UNIVERSITÀ DEGLI STUDI DI SALERNO

DIFARMA

UNIVERSITÀ DEGLI STUDI DI SALERNO
Dipartimento di Farmacia

Dottorato di ricerca
in Scienze Farmaceutiche
Ciclo XIV — Anno di discussione 2016

Coordinatore: Chiar.mo Prof. *Gianluca Sbardella*

Design, synthesis and biological evaluation of new small-molecule modulators of Protein Methyltransferases (PMTs)

settore scientifico disciplinare di afferenza: CHIM/08

Dottorando	Tutore
Dott. *Agostino Cianciulli*	**Chiar.mo Prof.** *Gianluca Sbardella*

INDEX

ABSTRACT

A large amount of evidences indicate that dysregulation of protein methylation is linked to the genesis and progression of several human diseases, including cancer. Therefore over the past years small-molecule modulators targeting Protein Methyltransferases (PMTs) have been actively developed as anticancer drugs as well as chemical tools to better understand the biological and physiological roles of protein methylation. In this thesis the design, synthesis and biological evaluation of three different classes of compounds, designed as isozyme-selective PRMT inhibitors, are reported. (1) Inspired by the structure-activity relationships (SAR) of pyrazole and indole compounds, the most potent PRMT4 inhibitors, we developed a series of pyrrole-based compounds, designed as inhibitors of PRMT4. A potent inhibition was observed when testing pyrrole derivatives against PRMT4 (i.e. **EML 438**, IC_{50} = 2.42 µM), nevertheless they didn't prove a significant cellular activity, due to their poor transcellular permeability. Therefore, in order to increase the activity and the lipophilicity of these compounds and supported by computational studies, we started a process of structural optimization of this class of compounds. Novel derivatives have been designed and selected. (2) Furthermore, in collaboration with Professor Nathaniel Martin from the University of Utrecht, we have successfully synthesized a set of novel PRMT4 bisubstrate inhibitors. Preliminary screening of their biological activity revealed a nanomolar inhibition of PRMT4 (**P2-C3-unsat**, IC_{50} = 43 nM) with about 900-fold selectivity for CARM1 over PRMT1. This data confirm our hypothesis that a potent and selective enzymatic inhibition is achieved by compounds characterized by structural features able to bind both the enzymatic substrate binding sites, thereby mimicking the transition state. (3) We also started a program aiming at developing inhibitors of PRMT3. Three series of indole-based compounds were prepared and their activity against a panel of PRMTs was assessed. **EML598** and **EML599** showed a selective inhibition of PRMT3 at fixed dose (70% at 100 µM). Further biophysical assays are still ongoing in order to better evaluate their biological activity.

CHAPTER 1

INTRODUCTION

1.1. Epigenetics

Epigenetics is the study of heritable changes in genome function that occur without alterations to the DNA sequence. Most of these changes are established during differentiation and are stably maintained through multiple cycles of cell division, enabling cells to have distinct identities while containing the same genetic information. Unlike the genome itself, epigenetic modifications are cell specific, plastic, and responsive to environmental influences. There are many examples in our everyday lives where common inputs result in differential outcomes, depending on specific control mechanisms. Consider, for example, how a talented musician can produce a wide variety of music on a guitar. A particular piece of music is encoded by the sequence and timing of notes, that are typically transcribed onto sheet music. The guitarist translates the sheet music by spatial and temporal modifications to the guitar strings; that is, he/ she holds down different strings, at different frets, at different times to produce the music. In an analogous manner, gene expression is spatially and temporally controlled within cells by specific epigenetic modifications that can signal the cell to turn on or turn off the transcription of particular genes.

There are three reasons for the intense interest in Epigenetics: (1) epigenetic mechanisms are involved in many fundamental areas of biology, and are often underpinned by novel, interesting mechanisms. (2) Epigenetic regulation plays a central role in gene expression, so is involved in (and occasionally responsible for) initiating disease processes. (3) A large number of enzymes and protein complexes are involved in epigenetic processes, and are potential targets for small-molecule inhibition. To date, two classes of drugs with epigenetic targets have been approved for cancer chemotherapy, but similar 'epigenetic therapies' are likely to be developed for a broad range of diseases in the medium term.

1.2. Chromatin modifications and Histone code

The chromosomes within eukaryotic cell nuclei are packaged together with structural proteins, histones, to form the complex known as chromatin. This macromolecular complex has a straightforward modular structure based on a single repeating building block known as nucleosome, which consists of eight histone proteins (two copies of each histones H2A, H2B, H3 and H4), around which DNA is wound almost twice (**Figure 1.0**).[1]
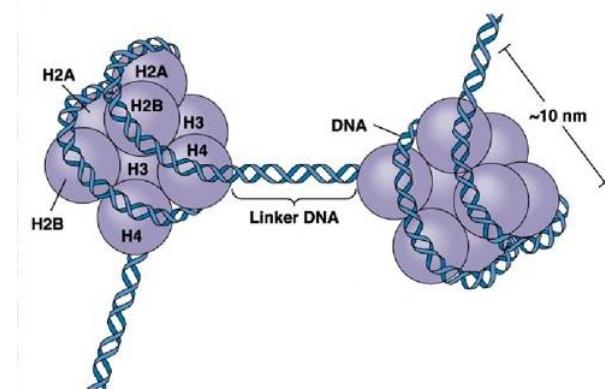
Figure 1.0. Structure of nuclesome

The chromatin exists in two main conformational states: heterochromatin and euchromatin (**Figure 1.1**). The chromatin can fold upon itself to compact the nucleosomes, forming a highly condensed structure, referred to as heterochromatin. This provides an effective mechanism for storing a large quantity of DNA in a very small volume. However, in order for genes to be transcribed into mRNA molecules, portions of the gene (referred to as promoter regions) need to interact with particular proteins, such as transcription factors that initiate the process of mRNA transcription. In the condensed heterochromatin structure these promoter regions are largely inaccessible to the transcription factors; hence protein production is inhibited. Therefore, to initiate protein production (also referred to as gene expression) the heterochromatin structure needs to be relaxed in a specific and controlled manner, so as to expose the promoter regions of particular genes that need to

be transcribed at specific times and under specific conditions during the lifetime of a cell. This more relaxed state is referred to as euchromatin.

Notably, the transition between the two conformational states of chromatin is accomplished by changing the structure of the nucleosomes by adding or removing small chemical units (e.g., acetyl groups, methyl groups, phosphate groups and ubiquitin groups)[2] at specific locations on the histones (**Figure 1.1**).
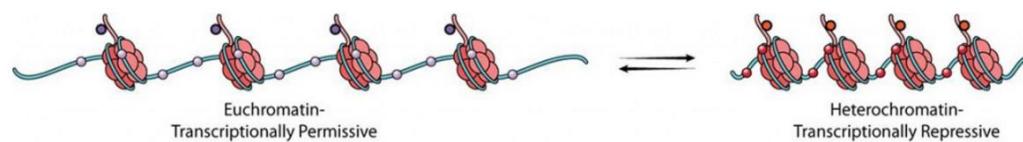
Figure 1.1. Two main conformational states: heterochromatin and euchromatin. The transition between the two states is regulated by post-translational modifications

Histones, the core components of the nucleosome, are basic proteins characterized by a comparable overall structure, with a globular hydrophobic internal region, around which the DNA sequence is rolled up, and a flexible charged amino termini (*histone tail*) protruding out the nucleosome, that represents the platform whereby the post-translational modifications (PTMs) occur (**Figure 1.2**).
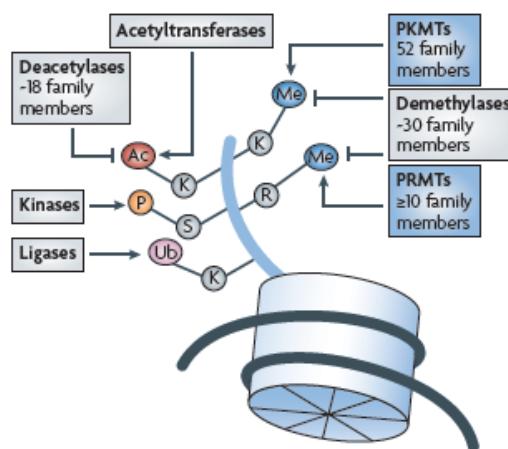
Figure 1.2. Representation of the post-translational modifications (PTMs) that occur on the histone tail

3

These reversible epigenetic modifications, performed by selective epigenetic enzymes within the cell, constitute the main features of the so-called "*histone code*".[2] The proposed function of this code is to integrate exogenous and endogenous signals into a diverse set of PTM patterns to enable the epigenetic control of gene expression, through a remodeling of the chromatin structure. Indeed, the chromatin modifications physically enhance or weaken the noncovalent interactions between histones or between histones and DNA, determining accessibility to specific DNA loci, providing also an informative platform for the recruitment of epigenetic regulators.

The epigenetic enzymes can be divided into distinct groups based on their broad functions: epigenetic *writers, erasers* and *readers* (**Figure 1.3**).[3]
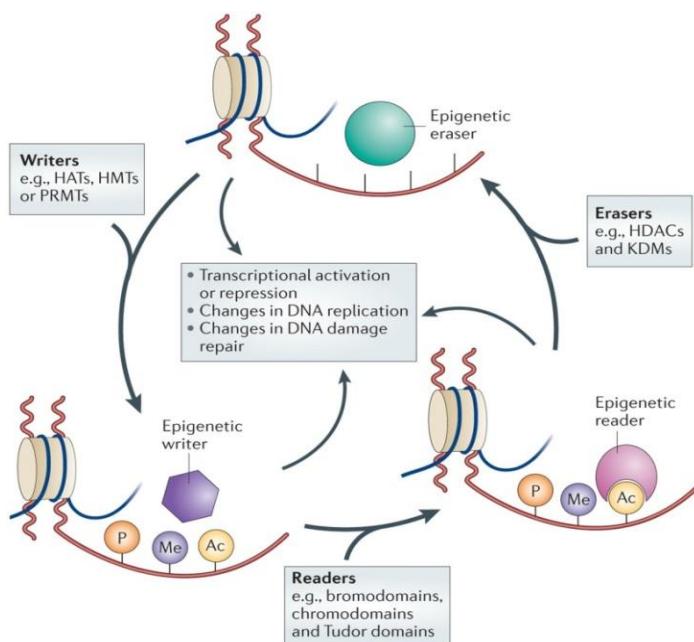
Figure 1.3. Epigenetic writers, readers and erasers

1.2.1. Epigenetic writers, readers and erasers

The epigenetic writers catalyze the addition of chemical groups onto either histone tails or the DNA itself. These modifications are known as epigenetic

marks, and are integral to gene expression and silencing. One such group of epigenetic writers is histone methyltransferases; these enzymes catalyze the transfer of a methyl group onto a lysine or arginine residue on histone tails. In addition to methyl marks, histone lysine residues may also undergo acetylation through the activity of histone acetyltransferases.[4] The transfer of an acetyl group from the cofactor acetyl-CoA to lysine residues on histone tails neutralizes the positive charge of lysine, thereby weakening the affinity of the histone tail for the DNA and reducing chromatin condensation. Since a more relaxed, open chromatin architecture enables the recruitment of transcription factors and polymerases, histone acetylation results in the promotion of gene expression.

Enzymes that catalyze the phosphorylation of histone tails are also important epigenetic writers.[5] For example, phosphorylation of histone H3 (H3Y41) by JAK2 disrupts binding of the heterochromatin protein HP1α to chromatin, leading to increased DNA accessibility and the transcription of the oncogene Lmo2. A further epigenetic mark that alters gene expression is ubiquitination.[6] Lysine residues on histone proteins H2A and H2B can undergo monoubiquitination through the concerted actions of E2 ubiquitin conjugases and E3 ubiquitin ligases.

DNA can also undergo methylation through different mechanisms. The addition of a methyl group to a nucleotide by DNA methyltransferases (DNMTs) occurs at the major groove of the DNA double helix, and prevents transcription by blocking the binding of transcription factors and polymerases.[7] There are two known types of DNA methylation – *de novo* and *maintenance methylation. De novo* methylation, predominantly carried out by DNA methyltransferases DNMT3A and DNMT3B, catalyzes the addition of methyl groups onto cytosine nucleotides. Since cell replication does not preserve such methylation, *maintenance methylation* copies these marks from the parent DNA onto the daughter DNA strands. The high affinity of DNMT1

for hemimethylated DNA *in vitro* suggests that this enzyme is primarily responsible for maintenance DNA methylation *in vivo.*

Epigenetic marks are not necessarily permanent modifications; instead, they can be removed by a group of enzymes known as epigenetic *erasers* in order to alter gene expression. There are multiple categories of epigenetic erasers that target histones; these include histone deacetylases,[8] histone serine/threonine/tyrosine phosphatases, histone deubiquitinases and histone lysine/ arginine demethylases.[9] The removal of acetyl groups through the actions of histone deacetylases (HDACs) is an important mechanisms for increasing chromatin condensation and therefore repressing gene transcription. HDACs can be divided into two groups termed group I and group II.

Histone phosphatases can target either phosphorylated serine, threonine or tyrosine residues on histone proteins.[10] Protein Serine/Threonine phosphatases PP1, PP2A and PP4, amongst others, have been reported to dephosphorylate histone proteins.

The removal of ubiquitin groups from histone lysine residues is catalyzed by proteases known as deubiquitinating enzymes (DUBs).[10] These proteins target histones H2A and H2B, where they regulate transcription, DNA repair, gene expression and cell cycle progression. Compared to other histone modifications, the functions of histone ubiquitination are less well understood, yet increasing evidence points to an important role for this epigenetic modification in the DNA damage response.

The first histone lysine demethylase to be discovered was lysine-specific demethylase 1 (LSD1),[11] also known as KDM1. LSD1 contains an amino oxidase domain that binds the cofactor, flavin adenine dinucleotide (FAD), crucial for demethylation. A further family of lysine demethylases have since been identified; these are termed Jumonji C domain-containing demethylases (JMJD). JMJD do not require FAD as a cofactor but instead are dependent on

Fe^{2+}/2-oxoglutarate (2-OG) for catalysis. As yet, only one enzyme with arginine demethylase activity has been identified, JMJD6, which is a 2-OG-dependent JMJD.

In contrast to the well-defined mechanisms of the removal of epigenetic marks from histones, the mechanism by which methyl groups are removed from DNA remains poorly understood. What is known, however, is that DNA demethylation can occur both actively and passively.[11] Passive DNA demethylation involves the failure of maintenance DNA methyltransferases to methylate newly synthesized DNA strands during mitosis, whilst the molecular machinery that catalyzes active DNA demethylation occurs is yet to be elucidated. Since demethylation of DNA is crucial for processes such as epigenetic reprogramming in germ cells, further research into the mechanisms of active DNA demethylation may identify novel targets for stem cell research.

Additionally, epigenetic *reader domains* can be thought of as effector proteins that recognize, and are recruited to specific marks on histones or nucleotides.[12] Enzymes which write or erase epigenetic marks may also contain such reader domains, leading to the coordination of 'read-write' or 'read-erase' epigenetic processes. The structure of reader domains typically provides a cavity or surface groove in which to accommodate a specific epigenetic mark.

Proteins that contain reader domains can be broadly classified into four groups:[12] chromatin architectural proteins; chromatin remodeling enzymes; chromatin modifiers and adaptor proteins that recruit other machinery involved in gene expression. The first group of these, chromatin architectural proteins, binds to nucleosomes and can either directly induce chromatin compaction or alternatively act as a shield to prevent the binding of proteins involved in RNA transcription.

In contrast to chromatin architectural proteins, chromatin remodeling enzymes prompt a more open chromatin architecture. The increased accessibility of chromatin facilitates DNA transcription, promoting gene expression.

Many other proteins that contain reader domains cannot directly influence chromatin architecture, but instead serve to recruit secondary chromatin modifiers to further modify chromatin or to reverse an existing chromatin modification.

The final class of reader domain-containing proteins is adaptor proteins: the principal function of these domains is to recruit factors that are linked to DNA metabolism processes including transcription, DNA damage repair, DNA recombination, DNA replication and RNA processing.

1.3. Protein Methyltransferase (PMTs)

Histone methylation by Protein Methyltransferases (PMTs) is one of the most studied post-translational modifications since it is implicated in heterochromatin formation and maintenance, transcriptional regulation, DNA repair, X-chromosome inactivation, and RNA maturation.[13]

Protein Methyltransferases (PMTs) catalyze the methylation of lysine or arginine residues on the histone or non-histone substrates.[14] [15] They can be divided into two classes: Protein Lysine Methyltransferase (PKMT) and Protein Arginine Methyltransferase (PRMT) (**Figure 1.4**). The PKMT is responsible for the generation of mono-, di-, and trimethylated lysine residues, while the PRMT is responsible for the generation of monomethyl arginine (MMA), symmetric dimethylarginine (SDMA), and asymmetric dimethylarginine (ADMA; Figure 2). The PKMT is a large class of enzymes with more than 50 members and can be divided into eight groups (KMT1-KMT8). Whereas, the PRMT can be distinguished into three subtypes based on the methylation products:[16] type I (PRMT1, 4, 6, 8) for ADMA and MMA

8

generation, type II (PRMT5 and PRMT9) for SDMA and MMA generation, and type III (PRMT7) for MMA generation only.
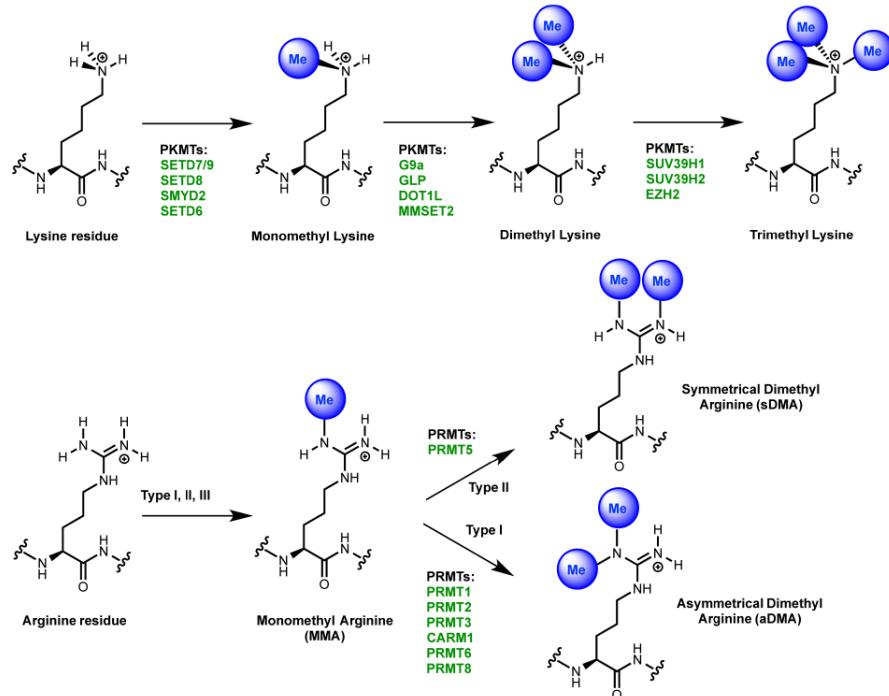
Figure 1.4. Methylation States of Lysine and Arginine Residues

Even though lysine or arginine methylation does not change the charge state of the residue side chain, the increased bulkiness and hydrophobicity of the methylated moiety would affect protein-protein interactions and protein recognition, thus gene expression and transcription regulation. Alterations of the post-translational modifications of arginine and lysine residues could have dramatic effects on cell signaling and, like other histone modifications, contribute to disease pathogenesis. Indeed, there are increasing experimental evidences to suggest that the deregulations of arginine and lysine modifying enzymes play pivotal roles in cancer, inflammatory diseases, neurodegenerative diseases, and other conditions.[9] On these bases, PMTs have been increasingly recognized as potential therapeutic targets and there is

growing interest in the scientific community to discover and develop selective small-molecule modulators of these enzymes.

Yet today there are tool compounds and clinical candidates[17] reported for < 15% of these enzymes. The PMTs for which small molecule modulators have been reported represent those enzymes with some of the best pathobiological validation, as is appropriate in novel target space. Nevertheless, the relative paucity of chemical coverage of this target class severely limits chemical biology approaches for understanding the biological and pathobiological roles of these enzymes more fully. A breakthrough in this area will be truly exciting and is keenly awaited.

1.3.1 Protein Lysine Methyltransferases (PKMT)

As previously mentioned protein methylation may occur on lysines and specifically on the ε amino groups of their side chains. This covalent reversible modification is inserted by a distinct family of methyltransferases, called Protein Lysine Methyl Transferases (PKMTs). This enzymes are capable of a *substrate selectivity* being able to modify only distinguishing lysinic residues of their substrates, and moreover of a *product specificity* being endowed to methylate these residues only to a characteristic state as mono- (me1), di- (me2) and trimethylated (me3). As well as protein arginine methyltransferases, the PKMTs catalyze the methylation transfer using *SAM* or *S*-Adenosyl-L-Methionine (AdoMet) as cofactor and affording *SAH* or *S*-Adenosyl-L-Homocysteine (AdoHcy) and the methylated substrate as products (**Figure 1.4**).

More than 50 human lysine methyltransferases have been identified and characterized so far (**Figure 1.5**). According to structural and sequence criteria of their catalytic domain, these can be classified into two families: the DOT1-like (DOT1L) and the SET-domain-containing methyltransferases. To date

only the DOT1L enzyme belongs to the former family, which is distinguished by the presence of seven-stranded *β*-sheet motif, which also structurally characterizes the PRMTs family. [18]Another aspect that makes DOT1L unique is the fact that it is the only methyltransferase catalyzing the methylation reaction of Lysine 79 on histone H3 (H3K79), which differs from the other lysines as it is located in the inner part of the nucleosomes. H3K79 methylation is a transcriptional activating histone mark playing pivotal roles in early embryonic development, and DOT1L has been found involved in MLL translocation effects in mixed lineage leukemia.[19]

Most of the PKMTs belong to the SET-domain-containing protein subset which is in turn organized into different subfamilies according to their founding member: EZ, RIZ, SET1, SET2, SMYD3 and SUV39 (**Figure 1.5).**


Figure 1.5. Phylogenetic tree of Protein Methyl Transferases (PMTs).

The SET domain is a sequence of 130 amino-acids and it takes its name from the Drosophila genes Su(var)3-9, E(z) (Enhancer of zeste) and Tritorax where
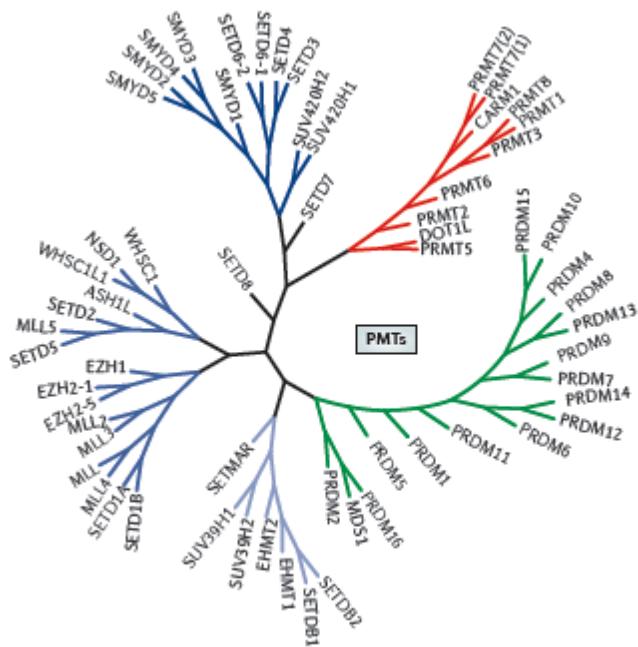
it was initially identified. Within the SET domain the substrate and the cofactor bind in different sites, flanking each other, and meet at the core of the domain through a narrow hydrophobic channel, where the methyl transfer takes place. Albeit SET domain is the responsible for the methyltrasferase activity, the catalysis also relies on the simultaneous presence of other domains that act as "framework" for the recruitment of the substrate or other partners within large multiunit complexes or modulate the catalytic step, like for example in the Polycomb Repressive Complex 2 (PRC2), where the catalytic subunit EZH2 displays a methyltransferase activity only in complex with the non-catalytic partners EED and SUZ12. All these surrounding domains can be divided into two subsets: the I-SET and post-SET motifs (respectively inserted within and *C*-terminal shifted respect to the SET domain) that participate to the substrate recognition and extent to the cofactor binding; and the Pre-SET (cysteine-rich also referred to as a CXC domain), MYND (between I-SET and SET) and CTD (C-terminal to Post-SET) domains, which are thought to be interfaces for the interaction with other proteins and DNA. The pre-SET motif seems to play an important role also in the catalytic event.[20] According to this, distinct domains with different sequence and structure may decorate the SET fold in assorted combinations to obtain selective recruitment of interaction partners and therefore specific functions. Being involved in the substrate recognition, the I-SET and post-SET domain are always found in the SET methyltransferases. Even if the amino acidic sequence can change, the I-SET sequence has a overall conserved structural organization, while the Post-SET domain is more dynamic and can adopt different conformations.

As far as the substrate recognition is concerned, different mechanistic analysis has displayed that the Lys- and Arg-rich electropositive histone tails and the overall electronegative substrate-binding groove, spatially defined by pre-formed I-SET and the open folded Post-SET, initially interplay through a long

range electrostatic interactions. Once this loose complex has been formed, SAM binding induces a partially folded Post-SET conformation which keeps close the PKMT and the substrate, allowing the sequence specific recognition to start. In this proposed model, the PKMT slides along the substrate peptide, with I-SET domain acting as a reading template. Once specific interactions have been established, the lysine substrate looses a proton to the solvent and fits into the hydrofobic channel. The loose complex thus turns into a catalytically competent conformation and the substrate peptide is further fasten by the Post-SET domain, which closes on it and shields the active site from the surrounding solvent. Owing to a tyrosine residue, the deprotonated ε-amino group of lysine substrate aligns its lone pair with the assailable methyl-sulfur bond and methyltransfer takes place.[21]

The structure of the active site are peculiar of each PKMTs and suggest that is theoretically possible to develop selective PKMTs inhibitors. As previously stated SET-domain containing methyltransferases not only exhibit a substrate selectivity but also a product specificity. EZH2 has the greatest catalytic activity in mono-methylating the H3K27 but can achieve all three methylated states of this lysine, SET7/9 affords only monomethylated H3K4 (H3K4me1) after one round of catalysis, while G9a and GLP are either mono- and dimethyltransferases for H3K9 (H3K9me1/me2).[22] This product specificity can be realized on the basis of structural and sequence explanations: the ability to mono-, di- and trimethylate lysine substrates seems to be correlated to a "tyrosine-phenylalanine switch" and in general to the steric crowding of the hydrophobic channel, since some PKMTs (Dim5, G9a) change their specificity following point mutations from tyrosine to phenylalanine, and vice versa, nearby the active site.[23]

1.3.1.2. *Physiopathological role of PKMT*

The protein lysine methyl transferases exert their activity mainly on histones even if other proteins have been identified as PKMTs substrates.[24, 25, 26]A extensive example may be represented by the tumour suppressor protein p53, which can be targeted by different PKMTs with opposing cellular outcomes depending on the site of lysinic methylation: p53-responsive genes may be transcriptionally repressed after methylation at K370 by SMYD2[27] or on K382 by SETD8,[14] while methylation of K372 by SET7/9 result in transcriptional activation;[15] p53 may be also inactivated because of the methylation on K373 by G9a and GLP.[28]

Histones represent the main site of action of lysine methyltransferases which are therefore referred to as Histones Lysine Methyl-Transferases (HKMTs). As epigenetic "*writers*" they are correlated with both transcriptional activation and silencing, depending on the specific residue involved and on the extent of methylation, since lysine may be methylated up to three times. Nonetheless, generally methylation on H3K4, H3K36 and H3K79 is associated with activation of transcription whereas methylation at H3K9, H3K27 and H4K20 is correlated with transcriptional silencing.[29] Differently from the lysinic acetylation, whose transcription-activating mechanism is essentially based on the neutralization of the positive charge of the lysine side chains (which in turn interrupts the electrostatic interaction with phosphate groups of DNA backbone); lysine methylation doesn't affect the overall charge of the residues. Methyl groups rather seem acting as "*hubs*" for the recruitment of highly evolved methyl-lysine-binding proteins able to distinguish distinct level of methylation and the surrounding amino acid sequence, especially along histone tails. The most remarkable biological consequences of the enhanced specificity of HKMTs is that each component of the former family of enzymes controls genes expression in a narrower group of cell types, being them normal or cancerous, thus proving to be more promising targets for drug development.

In this concern a multitude of evidences draw attention to the wide spreading involvement of HKMTs in the human diseases related biochemical pathways including oncogenic transformation, inflammation, metabolic and neuropsychiatric disorders, finally in the handling of stem cells as tools for regenerative medicine.[30]

HKMTs may contribute to the diseases development or maintenance by gaining aberrant activity due to mutations, altered expression or translocations that directly affect cellular genes expression, or being involved in altered upstream cellular signals.

In cancer, MLL1 (mixed –lineage leukemia 1) methyltransferase is subject of more than 50 chromosomal translocations, especially in human lymphoid and myeloid leukemias, where these are associated with very poor prognosis.[31] Despite the translocations are responsible for the loss of the carboxy-terminal SET domain, the N-terminal fusion proteins upregulate expression of several target genes including HoxA7, HoxA9 and the Hox factor MEIS1 that are important for proliferation and final differentiation of hematopoietic cells.[32] During normal hematopoiesis, expression of HoxA7, HoxA9 and HoxA10 promote stem cell self-renewal, and the downregulation of these genes correlates with terminal differentiation.[33] The inappropriate transcriptional activation seems to arise from the interplay with transcriptional elongation partners like AFF4, AFF1, AF9, and ENL, and other epigenetic factors such as the methyltransferase DOT1-like protein (DOT1L). MLL-DOT1L and MLL-AF10 fusion proteins lead to the abnormal expression of leukemia-relevant genes, like Hox9A, and the transcriptional "prime" due by H3K79 instead of H3K4 methylation, respectively own of DOT1L and MLL1 methyltransferases, could be the prompting signal for oncogenic transformation in hematopoietic cells.[34] Other MLL fusion proteins can interact with factors that promote malignancy: MLL-AF9 fusion protein holds AF9 (also know MLLT3) capacity to recruit YEATs domain-containing 4

(YEATS4), which is upregulated in neuroblastoma and is required for the aberrant inactivation of p53106.[35] Being composed of multipotent precursor cells interested by cellular differentiation and clonal expansion following a triggering stimulus (antigens), which then turn into specialized lymphocytes, the adaptive immune system shows all the hallmarks of a cellular process that can be regulated by epigenetic pathways. Though acetylation is the most firmly recognised histone modification playing important role in these mechanisms, the immune system regulation represents a new field of application of histone lysine methylation; in particular, G9a has been found to silence specific genes in the endotoxin shock through its H3K9 dimethylation.[36]

SET-domain-containing protein 7 (SETD7) and Suppressor of Variegation 3-9 homolog 1 (SUV39H1), two other lysine methyltransferases, contribute to hyperglycaemic memory. Together with LSD1, they maintain the glucose response-related upregulation of p65 gene (a subunit of NF-kB), which is associated with diabetic vascular injury.[37]

Modulation of epigenetic proteins find application also in the newest regenerative medicine, as they can direct the differentiation of embryonic stem cells and induce a pluripotent stem cellular state starting from somatic cells. Small modulators of epigenetic proteins, including HDACs, PKMTs and Histone Lysine Demethylases (HDMs), when associated with biological techniques like transduction with transcription factors (OCT3, SOX2, MYC and KLF4), give the same cellular output, improve the reprogramming efficacy and avoid the risk of carcinogenesis. For example the G9a inhibitor BIX-01294 improves the stem cell reprogramming process in neural progenitors when in association with only OCT3 and KLF4.[38]

1.3.2. Protein Arginine Methyltrasferase (PRMT)

Protein arginine methylation is a common post-translational modification that regulates numerous cellular processes, including gene transcription, mRNA splicing, DNA repair, protein cellular localization, cell fate determination, and signaling.[39, 40, 41, 40] Notably, it was shown that about 2% of the total arginine residues isolated from rat liver nuclei are dimethylated.[42] The methylation of arginine residue does not perturb the overall positive charge of the arginine guanidinium group, but changes potential hydrogen bond interactions, since the number of added methyl groups reduces the hydrogen bond donor sites, as well as imparts bulkiness to the arginine side chain, thus influencing the protein-protein interactions both positively and negatively.

The formation of this PTM is catalyzed by a family of methyltransferases referred to as PRMT (Protein Arginine Methyltransferase). In mammal cells, nine protein arginine methyltransferases (PRMTs), divided into three different classes, have been identified so far.[43] Type I (PRMT1, PRMT2, PRMT3, PRMT4/CARM1, PRMT6 and PRMT8) and Type II (PRMT5 and PRMT7) enzymes both catalyze the formation of monomethylarginine (MMA) as an intermediate, but while Type I enzymes facilitate the formation of ω-*NG,NG*-asymmetric dimethylarginine (ADMA), Type II members lead to the formation ω-*NG,NG*-symmetric dimethylarginine (SDMA) (**Figure 1.4**). Upon certain substrates PRMT7 seems to generate only ω-*NG*-monomethylation (MMA) arginines thus being classified as a Type III enzyme too (**Figure 1.4**); [44] while PRMT9 activity has not been well characterized yet.[45] To date, no enzymes forming both asymmetric and symmetric dimethylarginines have been identified.

The different PRMT subfamilies are characterized by common structural features. All of them possess a single catalytic core region of approximately 310 amino acids, which comprises five highly characteristic signature motifs including (1) motif I (VLD/EVGXGXG), which forms the base of the SAM-

binding site and is structurally homologous to sequences found in other nucleotide-binding proteins, (2) post I (L/V/IXG/AXD/E), which is important for hydrogen bond formation to each hydroxyl of the ribose part of SAM via the carboxylate of the acidic residue, (3) motif II (F/I/VDI/L/K), which stabilizes motif I by the formation of a parallel β-sheet, (4) motif III (LR/KXXG), which forms a parallel β-sheet with motif II, and (5) the THW loop, which is close to the active site cavity and helps stabilize the N-terminal helix, which is important for substrate recognition.[16] Moreover several PRMTs possess additional domains (e.g., SH3, Zn finger, TIM barrel, and TPR) that have been suggested to diversify the substrate specificity of the enzymes and to regulate their activity (**Figure 1.6**).
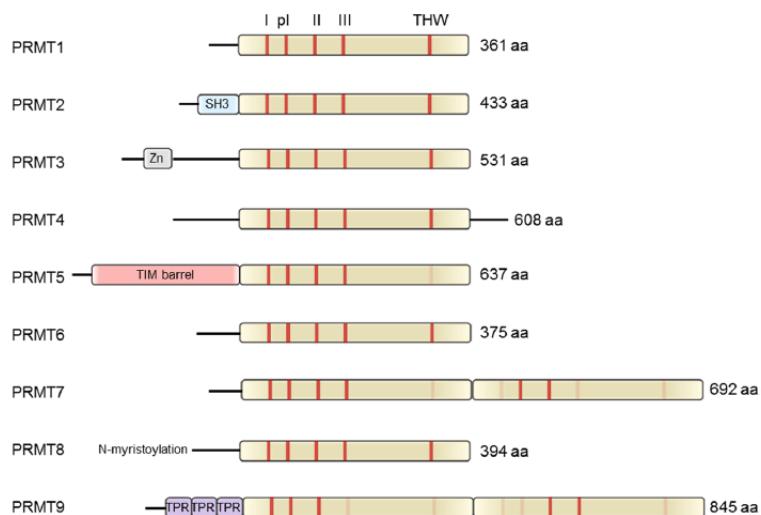

Figure 1.6. Schematic depiction of the human PRMT family. The SAM-binding methyltransferase region is highlighted in olive green. All family members contain the methyltransferase signature motifs I, post-I, II, and III and the conserved THW loop, labeled as red bars, respectively. Sequence motifs with low or no sequence similarity are depicted in light red. Abbreviations: SH3, SH3 domain; Zn, zinc finger motif; TPR, tetratricopeptide repeat

The crystal structures of several PRMTs revealed that these proteins mainly exist as homodimeric head-to-tail protein complexes (Figure 25A).[46, 47, 48, 49] It was proposed that the dimer is critical for proper substrate binding and therefore is required for activity.[35] By contrast, PRMT7, the only known type III methyltransferase, is unusual in that. It, indeed, contains two PRMT core units arranged in tandem. The catalytic core region consists of three structurally and functionally distinguishable regions as exemplified by the structure of PRMT1 (**Figure 1.7**). The most critical is the SAM-binding domain, which is highly conserved in other SAM-dependent methyltransferases.[50] This SAM-binding domain adopts a typical Rossmann fold and is followed by the β-barrel domain, which is quite unique to the PRMT family and is thought to be important for substrate binding.[51] Moreover, the β-barrel domain contains an α-helical insertion that acts as a dimerization arm. Despite the variation in amino acid sequences, the crystal structures of several PRMTs reveal highly similar general folds. In addition, key structural features such as the active site double E-loop, which is critical for guanidinium binding, as well as the SAM-binding residues and several β-strand-forming signature motifs, are conserved among all PRMTs.
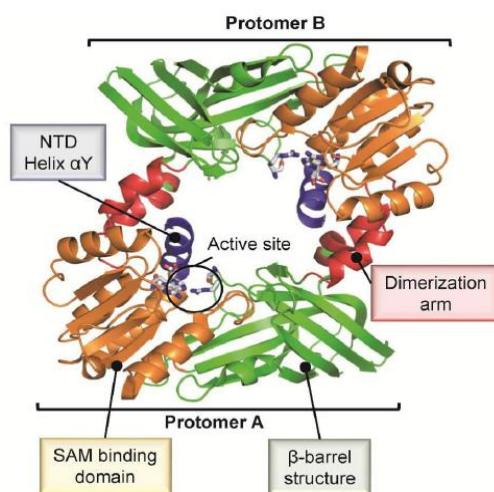

Figure 1.7. PRMT1 exists as a head-to-tail dimeric protein that comprises four characteristic functional regions, as indicated

PRMTs employ a bisubstrate mechanism,[52] transferring the methyl group of SAM to specific arginine residues in histone and nonhistone protein substrates, resulting in mono- and dimethylated arginine residues and the byproduct S-adenosyl-L-homocysteine (SAH or AdoHcy). Once activated, these enzymes endow themselves with an ordered sequential kinetic mechanism, in which SAM binds prior to the substrate.[48] The substrate binding pocket and the cofactor binding site, almost flanking each other, are joined by a narrow hydrophobic channel, large enough to allow the terminal amino group of the substrate to come within bonding distance of the cofactor. Two conserved residues (Glu-100 and Arg-54 in PRMT1) interact with the two ribose hydroxyls and carboxylate from SAM, respectively, placing the methylsulfonium group of the cofactor at the base of the channel where arginine side chain of the substrate raises. In the same manner two hydrogen bonds by other invariant glutamate residues (Glu-144 and Glu-153 of PRMT1) are established with the guanidinium side chain of the arginine substrate. These two H-bonds are critical interactions as they are supposed to concentrate the delocalized positive charge onto an only one nitrogen of the guanidinium group, leaving the lone pair of the other one available for the nucleofilic substitution (SN2 mechanism) to the methylsulfonium group(**Figure 1.8**).[51]
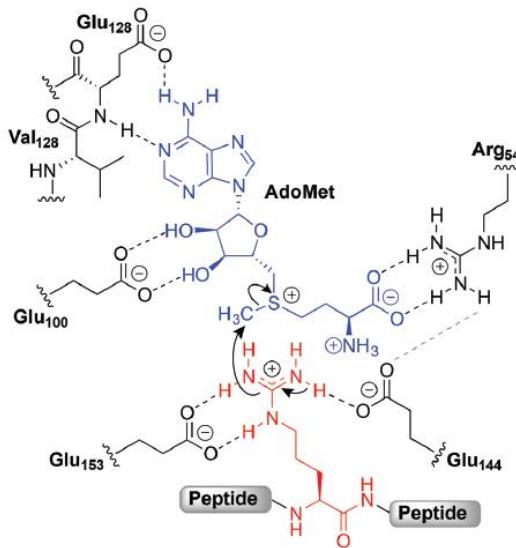
Figure 1.8. Proposed PRMT "SN2-like" bisubstrate mechanism with key active site residues shown (PRMT1 numbering scheme). AdoMet is presented in blue and the arginine residue of a target peptide is presented in red.

Interestingly, the catalytic domain of PRMT5 which catalyzes SDMA formation, is also highly similar to that of the ADMA-generating type I enzymes.[53, 54] However, the molecular basis for their distinct product formation paths is largely unknown. Recent reports proposed that a conserved phenylalanine (F327) in the active site of PRMT5 is important for directing symmetric dimethylation, while a methionine (M48) residue in Type I PRMT confers specificity toward asymmetric dimethylation.

1.3.2.1. *Physiopathological role of PRMTs*

PRMTs are generally ubiquitously expressed and govern important cellular processes that affect cell growth, proliferation and differentiation. They methylate numerous cellular protein substrates, including nuclear proteins such as transcription factors, other coregulators, and histones. The importance of this PTM to cellular growth is probably best exemplified by the fact that both PRMT1 and CARM1 mouse knockouts are embryonically lethal. The

diversity of arginine methylation sites on histone proteins provides multiple routes to directly link arginine methylation to the epigenetic regulation of gene expression. In addition, the methylation of arginine residues in several transcriptional coactivators (e.g., the histone acetyltransferases p300 and CBP) provides an indirect route to influence the epigenetic state of affected genes. Originally, it was assumed that the methylation of histone arginine residues was associated with gene activation, as the dimethylation of histone H4R3 by PRMT1 facilitates transcriptional activation by a variety of nuclear hormone receptors.[55] More recently, however, it became apparent that arginine methylation can be either an activating or a repressing mark that regulates the expression of multiple genes. Each members of PRMT family present a characteristic distribution patterns and has been demonstrate that deregulation of these enzymes seems to be implicated in the pathogenesis of several different diseases, including cancer. Specifically, PRMT1, PRMT2, PRMT3, PRMT4, PRMT5, PRMT6 and PRMT7 have been shown to be overexpressed or otherwise contribute to tumorigenesis. Below are highlighted the physiopathological roles of each isoforms.

PRMT1 was the first mammalian PRMT to be cloned[56] and is the predominant type I enzyme[57] found in all embryonic and adult tissues examined so far.[58] Due to alternative splicing, there are seven isoforms of the protein, all varying in their N-terminal domain, which are expressed in a tissue-specific manner and have distinct subcellular localization patterns.[59] PRMT1 has broad substrate specificity with over 40 targets, most being RNA processing proteins, and multiple interacting partners.[60]As it is previously mentioned, targeted PRMT1 knockout in mice results in embryonic lethality, thus showing that loss of PRMT1 activity is incompatible with life.[61] Nevertheless, alteration of PRMT1 expression, mostly upregulation, has been observed in various types of human cancers. For example aberrant methylation

of arginine 3 of Histone 4 (H4R3me2a) is positively correlated with increasing tumour grade and can be used to predict the risk of prostate cancer recurrence.[62] Although the underlying mechanisms that link H4R3me2a to active chromatin states have not been completely elucidated, is well known that the methylarginine effector molecule TDRD3 interacts with this mark.[63] Interestingly, TDRD3 (Tudor Domain Containing 3) has been identified as one of the candidate genes for a prognosis scoring system of breast cancer.[64] Higher expression of TDRD3 is linked to poor prognosis for the survival of patients with breast cancer. Additionally the methylation of the DNA repair pathway proteins MRE11[65] and p53 binding protein (53BP1)[66] might also contribute to the linkage of PRMT1 and cancer. The mammalian MRE11–RAD50–Nijmegen breakage syndrome 1 (NBS1) complex, known as the MRN complex, has crucial roles in homologous recombination repair (HRR) of DNA double-strand breaks (DSBs) that are induced by ionizing radiation or that occur during DNA replication. Methylation of MRE11 within its GAR motif regulates its exonuclease activity on double-stranded DNA. Further investigation showed that cells containing hypomethylated MRE11 displayed intra-S phase DNA damage checkpoint defects. Likewise, 53BP1, which is involved in the early events of detection, signaling and repair of damaged DNA, also has a GAR motif and is methylated by PRMT1. Blocking of 53BP1 methylation by treating cells with methyltransferase inhibitors perturbs 53BP1 localization to damaged DNA and forms fewer γH2AX foci.

