detector. The stationary phase consisted of a Supelco Ascentis® C18 column (10 μm, 25cm x 21.2 mm). The mobile phase consisted of solvent A: 100% acetonitrile in 0.1% TFA, and solvent B: 100% H_2O in 0.1% TFA. The operational conditions involved a linear gradient of 5-70% acetonitrile with 0.1% TFA, which developed in 30 minutes at a flow rate of 30 mL/min. Product purity was assessed by analytical RP-HPLC that used a Macherey-Nagel Nucleosil 100-5 C18 column (5 μm, 4 x 125 mm). The column was connected to a Rheodyne model 7725 injector, a Shimadzu-10 ADsp HPLC system, a Shimadzu SPD-20 A/SPD-20 AV UV-VIS detector set to 254 nm.

The analytical determinations employed solvent A: 100% acetonitrile in 0.1% TFA, and B: 100% H_2O in 0.1% TFA. The operational conditions involved a linear gradient of 5-70% acetonitrile with 0.1% TFA, which developed in 30 minutes at a flow rate of 1 mL/min. Purity of all final compounds was ≥95%.

Mass spectrometric analyses of the final products were performed on API 2000 Applied Biosystem triple-quadrupole mass spectrometer.

NMR spectra were recorded on a Varian Mercury Plus 400 instrument. Chemical shifts are given as δ and are referred to Me_4Si as internal standard. The following abbreviations are used to describe peak patterns, when appropriate: s (singlet), bs (broad singlet), d (doublet), dd (double doublet), t (triplet), m (multiplet).

All oligonucleotides were synthesized by Metabion International AG (Martinsried, Germany) and stored at -20°C in 10 mM Tris-HCl pH 8.0. Dilutions were made in DEPC-treated water (Ambion). The TAR construct consisted of 29-mer RNA oligonucleotide with sequence:

5’-GGCAGAUCUGAGCCUGGGAGCUCUCUGCC-3’.

The cTAR construct consisted of its DNA complementary sequence:

5’-GGCAGAGAGCTCCCAGGCTCAGATCTGCC-3’.

When specified, TAR and cTAR were labeled at the 5’ and 3’ ends respectively by the fluorophore 5-carboxyfluorescein (FAM) and the dark quencher 4-(4’-dimethylaminophenylazo)benzoic acid (Dabcyl).

Full-length recombinant NC protein was obtained as previously reported [45]. Protein concentration was determined by using a Lambda 20 (Perkin Elmer) UV-vis spectrophotometer using an extinction coefficient at 280 nm of 6410 M^{-1} cm^{-1}.

(9H-fluoren-9-yl)methyl 2-chloro-2-oxoethylcarbamate (2a; Fmoc-Gly-Cl).

Commercially available Fmoc-Gly-OH (Aldrich®, 2.5g, 8.40 mmol) was dissolved in anhydrous dichloromethane (40 mL) in a two-neck flask. Excess $SOCl_2$ (5 mL; 68.54mmol) was added dropwise. The mixture was heated at reflux for 30 minutes. Solvent was evaporated under reduced pressure and the residue was purified by crystallization from *n*-hexane, which afforded 2.57 g of pure **2a** as a white solid. Yield: 97%.

(S)-(9H-fluoren-9-yl)methyl 1-chloro-1-oxopropan-2-ylcarbamate (2b; Fmoc-L-Ala-Cl).

This compound was obtained by adopting the same procedure used for **2a**, starting from Fmoc-L-Ala-OH and $SOCl_2$. Yield: 98%. White solid.

(R)-(9H-fluoren-9-yl)methyl 1-chloro-1-oxopropan-2-ylcarbamate (2c; Fmoc-D-Ala-Cl).

This compound was obtained by adopting the same procedure used for **2a**, starting from Fmoc-D-Ala-OH and $SOCl_2$. Yield: 99%. White solid.

N,N'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(2-aminoacetamide)-bis-trifluoroacetate (3a).

Fmoc-Gly-Cl (**2a**, 7 g, 19.4 mmol) and pyridine (1.27 mL, 15.8 mmol) were slowly added to a solution of 2,6-diaminoanthraquinone (**1**, 350 mg, 1.94 mmol) in DMF (80 mL). The reaction mixture was stirred at room temperature for 24 hours. Solvent was removed by reduced pressure distillation. The residue was purified by crystallization from ethyl acetate, thus obtaining 2,6-bis[N-(2-Fmoc-amino)-acetamide]anthracene-9,10-dione (2,6-Fmoc-Gly-anthraquinone) as an orange powder (1.4 g, 1.68 mmol, yield 88%), which was used in the next step without any further purification. This compound was reacted with a 33% diethylamine solution in THF (60 mL) for 2 hours. Solvent was again distilled off under reduced pressure and the residue was finally reacted with a solution of trifluoroacetic acid in water (9:1, v/v, 20 mL) for 1 hour. Reaction mixture was then added with diethyl ether (60 mL). The precipitate was collected by centrifugation and dried to obtain 920 mg of intermediate **3a** as an intense orange solid. Yield 78% (overall). ^{1}H-NMR, ^{13}C-NMR and MS were consistent to that already reported in literature [46].

(2S,2'S)-N,N'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(2-amino-propanamide)-bis-trifluoroacetate (3b).

This compound was obtained by adopting the same procedure used for **3a**, starting from Fmoc-L-Ala-Cl and 2,6-diamminoanthraquinone. Orange solid. Yield: 75% (overall). ^{1}H-NMR (DMSO-d6): δ 11.26 (bs, 2H), 8.49 (d, 2H), 8.31 - 8.21 (m,

8H), 8.08 (dd, 2H), 2.70 (d, 6H), 2.52 (m, 2H). ^{13}C NMR (DMSO- d6): δ 181.1, 165.8, 143.6, 134.3, 128.6, 128.4, 123.6, 115.8, 49.5, 16.8. ESI-MS: 380.15[M + H]+; 191,18[M + 2H]++.

(2R,2'R)-N,N'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(2-amino-propanamide)-bis-trifluoroacetate (3c).

This compound was obtained by adopting the same procedure used for **3a**, starting from Fmoc-D-Ala-Cl and 2,6-diamminoanthraquinone. Orange solid. Yield: 73% (overall). ^{1}H-NMR (DMSO-d6): δ 11.28 (bs, 2H), 8.49 (d, 2H), 8.30 - 8.20 (m, 8H), 8.06 (dd, 2H), 2.69 (d, 6H), 2.50 (m, 2H). ^{13}C NMR (DMSO- d6): δ 181.2, 165.6, 143.6, 134.2, 128.4, 128.2, 123.4, 115.5, 49.4, 16.6. ESI-MS: 380.15[M + H]+; 191,18[M + 2H]++.

N,N'-(2,2'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(azanediyl)bis-(2-oxoethane-2,1-diyl))dipiperidine-4-carboxamide-bis-trifluoroacetate (4a).

Intermediate **3a** (250 mg, 0.42 mmol) was dissolved in dry DMF (18.6 mL), and DMAP (545.56 mg, 3.24 mmol), 1-Fmoc-piperidine-4-carboxylic acid (Fmoc-Inp-OH, 1.00 g, 3.20 mmol), and HBTU (1.36 g, 3.6 mmol) were added. The resulting solution was stirred at room temperature for 24 hours, then poured into Et_2O and centrifuged. The solid obtained was washed with Et_2O. The final product was dried *in vacuo* and used in the next step without any further purification. In order to remove the Fmoc- protecting group, the solid thus obtained was reacted with a 33% diethylamine solution in THF (30 mL) for 2 hours, poured into Et_2O and centrifuged.

The solid obtained was washed with Et_2O and water and then dried *in vacuo*. Subsequently, the obtained solid was reacted with a solution of trifluoroacetic acid in water (9:1, v/v, 20 mL) for 1 hour and then poured into Et_2O. The resulting suspension was centrifuged. The solid obtained was washed with Et_2O and dried *in vacuo*. Purification by preparative RP-HPLC afforded the pure compound **4a** as an intense orange solid. Yield: 79%. Orange solid. K' (HPLC): 9.10. ^{1}H-NMR (DMSO-d6, 400MHz): δ 10.71 (s, 2H), 8.67 (t, 2H), 8.46 (d, 2H), 8.36 (bs, 4H), 8.16 (d, 2H), 8.02 (dd, 2H), 3.76 (d, 4H), 3.27 (m, 4H), 2.91 (m, 4H), 2.54 (m, 2H), 1.90 (m, 4H), 1.75 (m, 4H). ^{13}C-NMR (DMSO-d6, 400MHz): δ 181.78, 174.15, 169.25, 158.64, 144.90, 134.85, 129.05, 128.54, 123.87, 116.24, 43.23, 42.89, 38.98, 25.58. ESI-MS: 575.2 [M + H]+; 288.3 [M + 2H]++.

(1R,1'R,4R,4'R)-N,N'-(2,2'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis-(azanediyl)bis(2oxoethane2,1diyl))bis(4(aminomethyl)cyclohexanecarboxamide)-bis-trifluoroacetate (4b).

The compound was obtained by using the same coupling procedure adopted for **4a,** starting from intermediate **3a** and trans-4-(Fmoc-aminomethyl)-cyclohexanecarboxylic acid (N-Fmoc-tranexamic acid). Yield: 57%. Orange solid. K' (HPLC): 9.4 ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.68 (s, 2H), 8.46 (d, 2H), 8.13 (d, 2H), 8.01 (dd, 2H), 7.91 (t, 2H), 7.76 (bs, 6H), 3.98 (d, 4H), 3.33 m (2H), 2.54 (s, 4H), 1.69 (m, 2H), 1.14 (m, 8H), 0.88 (m, 8H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ

182.09, 173.90, 171.23, 145.29, 135.05, 129.21, 128.42, 124.13, 116.31, 45.72, 43.36, 37.82, 35.63 29.55, 29.06. ESI-MS: 630.8 [M + H]+; 315.9 [M + 2H]++.

N,N'-(2,2'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(azanediyl)bis(2-oxo-ethane-2,1-diyl))bis(4-aminopiperidine-4-carboxamide)-tetratrifluoroacetate (4c).

The compound was obtained by using the same coupling procedure adopted for **4a**, starting from the intermediate **3a** and 1-Fmoc-4-(Fmoc-amino)-piperidine-4-carboxylic acid (Fmoc-Pip(Fmoc)-OH). Yield: 70%. Orange solid. K' (HPLC): 5.8. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.68 (s, 2H), 8.94 (bs, 4H), 8.65 (bs, 6H), 8.53 (t, 2H), 8.45 (d, 2H), 8.13 (d, 2H), 8.03 (dd, 2H), 3.98 (d, 4H), 3.24 (m, 4H), 2.85 (m, 4H), 2.40 (m, 4H), 1.96 (m, 4H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.98, 171.07, 168.92, 145.32, 135.04, 129.20, 128.68, 124.19, 116.42, 56.88, 43.25, 35.71, 28.96. ESI-MS: 604.9[M + H]+; 303.1 [M + 2H]++.

N,N'-(2,2'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(azanediyl)bis(2-oxo-ethane-2,1-diyl))bis(3-aminopropanamide)-bis-trifluoroacetate (4d).

The compound was obtained by using the same coupling procedure adopted for **4a**, starting from the intermediate **3a** and Fmoc-β-Alanine-OH (Fmoc-β-Ala-OH). Yield: 79%. Yellow solid. K' (HPLC): 3.5. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.68(s, 2H), 8.54 (t, 2H), 8.45 (d, 2H), 8.16 (d, 2H), 8.04 (dd, 2H), 7.45 (bs, 6H), 3.98 (d, 4H), 2.97 (t, 4H), 2.52 (t, 4H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.78,

170.76, 169.17, 144.83, 134.86, 129.06, 128.61, 123.96, 116.35, 43.33, 36.03, 33.01. ESI-MS: 494.9 [M + H]+; 248.2 [M + 2H]++.

N,N'-(2,2'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(azanediyl)bis(2-oxo-ethane-2,1-diyl))bis(2,3-diaminopropanamide)-tetra-trifluoroacetate (4e).

The compound was obtained by using the same coupling procedure adopted for **4a,** starting from intermediate **3a** and N^{α}-Fmoc-N^{β}-Boc-L-2,3-diaminopropionic acid (Fmoc-Dap(Boc)-OH). Yield: 83%. Yellow solid. K' (HPLC): 3.40. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.88 (s, 2H), 8.97 (bs, 8H), 8.46 (t, 2H), 8.19 (d, 2H), 8.17 (d, 2H), 8.04 (dd, 2H), 4.12 (d, 4H), 3.16 (m, 4H), 2.47 (m, 2H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.75, 168.69, 158.76, 144.63, 134.86, 129.09, 128.75, 124.08, 116.43, 62.23, 44.18, 41.61. ESI-MS: 525.2 [M + H]+; 263.2 [M + 2H]++.

N,N'-(2,2'-(9,10-dioxo-9,10-dihydroanthracene-2,7-diyl)bis(azanediyl)bis(2-oxo-ethane-2,1-diyl))bis(2,4-diaminobutanamide)-tetra-trifluoroacetate (4f).

The compound was obtained by using the same coupling procedure adopted for **4a,** starting from intermediate **3a** and N^{α}-Fmoc-N^{γ}-Boc-L-2,4-diaminobutyric acid (Fmoc-Dab(Boc)-OH). Yield: 86%. Yellow solid. K' (HPLC): 3.26. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.90 (s, 2H), 9.03 (t, 2H), 8.48 (bs, 6H), 8.21 (d, 2H), 8.19 (bs, 6H), 8.17 (d, 2H), 7.99 (dd, 2H), 3.99 (d, 4H), 2.98 (m, 2H), 2.05 (m, 4H), 1.24 (m, 4H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.80, 168.78, 168.56, 144.72, 134.88, 129.10, 128.70, 123.90, 116.28, 50.32, 43.18, 35.43, 29.47. ESI-MS: 553.1 [M + H]+; 277.2 [M + 2H]++.

N,N'-(2,2'-(9,10-dioxo-9,10-dihydroanthracene-2,7-diyl)bis(azanediyl)bis(2-oxo-ethane-2,1-diyl))bis(2,5-diaminopentanamide)-tetra-trifluoroacetate (4g).

The compound was obtained by using the same coupling procedure adopted for **4a,** starting from intermediate **3a** and N^{α}-Fmoc-N^{δ}-Boc-L-ornithine (Fmoc-Orn(Boc)-OH). Yield: 80%. Yellow solid. K' (HPLC): 3.94. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.78 (s, 2H), 8.66 (t, 2H), 8.52 (d, 2H), 8.25 (bs, 6H), 8.16 (d, 2H), 8.04 (dd, 2H), 7.96 (bs, 6H), 3.98 (d, 4H), 3.42 (m, 4H), 2.68 (m, 2H), 1.65 (m, 4H), 1.53 (m, 4H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.88, 169.38, 168.57, 144.55, 134.72, 129.04, 128.66, 124.22, 116.45, 52.12, 43.20, 38.85, 28.47, 22.78. ESI-MS: 581.1 [M + H]+; 291.3 [M + 2H]++.

N,N'-(2,2'-(9,10-dioxo-9,10-dihydroanthracene-2,7-diyl)bis(azanediyl)bis(2-oxo-ethane-2,1-diyl))bis(2,6-diaminohexanamide)-tetra-trifluoroacetate(4h).

The compound was obtained by using the same coupling procedure adopted for **4a,** starting from intermediate **3a** and N^{α}-Fmoc-N^{ε}-Boc-L-Lysine(Fmoc-Lys(Boc)-OH. Yield: 84%. Yellow solid. K' (HPLC): 3.81. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.85 (s, 2H), 8.92 (t, 2H), 8.21 (bs, 6H), 8.18 (d, 2H), 8.16 (d, 2H), 8.03 (dd, 2H), 7.82 (bs, 6H), 4.04 (d, 4H), 3.87 (m, 2H), 2.76 (m, 4H), 1.75 (m, 4H), 1.56 (m, 4H), 1.41 (m, 4H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.88, 171.12, 169.08, 145.23, 135.23, 129.12, 128.78, 124.04, 116.53, 52.20, 38.63, 35.38, 28.78, 23.02. ESI-MS: 609.01 [M + H]+; 305.3 [M + 2H]++.

N,N'-(2,2'-(9,10-dioxo-9,10-dihydroanthracene-2,7-diyl)bis(azanediyl)bis(2-oxo-ethane-2,1-diyl))bis(2-amino-3-(4-aminophenyl)propanamide)-tetra-trifluoroacetate(4i).

The compound was obtained by using the same coupling procedure adopted for **4a,** starting from intermediate **3a** and Boc-4-(Fmoc-amino)-L-phenylalanine (Boc-4-(NH-Fmoc)-Phe-OH). Yield: 77%. Orange solid. K' (HPLC): 3.85. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.82 (s, 2H), 8.96 (t, 2H), 8.47 (bs, 6H), 8.19 (d, 2H), 8,17 (d, 2H), 8.15 (bs, 6H), 8.04 (dd, 2H), 7.17 (d, 4H), 6.91 (d, 4H), 4.07 (d, 4H), 3.07 (m, 2H), 2.91 (m, 4H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.79, 169.20, 168.57, 158.72, 144.76, 134.88, 130.92, 129.13, 128.68, 128.61, 123.99, 118.97, 116.36, 53.94, 43.36, 36.87. ESI-MS: 677.2 [M + H]+; 339.2 [M + 2H]++.

N,N'-(((9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(azanediyl))bis(2-oxo-ethane-2,1-diyl))bis(2-amino-3-(4-hydroxyphenyl)propanamide), 2,2,2-trifluoro-acetate salt (4l).

The compound was obtained by using the same coupling procedure adopted for **4a**, starting from the intermediate **3a** and *N*-(9-Fluorenylmethoxycarbonyl)-*O*-tert-butyl-L-tyrosine (Fmoc-Tyr(tBU)-OH. Yield: 69%. Orange solid. K'(HPLC): 3.96. ^{1}H-NMR (DMSO-d_6, 400MHz): δ ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.85 (s, 2H), 8.22 (t, 2H), 8.18 (d, 2H), 8.17 (d, 2H), 8.02 (dd, 2H), 7.19 (d, 4H), 6.82 (d,4H), 6.74 (bs, 6H), 5.15 (bs, 2H), 4.04 (d, 4H), 3.87 (m, 2H), 2.76 (m, 4H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.88, 171.12, 169.08, 155.7, 145.23, 135.23, 129.12, 128.78, 128.68,

128.61, 124.04, 118.83, 116.53, 52.20, 38.63, 35.38. ESI-MS: 679.2 [M + H]+; 340.2 [M + 2H]++.

N,N'-(2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(azanediyl)-bis(1-oxopropane-2,1-diyl)dipiperidine-4-carboxamide-bis-trifluoroacetate (5a).