Moreover, PRMT1 methylation of estrogen-receptor α (ERα) cause an activation of signal transduction pathway, through an activation of the SRC-PI3K-FAK (respectively Proto-oncogene tyrosine-protein kinase Src - Phosphoinositide 3-kinase –Focal Adhesion Kinase) cascade thus coordinating cell proliferation and survival. It has been demonstrated that hypermethylation of ERα in breast cancer might cause hyperactivation of this signalling

pathway, thus ascribing a selective survival advantage to tumor cells, even in the presence of anti-oestrogen drugs. [67]

Two independent studies have implicated PRMT1 in the regulation of leukemia development. First, it was demonstrated that PRMT1 has a role in mixed lineage leukaemia (MLL) by transducing mouse primary haematopoietic cells with the MLL–EEN gene fusion product. These transduced cells display enhanced self-renewal abilities and can form compact colony-forming unit–granulocyte, erythrocyte, monocyte, megakaryocyte (CFU–GEMM)-like colonies in vitro.[68]

A second study involved the acute myeloid leukaemia 1 (AML1; also known as RUNX1–ETO) fusion gene that encodes a transcription factor that has a role in the development of acute myeloid leukaemia.[68] PRMT1 interacts with this fusion protein and is crucial for AML1–ETO-mediated transcriptional activation, and knockdown of PRMT1 suppresses the proliferative potential of this fusion in mouse bone marrow cells. These studies suggest that PRMT1-specific inhibitors might have a therapeutic effect on leukaemia development.

PRMT2 transcripts have been found in most human tissues, with highest levels in heart, prostate, ovary and neuronal system.[69] PRMT2 was found predominantly in the nucleus and to a lower degree in the cytoplasm of mammalian cells.[70] Based on the observation that it catalyzes the formation of MMA and ADMA residues on histone H4, PRMT2 was recognized as a type I enzyme.[71] PRMT2 may act in cooperation with PRMT8, since its SH3 domain binds the N-terminal domain of the latter enzyme.[72] Some of the known targets of PRMT2, besides the aforementioned histone H4, are the STAT3 (signal transducer and activator of transcription 3) protein,[73] the estrogen receptor alpha,[74] the androgen receptor, the retinoblastoma gene product,[75] and the

heterogeneous nuclear ribonucleoprotein (hnRNP) E1B-AP5.[76] In contrast with PRMT1 knockouts, PRMT2 null mice are viable and grow normally.

PRMT2 is also implicated in breast cancer relating to its ability to act as a transcriptional coactivator of ERα. In breast cancer cell lines, the levels of PRMT2 and a splice variant, PRMT2L2, were shown to be increased in ER+ lines.

PRMT3 was first identified as a PRMT1-binding partner. It is widely expressed in human tissues and has a predominantly cytosolic subcellular localization. PRMT3 possesses a zinc finger domain that assists in its binding to ribosomal proteins, including the S2 protein of the small ribosomal subunit (RPS2). [77] Although RPS2 is the primary substrate for PRMT3, it is not its only substrate. The von Hippel–Lindau (VHL) tumour suppressor protein can interact with PRMT3, and in the presence of ARF (Alternate Reading Frame) this protein complex methylates p53,[78] however the biological importance of these interactions and methylation event is unclear.

Another tumour suppressor protein, DAL1, interacts with PRMT3 and inhibits its methyltransferase activity.[79] Epigenetic down regulation of DAL1 has been reported in several cancers,[80] and in these tumors PRMT3 displays an increased activity.

Additionally, PRMT3 with PRMT1 methylates the recombinant mammalian nuclear poly(A)-binding protein (PABPN1)[81] and has been implicated in oculopharyngeal muscular dystrophy, which is caused by polyalanine expansion in PABPN1.[82] Moreover PRMT3 has been shown to be overexpressed in human myocardial tissues from patients with coronary heart disease.[83]

CARM1 also known as PRMT4 was the first PRMT to be functionally linked to transcriptional regulation. It regulates a number of additional cellular

processes, including mRNA splicing,[84, 85] cell cycle progression,[86] and the DNA damage response.[87] CARM1 achieves these functions by methylating histones,[88,89] transcriptional factors [89] , co-regulators, [90, 91, 92, 93] splicing factors and RNA polymerase II.[94] CARM1 levels are elevated in castration-resistant prostate cancer,[95, 96] as well as in aggressive breast tumours. It also has a role in determining ERα-dependent breast cancer, cell differentiation and proliferation.[97] The aggressive breast tumours that overexpress CARM1 also have high levels of the oncogenic co-activator AIB1 (amplified in breast cancer) (also known as NCoA3). Overexpression of AIB1 has been reported in breast cancer and is clearly associated with poor prognosis.[98] Importantly, CARM1 methylates AIB1, thereby regulating its activity and stability. CARM1 recruitment to ERα-regulated promoters relies on the presence of AIB1, and is essential for estrogen induced proliferation of the MCF-7 breast cancer cell line.[99] Thus, the oncogenic properties of AIB1 may be attenuated by inhibiting CARM1 activity. The histone acetyltransferases (HATs) CREB-binding protein (CBP; also known as CREBBP) and p300 are also arginine methylated by CARM1.[100] Notably, methylation increases the HAT activity of CBP and stimulates its autoacetylation. Whereas methylation of p300-R2142 inhibits its interaction with glucocorticoid receptor-interacting protein 1 (GRIP1; also known as NCoA2)130, which is required for responses to DNA damage. The absence of CARM1 methyltransferase activity leads to the loss of cell cycle arrest in response to DNA damage. CARM1-mediated methylation of p300, specifically at R754 in the KIX domain, is required for the induction of cell cycle regulators such as p21 and growth arrest and DNA damage-inducible protein-α (GADD45α).

CARM1 is also recruited to the promoter of E2F1 (which encodes a cell cycle regulator) as a transcriptional co-activator. CARM1-mediated upregulation of E2F1 expression might also contribute to the oncogenic DNA damage and tumour suppressive apoptosis that are often observed in cancer. CARM1 is

also an important positive modulator of WNT–β-catenin-driven transcription.[101] Depletion of CARM1 expression in colorectal cancer cells suppresses clonal survival and anchorage-independent growth,[90] supporting the observations from clinical samples which show that 75% of colorectal cancers have CARM1 over- expression.[102]

PRMT5 was the first PRMT type II to be identified. PRMT5 is widely expressed in mammals, and more extensively in heart, muscle and testis. PRMT5 is located both in the nucleus and in the cytoplasm.[103] Interestingly, PRMT5 is localized in the Golgi apparatus (GA) and its loss disrupts the GA structure, thereby suggesting that this enzyme plays a role in the maintenance of GA architecture. In addition, PRMT5 is involved in the maintenance of the spliceosome integrity.[104] PRMT5 targets histones, transcriptional elongation factors, chromatin remodelers and co-repressors.[105] In mice, loss of PRMT5 results in early embryonic lethality.

H3R8me2s and H4R3me2s are the key repressive histone methylation marks that are deposited by PRMT5. It is in this context as a transcriptional repressor that PRMT5 has oncogene-like properties because of its ability to repress the expression of tumour suppressor genes. Indeed, the levels of ST7 and NM23 (also known as NDKA) expression are clearly reduced when PRMT5 is overexpressed.[106] The transcriptional repressor functions of PRMT5 are also crucial for the process of *epithelial–mesenchymal transition* (EMT), which has an important role in tumour progression. As a consequence of EMT, cancer cells can migrate and invade much more efficiently, and this process probably has an active role in metastasis. A hallmark of EMT is the loss of E-cadherin expression; E-cadherin expression is actively repressed by the transcription factor SNAIL, and PRMT5 interacts with SNAIL through a bridging molecule, AJUBA.[107] Importantly, small interfering RNA (siRNA)-mediated PRMT5 knockdown results in increased E-cadherin expression. These data

indicate that PRMT5 is a crucial co-repressor of SNAIL. In addition, PRMT5 overexpression promotes anchorage-independent cell growth, supporting the idea that PRMT5 might be an oncoprotein. Consistently, PRMT5 levels are elevated in a number of cancer types.

High expression levels of the tumour suppressor protein programmed cell death protein 4 (PDCD4) are generally correlated with a better outcome in a number of different cancer types.[108] However, there are exceptions to this rule: under certain circumstances PDCD4 displays oncogenic properties, and it seems that PRMT5 is a key player in switching this function. PRMT5 interacts with PDCD4 and methylates GAR motifs that are present in its N terminus. In an orthotopic breast tumour model, overexpression of PRMT5 enhances tumour growth; this depends on the presence of PRMT5 enzymatic activity and an intact PDCD4 methylation site.[109] Thus, high levels of PDCD4 are protective against tumorigenesis, unless PRMT5 is also highly expressed.

Additionally, PRMT5 was originally described as JAK2-binding protein 1[110] and the functional importance of this interaction in cancer has been realized. Most patients with myeloproliferative neoplasms express a constitutively activated form of JAK2: JAK2-V617F150. This mutant form of JAK2 interacts with PRMT5 more strongly than does the wild-type form, and furthermore JAK2-V617F phosphorylates PRMT5.[111] Phosphorylation of PRMT5 greatly diminishes its arginine methyltransferase activity by blocking its interaction with its co-regulator MEP50. Knockdown of PRMT5 in human CD34+ cells increases colony formation and PRMT5 overexpression significantly decreases colony formation, providing evidence that the phosphorylation of PRMT5 (and thus its inactivation) contributes to the JAK2-V617K-induced myeloproliferative phenotype. In a second example of PRMT5–MEP50 regulation, it was found that cyclin D1–CDK4 phosphorylates MEP50 at T5, and this phosphorylation increases the activity of PRMT5.[112] This increase in PRMT5 activity results in an increase in the

repressive epigenetic marks deposited by PRMT5 and in the silencing of CUL4A and CUL4B expression, thus providing a key mechanistic insight into the neoplastic activities of the cyclin D1–CDK4 kinase. CUL4A and CUL4B are E3 ubiquitin ligases that direct the degradation of the replication licensing protein CDT1. Thus, increased PRMT5 activity will cause the accumulation of CDT1 and will facilitate DNA replication, which activates DNA damage checkpoints and predisposes to malignant transformation

PRMT6 is the smallest family member (316 amino acids) and is localized predominantly in the nucleus. PRMT6 functions are not completely understood, however, a role for PRMT6 in the regulation of cell proliferation and senescence, through transcriptional repression of tumor suppressor genes, has been reported.[113] Its primary targets are histone H3 (generating H3R2me2a)[114, 115] and DNA polymerase β.[116] As a transcriptional repressor, PRMT6 suppresses the expression of thrombospondin 1 (TSP1; also known as THBS1), which is a potent natural inhibitor of angiogenesis and endothelial cell migration.[117]

PRMT7 is present mainly in thymus, dendritic cells and testis. It has both nuclear and cytosolic localizations. PRMT7 is involved in the modulation of sensitivity to DNA damaging agents,[118] methylation of male germline imprinted genes, and embryonic stem cell pluripotency.[119]
PRMT7-mediated methylation of H2AR3 affects DNA damage repair. The effects of inhibiting PRMT7 are complex and could result in resistance to certain chemotherapeutics.[120]

1.4. Chemical probes – inhibitors of PRMTs

Since PRMTs affect a plethora of different target genes, it is unsurprising that, when dysregulated, they play an important role in human disease. In fact, dysregulation of PRMT activity has been causally linked to the development and progression of numerous cancers, as well as to viral replication and cardiovascular disease. Therefore, PRMTs constitute promising targets for drug discovery, and inhibitor development is the frontline of current PRMT research. Selective PRMT modulators not only serve as potential therapeutics but also are useful chemical probes for further investigations of PRMTs physiopathological roles.

1.4.1. SAM analogues as inhibitors of PRMTs

One of the earliest described PRMT inhibitors to be discovered is autogenerated by the enzyme during catalysis. As described above (**paragraph 1.3.**), the methyl donor substrate SAM is converted to SAH (**1**) (**Figure 1.9**), which represents a potent feedback inhibitor of PRMT activity, and accumulates as an inevitable byproduct of protein methylation. However, it can be degraded into adenosine and homocysteine by SAH hydrolase. Therefore, in cellular studies, SAH hydrolase inhibitors such as adenosine dialdehyde (**2**) are used as SAM-dependent inhibitors by blocking the degradation of SAH. Moreover, SAM analogues, such as methylthioadenosine (**3**) and sinefungin (**4**), also function as general PRMT inhibitors (**Figure 1.9**). The SAM analogues and SAH exhibit limited specificity due to their structural homology to SAM. Thus, they inhibit numerous other SAM dependent methyltransferases, thereby displaying limited specificity among all methyltransferase enzymes.
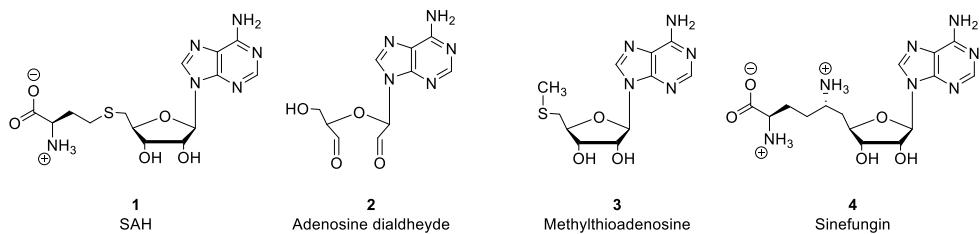
Figure 1.9. SAM analogues, a non selective inhibitors of PRMTs

1.4.2. Small molecule inhibitors of PRMTs

To obtain PRMT-selective inhibitors, several approaches, including both virtual and high-throughput screening, as well as substrate analogue inhibitor design were conducted. One of the first high-throughput screens directed against the yeast type I PRMT, Hmt1p, resulted in the identification of several arginine methyltransferase inhibitors, denoted as AMI.[121] Most of these inhibitors were highly nonspecific and also blocked the activity of the protein lysine methyltransferases. AMI-1 (**5**), however, a symmetric naphthalenesulfonate molecule, inhibited PRMT1 with an IC_{50} of 8.8 μM, but did not diminish the activity of distinct lysine methyltransferases (**Figure 1.10**). Subsequently the EMCL Group (Epigenetic Medicinal Chemistry Laboratory) of Professor Gianluca Sbardella, where I performed the current PhD project, being interested in discovery of small-molecule modulators of epigenetic targets, [122, 123, 124, 125, 126, 127,128, 129, 130, 131, 132, 133,134, 135, 136] focused its attention on AMI compounds and realized that AMI-1 should be optimized as it is likely to have low bioavailability and would probably not penetrate the blood–brain barrier due to the bisanionic structure. Therefore, were synthesized a number of derivatives, structurally related to AMI-1 and was found that the substitution of the sulfonic groups, involved in pleiotropic interactions, with the bioisosteric carboxylic group together with the formal shift of the ureidic function from the C-7 to the C-6 position of the naphthalene ring, resulted in the generation of even more potent SAM

31

noncompetitive inhibitors, that showed specificity for PRMTs over PKMTs (**Figure 1.10**).[137]
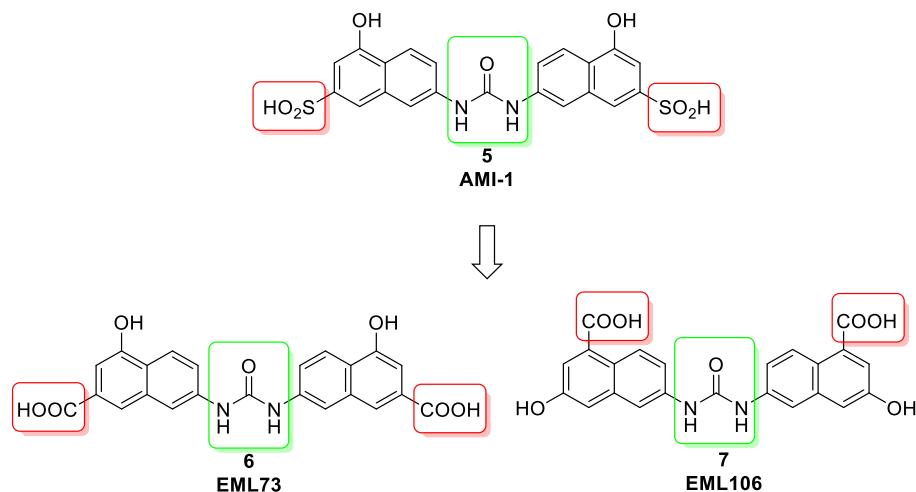
Figure 1.10. Inhibitors of PRMTs

Moreover, pursuing our efforts toward the identification of potent and selective PRMT inhibitors, we replaced the hydroxynaphthalene moiety of **EML73** with the bioisosteric indole framework, with the aim to obtain a more versatile and synthetically handy scaffold for the future development of more potent inhibitors (**Figure 1.11. A**).[126] Thus, we prepared a series of ureido acetamido indole derivatives (dubbed Uracandolates) that, however, showed a peculiar activity against PRMTs. They selectively enhanced the methylation activity of CARM1 both in in vitro and cellular settings (**Figure 1.11. B**). We also assessed the activity of **EML105** over a panel of PRMTs and SET7, a lysine methyltransferase. This compound together with a positive modulation of CARM1, showed a scarcely, but defined inhibitory activity against PRMT3, as it is illustrated in **Figure 1.11**. Notably, we have taken into account this data for developing PRMT3 inhibitors (see Chapter 2, Aim of the work).

A)

B)

	% inhibition (at 50 μM, mean ± SD)				
	PRMT1 (H4)	PRMT3 (GAR)	CARM1 (H3)	PRMT6 (H3)	SET7 (H3)
EML105	27.7 ± 0.5	54.7 ± 1.7	-175.5 ± 2.6	-19.3 ± 0.3	10.2 ± 0.2
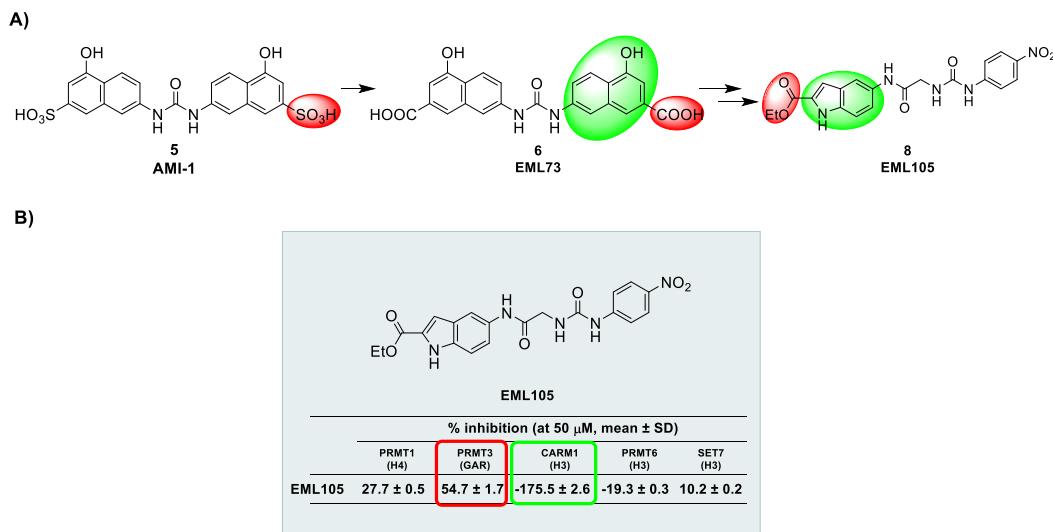
Figure 1.11. A) Structural optimization of **EML73** that led to **EML105**, an activator of CARM1. **B)** Activity of **EML105** over a panel of PRMTs and SET7, a lysine methyltransferase.

By Fragment-based virtual screening, Spannhoff and coll. identified RM65 (**9**) (**Figure 1.12**) as a cell-permeable PRMT1 inhibitor.[138] Docking studies suggest that compound **9** is a competitive inhibitor, occupying both the SAM binding and the substrate arginine binding sites. RM65 was also shown to reduce histone H4R3 methylation in HepG2 cancer cells at concentrations above 100 μM. Using a similar virtual docking and pharmacophore-based filtering approach, the Jung group identified the diamidine stilbamidine (**10**) and allantodapsone (**11**) as PRMT1 inhibitors (**Figure 1.12**).[138] Both compounds are competitive for the protein substrate, but not for the cofactor SAM. Moreover, these inhibitors are active in a functional assay of estrogen receptor activation and also reduced the cellular methylation of R3 in histone H4 at concentrations below 50 μM, while having minor effects on the lysine methylation at H3K4. Further optimization resulted in the generation of even more potent inhibitors, such as thedapsone derivative **12**, which inhibits PRMT1 with an IC$_{50}$ of 1.5 μM (**Figure 1.12**).[139]
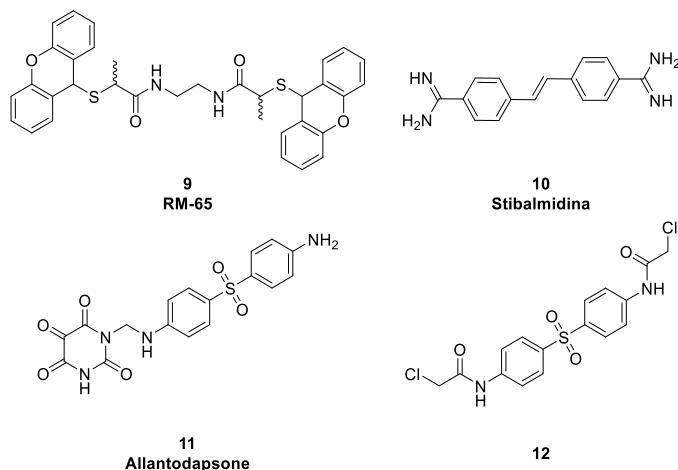
Figure 1.13. Non selective inhibitors of PRMT identified by Fragment-based virtual screening and synthetic optimization.

1.4.3. Bisubstrate inhibitors of PRMTs

The ability of PRMTs to chemoenzymatically generate an effective bisubstrate analogue inspired the development of defined bisubstrate derivatives (**Figure 1.13**). In this respect, the partial-bisubstrate analogue **13**, which comprises an arginine containing peptide fragment that was linked to the amino acid moiety of SAM, was shown to block PRMT1 activity with an IC_{50} of 14 μM (**Figure 1.13**).[140] This compound shows limited PRMT selectivity and also blocks PRMT4 and PRMT6 with similar efficiency. In addition, Dowden and colleagues reported the development of SAM derivatives conjugated to a guanidinium group via varying carbon linkers (**14−16**).[141] Comparison of the generated derivatives revealed that a four carbon spacer, present in **15** between the guanidinium group and the SAM analogue, is most suited for efficient PRMT1 inhibition (IC_{50} = 2.9 μM). Though the selectivity of this inhibitor for other PRMTs was not evaluated, it did not show substantial inhibitory activity against the lysine methyltransferase SET7. The Martin group developed a

partial bisubstrate inhibitor where the SAM adenosine moiety is connected to the guanidinium group (**17**).[142] Interestingly, **17** was more potent than the previously described (partial) bisubstrate inhibitors with IC_{50} values of 1.3 μM, 560 nM, and 720 nM for PRMT1, PRMT4, and PRMT6, respectively. Moreover, this compound did not display any measurable inhibitory effect on the lysine methyltransferase G9a; however, it also did not show any inhibitory effect on cell proliferation using MCF7 and Caco2 cells.
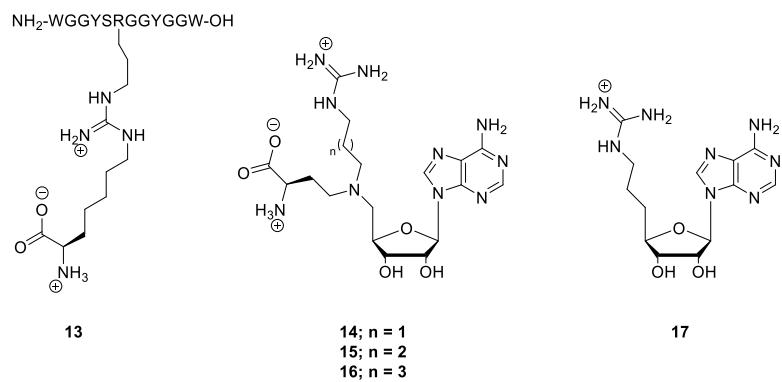
Figure 1.13. SAM analogue inhibitors

1.4.4. Selective inhibitors of type I PRMTs

Although bisubstrate analogues represent interesting tools to analyze PRMTs activity, they comprise several limitations, including lack of selectivity within the PRMT family. Therefore, recently, great efforts have been taken to yield potent and isozyme-selective PRMT inhibitors.

In 2014, Yan et al. reported compound **18**, a diamidine containing PRMT1 inhibitor (**Figure 1.14**).[143] It inhibited PRMT1 with an IC_{50} of 9.4 μM and was selective for PRMT1 over CARM1 (more than 42-fold), PRMT5 (around 18-fold), and PRMT6 (around 30-fold). Compound **18** inhibited the proliferation of several leukemia cell lines. It was also found that cell lines derived from Down syndrome patients and MLL-AF9 patients (CMY, CHRF-288-1, and

MOLM-13 cells) were more sensitive to compound **18** treatment than other cell lines tested (HEL, Jurkat, and HL-60).

In 2015, Eram et al. discovered MS023 (compound **19**)[144] as a potent and cell-active inhibitor of type I PRMTs, including PRMT1, -3, -4, -6, and -8 (**Figure 1.14**). In biochemical assays, MS023 was highly potent and selective for type I PRMTs, over a broad range of epigenetic modifiers including type II and type III PRMTs, PKMTs, DNMTs, histone lysine demethylases, and methyl-lysine and methylarginine reader proteins. Moreover, after treating MCF7 and HEK293 cells with compound **19** for 2 days, was observed a significant decrease in global levels of arginine asymmetric dimethylation (Rme2a) and a concurrent increase in global levels of arginine monomethylation (Rme1) and arginine symmetric dimethylation
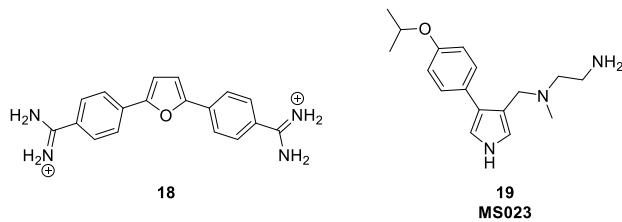
Figure1.14. Compound **18** is selective inhibitors of PRMT1, whereas a compound **19** is a selective inhibitor of type I PRMTs

1.4.5. Allosteric inhibitors of PRMT3

With respect to PRMT3, Siarheyeva et al. employed a library screening approach to identify selective inhibitors targeting this enzyme and reported the discovery of 1-(benzo[d] [1,2,3]thiadiazol-6-yl)-3-(2-cyclohexenylethyl)urea (compound **20**, **Figure 1.15 A**), the first selective inhibitor of PRMT3.[144b] Notably, this inhibitor is noncompetitive with respect to both SAM and the peptide substrate, but it occupies an allosteric binding site revealed by the X-ray crystal structure of compound **20** in complex with PRMT3. There the benzothiadiazole moiety of **20** forms a buried hydrogen bond with T466, the

urea group forms hydrogen bonds with the side chains of E422 and R39, whereas the cyclohexyl arm extends out of the allosteric pocket and makes hydrophobic interactions with the α-Y-helix, inducing conformational constraints on this motif that prevent the formation of a catalytically competent state (**Figure 1.15 B**). Subsequent mechanism of action (MOA) studies confirmed that this inhibitor is noncompetitive with both the peptide substrate and the cofactor.

A)

B)
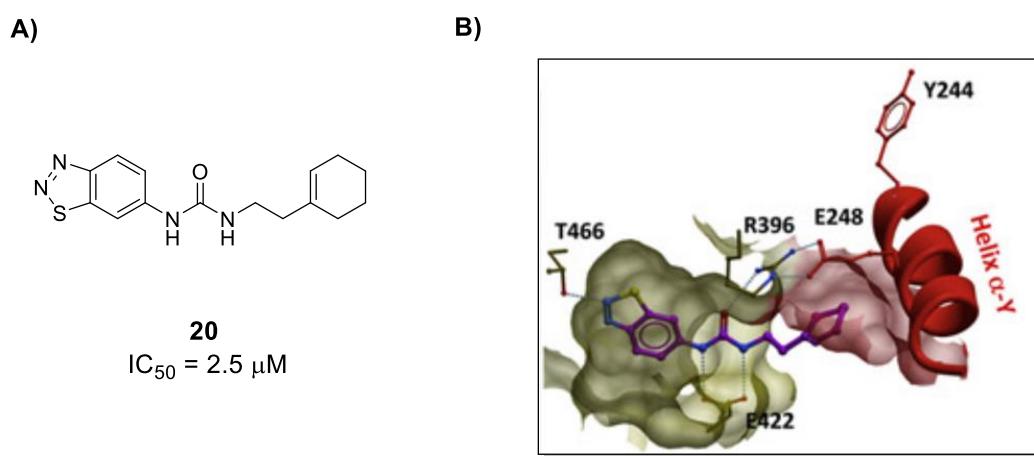
Figure 1.12. A) PRMT3 inhibitor. **B**) Key interactions of compound **20** within the allosteric binding pocke of PRMT3.

To improve the potency of compound **20** (PRMT3, IC_{50} = 2.5 μM) and demonstrate that potent and selective inhibitors can be generated by exploiting the allosteric binding site of PRMT3, Jin and collaborators in 2013 reported extensive structure-activity relationship studies of benzothiadiazolic compounds, which resulted in discovery of several PRMT3 inhibitors with submicromolar potencies. [145] Below are analyzed in detail the efforts of Hit to Lead optimization of this series of compounds, given the important role that these SAR studies played for us regarding the development of PRMT3 indole inhibitors (see Chapter 2 Aim of the work).

1.4.5.1 SAR of benzothiadiazole derivatives

The replacement of the uncommon cyclohexenylethyl group of compound **20** with a more common and potentially stable group, the cyclohexylethyl moiety (Compound **21**, IC_{50} = 1.0 µM), led to an improvement of PRMT3 inhibition (**Figure 1.13**).

20
IC_{50} = 2.5 µM

21
IC_{50} = 1 µM
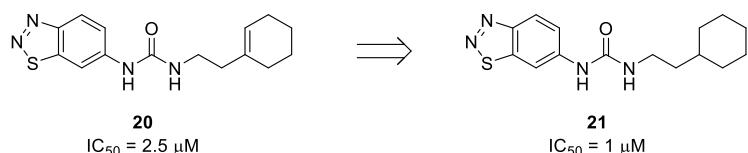
Figure 1.13. The replacement of cyclohexenylethyl group with cyclohexylethyl moiety led to an improvement of PRMT3 inhibition.

On the other hand the replacement of the benzothiadiazole ring with other fused bicyclic aromatic groups (**Figure 1.14**), like benzothiazole (compound **22**), benzotriazole (compound **23**) and indazole (compounds **24** and **25**), led to total loss of potency, suggesting that, in addition to making the hydrogen bond interaction with T466, the benzothiadiazole moiety is particularly well accommodated by the PRMT3 allosteric binding pocket due to its unique electronic nature and steric fit. Interestingly, has not been taken into account the replacement effect of the benzothiadiazole ring with the indole, which is considered a privileged scaffold for drug discovery.

22 IC_{50} > 100 µM

23 IC_{50} > 100 µM

24 IC_{50} > 100 µM

25 IC_{50} > 100 µM

21
IC_{50} = 1 µM
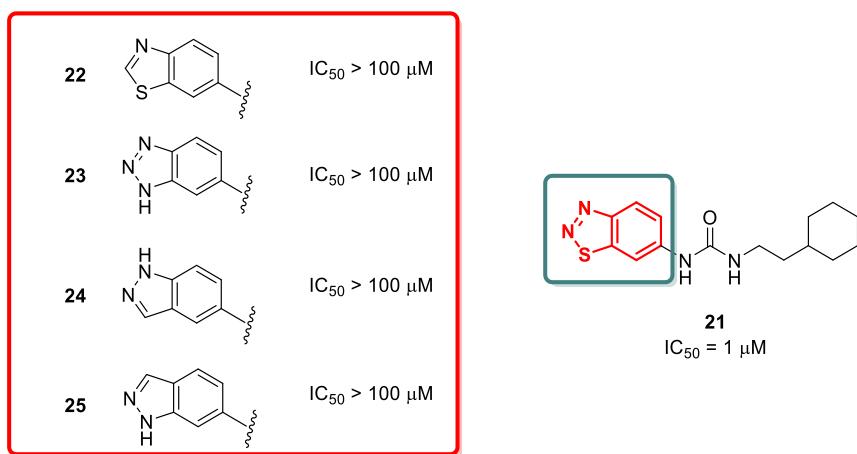
Figure 1.14. Replacement of benzothiadiazole moiety with other fused bicyclic aromatic groups

Further investigations involved the middle urea moiety of compounds **21**. The X-ray crystal structure of the inhibitor **20**−PRMT3 complex reveals that the carbonyl group of the middle urea moiety forms a hydrogen bond with the guanidinium group of R396 whereas the two amino groups of the urea moiety form two hydrogen bonds with the carboxylate of E422 (glutamate 422) (**Figure 1.12 B**). The replacement of the ureidic moiety with urea bioisosteres and more rigid analogues (**Figure 1.15**), as diaminosquarate (compound **26**), cyanoguanidine (compound **27**) and aminotriazole (compound **28**), led to total loss of potency, suggesting that this group is ideal for interacting with the allosteric binding pocket of PRMT3.
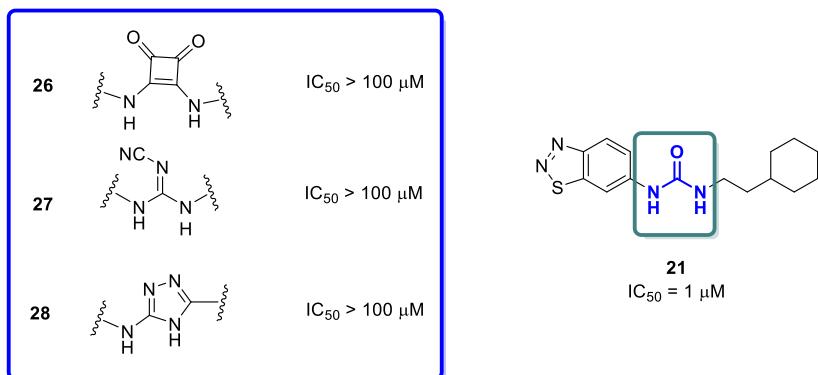
Figure 1.15. Effect of the substitution of the ureidic moiety.

As far as the cyclohexylethyl group is concerned, replacements of this moiety are well tolerated. Indeed aliphatic or aromatic moieties, with varying steric and electronic characteristics, were introduced instead of the cyclohexylethyl group; leading in same case to a significant increase in potency (for example compounds **42** and **43, Figure 1.16**).
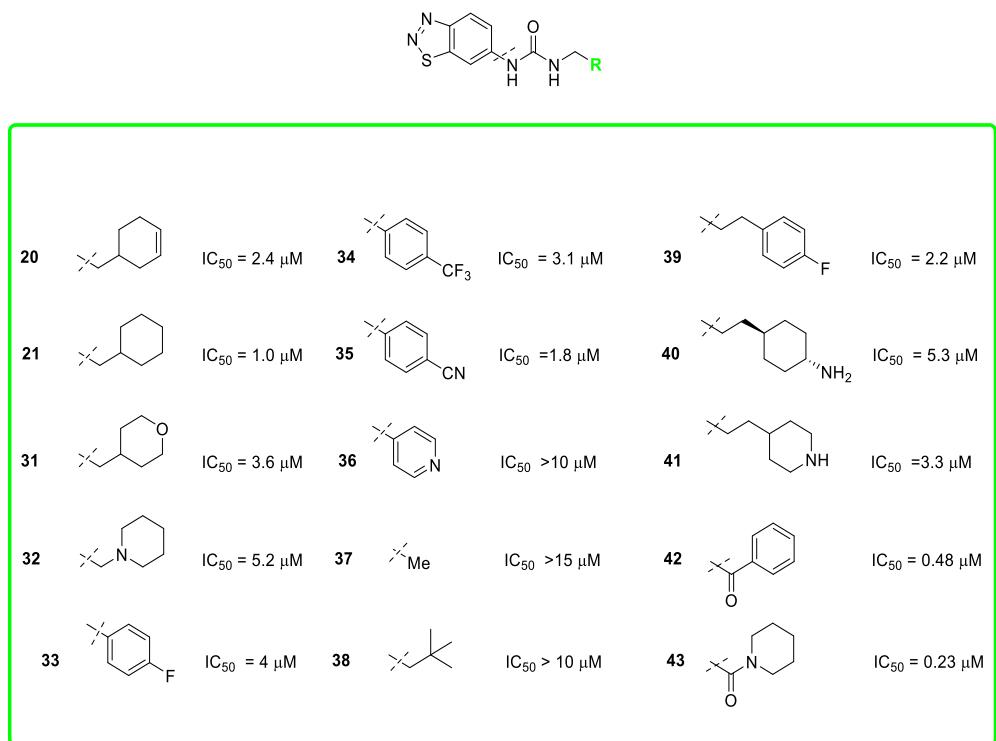
Figure 1.16. Effect of the replacement of the cyclohexyl ring

The replacement of the cyclohexyl ring with different heterocycles such as tetrahydropyran (compound **31,** IC_{50} = 3.6 µM), and 1-piperidine (compound **32,** IC_{50} = 5.2 µM) generally resulted in a slight decrease in potency, allowing to conserve a good potency. Similarly, the introduction of *para* substituted aromatic ring (compounds **33**, **34** and **35** with IC_{50} values of 4, 3.1 and 1.8 µM respectively) gave slightly lower or either similar potency compared with the cyclohexylethyl group (compound **21,** IC_{50} = 2.4 µM). On the other hand the introduction of pyridine (compound **36**, IC_{50} > 10 µM)**,** methyl (compound **37g,** IC_{50} > 15 µM) or tert-butyl groups (compound **38,** IC_{50} > 10 µM) led to a significant decrease in potency. Interestingly, the introduction of a longer chain (three carbons versus two carbons) between the ureidic moiety and the aliphatic or aromatic groups allowed to conserve a good potency against PRMT3 (compounds **39 – 41,** IC_{50} = 2.2, 5.3, 3.3 µM), suggesting that the

allosteric pocket of PRMT3 is able to accommodate compounds endowed an endowed a greater length.

Interestingly, the introduction of an additional carbonyl group on the ethylene chain led to the identification of the most active compounds, **42** ($IC_{50} = 0.48$ µM) **and 43** ($IC_{50} = 0.23$ µM), which contain a ketone and amide moieties respectively.

In 2015 the group of Jin, pursuing the efforts toward the identification of potent and selective PRMT3 allosteric inhibitors, performed further structural optimization of benzothiadiazole derivatives revealing a potent, selective, and cell-active allosteric inhibitor of PRMT3 (compound **44, Figure 1.17**).[146] This compound is highly selective for PRMT3, and detailed structural analyses showed that binds to an allosteric pocket at the base of the dimerization arm between the two PRMT3 subunits, as the others benzothiadiazolic compounds. Compound **44** engages PRMT3 and potently inhibits its methyltransferase activity in cells. It is also bioavailable and suitable for animal studies. This well characterized chemical probe is an excellent tool to further study the role of PRMT3 in health and disease.
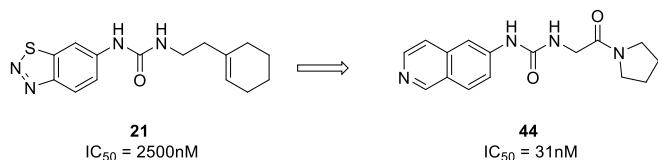
Figure 1.17. Optimization of the initial hit compounds revealed inhibitor **20**

1.4.6. Selective inhibitors of CARM1

High-throughput screening efforts led to the identification of pyrazole and Benzo[d]imidazole derivatives as selective PRMT4 inhibitors (**Figure 1.18**).[147, 148] Based on these derivatives, Hits to Lead efforts were done to further improve the potency of these derivatives.
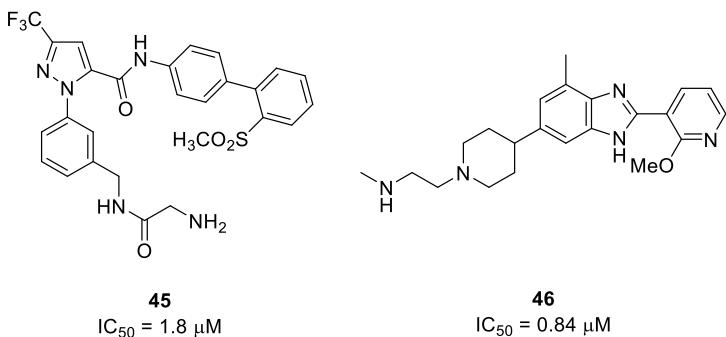
Figure 1.18. Pyrazole (45) and Benzo[d]imidazole (**46**) compounds identified as selective inhibitors of CARM1 by High-throughput screening

Inspired by the structural information of these compounds (Pyrazole and Benzo[d]imidazole derivatives), we set out to design pyrrole inhibitors of CARM1 (Chapter 2, Aim of the work). Therefore, given the important role that the SAR information of Pyrazole and Benzo[d]imidazole derivatives have played for us, below are analyzed in detail their structure activity relationships.

1.4.6.1. SAR of pyrazole inhibitors of Coactivator Associated Arginine Methyltrasferase 1 (CARM1)

As it is previously mentioned Purandare et al. identified a pyrazole amide inhibitor of CARM1 by High-throughput screening (Compound **45**, **Figure 1.18**). This compound showed a modest activity in CARM1 mediated methylation assay, therefore, to further improve the potency of this compound structural optimization efforts were performed.[149, 150, 151]

First the glycinamide portion of **45** was examined. As it is shown in **Figure 1.19,** α-amino acids (compound **47**, $IC_{50} = 0.16$ μM) were found to be preferred over either β (compound **50**, $IC_{50} > 20$ μM) or α, α' di-substituted amino acid amides (compound **49**, $IC_{50} > 30$ μM). Within the subset of analogs prepared from α -amino acids, the (S)-enantiomer (**47**, $IC_{50} = 0.16$ μM) was about 150-fold more active than the corresponding (R)-enantiomer (**48**, $IC_{50} = 20.1$ μM), suggesting that the (S)-alanine benzyl-amide is optimal for CARM1

enzymatic activity. Moreover, simple modifications as an introduction of ethyl group between the phenyl ring and the aminoacidic portion (**51**, $IC_{50} > 20$ μM) as well as the methylation of the terminal amino group of (S)-alanine (**52**, $IC_{50} > 20$ μM) were not tolerated.
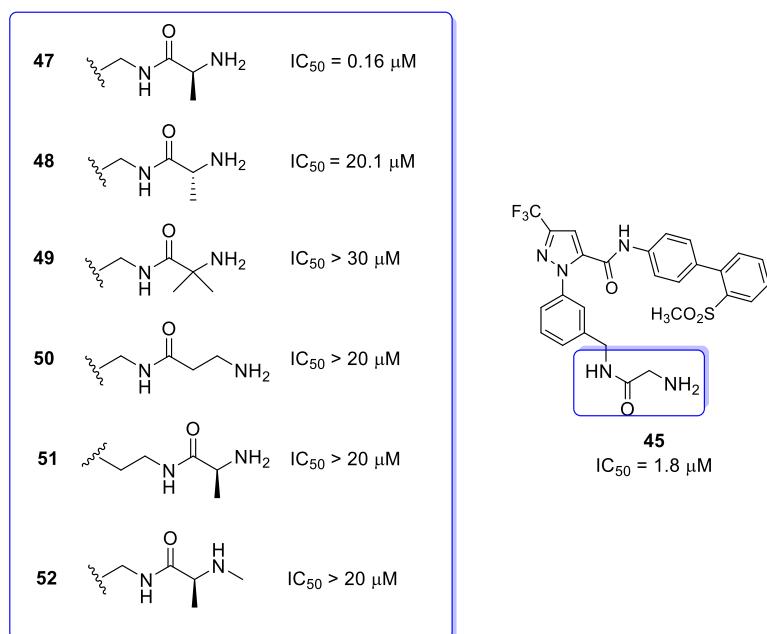
Figure 1.19. Substitution of glycinamide portion

Next, retaining the (S)-alanine benzyl-amide portion on the phenyl ring (Compound **47**, **Figure 1.19**) was evaluated the effect of the substitution of the biphenyl amide moiety (**Figure 1.20**). The introduction of a benzyl-amide portion (compound **54,** $IC_{50} = 0.08$ μM) was found to be superior to phenyl-amide portion (compound **53**, $IC_{50} = 1.32$ μM) in term of potency, indeed compound **54** was about 20-fold more potent than compound **45** ($IC_{50} = 1.8$ μM) and **53** ($IC_{50} = 1.32$ μM). Interestingly, the functionalization of the benzylamide group had different effect on the enzymatic activity. The *o*-OMe insertion on the aromatic ring (Compound **55**, $IC_{50} = 0.06$ μM) provided an improvement of activity, whereas the introduction of *o*-NH$_2$ (compound **57**, $IC_{50} > 10$ μM) or *m*-OMe (compound **57,** $IC_{50} = 0.37$ μM) moieties had a negative effect, reducing the CARM1 activity. In addition, alternative linkage

43

for the benzyl-amide side chain (compound **59,** $IC_{50} = 1.1$ µM) was evaluated, assessing that a major length of this moiety was detrimental for the activity.

Compounds **54** ($IC_{50} = 0.08$ µM) and **55** ($IC_{50} = 0.06$ µM) bearing the benzylamide group at C-5 position of pyrazole ring provided the best activity toward CARM1 in vitro assay, however these compounds lack of cellular effect on CARM1 due to a modest permeability (PAMPA assay, Pc of **54** < 15 nm/s). Therefore in order to improve the pharmacokinetic properties of pyrazolic derivatives the benzylamide moiety was replaced with the 1, 3, 4-oxadiazole ring, a potentially bioisosteric group of the amide moiety (Compounds **60** and **61**). This modification improved the cytopermeability (PAMPA assay Pc of **60** $= 106$ nm/s) versus benzylamide, also leading to a slight improvement of the in vitro potency (Compound **61**, $IC_{50} = 0.04$ µM), suggesting that the introduction of bulky substituents was well tolerated by the CARM1 binding site. However, also these compounds didn't revealed activity against CARM1 in cell-based assays.

Figure 1.20. Replacement of biphenyl- amide moiety
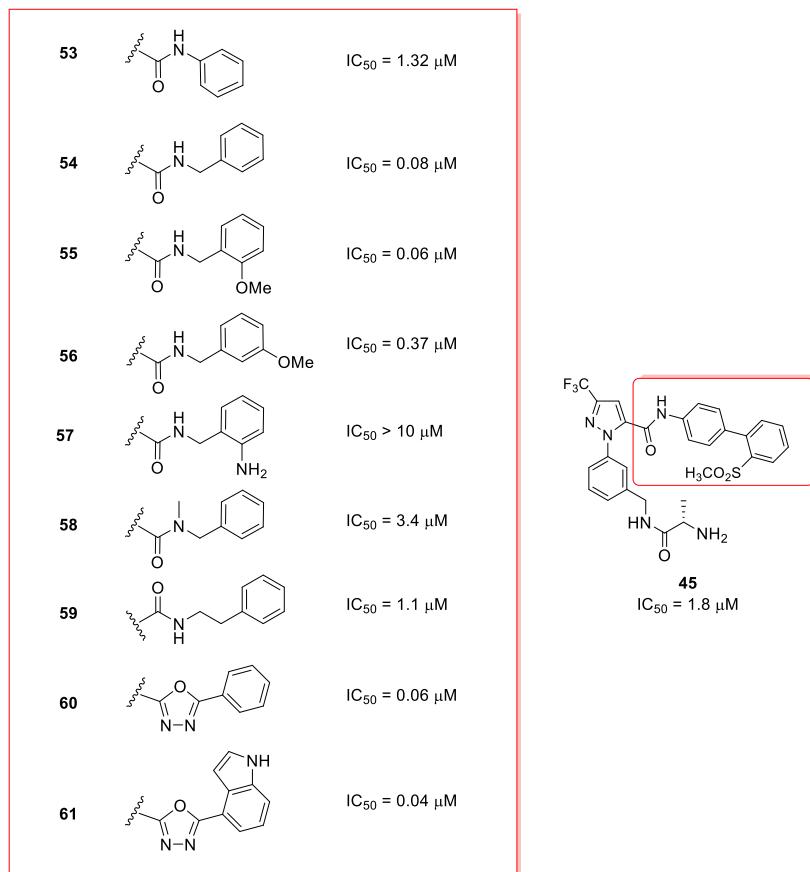
Further SAR exploration centered on compounds **54** and **55** (**Figure 1.21**) involved the replacement of the central core phenyl ring with various heterocycles (**Figure 1.21**). 2, 4- and 2, 5-disubstituted thiophene analogues (compounds 62 and **63** respectively) were synthesized. These modifications didn't lead to a CARM1 cellular inhibition, however, compared with **55** (IC_{50} = 0.06 μM), an equipotent PRMT4 inhibition in vitro was observed for compound **62** (IC_{50} = 0.06 μM). Other heterocycles were introduced instead of the phenyl ring, like thiazole (compound **64**, IC_{50} = 2.20 μM) and bicyclic benzofurane (compounds **65**, IC_{50} > 10 μM). The thiazole analogue was 36-fold less active than **55**, whereas compound **65** was devoid of activity.

Figure 1.21. SAR exploration of central core phenyl ring
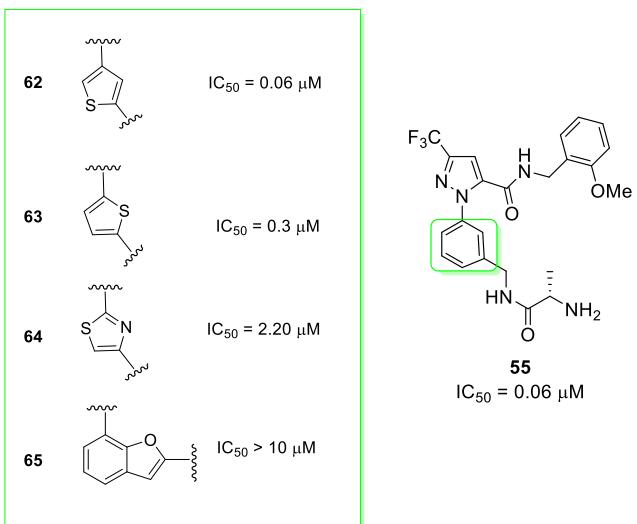
In conclusion, was investigated the role of the 3-trifluoromethyl-pyrazole scaffold. Retaining the same substituents and dispositions of compound **60** (**Figure 1.22**), various five-membered heterocycles were evaluated (Compounds **66 – 68, Figure 1.22**). The 3-trifluoromethyl substituted pyrazole was found to be the preferred core from other five-membered-heterocycles, suggesting that this scaffold is particularly well accommodated by the PRMT4 binding pocket, due to its unique electronic nature and steric fit. Notably, it was not investigated the effect of the pyrrole ring, probably due to a not immediate synthetic accessibility.

Figure 1.22. SAR exploration of 3-trifluoromethyl- pyrazole scaffold
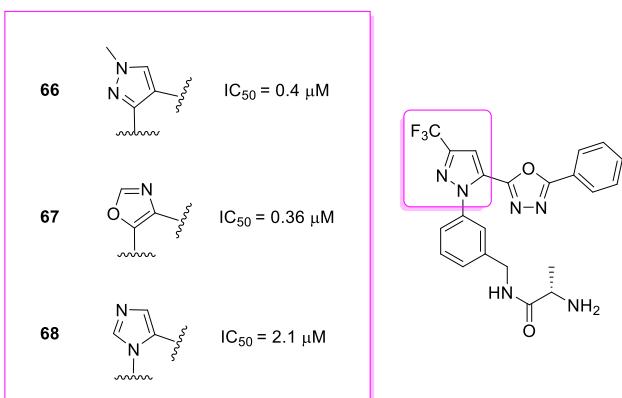
1.4.6.2. SAR of Benzo[d]imidazole inhibitors of Coactivator Associated Arginine Methyltrasferase 1 (CARM1)

During the High-throughput screening campaign, Wan et al. in 2009 identified a benzo[d]imidazole derivative as Hit with modest activity (IC_{50} 0.84 μM) in the CARM1 mediated methylation assay (**Figure 1.18**, compound **46**). Based on this hit, was started Hit to Lead optimization efforts to further improve the in vitro potency of this series of compounds.

46
IC_{50} = 0.84 μM
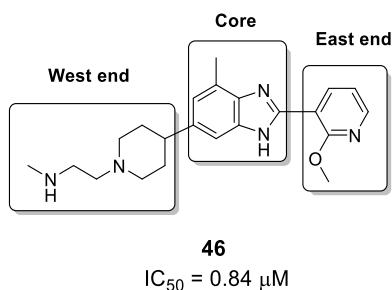
Figure 1.23. Benzo[d]imidazole Hit

SAR exploration of the west end of the hit (Figure 1.23) was performed. Very close modifications such as introduction of α-methyl group (as in **70** and **71**) as well as removal of internal nitrogen atom or abolishing basicity of either of nitrogen atoms (data not shown) led to a significant loss of potency (**Figure 1.24**). This observation suggests that the binding pocket near the West end is tight, and there is very little tolerance for any changes.
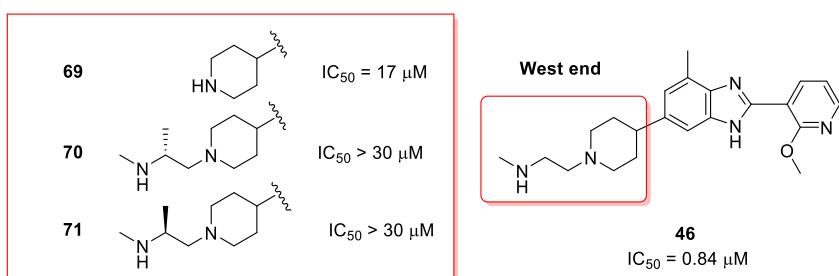
Figure 1.24. SAR exploration of west end of compound **46**

Then was explored the SAR of the East end of **46**. SAR from this study showed that almost all of the substituted phenyl and pyridyl groups were tolerated (only partial data displayed). Interestingly, an unsubstituted phenyl group was not tolerated (compound **72,** IC_{50} = 4.6 µM). Within the set of substituted phenyls, compounds with 2, 6-disubstituted phenyl group were preferred (as in **74,** IC_{50} = 0.07 µM and **75,** IC_{50} = 0.26 µM), indicating conformational preferences for the activity. Generally, phenyl ring bearing at least one hydrogen bond accepting ortho-substituents were favored.
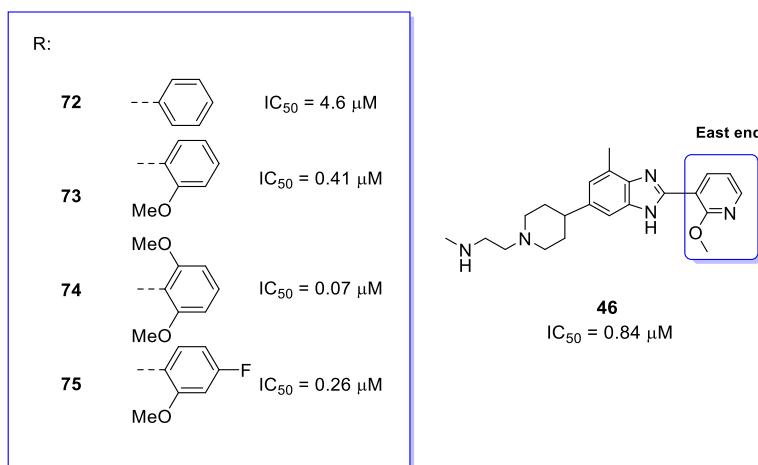
Figure 1.25. SAR exploration of east end of compound **46**

Furthermore, was investigated the core region of **46** (**Figure 1.26**). Towards this end, were synthesized analogs **76** and **77** to understand the effect of a methyl group at the 4- and 1-position of the benzo[d]imidazole. The 4-desmethyl analog (**77,** IC_{50} = 1.2 µM)) was slightly less potent while 1-methyl substitution at the nitrogen of the benzo[d]imidazole (**76,** IC_{50} = 2.9 µM) resulted in further loss of activity. Introduction of a nitrogen atom (**78,** IC_{50} = 0.7 µM) however was tolerated but did not offer any additional advantage.

48
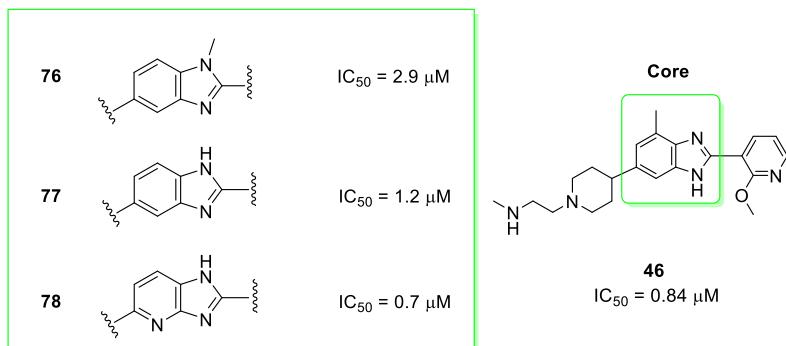
Figure 1.25. SAR exploration of core region of compound **46**

Further structural optimization of **46** led to the identification of indole derivative, the most active compounds of this series, with submicromolar activity in vitro methylation assay ($IC_{50} = 0.04$ μ) (**Figure 1.26**).[1] The benzo[d]imidazole/indole derivatives, like pyrazole inhibitors, lack of cellular activity, probably due to cytotoxic problems. Therefore further developments of these two classes of compounds were not pursued.
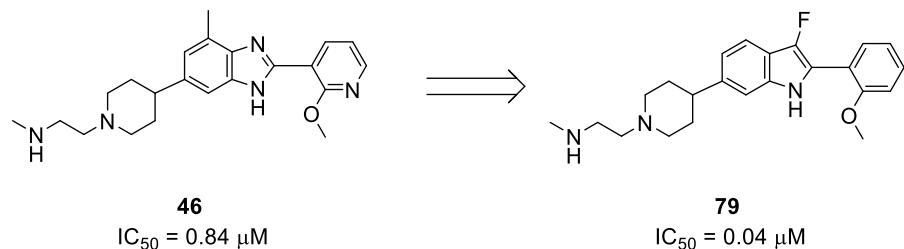
Figure 1.26. Hit to lead optimization

1.4.6.3. Structural basis for CARM1 inhibition by indole and pyrazole inhibitors

In 2012 Sack et al. reported crystal structures of the CARM1 catalytic domain in complex with cofactors [SAH (*S*-adenosyl- L-homocysteine) or SNF (sinefungin)] and indole (compound **79, Figure 1.26**) or pyazole (compound **61**, **Figure 1.20**) inhibitors, also known in literature as CMPD1 and CMPD2 respectively (**Figures 1.27** and **1.28**). Interestingly, both of these inhibitors bind to the substrate arginine-binding cavity of PRMT4 and require the presence of bound cofactor SAH or sinefungine. [152]

Structural studies of PRMT4 in complex with sinefungin and CMPD1 (**Figures 1.27**) revealed that the N-methylethanamine moiety of the inhibitor is directed toward the bottom of the arginine-binding cavity and directly interacts with the active site residue E258. The piperidine group is positioned at the entrance of the active site cavity and hydrogen bonds to H415 of the THW loop, whereas the indole moiety makes hydrophobic interactions with several aromatic side chains of PRMT4 and forms a water-mediated hydrogen bond to the main chain carbonyl of K471 and the side chain of N266.
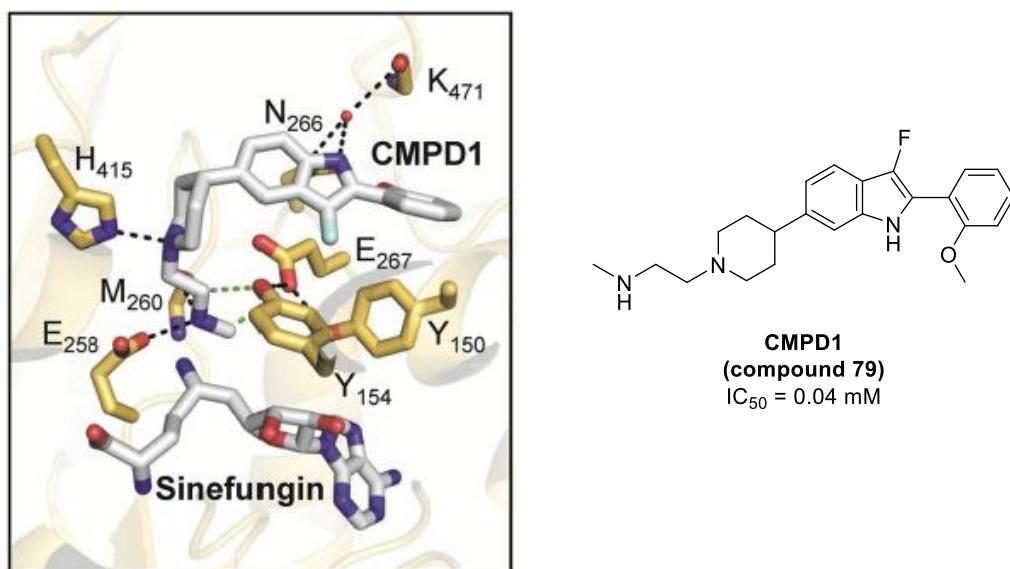
Figures 1.27. View of CMPD-bound in the active site of CARM1

The crystal structure of PRMT4 bound to CMPD2 (**Figures 1.28**) revealed that the terminal L-alaninamide moiety mimics the arginine guanidinium group and makes several polar interactions with PRMT4, including the carboxyl groups of the double E-loop residues E258 and E267. In addition, the alanylmethyl group of CMPD2 forms CH···O hydrogen bonds to the hydroxyl oxygen of Y154 and the backbone carbonyl oxygen of M260, while the carbonyl oxygen of the L-alaninamide hydrogen bonds with one of the imidazole nitrogens of H415. Notably, the (trifluoromethyl)- pyrazole, 1,3,4-oxadiazole, and indole scaffolds are thought to mainly interact with PRMT4 via shape complementarity rather than by polar interactions, except for two hydrogen bonds formed between the side chain of Y262 and the oxadiazole oxygen atom and between the hydroxyl of Y477 and a fluorine of the trifluoromethyl group.
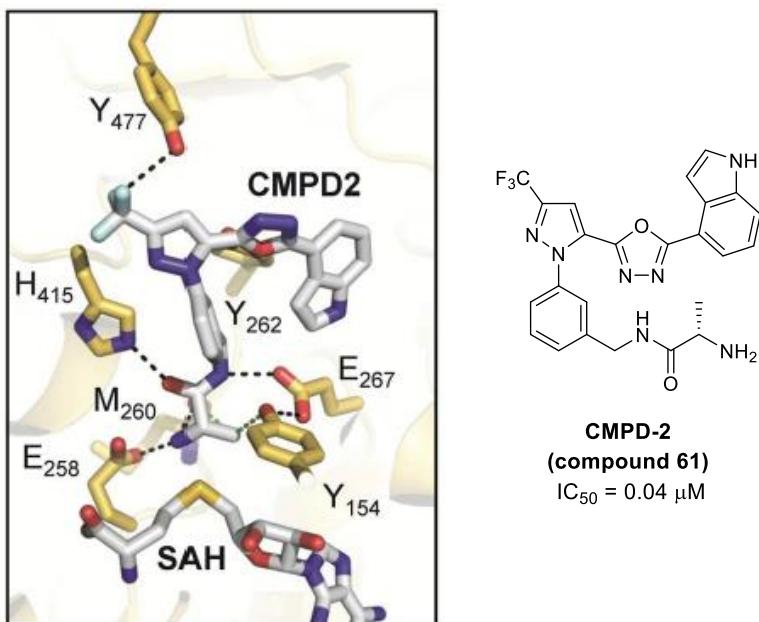
Figures 1.28. View of CMPD-2 bound in the active site of CARM1

CHAPTER 2

AIM OF THE WORK

2.1. Aim of the Work

The human Protein Methyltransferases (PMTs) constitute a large class of epigenetic enzymes, which include two different families: the Protein Lysine Methyltransferases (PKMTs) and the Protein Arginine Methyltransferases (PRMTs). They are involved in many cellular processes and it has been demonstrated that dysregulation of PMT activity contributes to different pathological states, including cancer.

Yet today there are clinical candidates [17] reported only for PKMTs and not for PRMTs, although the selective PRMT inhibitor discovery field is gaining momentum. A breakthrough in this area will be truly exciting and is keenly awaited. In this regard the present PhD project was focused on design, synthesis and biological evaluation of new small-molecule modulators of two PRMT isozymes: PRMT3 and CARM1 (Coactivator associated Arginine Methyltransferase 1), also known as PRMT4.

The research activities of my PhD project could be divided in three main topics:

1) Design, synthesis and biological evaluation of pyrrole inhibitors of CARM1.
2) Design, synthesis and biological evaluation of bisubstrate inhibitors of CARM1.
3) Design, synthesis and biological evaluation of indole inhibitors of PRMT3.

2.2. Design of CARM1 pyrrole inhibitors

For the design of CARM1 pyrrole-based inhibitors we have taken into account the information obtained from the structure-activity relationships (SAR) of indole and pyrazole compounds (see Chapter 1, paragraphs 1.4.6, pag. 41), the most potent and selective enzymatic inhibitors of PRMT4, which, however, lack activity in cell-based assays, due to a low cytopermeability and probably for cytotoxicity problems. Therefore, in an attempt to developed a new class of CARM1 inhibitors characterized by inhibition activity both in vitro and in cellular setting, we designed and synthesized a small set of pyrrole derivatives structural related to the pyrazole inhibitors of CARM1. With the aim to obtain a versatile and synthetically handy scaffold, we identified the pyrrole ring as the central core of our derivatives. This "privileged" structural motif, commonly found in pharmaceutical drugs and natural products, could be considered as a result of isosteric substitution of pyrazole together with opening and simplification of the indole ring, which represent the two central scaffolds of the aforementioned classes of CARM1 inhibitors (**Figure 2.1**).
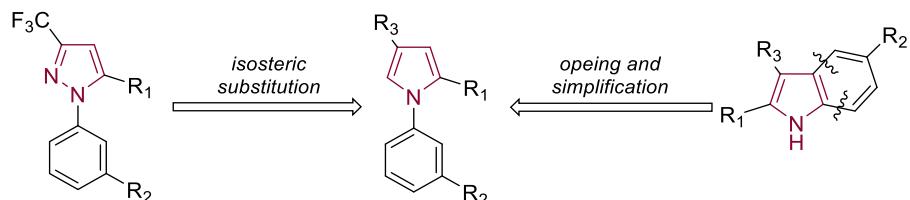
Figure 2.1. Selection of pyrrole scaffold

Moreover, considering the low cytopermeability of some pyrazole inhibitors (*i. e.* compounds **54** and **55,** Chapter 1, paragraph 1.4.6.1, pag.42), we hypothesized that the replacement of the nitrogen at position 2' of the pyrazole ring (which by crystallographic studies seems to be not involved into interactions with the enzyme) with a carbon should increase the lipophilicity and therefore the cytopermeability of the molecules. In addition, within the extensive SAR studies of pyrazole compounds (reported in Chapter 1,

paragraph 1.4.6.1) the effect of the substitution of the pyrazole core with a pyrrole ring was not investigated, probably due to a not immediate synthetic accessibility.

Once identified the central scaffold we started the process of scaffold decoration.

2.2.1. Design of pyrrole analogue of pyrazole lead I

The first pyrrole compound (**analogue of lead compound I**) was obtained by the replacement of the pyrazole scaffold with a pyrrole, retaining the same substitution patterns of pyrazole derivative **54**, identified as our lead compound (**Figure 2.2**). This modification was performed in an attempt to verify whether the isosteric substitution of the two heterocycles was also bioisosteric.

Concerning the substituent at the south end of **analogue of lead compound I** (**Figure 2.2**) we decided to retain the (S)-alanine moiety of **54,** because, as it is demonstrated by SAR studies of pyrazole compounds (Chapter 1, paragraph 1.4.6.1, pag 43), this group is essential for a potent inhibition activity toward CARM1. This warhead exploits the arginine binding cavity of CARM1 and has been verified that its modifications are not tolerated. Also the replacement with its enantiomer is detrimental for the activity; hence, in the new scaffold contemplated, we introduced the (S)-Alanine at the South end of the derivative.

Furthermore, at position 2 of the pyrrole ring (East end), the benzyl-amide moiety was selected as substituent (**Figure 2.2**). It has been demonstrated by SAR studies of pyrazole compounds (widely described in Chapter 1, paragraph 1.4.6.1, pag. 45) that this moiety allowed to obtain a high potency, preserving an ease of synthesis.
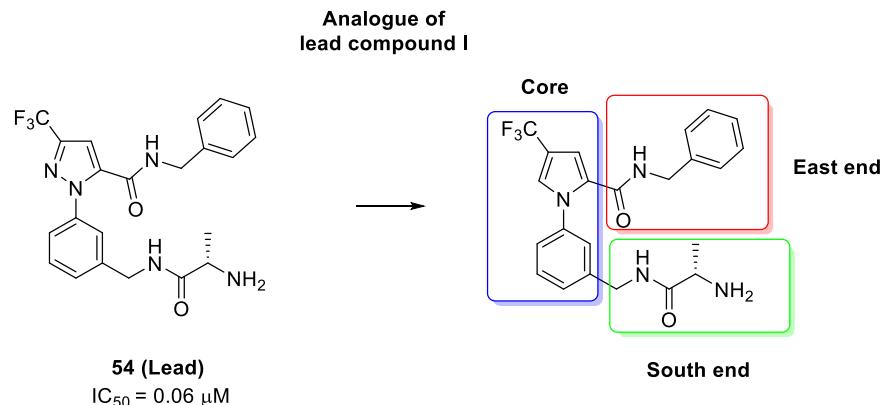
Figure 2.2. The replacement of pyrazole ring with pyrrole one led to the design of the pyrrole analogue of lead compound I

2.2.2. Design of Series II derivatives

Next, maintaining the East and South ends of the **analogue of lead compound I**, we explored the position 4 of the pyrrole core. In order to investigate the nature of the interactions toward CARM1, the trifluoromethyl moiety was replaced with COOEt, COOH or H groups, thus obtaining a series of N-Phenyl pyrrole derivatives (**Series II, Figure 2.3**). In the framework of this series, unlike the **analogue of lead compound I**, we considered both the enantiomeric forms of the alanine moiety, in order to assess if the stereospecific activity found for pyrazole compounds (Chapter 1, **paragraph 1.4.6.1**) also applies to pyrrole derivatives.

In addition, compounds of Series II were divided in two subsets, **IIa** and **IIb**, according to the presence or absence of a methylene group (highlighted in green in **Figure 2.3**) between the phenyl and aminoacidic moieties at the south end of the molecules. As it is depicted in **Figure 2.3**, the elimination of this short spacer provided the compounds of **Subseries IIb**. We assumed that the elimination of this methylene unit could positively enhance the positioning of the (S)-alanine group within the enzyme pocket. Indeed, biological evaluation of these compounds (Chapter 4) confirmed our hypothesis.

59
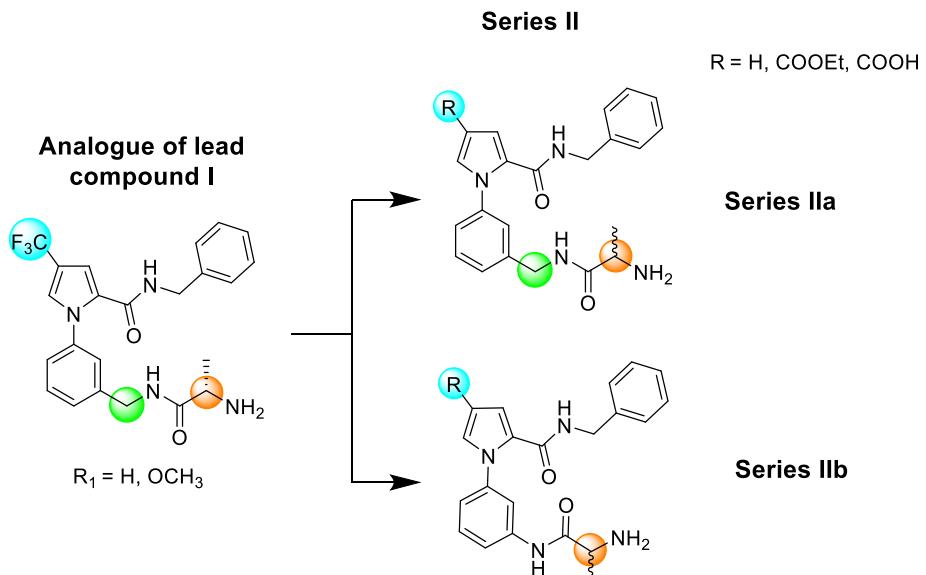
Figure 2.3. Design of N-Phenyl pyrrole derivatives (Series II)

2.2.3. Design of Series III derivatives

Furthermore, in order to evaluate the effect on the activity resulting from an increase in conformational flexibility, another set of molecules was designed by replacing the N-Phenyl pyrrole scaffold of Series II derivatives with N-Benzyl pyrrole one (**Figure 2.4**). The aim of this approach was to evaluate if a higher degree of flexibility could lead the side chains of the molecules to match the binding sites of the protein in a most optimal way. As it is possible to see in **Figure 2.4**, also this series of compounds was divided in two subseries (**Series IIa** and **IIb**), according to the presence or absence of the methylene unit between the phenyl ring and the basic moiety at the south end.
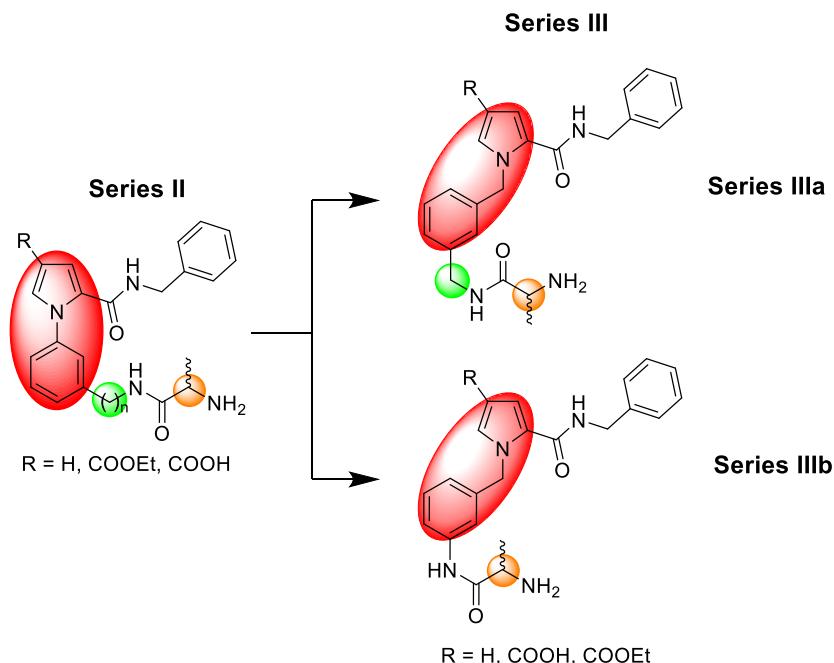
Figure 2.4. Design of N-benzyl pyrrole derivatives

2.2.4. Design of new pyrrole derivatives

Once that the biological investigations of pyrrole compounds were completed (Chapters 4 and 5), we realized that the replacement of the pyrazole ring with the pyrrole one allowed to conserve a good activity; indeed a potent inhibition was observed when testing pyrrole derivatives against CARM1 (i.e. **EML 438**, IC_{50} = 2.42 μM). However, although our compounds were more lipophilic than pyrazole lead, they didn't prove a significant cellular activity, due to their low transcellular permeability. Therefore, in order to improve the lipophilicity of compounds here described, we employed structural optimization efforts, according to the computational studies performed by the research group of Professor Giuseppe Bifulco from the University of Salerno. This approach should also ensure an improvement of the enzyme binding properties, since binding pose investigations of pyrrolic compounds demonstrated that they were able to occupy only a portion of the binding site of the enzyme. To this

end, starting from the most active pyrrole-based derivatives (**EML 438** and **EML 446**), we introduced an alkyl moiety at the East end of the molecules (**Series A**) or a second heterocycle at right or left sides of the pyrrolic central core (**Series B** or **C**, **Figure 2.5**).
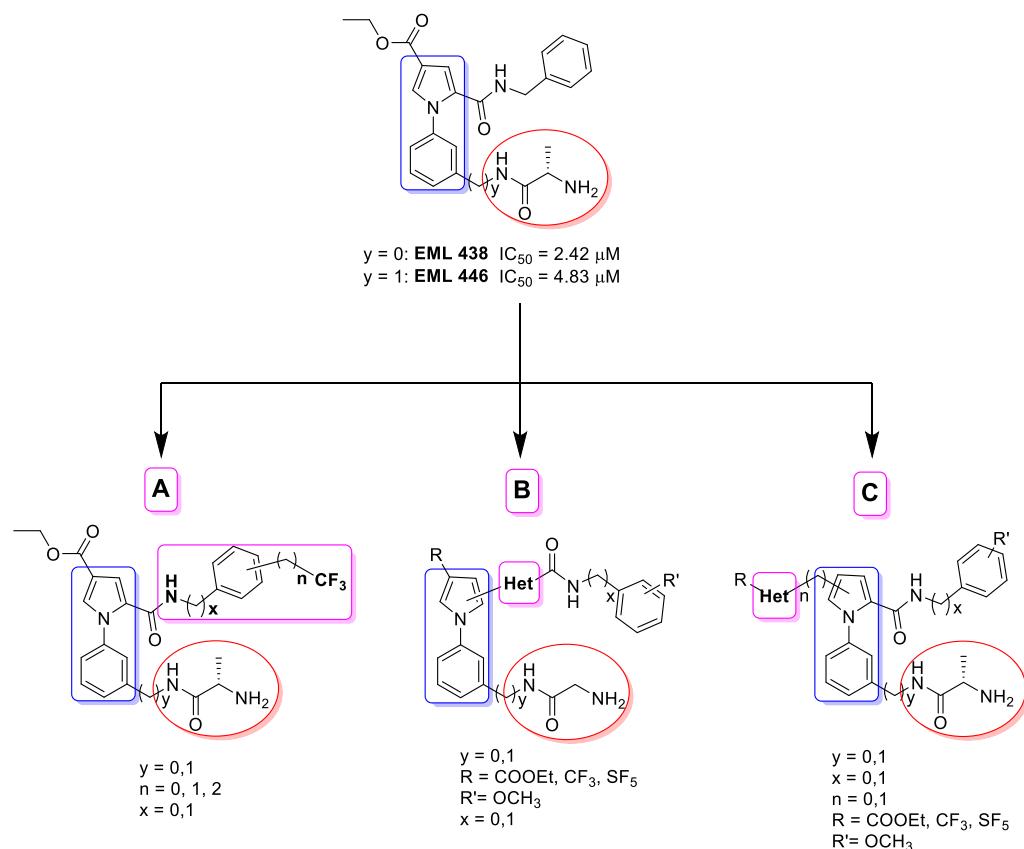
Figure 2.5. Development of new pyrrole derivatives

2.3. Design of bisubstrate inhibitors of CARM1

Pursuing our efforts toward the development of selective inhibitors of CARM1, in collaboration with the group of Professor Nathaniel Martin from the University of Utrecht, where I performed a short research period as visiting PhD student, we designed and synthesized a set of novel PRMT4 bisubstrate inhibitors. It is conceivable that bisubstrate analogs have the potential to afford increased potency and selectivity by combining the free binding energies of the cofactor and substrate interactions without suffering entropic reaction (**Figure 2.6**). Ideal bisubstrate inhibitors should have an effective linker between the two fragments that mimic the transition state of the methyl transfer reaction, with one fragment of the inhibitor targeting the SAM-binding site and the other targeting the substrate-binding site.
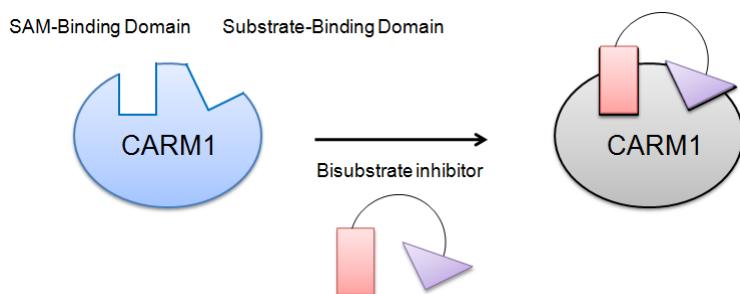
Figure 2.6. Schematic illustration of CARM1 bisubstrate inhibitor concept

As illustrated in **Figure 2.7**, CARM1, like other PRMTs, operates via a bisubstrate mechanism wherein the guanidine group of an arginine-containing peptide substrate is precisely orientated (via a hydrogen-bond network with Glu257 and Glu266) in proximity to the electrophilic methylsulfonium group of AdoMet leading to an "SN2-like" substitution reaction. Key structural features such as the active site double E-loop, which is critical for guanidinium binding, as well as the SAM-binding pocket residues, are conserved among all PRMTs. Within the active site there are a number of conserved residues that are important for SAM binding, catalysis, and maintaining the overall

architecture of the CARM1 active site (**Figure 2.7**). These residues include Glu244 and Val243, which directly interact with the SAM adenine ring and Glu160 which forms a bidentate hydrogen bond to the ribose moiety. Regarding the guanidine group of the arginine-containing peptide substrate (**Figure 2.7**), it makes a bidentate interaction with Glu266 and single hydrogen bond with Glu257. Both of these residues are thought to recognize and align the arginine guanidinium substrate for proper catalysis to occur.
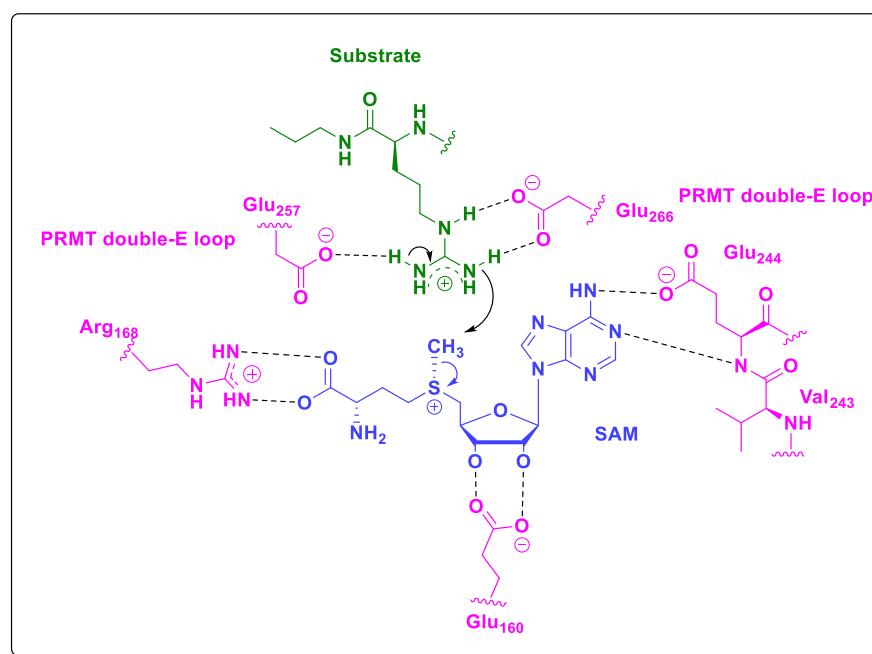
Figure 2.7. Hypothesized PRMT bisubstrate transition state.

The remaining interactions between CARM1 and its peptidic substrate are distinct from others PRMTs and are responsible for the substrate recognition. CARM1 is able to mono- and di- methylate multiple histone (H3R17 and H3R26[153]) and non histone substrates (SRC-3,[91] NCOA2,[99] PABP1[154] and SmB.[84]) with diverse sequences, exhibiting a preference for distinct substrate motifs compared to those preferred by other type I PRMTs. Understanding how CARM1 achieves this catalytic promiscuity and what the common recognition determinants are for its various substrates is a critical unanswered

question. However, have been reported by Copeland at al. [155] three crystal structures of human CARM1 with the SAM mimic sinefungin and different peptide sequences from histone H3 and PABP1, in which are elucidated the main interactions between the enzyme and the peptide sequences. In particular it has been shown that CARM1 makes several hydrogen bonds with the backbone residues of the peptide substrates. These interactions seem to be largely substrate sequence independent and involve Glu266, Asn265, Tyr476, Asn161 and Tyr416 residues (**Figure 2.8**).
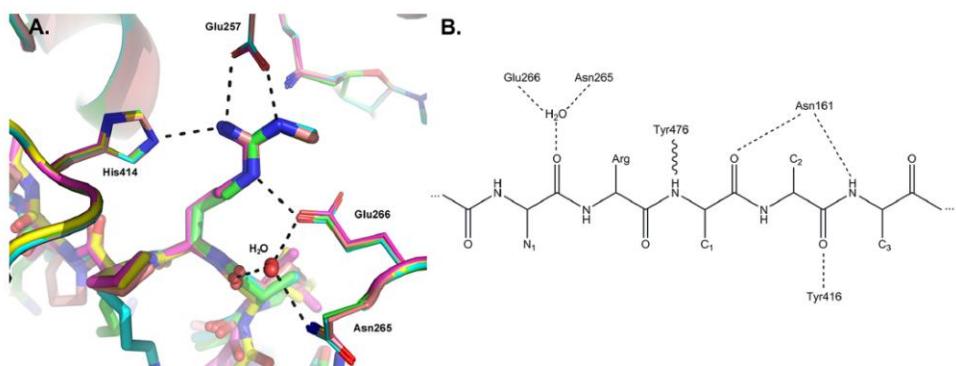
Figure 2.8. Schematic illustration of CARM1-peptide interactions.

Considering these structural features of CARM1 binding sites, we designed a series of novel bisubstrate inhibitors of CARM1. These compounds are composed by an Arginine containing peptide fragments covalently linked, through a variable guanidine linkers, to adenosine ring (**Compounds 80 – 85**, **Figure 2.9**).
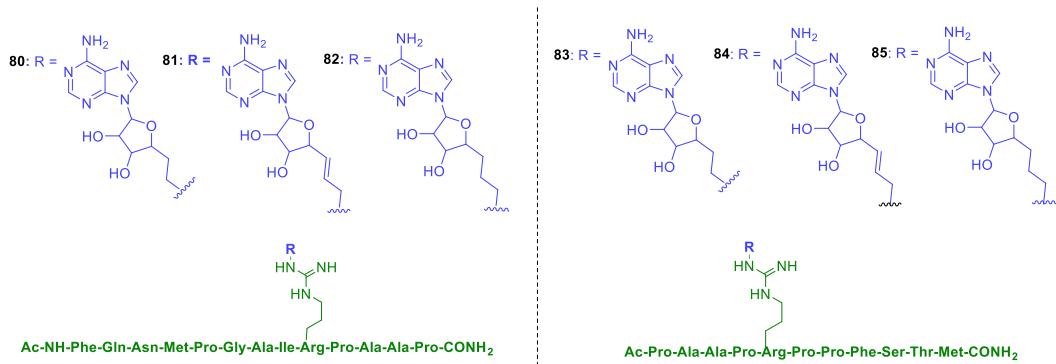
Figure 2.9. Illustration of CARM1 bisubstrate inhibitors.