Intermediate **3b** (250 mg, 0.42 mmol) was dissolved in dry DMF (18.6 mL), and DMAP (545.56 mg, 3.24 mmol), 1-Fmoc-piperidine-4-carboxylic acid (Fmoc-Inp-OH, 1.00 g, 3.20 mmol), and HBTU (1.36 g, 3.6 mmol) were added. The resulting solution was stirred at room temperature for 24 hours, then poured into Et_2O and centrifuged. The solid obtained was washed with Et_2O and the final product was dried *in vacuo* and used in the next step without any further purification. In order to remove the Fmoc- protecting group, the solid obtained was reacted with a 33% diethylamine solution in THF (30 mL) for 2 hours, poured into Et_2O and centrifuged. The solid obtained was washed with Et_2O and water and then dried *in vacuo*. Subsequently, the obtained solid was reacted with a solution of trifluoroacetic acid in water (9:1, v/v, 20 mL) for 1 hour and then poured into Et_2O. The resulting suspension was centrifuged. The solid obtained was washed with Et_2O and dried *in vacuo*. Purification by preparative RP-HPLC afforded the pure compound **5a** as an intense orange solid. Yield: 79%. Orange solid. K' (HPLC): 6.3. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.70 (s, 2H), 8.71 (t, 2H), 8.45 (bs, 4H), 8.34 (d, 2H), 8.14 (bs, 4H), 8.04 (dd, 2H), 4.37 (m, 2H), 3.26 (m, 4H), 2.85 (m, 4H), 1.82 (m, 4H), 1.69 (m, 4H),

1.31 (d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.78, 172.06, 168.72, 144.89, 134.82, 129.05, 128.69, 124.11, 116.45, 51.90, 50,01, 40.58, 38.72, 28.69, 23.01, 18.31. ESI-MS: 603.3 [M + H]+; 302.1 [M + 2H]++.

(1R,1'R,4S,4'S)-N,N'-((2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)-bis(azanediyl)bis(1oxopropane2,1diyl))bis-(4(aminomethyl)cyclohexane-carboxamide)-bis-trifluoroacetate (5b).

The compound was obtained by using the same coupling procedure adopted for **5a,** starting from intermediate **3b** and 4-(Fmoc-aminomethyl)cyclohexanecarboxylic acid (N-Fmoc-tranexamic acid). Yield: 67%. Orange solid. K' (HPLC): 6.2. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.67 (s, 2H), 8.46 (t, 2H), 8,16 (d, 2H), 8.14 (d, 2H), 8.06 (dd, 2H), 7.76 (bs, 6H), 4.38 (m, 2H), 2.84 (m, 4H), 2.64 (m, 2H), 1.77 (m, 2H), 1.68 (m, 8H), 1.31 (d, 6H), 0.89 (m, 8H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 182.04, 172.44, 171.85, 144.69, 135.01, 129.17, 128.52, 124.13, 116.41, 51.84, 43.96, 38.02, 35.69 29.65, 29.13, 17.96. ESI-MS: 659.1 [M + H]+; 330.2 [M + 2H]++.

N,N'-(1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(azanediyl)bis(1-oxo-propane-2,1-diyl))bis(4-aminopiperidine-4-carboxamide)-tetra-trifluoroacetate (5c).

The compound was obtained by using the same coupling procedure adopted for **5a,** starting from intermediate **3b** and 1-Fmoc-4-(Fmoc-amino)-piperidine-4-carboxylic acid (Fmoc-Pip(Fmoc)-OH). Yield: 66%. Orange solid. K' (HPLC): 4.7. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.69 (s, 2H), 8.45 (t, 2H), 8.32 (d, 2H), 8.18 (bs,

4H), 8.15 (d, 2H), 8.02 (dd, 2H), 7.92 (bs, 6H), 4.32 (m, 2H), 3.57 (m, 4H), 3.14 (m, 4H), 2.86 (m, 4H), 2.70 (m, 4H), 1.31 (d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 182.05, 171.33, 169.12, 145.23, 134.99, 129.24, 128.62, 124.23, 116.46, 56.88, 51.23, 35.71, 28.96, 17.93. ESI-MS: 633.4 [M + H]+; 317.2 [M + 2H]++.

N,N'-(2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6diyl)bis(azanediyl)bis-(1-oxopropane-2,1-diyl)bis(2,4-diaminobutanamide)-tetra-trifluoroacetate (5d).

The compound was obtained by using the same coupling procedure adopted for **5a**, starting from the intermediate **3b** and Fmoc-β-Alanine-OH (Fmoc-β-Ala-OH). Yield: 82%. Yellow solid. K' (HPLC): 3.9. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.86 (s, 2H), 8.94 (t, 2H), 8.48 (d, 2H), 8.19 (d, 2H), 8.06 (dd,, 2H), 7.78 (bs, 6H), 4.56 (m, 2H), 2.81 (m, 4H), 1.75 (m, 4H), 1.39 (d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.98, 171.16, 169.44, 144.98, 134.86, 129.26, 128.61, 124.06, 116.53, 51.33, 36.03, 33.01, 18.13. ESI-MS: 523.5 [M + H]+; 262.1 [M + 2H]++.

N,N'-(2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6diyl)bis(azanediyl)-bis(1-oxopropane-2,1-diyl)bis(2,3-diaminopropanamide)-tetra-trifluoroacetate (5e).

The compound was obtained by using the same coupling procedure adopted for **5a,** starting from intermediate **3b** and N^{α}-Fmoc-Nβ-Boc-L-diaminopropionic acid (Fmoc-Dap(Boc)-OH). Yield: 86%. Orange solid. K' (HPLC): 4.8. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.92 (s, 2H), 9.01 (t, 2H), 8.47 (bs, 6H), 8.20 (bs, 6H), 8.17 (d, 2H), 8.06 (d, 2H), 8.04 (dd, 2H), 4.54 (m, 2H), 4.20 (m, 2H), 3.15 (m, 4H), 1.42

(d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.81, 172.19, 158.77, 145.65, 134.81, 129.07, 128.85, 124.26, 116.61, 50.82, 50.36, 40.15, 18.30. ESI-MS: 553.5 [M + H]+; 277.2 [M + 2H]++.

N,N'-(2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(azanediyl)-bis(1-oxopropane-2,1-diyl)bis(2,4-diaminobutanamide)-tetra-trifluoroacetate (5f).

The compound was obtained by using the same coupling procedure adopted for **5a,** starting from intermediate **3b** and N^{α}-Fmoc-N^{γ}-Boc-L-diaminobutyric acid (Fmoc-Dab(Boc)-OH). Yield: 78%. Orange solid. K' (HPLC): 4.6. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.84 (s, 2H), 8.86 (t, 2H), 8.45 (bs, 6H), 8.17 (d, 2H), 8.16 (d, 2H), 8.05 (dd, 2H), 7.78 (bs, 6H), 4.51 (m, 2H), 3.84 (m, 2H), 2.82 (m, 4H), 1.75 (m, 4H), 1.40 (d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.87, 171.98, 168.66, 145.32, 134.88, 129.09, 128.77, 124.24, 116.55, 50.85, 50.32, 35.43, 29.47, 18.13. ESI-MS: 581.4 [M + H]+; 291.3 [M + 2H]++.

N,N'-(1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(azanediyl)bis(1-oxo-propane-2,1-diyl))bis(2,5-diaminopentanamide)-tetra-trifluoroacetate (5g).

The compound was obtained by using the same coupling procedure adopted for **5a,** starting from intermediate **3b** and N^{α}-Fmoc-N^{δ}-Boc-L-ornithine (Fmoc-Orn(Boc)-OH). Yield: 82%. Orange solid. K' (HPLC): 4.7. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.87 (s, 2H), 8.87 (t, 2H), 8.46 (bs, 6H), 8.18 (d, 2H), 8.16 (d, 2H), 8.07 (dd, 2H), 7.86 (bs, 6H), 4.50 (m, 2H), 3.86 (m, 2H), 3.16 (m, 4H), 2.82 (m, 4H), 1.77 (m, 4H),

1.40 (d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 182.02, 169.78, 168.17, 144.75, 134.23, 129.09, 128.68, 124.24, 116.47, 52.12, 50.24, 38.85, 28.47, 22.78, 18.19. ESI-MS: 609.4 [M + H]+; 305.4 [M + 2H]++.

N,N'-(2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydro-anthracene-2,6diyl)bis(azanediyl)-bis(1-oxopropane-2,1-diyl)bis-(2,6-diaminohexanamide)-tetra-trifluoroacetate (5h).

The compound was obtained by using the same coupling procedure adopted for **5a,** starting from intermediate **3b** and N^{α}-Fmoc-N^{ε}-Boc-L-Lysine (Fmoc-Lys(Boc)-OH. Yield: 84%. Orange solid. K' (HPLC): 4.7. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.85 (s, 2H), 8.86 (t, 2H), 8.45 (bs, 6H), 8.18 (d, 2H), 8.16 (d, 2H), 8.06 (dd, 2H), 7.79 (bs, 6H), 4.49 (q, 2H), 3.86 (m, 2H), 2.74 (m, 4H), 1.73 (m, 4H), 1.54 (m, 4H), 1.39 (d, 6H), 1.34 (m, 4H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.88, 172.20, 168.92, 158.72, 144.92, 134.83, 129.08, 128.67, 124.06, 116.43, 52.22, 50.03, 38.96, 30.88, 26.91, 21.49, 18.10. ESI-MS: 636.9 [M + H]+; 319.1 [M + 2H]++.

N,N'-(2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6diyl)bis(azanediyl)bis-(1-oxopropane-2,1-diyl)bis(2-amino-3-(4-aminophenyl)propanamide)-tetra-trifluoroacetate (5i).

The compound was obtained by using the same coupling procedure adopted for **5a,** starting from intermediate **3b** and Fmoc-4-(Boc-amino)-L-phenylalanine (Fmoc-4-(NH-Boc)-OH). Yield: 72%. Orange solid. K' (HPLC): 4.8. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.78 (s, 2H), 8.75 (t, 2H), 8.47 (bs, 6H), 8.18 (d, 2H), 8.15 (d, 2H), 8.04

(dd, 2H), 7.01 (d, 4H), 6.97 (d, 4H), 6.74 (bs, 6H), 4.45 (m, 2H), 4.00 (m, 2H), 2.91 (m, 4H), 1.23 (d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 182.09, 169.72, 168.37, 158.33, 144.46, 134.98, 130.96, 129.03, 128.88, 128.61, 123.99, 118.87, 116.52, 53.94, 50.33, 36.87, 18.03. ESI-MS: 705.4 [M + H]+; 353.3 [M + 2H]++.

N,N'-(2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydroanthracene-2diyl)bis(azanediyl)bis(1-oxopropane-2,1-diyl)bis(azanediyl)bis(3-(4-hydroxyphenyl)-1-oxopropan-2-aminium) 2,2,2-trifluoroacetate (5l).

The compound was obtained by using the same coupling procedure adopted for **5a,** starting from intermediate **3b** and *N*-(9-Fluorenylmethoxycarbonyl)-*O*-tert-butyl-L-tyrosine (Fmoc-Tyr(tBU)-OH. Yield: 57%. Orange solid. K' (HPLC): 5.79. ^{1}H-NMR (DMSO-d_6, 400MHz): δ ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.85 (s, 2H), 8.22 (t, 2H), 8.18 (d, 2H), 8.17 (d, 2H), 8.02 (dd, 2H), 7.19 (d, 4H), 6.82 (d,4H), 6.74 (bs, 6H), 5.14 (bs, 2H), 4.23 (m, 2H), 3.84 (m, 2H), 2.76 (m, 4H), 1.23 (d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.86, 170.89, 168.98, 155.6, 145.25, 135.33, 129.11, 128.78, 128.68, 128.60, 124.02, 118.78, 116.49, 51.20, 37.63, 35.40, 18.4. ESI-MS: 707.9 [M + H]+; 354.8 [M + 2H]++.

N,N'-(2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(azanediyl)-bis(1-oxopropane-2,1-diyl)dipiperidine-4-carboxamide-bis-trifluoroacetate (6a).

Intermediate **3c** (250 mg, 0.42 mmol) was dissolved in dry DMF (18.6 mL), and DMAP (545.56 mg, 3.24 mmol), 1-Fmoc-piperidine-4-carboxylic acid (Fmoc-Inp-OH, 1.00 g, 3.20 mmol), and HBTU (1.36 g, 3.6 mmol) were added. The

resulting solution was stirred at room temperature for 24 hours, then poured into Et_2O and centrifuged. The solid obtained was washed with Et_2O and the final product was dried *in vacuo* and used in the next step without any further purification. In order to remove the Fmoc- protecting group, the solid obtained was reacted with a 33% diethylamine solution in THF (30 mL) for 2 hours, poured into Et_2O and centrifuged. The solid obtained was washed with Et_2O and water and then dried *in vacuo*. Subsequently, the obtained solid was reacted with a solution of trifluoroacetic acid in water (9:1, v/v, 20 mL) for 1 hour and then poured into Et_2O. The resulting suspension was centrifuged. The solid obtained was washed with Et_2O and dried *in vacuo*. Purification by preparative RP-HPLC afforded the pure compound **6a** as an intense orange solid. Yield: 76%. Orange solid. K' (HPLC): 5,5. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.71 (s, 2H), 8.71 (t, 2H), 8.44 (bs, 4H), 8.33 (d, 2H), 8.14 (bs, 4H), 8.03 (dd, 2H), 4.37 (m, 2H), 3.24 (m, 4H), 2.84 (m, 4H), 1.83 (m, 4H), 1.68 (m, 4H), 1.31 (d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.79, 172.07, 168.73, 144.89, 134.82, 129.04, 128.68, 124.11, 116.44, 51.90, 50,02, 40.59, 38.72, 28.68, 23.02, 18.31. ESI-MS: 603.3 [M + H]+; 302.1 [M + 2H]++.

(1R,1'R,4S,4'S)-N,N'-((2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)-bis(azanediyl)bis(1oxopropane2,1diyl))bis-(4(aminomethyl)cyclohexane-carboxamide)-bis-trifluoroacetate (6b).

The compound was obtained by using the same coupling procedure adopted for **6a,** starting from intermediate **3c** and 4-(Fmoc-aminomethyl)cyclohexanecarboxylic

acid (N-Fmoc-tranexamic acid). Yield: 65%. Orange solid. K' (HPLC): 6.4. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.69 (s, 2H), 8.47 (t, 2H), 8,18 (d, 2H), 8.16 (d, 2H), 8.05 (dd, 2H), 7.77 (bs, 6H), 4.36 (m, 2H), 2.85 (m, 4H), 2.64 (m, 2H), 1.76 (m, 2H), 1.67 (m, 8H), 1.31 (d, 6H), 0.88 (m, 8H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 182.02, 172.43, 171.83, 144.69, 135.03, 129.18, 128.52, 124.11, 116.41, 51.82, 43.97, 38.02, 35.67 29.66, 29.13, 17.95. ESI-MS: 659.1 [M + H]+; 330.2 [M + 2H]++.

N,N'-(1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(azanediyl)bis(1-oxo-propane-2,1-diyl))bis(4-aminopiperidine-4-carboxamide)-tetra-trifluoroacetate (6c).

The compound was obtained by using the same coupling procedure adopted for **6a,** starting from intermediate **3c** and 1-Fmoc-4-(Fmoc-amino)-piperidine-4-carboxylic acid (Fmoc-Pip(Fmoc)-OH). Yield: 63%. Orange solid. K' (HPLC): 4.7. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.69 (s, 2H), 8.46 (t, 2H), 8.31 (d, 2H), 8.18 (bs, 4H), 8.15 (d, 2H), 8.03 (dd, 2H), 7.90 (bs, 6H), 4.34 (m, 2H), 3.58 (m, 4H), 3.14 (m, 4H), 2.85 (m, 4H), 2.71 (m, 4H), 1.31 (d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 182.04, 171.37, 169.08, 145.22, 135.00, 129.21, 128.59, 124.19, 116.45, 56.87, 51.23, 35.72, 28.96, 17.94. ESI-MS: 633.4 [M + H]+; 317.2 [M + 2H]++.

N,N'-(2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6diyl)bis(azanediyl)bis-(1-oxopropane-2,1-diyl)bis(2,4-diaminobutanamide)-tetra-trifluoroacetate (6d).

The compound was obtained by using the same coupling procedure adopted for **6a**, starting from the intermediate **6c** and Fmoc-β-Alanine-OH (Fmoc-β-Ala-OH).

Yield: 81%. Yellow solid. K' (HPLC): 4.1. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.85 (s, 2H), 8.91 (t, 2H), 8.48 (d, 2H), 8.18 (d, 2H), 8.07 (dd,, 2H), 7.78 (bs, 6H), 4.54 (m, 2H), 2.82 (m, 4H), 1.75 (m, 4H), 1.38 (d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.99, 171.16, 169.46, 144.96, 134.87, 129.25, 128.61, 124.07, 116.55, 51.33, 36.02, 33.03, 18.13. ESI-MS: 523.5 [M + H]+; 262.1 [M + 2H]++.

N,N'-(2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6diyl)bis(azanediyl)-bis(1-oxopropane-2,1-diyl)bis(2,3-diaminopropanamide)-tetra-trifluoroacetate (6e).

The compound was obtained by using the same coupling procedure adopted for **6a,** starting from intermediate **3c** and N^{α}-Fmoc-N^{β}-Boc-L-diaminopropionic acid (Fmoc-Dap(Boc)-OH). Yield: 86%. Orange solid. K' (HPLC): 4.8. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.92 (s, 2H), 9.00 (t, 2H), 8.47 (bs, 6H), 8.19 (bs, 6H), 8.18 (d, 2H), 8.06 (d, 2H), 8.04 (dd, 2H), 4.55 (m, 2H), 4.21 (m, 2H), 3.15 (m, 4H), 1.43 (d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.81, 172.169, 158.77, 144.65, 134.81, 129.07, 128.85, 124.26, 116.61, 50.82, 50.36, 40.15, 18.30. ESI-MS: 553.5 [M + H]+; 277.2 [M + 2H]++.

N,N'-(2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(azanediyl)-bis(1-oxopropane-2,1-diyl)bis(2,4-diaminobutanamide)-tetra-trifluoroacetate (6f).

The compound was obtained by using the same coupling procedure adopted for **6a,** starting from intermediate **3c** and N^{α}-Fmoc-N^{γ}-Boc-L-diaminobutyric acid

(Fmoc-Dab(Boc)-OH). Yield: 79%. Orange solid. K' (HPLC): 4.8. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.81 (s, 2H), 8.87 (t, 2H), 8.45 (bs, 6H), 8.17 (d, 2H), 8.16 (d, 2H), 8.04 (dd, 2H), 7.78 (bs, 6H), 4.53 (m, 2H), 3.82 (m, 2H), 2.84 (m, 4H), 1.75 (m, 4H), 1.41 (d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.88, 172.01, 168.69, 145.33, 134.87, 129.11, 128.73, 124.24, 116.55, 50.86, 50.31, 35.41, 29.48, 18.13. ESI-MS: 581.4 [M + H]+; 291.3 [M + 2H]++.