As peptide fragments we used two short peptide sequences (**P1**: FQNMPGAIRPAAP and **P2**: PAAPRPPFSTM) taken from PABP1 (Poly (A)-binding protein 1), a non histonic substrate of PRMT4. It has been demonstrated by Bedford M. T. and collaborators[154] that those sequences are recognized and methylated efficiently by the enzyme.

In order to engage also the cofactor binding site, the two peptide sequences were coupled with a set of compounds containing the adenosine moiety, previously developed as partial bisubstrate inhibitors of PRMT4 by Martin N. I. and collaborators (**Figure 2.10**).[142] These compounds (**17a – c**), characterized by an adenosine ring connected through an aliphatic linkers with a guanidine group, showed a preferential activity toward CARM1. They are considered partial bisubstrate inhibitors because are able to exploit the SAM binding site of CARM1 and only a portion of the enzymatic pocket involved in binding interactions with the protein substrate .

Figure 2.10. Partial bisubtrate inhibitors identified by Martin et al.

2.3.1. Asymmetrically dimethylated peptides

It has been demonstrated by Bedford and coll.[154] that the two peptide sequences (FQNMPGAIRPAAP and PAAPRPPFSTM) are recognized as substrates by PRMT4. Therefore, taking into account the effect of product inhibition, we decided to synthesized the same sequences with the arginine residue asymmetrically dimethylated (**Figure 2.11**).
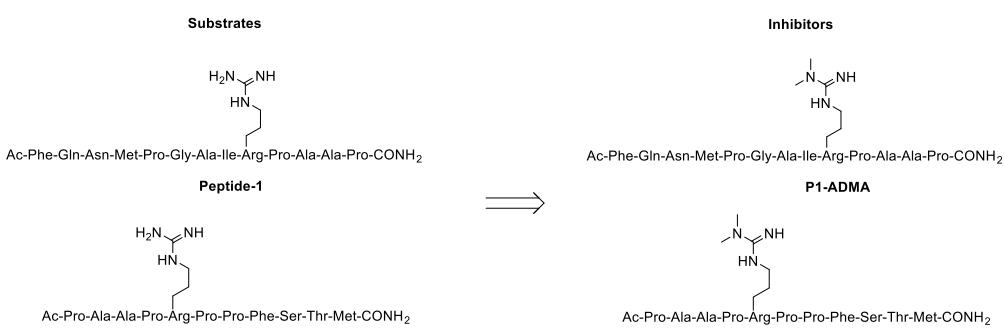
Figure 2.11. Conversion of peptide substrates into inhibitors

2.4. Design of PRMT3 indole inhibitors of PRMT3

In 2012, during a project with the aim to identify selective inhibitors of PRMTs, we were pleased to discover a Series of indole-based compounds referred to as "Uracandolates" (aryl ureido acetamido indole carboxylates), which selectively and potently enhanced the methylation activity of CARM1 both in in vitro and cellular settings (**Figure 2.12**). Moreover, as it is possible to see in **Figure 2.12**, these compounds showed a defined, albeit lower, inhibition activity against PRMT3. Therefore, taking into account this information, we decided to prepare novel indole derivatives, structural related to Uracandolate compounds, in order to develop novel PRMT3 inhibitors and simultaneously extent the structure-activity relationships of the CARM1activators.
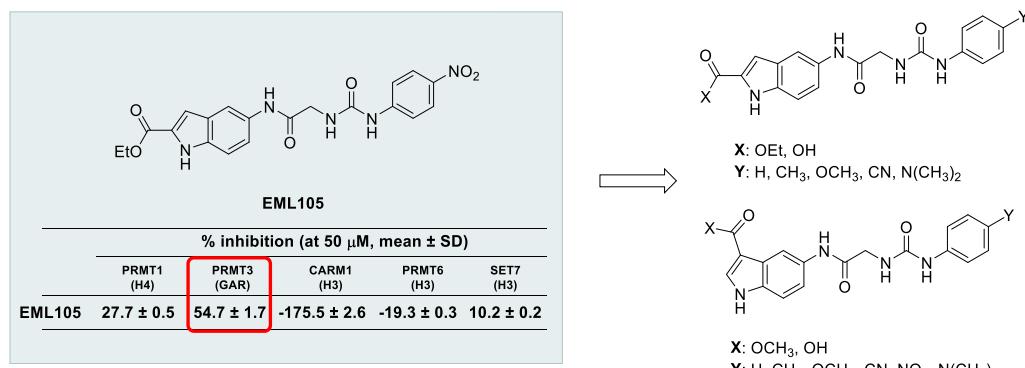
Figure 2.12. Development of Uracandolates

This project was then supported by the subsequent discovery of benzathiadiazole/isoquinoline derivatives, designed by Jin and coworkers as potent and selective allosteric inhibitors of PRMT3 (Chapter 1, see paragraph 1.4.5). Interestingly, both the classes of PRMTs modulators, Uracandolates and benzathiadiazole/isoquinoline compounds, display some common structural features; therefore we evaluated the possibility of integrating their structural information for the development of new modulators of PRMTs. In particular both the series of compounds are characterized by:

- ➢ A byciclic aromatic region (highlighted in red in **Figure 2.13**)

- ➢ A central linker of 5 or 6 elements, bearing a carbonyl moiety and an
  ureidic function (highlighted in blue)

- ➢ A right end of the molecule, bearing different substituents (showed in
  green in **Figure 2.13**).
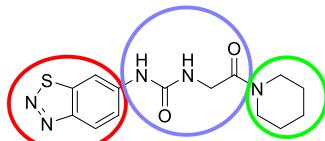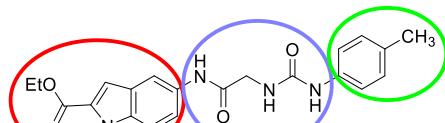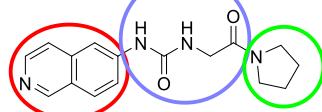
Figura 2.13. Structural correlation between Uracandolates and Benzothiadiazole/isoquinoline inhibitors. In red is highlighted the byciclic aromatic region, in blue the ureidic linker and in green the right end of the molecule

Hence, inspired by the SAR information of benzathiadiazole/isoquinoline compounds and considering the experimental data in our hands about the PRMT3 inhibition by Uracandolates, we decided, starting from the backbone of Uracandolate compounds, to integrate the structural information of both the

classes of PRMTs modulators in order to develop new indole-based compounds.

Three Series of derivatives were designed (**Figure 2.14**), all of them share the indolic scaffold and the same substituents at the right end of the molecule, instead, differ for the central ureidic linkers.

Figure 2.14. Design of new indolic derivatives
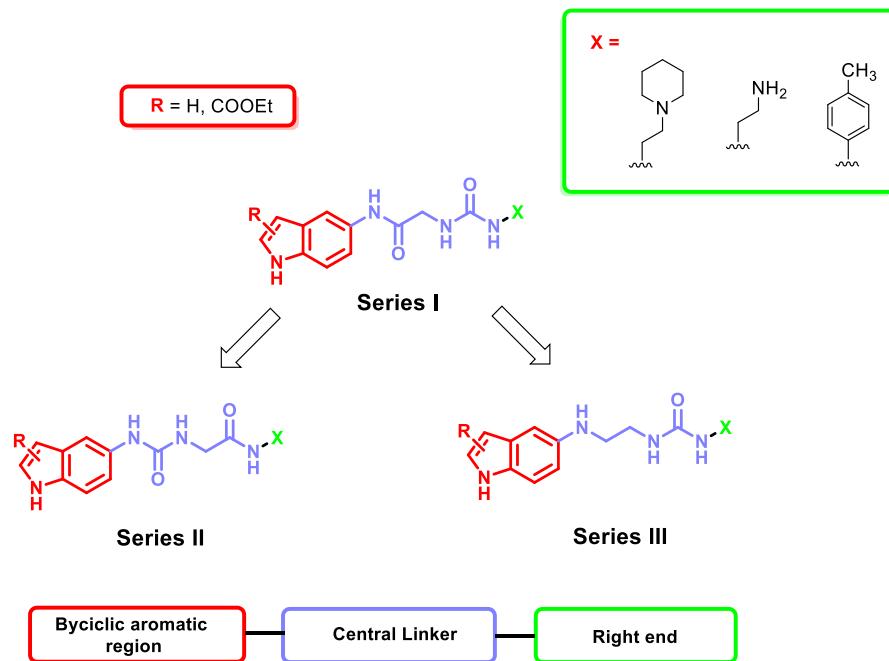
2.4.1. Indolic scaffold

The indole ring, substituted at positions 2 or 3 with COOEt or H, was selected as central scaffold of the new derivatives. This core, considered as privileged scaffold for drug discovery, is the central core of Uracandolate compounds; therefore we decided to use the same heterocycle for the new derivatives. Moreover, within the SAR studies of the benzathiadiazole/isoquinoline derivatives the effect of the replacemt of the central scaffold with an indolic ring was not investigated, despite SAR efforts focused on the substitution of the benzathadiazole and isoquinoline rings were performed.

2.4.2. Right end of the molecules

On the other hand, regarding the right end of the molecules (highlighted in greed in Figure **2.12**), inspired by the SAR information of benzathiadiazole/isoquinoline derivatives, we selected three different groups: a *para* substituted aromatic ring, a cyclic aliphatic amine (piperidine) and a primary amine. These groups were introduced in an attempt to investigate the nature of the interaction exploited by the right end of the indole-based derivatives with the enzyme (highlighted in greed in Figure **2.12**).

As it is reported in SAR studies of benzathiadiazole/isoquinoline derivatives (*paragraph 1.4.5.1,* pag. 40) this portion of the molecule has proven to be essential for PRMT3 inhibition activity, since prevents the formation of a catalytically competent state of PRMT3 through hydrophobic interactions with the dimerization arm of PRMT3.

2.4.3. Central linkers

The three series of compounds differ each others for the nature of the central linker between the indolic scaffold and the substituents at the right end. An aliphatic chain of 6 units was selected for all the derivatives, however the central spacer differs among the three series for the position of the ureidic moiety and for the presence of an amide function.

Compounds of **Series I** show the same aliphatic linker of Uracandolates, in particular they are characterized by a methylene unit between an amide moiety, directly bound to the byciclic aromatic region, and the ureidic function that is asymmetrically substituted (**Figure 2.15**).

71
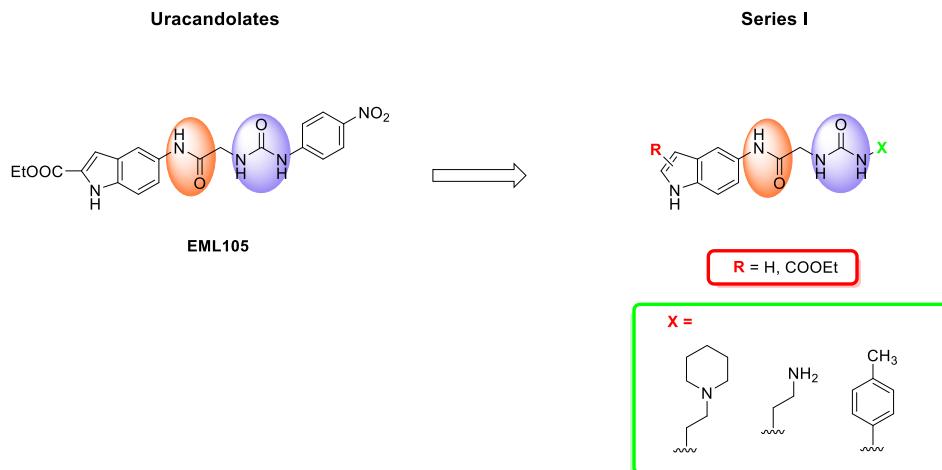
Figure 2.15. Compounds of Series I

Starting from compounds of **Series I**, the formal shift of the ureidic moiety to the indolic ring provided compounds of **Series II** (**Figure 2.16**). The aliphatic spacer of this set of compounds displays an arrangement similar to benzathiadiazolic derivatives.
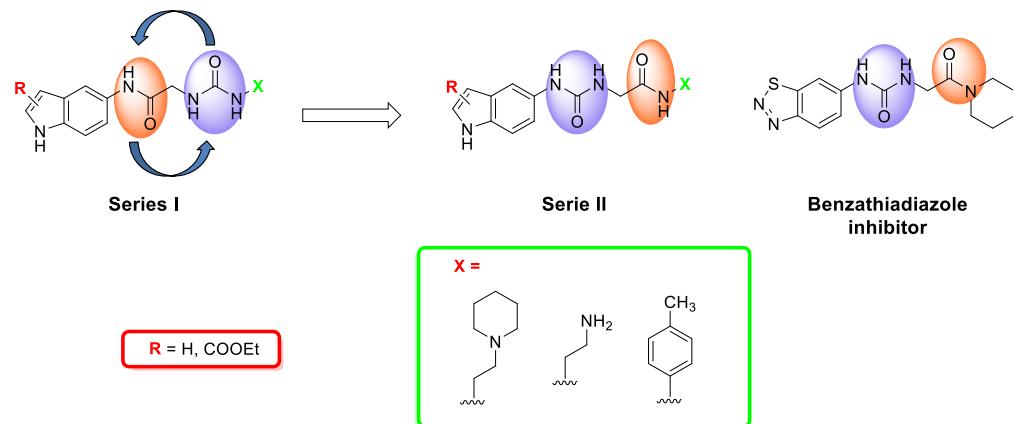

Figura 2.16 Design of the Series II

Finally the replacement of the amide group of **Series I** derivatives with an amino group led to the derivatives of **Series III** (**Figure 2.17**). This modification was performed in order to assess the role of the carbonyl group of the amide function within the aliphatic spacer. The role of this function was not well elucidated by the SAR studies of bezathaidiazole compounds, indeed

it has been reported that derivatives bearing only the ureidic function without the amide group still showed a good inhibition activity against PRMT3 (chapter 1, **Figure 1.1.6**).
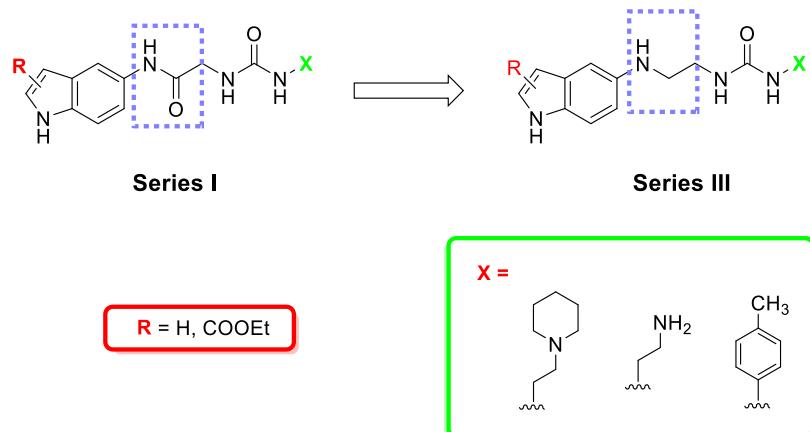
Figura 2.17. Design of Series III

CHAPTER 3

CHEMISTRY

3.1. Synthesis of CARM1 pyrrole inhibitors

From a synthetic point of view, the pyrrolic compounds were divided into 3 groups:

1) Pyrrole analogue of lead compound I.

2) N-phenyl pyrrole derivatives (**Series II**).

3) N-benzyl pyrrole derivatives (**Series III**).

The scaffold construction was the pivotal step for the preparation of the designed derivatives, indeed the different substitutions of the heterocycle required distinct methods for its construction. In **Figure 3.1** are schematically reported the procedures used for the scaffolds synthesis, which differ between each other for the retrosynthetic dissections
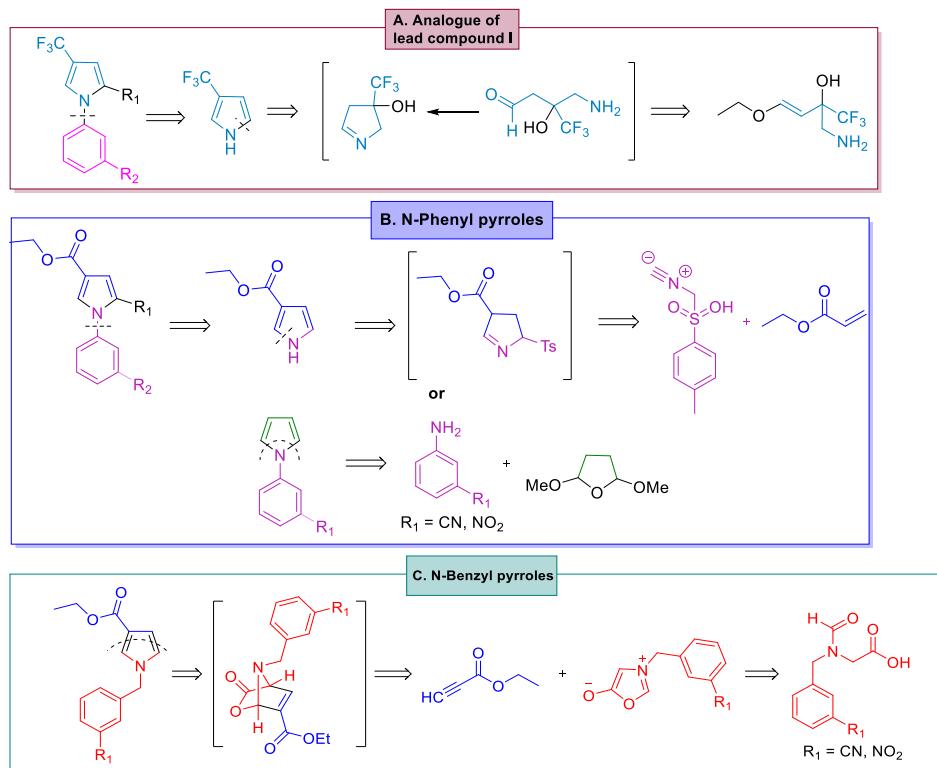
Figure 3.1. Three different synthetic approaches for the synthesis of the scaffolds: A**)** Intramolecular cyclization. B) van Leusen or Clauson-Kaas reaction. C**)** 1,3-dipolar cycloaddition reaction

The 3-trifluoromethyl-1H-pyrrole, precursor of the pyrrolic analogue of lead compound I, was obtained exploiting an intramolecular cyclization of the aminocarbonyl intermediate (**Figure 3.1. A**).[156] Whereas the ethyl 1H-pyrrole-3-carboxylate was obtained *via* van Leusen reaction (**Figure 3.1. B**),[157] in which tosylmethyl-isocyanide (TOSMIC) is cyclocondensed with the α,β-unsaturated ester (ethyl acrylate) in presence of base.

These procedures, unlike the others that are discussed below, did not afford to N-substituted pyrrole scaffolds, therefore a cross coupling reaction between the heterocycles and a proper aryl bromide was performed in order to obtain the derivatives N-Phenyl substituted.

Conversely the Clauson-Kaas reaction as well as the 1,3-dipolar cycloaddition reaction,[158] respectively used for the synthesis of the **Series II** and **III** compounds **(Figure 3.1 B** and **C)**, provided the pyrrolic scaffold N-Phenyl or N-Benzyl substituted.

Starting from proper aniline and 2, 5 dimethoxytetrahydrofuran the N-Phenyl pyrrole derivatives of the **Series II** were prepared by Clauson-Kaas reaction (**Figure 3.1. B).** While, the 1, 3 dipolar cycloaddition reaction between ethylpropiolate and 1, 3 diacylic intermediate led to the synthesis of N-benzyl derivatives (**Figure 3.1. C**).

Once obtained the pyrrolic central cores, the synthetic procedures used for the scaffolds decoration were common for all the designed derivatives and simple reactions were employed (see next paragraphs).

3.1.1. Synthesis of pyrrole analogue of lead compound I

3.1.1.1. Synthesis of 3-trifluoromethyl-1H-pyrrole.

The synthesis of the 3-trifluoromethyl-1H-pyrrole was accomplished following the synthetic procedure reported by Shaitanova et al. in 2008.[156] The reaction between ethyl vinyl ether (**86**) and trifluoroacetic anhydride in the presence of catalytic amount of DMAP yielded α, β unsaturated carbonylic compound (**87**), which after the addition of trimethylsilylcyanide (TMSCN) in presence of a catalytic amount of base, was converted to a silylated cyanohydrins compound (**88**). Finally, reduction of **88** with LiAlH$_4$ led to the synthesis of the amino alcohol intermediate (**89**) in high yield, **Scheme 3.1.**
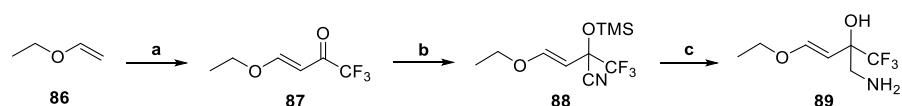
Scheme 3.1. Reagent and condition: (**a**) DMAP, trifluoroacetic anhydride, Ethyl vinyl ether, -10 °C, 30 min., then 0 °C, 10 h; 99%; (**b**) TMSCN, Et$_3$N, 0 – 10 °C, 95%; (**c**) LiAlH$_4$, diethyl ether, 0 °C, 88%.

At this point, according to the reported procedure, the last step of the 3-trifluoromethyl-1H-pyrrole synthesis should be the hydrolysis of the alkoxyvinyl group of **89** to aminocarbonyl compounds (**90**) using catalytic or stoichiometric amount of 5% aqueous hydrochloric acid. Intermediate **89** in turn, should cyclize to the pyrrole (**92**) via intramolecular Schiff base formation (**91**) with subsequent dehydration and proton migration (**Scheme 3.2**). Following this procedure, any attempts to obtain the aforementioned scaffold (**92**) failed. In particular, we found that the intermolecular polymerizations of **89,** after the hydrolysis of the alkoxyvinyl group, were favourited compared to the desired intramolecular cyclization. In addition, the 3-trifluoromethyl-1H-pyrrole has a very low boiling point (< 35 °C), therefore its failure recovery, should be also due to its loss during the time of reaction (7

days). Our hypotheses were then confirmed by the authors of this procedure, which revealed that the reaction suffers of scale-up problems too.

The selective protection of the amino group of the intermediate **89 (Scheme 3.2**) (compound **93**) with benzoyl chloride**,** followed by an intramolecular cyclization under acidic condition (mixture of DCM/TFA 7: 3), led to the pyrrole ring formation N-benzoyl protected (compound **94, Scheme 3.2**), overcoming the intermolecular polymerization and the problem of product volatility.

Final deprotection with LiOH in a mixture of dioxane/water allowed us to obtain the desidered pyrrolic scaffold (**92**), which, considering the low boiling point, was directly used in the next step without being isolated (Scheme **3.2**).
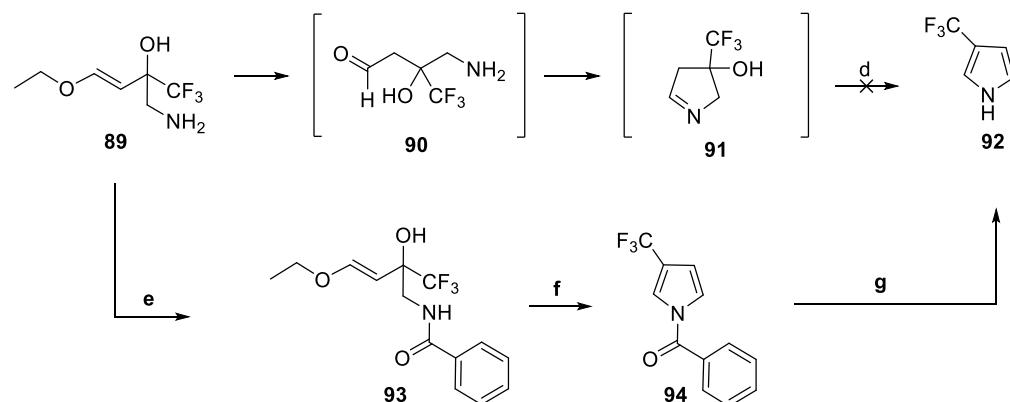
Scheme 3.2. Reagent and condition: (**d**) HCl (0.1 eq. 5% aq.), CH$_3$CN/H$_2$O 2:1, t. a., 1 week or HCl (1 eq. 5% aq.), CH$_3$CN/H$_2$O 2:1, t. a., 1 week; (**e**) Benzoyl chloride, TEA, DCM, 0°C, 5 min, r. t., overnight., 65%; (**f**) DCM/TFA 7:3, r. t., 24 h, 99%; (**g**) LiOH, dioxane/H$_2$O 1: 1, 4 h, r. t
.

3.1.1.2. Functionalization of the 3-trifluoromethyl-1H-pyrrole

As it is previously reported the key step for the synthesis of the designed pyrrole derivatives was the central core construction. The synthetic procedures for the scaffolds decoration were common for all the pyrrolic derivatives and consisted of simple reactions.

The C-2 functionalization of the pyrrole intermediate **92** with the benzylamide moiety was accomplished by Friedel-Crafts acylation with trichloroacetyl chloride (**95**), followed by nuclephilic substitution, employing benzylamine as reagent (derivative **96**, Scheme **3.3**). Notably, negative results were obtained in our attempt to acylate, under Friedel-Crafts conditions, the pyrrolic α-position of **94** (**Scheme 3.2**), since the presence of the benzoyl protecting group strongly deactivated the aforementioned position.
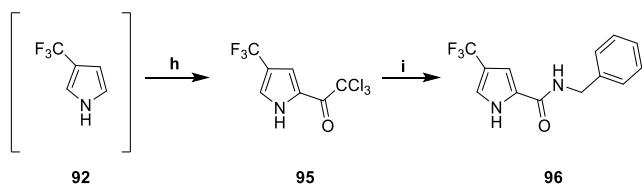
Scheme 3.3. Reagent and condition: (**h**) trichloroacetyl chloride, DCM, 70 °C, 36 h, 56 % over steps **f** and **g**; (**i**) benzylamine, K$_2$CO$_3$, DMF, MW, 80 °C, 30 min, 70%

Afterwards, the N-arylation of **96** was accomplished exploiting an Ullmann-type reaction (**Scheme 3.4**). [159] The microwave-assisted reaction between 3-Bromobenzonitrile and **96**, using CuI, N, N'-dimethylethylendiamine and K$_3$PO$_4$ at 130 °C in dioxane yielded the N-Phenyl substituted intermediate **97**. The subsequently reduction of nitrile group with NaBH$_4$ in the presence of catalytic amount of NiSO$_4$•6 H$_2$O (compound **98**), followed by coupling reaction with L-Boc-Ala-OH led to the introduction of the aminoacidic portion on the phenyl ring (intermediate **99**). Final removal of Boc protecting group provided the **pyrrole analogue of the lead compound I** (compound **100, Scheme 3.5**).
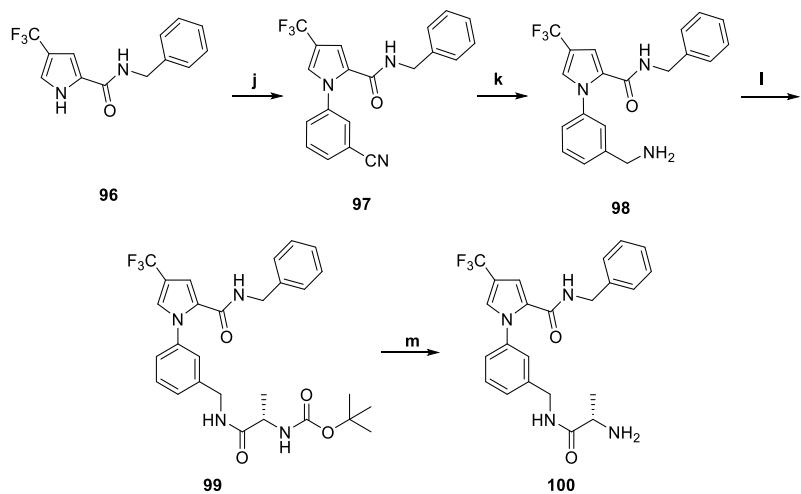
Scheme 3.4. Reagent and condition: (**j**) 3-Bromobenzonitrile, CuI, *N,N'*-Dimethylethylenediamine, K_3PO_4, dioxane dry, MW, 130 °C, 30 min, 80% ; (**k**) $NaBH_4$, $NiSO_4 \cdot 6 \ H_2O$, MeOH, 30 min., r. t.; (**l**) (L) - Boc - Ala - OH, DMAP, DCC, DCM, overnight, r.t., 75 % over steps **j** and **k**; (**m**) DCM/TFA 9: 1; r. t., 30 min, 99%.

3.1.2. Synthesis of N-Phenyl pyrrole derivatives (Series II)

Within this series of derivatives, two synthetic approaches were used for the preparation of the central core, since the different substitution of the pyrrole ring influenced its reactivity.

3.1.2.1. Pyrrole scaffold synthesis

Compounds **103a** and **103b** were obtained by Clauson-Kaas reaction,[160] starting from a proper substituted aniline (**101a** or **101b**) and 2, 5 dimethoxy tetrahydrofuran (**102a**) heating at 80 °C in AcOH (**Scheme 3.5**).

101a: R = NO_2 **102a:** R' = H
101b: R = CN   **102b:** R' = COOEt

103a: R = NO_2; R' = H
103b: R = CN; R' = H
103c: R = NO_2; R' = COOEt
103d: R = CN; R' = COOEt
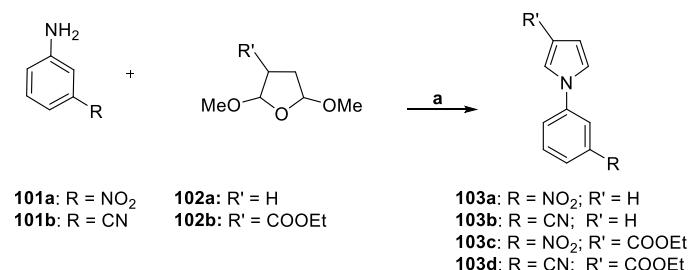
Scheme 3.5. Reagents and conditions: (**a**) 2,5-Dimethoxytetrahydrofuran, AcOH, 80 °C, 2 h, 70 – 84%.

The Clauson-Kaas reaction is a very advantageous procedure, because allow to obtain the N-Phenyl pyrrolic scaffold in one step with high yield. Therefore, we initially sought to use this approach for preparing all the derivatives of the **Series II**, including the 3-ethyl ester substituted N-Phenyl pyrroles (**Scheme 3.5**, compounds **103c, d**). Our initial attempts gave good results leading to the synthesis of **103c** and **103d** in good yield **(Scheme 3.5)** by reaction of **101a** and **101b** with ethyl 2,5-dimethoxytetrahydrofuran-3-carboxylate (**102b**), heating at 80°C in AcOH. Nevertheless, the following step of pyrrole electrophilic substitution of **103c, d** with trichloroacetyl chloride afforded several problems. For the acylation step of the ester substituted derivatives were required strong conditions, since the steric hindrance as well as the electronic deactivation of the C-2 pyrrolic position, made it less susceptible to the electrophilic substitution. In particular our attempts to carry out this reaction with trichoroacetyl chloride in presence of aluminum chloride, under Friedel-Crafts acylation conditions, resulted in a mixture of positional isomers obtained in low yields (< 20 %) difficult to separate (**Scheme 3.6**). Whereas the reaction by **103a, b** and trichloroacetyl chloride heating at 120 °C yielded the desired products in high yield (intermediates **104a – b, Scheme 3.6**).

103a: R = NO$_2$; R' = H
103b: R = CN; R' = H
103c: R = NO$_2$; R' = COOEt
103d: R = CN; R' = COOEt

104a: R = NO$_2$
104b: R = CN

104c: R = NO$_2$
104d: R = CN

104e: R = NO$_2$
104f: R = CN
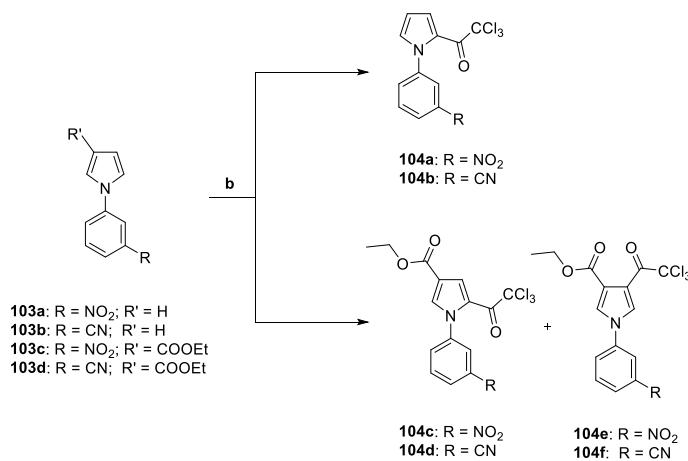
Scheme 3.6. Reagents and conditions: (**b**) ClCOCCl$_3$, neat, MW 120 °C, 1.5 h for **105b, c** or ClCOCl$_3$, AlCl$_3$, DCM, r. t., 12h, > 20%.

Therefore, taking into account these problems, we decided to employ the van Lausen reaction for obtaining the pyrrole scaffold bearing the ethyl ester moiety at the β position of the heterocycle. Ethyl 1H-pyrrole-3-carboxylate (**107**) was prepared starting from tosylmethyl-isocyanide (TOSMIC) which was cyclocondensed with ethyl acrylate in the presence of NaH in good yield (**Scheme 3.7**).

105   **106**   **107**
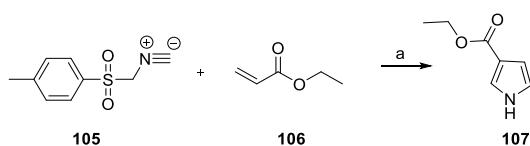
Scheme 3.7. Reagents and conditions: (**a'**) Tosmic, NaH, DMSO dry, Et$_2$O dry, 5 h, 60 %;

3.1.2.2. Scaffold decoration

Friedel-Crafts acylation with trichloroacety chloride, followed by nuclephilic substitution with benzylamine provided the introduction of the benzylamide moiety at pyrrole position C-2 of derivatives **108a, b** and **110** (**Schemes 3.8 and 3.9**). The N-Aryl derivatives **108c, d** were prepared by reaction of Ullman employing intermediate **110** (**Scheme 3.9**), a proper aryl bromine, K$_3$PO$_4$ as a base and CuI, N, N'- Dimethylethylenediamine in dioxane heating at 130 °C for 30 minutes (**Scheme 3.9**).

103a: R=NO$_2$ **104a**: R=NO$_2$ **108a**: R=NO$_2$
103b: R=CN   **104b**: R=CN   **108b**: R=CN
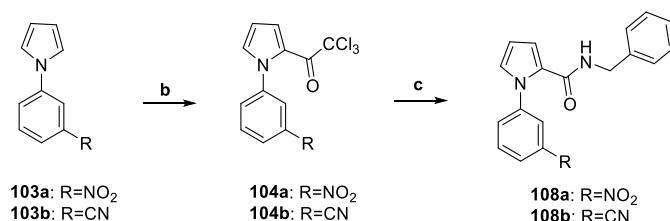
Scheme 3.8. Reagents and conditions: (**b**) ClCOCCl$_3$, neat, MW 120 °C, 1.5 h; (**c**) benzylamine, K$_2$CO$_3$, DMF, MW, 80 °C , 30 min, 65 – 70 %.

84
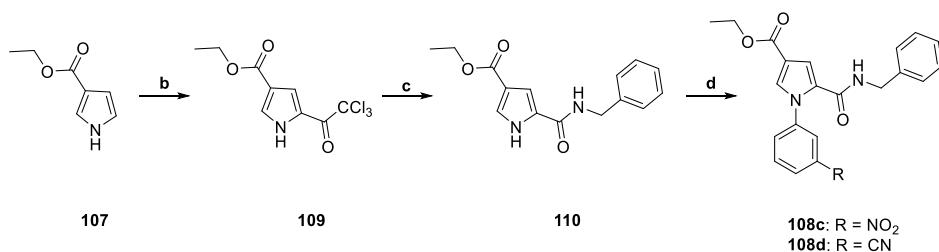
Scheme 3.9. Reagent and condition: (**b**) ClCOCCl$_3$, DCM dry, 80 °C, 5 h, 60 – 70 %; (**c**) benzylamine, K$_2$CO$_3$, DMF, MW, 80 °C , 30 min, 70 – 80 %; (**d**) 1-Bromo-3-nitrobenzene or 3-Bromobenzonitrile, CuI, N,N'- Dimethylethylenediamine, K$_3$PO$_4$, dioxane dry, MW, 130 °C, 30 min, 75 – 80 %

Reduction of nitro (**108a, c**) and cyano (**108b, d**) substituted derivatives, using Zn in AcOH and NaBH$_4$ in presence of catalytically amount of NiSO$_4$•6H$_2$O respectively, followed by coupling reaction with L-Boc-Ala-OH or D-Boc-Ala-OH provided the Boc-protected alanine-benzylamides **110 a** - **d** and *ent*-**110 a** - **d** (**Scheme 3.10**), which then were deprotected with trifluoroacetic acid (**compounds 111 a** - **d** and *ent*-**111 a - d**). Hydrolisis of esters **110c, d**; *ent*-**110c, 111c, d** and *ent*-**111c d** with LiOH gave the corresponding acids **110 e - f**, *ent*-**110 e - f**, **111 e - f** and *ent*-**111 e - f**.
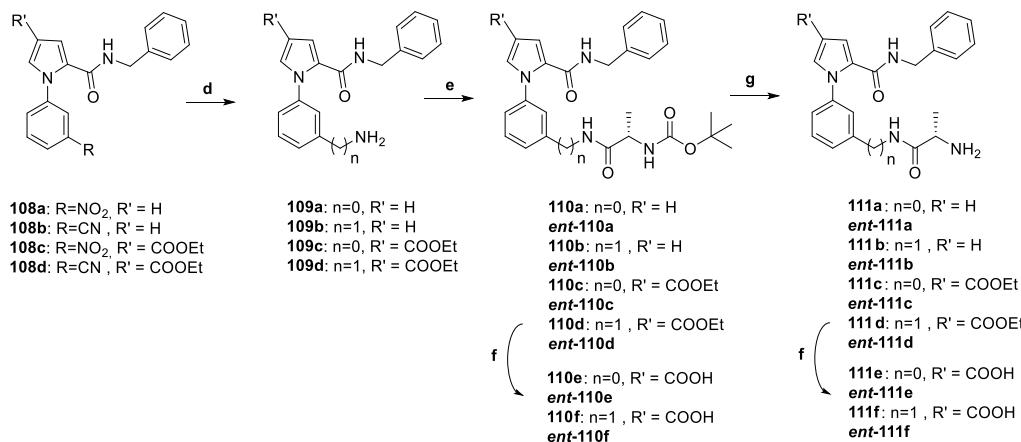
Scheme 3.10. Reagent and condition: (**d**) Zn, AcOH, r. t. 1 h, (for **108a, c**) or NaBH$_4$, NiSO$_4$•6H$_2$O, THF/iPrOH, r. t., 4 h (for **108c, d**); (**e**) Boc-L-Ala-OH (for **109a-d**) or Boc-D-Ala-OH (for *ent*-**109 a-d**), DCC, DMAP, DCM, r. t., overnight, 50 – 70% (over step **d** and **e**); (**f**) LiOH, 40 °C, EtOH/H$_2$O, overnight, 90 – 95%; (**g**) DCM/TFA (9:1), r. t., 30 min, 99%.

3.1.3. Synthesis of N-Benzyl derivatives (Series III)

3.1.3.1. Synthesis of N-Benzyl pyrrolic scaffold

The pyrrolic scaffold of **Series III** derivatives was obtained employing a 1, 3 dipolar cycloaddition reaction between ethylpropiolate and 1, 3 diacylic intermediates (**115 a, b**).[194] Apparently, this procedure compared to the Clauson-Kaas reaction seems to be not convenient due to a greater number of synthetic steps. However, it allowed to obtain the desidered N-benzyl pyrrole building blocks (**116a, b**) with high yields and starting from not expensive reagents (**Scheme 3.11**). Moreover this procedure did not require intermediate steps of purification.

The first reaction is a reductive amination, the reaction between a glycine ethyl ester and 3-nitro or 3-cyano benzaldheyde (**112a** or **112b**) led to the iminic intermediates, which in turns were reduced *in situ* with NaBH$_4$ providing the amino derivatives (**113a** and **113b**). Formylation of compounds **113a, b** accomplished by formic acid and acetic anhydride, followed by hydrolysis of the esteric moiety with NaOH 2N led to the 1, 3 diacylic derivatives **115a** and **115b**. The N-Benzyl pyrrolic scaffold (**116a** and **116b**) was obtained by reaction of 1, 3 dipolar cycloaddition between ethylpropiolate and compounds **118a** and **118b** in acetic anhydride heating at 140 °C (**Scheme 3.11**).
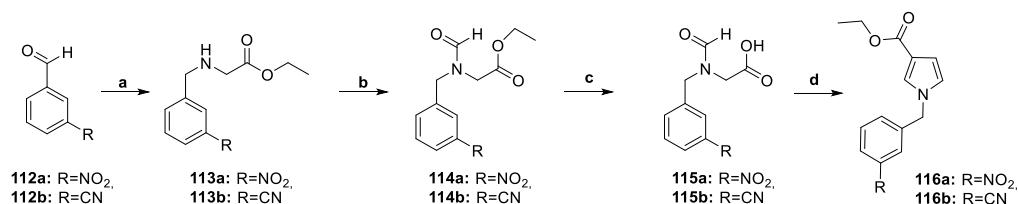
Scheme 3.11. Reagent and condition: (**a**) Gly-OEt•HCl, TEA, EtOH dry, r. t., overnight, NaBH$_4$, 0 °C, 1 h, 70 – 80 %; (**b**) Ac$_2$O, formic acid, r.t., overnight 80 – 90 %; (**c**) NaOH 2 N, EtOH, r. t., 1 h, 80 – 90 %; (**d**) ethylpropiolate, Ac$_2$O, 140 °C, 5 h, 60 – 70 %;

3.1.3.2. N-Benzyl pyrrolic scaffold decoration

The pyrrole scaffolds of derivatives **116a, b** was then functionalized employing the same synthetic approaches used for **pyrrole analogue of lead compound I** and for derivatives of the **Series II** (*paragraphs 3.1.1.2* and *3.1.2.2*). In particular the acylation with trochloroacetyl chloride (**117a, b**) followed by nuclephilic substitution with benzylamine provided the introduction of the benzylamide moiety at the C-2 position of the pyrrole ring (**118a, b**). Whereas the coupling reaction with L-Boc-Ala-OH or D-Boc-Ala-OH, after the reduction of the nitro or cyano groups with Zn in AcOH (**119a**) or NaBH$_4$ in presence of catalytically amount of NiSO$_4$•6H$_2$O (**119b**), led to the boc protected alanine-amides **120a, b** and *ent*-**120a**, **b**. Removal of the tert-butyloxycabonyl protecting group with DCM/TFA 9:1 yielded the amino derivatives **121a-d** and *ent*-**121a-d**, which in turn where hydrolysed with LiOH obtaining the corresponding acids **120c-d**, *ent*-**120c-d**, **121c-d** and **e***nt*-**121c-d** (**Scheme 3.12**).
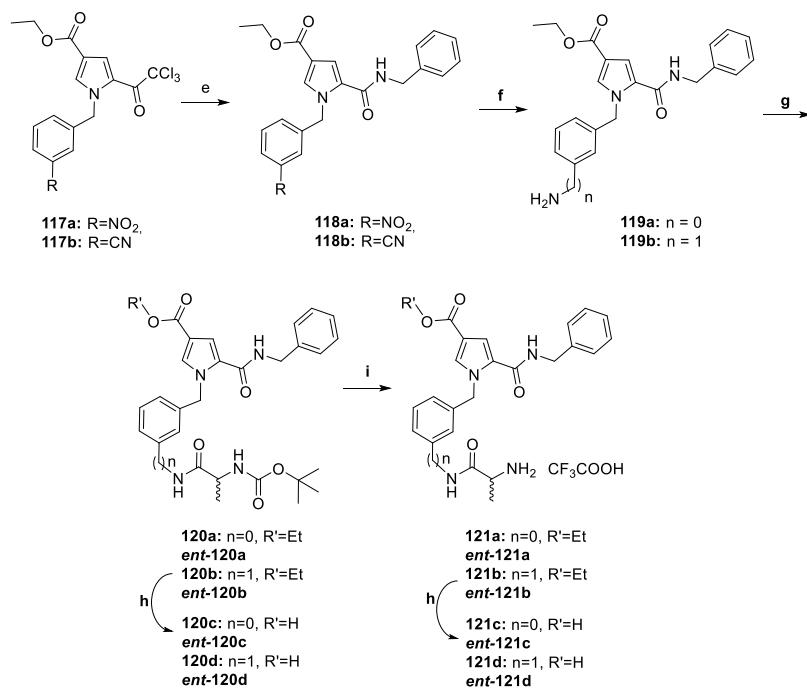
Scheme 3.12. Reagent and condition:(e) ClCOCCl$_3$**,** MW 120 °C, 1.5 h, 60% (**f**) Zn, AcOH, r. t., 1 h (for **118a**) or NaBH$_4$, NiSO$_4$•6 H$_2$O, THF/iPrOH, r. t., 4 h (for **118b**); (**g**) Boc-l-Ala-OH (for **120a, b)** or Boc-d-Ala-OH (for **ent-120a, b)**, DCC, DMAP, DCM, r. t., overnight, 60 – 70 % (over step **c** and **d**); (**h**) LiOH, EtOH/H$_2$O, 45 °C, 5 h, 90 – 95 %; (**i**) DCM/TFA 9:1, r. t., 30 min, 99 %.