N,N'-(1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(azanediyl)bis(1-oxo-propane-2,1-diyl))bis(2,5-diaminopentanamide)-tetra-trifluoroacetate (6g).

The compound was obtained by using the same coupling procedure adopted for **6a,** starting from intermediate **3c** and N^{α}-Fmoc-N^{δ}-Boc-L-ornithine (Fmoc-Orn(Boc)-OH). Yield: 81%. Orange solid. K' (HPLC): 4.8. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.88 (s, 2H), 8.86 (t, 2H), 8.46 (bs, 6H), 8.18 (d, 2H), 8.16 (d, 2H), 8.06 (dd, 2H), 7.85 (bs, 6H), 4.51 (m, 2H), 3.85 (m, 2H), 3.16 (m, 4H), 2.83 (m, 4H), 1.77 (m, 4H), 1.41 (d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.99, 169.76, 168.13, 144.74, 134.24, 129.09, 128.69, 124.24, 116.49, 52.13, 50.24, 38.84, 28.44, 22.79, 18.18. ESI-MS: 609.4 [M + H]+; 305.4 [M + 2H]++.

N,N'-(2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydro-anthracene-2,6diyl)bis(azanediyl)-bis(1-oxopropane-2,1-diyl)bis-(2,6-diaminohexanamide)-tetra-trifluoroacetate (6h).

The compound was obtained by using the same coupling procedure adopted for **5a,** starting from intermediate **3c** and N^{α}-Fmoc-N^{ε}-Boc-L-Lysine (Fmoc-Lys(Boc)-OH.

Yield: 83%. Orange solid. K' (HPLC): 4.9. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.86 (s, 2H), 8.86 (t, 2H), 8.44 (bs, 6H), 8.18 (d, 2H), 8.16 (d, 2H), 8.06 (dd, 2H), 7.78 (bs, 6H), 4.50 (q, 2H), 3.85 (m, 2H), 2.75 (m, 4H), 1.72 (m, 4H), 1.54 (m, 4H), 1.40 (d, 6H), 1.33 (m, 4H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.87, 172.19, 168.90, 158.75, 144.93, 134.84, 129.09, 128.67, 124.08, 116.44, 52.22, 50.04, 38.97, 30.88, 26.93, 21.48, 18.12. ESI-MS: 636.9 [M + H]+; 319.1 [M + 2H]++.

N,N'-(2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydroanthracene-2,6diyl)bis(azanediyl)bis-(1-oxopropane-2,1-diyl)bis(2-amino-3-(4-aminophenyl)propanamide)-tetra-trifluoroacetate (6i).

The compound was obtained by using the same coupling procedure adopted for **6a,** starting from intermediate **3c** and Fmoc-4-(Boc-amino)-L-phenylalanine (Fmoc-4-(NH-Boc)-OH). Yield: 69%. Orange solid. K' (HPLC): 5.1. ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.74 (s, 2H), 8.74 (t, 2H), 8.44 (bs, 6H), 8.18 (d, 2H), 8.15 (d, 2H), 8.02 (dd, 2H), 7.01 (d, 4H), 6.94 (d, 4H), 6.72 (bs, 6H), 4.44 (m, 2H), 4.00 (m, 2H), 2.88 (m, 4H), 1.22 (d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 182.05, 169.72, 168.36, 158.32, 144.45, 134.95, 130.94, 129.03, 128.84, 128.58, 123.93, 118.84, 116.52, 53.93, 50.31, 36.87, 18.03. ESI-MS: 705.2 [M + H]+; 353.1 [M + 2H]++.

N,N'-(2S,2'S)-1,1'-(9,10-dioxo-9,10-dihydroanthracene-2diyl)bis(azanediyl)bis(1-oxopropane-2,1-diyl)bis(azanediyl)bis(3-(4-hydroxyphenyl)-1-oxopropan-2-aminium) 2,2,2-trifluoroacetate (6l).

The compound was obtained by using the same coupling procedure adopted for **6a,** starting from intermediate **3c** and *N*-(9-Fluorenylmethoxycarbonyl)-*O*-tert-butyl-L-tyrosine(Fmoc-Tyr(tBU)-OH. Yield: 62%. Orange solid. K' (HPLC): 5.8. ^{1}H-NMR (DMSO-d_6, 400MHz): δ ^{1}H-NMR (DMSO-d_6, 400MHz): δ 10.86 (s, 2H), 8.24 (t, 2H), 8.18 (d, 2H), 8.17 (d, 2H), 8.02 (dd, 2H), 7.18 (d, 4H), 6.84 (d,4H), 6.73 (bs, 6H), 5.14 (bs, 2H), 4.22 (m, 2H), 3.85 (m, 2H), 2.74 (m, 4H), 1.24 (d, 6H). ^{13}C-NMR (DMSO-d_6, 400MHz): δ 181.87, 170.87, 168.96, 155.53, 145.24, 135.33, 129.09, 128.77, 128.69, 128.62, 124.04, 118.78, 116.47, 51.21, 37.63, 35.42, 18.3. ESI-MS: 707.9 [M + H]+; 354.8 [M + 2H]++.

4.2 High Throughput Screening (HTS)

The ability of each compound to stabilize nucleic acids was measured by the increase of melting temperature of the oligonucleotide in presence of each anthraquinone. Melting temperature (*Tm*) is the temperature at which 50% molecules of oligonucleotide are denatured. TAR, cTAR or annealed TAR/cTAR formed by thermal denaturation [47] were folded at 10 μM in TNMg (Tris-HCl 10 mM, NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM pH 7.5) and then diluted to 1 μM concentration in ETN (EDTA 1 mM, Tris HCl 10 mM, NaCl 20 mM, pH 7.5). In each microplate well, the nucleic acid solutions were mixed with the anthraquinones solutions to the final concentrations of 1, 10 and 100 μM. Nucleic acid solutions without compound were used to measure the reference *Tm* value. The melting protocol consisted of a melting phase, in which the temperature increased from 25 °C to 99 °C in 1 h (0.02 °C/s).

Fluorescence emission of FAM was read by using a Light Cycler 480 II (Roche) with emission at λ=510 nm and correlated to the melting temperature of the oligonucleotide. The *Tm* value was mathematically derived from the thermal denaturing profile by using LC480 software. *ΔTm* was calculated by using the following equation: $\Delta Tm = Tm_2 - Tm_1$, where Tm_2 and Tm_1 are the *Tm* values measured by testing the oligonucleotides (or the hybrid) in the presence or absence of compound, respectively.

4.3 Fluorescence Quenching Assay (FQA)

High Throughput Screening (HTS) was performed to identify inhibitors of NC chaperone activity on either TAR or cTAR. We used a Victor III (Perkin Elmer) microplate reader with 485 and 535 nm as excitation and emission wavelengths. A 1 μM aliquot of either TAR or cTAR, which bore 5'-FAM and 3'-DAB modifications, was heat refolded in TNMg (Tris-HCl 10 mM, NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM pH 7.5). Briefly, the oligonucleotides were denatured at 95 °C for 5 min and then left to cool to room temperature in order to assume their *stem-bulge-loop* structure. The samples were then diluted to 0.1 μM in TN (Tris-HCl 10 mM, NaCl 20 mM pH 7.5). Increasing concentrations of compound (0, 0.1, 0.5, 1, 5, 10, 50, 100 μM final) were added to each sample before introducing NC to a final concentration of 0.8 μM (for a 1:8 oligo to NC molar ratio). The plate was read three times with 1 min intervals, unless differently specified. The experimental data were fitted as reported earlier to enable calculation of the respective IC_{50} value. Each experiment was performed in triplicate to calculate an average and standard deviation.

4.4 Nucleocapsid Annealing Mediated Electrophoresis (NAME) assay

Nucleocapsid Annealing Mediated Electrophoresis (NAME) assay was developed to investigate the ability of compounds to impair the biological activity of the full-length NC protein, which involved monitoring the annealing of TAR with cTAR. Solutions containing 1μM concentrations of TAR, cTAR and the hybrid TAR/cTAR were heat re-folded in TNMg (Tris-HCl 10 mM, NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM pH 7.5). Briefly, the oligonucleotides were denatured at 95 °C for 5 min and then left to cool to room temperature in order to assume their *stem-bulge-loop* (TAR and cTAR) or double-stranded (hybrid TAR/cTAR) structure. Two different assay formats were employed to evaluate the inhibition of NC-mediated TAR/cTAR hybrid formation:

NC-preincubation. Full-length recombinant NC protein (8 μM) was preincubated with increasing concentrations (0, 1, 10, 20, 50, 100 μM final concentrations) of each anthraquinone derivative for 15 min at room temperature. Folded TAR (1 μM) and cTAR (1 μM) were mixed with the NC-anthraquinone solutions and then incubated for 15 min at room temperature. The samples were added with Gel Loading Buffer containing SDS (GLB_{SDS}: Tris-HCl 100 mM, EDTA 4 mM, 50% w/v glycerol, 2% w/v SDS, 0.05% w/v bromophenol blue), kept on ice and finally resolved on a 12% native PAGE (Acrylamide:Bisacrylamide = 19:1) in TBE buffer (Tris 89 mM, Boric acid 89 mM and EDTA 2 mM, pH 8) for 3 h at 200 V.

Oligo-preincubation. TAR (1 μM) and cTAR (1 μM) were folded separately as described above and then individually preincubated with increasing concentrations of

compound (each oligo with 0, 1, 10, 20, 50, 100 μM final concentrations) for 15 min at room temperature. Aliquots of each construct were mixed together, added with NC solution (8 μM), and then incubated for other 15 min at room temperature. The samples were added with Gel Loading Buffer containing SDS (GLB_{SDS}: Tris-HCl 100 mM, EDTA 4 mM, 50% w/v glycerol, 2% w/v SDS, 0.05% w/v bromophenol blue), kept on ice, and finally analyzed on a 12% native PAGE (19:1 acrylamide to bisacrylamide ratio) in TBE buffer (Tris 89 mM, Boric acid 89 mM and EDTA 2 mM, pH 8) for 3 h at 200 V.

After electrophoresis, all gels were stained with SybrGreen II and analyzed on a Geliance 600 Imaging System (PerkinElmer). Gene Tools software (PerkinElmer) was employed to quantify the percentage of hybrid formation, which enabled the calculation of the sought-after IC_{50} value, namely the concentration of compound required to inhibit hybrid formation by half.

4.5 ESI-MS analysis

TAR and cTAR were separately folded in 150 mM ammonium acetate (pH 7.5) containing 1μM $MgCl_2$. The oligonucleotides were denatured at 95 °C for 5 min and then left to cool to room temperature in order to assume their *stem-bulge-loop* structure. Before mixing with compounds, folded TAR and cTAR were filtered by using 3K NMWL (Millipore Corporation, MA, USA) centrifugal filters to minimize the presence of magnesium salts that can adversely interfere with ESI performance.

Samples for binding studies were prepared by mixing appropriate volumes of folded TAR or cTAR (1 μM final) with each compound in 150 mM ammonium acetate (pH 7.5). The final mixtures contained up to a 10:1 of 2,6 dipeptidyl-anthraquinone: substrate molar ratio. In order to establish the binding equilibrium in solution, samples were incubated for 15 min at room temperature before the analysis. Control experiments were performed on 1 μM solutions of TAR or cTAR in 150 mM ammonium acetate, and 10 μM solutions of each compound in 150 mM ammonium acetate.

The **4a**, **4g**, **4h** and **5a**, **5g**, **5h** compounds were selected as representatives of the 2,6-AQ-Gly-X and of the 2,6-AQ-L-Ala-X series, respectively. All samples were analyzed in negative ion mode by direct infusion electrospray ionization (ESI) on a Thermo Fisher Scientific (West Palm Beach, CA) LTQ-Orbitrap Velos mass spectrometer. The analyses were performed in nanoflow ESI mode by using quartz emitters produced in-house by a Sutter Instruments Co. (Novato, CA) P2000 laser pipette puller. Up to 6 μL samples were loaded onto each emitter by using a gel loader pipette tip. A stainless steel wire was inserted in the back-end of the emitter and used to supply an ionizing voltage ranged around 0.8-1.0 kV. Source temperature and desolvation conditions were adjusted to decrease the incidence of salt adducts, with typical source temperature of 200 °C. Data were processed by using Xcalibur 2.1 software (Thermo Scientific).

Chapter 5

Results and Discussion

The new series of compounds were tested at the Department of Pharmaceutical Sciences of the Padua University, by the research group of Prof. Barbara Gatto, in order to verify their ability to inhibit the formation of the Tat-TAR complex and the chaperonic activity of NC protein, included its role of mediator in the annealing reaction of TAR with cTAR.

Compounds were evaluated by different assays:

- *High Throughput Screening* (HTS);
- *Fluorescence Quenching Assay* (FQA);
- NAME Assay (*Nucleocapsid Annealing-Mediated Electrophoreis*).

Moreover, in order to extend the spectrum of information related to the interaction between the synthesized compounds and nucleic acids of interest, compounds were evaluated in ESI-MS studies performed at the RNA Institute-University of New York, under the supervision of Prof. Dan Fabris.

For compounds belonging to 2,6-AQ-D-Ala-X series, as well as for the derivatives characterized by the presence of a tyrosine residue (compounds **4l**, **5l** and **6l**), we didn't receive to date biological results and therefore this PhD thesis reports the results related to the compounds of 2,6-AQ-Gly-X (compounds **4a**-**4i**) and 2.6-AQ-L-Ala-X(**5a**-**5i**).

5.1 2,6-Dipeptidyl-anthraquinones intercalate into TAR and cTAR dynamic structures (*High Throughput Screening*)

The research group to which I belong recently reported that the TAR and cTAR stabilization effect of the previously synthetized 2,6-AQ-βAla-X series of anthraquinonic derivatives leads to the inhibition of NC [43].

On this background, novel compounds were evaluated by High Throughput Screening studies, that have been performed to achieve preliminary identification of compounds able to stabilize the strands of TAR and cTAR.

Assuming a direct correlation between melting temperature and nucleic acid-binding activity, this setup allowed us to investigate the ability of compounds **4a**-**4i** and **5a**-**5i** to modify the thermal denaturing profile of initial TAR and cTAR. Reference *Tm* values were determined in the absence of ligand, which amounted to 69.3 °C for TAR, 53.8 °C for cTAR, and 69.4 °C for the hybrid. The melting profile of TAR, cTAR, and annealed heteroduplex were obtained to evaluate the shift of melting temperature (*ΔTm*) in the presence of increasing concentrations of anthraquinones. Sample solutions contained 1 μM of each nucleic acid substrate and 1, 10 or 100 μM of each ligand. Figure 12 and 13 summarize the values of *ΔTm* obtained respectively from the 2,6-AQ-Gly-X and 2,6-AQ-L-Ala-X series in the presence of a 10 μM ligand concentration, which displayed the most significant deviations.

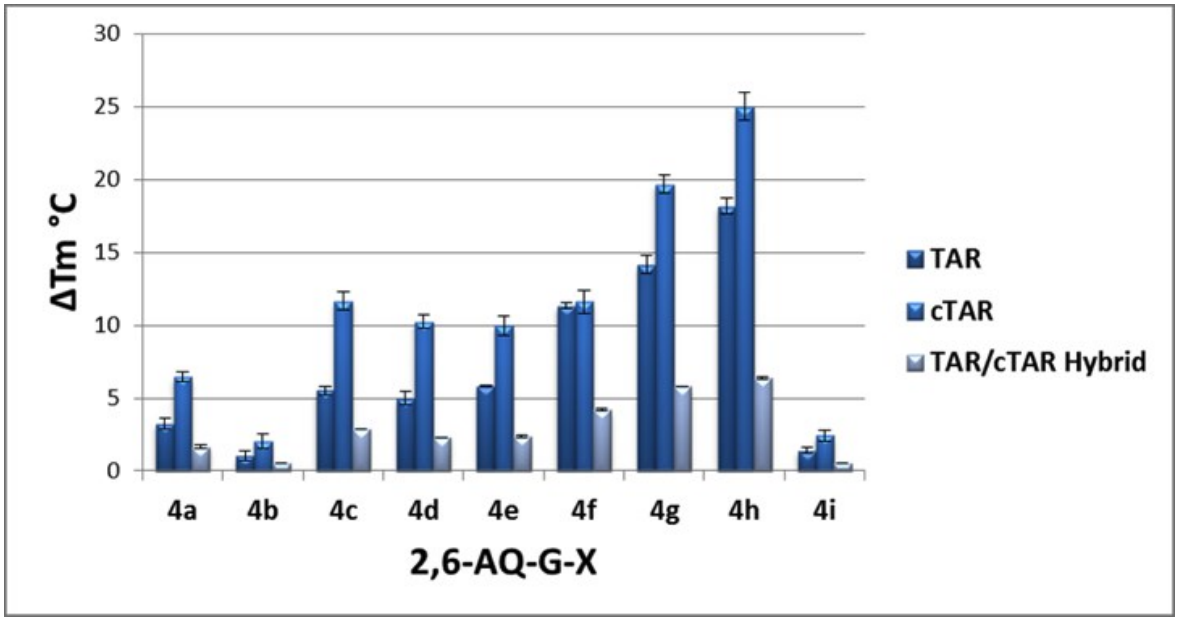


Figure 12: Variations of melting temperature (*ΔTm*) manifested by TAR, cTAR and TAR/cTAR hybrid in the presence of peptidyl-anthraquinones **4a**-**4i**. Reported values are the mean ± standard error of the mean (SEM) of triplicate experiments performed on samples containing 1 μM of oligonucleotide and 10 μM of ligand.

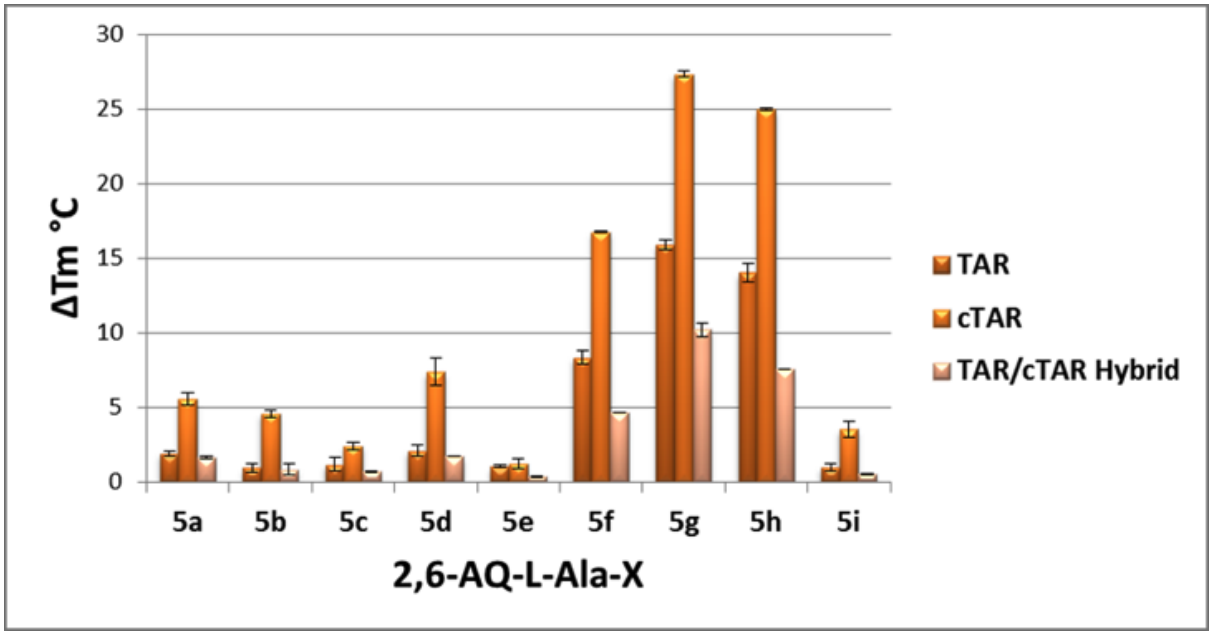


Figure 13: Variations of melting temperature (*ΔTm*) manifested by TAR, cTAR and TAR/cTAR hybrid in the presence of peptidyl-anthraquinones **5a**-**5i**. Reported values are the mean ± standard error of the mean (SEM) of triplicate experiments performed on samples containing 1 μM of oligonucleotide and 10 μM of ligand.

The results clearly revealed that derivatives of either series stabilized to different extent the structures of tested nucleic acids. In the case of the 2,6-AQ-Gly-X series, the length of the alkyl chain with two protonable amino groups correlated with the increased melting temperature according to the **4e** < **4f** < **4g** < **4h** ranking. The increases displayed by the TAR and cTAR constructs were particularly significant in the presence of **4g** and **4h**, whereas weaker stabilization was achieved by **4a**, **4c**, **4d**, **4e** and **4f**. In contrast, **4b** and **4i** produced no detectable effect on *Tm*, thus suggesting that their putative binding may not induce structure stabilization. Overall, this series induced slightly lower *Tm* shifts on TAR than cTAR, while those induced on the TAR/cTAR duplex were significantly lower, thus suggesting that the hybrid-stabilizing effects were less prominent.

The greater stabilization of TAR and cTAR is consistent with the hypothesis that disubstituted anthraquinones selectively recognize dynamic structures of nucleic acids, such as the *bulge-loop* features of these constructs, as expected from threading intercalators [49-51].

Compounds **5a-5i** (2,6-AQ-L-Ala-X) were analyzed in the same manner to investigate whether the introduction of a methyl group in side chains would affect nucleic acids recognition and stabilization.

In this case, **5f**, **5g** and **5h** were identified as strong TAR and cTAR stabilizers, while the other analogues did not produce significant effects. The **5e** compound displayed no detectable activity. In analogy with glycine supporting derivatives, the

L-alanine compounds showed greater stabilization as the length of the side chains increased, with the following rank: **5e** < **5f** < **5g**. However, this trend was not present when proceeding from the ornithine (i.e., **5g**) to the lysine anchor (i.e., **5h**), since the latter introduced lower stabilization than the former.

5.2 2,6-Dipeptidyl-anthraquinones inhibit NC-melting of TAR and cTAR structure (*Florescence Quenching Assay*)

Because the two series of derivatives stabilize the structures of nucleic acids, which represent the substrates of NC protein during the transcription, we analyzed their anti-NC properties *in vitro* through an assay based on fluorescence, consisting of a modification of the FRET protocol.

In particular, Florescence Quenching Assay is a FRET-based competition assay performed using the Tat-derived peptide labeled with the donor "fluorescein" at its N-terminus and the 29 nt TAR labeled at its 3' end with the dark quencher "dabcyl".

This assay is based on the possibility that protein NC, inducing the fusion of the lower half of the TAR strand, increases the distance between the fluorophore and quencher, leading to an increase in fluorescence. So, every inhibition of destabilization activities of NC protein, related to interaction with the ligand, can be directly assessed by fluorescence quenching.

After preliminary studies, designed to verify the absence of direct quenching of anthraquinonic compounds, it was possible to assess their inhibitory activity in the presence of the full length of recombinant NC protein of HIV-1 virus.

Tables 3 and 4 report the concentrations that produced half maximal inhibition of TAR and cTAR destabilization mediated by NC, which were observed for **4a**−**4i** and **5a**−**5i**, respectively.

Table 1: IC_{50} values observed for the **4a-4i** inhibition of NC-mediated melting of TAR and cTAR.

Compound	IC_{50} TAR (µM)	IC_{50} cTAR (µM)
4a	10.71±1.6	10.67±0.12
4b	11.33±0.34	9.75±1.72
4c	7.79±0.80	5.13±1.21
4d	6.42±0.05	4.06±1.49
4e	4.22±0.35	3.81±1.34
4f	3.87±0.18	4.61±0.17
4g	2.62±0.32	1.81±0.23
4h	2.12±0.16	1.11±0.19
4i	7.72±1.12	7.78±0.86

Table 2: IC_{50} values observed for the **5a-5i** inhibition of NC-mediated melting of TAR and cTAR.

Compound	IC_{50} TAR (µM)	IC_{50} cTAR (µM)
5a	20.80±6.75	11.97±4.27
5b	14.80±3.83	14.64±2.20
5c	40.25±1.98	14.59±0.42
5d	39.50±2.08	12.71±1.68
5e	67.12±1.61	44,15±1.43
5f	6.90±0.51	4.28±0.27
5g	4.91±0.48	1.93±0.50
5h	6.61±0.51	2.68±0.47
5i	7.29±1.14	7.53±0.85

The results obtained from the 2,6-AQ-Gly-X series evidenced that, with the sole exception of **4a** and **4b**, all derivatives with glycine spacers were potent inhibitors of the NC-mediated melting, which exhibited IC_{50} values lower than 10 μM. In particular, **4g** and **4h** were the most active inhibitors of NC, with IC_{50}s of 2.62 μM and 2.12 μM with TAR, and 1.81 μM and 1.11 μM with cTAR.

These results corroborated the strong stabilization of these constructs revealed by the HTS experiments. Consistent with the HTS results, the length of the doubly charged side chains correlated well with the potency of inhibition with both TAR and cTAR, according to the **4h** > **4g** > **4f** > **4e** rank.

Although members of the 2,6-AQ-L-Ala-X series were also found capable of inhibiting NC-mediated melting, they displayed weaker activity than the corresponding 2,6-AQ-Gly-X counterparts. In this case, the ornithine (**5g**) and lysine (**5h**) derivatives were among the most active inhibitors in the series, either when TAR or cTAR were assayed, but their activities were lower than those of the corresponding **4g** and **4h** analogues.

These results were unexpected. Since **5a-5i** compounds showed a greater ability in binding and stabilizing TAR and cTAR construct compared to their glycine-spacer analogues, we expected a greater NC inhibition. This expectation was not verified and this apparent discrepancy could be rationalized on the basis of the location of the putative binding sites for the threading intercalators onto the nucleic acid constructs.

Our experimental results suggest that ligand binding did not necessarily compete with NC interactions with the nucleic acid constructs. Hence, the introduction of a methyl group, which changed side chains maintaining the same length, leads to lower NC-mediating TAR and cTAR melting inhibition.

5.3 2,6-Dipeptidyl-anthraquinones inhibit NC-mediated TAR/cTAR annealing (*Nucleocapsid Annealing-Mediated Elettrophoresis*)

NC possesses more complex activities in addition to the destabilization of nucleic acid structures: NC promotes the annealing of nucleic acid strands. While the destabilization phase was appropriately analyzed by using the FRET-based platform, the strand re-annealing process was evaluated by using our recently developed Nucleocapsid Annealing-Mediated Electrophoresis (NAME) assay [52].

This type of experiment can be used to identify molecules capable of inhibiting the NC-mediated melting and annealing of viral nucleic acids folded into stable 3D conformations [53].

Oligonucleotides without the need of their previous labeling are used to perform this electrophoretic assay, avoiding unspecific compounds-fluorophore interactions.

The NAME protocol allows the analysis of the TAR/cTAR annealing reaction in the presence of NC and was used here to evaluate the impairment of the TAR/cTAR hybrid formation by compounds that strongly stabilize TAR and cTAR constructs.

The NAME assay was performed in two different modes. The first relied on preincubation of compounds with the full-length NC for 15 min, followed by the addition of the folded oligonucleotides for other 15 min (NC-preincubation mode).

The second involved preincubating the compounds for 15 min with each folded nucleic acid substrate, followed by addition of NC and further 15 min incubation (Oligo-preincubation mode).

In this way, the overall incubation time remained constant in either mode, but alternative preincubation with protein or folded constructs gave us the opportunity to detect possible differences in the mechanisms of action of related compounds [54].

Compounds **4a**-**4i** and **5a**-**5i** were analyzed in both modes. Representative data obtained by performing NAME assays in the presence of **4h** and **5g** are reported in Figures 14 and 15, respectively.

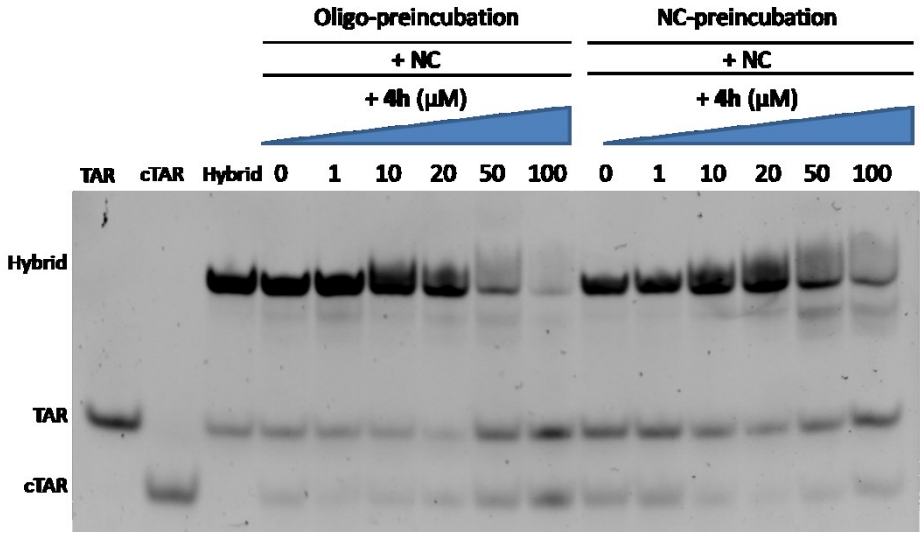


Figure 14: Inhibition of the TAR/cTAR annealing reaction by **4h** derivative of the 2,6-AQ-Gly-X series. NAME assay was completed in Oligo-preincubation mode on the left and in NC-preincubation mode on the right as above described. TAR, cTAR and the hybrid TAR/cTAR formed by thermal denaturation were used as control. Increasing concentrations (0, 1, 5, 10, 20, 50, 100 µM final) of **4h** were used to inhibit the activity of the full-length recombinant NC protein.

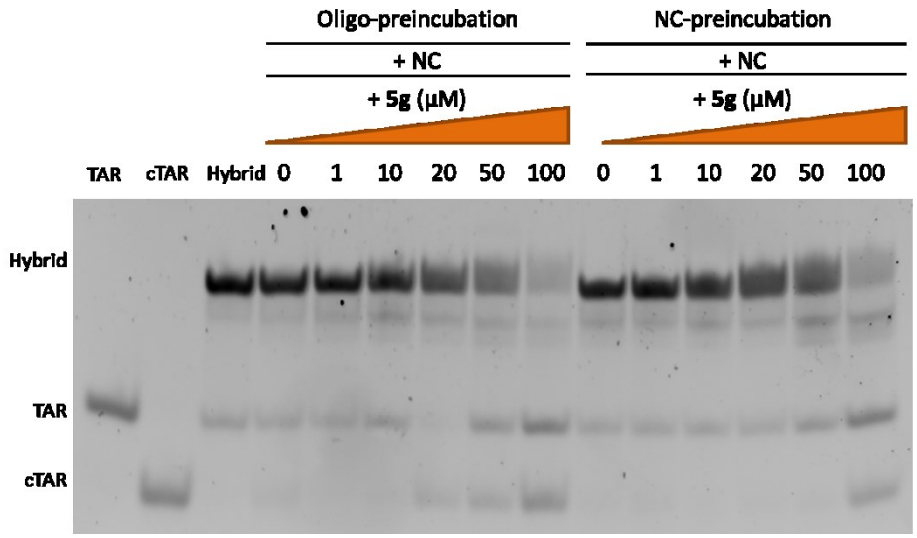


Figure 15: Inhibition of the TAR/cTAR annealing reaction by **5g** derivative of the 2,6-AQ-L-Ala-X series. NAME assay was completed in Oligo-preincubation mode on the left and in NC-preincubation mode on the right as above described. TAR, cTAR and the hybrid TAR/cTAR formed by thermal denaturation were used as control. Increasing concentrations (0, 1, 5, 10, 20, 50, 100 µM final) of **5g** were used to inhibit the activity of the full-length recombinant NC protein.

This first analysis by NAME allowed us to obtain indications on the putative mechanism of action of the identified NC inhibitors. Comparing the results obtained in the two different NAME assay modes, it is evident that the Oligo-preincubation favored the inhibitory activity of anthraquinones, thus suggesting that their mechanism of action relied on the direct interaction with the nucleic acids substrates, rather than with NC. A thorough examination of these data allowed us to group together compounds that exhibited similar activity. In particular, **4g** and **4h** (Figure 16) resulted to be highly active in the 20-50 μM concentration range, while **4c**, **4d**, **4e**, **4f**, **5g** and **5h** exhibited moderate NC inhibition in the 50-100 μM range (Figure 17); all the other derivatives resulted to be inactive up to 100 μM concentration (data not shown).

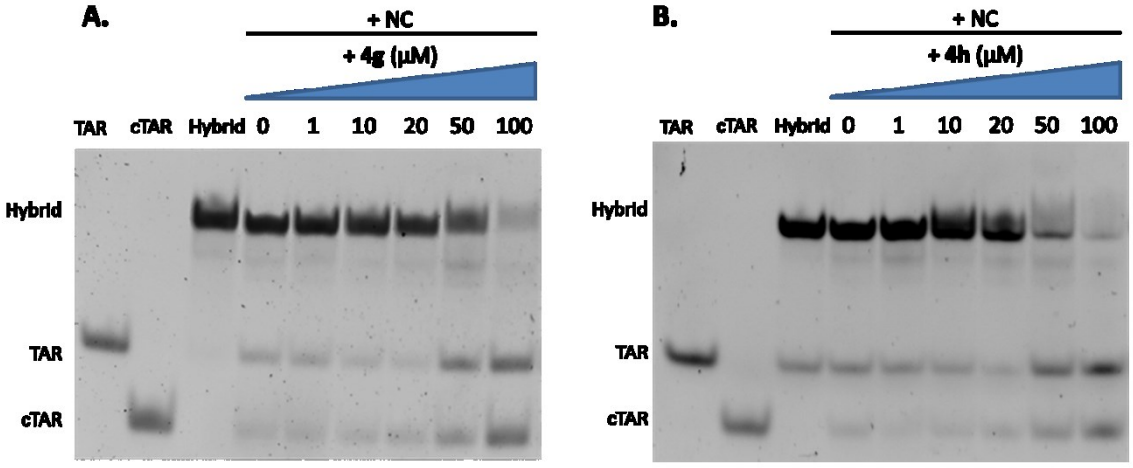


Figure 16: Inhibition of the TAR/cTAR annealing reaction by compounds exhibiting activity in the 20-50 μM concentration range**.** NAME assay performed in the Oligo-preincubation mode in the presence of **4g** (**A**) and **4h** (**B**). TAR, cTAR and the hybrid TAR/cTAR formed by thermal denaturation were used as control. Increasing concentrations (0, 1, 5, 10, 20, 50, 100 μM final) of anthraquinone were used to inhibit the activity of the full-length recombinant NC protein.

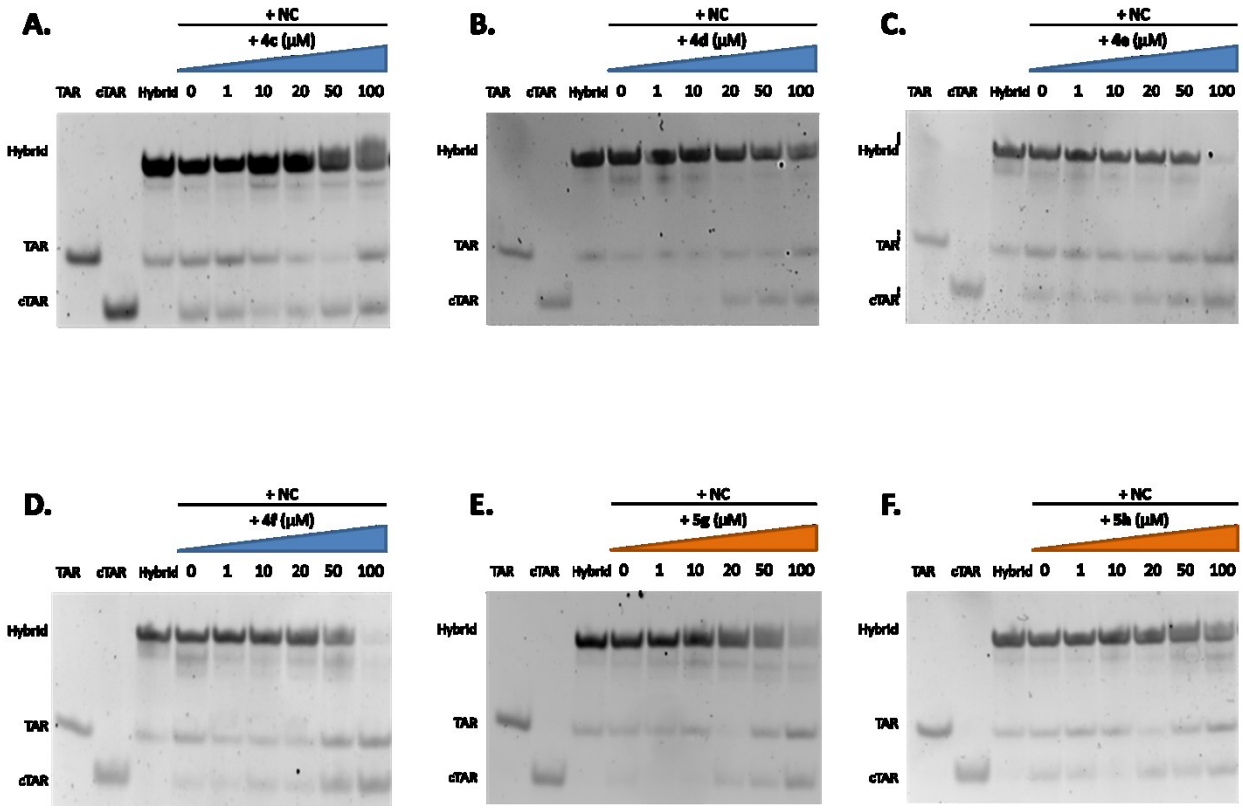


Figure 17: Inhibition of the TAR/cTAR annealing reaction by compounds exhibiting activity in the 50-100 µM concentration range**.** NAME assay performed in the Oligo-preincubation mode in the presence of **4c** (**A**), **4d** (**B**), **4e** (**C**), **4f** (**D**), **5g** (**E**) and **5h** (**F**). TAR, cTAR and the hybrid TAR/cTAR formed by thermal denaturation were used as control. Increasing concentrations (0, 1, 5, 10, 20, 50, 100 µM final) of anthraquinone were used to inhibit the activity of the full-length recombinant NC protein.