3.2. Synthesis of bisubtrate inhibitors of CARM1

From a synthetic point of view the assembly of the bisubstrate inhibitors required a convergent synthesis, based on the initial preparation of both peptides and Adenosine fragments. Subsequently, the obtained building blocks were coupled employing, on solid phase, a synthetic procedure reported by Martin N. I. and coworkers for the preparation of N^G – substituted guanidines.[161]

3.2.1. Synthesis of peptide sequences

The peptide fragments (**124** and **127**) were obtained employing an acid sensitive resin (TentaGel® S RAM Resin) and using standard Fmoc SPPS techniques (Solid Phase Peptide Synthesis) (**Scheme 3.13 A** and **B**). The synthesis of the two sequences began respectively at the Proline (**122**) and Methionine (**125**) residues and an Alloc (allyloxycarbonyl) protected L-Ornithine was incorporated instead of L-Arginine. Subsequently, the protecting group on the side chain of Ornhitine residue was selectively removed [162] with Pd(PPh₃)₄ and PhSiH₃ in CH₂Cl₂ under argon atmosphere providing an attachment point for the adenosine portions.

A)

FmocN-Pro—○ →(a) Ac-Phe-Gln(Trt)-Asn(Trt)-Met-Phe-Gly-Ala-Ile-Orn-Pro-Ala-Ala-Pro—○ →(b)

122   **123**

Ac-Phe-Gln(Trt)-Asn(Trt)-Met-Phe-Gly-Ala-Ile-Orn-Pro-Ala-Ala-Pro—○

124

B)

FmocHN-Met—○ →(a) Ac-Pro-Ala-Ala-Pro-Orn-Pro-Pro-Phe-Ser(tBu)-Thr(tBu)-Met—○ →(b)

125   **126**

Ac-Pro-Ala-Ala-Pro-Orn-Pro-Pro-Phe-Ser(tBu)-Thr(tBu)-—○
Met

127
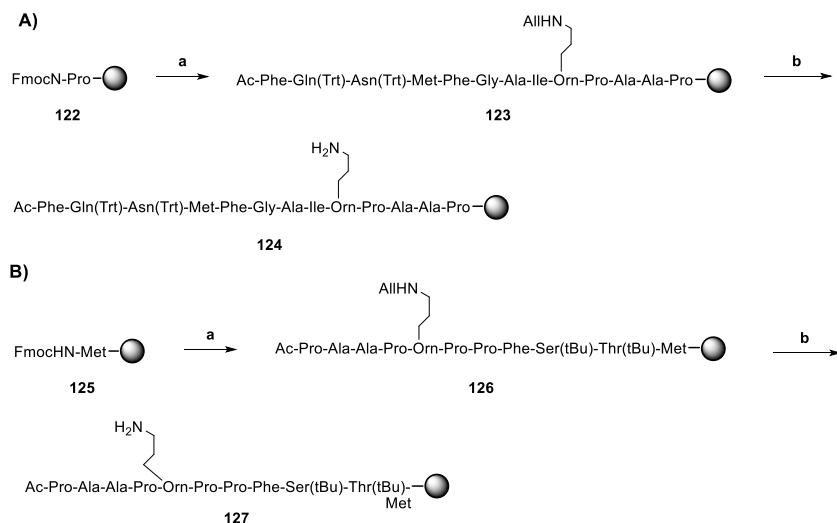
Scheme 3.13. Reagents and conditions: (**a**) Fmoc-Xaa-OH, BOP, DIPEA, DMF; (**b**) Pd(PPh$_3$)$_4$, PhSiH$_3$, CH$_2$Cl$_2$.

3.2.2. Synthesis of Adenosine building blocks

The synthetic routes used for the preparation of the Adenosine Pbf-protected thioureas **131**, **137** and **138** all started from the commercially available 2,3-O-isopropylideneadenosine building block **128**.

The synthesis of **131** (**Scheme 3.14**) began with the one carbon homologation of **128** via a Mitsunobu reaction employing PPh$_3$, DEAD, and acetonecyanohydrin to yield the nitrile **129**. Subsequent reduction to the amine (**130**) followed by treatment with Pbf-NCS provided the Pbf-protected thiourea species (compound **131, Scheme 3.14**).

128   **129**   **130**   **131**
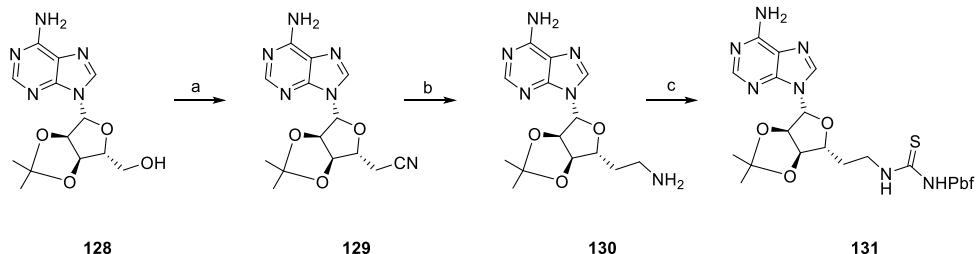
Scheme 3.14. Reagents and conditions: (**a**) acetone cyanohydrin, PPh$_3$, DEAD, THF, 90%; (**b**) PtO$_2$, AcOH, H$_2$, quant.; (**c**) PbfNCS, Et$_3$N, CH$_2$Cl$_2$, 50%.

On the other hand the synthesis of **137** and **138** (**Scheme 3.15**) began with a one pot IBX oxidation of **128** to yield the intermediate aldehyde which was directly converted to alkene **132** via a Wittig reaction with triphenylcarbethoxymethylenephosphorane. Reduction of **132** with DIBAL yielded an alcohol which was transformed into intermediate **134** by a Mitsunobu reaction employing PPh3, DEAD, and phthalimide.[142] Removal of the phthalimide group with methylamine provided amino compound **135,** which was converted into Pbf-protected thiourea species **137** by treatment with Pbf-NCS. Hydrogenation of **135**[163] followed by reaction with Pbf-NCS using the procedure described above for **131** and **137** yielded compound **138.**
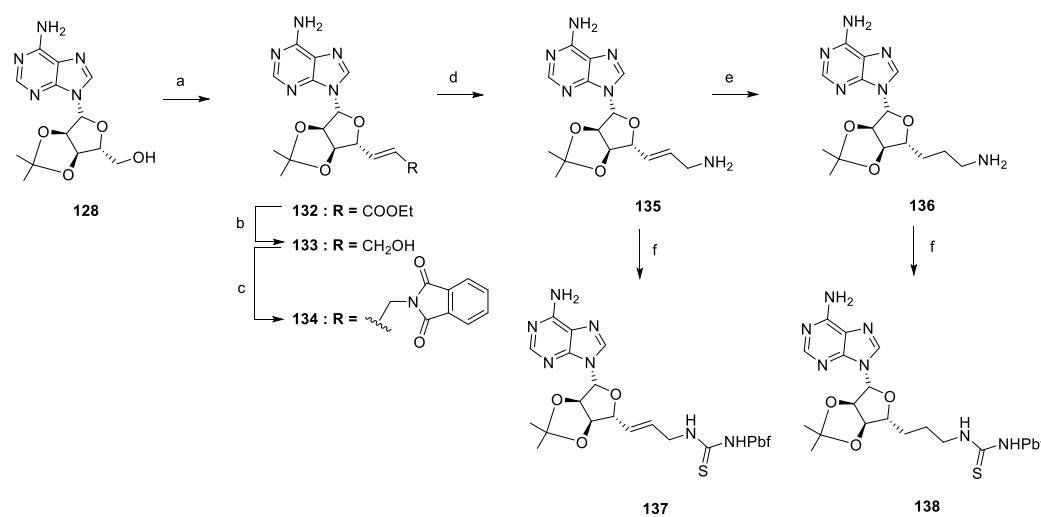
Scheme 3.15. Reagents and conditions: (**a**) IBX, Ph3P=CHCO2Et, DMSO, 79%; (**b**) DIBAL-H, hexane, CH2Cl2, 78%; (**c**) phthalimide, PPh3, DEAD, THF, 83%; (**d**) MeNH2, EtOH, 94%; (**e**) 10 % Pd/C, H2EtOH, quant.; (**f**) PbfNCS, Et3N, CH2Cl2, 50 – 60%.

3.2.3. Synthesis of bisubstrate inhibitors

The obtained fragments (adenosine and peptide building blocks) were coupled through guanidine moiety, using a synthetic methodology previously described by Martin N. I. and co-workers.[161] This procedure, reported for the synthesis

of N^G-substituted guanidines, proved to be suitable also for the solid phase techniques. The coupling reaction between the proper ornithine containing peptide on resin (**124** or **127**) and Pbf-protected adenosine thiourea species (**131**, **137** and **138**), was performed using EDCI as a coupling reagent (**Schemes 3.16**). Final deprotection/cleavage and purification by using RP-HPLC allowed us to obtain all of the designed compounds (**80 – 85**).
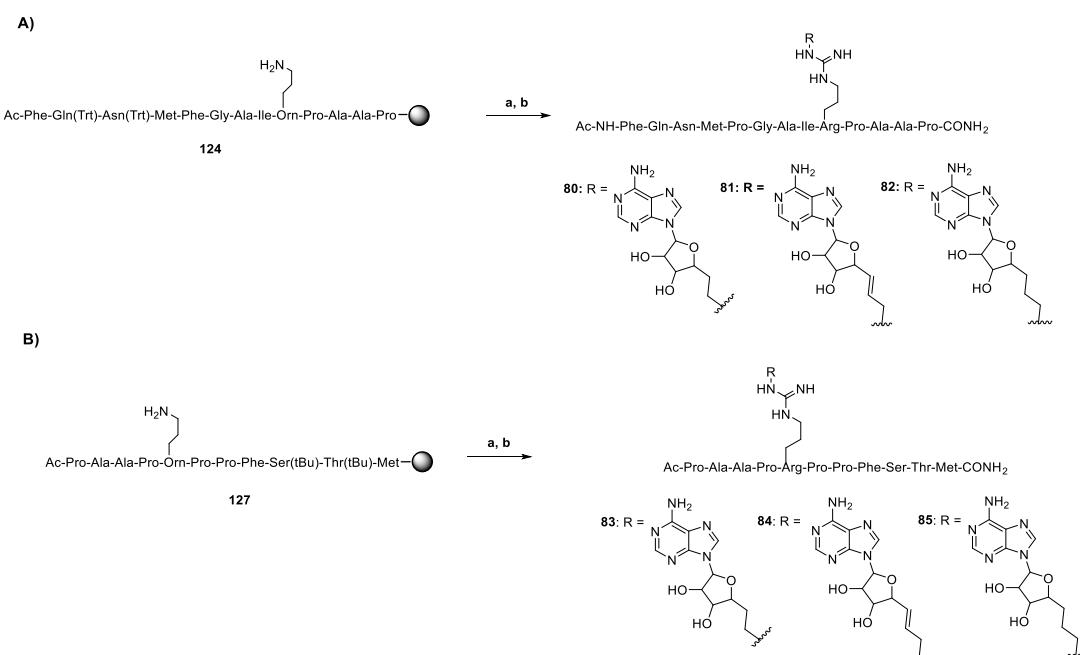
Scheme 3.16. Reagents and conditions: (**a**) EDCI, proper Pbf-protected adenosine thiourea (**131**, **137** or **138**), CH$_2$Cl$_2$; (**b**) TFA–TIS–H$_2$O.

3.2.4. Synthesis of control compounds

Moreover, we decided to synthesize the two selected peptide sequences bearing an unmethylated (compounds **139** and **142, Scheme 3.17**), singly methylated (compounds **140** and **143**, **Scheme 3.18**) or asymmetric dimethylated (compounds **141** and **144**, **Scheme 3.17**) arginine residues.

Peptides **139 – 144** were obtained employing a standard Fmoc SPPS technique (Solid Phase Peptide Synthesis) using the procedure described above for **124** and **127 (Scheme 3.13)**.
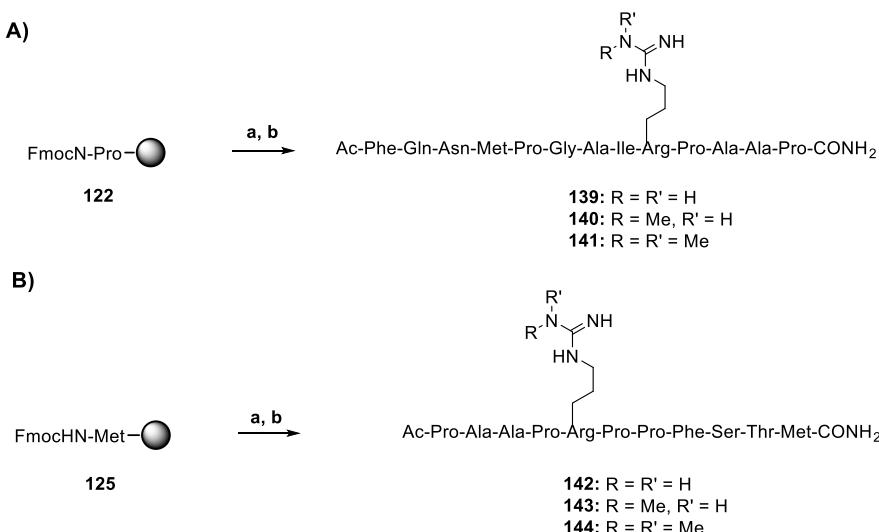
Scheme 3.17. Reagents and conditions: (**a**) Fmoc-Xaa-OH, BOP, DIPEA, DMF; (**b**) TFA–TIS–H$_2$O.

The modified Arginine building block used for the peptide synthesis was obtained employing a synthetic procedure previously describes by the group of Martin and coll. (**Scheme 3.18**).[161] The Pbf/Fmoc-protected thiourea precursor (**146**) was obtained after treatment of deprotected L-ornithine (**145**) with PbfNCS in the presence of TEA. The reaction between **146** and methylamine or dimethylammine with EDCI yielded the expected modified- L-arginine (**147 - 148**).
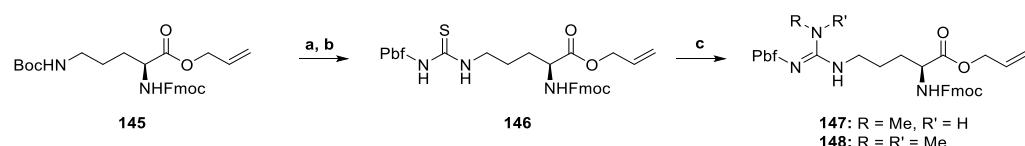
Scheme 3.18. Reagents and conditions: (**a**) TFA/CH$_2$Cl$_2$ 3: 7, r. t. 30 min; (**b**) PbfNCS, CH$_2$Cl$_2$, TEA, r. t., 30 min.; (**c**) EDCI, methylamine or dimethylamine, CH$_2$Cl$_2$, r. t.; 1h.

3.3. Indole inhibitors of PRMT3

The indole inhibitors of PRMT3 have been structurally categorized in three Series according to the nature of the central linker (**Figure 3.2**)
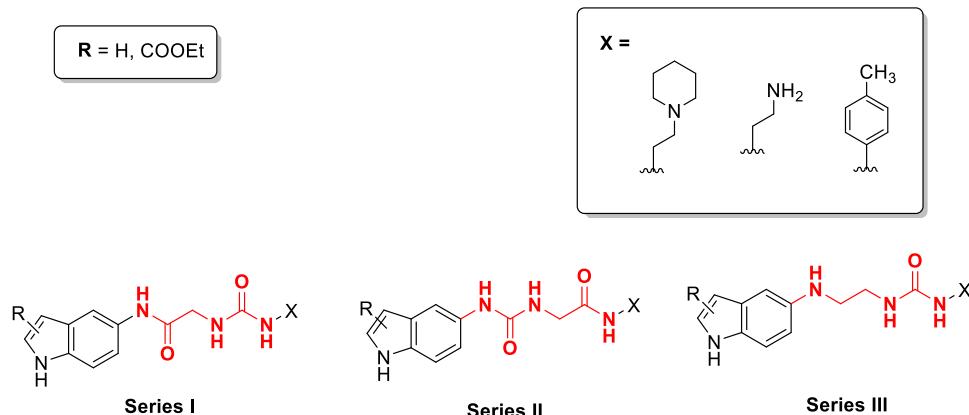
Figure 3.2. Three Series of indole based derivatives

3.3.1. Indole scaffold synthesis

Derivatives of **Series I**, **II** and **III** were prepared starting from 5-nitro-1H-indole (**149a**), ethyl 5-nitro-1H-indole-3-carboxylate (**149b**) and ethyl 5-nitro-1H-indole-2-carboxylate (**149c**) scaffolds (**Scheme 3.19** and **3.20**). Compound **149a** is commercially available, whereas the 3,5 and 2,5 substituted indole scaffolds were synthesized following the previously described procedures.[126]

Ethyl 5-nitro-1H-indole-3-carboxylate (**149b**) was prepared according to the synthetic approach reported in **Scheme 3.19**. The first step was a regioselective Friedel-Crafts acylation of 5-nitroindole (**149a**) with oxalyl chloride (**150**) followed by alkaline hydrolysis (**151**). Oxidative decarboxylation (**152**) and final etherification yielded the 3, 5 substituted indole scaffold (**149b**).

On the other hand, the 2,5-substituted analogue (**149c**) was obtained employing one of the best method for preparing indoles: the Fischer indole synthesis (**Scheme 3.20**).[164] The reaction by 4-nitrophenylhydrazine with

ethyl-piruvate led to the formation of the phenylhydrazone (**154**), which in turn was converted to ethyl 5-nitro-1H-indole-2-carboxylate (**149c**) under acid condition with polyphosphoric acid.
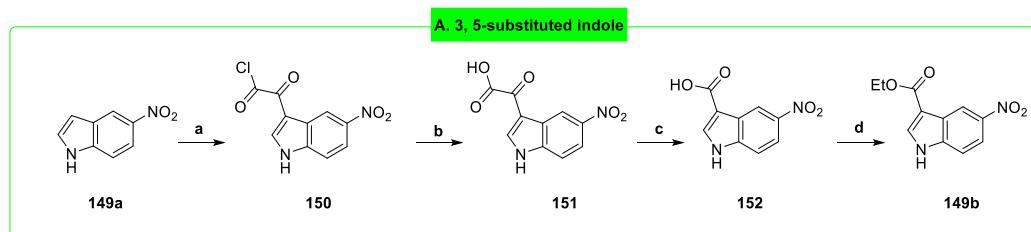
Scheme 3.19. A. Reagents and conditions: (**a**) (COCl)$_2$, Et$_2$O, 0 °C, overnight; (**b**) KOH, H$_2$O, reflux, 2 h; (**c**) H$_2$O$_2$ (10 % wt in H2O), reflux, 4 h; (**d**) H$_2$SO$_4$, MeOH, reflux, 6 h.
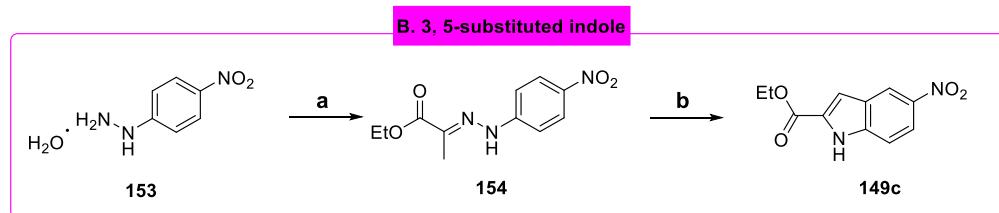
Scheme 3.20. B. Reagents and conditions: (**a**) AcOH/EtOH 1:2, 55°C, 1 h; ethyl-piruvate, r. t., overnight; (**b**) PPA, toluene, 120°C, 5 h.

3.3.2. Aromatic nitro reduction

Reduction of nitro compounds **149a – c** to the corresponding 5-amino derivative **155a – c** (**Scheme 3.21**), under continuous-flow catalytic hydrogenation conditions, provided an attachment point for the scaffold decoration. The continuous flow aromatic nitro reduction was preferred over traditional metal mediated non-catalytic reduction procedures, due to its efficiency, high yielding and ease in scale up. Moreover, since the products were delivered in high intrinsic purities, no purification steps were required.

149a: R= H; R'= H
149b: R= H; R'= COOEt
149c: R= COOEt; R'= H

155a: R= H; R'= H
155b: R= H; R'= COOEt
155c: R= COOEt; R'= H
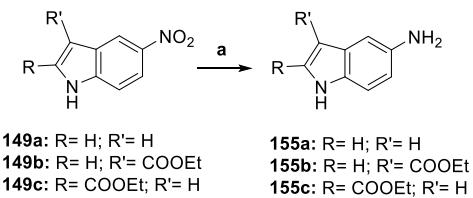
Scheme 3.21. Reagents and conditions: (**a**) H$_2$, 10% Pd/C, EtOH/AcOEt 1:1, 30 °C, 10 bar, 1mL/min, H-Cube™ (ThalesNano©).

3.3.3. Synthesis of Series I derivatives

Once obtained the 5-amino indolic scaffolds (**155a – c**), compounds of **Series I** were prepared according to the previously synthetic procedures reported by us for the synthesis of Uracandolates (Chapter 2 Aim of the work)[126]. Coupling reaction followed by deprotection and ureidic moiety formation allow us to obtaine the designed derivatives of **Series I**.

3.3.3.1. Coupling reaction and removal of protecting group

The reaction of **155a - c** with Fmoc-Gly-OH or Boc-Gly-OH in the presence of the peptide coupling reagents hydroxybenzotriazole (HOBt) and O-benzotriazole-N,N,N′,N′-tetramethyl-uronium-hexa-fluoro-phosphate (HBTU) and N,N-diisopropylethylamine, DIEA yielded the Fmoc or Boc protected glycinamides (**156 a – c**), which were deprotected with piperidine or trifluoroacetic acid yielding the corresponding amino derivatives (**157a – c**), **Scheme 3.22**.

155a: R= H; R'= H
155b: R= H; R'= COOEt
155c: R= COOEt; R'= H

156a: R= H; R'= H; R" = Fmoc
156b: R= H; R'= COOEt; R"= Boc
156c: R= COOEt; R'= H; R"= Boc

157a: R= H; R'= H; R" = H
157b: R= H; R'= COOEt; R"= H
157c: R= COOEt; R'= H; R"= H
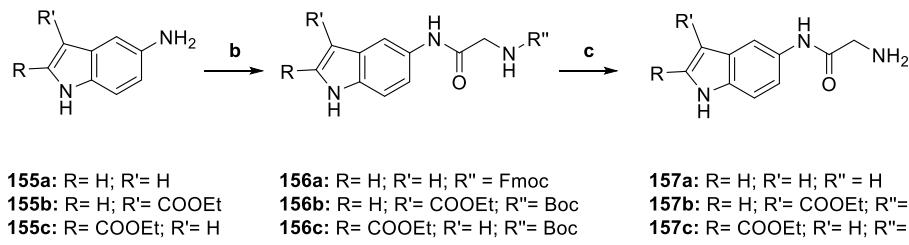
Scheme 3.22. Reagents and conditions: (**b**) Fmoc-Gly-OH or Boc-Gly-OH, DIPEA, HOBt, HBTU, THF/DMF 4:1, r.t., overnight; (**c**) CH$_2$Cl$_2$/piperidine 8:2, r.t., 30 min or CH$_2$Cl$_2$/TFA 8:2, r.t., 30 min.

Notably, it was not possible to couple compound **155a** with Boc-Gly-OH, since the traditional acidic conditions used for removing the Boc protecting group led to the degradation of the 5-substituted indole scaffold (**155a**), which conversely resulted to be stable in basic condition. Therefore we used two different protecting groups, Fmoc (**156a**) or Boc (**156b, c**) for the glycinamide moieties.

3.3.3.2. Ureidic moiety formation

Derivatives **157a – c** by reaction with commercially available 4-methyl phenylisocyanate (**158**) in the presence of triethylamine yielded final compounds **159a – c** bearing an aromatic moiety at the right end of the molecules (**Scheme 3.23**).
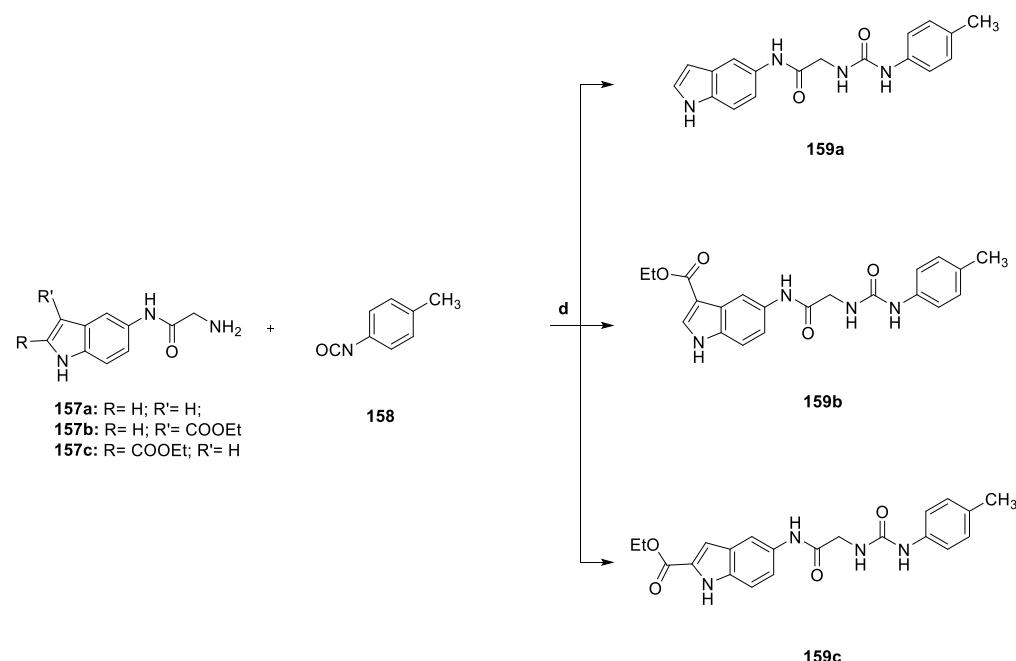
Scheme 3.23. Reagents and conditions: (**d**) TEA, THF, r. t., 8 h.

Instead, the formation of the ureidic moieties of compounds **161a – c** and **163a – c** was accomplished by reaction of **157 a – d** with a proper amine (1-(2-Aminoethyl)-piperidine **160** or protected ethylendiamines (**162a, b**) in the

presence of CDI and TEA (**Schemes 3.24** and **3.25**). Removal of Boc protecting group of **163b, c** with TFA led to the corresponding amino compounds **164b, c (Scheme 3.25)**. On the other hand, being the 5 substituted scaffold (**157a**) unstable under acidic conditions, the N-Boc ethylenediamine moiety (**162b**), used for compounds **163b, c**, was replaced with 2-azidoethylamine (**162a**), which by reaction with **157a** in the presence of CDI and TEA, followed by continuous-flow catalytic hydrogenation yielded the final compound **164a.**
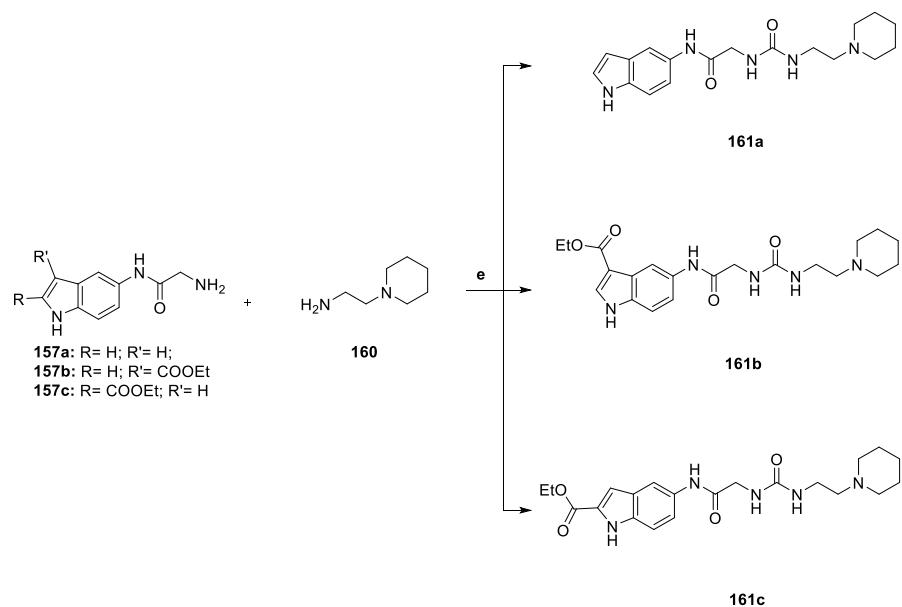
Scheme 3.24. Reagents and conditions: (**e**) CDI, TEA, 0 °C, 5 min, **9a**, **9b** o **9c** THF, 2 die
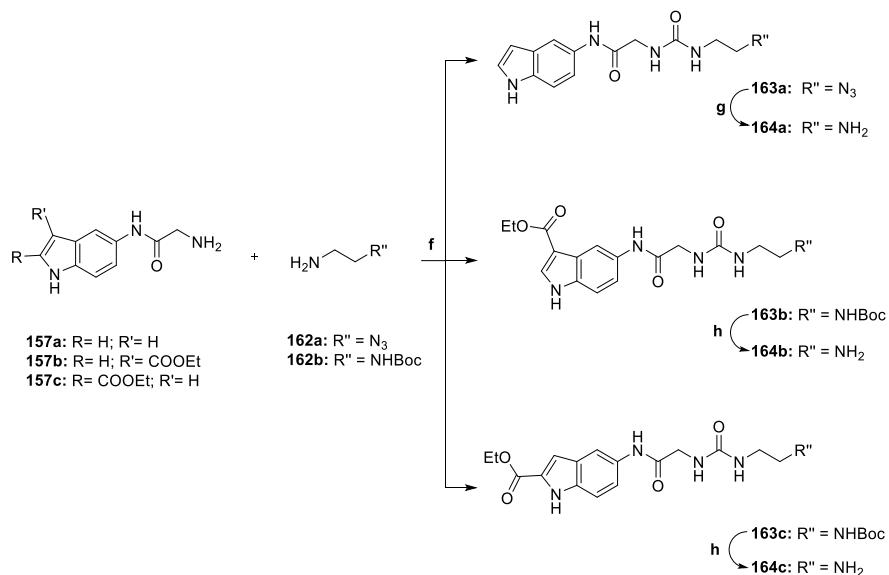
Scheme 3.25. Reagents and conditions: (**f**) CDI, TEA, 0 °C, 5 min, THF, 2 die; (**g**) H$_2$, 10% Pd/C, MeOH, 30 °C, 1 bar, 1mL/min, H-Cube™ (ThalesNano$^©$); (**h**) DCM/TFA 8:2, r. t., 30 min.

3.3.4. Synthesis of Series II derivatives

The synthesis of **Series II** compounds required the initial formation of the ureidic function on the heterocycles, followed by coupling reactions for introducing aliphatic or aromatic substituents at the right end of the molecule.

3.3.4.1. Urea formation and deprotection

The ureidic derivatives (**165 a – c**, **Scheme 3.26**) were easily prepared by reaction between 5-amino indolic derivatives (**158 a - c**) and glycine tert-butyl or ethyl esters in the presence of CDI and TEA. The corresponding acids (**166 a – c**) were then obtained by hydrolysis under basic (**for 165a**) or acidic conditions (for **165b, c**) of the esters. Also in this case were used two different protecting groups for the aliphatic carboxylic acid of the indole derivatives (**165a – c**). The instability of **155a** under acidic conditions didn't allow us to protect the aliphatic carboxylic moiety with an acid labile tert-butyl ester. In addition could not be used the glycine ethyl ester for 3, 5 and 2, 5 disubstituted

indole scaffolds (**155b** and **155c**), since the presence of an aromatic ethyl ester group limited the selective hydrolysis of the aliphatic ester.
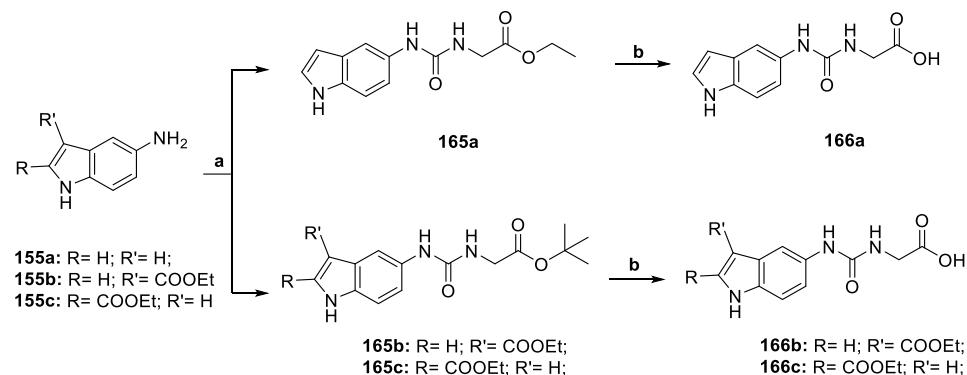
Scheme 3.26. Reagents e conditions: (**a**) CDI, TEA, glicyne ethyl ester o glycine tert-butyl ester, THF, overnight; (**b**) NaOH 2N, EtOH, r. t., 2 h or DCM/TFA 8:2, r. t., 30 min.

3.3.4.2. Coupling reactions

Final coupling reaction between compounds **166a – c** and a proper amine, p-toluidine (**Scheme 3.27**, compounds **168 a – c**), 1-(2-Aminoethyl)-piperidine (**Scheme 3.28**, compounds **169 a – c**) and **162a, b** (**Scheme 3.29**, compounds **170 a – c**) respectively, in the presence of coupling reagent TBTU accomplished the derivatization of the aliphatic carboxylic moieties. Compounds **170 b – c** were converted to the corresponding amino compounds (**171 b – c**) by removal Boc protecting group with TFA (**Scheme 3.29**), whereas derivative **171a** was obtained after catalytic hydrogenation of the azido intermediate **170a** (**Scheme 3.29**).

Scheme 3.27. Reagents e conditions: (**c**) TBTU, TEA, DMF, r. t., overnight.

Scheme 3.28. Reagents e conditions: (**d**) TBTU, TEA, DMF, r. t., overnight.
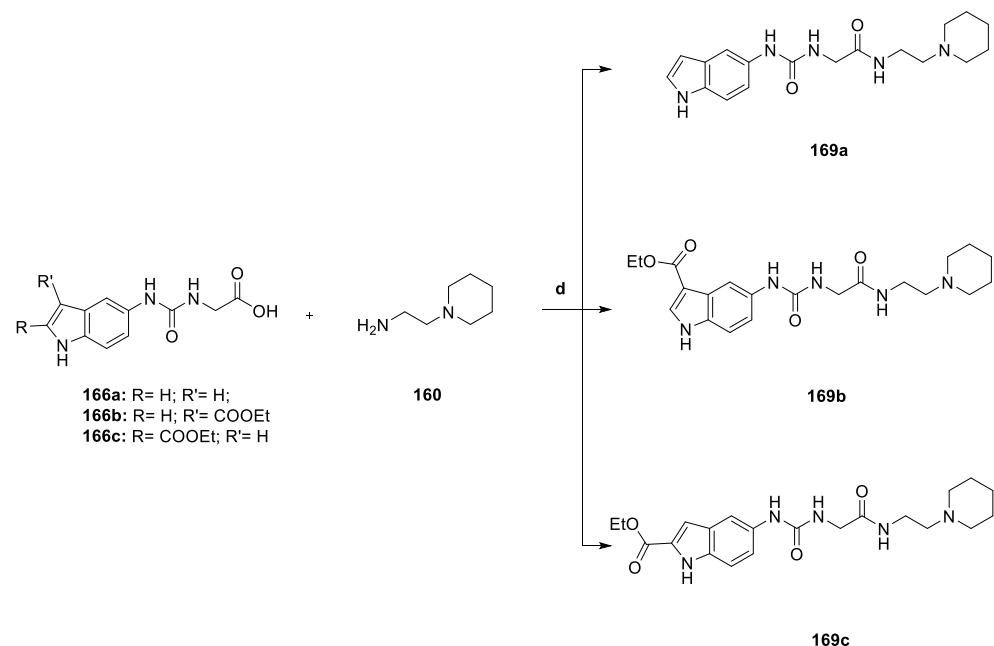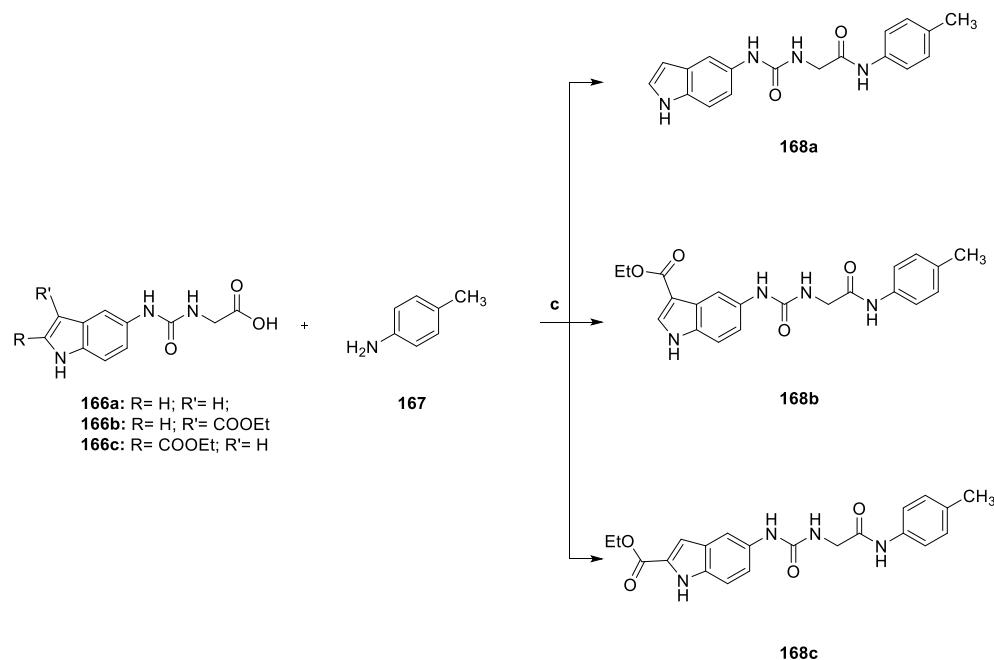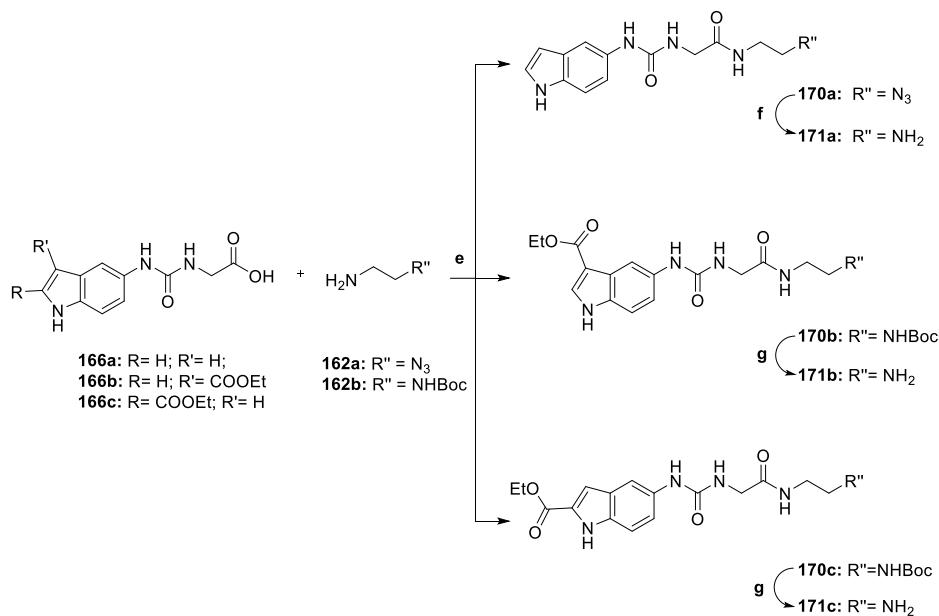
Scheme 3.29. Reagents e conditions: (**e**) TBTU, TEA, DMF, r. t., overnight; (**f**) H₂, 10% Pd/C, MeOH, 30 °C, 1 bar, 1mL/min, H-Cube™ (ThalesNano©); (**g**) DCM/TFA 8:2, r. t., 30 min.

3.3.5. Synthesis of Series III derivatives

Compounds of Series III were prepared by alkylation of 5 amino-indoles **155a – c** with methanesulfonates ureidic intermediates (**174 a – c**), previously synthesized.

3.3.5.1. Synthesis of 174a - c building blocks

The mesyl activated ureidic intermediates were synthesized according to the synthetic route depicted below (**Scheme 3.30**). Compound **174a** was prepared by reaction of 4-methy-phenylisocyanate (**158**) with 2-aminoethanol, followed by OH-activation with mesyl chloride. While the synthesis of compounds **174b, c** required one more step. Initially, the amino building blocks **162a, b** were treated with phenyl-chloroformate in order to obtain the carbamate intermediates (**172a, b**), which were converted into the corresponding ureidic compounds by reaction with 2-aminoethanol in 1,2-dichloroetane at 120 °C for

8 h. Finally the methanesulfonates ureidic intermediates **174b, c** were prepared by reaction with mesyl chloride.
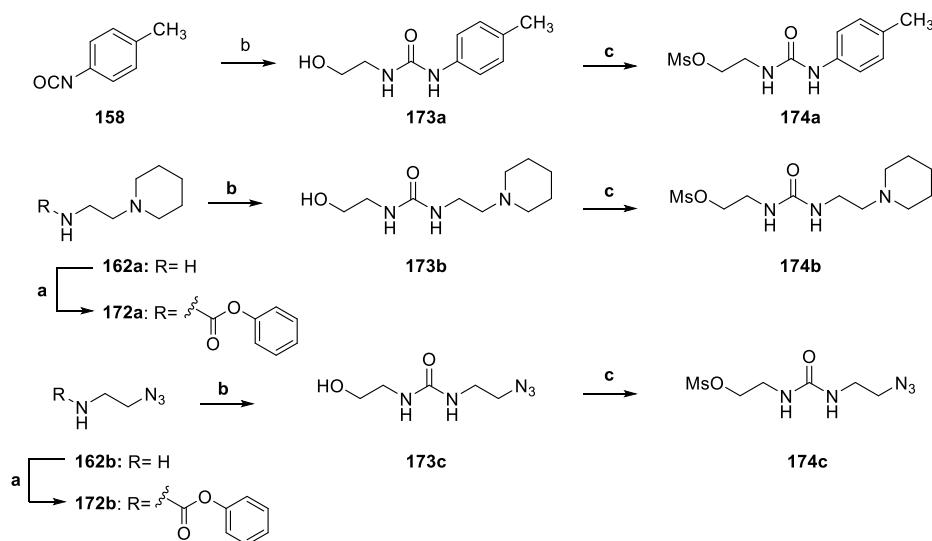
Scheme 3.30. Reagents e conditions: (**a**) phenyl-chloroformate, TEA, AcOEt, 0 °C, 4 h, 70%; (**b**) for **173a** , 2-aminoethanol, TEA, THF dry, 0 °C, 8 h, 70%; for **173b, c,** 2-aminoethanol, TEA, 1,2-dichloroetane, 120 °C, 8 h, 80%; (**c**) MsCl, TEA, THF dry, 0°C, 1 h, 100%.