We therefore focused on active compounds and the more active ones were further investigated in the Oligo-preincubation mode to evaluate the influence of concentration.

In order to determine accurately the IC_{50} values afforded by the various anthraquinones, two different sets of concentrations were selected for highly active compounds (0, 1, 10, 20, 30, 40, 50, 75, 100 µM final) or moderately active compounds (0, 1, 10, 25, 50, 60, 70, 80, 100 µM final).

Among the inactive compounds identified by the preliminary screening, **4i** and **5f** were employed in these experiments as negative controls.

Results from three independent analyses were assessed to calculate the IC_{50}, as a measure of the ability of these compounds to slow down the annealing activity of NC under the selected experimental conditions (Table 3).

Table 3: IC_{50} values observed for the inhibition of NC-mediated annealing of TAR and cTAR (NAME assay).

Compound	IC_{50} NAME (μM)	Compound	IC_{50} NAME (μM)
4a	ND	**5a**	ND
4b	ND	**5b**	ND
4c	>100	**5c**	ND
4d	>100	**5d**	ND
4e	102.7±7.5	**5e**	ND
4f	81.4±2.3	**5f**	>100
4g	54.1±12.5	**5g**	49.1±15.9
4h	23.5±9.8	**5h**	80.4±8.6
4i	>100	**5i**	ND

The results provided by these NC-mediated annealing experiments proved that tested dipeptidyl-anthraquinones are annealing inhibitors. Interestingly, these results matched those achieved from the NC-mediated melting assays and the relative potencies of annealing inhibition follow exactly the same order obtained in the melting inhibition.

Indeed, as the length of the doubly-charged side chains increased in the 2,6-AQ-Gly-X series, the efficacy of inhibition increased according to the **4h** < **4g** < **4f** < **4e** rank. In contrast, this trend was verified in the 2,6-AQ-L-Ala-X series only for **5e** < **5f** < **5g**, whereas the NC-inhibition potency decreases again when proceeding from **5g** to **5h**. In addition, the NAME assays confirmed that 2,6-AQ-Gly-X anthraquinones were in general more potent than the 2,6-AQ-L-Ala-X derivatives, suggesting the importance of the spacer length related to the activities of these compounds, highlighted **4h** and **5g** as the strongest inhibitors within either series.

5.4 2,6-Dipeptidyl-anthraquinones bind to TAR and cTAR (ESI-MS)

Additional insights into the specific interactions of 2,6-dipeptidyl-anthraquinones with nucleic acids were obtained by evaluating the binding properties of selected compounds identified by the FQA determinations. With the goal of assessing the stoichiometry and overall affinity of such compounds for TAR and cTAR, we analyzed ligand-substrate mixtures by electrospray ionization mass spectrometry (ESI-MS) under conditions that enable the detection of intact non-covalent complexes of oligonucleotides [55-59].

All the samples were analyzed in negative ionization mode by direct infusion.

Preliminary control determinations were performed on TAR and cTAR samples in negative ion mode, which afforded experimental masses of 9286.2 and 8884.5 u,

respectively. These values matched very closely the monoisotopic values calculated from the respective sequences (i.e., 9286.3 Da and 8884.5 Da).

Increasing amounts of ligand were added to a 1 μM solution of either construct, in such a way as to obtain samples containing up to a 10:1 molar ratio of ligand versus substrate.

Based on the strong stabilizing activity revealed by the FQA experiments, the **4g**, **4h** and **5g**, **5h** compounds were selected as representatives of the 2,6-AQ-Gly-X and 2,6-AQ-L-Ala-X series, respectively. In addition, **4a** and **5a** were also included as possible negative controls, owing to their weak stabilization effects.

The representative data reported in Figure 18 show only the regions containing the 7- and 6- charge states of TAR and cTAR in the presence of **4h** and **5g**. The spectra displayed signals corresponding to free unbound constructs, as well as those of stable complexes with different stoichiometries.

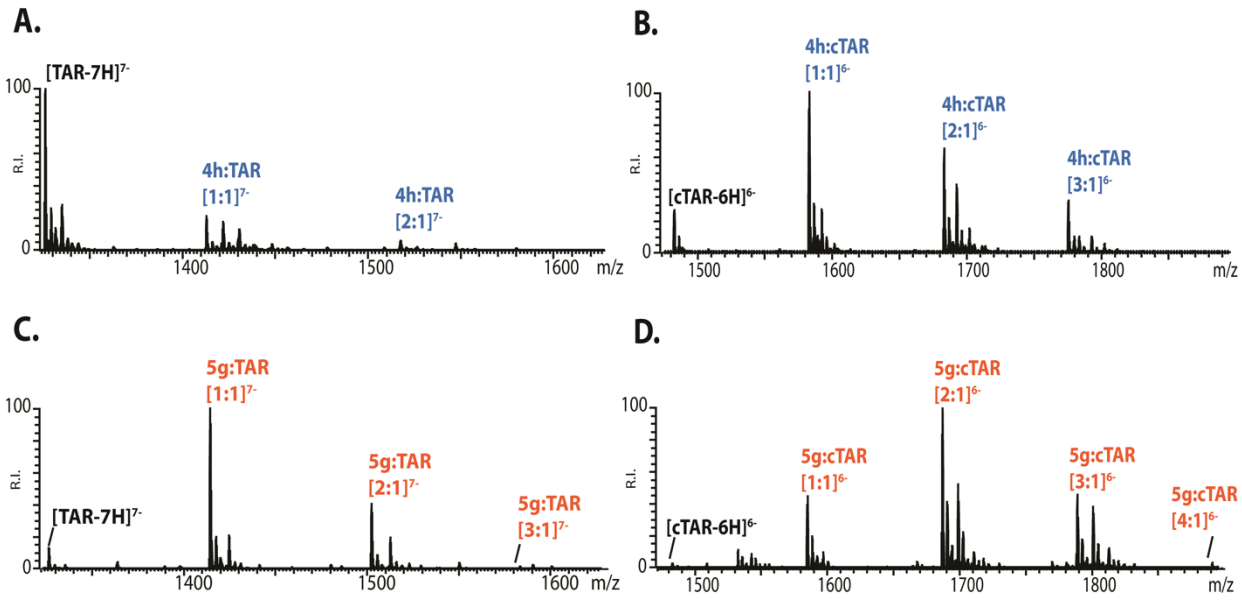


Figure 18: Representative ESI-MS spectra of samples obtained by mixing 1 μM of either TAR or cTAR with 10 μM of 2,6 dipeptidyl-anthraquinone in 150 mM ammonium acetate. Panel **A)** was obtained from a sample containing **4h** and TAR; **B) 4h** and cTAR; **C) 5g** and TAR; and **D) 5g** and cTAR. The stoichiometries of the observed complexes are indicated in each spectrum. Signals of lower intensity detected near free/bound species consist of typical sodium and ammonium adducts.

These results confirmed the ability of ESI-MS to preserve the non-covalent interactions established by peptidyl-anthraquinonic derivatives with dynamic nucleic acid structures.

Experimental and calculated masses for all the species detected in these experiments are provided in Table 4 and 5.

Table 4: Monoisotopic mass (Da) of 2,6 dipeptidyl-anthraquinone: TAR complexes by ESI-MS analysis.

Compound	1:1 compound:TAR		2:1 compound:TAR		3:1 compound:TAR		4:1 compound:TAR	
	Exp mass (Da)	Calc Mass (Da)	Exp mass (Da)	Calc Mass (Da)	Exp mass (Da)	Calc Mass (Da)	Exp mass (Da)	Calc Mass (Da)
4a	9860.528	9860.499	-	-	-	-	-	-
4g	9866.548	9866.520	10446.813	10446.796	11027.078	11027.072	-	-
4h	9894.583	9894.552	10502.625	10502.86	-	-	-	-
5a	9888.546	9888.528	-	-	-	-	-	-
5g	9894.562	9894.551	10502.876	10502.858	11111.169	11111.165	-	-
5h	9922.618	9922.583	10558.946	10558.922	11195.239	11195.261	-	-

Table 5: Monoisotopic mass (Da) of 2,6 dipeptidyl-anthraquinone: cTAR complexes by ESI-MS analysis.

Compound	1:1 compound:cTAR		2:1 compound:cTAR		3:1 compound:cTAR		4:1 compound:cTAR	
	Exp mass (Da)	Calc Mass (Da)	Exp mass (Da)	Calc Mass (Da)	Exp mass (Da)	Calc Mass (Da)	Exp mass (Da)	Calc Mass (Da)
4a	9458.765	9458.755	10033.67	10033.01	-	-	-	-
4g	9464.789	9464.776	10045.052	10045.050	-	-	-	-
4h	9492.824	9492.807	10101.124	10101.114	10709.429	10709.421	-	-
5a	9486.803	9486.784	10089.09	10089.07	-	-	-	-
5g	9492.809	9492.807	10101.139	10101.114	10709.399	10709.421	11317.689	11317.728
5h	9520.849	9520.838	10157.169	10157.176	10793.504	10793.514	11429.844	11429.852

The experimental mass is obtained considering the monoisotopic peak of each complex; the theoretical calculated mass is obtained from the oligonucleotide sequence and the elemental composition of the ligand. The experimental masses are calculated from the negative ion mode spectra obtained for 10:1 2,6 dipeptidyl-anthraquinone: TAR or 2,6 dipeptidyl-anthraquinone: cTAR molar ratio in 150 mM ammonium acetate.

The determination of free and bound RNA or DNA, necessary to evaluate the binding affinity of compounds to TAR and cTAR, was directly assessed from the intensity of the respective peaks divided by the charge state. Taking into consideration all the charge states observed for each species offsets the effects of a

shift of charge state distribution, which manifests itself as compensatory changes in the relative signal intensities of successive charge states. Normalizing the signal intensity by the respective charge state constitutes a simple way to reduce the possible bias due to the fact that different charge states induce different image currents during ion detection in ESI-MS. Free and complexed TAR or cTAR abundances were finally expressed as percentage and compared.

The results obtained by ESI-MS clearly revealed that the selected 2,6-dipeptidyl-anthraquinones could bind either substrate with different stoichiometries, thus suggesting the presence of multiple binding sites on the selected constructs. In particular, the coexistence of free TAR or cTAR with complexes of different stoichiometries indicated that binding to the second site was initiated before saturation of the first was complete, which is consistent with independent sites sharing similar affinities and negligible cooperativity. Similar binding modes were detected for all the TAR and cTAR complexes observed for the analyzed compounds.

The partitioning between free and bound species in each sample served not only to understand the putative binding modes supported by the constructs, but also to estimate their overall affinity towards the different ligands.

The percentage of complexed over total substrate in each sample was calculated to obtain relative scales of binding affinities for TAR and cTAR, respectively: **4g** ≈ **5g** ≈ **5h** > **4h** >> **4a** ≈ **5a** and **5g** ≈ **5h** > **4h** > **4g** >> **4a** ≈ **5a**. In addition, the fact that the percentage of complexed substrate was consistently lower in the TAR than in the cTAR experiments at the same ligand concentrations indicated that all compounds

possessed greater affinities for the DNA than for the RNA construct, in accordance with HTS results (Figure 12 and 13).

These experiments showed also that stronger ligands produced greater binding stoichiometries. The strongest TAR binders **4g**, **5g** and **5h** displayed up to 3:1 ratios, whereas only 1:1 complexes were detected for **4a** and **5a** at the same ligand concentrations (Figure 19).

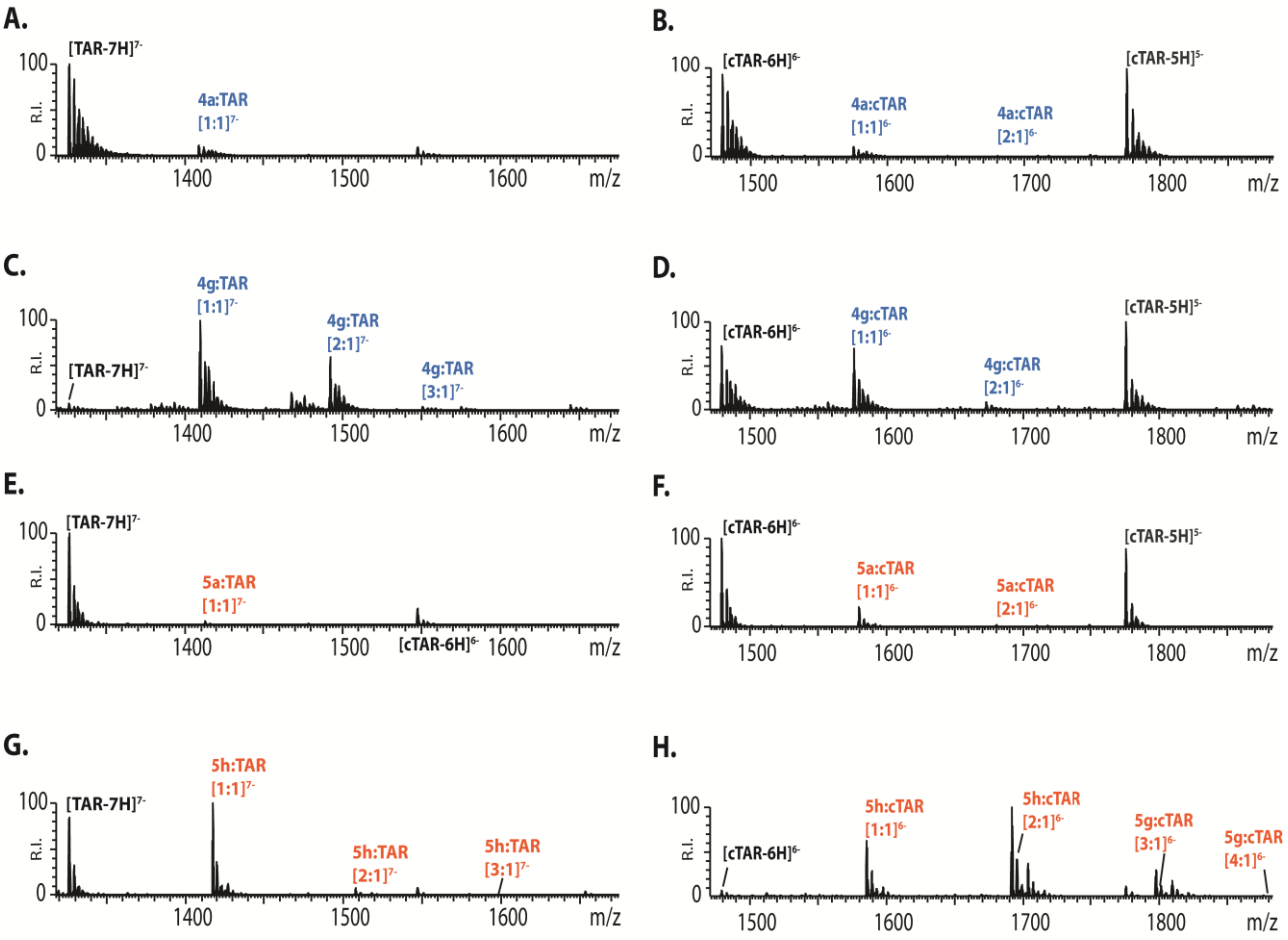


Figure 19: ESI-MS spectra of samples obtained by mixing either TAR or cTAR with 2,6 dipeptidyl-anthraquinones**.** Panel **A)** was obtained from a sample containing **4a** and TAR; **B) 4a** and cTAR; **C) 4g** and TAR; **D) 4g** and cTAR; **E**) **5a** and TAR; **F**) **5a** and cTAR; **G**) **5h** and TAR; **H**) **5h** and cTAR. The stoichiometry of bound species is reported in each spectrum. Lower intensity signals near free/bound species consist of typical sodium and ammonium adducts.

In the case of cTAR construct, the strongest binders **5g** and **5h** displayed up to 4:1 complexes, whereas complexes up to only 2:1 were detected for the weaker **4a** and **5a** binders (Figure 19). Analyzing in details the data it is evident how binding stoichiometries and affinities observed by ESI-MS experiments possessed excellent correlation with the stabilization of TAR and cTAR folding detected by *Fluorescence Quencing Assay* (FQA).

Chapter 6

Conclusion

The results obtained during my PhD at the Deparment of Pharmacy of the University of Naples Federico II have led to the identification of new promising classes of NC inhibitors. Having the goal of finding NC inhibitors, we identified novel 2,6-diaminoanthraquinonic derivatives able to bind and stabilize preferentially dynamic regions of nucleic acids like TAR and cTAR over annealed duplexes [60].

Based on precise structural requirements placing charged substituents at the 2,6-positions of the anthraquinones ring, the presence of two charged residues in each flexible side chain at the distance of three to four methylene groups, we have discovered potent *in vitro* inhibitors of NC, as demonstrated through all the *in vitro* studies with full length NC.

The new derivatives differ from previously reported ones by the length and identity of the linker between the anthraquinone ring and the charged groups on the side chains. The compounds were obtained by following an excellent synthetic scheme that provided high yields and purity.

Comparing the structures of the two series with different linkers, we can conclude that the presence of an Orn residue in the side chain grants very favorable features to enhance the inhibitory activity of both anthraquinone series tested.

This conclusion is supported by the fact that **4g** and **5g** exhibited high activity in both HTS and NAME assays. Consistent with this, the 2,6-AQ-β-Ala-Orn compound was also the most active inhibitor in the series described previously [43]. It is important to note, however, that the rank of potencies in the new series 2,6-AQ-Gly-

X was **4h** > **4g** > **4f** > **4e**, which makes **4h,** with a Lys as terminal, the lead compound in this series. This analogue was highly active in all assays at concentrations comparable to those observed for the 2,6-AQ-β-Ala-Orn derivative [43].