3.3.5.2. Alkylation steps

Final compounds (**175a − c**, **176a − c**, **178b, c**) were obtained after treatment of 5-amino indole derivatives **155a − c** with methanesulfonates ureidic intermediates (**174a − c**), **Schemes 3.31**, **3.32**, **3.33**.

155a: R= H; R'= H;
155b: R= H; R'= COOEt
155c: R= COOEt; R'= H

174a

175a

175b

175c
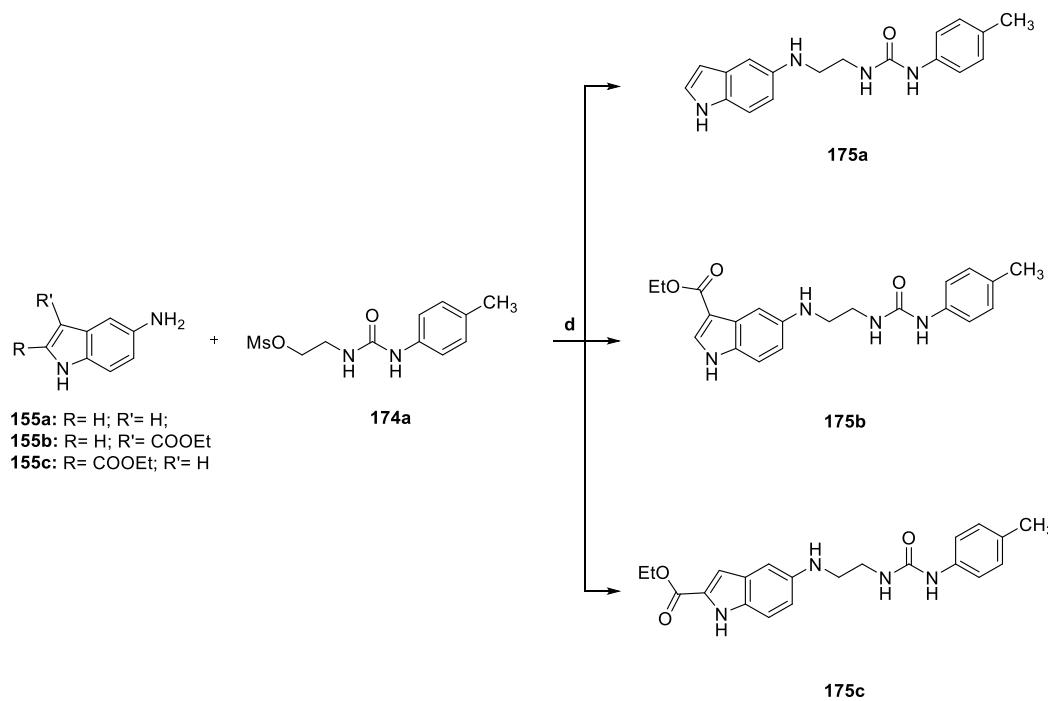
Scheme 3.31. Reagents e conditions: (**d**) TEA, THF dry, 50 °C, 2 h, 50-70 %.

155a: R= H; R'= H;
155b: R= H; R'= COOEt
155c: R= COOEt; R'= H

174b

176a

176b

176c
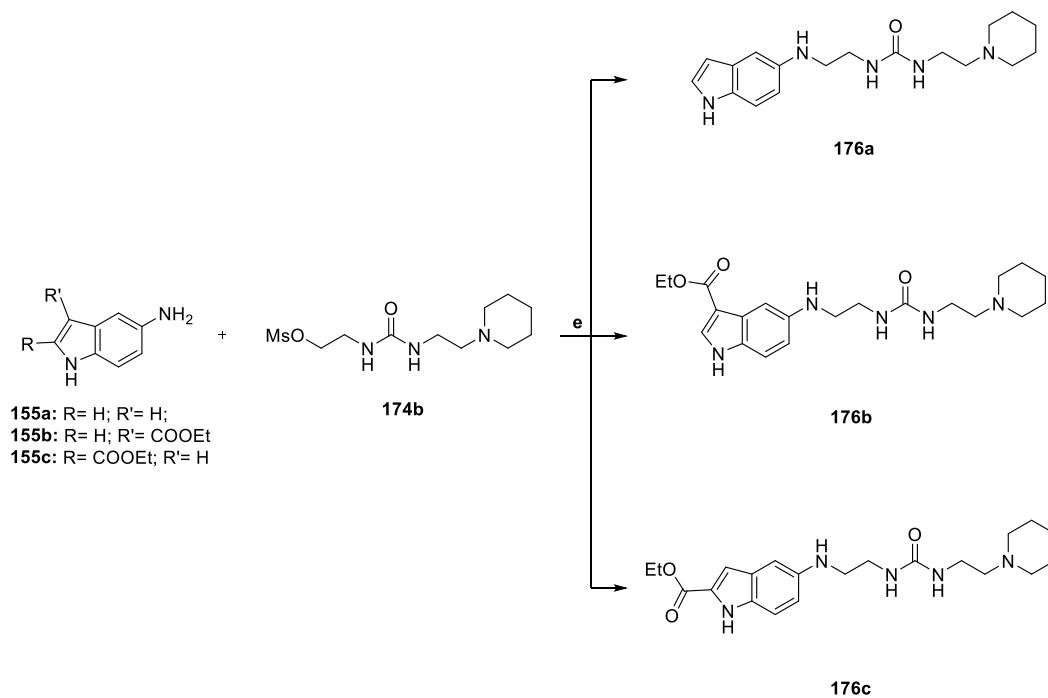
Scheme 3.32. Reagents e conditions: (**e**) TEA, THF dry, 50 °C, 2 h, 50-70 %.
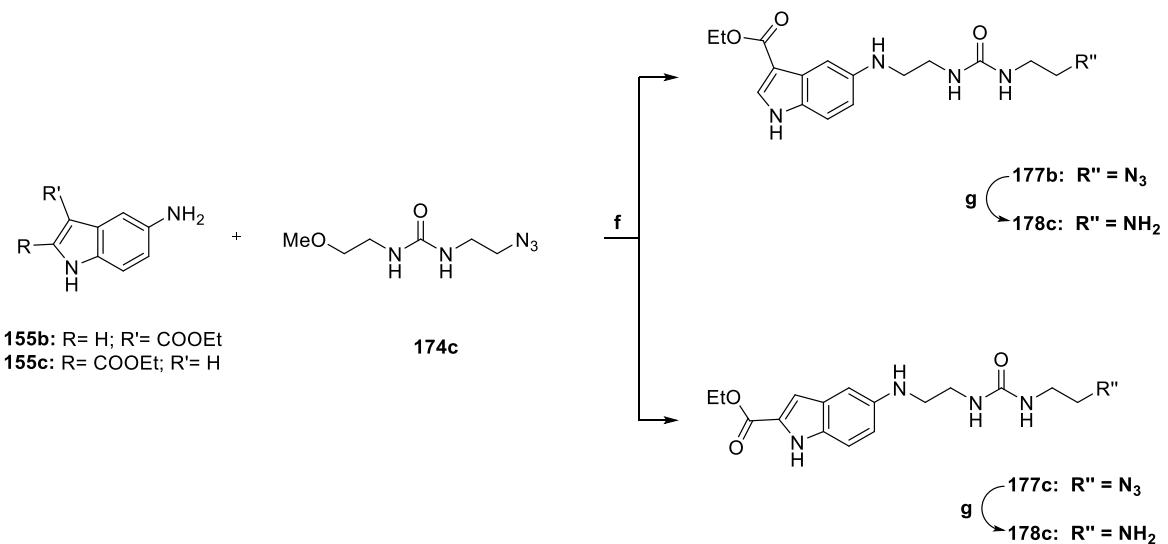
Scheme 3.33. Reagents e conditions: (**f**) TEA, THF dry, 50 °C, 2 h, 50-70 %; (**g**) H$_2$, 10% Pd/C, MeOH, 30 °C, 1 bar, 1mL/min, H-Cube™ (ThalesNano©);

CHAPTER 4

BIOLOGY

4.1. Biological evaluation of pyrrole derivatives

Pyrrole derivatives were preliminarily tested at fixed dose (50 μM) in duplicate by Scintillation proximity assay against human CARM1, using as substrates histone H3 at 5 μM and ^{3}H-SAM at 1 μM. The inhibition (%) at fixed dose was first determined, and then the IC_{50} values for the most active compounds were established. Pyrazole inhibitor (compound **54**) was used as our internal reference compound. All the enzymatic assays were performed by Reaction Biology corporation.

4.1.1. Results

In **Figure 4.1** are reported the results of CARM1 enzymatic inhibition assay at fixed dose by pyrrole compounds; whereas in **Table 4.1** are reported the structures of the tested pyrrole derivatives.
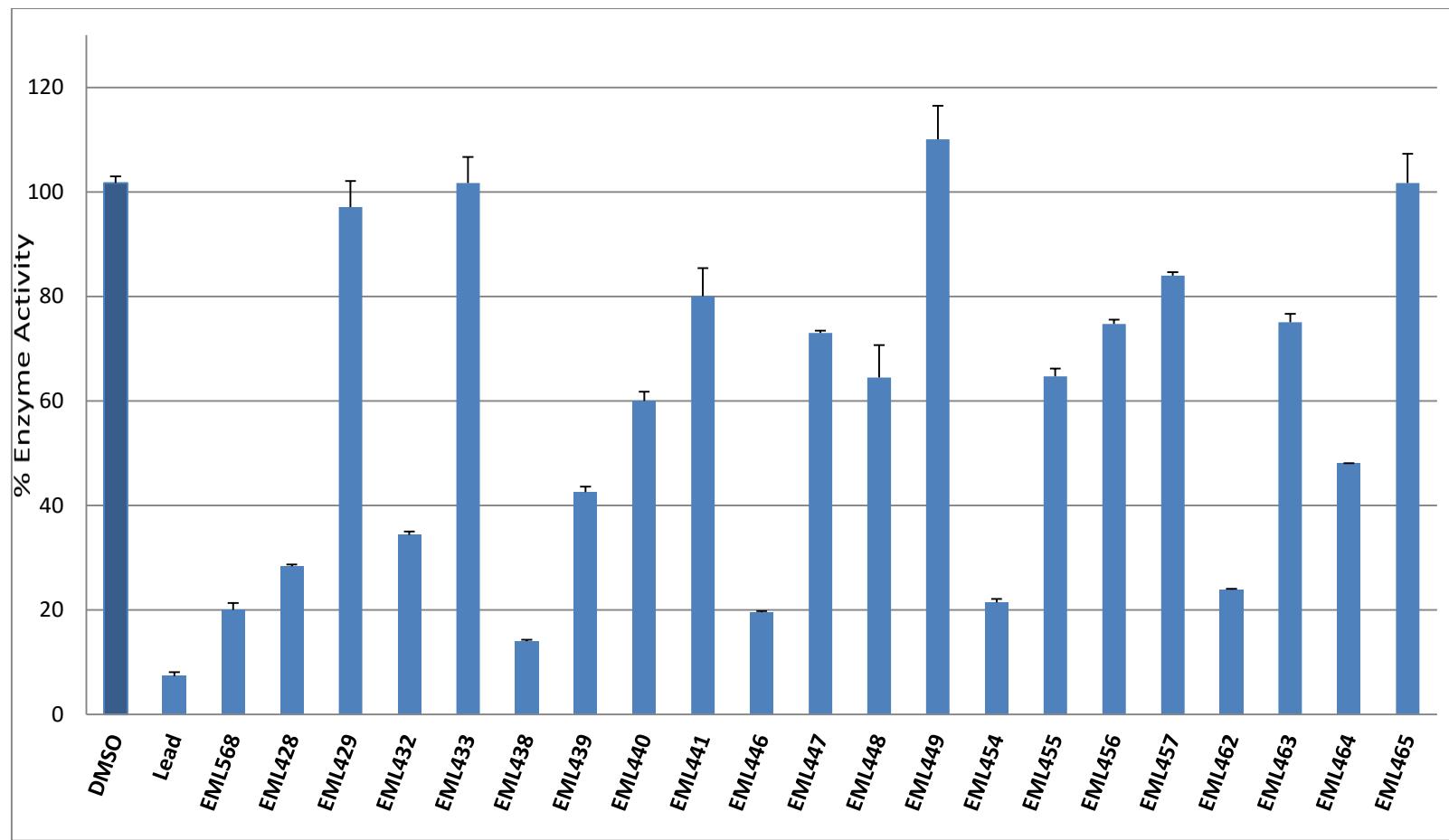
Figure 4.1. CARM1 inhibition by pyrrole inhibitors at fixed dose.

Table 4.1. Tested compound

Lead		Pyrrole analogue of lead I	
	54 Lead		**100** EML568
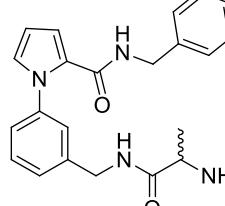
Series IIa					
	111b EML432 *ent*-**111b** EML433		**111d** EML446 *ent*-**111d** EML447		**111f** EML448 *ent*-**111f** EML449
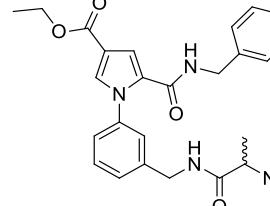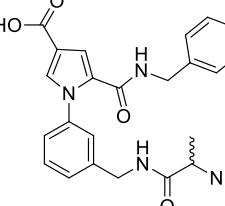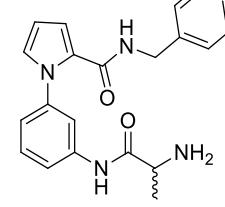
Series IIb					
	111a EML428 *ent*-**111a** EML429		**111c** EML438 *ent*-**111c** EML439		**111e** EML440 *ent*-**111e** EML441
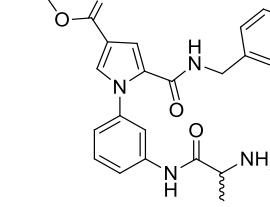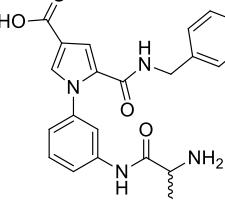
Series IIIa			
	121b EML462 ***ent*-121b** EML463		**121d** EML464 ***ent*-121d** EML465
Series IIIb			
	121a EML454 ***ent*-121a** EML455		**121c** EML456 ***ent*-121c** EML457
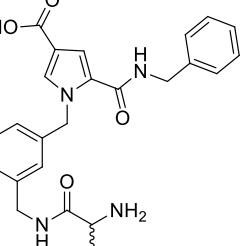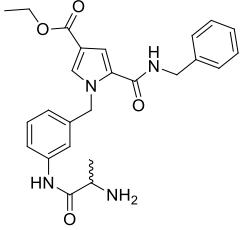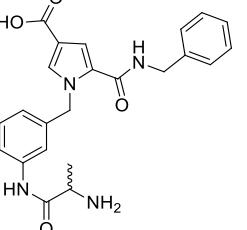
In Table **4.2** are reported the IC_{50} values of compounds that showed a CARM1 inhibition > 50% at 50 µM.

Compounds	IC_{50} (µM) CARM1
EML568	7.23
EML428	9.28
EML432	9.72
EML438	2.42
EML439	25.2
EML446	4.83
EML454	4.83
EML462	11.5
EML464	65.4
Lead	0.33

Table 4.2. IC_{50} values of selected compounds. Compounds were tested in a 10 dose IC_{50} mode, in singlet, with 3-fold serial dilution starting at 100µM.

4.1.2. SAR of pyrrole inhibitors of CARM1

These preliminary studies allowed us to delineate the structure-activity relationships of CARM1 pyrrole inhibitors.

The most important result that emerge from this preliminary screening is that the structure-activity relationships of pyrrole derivatives and pyrazolic inhibitors are not overlapping, suggesting that has been developed and identified a new class of CARM1 inhibitors, which, due to the high synthetic versatility of the pyrrole ring, showed several possibilities for further structural optimizations, whom are precluded for pyrazole compounds.

The activity revealed by the pyrrolic analogue of lead compound I (**EML 568**) confirmed our assumptions that the isosteric replacement of the pyrazolic ring with the pyrrole one could be also considered as a bioisosteric substitution. Indeed the aforementioned pyrrole derivative shows a good enzymatic inhibition against CARM1 (**EML 568**, $IC_{50} = 7.23$ µM, **Table 4.2**), however the pyrazole compound remains the most potent inhibitors of CARM1 (**Lead** $IC_{50} = 0.5$ µM) **Figure 4.2.**
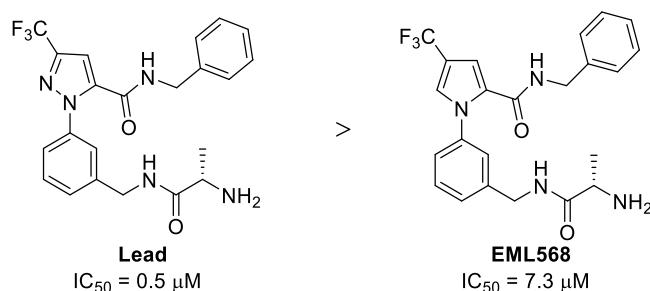
Figure 4.2. Effect of the substitution of the heterocycle central core of pyrazolic inhibitor

The only thing in common between the SAR of pyrrole and pyrazole derivatives is the stereospecific activity of compounds. Indeed considering the enantiomeric couples of Series II and III derivatives, the CARM1 inhibition is achieved by compounds bearing the (S)-alanine moiety at the south end of the

molecules, its replacement with the optical isomer (R)-alanine leads to a loss of activity (**Figure 4.3**). [71.6 % of enzymatic inhibition (**EML428**) *vs* 2.89 % of enzymatic inhibition (**EML429**); IC_{50} = 2.42μM (**EML438**) *vs* IC_{50} = 25.2μM (**EML439**), 80.5 % of enzymatic inhibition (**EML446**) *vs* 27.0 % of enzymatic inhibition (**EML447**)].
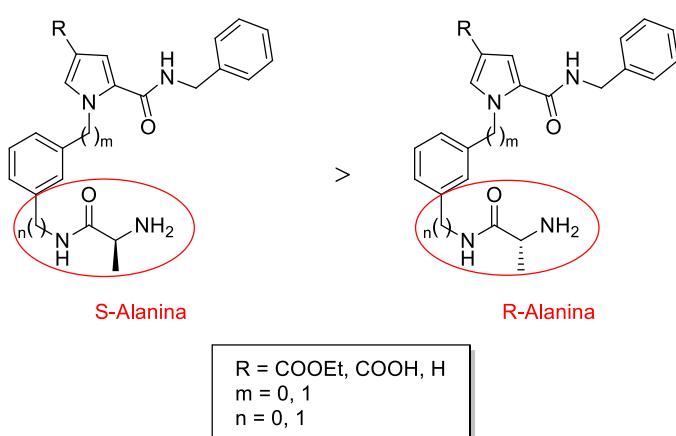
Figure 4.3. Comparison between the enantiomeric couples of derivatives

Regarding the effect of the different functionalization of the pyrrolic C-4 position emerge that the replacement of the trifluoromethyl group, present in analogue of lead compound I (**EML568**), with the ethyl ester moiety leads to an improvement of the inhibition activity [IC_{50} = 7.23 μM (**EML568**) *vs* IC_{50} = 4.83 μM (**EML446**)] (**Figure 4.4**), whereas its elimination determines a slight decrease of potency [IC_{50} = 7.23 μM (**EML568**) *vs* IC_{50} = 9.72 μM (**EML432**)] (**Figure 4.4**). Interestingly, the hydrolysis of the ester derivatives to the corresponding acid results in a total loss of activity. The carboxylic derivative (**EML448**) show the lowest inhibitory potency compared to the unsubstituted pyrrole derivative (**EML432**) as well as to the 4- trifluoro (**EML568**) and 4-ethyl ester (**EML446**) substituted inhibitors (**Figure 4.4**).
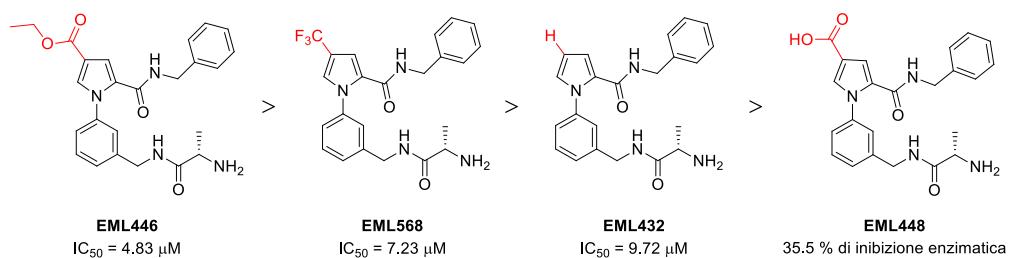
Figura 4.4 Effect of different substitution at pyrrole position 4 on the activity

Next the role of the central core was investigated. The replacement of the N-phenyl pyrrole scaffold of **Series II** derivatives with N-benzyl pyrrole one (**Series III**) negatively influences the activity of compounds. Comparing the IC_{50} values reported in **Table 4.2**, the N-benzyl pyrrole derivatives are less active than N-phenyl pyrroles [IC_{50} = 4.83 μM (**EML446**) *vs* IC_{50} = 11.5 μM (**EML462**) and IC_{50} = 2.42 μM (**EML438**) *vs* IC_{50} = 4.83 μM (**EML454**)] (**Figure 4.5 A** and **B**).
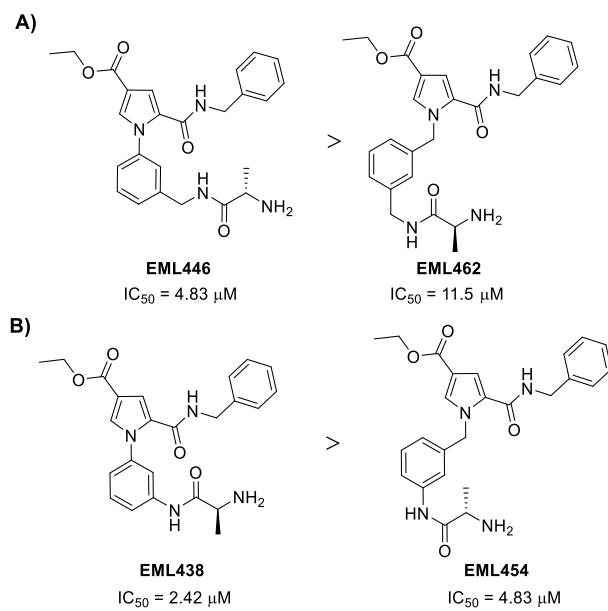
Figure 4.5. N-phenyl derivatives *vs* N-benzyl derivatives

In addition, regarding the different linkages between the aminoacidic moiety and the aromatic ring at the south end of the molecules, it has been assessed

that the elimination of the methylene group between the aforementioned groups provides an improvement of the activity in both the Series II and III of pyrrole derivatives [IC_{50} = 4.83 μM (**EML446**) *vs* IC_{50} = 2.42 μM (**EML438**) e IC_{50} = 11.5μM (**EML462**) *vs* IC_{50} = 4.83μM (**EML454**)] (**Figure 4.6 A** and **B**).
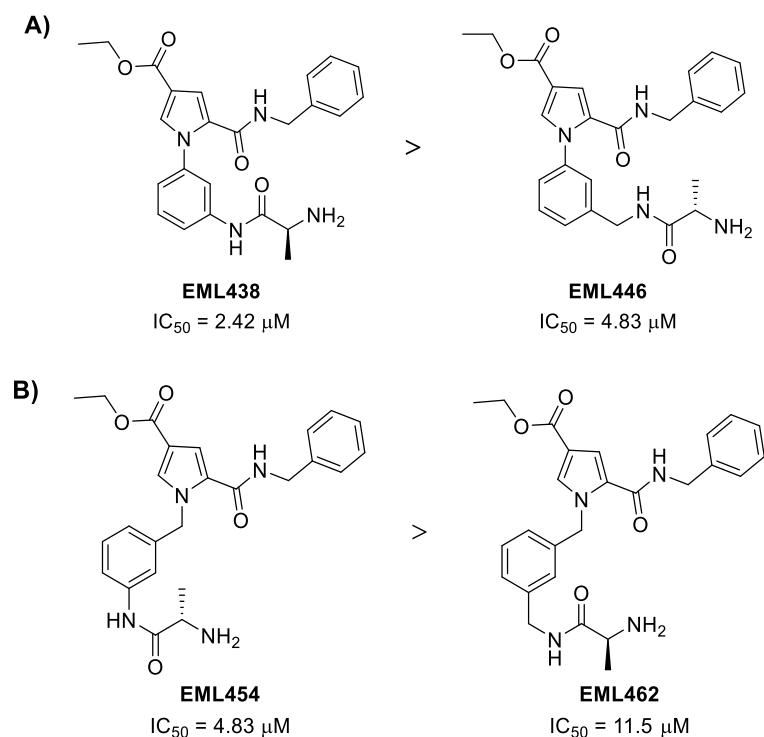
Figura 4.6 Effect of different linkages between the aminoacidic moiety ant the aromatic group at the south end of the molecules

Summarizing all the information reported above the most potent pyrrole inhibitor is **EML438** (IC_{50} = 2.42 μM).

117

4.1.3. Evaluation of cellular activity of pyrrole derivatives

To determine whether the pyrrole inhibitors were able to inhibit CARM1 within a cellular context, in collaboration with Professor Wei Xu from the University of Wisconsin, a homogenous cellular assay (LanthaScreen), that uses time-resolved (TR) FRET technology for monitoring CARM1 activity after treatment with pyrrole inhibitors, has been performed.

4.1.3.1. General principles of FRET

FRET (Fluorescence Resonance Energy Transfer) is based on the transfer of energy between two fluorophores, a donor and an acceptor, when in close proximity. Molecular interactions between biomolecules can be assessed by coupling each partner with a fluorescent label and by detecting the level of energy transfer. When two entities come close enough to each other, excitation of the donor by an energy source (e.g. a flash lamp or a laser) triggers an energy transfer towards the acceptor, which in turn emits specific fluorescence at a given wavelength (**Figure 4.7**). The donor and acceptor can be grafted covalently onto multiple partners that can associate, among others, two dimerizing proteins, two DNA strands, an antigen and an antibody, or a ligand and its receptor. Because of these spectral properties, FRET, a donor-acceptor complex, can be detected without the need for physical separation from the unbound partners. Fully homogeneous assays do not require separation steps such as centrifuging, washing, filtration, or magnetic partitioning. Traditional FRET chemistries are hampered by background fluorescence from sample components such as buffers, proteins, chemical compounds and cell lysate. This type of background fluorescence is extremely transient (with a lifetime in the nanosecond range) and can therefore be eliminated using time-resolved methodologies.
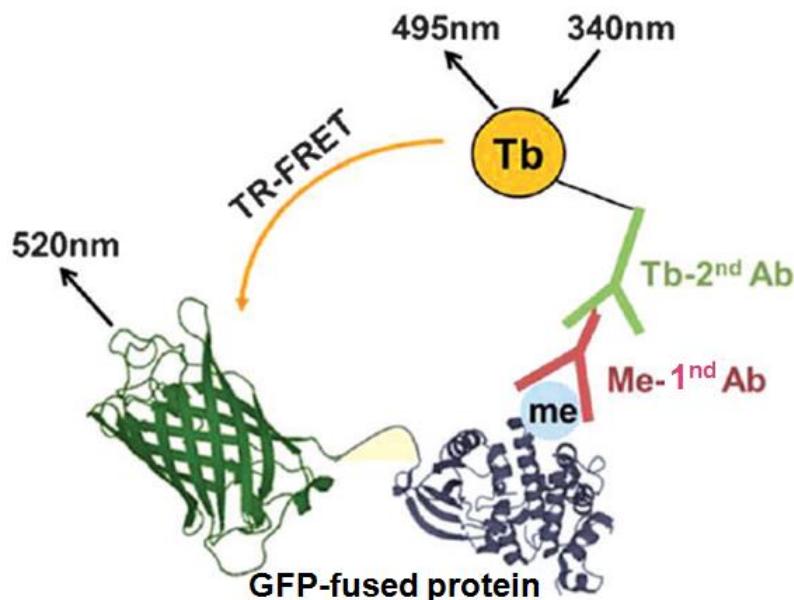
Figure 4.7. Exciting the donor (Tb-2nd Ab) at 340 nm led to FRET-mediated light emission of the acceptors (GFP-fused protein), which in turn emits specific fluorescence at 520 nm

4.1.3.2. TR-FRET cell-based assay and results

Time-resolved Fluorescence resonance energy transfer (TR-FRET) LanthaScreen® cellular assay has been performed for the interrogation of CARM1 cellular inhibition by pyrrole derivatives **EML426 – 466** (**Figure 4.8**), using Tb-labeled antibody as a donor fluorophore and GFP-BAF155 (Chromatin Remodeling Factor) as acceptor fuorophore.

 BAF 155 is one of the non histonic substrates of CARM1 and the methylation occurs at R1064 residue.

HEK 293 cells, stably expressing the fusion protein GFP-BAF155, were treated with 20 μM of each pyrrole compounds for 24h, using DMSO and Adox (20 μM) as negative and positive controls respectively. After cellular lysis, the buffer was supplemented with Me-BAF155 Ab and Tb-2nd Ab and incubated for 1 h. The ternary complex (Me-GFP-BAF155, Me-BAF155 Ab,

and Tb-2nd Ab) was detected by FRET between the Tb donor and the GFP acceptor. The TR-FRET ratio of the GFP-specific signal (measured at 520 nm) to the Tb-specific signal (measured at 495 nm) reflects the level of Me-GFP-BAF155.

The data emerging from this assay were not exciting (**Figure 4.8**). Due to their poor lipophilicity, pyrrolic derivatives (**EML426 – 466**) didn't show any cellular activity in the tr-FRET cell-based assay as well as the pyrazolic lead (**EML327**). **EML458** and **459** showed cytotoxic properties.

The hypothesis of low cytopermeability of pyrrole derivatives was then confirmed by PAMPA assay (*Parallel Artificial Membrane Permeation Assay*). This non-cell based assay is designed to predict passive transcellular permeability through an artificial membrane of drugs in early drug discovery (**Figure 4.9 A**). As it is possible to see in Figure **4.9 B**, the pyrrolic derivatives, although more lipophilic than pyrazolic lead, didn't show a good transcellular permeability [P_a (x 10^{-6} cm/s) = 0.15 (**Lead**) *vs* P_a (x 10^{-6} cm/s) = 0.75 (EML438)] (**Figure 4.9 B**).
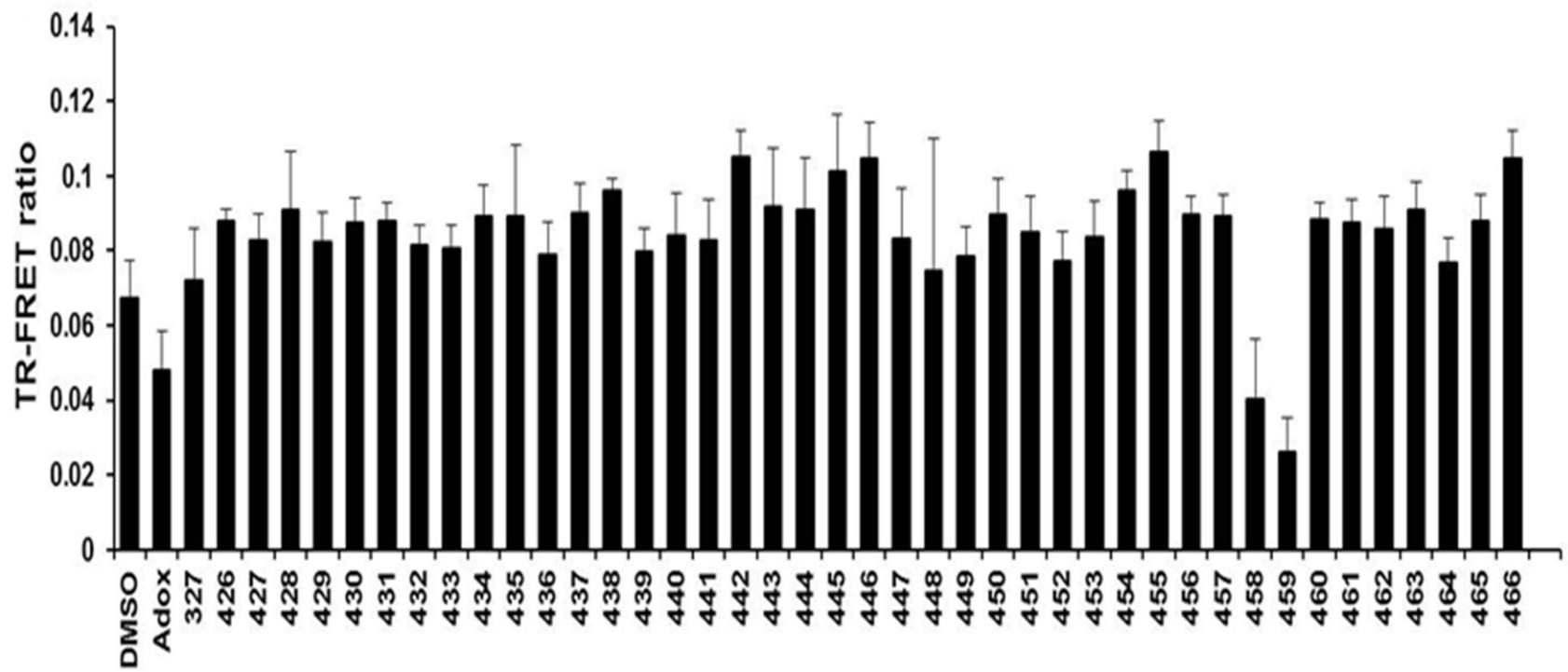
Figure 4.8. TR-FRET measurement of cellular activity of pyrroles inhibitors (**EML426 – 466**). EML327 is the pyrazole lead compound.
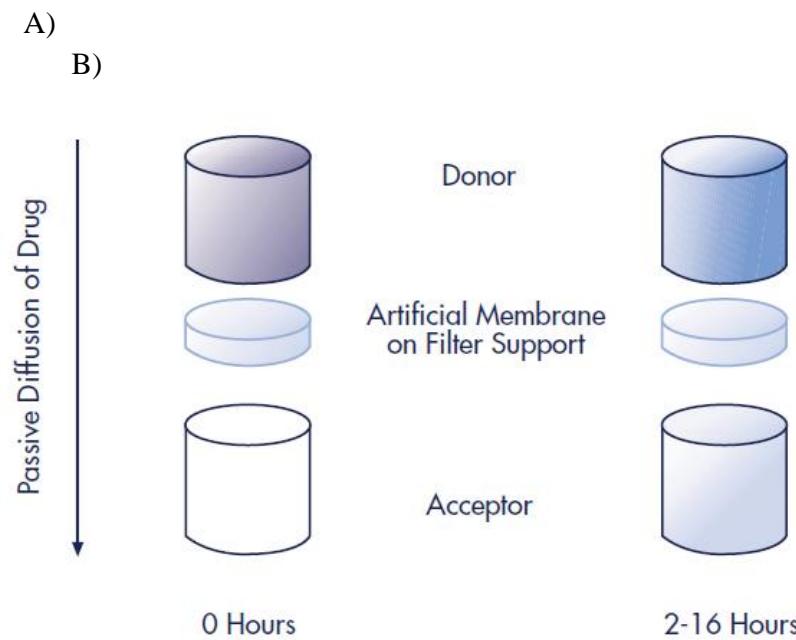
Compound	P_a (x 10^{-6} cm/s) at pH 7.4
Propranolol	7.13 ± 0.20
Furosemide	0.10 ± 0.01
EML568	0.35 ± 0.08
EML428	0.15 ± 0.03
EML432	0.05 ± 0.06
EML438	0.76 ± 0.03
EML446	0.48 ± 0.05
EML454	0.30 ± 0.10
EML462	-
Lead	0.15 ± 0.03

Figure 4.9. **A**) Individual Donor/Acceptor Well Assembly Before and After Incubation. **B**) All compound were tested at 500 μM (5% DMSO/PBS) except for EML 438 which was tested at 250 μM 5% DMSO/PBS. EML 462 which was not soluble. Propanolol and Furosemide were used as positive and negative control respectively

4.1.3.3. Future plans

Taking into account the good activity of pyrrolic inhibitors and the greater lipophilicity and versatility compared to the pyrazolic modulator, supported by computational studies performed by the group of Professor Giuseppe Bifulco of the University of Salerno, we operated structural optimization of the pyrrole derivatives, in order to obtain new derivatives, more permeable and potent than the molecules here described (see chapter 5). The synthesis of novel pyrrole inhibitors is still ongoing, and then their activity toward CARM1 will be assessed through in vitro and cell-based assays.

4.2. Biological evaluation of bisubstrate inhibitors of CARM1

Bisubstrate inhibitors and asymmetric dimetilated peptide **141** and **144** (ADMA) (**Figure 4.11**) were evaluated as inhibitors of PRMT1 and 4. Enzyme activity was measured using an established chemiluminescence-based assay wherein a substrate peptide derived from histone H3 tail is treated with AdoMet (1 µM) and the methyltransferase of interest (PRMT1 80 ng per reaction, PRMT4 200 ng per reaction). After incubation, a primary antibody is then added to specifically bind the methylated arginine residue followed by washing and treatment with a horseradish peroxidase (HRP) conjugated secondary antibody. In the final step, an HRP substrate is added to generate a chemiluminescent signal that is measured using a standard microplate reader. Inclusion of potential inhibitors in the first step of this process in turn allows for the convenient detection and quantization of enzyme inhibition (**Figure 4.10**).
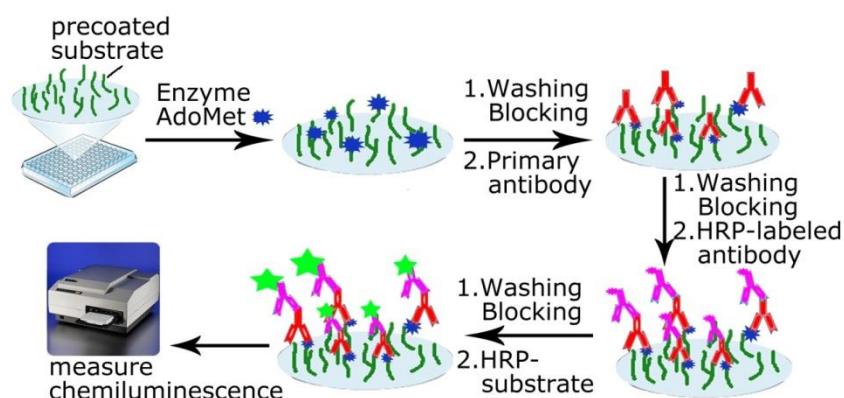
Figure 4.10. Chemiluminescent Assay

Bisubstrate inhibitors and the asymmetrically dimethylated peptide (**Figure 4.11**), were initially screened against PRMT1 and PRMT4 at a threshold concentration of 50 µM, in order to assess the biological activity. The known methyltransferase inhibitor S-adenosyl-L-homocysteine (AdoHcy), a by-product of the AdoMet cofactor resulting from methylation, was also included

>50% inhibition was observed at the threshold inhibitor concentration of 50 μM, complete IC$_{50}$ curves were generated from at least seven unique inhibitor concentrations (from 250 μM to 1 nM) (**Table 4.3**).

A)
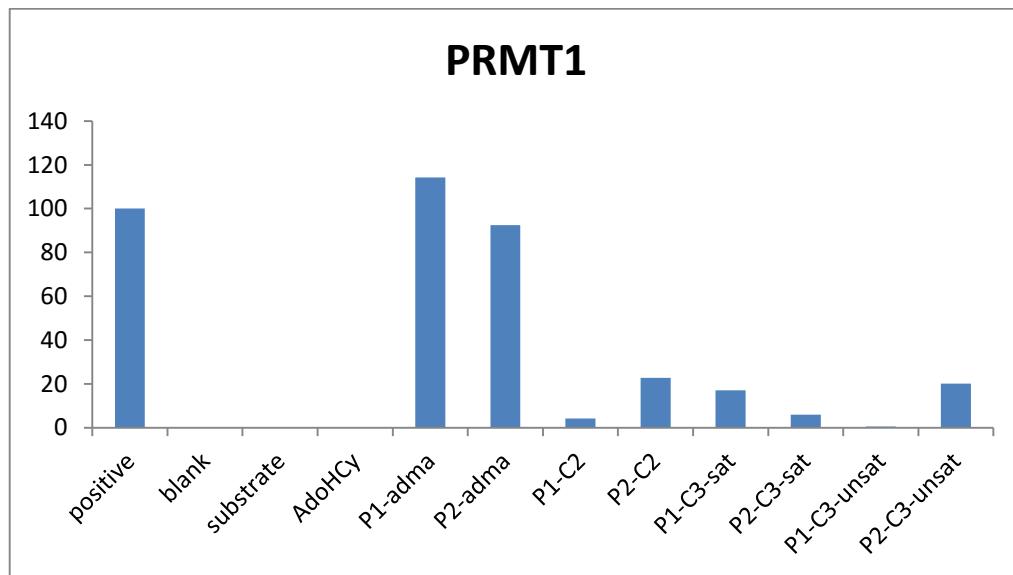
B)
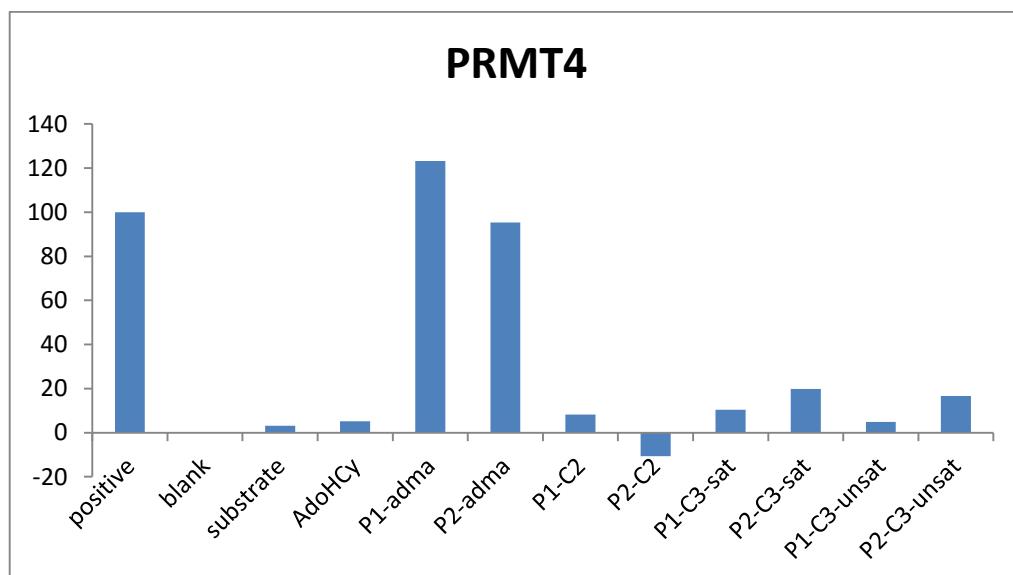
Figure 4.12. A) PRMT1 enzymatic inhibition assays at fixed dose 50μM. B) PRMT4 enzymatic inhibition assays at fixed dose 50μM

#	PRMT1 IC_{50} (µM)[a]	PRMT4 IC_{50} (µM)[a]	PRMT1/PRMT4 IC_{50} ratio
AdoHcy	8.0	0.24	33
P1-C2	25	1.8	13
P2-C2	26	3.2	8
P1-C3-sat	24	0.078	307
P2-C3-sat	26	0.070	370
P1-C3-unsat	3.7	0.037	100
P2-C3-unsat	39	0.043	906

[a]IC_{50} values (µM) based on the best fit of dose-response curves generated from at least seven unique inhibitor concentrations (from 250 µM to 1 nM). Curve fitting performed with GraphPad Prism 5.0

Table 4.3. IC_{50} values of tested compounds

Our preliminary screening confirms our hypothesis that selective and potent inhibition of PRMT4 is achieved by compound characterized by structural features able to bind both the substrate binding sites (**Table 4.3**). The bisubstrate inhibitors generally display a potent and selective inhibition of PRMT4 with IC_{50} values between 3.2 µM and 43 nM. Based on these preliminary results it is possible to speculate that the linkers between the adenosine ring and the guanidine function of the peptide fragments [2 carbons spacer (**P1-C2** and **P2-C2** compouds), saturated or unsaturated 3 carbons spacer (**P1-C3-sat**, **P2-C3-sat**, **P1-C3-unsat** and **P2-C3-unsat** compounds)] play a pivotal role for the correct positioning of the inhibitors warheads, influencing the potency of compounds. On the other hand the selectivity of compounds seems to be regulated by the peptide sequences,

(P1 =Ac-Phe-Gln-Asn-Met-Pro-Ala-Ile-**Arg**-Pro-Ala-Ala-Pro-NH₂; P2 = Pro-Ala-Ala-Pro-**Arg**-Pro-Pro-Phe-Ser-Thr-Met). As it is reported in Chapter 2 (paragraph 2.3, pag. 66) these peptide fragments have been taken from PABP1

and contain respectively the residues of R455 and R460, which are recognized and methylated by PRMT4.