The fact that the longer Lysine, bearing an extra methylene group compared to ornitine in the side chain, is better suited than Orn to confer inhibitory properties when placed in the 2,6-AQ-Gly-X series suggested that the overall length of the peptidic substituent played a determinant role.

This possibility was supported by examining the distance between the terminal amino groups, which are chargeable, and the anthraquinonic scaffold in compounds with either β-Ala or Gly linkers. As shown in Figure 20 this examination showed that the lengths of the entire peptidic portions of 2,6-AQ-β-Ala-Orn and 2,6-AQ-Gly-Lys (**4h**) were quite similar.

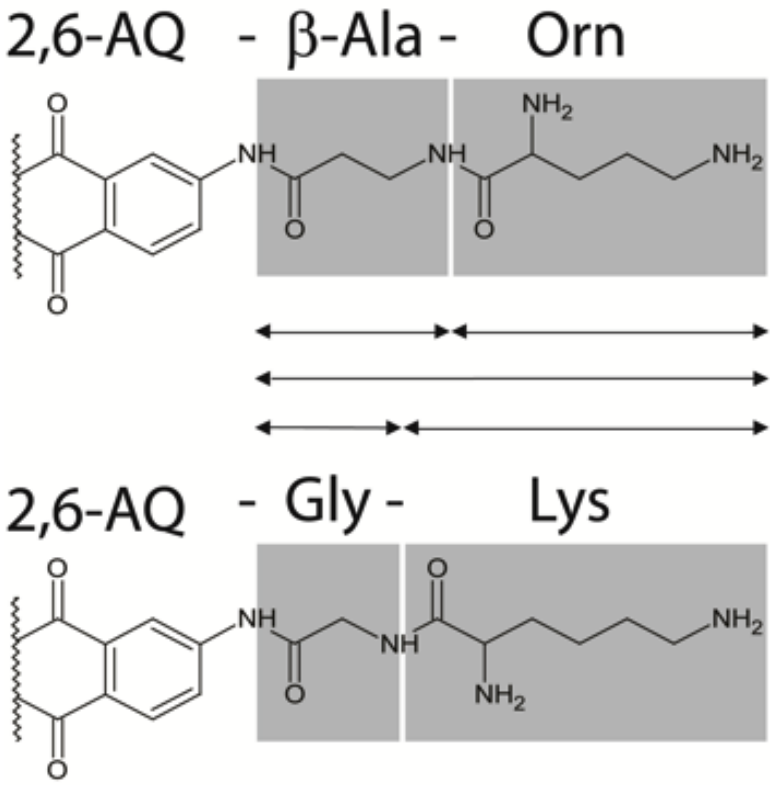


Figure 20: Details of the structures of 2,6-AQ-β-Ala-Orn and **4h** (2,6-AQ- Gly-Lys), which enable a direct comparison of the lengths of the respective side chains.

This observation suggests that proper positioning and orientation of the terminal protonated side chains within the anthraquinone-nucleic acid complex must play a fundamental role in the inhibition of NC's chaperone activity.

In summary, we demonstrated that:

- the presence of two charged residues in each flexible side chain, with mutual distance of three to four methylene groups, is necessary for the development of potent NC inhibitors;

- Gly linker increases the activity of the 2,6-disubsituted anthraquinones, while the presence of an additional methyl group in the L-Ala linker decreases activity;
- in addition to the length, also the orientation of the positively charged side chains is important for NC inhibition.

In this direction, additional experiments are underway to explore the possible effects of D-Ala linkers (series 2,6-AQ-D-Ala-X compounds **6a-6l**) on the activity of these anthraquinones.

In conclusion, 2,6-dipeptidyl-anthraquinones are a promising class of nucleic acid-binding compounds that act as NC inhibitors *in vitro*. We designed, synthesized and tested new series of 2,6-disubstituted-anthraquinones, which are able to bind viral nucleic acid substrates of NC. We demonstrated that these novel derivatives interact preferentially with non-canonical structures of TAR and cTAR, stabilize their dynamics structures (such as bulge and loop), and interfere with NC chaperone activity.

The most notable compound consisted of **4h**, which acted as TAR and cTAR binder, as well as NC inhibitor with activity in the micromolar range. For this reasons, **4h** can be considered as a "lead compound, with a IC_{50} value of 2.12 μm.

Abstract Unit 2

IGFBP6 is an innovative interesting target: it's a protein involved in the pathogenesis of cancer (breast, prostate, colon, neck and head of the pancreas, ovarian cancer and neuroblastoma). It's chemical features are still not completely disclosed, because its crystallographic structure is not available, but several informations are accessible from NMR studies: it is composed by three structural domains, and either N-terminal residue or C-terminal residue are rich in cysteines.

Recent *in vivo* studies performed in transgenic mice have highlighted the role of IGFBP6 in the CNS: in particular, its pathogenic role in the demyelination process, certainly interesting, due to its correlation with important pathologies, such as multiple sclerosis.

IGFBP6 can be O-glycosilated and the glycosilation site in represented by C-terminal residue.

During my experience at University of Vienna I synthesized a peptide that is part of IGFBP6 ([126-153] fragment), suitably modified with Fmoc-Thr(GalNAc-Ac_3)-OH and then iodoacetylated. Moreover my attention was focused to preliminary tests aimed to link a photo-cleavable auxiliary to the peptide in order to obtain a glycosilated analogue of IGFBP6 (insulin-like growth factor-binding protein 6).

These studies allowed to check if the insertion of the auxiliary could interfere in the glycosylation due to the steric hindrance caused by the proximity of the glycosylation site and the N-terminal residue of the peptide.

UNIT 2: Synthesis of a PEGylated ligation auxiliary and its application in the preparation of homogeneously glycosylated linker domain of IGFBP6 protein

Chapter 1

Introduction

Insulin-like growth factors (IGFs), are key elements related to the physiological growth and are also implicated in several diseases including cancer.

There are two different IGFs: IGF-1 and IGF-2, whose activity is modulated by a family of high affinity proteins, called IGF Binding Proteins (IGFBP): in particular, IGFBP-6 is notable for its marked higher binding affinity for IGF-2 compared to IGF-1.

The main role of IGFBP6 protein consists in the carrier's function: indeed, they modulate the activity and half-life of IGFs, but recent studies have shown that this protein is also able to inhibit angiogenesis and to induce tumor cell migration (Figure 1) [60].

IGFBP6 owns peculiar structural features: is composed by three structural domains, and either N-terminal or C-terminal domains are rich in cysteines.

All IGFBPs are subject to a large number of post-translational modifications that affect their actions, mediated by a large number of protease [61].

IGFBP3 and IGFBP4 can be *N*-glycosylated, whereas IGFBP5 and IGFBP6 can be *O*-glycosylated. The glycosylation does not act directly influencing the actions of these proteins, but decreases their interaction with glycosaminoglycans, inhibiting their proteolytic process. More recently it has been highlighted that the IGFBP6 can also be phosphorylated, although the consequences of these changes are still unknown [62].

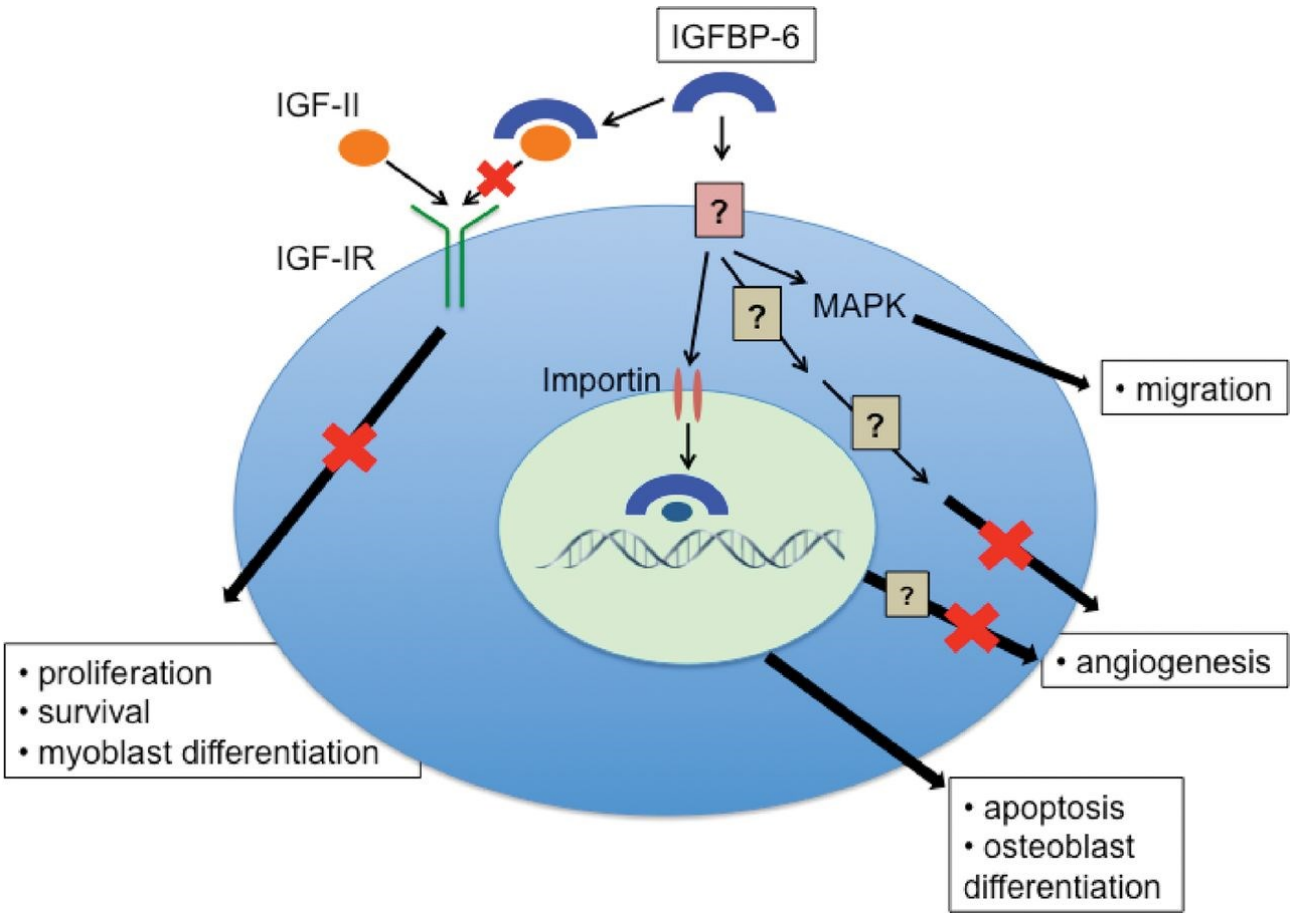


Figure 1: Insulin-like growth factor-binding protein-6 and cancer [62].

Through *in vivo* studies performed in transgenic mice, it has been also highlighted the role of IGFBP6 in the Central Nervous System (CNS): in particular its pathogenic role in the demyelination process, certainly interesting, due to its linkage with important pathologies, such as multiple sclerosis [62].

Moreover, IGFBP6 is involved in the pathogenesis of cancer: breast cancer, prostate, colon, neck and head, pancreas, ovarian and neuroblastoma are the diseases in which the role of this protein has been explained cleary [63].

Therefore its inhibition can be an interesting target for cancer therapy, above all as regards children or young adolescent patients, where a specific IGF-2 receptor inhibition appears to be advantageous to block the high IGF-2 autocrine without going to inhibit the physiological growth mediated by the presence of IGF-1 [63].

IGFBP6 is a O-glycosilated glycoprotein: its structural features derived only from NMR studies, in fact are not available yet crystallographic structure: also for this reason could be interesting to deepen the studies about this target.

Therefore it could be useful to have available analogues of this domain in homogeneous form to be able to study features and the role of glycosylation.

Chapter 2

Aim of the research

During my second year of PhD, I was awarded the STAR Program fellowship from “Compagnia di San Paolo” in collaboration with the University of Naples Federico II. In the frame of this program, I had the opportunity to spend six months working at Department of Biological Chemistry, in the University of Vienna, headed by Prof. Christian Becker.

One of the main interests of the research group where I carried out my internship is protein homogeneous, site-selective glycosylation. In particular, Dr. Claudia Bello, a postdoc researcher in Prof Becker’s group, developed an innovative method using a PEGylated photoclevable ligation auxiliary for the homogeneous, sequential glycosilations of Mucin 1 peptides [64].

In particular this work allows the utilization of a photocleavable auxiliary that leads easily to attach the PEG polymer to obtain efficient enzymatic site specifically O-glycosylated mucin 1 (MUC1) polypeptides and also it is easily removed by UV radiation at 365nm in few minutes.

From this encouraging scientific starting point, it was decided to apply this method to the synthesis and glycosylation of a peptide that is part of IGFBP6, our target of interest.

To this purpose, during the first period of my internship, I focused my attention to the synthesis of this auxiliary in higher yields in order to be able to apply this

strategy to a new target, the IGFBP6 protein [64]. This synthetic procedure proved to be very difficult either in purification and in the characterization steps.

I then synthesized successfully the [126-153] fragment of the linker domain of IGFBP6 and I carried out preliminary experiments for the attachment of the auxiliary to the peptide.

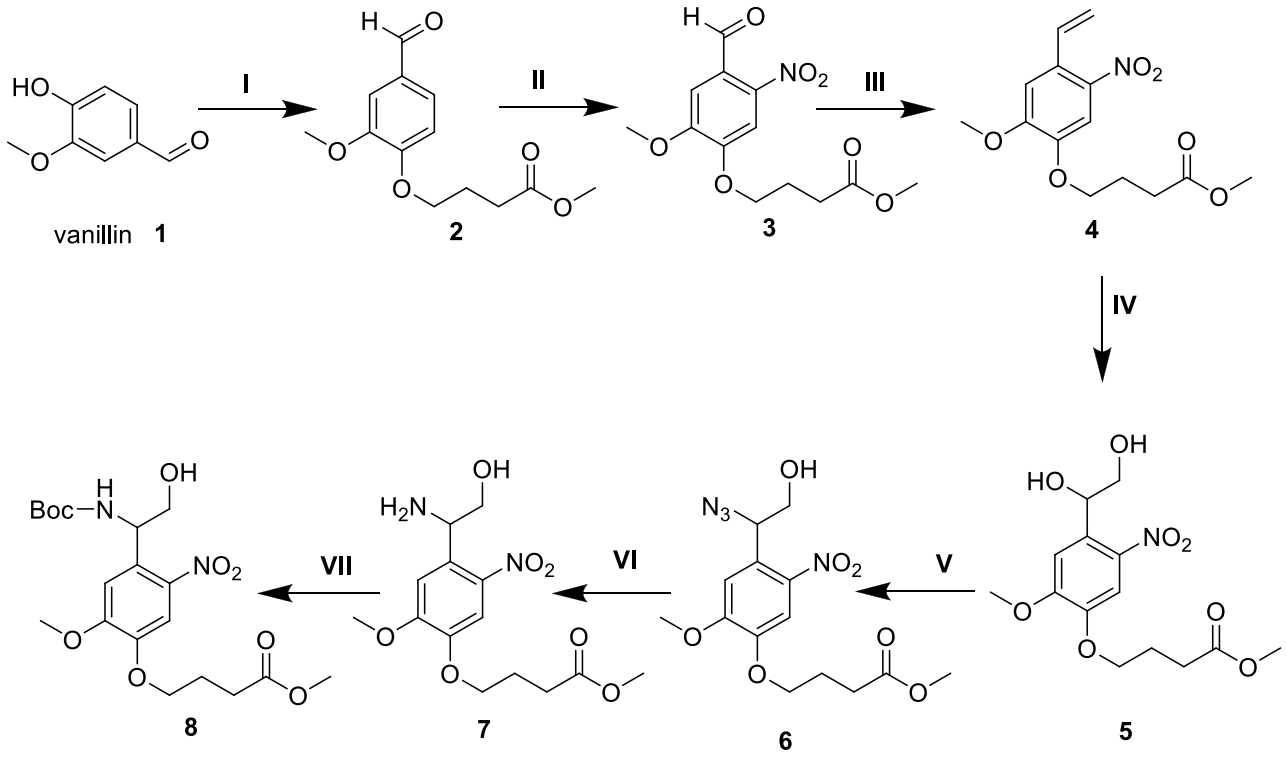
Because of the limited time available, it wasn't possible to perform the enzymatic glycosylation reaction. The project is currently been carried on in the Department of Biological Chemistry at the University of Vienna.

Chapter 3.

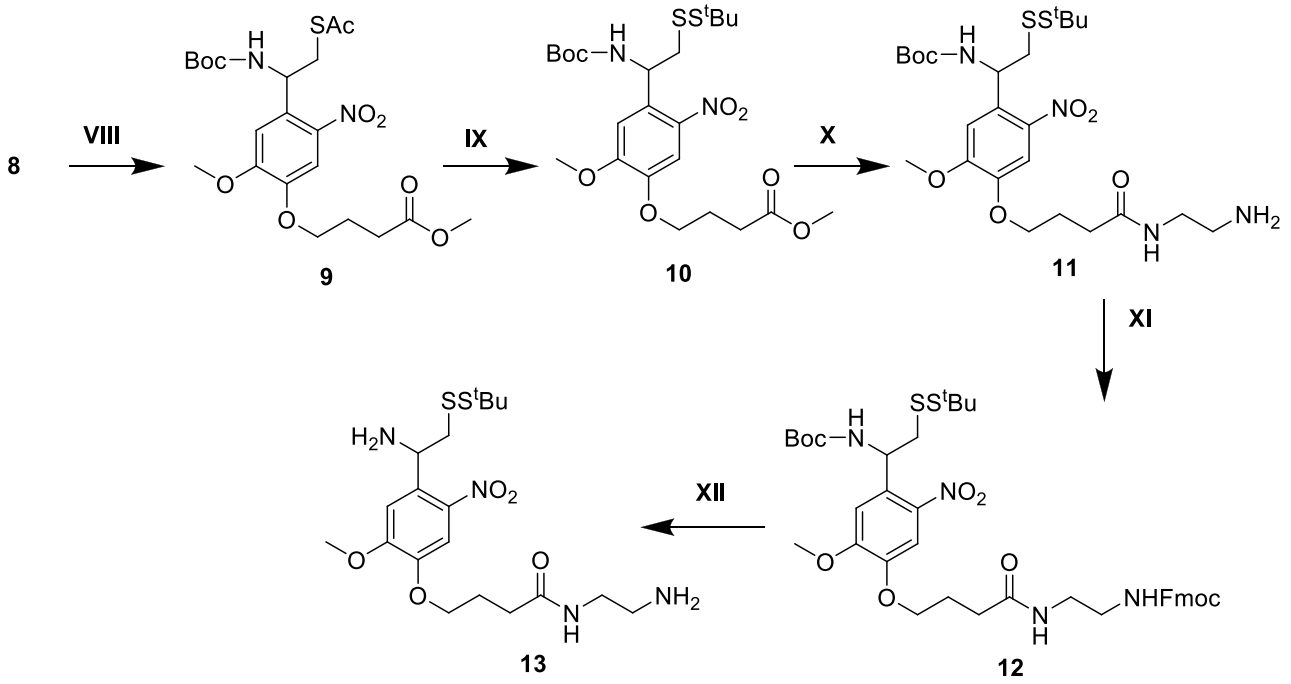
Results and discussion

3.1 Synthesis of the auxiliary

The synthesis of the auxiliary (compound **13**) is summarized in the following schemes (Scheme 1 and Scheme 2):



Scheme 1. Synthesis of intermediate **8**. Reagents and conditions: I) Methyl 4-chloro butanoate, K_2CO_3, Bu_4NI, CH_3CN, reflux, 80%. II) HNO_3, CH_3COOH, 0°C to rt, 93%. III) $MePPh_3Br$, TMS_2NNa, THF, 10°C to 0°C, 65%. IV) AD-mixα, *t*BuOH/CH_2Cl_2/H_2O, 90%. V) $SOCl_2$, NEt_3,CH_2Cl_2, 0°C. VI) NaN_3, DMF, 60°C. VII) PPh_3, H_2O, THF, 70°C.



Scheme 2. Conversion of intermediate **8** in the auxiliary (final compound **13**). Reagents and conditions: VIII) Boc_2O, DMAP, CH_3CN, 0°C, 63%. IX) PPh_3, DIAD, CH_3COSH, THF, 0°C to rt, 84%. X) MeONa, 1-(butylthio)-1,2-hydrazine carboxymorpholide, MeOH, rt, 77%. XI) 1,2-ethylenediamine, $Zr(OtBu)_4$, HOBt, CH_3CN, rt, 70%. XII) a. Fmoc-Cl, DIEA, CH_2Cl_2, 0°C to rt; b. TFA, CH_2Cl_2, 92% (two steps).

The synthesis is composed by twelve reaction steps: in particular 7 synthetic steps are referred to Scheme 1 and lead to the key intermediate **8**. The next five synthetic steps (Scheme 2) lead to the desired auxiliary (compound **13**).

All the synthetic details are reported in a previously published paper by Claudia Bello and Christian Becker [64].

All intermediates were purified by silica gel chromatography and were fully characterized by H^1-NMR, C^{13}-NMR and also by bidimensional experiments such as COSY; NOESY; HSCQ; HMBC to obtain a complete structural characterization.

The final derivative (**13**) was obtained in high reaction yield and in high purity, as it is clear evidenced by NMR spectra (H^1 and C^{13} NMR) represented in Figure 2.

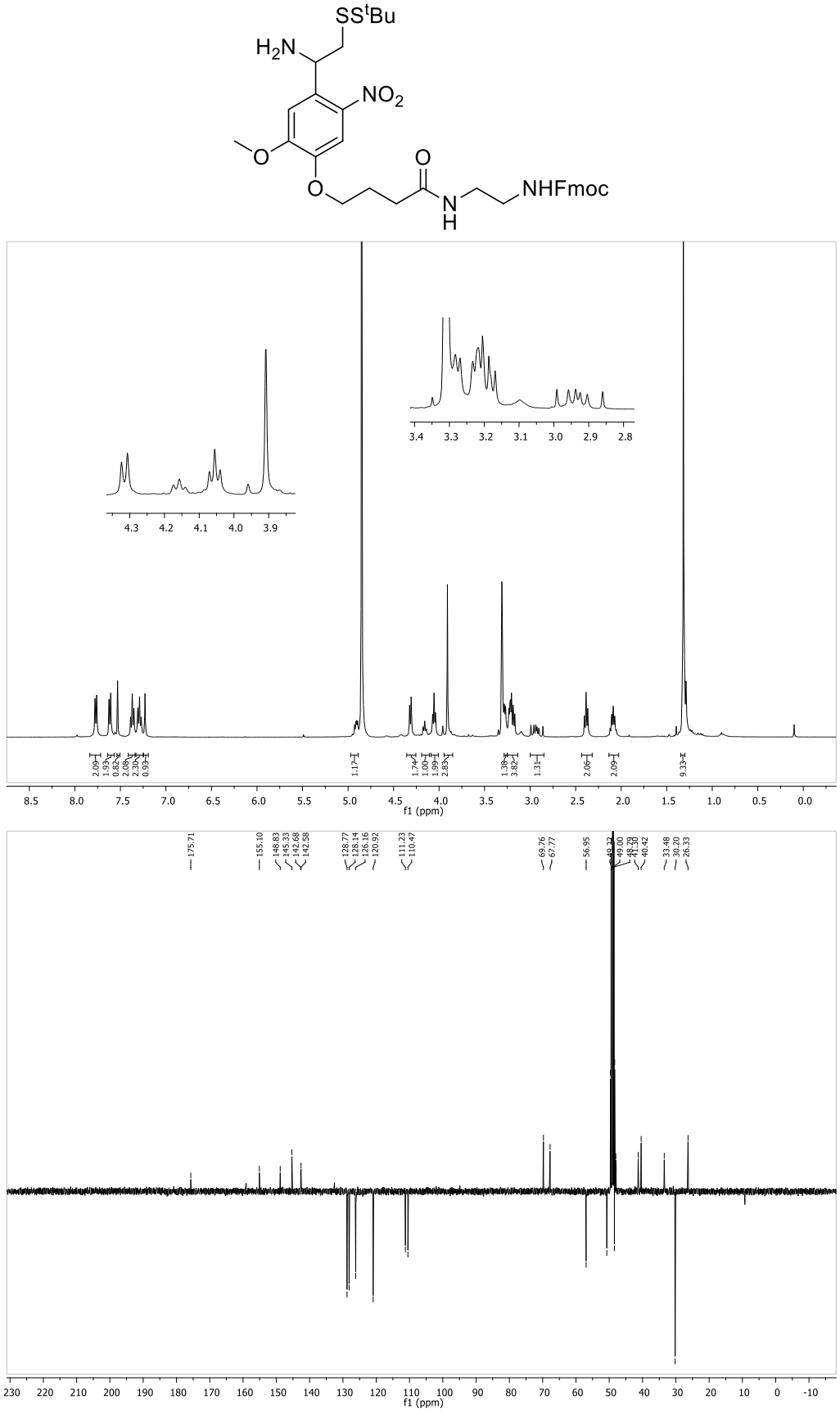


Figure 2: H^1 and C^{13} NMR spectra of auxiliary (final compound **13**).

3.2 Synthesis of IGFBP-6 [126-153]

The synthesis target, a 27 residues peptide part of IGFBP6, was synthesized on a solid support using the 9-fluorenylmethoxycarbonyl (Fmoc)–based Solid Phase Peptide Synthesis (SPPS) strategy. The preloaded Alanyl-TentaGel resin Wang type was used for the synthesis of the peptide on a 0.1 mmol scale; the synthesis was carried out manually and repeated using the Liberty Blue automated microwave peptide synthesizer.

The resin was swollen with DMF for 2h before starting of the synthesis. After each coupling the resin was washed with DMF, MeOH and DCM dried under vacuum.

After the complete synthesis of the peptide:

ARPQDVNRRDQQRNPGTSTTPSQPNSA

I did test cleavage to verify the efficiency of the synthesis: 10 mg of peptidyl-resin were incubated with a mixture of TFA/TIS/H_2O (92.5:5:2.5, v/v) for 2h at room temperature.

The free peptide was precipitated by addition of cold diethyl ether and centrifugation. Finally, the supernatant was removed and the peptide washed with cold diethyl ether, dissolved in a mixture of H_2O/CH_3CN 1:1(v/v) and analyzed by RP-HPLC and HR-ESI-MS (Figure 3).

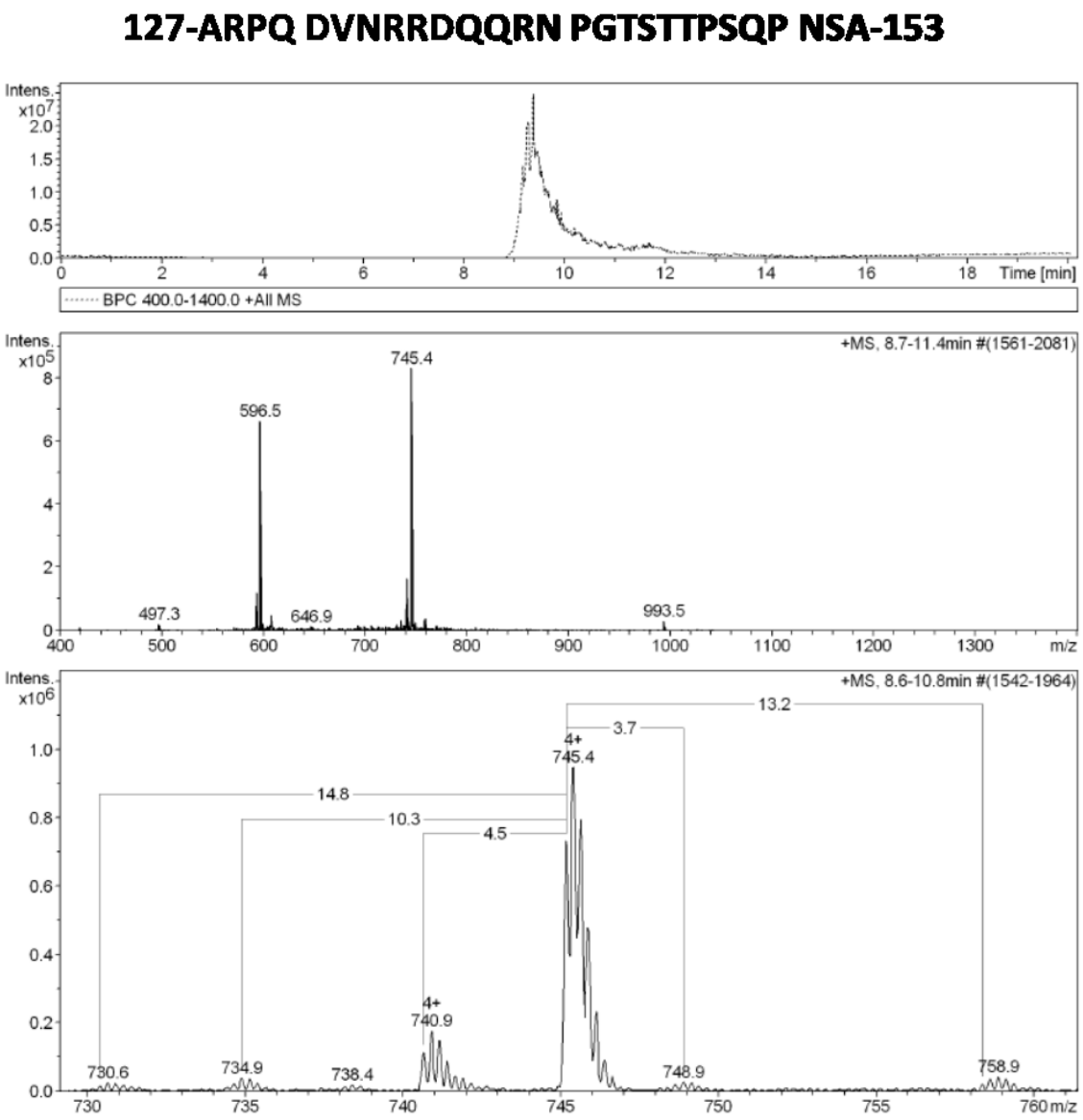


Figure 3: Mass spectrum of the peptide.

The base pick was consistent with the desired peptide.

After last alanine coupling, the glycosylated building block represented by Fmoc-Thr(GalNAc-Ac_3)-OH was coupled to the peptide: the reaction was performed using HATU (1.2 eq) as coupling reagent and DIEA(2.5 eq). The resin was incubated with the solution of activated modified amino acid for 2h and then washed with DMF.

Fmoc-Thr(GalNAc-Ac_3)-OH was inserted as N-terminal residue, in order to be linked with the PEG to perform the desired glycosylation. After Fmoc-Thr(GalNAc-Ac_3)-OH coupling, the resin was washed with DMF, MeOH and DCM and dried under vacuum. A test cleavage was performed to asses the efficiency of coupling of the modified amino acid. HR-ESI-MS allowed the characterization of the intermediate (Figure 4).

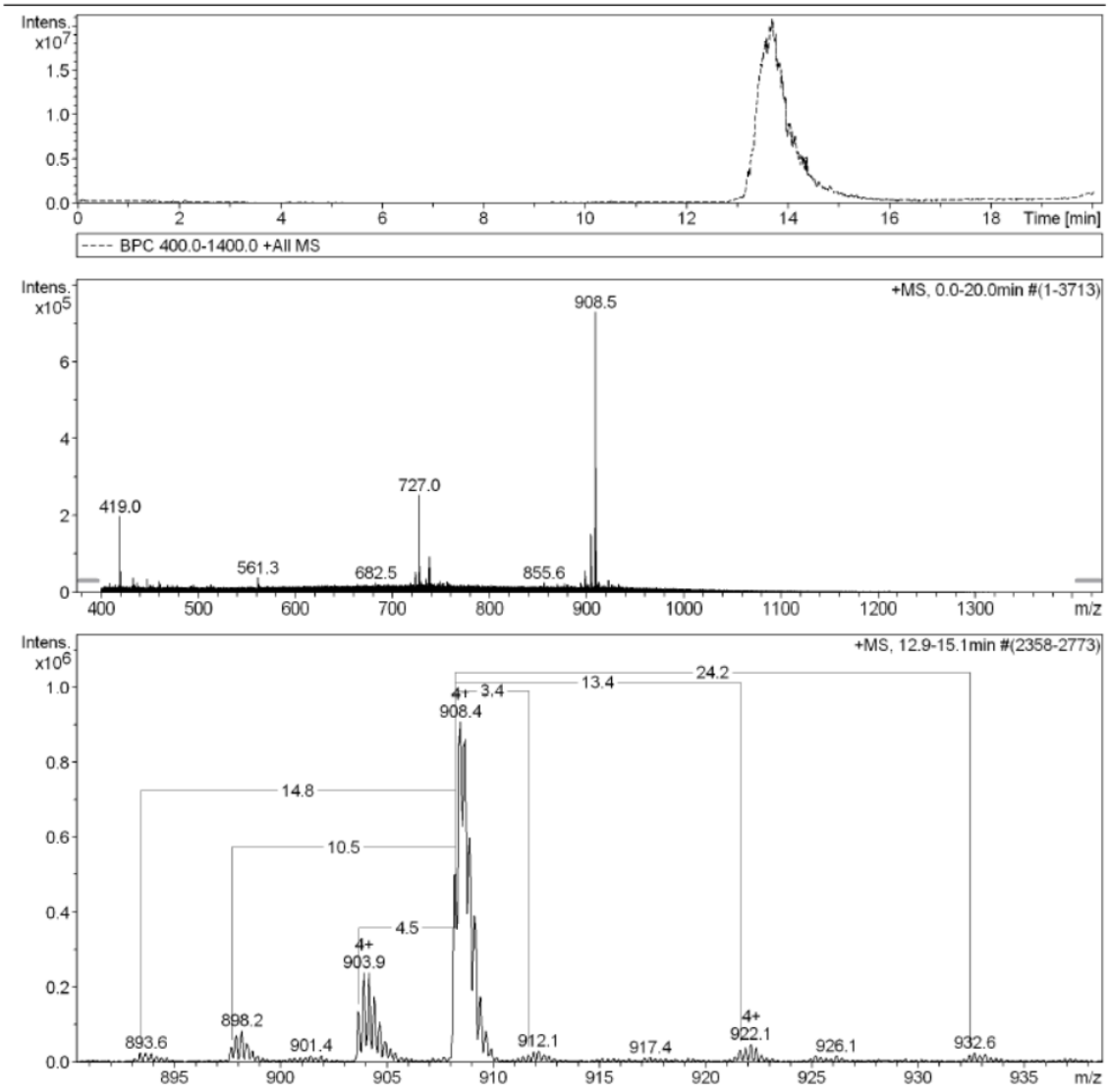


Figure 4: MS spectrum of peptide IGFBP6 [126-153] coupled with Fmoc-Thr(GalNac-Ac_3)-OH

N-terminal residue of the peptide was iodoacetylated in order to provide a binding site for the auxiliary: the resin was swollen in DMF for 1h, then the Fmoc protecting group was removed treating the resin twice with a solution of piperidine (20% v/v) in DMF for 3 min and 7 min, respectively.

After extensive washing with DMF, the resin was incubated with a solution of iodoacetic acid and diisopropylcarbodiimide in DMF for 30 min. After removal of the supernatant and washing with DMF, the coupling was repeated for 20 min (double coupling was performed to obtain the compound in higher yield).

The test cleavage of the final peptide showed that the desired product was obtained in a high purity and in high yield. Therefore, the synthetic strategy proved to be excellent (Figure 5).

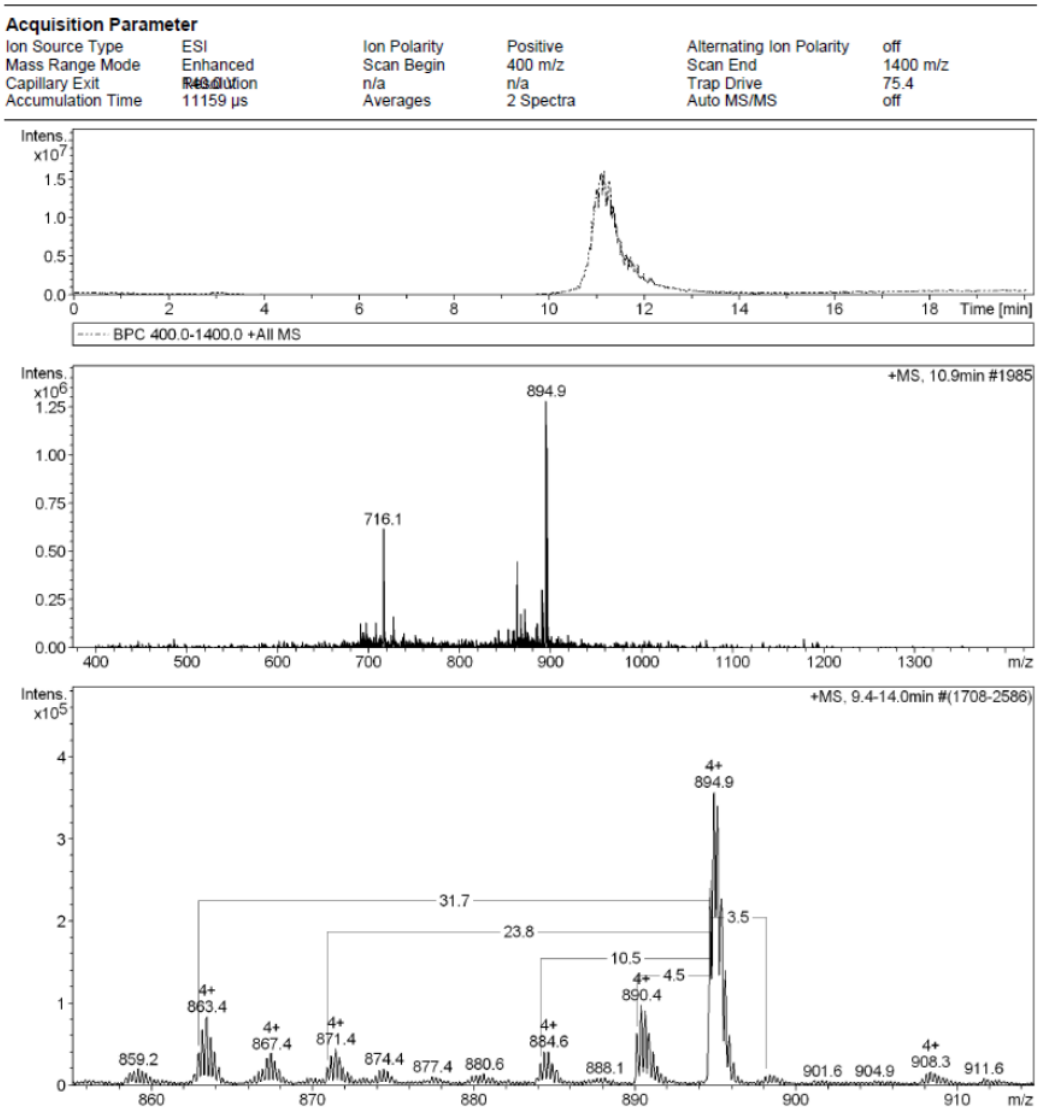


Figure 5: Mass spectrum of the final peptide.