Exception to this trend is represented by compounds **P1-C2** and **P2-C2**, characterized by a two carbon spacer between the guanidine moiety and the adenosine ring. Indeed, they display the lowest potency and selectivity against PRMT4 (**P1-C2** $IC_{50 =}$ 1.8 μM, **P2-C2** $IC_{50 =}$ 3.2 μM), suggesting that their shorter linker between the two fragments is not optimal for mimicking the transitional state of the methyl transfer reaction. The two carbon spacer of compounds **P1-C2** and **P2-C2** probably does not allow to a correct positioning of the adenosine and peptidic fragments within the substrate binding pockets. This could also explain their lower selectivity toward PRMT4 vs PRMT1 and their lower potency compared to the shorter analogue **17a**, a partial bisubstrate inhibitor developed by Martin et al.[142] (**Figure 4.13**).
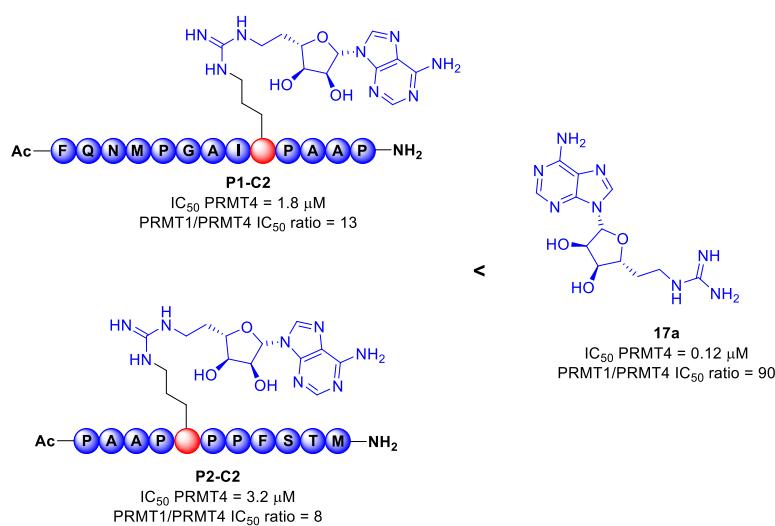
Figure 4.13. Comparison of bisubstrate inhibitors P1-C2 and P2-C2 with their shorter analogue compound **17a**.

Conversely bisubstrate inhibitors bearing the saturated and unsaturated 3-carbons linkers seem to be able to most closely approximate the PRMT4 transition state. Indeed **P1-C3-sat**, **P2-C3-sat**, **P1-C3-unsat** and **P2-C3-unsat** show a nanomolar inhibition of CARM1 with IC_{50} values of 78, 70, 37, 43 nM

respectively (**Table 4.3**). These compounds are largely more potent and selective than their shorter analogues **17b** and **17c**[142] (**Figure 4.14**), which are characterized by an adenosine ring linked to a guanidine moiety through a saturated or unsaturated three carbons linkers (**Figure 4.14** and **4.15**). In these cases the presence of the peptide fragment leads to a net improvement of activity [IC_{50} = 0.078 µM and 0.070 µM (**P1-C3-sat** and **P2-C3-sat**) *vs* IC_{50} = 0.15 µM (**17c**); IC_{50} = 0.037 µM and 0.043 µM (**P1-C3-unsat** and **P2-C3-unsat**) *vs* IC_{50} = 0.56 µM (**17b**)].
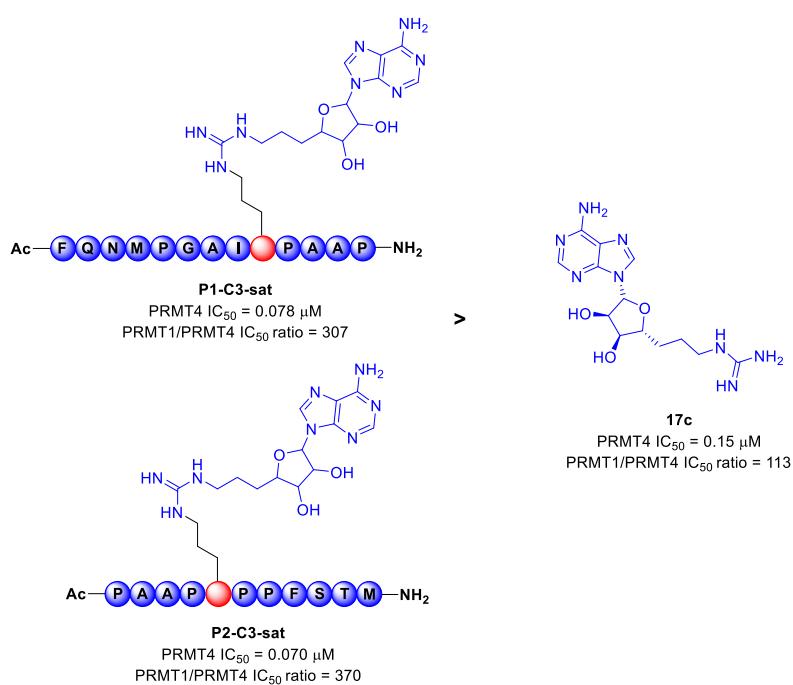
Figure 4.14. Comparison between bisubstrate inhibitors **P1-C3-sat** and **P2-C3-sat** with the partial bisubstrate inhibitor **17c**
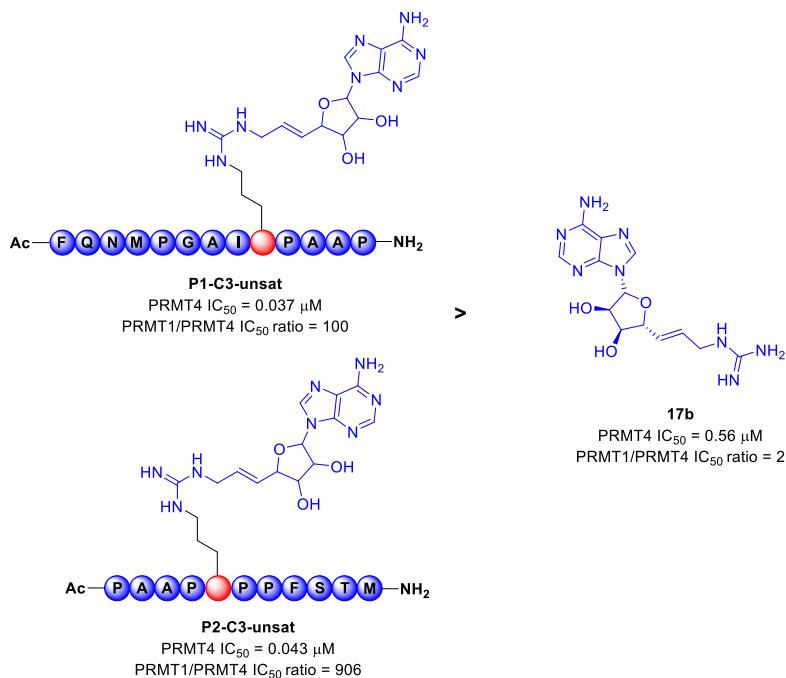
Figure 4.15. Comparison between bisubstrate inhibitors **P1-C3-unsat** and **P2-C3-unsat** with the partial bisubstrate inhibitor **17b**

Comparing the bisubstrate inhibitors bearing the three carbons linkers, the results of the inhibition studies revealed that compounds **P1-C3-unsat** and **P2-C3-unsat** which show a double bond in their alkyl spacer are more potent PRMT4 inhibitors than compounds **P1-C3-sat** and **P2-C3-sat** which contain a fully saturated spacer [($IC_{50} = 37$ nM (**P1-C3-unsat**) *vs* $IC_{50} = 78$ nM (**P1-C3-sat**); $IC_{50} = 43$ nM (**P2-C3-unsat**) *vs* $IC_{50} = 70$ nM (**P2-C3-sat**)] (**Figure 4.16**).
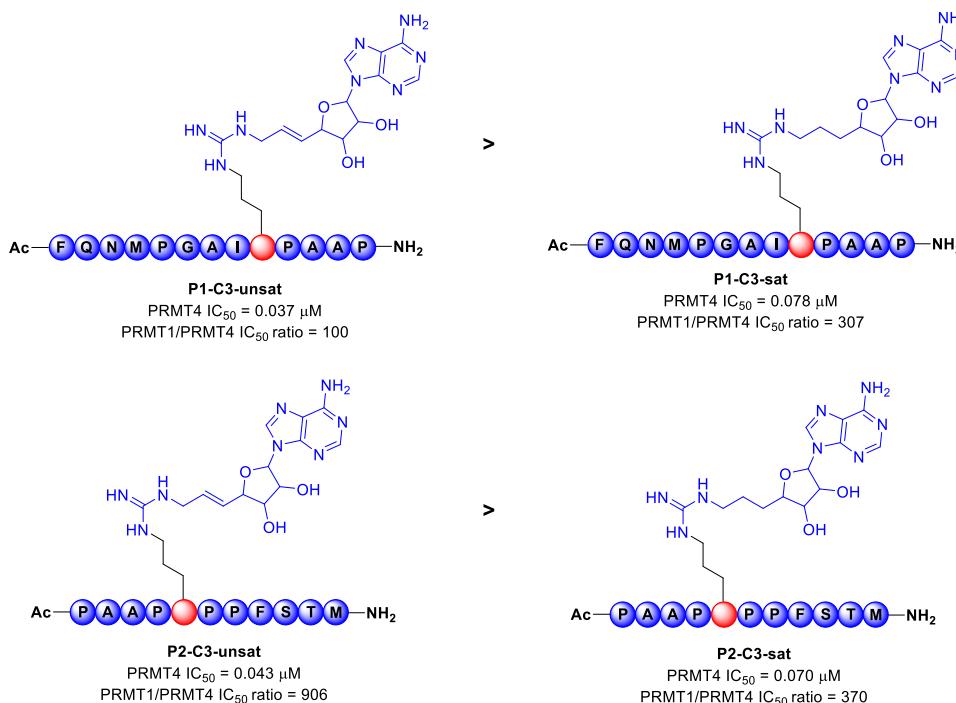
Figure 4.16. Derivatives **P1-C3-unsat** and **P2-C3-unsa**t are more PRMT4 inhibitors than their saturated analogues P1-C3-sat and P2-C3-sat

Interestingly, the peptide sequences seem to influence the selectivity of compounds toward PRMT4 vs PRMT1 without affecting dramatically the potency. In particular compounds bearing the shorter peptide sequence (P2-C3-sat and P2-C3-unsat) showed the highest selectivity (**Figure 4.17**). Compound **P2-C3-unsat** shows a near 900 fold selective inhibition towards CARM1 over PRMT1, whereas its analogue with a longer peptide sequence **P1-C3-unsat** displays a similar potency, but less selectivity toward CARM1 vs PRMT1(PRMT1/PRMT4 IC$_{50}$ ratio = 100) (**Figure 4.17**). Although to a lesser extent, also for compounds bearing a fully saturated spacer (P1-C3-sat and P2-C3-sat) is possible to verify a difference of selectivity [PRMT1/PRMT4 IC$_{50}$ ratio = 370 (**P2-C3-sat**) *vs* PRMT1/PRMT4 IC$_{50}$ ratio = 307 (**P1-C3-sat**)] (**Figure 4.17**).
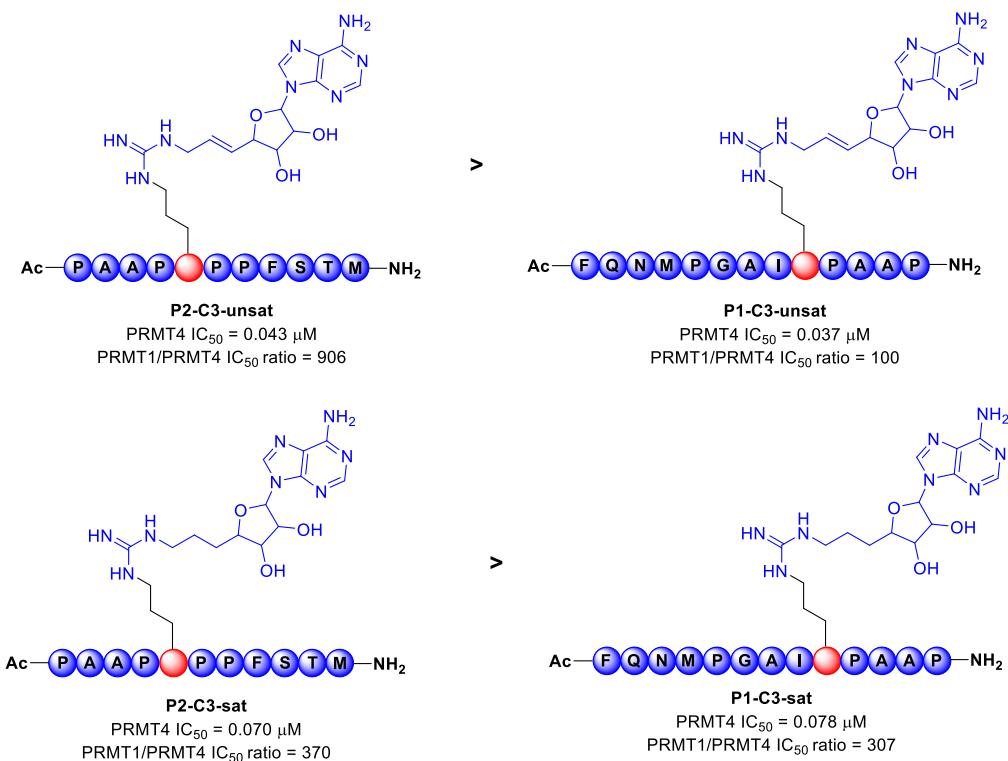
Figure 4.17. Effect of the peptide sequences on the selectivity against PRMT4 over PRMT1

The potent and selective CARM1 inhibition activities displayed by compounds **P1-C3-sat**, **P2-C3-sat**, **P1-C3-unsat** and **P2-C3-unsat** in vitro assay, place them amongst the most active PRMT4 inhibitors reported to date. Further biochemical and biophysical assays are necessary in order to better characterize their activity and their inhibition mechanism.

Crystal structures of CARM1 and bisubstrate inhibitors will be also solved in order to elucidate the accommodation of these molecules within the CARM1 binding sites. The information obtained from these studies could provide meaningful guidance for designing new small molecule modulators.

4.3. Biological evaluation of PRMT3 indole inhibitors

The indole derivatives of **Series I** and **II** (reported in **Figure 4.13**) were preliminarily tested at fixed dose (100 μM) in duplicate against PRMT1 and PRMT3 using biotinylated histone H4 as substrate and SAM as cofactor, employing Alphascreen technology.

4.3.1 General principles of Alphascreen technology

AlphaScreen is a bead-based, non-radioactive Amplified Luminescent Proximity Homogeneous Assay. When a biological interaction brings the beads together, a cascade of chemical reactions acts to produce a greatly amplified signal. On laser excitation, a photosensitizer in the "Donor" bead converts ambient oxygen to a more excited singlet state. The singlet state oxygen molecules diffuse across to react with a thioxene derivative in the Acceptor bead generating chemiluminescence at 370 nm that further activates fluorophores contained in the same bead. The fluorophores subsequently emit light at 520-620 nm (**Figure 4.18**). In the absence of a specific biological interaction, the singlet state oxygen molecules produced by the Donor bead go undetected without the close proximity of the Acceptor bead. As a result, only a very low background signal is produced.
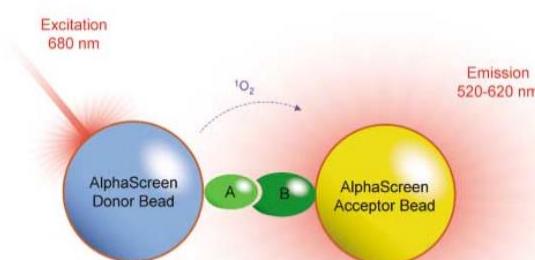
Figure 1.18. When biological interactions bring the Donor and Acceptor beads into close proximity, a highly amplified signal with output in the 520-620 nm range is generated.

4.3.2. Results

Indole compounds of Series I and II (**Figure 1.19**) were tested as inhibitors at 100 μM against PRMT1 and PRMT3 using a biotinylated histone H4-derived peptide (1-21) as substrate, SAM as cofactor and AMI-1 as reference inhibitor. Detection of the histone H4 arginine 3 methylated product is achieved through the addition of Streptavidin (SA) Alpha Donor beads and AlphaLISA Acceptor beads conjugated to an antibody (Ab) directed against the mark of interest. Upon laser irradiation of the beads-target complexes at 680 nm, short-lived singlet oxygen molecules produced by the Donor beads can reach the Acceptor beads in proximity to generate an amplified chemiluminescent signal at 615 nm. The intensity of the light emission is proportional to the methylation activity of the PRMT enzymes.
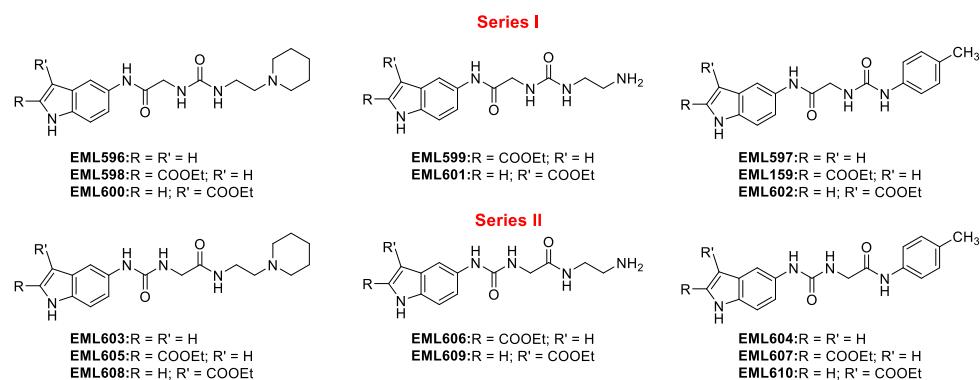
Figure 4.19. Structures of tested compounds

Compounds of **Series I** and **II** bearing the aromatic moiety at the right end of the molecule (**Figure 4.20**) were not tested for problems of solubility. Also the amino derivatives **164a** and **171a** were not tested since they showed high instability (**Figure 4.20**).
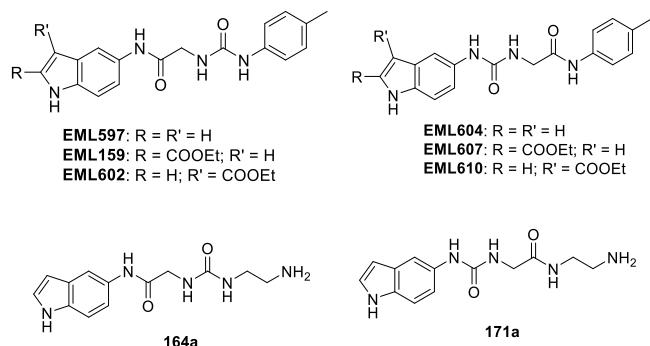
Figure 4.20. Indole compounds bearing the aromatic moiety at the right and of the molecules as well as amino derivatives **164a** and **171a** were not tested against PRMT1and PRMT3 for problems of solubility or instability.

In **Figure 4.21** are reported the results of PRMT1 and PRMT3 enzymatic assays. **EML598** and **599** show an appreciable inhibition of PRMT3 at 100 µM, without affecting the activity of PRMT1. To date is ongoing the screening of the biological activity of Series III compounds against PRMT1 and PRMT3.
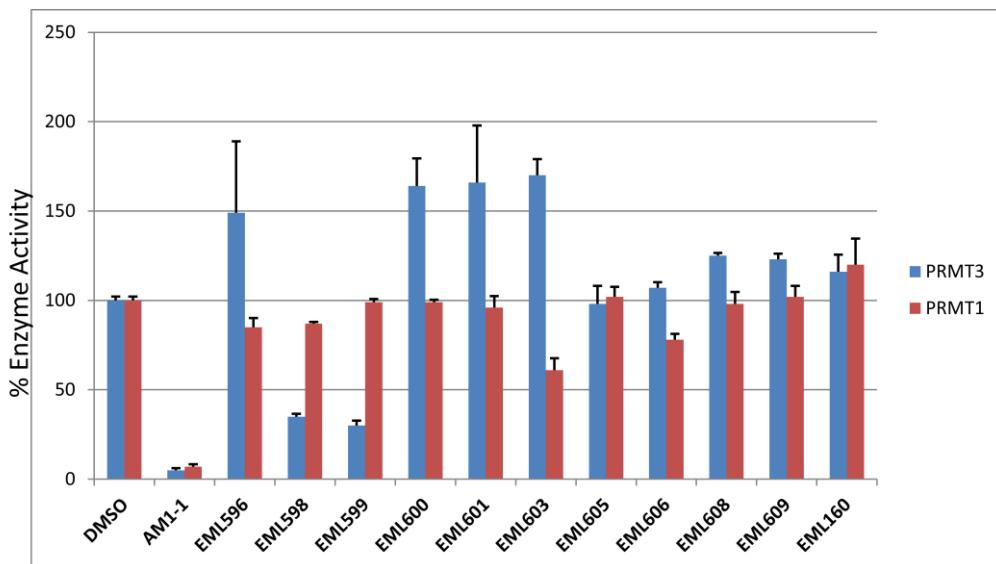
Figure 4.15. Enzymatic in vitro inhibion assay against PRMT1 and PRMT3, using a biotinylated histone H4-derived peptide (1-21) as substrate, SAM as cofactor and AMI-1 as reference inhibitor

In addition, considering that these derivative have been developed starting from activators of PRMT4 (Uracandolates), we decided to assess the activity of some selected indole compounds against CARM1 through enzymatic inhibition assay at fixed dose (50 µM), in order to verify the presence of possible interferences. As it is shown in **Figure 4.16**, indole derivatives didn't show any activity against CARM1.
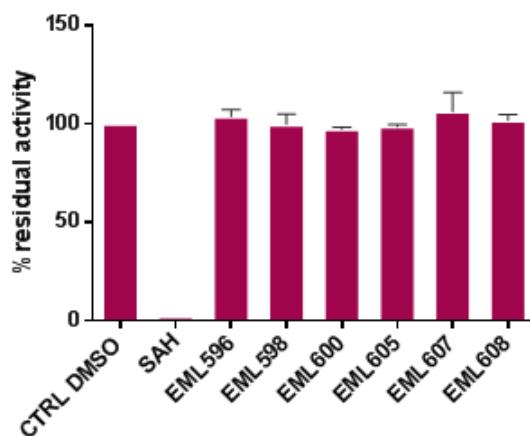
Figure 4.16. CARM1 in vitro inhibition assay at fixed dose (50 µM)

Once assessed the activity of all the synthesized indole derivatives through enzymatic assay at fixed dose (100 µM), the IC_{50} values of the most potent derivatives will be calculated.

CHAPTER 5

COMPUTATIONAL STUDIES AND

STRUCTURAL OPTIMIZATION OF

PYRROLE INHIBITORS

137

5.1. Introduction

As it is discussed in Chapter 4, pyrrole derivatives, designed as inhibitors of CARM1, didn't prove any effects against the enzyme target in cell-based assay, due to their poor cell permeability. On these basis, being interested on developing more potent and cell permeable derivatives, supported by computational studied performed by Professor Giuseppe Bifulco from the University of Salerno, we decided to start a structural optimization process of the synthesized pyrrole inhibitors. In fact, the high synthetic versatility of the pyrrolic ring could allow us to do different structural modifications of the compounds here described (Chapters 2, 3 and 4), in order to enhance enzymatic inhibition and cell permeability.

First of all the binding modes of pyrrolic derivatives were investigated then novel derivatives with different substitution patterns were designed.

5.2. Binding mode of EML438 and design of new derivatives

The computational studies were performed by molecular docking (Glide software version 6.1, standard and extra precision level) using as a model the X–ray structures of human CARM1 co-crystallized with indole CMPD1 or pyrazole CMPD2 inhibitors in complex with cofactors [SAH (*S*-adenosyl- L-homocysteine) or SNF (sinefungin)] (**Figure 5.1**). [152] Comparison of the binding mode of CMPD-1 and CMPD-2 reveals that the two inhibitors, since use different regions of the binding pocket that exists at the interface between the two domains of CARM1 (**Figures 5.1**), both exploit the arginine-binding cavity; site where the transfer of methyl group from the cofactor to the peptidic substrate takes place. Here they make polar interactions (H-bond) with Met260, Glu266 and Glu257, residues known to be critical for substrate recognition; π- π interaction with His414 and van der Waals interaction with Phe153, indicative of a good accommodation within the CARM1 binding site.

These interactions play a major role in influencing CARM1 functionality, guaranteeing the ligand inhibitory character.
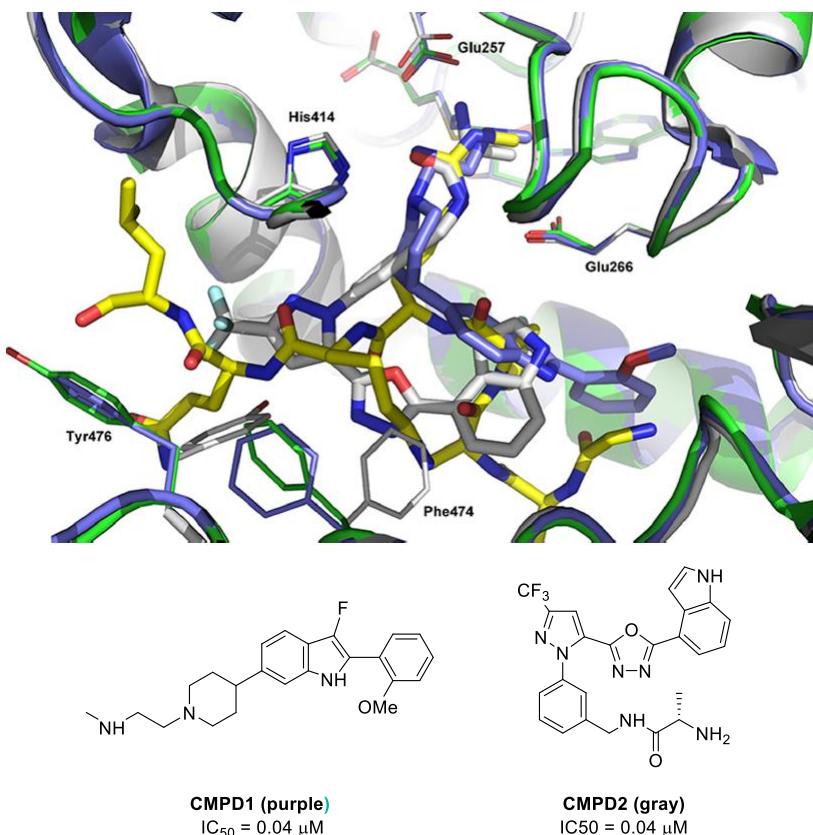
Figure 5.1. Superpositions of the CARM1−H3 R17(Me) complex (green/yellow) with CMPD1 (purple) and CMP2 (gray) inhibitors..

In light of the elucidated structural insights, we evaluated the binding mode of our best pyrrole inhibitors, **EML438**, by molecular docking experiments. These studies revealed that the pyrrole compound, although maintains a similar contact patterns of CMPD1 and CMPD2 within the arginine-binding cavity, is not able to completely exploit the binding pocket that exists at the interface between the two domains of CARM1. This is clearly demonstrated

by the structures superimposition of CMPD1, CMPD2, pyrazole inhibitor **54** and **EML438** docked into PRMT4 binding site (**Figure 5.2**).
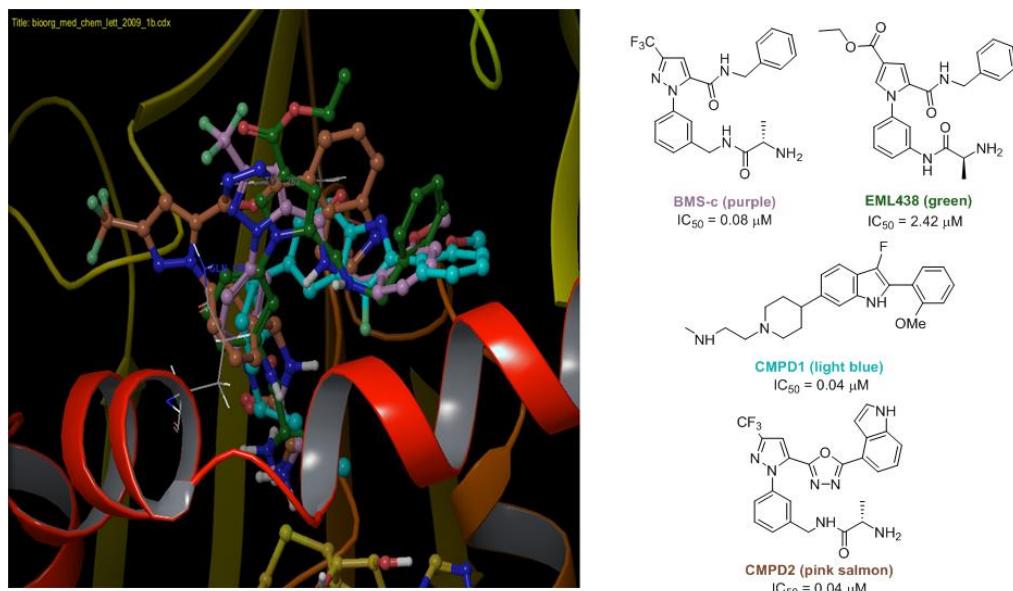
Figure 5.2. Overlay of CMPD1 (blue), CMPD2 (pink salmon), pyrazole inhibitor **54** (purple) and **EML438** (green) structures within the CARM1 binding pocket

This data suggests the possibility of decorating the pyrrole scaffold with additional substituents, in order to improve the binding properties of pyrrole derivatives. To this end, the selection of bulky moieties will also ensure an improvement of the lipophilicity of the molecules, guaranteeing the development of new derivatives suitable for cell-based assays.

Therefore, we focused on designing three class of related analogues of **EML438** and **EML446**, preserving the N-phenyl-pyrrole scaffold, because of its good shape complementarity with the enzyme, and the S-alanine moiety, essential for the CARM1 inhibition, since it is interested in key interactions within the arginine binding cavity.

More than 100 compounds were designed making precise and accurate modifications. All compounds were divided in three classes (A, B and C) according to the different substitution patterns (**Figure 5.3**).
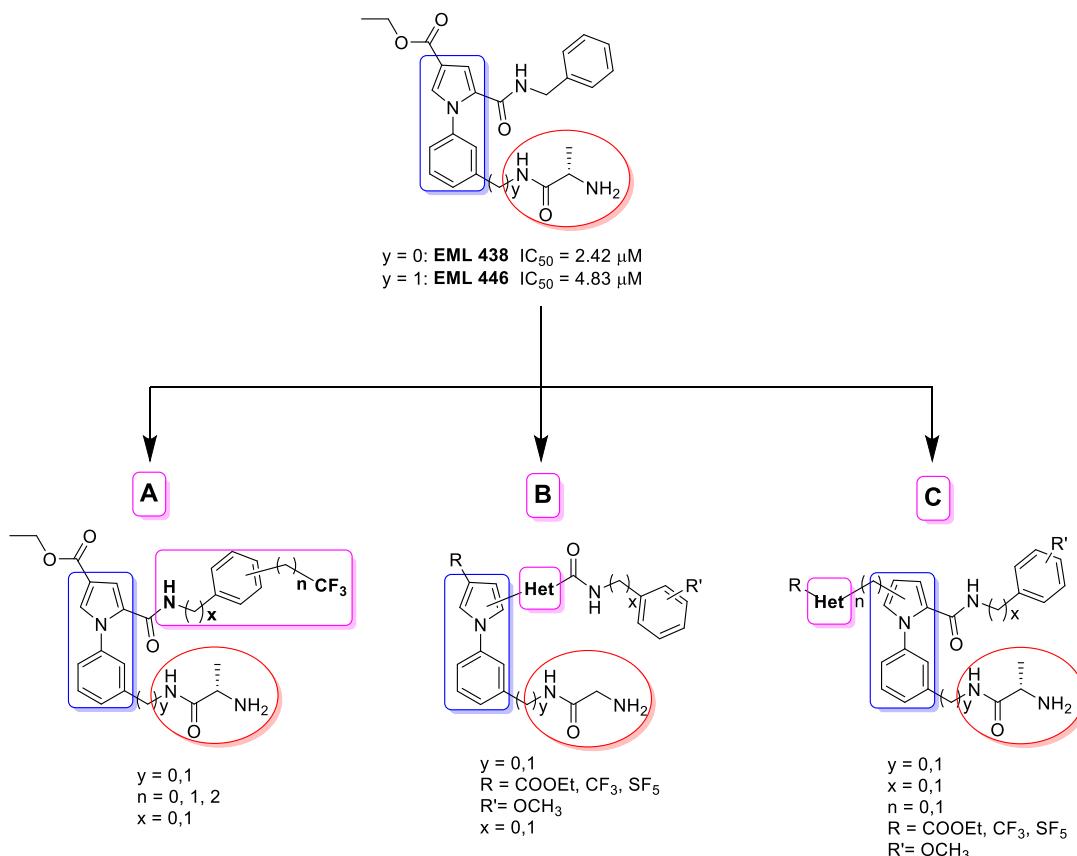
Figure 5.3. Structural optimization of EML438 and 446.

As it is highlighted in **Figure 5.3** compound of series A are characterized by the functionalization of the benzylamidic aromatic ring (highlight in violet), whereas the addition of a second heterocycle at the right or left side of the N-phenyl-pyrrolic scaffold led to the design of Series B and C derivatives.

Then, the ability of these novel derivatives to bind the targets, expressed by a Glide score (kcal/mol) and by the number of key interactions established with the arginine binding cavity of CARM1, were considered. For the calculated binding mode the interactions with the following amino acids were considered discriminant: (a) H-bond with Met 260, Glu266 and Glu257; (b) π-π interaction with His414; (c) van der Waals interaction with Phe153.

Furthermore, employing the QuikProp program, the prediction of apparent Caco-2 cell permeability (QikProp program) was calculated for all the designed derivatives.

According to the score of the aforementioned properties, were selected derivatives with: (a) Glide score $\geq$ 8 kcal/mol, (b) number of interaction with key residues of the arginine binding cavity $\geq$ 3 and (c) apparent Caco-2 permeability $\geq$ 2 nm/sec.

Notably, a first set of molecules (**Table 5.1**), structural related to **EML438** and **EML446**, has been selected in order to validate the computational model. These compounds, although not suitable for cellular assays due to their low predicted apparent permeability, display the highest binding properties and, compared to compounds of Series B and C, a greater ease of synthesis.

In accordance with the computational outcomes, the selected molecules of Series A should have the best in vitro potency. If biological evaluation of these molecules confirm our hypothesis, furher efforts will be focused on the synthesis of series B and C compounds, which, although less simple from a synthetic point of view, display interesting properties of binding mode as well as good values of predicted apparent permeability (**Table 5.2**).

The synthesis of Series A compounds is almost finished.

#	Structure	Energy (kcal/mol)		Number of key interactions		Permeability P_a (x 10^{-6} cm/s)	
A-I-7		-7.807	✅	3.4	✅	0.42	❌
A-I-8		-10.95	✅	3.4	✅	0.98	❌
A-II-7		-10.691	✅	4.2	✅	0.38	❌
A-II-8		-9.530	✅	3.4	✅	0.34	❌

Table 5.1. Compound selected for validating the computational model

#	Structure	Energy (kcal/mol)		Number of key interactions		Permeability P_a (x 10^{-6} cm/s)	
B-I-27		-9.573	✓	3.4	✓	3.64	✓
B-I-28		-9.727	✓	2.6	⚠	4.79	✓
C-I-18		-10.065	✓	4.2	✓	4.19	✓
C-II-18		-8	✓	3.4	✓	4.30	✓
C-I-23		-10.184	✓	4.2	✓	2.39	✓
C-II-22		-8.968	✓	4,2	✓	2.00	✓

Table 5.1. Most promising pyrrole derivatives.

CHAPTER 6

CONCLUSIONS

The present PhD project was focused on design, synthesis and biological evaluation of new small-molecule modulators of Protein Methyltransferases, with particular interest for PRMT3 and CARM1 (Coactivator associated Arginine Methyltransferase 1), also known as PRMT4.

The first objective of my PhD project was the design, synthesis and biological evaluation of a small set of novel pyrrole-based CARM1 inhibitors. These compounds were designed taking into account the information obtained from the structure-activity relationships (SAR) of pyrazole and indole compounds, the most potent PRMT4 inhibitor, which, however, lack activity in cell-based assays. We hypothesized that the pyrrole ring, considered as a result of isosteric substitution of pyrazole together with opening and simplifation of the indole ring, would be a good scaffold for targeting CARM1. Indeed, a potent inhibition was observed when testing pyrrole derivatives against CARM1 (i.e. **EML 438**, IC_{50} = 2.42 μM), nevertheless they didn't prove a significant cellular activity, due to their poor transcellular permeability. Docking studies were also performed with pyrrole compounds and PRMT4, in collaboration with the group of Professor Giuseppe Bifulco from the University of Salerno. These studies suggest many possibilities for further improving the inhibitor affinity and the cell permeability of the pyrrole modulators by structural optimization. To this end, facilitated by the high synthetic versatility of the pyrrole ring and considering the structure-activity relationships of the inhibitors, further pyrrole derivatives have been designed and identified. To date the synthesis and biological evaluation of the new selected compounds are still ongoing.

Furthermore, in collaboration with Professor Nathaniel Martin from the University of Utrecht, we have successfully synthesized a set of novel potential PRMT4 bisubstrate inhibitors, providing a convenient synthetic route to link SAM analogues with peptide substrate portions, through variable guanidine linkers. Preliminary screening of their biological activity revealed a

nanomolar inhibition of PRMT4 (**P2-C3-unsat**, IC_{50} = 43 nM) with about 900-fold selectivity for CARM1 over PRMT1. These results confirm our hypothesis that selective and potent inhibition is achieved by compounds characterized by structural features able to bind both the substrate binding sites and mimic the transition state. The potency and the selectivity displayed by bisubstrate inhibitors against PRMT4 place them as one of the most interesting set of chemical tools for studying the enzyme target. Further biological and biophysical assays are ongoing in order to better asses the selectivity, the binding mechanism and the pharmacokinetic properties of these compounds. Moreover crystal structures of bisubstrate inhibitors with PRMT4 will be solved. These christallografic structures may provide a meaningful guidance in designing novel CARM1 inhibitors.

We also started a program aiming at developing inhibitors of PRMT3. In 2012 we discovered a class of indole derivatives, with aryl ureido acetamido indole carboxylate structures (dubbed Uracandolates), characterized by a peculiar activity against the PRMTs. They were able to enhance the activity of CARM1 both in vitro and cellular setting, providing, at the same time, an inhibitory activity, although low, against PRMT3. Therefore with the aim to develop novel inhibitors of PRMT3, we decided to properly modify the structure of Uracandolate compounds. Ispired by the SAR information of benzothiadiazole and isoquinoline allosteric inhibitors developed by Jin and coll., we designed and synthesized three series of indole-based compounds. Their inhibitory activities were assessed against a panel of PRMTs (PRMT1, PRMT3 and PRMT4), testing compounds at fixed dose (100 µM) in duplicate. **EML598** and **EML599** display a selective inhibition of PRMT3 (70 % of PRMT3 inhibition at 100 µM). To date, further biochemical assays are ongoing in oreder to better characterize the activity of indole derivatives.

CHAPTER 7

EXPERIMENTAL SECTION

7.1. General information

All chemicals were purchased from Aldrich Chimica (Milan, Italy) and were of the highest purity. All solvents were reagent grade. Reactions were routinely monitored by TLC performed on aluminum-backed silica gel plates (Merck DC, Alufolien Kieselgel 60 F254) with spots visualized by UV light (l ¼ 254, 365 nm). Solvents were removed using a rotary evaporator operating at a reduced pressure of _10 Torr. Organic solutions were dried over anhydrous Na_2SO_4. Chromatographic purifications were done on an automated flash-chromatography system (Isolera ™One, Biotage) using cartridges packed with KP-SIL, 60 °A (40–63 mm particle size). High performance liquid chromatography (HPLC) was performed on a Shimadzu SPD 20A UV/VIS detector (l ¼ 215 nm) using C-18 column Phenomenex Synergi Fusion – RP 80A (75 _ 4.60 mm; 4 mm) at 25 _C using a mobile phase A (water + 0.1% trifluoracetic acid (TFA)) and B (MeCN + 0.1% TFA) at a flow rate of 1 mL min_1. The following gradient was applied: isocratic elution for 1 min at 10% of solvent B, linear increase from 10% to 95% of solvent B over 10 min, hold at 95% solvent B for 3 min.

Melting points were determined on a Stuart SMP30 melting point apparatus in open capillary tubes and are uncorrected. 1H spectra were recorded at 300 MHz with a Bruker Avance 300 spectrometer. Chemical shifts are reported in δ (ppm) relative to the internal reference tetramethylsilane (TMS). When the NMR analysis is not included, compounds were used in the next step without further purification. Mass spectra were recorded on a Finnigan LCQ DECA TermoQuest (San Jose, CA) mass spectrometer in electrospray positive and negative ionization modes (ESI-MS).

7.2. Pyrrole derivatives

Synthesis of Synthesis of (E)-4-ethoxy-1,1,1-trifluorobut-3-en-2-one (87)[156]
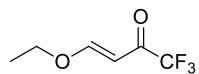
To a stirred solution of 4-dimethylaminopyridine (12.0 mg, 0.10 mmol) and trifuoro acetic anhydride (3.28 gm, 15.62 mmol) in dichloromethane (20.00 ml), ethyl vinyl ether (1.06 g, 14.70 mmol) was added dropwise at -10 °C. The reaction mixture was stirred for 19 h at 0°C, allowed to warm to room temperature and the solvent was evaporated in vacuo. The reaction mixture was poured into sodium bicarbonate solution (the deep violet colour changed to a yellow colour), the two phases were separated and the organic phase was washed with water, dried over anhydrous sodium sulfate and the solvent was evaporated in vacuo to give the crude title compound (**87**) as a yellow oil (2.00 gm, 81%). The crude product was used directly in subsequent step without further purification. ^{1}H NMR (300 MHz, CDCl$_3$): δ 7.89 (d, *J* = 12.3 Hz, 1H), 5.85 (d, *J* = 12.3 Hz, 1H), 4.10 (q, *J* = 7.1 Hz, 2H), 1.41 (t, *J* = 7.1 Hz, 3H). ESI-MS m/z: 169.3 (M + 1)$^{+}$.