In the latter part of my internship at the University of Vienna, I focused my research activity to preliminary tests aimed to link the auxiliary (**13**) to the final peptide: iodoacetylated peptidyl resin was swollen in CH_3CN for 3h. DIEA was added, followed by the auxiliary (dissolved in a mixture of CH_3CN and DMF).

The suspension was stirred at 26°C overnight, then the supernatant was removed (the unreacted auxiliary was recovered by recrystallization) and the resin washed with DMF, DCM and MeOH. Once again a test cleavage was performed to check the efficiency of the S_N2 reaction.

The final compound was obtained only in low percentage probably due to steric hindrance caused by the proximity of N-terminal of peptide and glycosilation site.

On this hypothesis we decided to further optimize and investigate the linkage procedure of the auxiliary to the peptide.

Prof. Christian Becker and Dr. Claudia Bello are continuing this research project with the aim to prepare homogeneously glycosilated analogues of the linker domain of IGFBP6 and, at the same time, discover if proximity of the auxiliary to the glycosylation site might cause problems during the enzymatic glycosylation reaction.

References

1. Sharp P.M., Hahn BH. Origins of HIV and the AIDS Pandemic.*Cold Spring Harb Perspect Med*. **2011.** a006841.
2. Merson M.H., O'Malley J., Serwadda D., Apisuk C. The history and challenge of HIV prevention. *Lancet* **2008**, 421-423.
3. Clavel F., Guetard D., Brun-Vezinet F., Chamaret S., Isolation of a new human retrovirus from West African patients with AIDS. *Science* **1986**,343-346.
4. Gao F., Bailes E.,Robertson DL., Chen Y. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature* **1999**, 436-41.
5. Vanden Haesevelde M., Peeters M., Jannes G. Sequence analysis of highly divergent HIV-1-related lentivirus isolated from a wild captured chimpanzee. *Virology*. **1996**, 346-50.
6. Van Heuverswyn F., Li Y., Neel C., Bailes E., Keele BF. Human immunodeficiency viruses: SIV infection in wild gorillas. *Nature.* **2006**, 164.
7. Van Heuverswyn F., Decker JM., Keele BF. Adaptation of HIV-1 to its human host. *Mol Biol Evol.* **2007**, 1853-1860.
8. De Leys R., Vanderborght B., Haesevelde MV. Isolation and partial characterization of an unusual human immunodeficiency retrovirus from two persons of west-central African origin. *J Virol.* **1990**, 1207-1216.
9. Plantier JC., Leoz M., Dickerson J., De Oliveira F. A new human immunodeficiency virus derived from gorillas. *Nat Med*. **2009**, 871-872.
10. Vallari A., Bodelle P., Ngansop C. Four new HIV-1 group N isolates from Cameroon: Prevalence continues to be low. *AIDS Res Hum Retrovirues*. **2010,** 109-115.
11. Wu, Y., HIV-1 gene expression: lessons from provirus and non-integrated DNA. Retrovirology, **2004**, 1-13.

12. Feng, S. and E.C. Holland, HIV-1 tat trans-activation requires the loop sequence within tar. *Nature*, **1988,**165-167.
13. Levin G.J., Mitra M., Mascarenhas A., Musier-Forsyth K. Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription.*RNA Biol*. **2010**, 754–774.
14. Gougeon M.L. Apoptosi as an HIV strategy to escape immune attack. *Nature Reviews Immunolog*y. **2003**, 392-404.
15. Moroni M. 25 anni di AIDS. Treccani Editore. **2007.**
16. Arts E.J. HIV-1 Antiretroviral Drug Therapy. *Cold Spring Harb Perspect Med. **2012***
17. Dalsleish A.G., Weiss R.A. HIV and the New Viruses.**1999**, 194-201.
18. Worobey M. Gemmel M., Teuwen DE., Haselkorn T. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature.* **2008**, 661-664.
19. Alan D. Frankel. , Young JAT. HIV-1. Fifteen Proteins and an RNA. *Annual Review of Biochemistry.***1998**, 1-25.
20. Eric J. Arts., Hazuda DJ. HIV-1 Antiretroviral Drug Therapy. *Cold Spring Harb Perspect Med*. **2012,** 2-4.
21. Young FE. The role of the FDA in the effort against AIDS. *Public Health Rep*. **1988**, 242–245.
22. Tantillo C, Ding J, Jacobo-Molina A, Nanni RG, Boyer PL, Hughes SH, Pauwels R, Andries K, Janssen PA, Arnold E. Locations of anti-AIDS drug binding sites and resistance mutations in the three-dimensional structure of HIV-1 reverse transcriptase. Implications for mechanisms of drug inhibition and resistance. *J Mol Biol.***1994,** 369-87.
23. Espeseth AS, Felock P, Wolfe A, Witmer M, Grobler J, Anthony N, Egbertson M, Melamed JY, Young S, Hamill T, Cole JL, Hazuda DJ. HIV-1 integrase inhibitors that compete with the target DNA substrate define a unique strand transfer conformation for integrase. *Proc Natl Acad Sci U S A. **2000,*** 11244-11249.

24. Hare S, Vos AM, Clayton RF, Thuring JW, Cummings MD, Cherepanov. Molecular mechanisms of retroviral integrase inhibition and the evolution of viral resistance. *Proc Natl Acad Sci*. **2010**, 20057–20062.

25. Miller V. International perspectives on antiretroviral resistance. Resistance to protease inhibitors. *J Acquir Immune Defic Syndr.* **2001***,* 34-50.

26. Wild C, Greenwell T, Matthews T. A synthetic peptide from HIV-1 gp41 is a potent inhibitor of virus-mediated cell-cell fusion. *AIDS Res Hum Retroviruses.* **1993.** *1051-3.*

27. Tsamis F, Gavrilov S, Kajumo F, Seibert C, Kuhmann S, Ketas T, Trkola A, Palani A, Clader JW, Tagat JR, McCombie S, Baroudy B, Moore JP, Sakmar TP, Dragic T. Analysis of the mechanism by which the small-molecule CCR5 antagonists SCH-351125 and SCH-350581 inhibit human immunodeficiency virus type 1 entry. *J Virol.* **2003***,* 5201-8.

28. Veazey RS, Klasse PJ, Ketas TJ, Reeves JD, Piatak M Jr, Kunstman K, Kuhmann SE, Marx PA, Lifson JD, Dufour J, Mefford M, Pandrea I, Wolinsky SM, Doms RW, DeMartino JA, Siciliano SJ, Lyons K, Springer MS, Moore JP. Use of a small molecule CCR5 inhibitor in macaques to treat simian immunodeficiency virus infection or prevent simian-human immunodeficiency virus infection. *J Exp Med.* **2003,** *1551-62.*

29. Pinar I., Karen S. A. Current Perspectives on HIV-1 Antiretroviral Drug Resistance. *Viruses.***2014***, 4095–4139.*

30. Min-Jung Kim, Seon Hee Kim, Jung Ae Park, Kyung Lee Yu, Soo In Jang, Byung Soo Kim, Eun Soo Lee, and Ji Chang You. Identification and characterization of a new type of inhibitor against the human immunodeficiency virus type-1 nucleocapsid protein. *Retrovirology.* **2015,** 218-219.

31. Wang J, Wang Y, Li Z, Zhan P, Bai R, Pannecouque C, Balzarini J, De Clercq E, Liu X. Design, synthesis and biological evaluation of substituted guanidine indole derivatives as potential inhibitors of HIV-1 Tat-TAR interaction. *Med Chem.* 2014*,* 738-746.

32. Lansdon EB, Brendza KM, Hung M, Wang R, Mukund S, Jin D, Birkus G, Kutty N, Liu X Crystal structures of HIV-1 reverse transcriptase with etravirine (TMC125) and rilpivirine (TMC278): Implications for drug design. *J Med Chem*. **2010,** 4295–4299.
33. Turner B.G., Summer M.F. Structural biology of HIV. *J Mol Biol*, **1999,** 1-32**.**
34. Davis, W.R., Davis WR., Gabbara S., Hupe D., Peliska JA. Actinomycin D inhibition of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase and nucleocapsid protein. *Biochemistry*, **1998,** 14213-21.
35. Hsiao-Wei Liu, B.S.C., Single-Molecule Studies on the Role of HIV-1 NucleocapsidProtein/Nucleic Acid Interaction in the Viral Replication Cycle. **2007**, University of Texas Libraries, Austin.
36. Liu H.W. , Landes CF., Zeng Y., Liu HW. Single- Molecule Studies on the Role of the HIV-1 Nucleocapsid Protein/ Nucleic Acid Interaction in the Viral Replication Cycle. *Journal of Biological Chemistry,* **2007,** 10181-10188.
37. Fisher, R.J., Rein A., Fivash M., Urbanejaet MA. Sequence-specific binding of human immunodeficiency virus type 1 nucleocapsid protein to short oligonucleotides. *J Virol,* **1998,** 1902-9.
38. Druillenec S. Dong CV., S Escaich S. A mimic of HIV-1 nucleocapsid protein impairs reverse transcription and displays antiviral activity. *Biochemistry.* **1999**, 4886- 4891.
39. Kim M.Y., Jeong S. Inhibition of the function of the nucleocapsid protein of human immunodeficiency virus 1 by an RNA aptamer. *Biochemical and Biophysical Research Communications*. **2004**, 1181–1186.
40. Stephen, A.G., et al., Identification of HIV-1 nucleocapsid protein: nucleic acid antagonists with cellular anti-HIV activity. *Biochem Biophys Res Commun*, **2002**, 1228-37.
41. Vercruysse T., Kim JM., Kim SH., Park JA., Yu KL. Identification and characterization of a new type of inhibitor against the human immunodeficiency virus type 1 nucleocapsid protein. *Retrovirology* .**2015**, 218-219.

42. Davis W., Gabbara S., Hupe D., Peliska JA. Actinomycin D Inhibition of DNA Strand Transfer Reactions Catalyzed by HIV-1 Reverse Transcriptase and Nucleocapsid Protein. *Biochemistry*.**1999**, 14213-14221.

43. Sosic A., Sinigaglia. L., Cappellini M., Carli M., Parolin C., Zagotto G., Sabatino G., Rovero P., Fabris D., Gatto B. Mechanism of HIV-1 Nucleocapsid Protein Inhibition By Lysyl-Peptidyl-Anthraquinone Conjugates. *Bioconjugate Chemistry*. **2016**, 247-256.

44. Sosic A., Frecentese F., Perissutti E., Sinigaglia E., Santagada V., Caliendo G., Magli E., Ciano A., Zagotto G., Parolin C., Gatto B. Design, synthesis and biological evaluation of TAR and cTAR binders as HIV-1 nucleocapsid inhibitors. *MedChemComm,* **2013,** *4*, 1388-1393.

45. Mason, S., Goudreau N., Hucke O., Faucher AM. Discovery and structural characterization of a new inhibitor series of HIV-1 nucleocapsid function: NMR solution structure determination of a ternary complex involving a 2:1 inhibitor/NC stoichiometry. *J Mol Biol* **2013,** 425, 1982-1998.

46. Guo, J.; Henderson, L. E.; Bess, J.; Kane, B.; Levin, J. G. Human immunodeficiency virus type 1 nucleocapsid protein promotes efficient strand transfer and specific viral DNA synthesis by inhibiting TAR-dependent self-priming from minus-strand strong-stop DNA. *J Virol* **1997,** 71, 5178-5188.

47. Mori, M.; Kovalenko, L.; Lyonnais, S.; Antaki, D.; Torbett, B. E.; Botta, M.; Mirambeau, G.; Mely, Y. Nucleocapsid Protein: A Desirable Target for Future Therapies Against HIV-1. *Curr Top Microbiol Immunol* **2015,** 389, 53-92.

48. Gooch, B. D.; Beal, P. A. Recognition of duplex RNA by helix-threading peptides. *J Am Chem Soc* **2004,** 126, 10603-10610.

49. Gooch, B. D.; Krishnamurthy, M.; Shadid, M.; Beal, P. A. Binding of helix-threading peptides to E. coli 16S ribosomal RNA and inhibition of the S15-16S complex. *Chembiochem* **2005,** 6, 2247-2254.

50. Krishnamurthy, M.; Gooch, B. D.; Beal, P. A. Peptide quinoline conjugates: a new class of RNA-binding molecules. *Org Lett* **2004,** 6, 63-66.

51. Krishnamurthy, M.; Schirle, N. T.; Beal, P. A. Screening helix-threading peptides for RNA binding using a thiazole orange displacement assay. *Bioorg Med Chem* **2008,** 16, 8914-8921.

52. Krishnamurthy, M.; Simon, K.; Orendt, A. M.; Beal, P. A. Macrocyclic helix-threading peptides for targeting RNA. *Angew Chem Int Ed Engl* **2007,** 46, 7044-7047.

53. Turner, K. B.; Hagan, N. A.; Fabris, D. Inhibitory effects of archetypical nucleic acid ligands on the interactions of HIV-1 nucleocapsid protein with elements of Psi-RNA. *Nucleic Acids Res* **2006,** 34, 1305-1316.

54. Turner, K. B.; Hagan, N. A.; Kohlway, A. S.; Fabris, D. Mapping noncovalent ligand binding to stemloop domains of the HIV-1 packaging signal by tandem mass spectrometry. *J Am Soc Mass Spectrom* **2006,** 17, 1401-1411.

55. Turner, K. B.; Kohlway, A. S.; Hagan, N. A.; Fabris, D. Noncovalent probes for the investigation of structure and dynamics of protein-nucleic acid assemblies: the case of NC-mediated dimerization of genomic RNA in HIV-1. *Biopolymers* **2009,** 91, 283-296.

56. Sosic, A.; Cappellini, M.; Scalabrin, M.; Gatto, B. Nucleocapsid Annealing-Mediated Electrophoresis (NAME) Assay Allows the Rapid Identification of HIV-1 Nucleocapsid Inhibitors. *J Vis Exp* **2015,** 95, 52474.

57. Sosic, A.; Cappellini, M.; Sinigaglia, L.; Jacquet, R.; Deffieux, D.; Fabris, D.; Quideau, S.; Gatto, B. Polyphenolic C-glucosidic ellagitannins present in oak-aged wine inhibit HIV-1 nucleocapsid protein. *Tetrahedron* **2015,** 71, 3020-3026.

58. Zagotto, G.; Sissi, C.; Lucatello, L.; Pivetta, C.; Cadamuro, S.A.; Fox, R. K.; Neidle, S.; Palumbo, M Aminoacyl-anthraquinone Conjugates as Telomerase Inhibitors: Sinthesis, Biophysical and Biological Evaluation. *J. Med. Chem*. **2008**, *51*, 5566-5574.

59. Frecentese F.; Sosic A.; Saccone I.; Gamba E.; Link K.; Miola A.; Severino B.; Fiorino F.; Magli E.; Corvino A.; Perissutti E.; Fabris D.; Gatto B.; Caliendo G.; Santagada V. Synthesis

and *in vitro* screening of new series of 2,6-dipeptidyl-anthraquinones: influence of side chain length on HIV-1 nucleocapsid inhibitors. *J. Med. Chem.* **2016**, 59, 1914-1924.

60. Vivian H., Youngman O., Rosenfeld G.R. The Insulin-Like Growth Factor-Binding Protein (IGFBP) Superfamily.*Endocrine Review*. **2013,** 761-786.
61. Beauchamp M.C., Yasmeen A., Knafo A., H. Gotlieb W.H. Targeting Insulin and Insulin-Like Growth Factor Pathways in Epithelial Ovarian Cancer. *Journal of Oncology.* **2010,** 1-11.
62. Briony E. F., P., Norton R.S. Insulin-Like Growth Factor Binding Proteins: A Structural Perspective. *Front Endocrinol*. **2012,** 1-13.
63. Bach L.A. Recent insights into the actions of IGFBP-6. *J Cell Commun Signal.* **2015,** 189-200.
64. Bello C., Wang S., Meng L., Moremen K.W. PEGylated Photoclevable Auxiliary Mediates the Sequential Enzymatic Glycosilation and Native Chemical Ligation of Peptides.**2015**, 7711-7715.

I would like to thank all the people who contributed in different way to my PhD thesis.

First of all I want thank my supervisor, ***Prof. Vincenzo Santagada*** *to give me this opportunity and for his positive attitude and passion for his work that provided a strong motivation to my PhD work.*

I would like to thank ***Prof Christian Becker*** *and* ***Dott. Claudia Bello*** *for their great support during my internship in Vienna, giving to me the possibility to learn how to handle alone a research project and the opportunity to meet people from all over the world, with their respective cultures. In particular* ***Dott. Claudia Bello,*** *a great researcher and person, for her excellent supervision for teaching me so many things, making me independent from the scientific point of view.*

Prof Caliendo *for his proverbs related to the life and work and all the research team for their support.*

I would like to thank my colleague and my dear friend Dott. ***Angela Corvino*** *because she allowed me to trust in her since the first day we met and has become an essential person in my life during my PhD. For our daily help and to be always a shoulder to each other in every moments, keeping always open our minds.*