Synthesis of (E)-4-ethoxy-2-(trifluoromethyl)-2-((trimethylsilyl)oxy)but-3-enenitrile (88)[156]
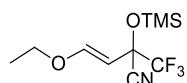
To a mixture (0.5 mmol, 50 mg) Et$_3$N and TMSCN (13 mmol, 1.29 g) was added 10 mmol of ketone **85** at 0-5 °C under stirring. Then the mixture was stirred for 12h at RT, and the product was isolated by vacuum distillation (80%).

^{1}H NMR (300 MHz, CDCl$_3$) 5:0.26 (s, 9 H), 1.32 (t, 3 H, J = 7.1 Hz), 3.85 (q, 2 H, J=7.1 Hz), 4.83 (d, 1 H, J = 12.5 Hz), 6.98 (d, 1 H, J = 12.5 Hz). ESI-MS m/z: 268.0 (M + 1)$^{+}$.

Synthesis of (E)-2-(aminomethyl)-4-ethoxy-1,1,1-trifluorobut-3-en-2-ol (89)[156]
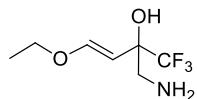
To a suspension of LiAlH$_4$ (2.03 g, 53.5 mmol) in dry ether (50 mL) **88** (12.99 g, 48.6 mmol) was added dropwise with stirring for 30 min at 0–5 °C. The mixture was stirred overnight at room temperature. Excess LiAlH$_4$ was decomposed with 30% aq NaOH (10 mL) with stirring at 0° C, then precipitated alumina was filtered and washed thoroughly with ether (3 x 50 mL). The filtrate was dried (Na$_2$SO$_4$) and concentrated under reduced pressure to give a brown oil (5.8 g, 60%). The productwas used directly in subsequent step without further purification. ^{1}H NMR (300 MHz, CDCl$_3$): δ 6.75 (d, *J* = 12.6 Hz, 1H), 5.32 (m, 2H, exchangeable with D$_2$O) 4.71 (d, *J* = 12.6 Hz, 1H), 3.77 (q, *J* = 7.0 Hz, 2H), 3.13 (s, 1H, exchangeable with D$_2$O), 2.69 (d, *J* = 13.1 Hz, 2H), 1.29 (t, *J* = 7.0 Hz, 3H). ESI-MS m/z: 200.5 (M + 1)$^+$.

Synthesis of (E)-N-(4-ethoxy-2-hydroxy-2-(trifluoromethyl)but-3-en-1-yl)benzamide (93)[156].
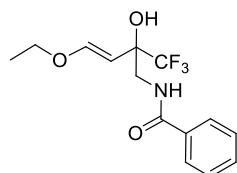
To a mixture of amino alcohols **88** (0.8 g, 4 mmol) and Et$_3$N (0.404 g, 4 mmol) in dry DCM (30 mL), benzoyl chloride (0.562 g, 4 mmol) was added at 0 C. The mixture was stirred for 2 h at room temperature and water (15 mL) was added. The organic phase was washed with saturated aqueous solution of NaHCO$_3$ (3 x 20 mL) and water (3 x 25 mL). The aqueous layer was extracted

with DCM (3 x 25 mL) and the combined organic layers were dried (Na_SO_4) and then evaporated. Compound **93** was obtained as a yellow oil after purification by column chromatography on silica gel (eluent EtOAc/hexane 1:6). ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.76 (d, *J* = 7.2 Hz, 2H), 7.59 – 7.50 (m, 1H), 7.44 (t, *J* = 7.4 Hz, 1H), 6.80 (d, *J* = 12.4 Hz, 1H), 6.64 (bs, 1H), 5.10 (bs, 1H), 4.73 (d, *J* = 12.4 Hz, 1H), 3.99 – 3.86 (m, 1H), 3.76 (q, *J* = 7.0 Hz, 2H), 3.69 – 3.53 (m, 2H), 1.25 (t, *J* = 7.0 Hz, 3H). ESI-MS m/z: 304 (M + 1)$^+$.

Synthesis of phenyl(3-(trifluoromethyl)-1H-pyrrol-1-yl)methanone (94).

A mixture of DCM/TFA 7:3 (2mL) was added to compound **93** (0.191 g, 0.63 mmol). The reaction was stirred at room temperature overnight and then concentrated in vacuo to give the title compound as a yellow oil (0.150 g, 99%). ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.76 (d, *J* = 7.1 Hz, 2H), 7.71 – 7.63 (m, 1H), 7.63 – 7.51 (m, 3H), 7.37 – 7.31 (m, 1H), 6.56 – 6.49 (m, 1H). ESI-MS m/z**:** 478.8 [(M x 2) + 1] $^+$.

Synthesis of 2,2,2-trichloro-1-(4-(trifluoromethyl)-1H-pyrrol-2-yl)ethanone (95).
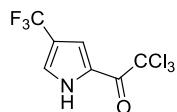
To a solution of **94** (0.500 g, 2.1 mmol) in Dioxane/H$_2$O 1: 1 (3 mL) was added LiOH (0.050 g, 2.1 mmol). The reaction was sealed and left to stir at room temperature for 4h. Water was added and the aqueous layer was extracted with DCM (3 x 3 mL). The combined organic layers were dried (Na$_2$SO$_4$) , filtered and placed into the vessel. Then trichloroacetyl chloride

was added (3.81 g ,21 mmol was stirred at 70 °C for 36 h. The mixture was diluted with DCM (100 mL) and washed with satured aqueous solution of NaHCO$_3$ (5 x 30 mL), Brine (1 x 10 mL), dried (Na$_2$SO$_4$), filtered and evaporated under vacuum. The residue was purified by silica gel chromatography (ethyl acetate : petroleum ether, 5 : 95) to give the title compound as a yellow solid (0.329 g, 56 %); m.p. 107 - 109 °C. ^{1}H NMR (300 MHz, CDCl$_3$) δ 9.61 (bs, 1H, exchangeable with D$_2$O), 7.53 – 7.48 (m, 1H), 7.47 – 7.42 (m, 1H). ESI-MS m/z **:** 277.85 (M – 1)$^-$.

Synthesis of N-benzyl-4-(trifluoromethyl)-1H-pyrrole-2-carboxamide (96).
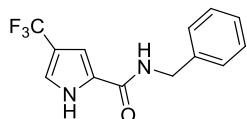
95 (0.100 g, 0.36 mmol), K$_2$CO$_3$ (0.148 g, 1.8 mmol) and benzylamine (0.148 g, 1.08 mmol) were dissolved in DMF (5 mL) at room temperature in a 35 mL CEM pressure vessel equipped with a stirrer bar. The vial was sealed and heated in a CEM Discover microwave synthesizer (max power 300W, max internal pressure 250 psi) to 80 °C (measured by the vertically focused IR temperature sensor) for 30 min. The cold reaction mixture was diluted with H$_2$O (100 mL) and extracted with AcOEt (3 x 30 mL). The combined organic layers were washed with 0.5 N HCl (3 × 30 mL), NaHCO$_3$ saturated aqueous solution and brine. The organic phase was dried (Na$_2$SO$_4$), filtered, and concentrated in vacuo. The residue was purified by silica gel chromatography (ethyl acetate : petroleum ether, 1:9 to 2:8) to give the title compound as a yellow solid (0.066 g, 68%), mp 128 – 130 °C. ^{1}H NMR (300 MHz, CDCl$_3$): δ 10.18 (bs, 1H, exchangeable with deuterium oxide), 7.42 – 7.29 (m, 5H), 7.23 – 7.15 (m, 1H), 6.77 – 6.67 (m, 1H), 6.22 (t, *J* = 5.8 Hz, 1H, exchangeable with deuterium oxide), 4.62 (d, *J* = 5.8 Hz, 2H). ESI-MS m/z **:** 267.8 (M + 1)$^+$.

Synthesis of N-benzyl-1-(3-cyanophenyl)-4-(trifluoromethyl)-1H-pyrrole-2-carboxamide (97).
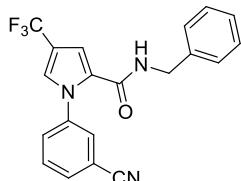
3-bromobenzonitrile (0.156 g, 0.82 mmol) and CuI (0.156 g, 0.82 mmol) were suspended in Dioxane dry (1.2 mL) at room temperature in a 10 mL CEM pressure vessel equipped with a stirrer bar. The vial was degassed and N1,N2-dimethylethane-1,2-diamine (1.65 mmol, 0.145 g) was added to a mixture. The reaction was stirred at room temperature for 5 minutes and then compound **96** (0.200 g, 0.75 mmol) and K_3PO_4 (0.316 g, 1.49 mmol) were added. The reaction was degassed and heated in a CEM Discover microwave synthesizer (max power 300W, max pressure 250 psi) to 130 °C for 30 min, then water (10 ml) was added and the mixture was extracted with ethyl acetate (3 x 10 mL). The combined organic layers were washed with brine, dried (Na_2SO_4), filtered and concentrated in vacuum. The residue was purified by silica gel chromatography (ethyl acetate : petroleum ether, 1:1) to give the title compound as a white solid (0.221 g, 80%); m. p. 162 – 164 °C. ^{1}H NMR (300 MHz, CDCl$_3$): δ 7.72 (d, J = 6.3 Hz, 1H), 7.67 – 7.49 (m, 3H), 7.44 – 7.27 (m, 5H), 7.20 (s, 1H), 6.87 (s, 1H), 6.22 (s, 1H,exchangeable with D$_2$O), 4.49 (d, J = 5.7 Hz, 2H). ESI-MS m/z **:** 370 (M + 1) $^+$.

Synthesis of 1-(3-(aminomethyl)phenyl)-N-benzyl-4-(trifluoromethyl)-1H-pyrrole-2-carboxamide (98).

158
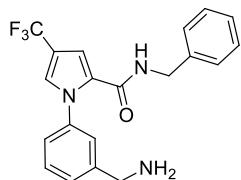
NaBH$_4$ (0.090g, 2.6 mmol) was added over 30 min to a solution of NiSO$_4$•6H$_2$O (0.012 g, 0.049 mmol), and Compound **97** (0.180 g, 0.49 mmol) in MeOH (15 ml) at 0°C. The mixture was stirred, at room temperature, for 3 h and then was added NaHCO$_3$ saturated aqueous solution (3 x 30 mL). The aqueous layer was extracted with EtOAc (3 x 10 m), the combined layers were washed with brine (3 x 10 mL), dried (Na$_2$SO$_4$), filtered and concentrated in vacuo. The residue was purified by silica gel chromatography (ethyl acetate : petroleum ether, 1:1) to give the title compound as a white gel (0.115 g, 63%). ^{1}H NMR (300 MHz, DMSO): δ 8.95 (t, *J* = 5.6 Hz, 1H, exchangeable with deuterium oxide), 7.74 – 7.58 (m, 1H), 7.41 – 7.23 (m, 8H), 7.21 – 7.13 (m, 2H), 4.33 (d, *J* = 5.6 Hz, 2H), 3.89 – 3.68 (m, 2H), 3.53 – 3.31 (m, 2H, exchangeable with deuterium oxide). ESI-MS m/z **:** 374 (M + 1)$^+$.

Synthesis of (S)-tert-butyl (1-((3-(2-((2-methoxybenzyl)carbamoyl)-4-(trifluoromethyl)-1H-pyrrol-1-yl)benzyl)amino)-1-oxopropan-2-yl)carbamate (99).
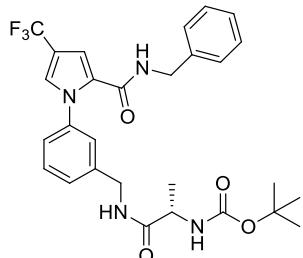
To a solution of **98** (0.115 g, 0.31 mmol), DMAP (0.004 g, 0.031 mmol) and Boc-L-Ala-OH (0.087 g, 0.46 mmol) in DCM dry (6.5 mL) was added DCC (0.095 g, 0.46 mmol). The mixture was left to stir at room remperature overnight, and then was filtered and concentrated under vacuum. The residue was dissolved with DCM (100 mL) and washed with saturated aqueous solution of NaHCO$_3$ (3 x 30 mL) and brine. The organic phase was dried over Na$_2$SO$_4$, filtered and concentrated under vacuum. The title compound **99** was obtained as white solid (0.147 g, yield 87 %) after purification by silica gel

chromatography (ethyl acetate : petroleum ether, 1:1); mp 165 – 166 °C. $[\alpha]_D$ -2.6 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO): δ 8.91 (t, *J* = 5.4 Hz, 1H, exchangeable with deuterium oxide), 8.44 – 8.27 (m, 1H, exchangeable with deuterium oxide), 7.72 – 7.55 (m, 1H), 7.38 – 7.14 (m, 10H), 7.04 – 6.90 (m, 1H, exchangeable with deuterium oxide), 4.40 – 4.20 (m, 4H), 4.04 – 3.85 (m, 1H), 1.34 (s, 9H), 1.19 (d, *J* = 7.0 Hz, 3H). ESI-MS m/z: 544.8 (M + 1) $^+$.

Synthesis of (S)-1-((3-(2-(benzylcarbamoyl)-4-(trifluoromethyl)-1H-pyrrol-1-yl)benzyl)amino)-1-oxopropan-2-aminium (100).
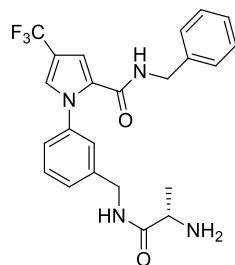
Compound **99** was dissolved with a mixture of DCM/TFA (9:1, 4 mL) and lefto to stir at room temperature for 30 min. The desidered product **100** was obtained as white solid after evaporation of the mixture of reaction under vacuum; mp 110 - 113 °C. $[\alpha]_D$ + 8.9 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO):δ 9.00 – 8.85 (m, 2H, exchangeable with deuterium oxide), 8.17 – 8.03 (m, 3H, exchangeable with deuterium oxide), 7.71 – 7.63 (m, 1H), 7.45 – 7.35 (m, 2H), 7.33 – 7.13 (m, 8H), 4.42 – 4.36 (m, 2H), 4.33 (d, *J* = 5.6 Hz, 2H), 3.93 – 3.84 (m, 1H), 1.38 (d, *J* = 6.8 Hz, 3H). ESI-MS m/z: 445.11 (M + 1)$^+$.

Synthesis of 1-(3-nitrophenyl)1H-pyrrole (103a) [165]**.**
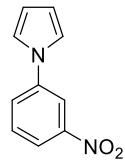
A solution of 3-nitroaniline (1.9 g, 13.7 mmol) and 2,5-dimethoxytetrahydrofuran (15.1 mmol, 2.0 g) in glacial acetic acid (10 mL) was heated at 80 °C for 2 h. After cooling, the solvent was evaporated under vacuo. The residue was made basic with a saturated aqueous solution of NaHCO$_3$ (150 mL) and extracted with ethyl acetate (3 x 50 mL). The organic layer was washed with brine (20 mL), dried over Na$_2$SO$_4$ and filtered. After evaporation of the solvent, the crude mixture was purified by flash silica gel chromatography (ethyl acetate: n-hexane = 1:9 as eluent) to furnish **103a** (2.26 g, 84%) as yellow solid, mp 73 – 74 °C. ^{1}H-NMR (300 MHz, CDCl$_3$): δ= 8.44 (t, 1H, *J* = 2.26 Hz), 8.30 (d, 1H, *J* = 8.30 Hz), 7.93-7.90 (m, 1H), 7.82-7.75 (m, 1H), 7.35 (t, 2H, *J* = 2.26 Hz), 6.60 (t, 2H, *J* = 2.26 Hz). ESI-MS m/z: 189.1 (M + 1)$^+$.

Synthesis of 3-(1H-pyrrol-yl)benzonitrile (103b) [166]**.**
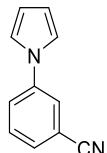
Compound 103**b** was obtained as a yellow solid (2.3 g, 80%) from compound **101b** (1.6 g, 13.7 mmol) and **102a** (2.0 g, 15.1 mmol), according to the procedure used for **103a.** 1H NMR (CDCl3) δ: 6.41 (t, J) 2.2 Hz, 2H), 7.09 (t, J) 2.2 Hz, 2H), 7.51-7.55 (m, 2H), 7.60-7.66 (m, 2H). ESI-MS m/z: 169 (M + H)

Synthesis of 2,2,2-trichloro-1-(1-(3-nitrophenyl)-1H-pyrrol-2-yl)ethanone (104a).
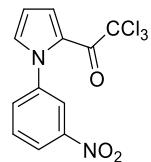
A mixture of trichloroacetyl chloride (5.4 g, 29.7 mmol) and **103a** (1.0 g, 2.97 mmol)) were placed into CEM pressure vessel equipped with a stirrer bar.

The vial was sealed and heated in a CEM Discover microwave synthesizer (max power 300W, max internal pressure 250 psi) to 140 °C (measured by the vertically focused IR temperature sensor) for 1.5 h. The cold reaction mixture was diluted with $CHCl_3$ (100 mL) and washed with a saturated aqueous solution of $NaHCO_3$ (3 x 20 mL) and Brine. The organic layer was dried, and filtered. Removal of a solvent gave a residue that was purified by column chromatography (silica gel, ethyl acetate: petroleum ether 1:9 as eluent) to give **104a** (1.40 g, 60%) as a yellow solid; mp 100 – 101°C. ^{1}H NMR (300 MHz, CDCl$_3$)**:** δ 8.35 – 8.29 (m, 1H), 8.21 – 8.14 (m, 1H), 7.73 – 7.68 (m, 1H), 7.68 – 7.63 (m, 2H), 7.15 – 7.10 (m, 1H), 6.52 – 6.45 (m, 1H). ESI-MS m/z: 355.04 $(M + Na)^+$

Synthesis of 3-(2-(2,2,2-trichloroacetyl)-1H-pyrrol-1-yl)benzonitrile (104b).
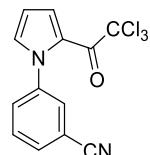
Compound **104b** was obtained as a yellow solid (0.56 g, 60 %) from compound **103b** (0.50 g, 2.97 mmol), according to the procedure used for **4a**; mp 120-123 °C. ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.78 – 7.71 (m, 1H), 7.71 – 7.66 (m, 1H), 7.65 – 7.51 (m, 3H), 7.12 – 7.03 (m, 1H), 6.50 – 6.40 (m, 1H). ESI-MS m/z: 314.7$(M + 1)^+$

Synthesis of ethyl 1H-pyrrole-3-carboxylate (107).
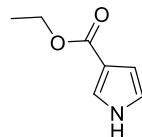
A suspension of TosMIC (10.5 g, 0.053 mol) and ethyl acrylates (0.056 mol) in dry ethyl ether/DMSO (100 mL/50 mL) was added dropwise under nitrogen to a suspension of NaH (3.75 g, 0.094 mol) in diethyl ether (50 mL) at room temperature. Then the mixture was stirred for 4–5 h. Ice water (200 mL) was added into the mixture. The aqueous phase was extracted with ethyl ether (3 × 100 mL). The organic phase was dried with anhydrous Na_2SO_4, concentrated under vacuum to get light brown liquid compounds which was purified by flash chromatography (AcOEt/hexane 2:8). Yellow oil (6.5 g, 88%); 1H-NMR (CDCl$_3$): δ 9.86 (s, 1H, exchangeable with D_2O), 7.45–7.46 (m, 1H), 6.76–6.79 (m, 1H), 6.65–6.67 (m, 1H), 4.31 (q, J = 7.1 Hz, 2H), 1.36 (t, J = 7.1 Hz, 3H).

Synthesis of N-benzyl-1-(3-nitrophenyl)-1H-pyrrole-2-carboxamide (108a).
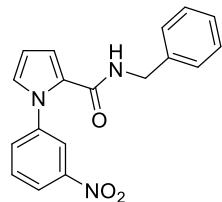
Compound **104a** (1.5 g, 4.37 mmol), benzylamine (1.4 g,13.1 mmol) and K_2CO_3 (21.8 mmol) were suspended in dry DMF (28 mL) under N_2 atmosphere. The vessel was sealed and the reaction was stirred at 120 °C for 2h, and then H_2O (50 mL) was added. The aqueous layer was extracted with EtOAc (3 × 50 mL), and the combined organic layers were washed with saturated aqueous solution of NaHCO$_3$ (3 x 30 mL), HCl 1N (3 x 30 mL), brine, dried over Na2SO4, filtered, and concentrated in vacuo. The residue was purified by silica gel chromatography (ethyl acetate : petroleum ether, 3:7) to give the title compound as a yellow solid (0.870 g, 62%), mp 98 – 100 °C. ^{1}H NMR (300 MHz, CDCl$_3$) δ 8.30 – 8.24 (m, 1H), 8.26 – 8.17 (m, 1H), 7.75 – 7.67 (m, 1H), 7.65 – 7.54 (m, 1H), 7.46 – 7.23 (m, 5H), 7.01 – 6.96 (m,

1H), 6.81 – 6.73 (m, 1H), 6.40 – 6.30 (m, 2H, 1H exchangeable with D_2O), 4.52 (d, J = 5.7 Hz, 2H). ESI-MS m/z: 322.20 (M + 1)$^+$.

Synthesis of N-benzyl-1-(3-cyanophenyl)-1H-pyrrole-2-carboxamide (108b).
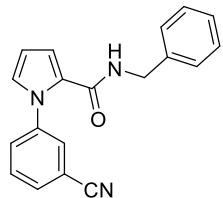
Compound **108b** was obtained as a yellow solid (0.36 g, 42 %) from compound **104b** (0.9 g, 2.8 mmol), according to the procedure used for **108a**, 98 – 100 °C. ^{1}H NMR (300 MHz, CDCl$_3$)**:** δ 7.70 – 7.47 (m, 4H), 7.44 – 7.23 (m, 5H), 6.89 (dd, J = 2.7, 1.6 Hz, 1H), 6.72 (dd, J = 3.8, 1.6 Hz, 1H), 6.29 (dd, J = 3.8, 2.7 Hz, 1H), 6.19 (t, J = 5.8 Hz, 1H, exchangeable with D_2O), 4.50 (d, J = 5.8 Hz, 1H). ESI-MS m/z**:** 301.97 (M + 1)$^+$

Synthesis of ethyl 5-(2,2,2-trichloroacetyl)-1H-pyrrole-3-carboxylate (109).
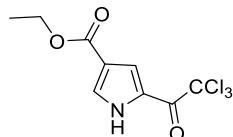
To a solution of **107** (2.15 mmol, 0.300 g) in dry DCM (3 mL) in sealed vessel was added dropwise ClCOCCl$_3$ (21.5 mmol, 3.9 g). The reaction was left to stir heating at 80°C for 36 h. The mixture of reaction was diluted with DCM (50 mL) and washed with saturated aqueous solution of NaHCO$_3$ (3 x 30 mL), brine and dried over Na$_2$SO$_4$. The organic layer was filtered and evaporated under reduced pressure. The title compound was obtained as yellow solid (1.13 g, 50%) after chromatographyc purification (AcOEt/hexane 2:8); mp 115-118 °C. ^{1}H NMR (300 MHz, CDCl$_3$): ^{1}H NMR (300 MHz, CDCl$_3$) δ 9.68

(bs, 1H, exchangeable with D_2O), 7.76 – 7.72 (m, J = 2.5 Hz, 1H), 7.72 – 7.69 (m, J = 3.3 Hz, 1H), 4.34 (q, J = 7.1 Hz, 2H), 1.37 (t, J = 7.1 Hz, 3H). ESI-MS m/z**:** 285.1 $(M + 1)^+$.

Synthesis of ethyl 5-(benzylcarbamoyl)-1H-pyrrole-3-carboxylate (110)
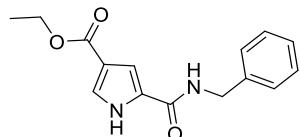
To a solution of 109 (3.20 mmol, 0.860 g) in DCM dry (16 mL) was added benzylamine (3.52 mmol, 0.377 g). The reaction was left to stir at room temperature for 12h and evaporated under under reduced pressure. The residue was purified by silica gel chromatography (ethyl acetate : petroleum ether, 3:7) to afford the desired product as white solid (0.740 g, 85%), mp 140 – 142 °C. ^{1}H NMR (300 MHz, CDCl$_3$): δ 10.00 (bs, 1H, exchangeable with D_2O), 7.53 – 7.46 (m, 1H), 7.41 – 7.29 (m, 5H), 7.07 – 6.92 (m, 1H), 6.29 (t, J = 5.8 Hz, 1H, exchangeable with D_2O), 4.61 (d, J = 5.8 Hz, 2H), 4.27 (q, J = 7.1 Hz, 2H), 1.33 (t, J = 7.1 Hz, 3H). ESI-MS m/z**:** 273.42 $(M + 1)^+$.

Synthesis of ethyl 5-(benzylcarbamoyl)-1-(3-nitrophenyl)-1H-pyrrole-3-carboxylate (108c).

Compound **108c** was obtained as yellow solid (0.346 g, 80%), from compound **110** (1.10 mmol, 0.300 g) and 3-bromonitrobenzene (2.64 mmol, 0.53 g), according to the procedure used for **97**; mp 156 – 157 °C. ^{1}H NMR (300 MHz,
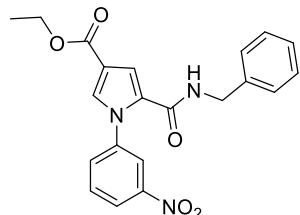
CDCl$_3$) δ 8.32 – 8.26 (m, , 1H), 8.23 – 8.19 (m, 1H), 7.74 – 7.59 (m, , 8.0 Hz, 2H), 7.54 – 7.51 (m, 1H), 7.40 – 7.25 (m, 5H), 7.17 – 7.13 (m, 1H), 6.23 (t, *J* = 5.8 Hz, 1H, exchangeable with D$_2$O), 4.49 (d, *J* = 5.8 Hz, 2H), 4.31 (q, *J* = 7.1 Hz, 2H), 1.35 (t, *J* = 7.1 Hz, 3H). ESI-MS m/z: 393.99 (M + 1)$^{+}$.

Synthesis of ethyl 5-(benzylcarbamoyl)-1-(3-cyanophenyl)-1H-pyrrole-3-carboxylate (108 d)
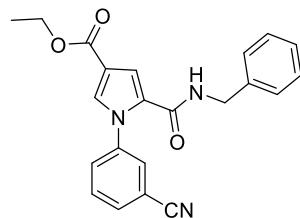
Compound **108c** was obtained as yellow solid (0.231 g, 85%), from compound **110** (0.73 mmol, 0.200 g) and 3-bromonitrobenzene (1.61mmol, 0.30 g), according to the procedure used for **97**, mp164 – 165 °C. ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.73 – 7.64 (m, 1H), 7.63 – 7.52 (m, 5H), 7.50 – 7.46 (m, 1H), 7.43 – 7.37 (m, 2H), 7.26 – 7.22 (m, 1H), 7.17 (m, 1H), 6.41 (t, *J* = 5.9 Hz, 1H, exchangeable with D$_2$O), 4.54 (d, *J* = 5.9 Hz, 2H), 4.28 (q, *J* = 7.1 Hz, 2H), 1.33 (t, *J* = 7.1 Hz, 3H). ESI-MS m/z: 374.04 (M + 1)$^{+}$.

General Procedure for reduction of nitro group 109a and 109c.

Example: Synthesis of 1-(3-aminophenyl)-N-benzyl-1H-pyrrole-2-carboxamide (109a).
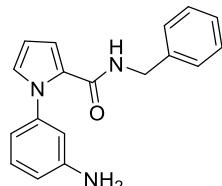
To a solution of **108a** (0.517 g, 17.20 mmol) in acetic acid (26 mL) was added a powder of Zn (1.12 g, 17.20 mmol). The reaction was stirred for 1 h at room

temperature, filtered, and concentrated in vacuo. The residue was dissolved in saturated aqueous solution of $NaHCO_3$ (100 mL) and extracted with EtOAc (3 × 50 mL). The combined organic layers were washed with brine, dried (Na_2SO_4), and filtered. Vacuum evaporation of the solvents gave the title compound (0.300 g, 65%) as a yellow solid which was directly used in the next step without further step of purification. ^{1}H NMR (300 MHz, CDCl$_3$): δ 8.55 (t, *J* = 5.7 Hz, 1H, exchangeable with deuterium oxide), 7.40 – 7.16 (m, 5H), 7.05 – 6.92 (m, 2H), 6.85 – 6.77 (m, 1H), 6.54 – 6.48 (m, 1H), 6.47 – 6.42 (m, 1H), 6.39 – 6.31 (m, 1H), 6.20 – 6.15 (m, 1H), 5.24 (s, 2H, exchangeable with deuterium oxide), 4.33 (d, *J* = 5.7 Hz, 2H). ESI-MS m/z: 582.69 (M x 2)$^+$

Ethyl 1-(3-aminophenyl)-5-(benzylcarbamoyl)-1H-pyrrole-3-carboxylate 109c:
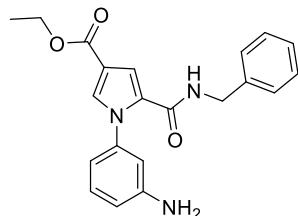
^{1}H NMR (300 MHz, DMSO) δ 9.83 (t, *J* = 5.5 Hz, 1H, exchangeable with deuterium oxide), 8.33 – 8.29 (m, 1H), 7.98 – 7.88 (m, 1H), 7.86 – 7.77 (m, 1H), 7.43 – 7.30 (m, 3H), 7.30 – 7.22 (m, 1H), 7.20 – 7.08 (m, 1H), 6.88 – 6.72 (m, 2H), 6.62 – 6.51 (m, 1H), 5.42 (s, 2H, exchangeable with deuterium oxide), 4.52 (d, *J* = 5.5 Hz, 2H), 4.27 (q, *J* = 7.0 Hz, 2H), 1.28 (t, *J* = 7.0 Hz, 3H). ESI-MS m/z: 364.12 (M + 1)$^+$

General Procedure for reduction of nitrile group 108b and 108d

Example: Synthesis of 1-(3-(aminomethyl)phenyl)-N-benzyl-1H-pyrrole-2-carboxamide (109b).
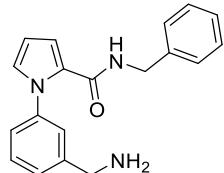
To a solution of $NiSO_4 \cdot 6H_2O$ (0.041 g, 0.16 mmol), and **5b** (0.470 g, 1.56 mmol) in a mixture of MeOH (30 ml) was added $NaBH_4$ (0.424 g, 10.92 mmol) portionwise. The mixture was left to stir, at room temperature, for 3 h and then water was added (50 mL). The aqueous layer was extracted with EtOAc (3 x 50 mL). The combined layers were washed with brine, dried (Na_2SO_4), and filtered. Vacuum evaporation of the solvent gave a compound **108b** (0.44 g, 90%) as a slurry which was directly used in the next step. ^{1}H NMR (300 MHz, DMSO) δ 8.71 – 8.54 (m, 1H, exchangeable with deuterium oxide), 7.40 – 7.16 (m, 9H), 7.16 – 6.99 (m, 3H), 6.98 – 6.83 (m, 1H), 6.27 – 6.15 (m, 1H), 4.43 – 4.29 (m, 2H), 3.83 – 3.62 (m, 2H). ESI-MS m/z: 306.08 $(M + 1)^+$.

Ethyl 1-(3-(aminomethyl)phenyl)-5-(benzylcarbamoyl)-1H-pyrrole-3-carboxylate (109d)
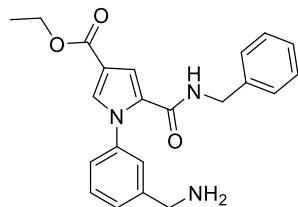
^{1}H NMR (300 MHz, DMSO): δ 9.88 – 9.78 (m, 1H), 8.16 – 8.06 (m, 1H), 8.09 – 7.84 (m, 1H), 7.84 – 7.66 (m, 1H), 7.65 – 7.58 (m, 1H), 7.53 – 7.17 (m, 7H),

4.53 (d, J = 4.9 Hz, 2H), 4.27 (q, J = 6.6 Hz, 2H), 3.83 – 3.78 (m, 2H), 1.30 (t, J = 6.6 Hz, 3H). ESI-MS m/z: 378.06 (M + 1)$^+$.

General procedure for coupling reaction 110 a – d and *ent*-110a – d

Example: Synthesis of (S)-tert-butyl (1-((3-(2-(benzylcarbamoyl)-1H-pyrrol-1-yl)phenyl)amino)-1-oxopropan-2-yl)carbamate (110a).
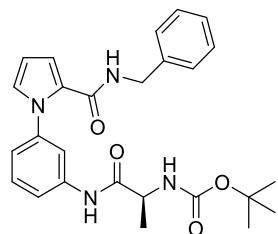
To a solution of **109a** (0.300 g, 1.30 mmol) , DMAP (0.013 g,0.130 mmol), and Boc-L-AlaOH (0.368 g,1.95 mmol) in dichlorometane dry (20 mL) was added DCC (0.402 g, 1.95 mmol) under N$_2$ atmosphere. The reaction was left to stir at room temperature overnight, filtered, and concentrated in vacuum. The residue was dissolved in dichlorometane (100 mL) and washed with saturated aqueous solution of NaHCO$_3$ (3 x 30 mL), and brine. The organic phase was dried (Na$_2$SO$_4$), filtered and concentrated in vacuo. The residue was purified by silica gel chromatography (ethyl acetate : petroleum ether, 1:1) to give the title compound as a yellow solid (0.420 g, 70 %), mp 116 – 118 °C. [α]$_D$ -40.3 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO): δ 10.05 (s, 1H, exchangeable with deuterium oxide), 8.66 (t, J = 5.6 Hz, 1H, exchangeable with deuterium oxide), 7.68 – 7.63 (m, 1H), 7.55 – 7.46 (m, 1H), 7.37 – 7.27 (m, 3H), 7.27 – 7.20 (m, 3H), 7.09 (d, J = 6.8 Hz, 1H), 7.06 – 7.00 (m, 1H), 6.96 – 6.85 (m, 2H), 6.26 – 6.20 (m, 1H, exchangeable with deuterium oxide), 4.32 (d, J = 5.6 Hz, 2H), 4.18 – 4.06 (m, 1H), 1.38 (s, 9H), 1.26 (d, J = 6.8 Hz, 3H). ESI-MS m/z: 462.83 (M + 1)$^+$

**Tert-butyl (R)-(1-((3-(2-(benzylcarbamoyl)-1H-pyrrol-1-l)phenyl)amino)-1-oxopropan-2-yl)carbamate (*ent*-110a)**
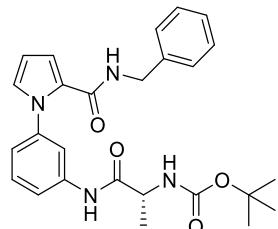
Yellow solid (0.420 g, 70 %), mp 116 – 118 °C, [α]$_D$ +39.0 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO): δ 10.05 (s, 1H, exchangeable with deuterium oxide), 8.66 (t, *J* = 5.6 Hz, 1H, exchangeable with deuterium oxide), 7.68 – 7.63 (m, 1H), 7.55 – 7.46 (m, 1H), 7.37 – 7.27 (m, 3H), 7.27 – 7.20 (m, 3H), 7.09 (d, *J* = 6.8 Hz, 1H), 7.06 – 7.00 (m, 1H), 6.96 – 6.85 (m, 2H), 6.26 – 6.20 (m, 1H, exchangeable with deuterium oxide), 4.32 (d, *J* = 5.6 Hz, 2H), 4.18 – 4.06 (m, 1H), 1.38 (s, 9H), 1.26 (d, *J* = 6.8 Hz, 3H). ESI-MS m/z: 462.83 (M + 1)[+]

Tert-butyl (S)-(1-((3-(2-(benzylcarbamoyl)-1H-pyrrol-1-yl)benzyl)amino)-1-oxopropan-2-yl)carbamate (110b)
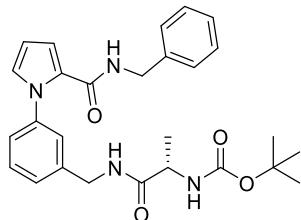
White solid (0.340 g, 50%), mp 120 – 122 °C. [α]$_D$ - 6.7 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO): δ 8.68 (t, *J* = 5.9 Hz, 1H, exchangeable with deuterium oxide), 8.34 (t, *J* = 5.1 Hz, 1H, exchangeable with deuterium oxide), 7.37 – 7.25 (m, 5H), 7.25 – 7.16 (m, 3H), 7.15 – 7.08 (m, 1H), 7.06 – 7.02 (m, 1H), 6.99 – 6.88 (m, 2H, 1H exchangeable with deuterium oxide),

6.26 – 6.22 (m, 1H), 4.38 – 4.27 (m, 4H), 4.00 (s, 1H), 1.38 (s, 9H), 1.21 (d, *J* = 7.0 Hz, 3H). ESI-MS m/z: 499.05 (M +Na)⁺, 476.81 (M + 1)⁺.

6.26 – 6.22 (m, 1H), 4.38 – 4.27 (m, 4H), 4.00 (s, 1H), 1.38 (s, 9H), 1.21 (d, *J* = 7.0 Hz, 3H). ESI-MS m/z: 499.05 $(M +Na)^+$, 476.81 $(M + 1)^+$.

Tert-butyl (R)-(1-((3-(2-(benzylcarbamoyl)-1H-pyrrol-1-yl)benzyl)amino)-1-oxopropan-2-yl)carbamate (*Ent*-110b)
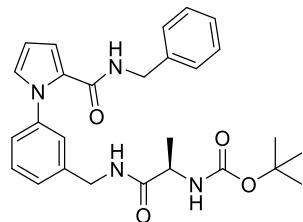
White solid (0.340 g, 50%), mp 120 – 122 °C, $[\alpha]_D$ + 6.9 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO): δ 8.68 (t, *J* = 5.9 Hz, 1H, exchangeable with deuterium oxide), 8.34 (t, *J* = 5.1 Hz, 1H, exchangeable with deuterium oxide), 7.37 – 7.25 (m, 5H), 7.25 – 7.16 (m, 3H), 7.15 – 7.08 (m, 1H), 7.06 – 7.02 (m, 1H), 6.99 – 6.88 (m, 2H, 1H exchangeable with deuterium oxide), 6.26 – 6.22 (m, 1H), 4.38 – 4.27 (m, 4H), 4.00 (s, 1H), 1.38 (s, 9H), 1.21 (d, *J* = 7.0 Hz, 3H). ESI-MS m/z: 499.05 $(M +Na)^+$, 476.81 $(M + 1)^+$.

Ethyl (S)-5-(benzylcarbamoyl)-1-(3-(2-((tert-butoxycarbonyl)amino) propanamido) phenyl)-1H-pyrrole-3-carboxylate (110c)
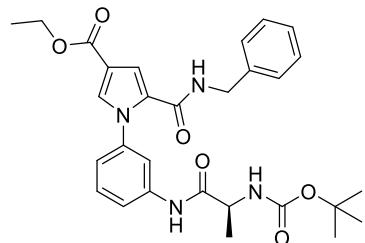
White solid (0.120 g, 83%), mp 150 – 151 °C. $[\alpha]_D$ – 38.3(c 0.3, MeOH). ^{1}H NMR (300 MHz, CDCl$_3$) δ 8.79 – 8.67 (m, 1H, exchangeable with D$_2$O), 7.79 – 7.76 (m, 1H), 7.50 – 7.48 (m, 1H), 7.38 – 7.25 (m, 8H), 7.15 – 7.12 (m, 1H),

7.09 – 7.04 (m, *J* = 7.1 Hz, 1H, exchangeable with D2O), 6.25 (t, *J* = 5.6 Hz, 1H, exchangeable with D$_2$O), 5.08 – 4.99 (m, 1H), 4.50 (d, *J* = 5.7 Hz, 2H), 4.30 (q, *J* = 7.1 Hz, 2H), 1.48 (s, 9H), 1.42 (d, *J* = 7.0 Hz, 3H), 1.35 (t, *J* = 7.1 Hz, 3H). ESI-MS m/z: 1068.91 (M x 2)$^+$.

Ethyl (R)-5-(benzylcarbamoyl)-1-(3-(2-((tert-butoxycarbonyl)amino) propanamido)phenyl)-1H-pyrrole-3-carboxylate (*Ent*-110c)

White solid (0.120 g, 83%), mp 150 – 151 °C. [α]$_D$ +29.7 (c 0.3, MeOH). ^{1}H NMR (300 MHz, CDCl$_3$) δ 8.79 – 8.67 (m, 1H, exchangeable with D$_2$O), 7.79 – 7.76 (m, 1H), 7.50 – 7.48 (m, 1H), 7.38 – 7.25 (m, 8H), 7.15 – 7.12 (m, 1H), 7.09 – 7.04 (m, *J* = 7.1 Hz, 1H, exchangeable with D2O), 6.25 (t, *J* = 5.6 Hz, 1H, exchangeable with D$_2$O), 5.08 – 4.99 (m, 1H), 4.50 (d, *J* = 5.7 Hz, 2H), 4.30 (q, *J* = 7.1 Hz, 2H), 1.48 (s, 9H), 1.42 (d, *J* = 7.0 Hz, 3H), 1.35 (t, *J* = 7.1 Hz, 3H). ESI-MS m/z: 1068.91 (M x 2)$^+$.

Ethyl-(S)-5-(benzylcarbamoyl)-1-(3-((2-((tert-butoxycarbonyl)amino) propanamido)methyl)phenyl)-1H-pyrrole-3-carboxylate110d:

White solid (0.073 g, 50%), mp 86 – 88 °C. [α]$_D$ – 6.00 (c 0.3, MeOH). ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.50 – 7.47 (m, 1H), 7.44 – 7.23 (m, 10H), 7.15 –

7.13 (m, 1H), 6.63 (t, *J* = 4.8 Hz, 1H), 6.21 (t, *J* = 6.1 Hz, 1H), 5.08 – 4.97 (m, 1H), 4.60 – 4.39 (m, 4H), 4.31 (q, *J* = 7.1 Hz, 2H), 4.22 – 4.10 (m, 1H), 1.41 (s, 9H), 1.40 – 1.32 (m, 6H). ESI-MS m/z: 549.07 (M + 1)$^+$.

Ethyl (R)-5-(benzylcarbamoyl)-1-(3-((2-((tert- butoxycarbonyl)amino) propanamido) methyl)phenyl)-1H-pyrrole-3-carboxylate *Ent*-110d:
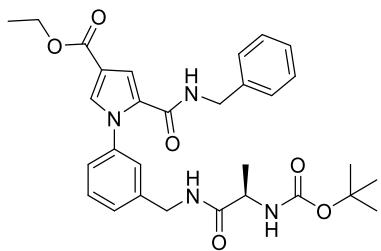
White solid (0.073 g, 50%), mp 86 – 88 °C. [α]$_D$ + 7.8 (c 0.3, MeOH). ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.50 – 7.47 (m, 1H), 7.44 – 7.23 (m, 10H), 7.15 – 7.13 (m, 1H), 6.63 (t, *J* = 4.8 Hz, 1H), 6.21 (t, *J* = 6.1 Hz, 1H), 5.08 – 4.97 (m, 1H), 4.60 – 4.39 (m, 4H), 4.31 (q, *J* = 7.1 Hz, 2H), 4.22 – 4.10 (m, 1H), 1.41 (s, 9H), 1.40 – 1.32 (m, 6H). ESI-MS m/z: 549.07 (M + 1)$^+$.

General procedures for acids 110e, f and ent-110e, f.

Example: synthesis   of   (S)-5-(benzylcarbamoyl)-1-(3-(2-((tert-butoxycarbonyl)amino)propanamido)phenyl)-1H-pyrrole-3-carboxylic acid (110e).
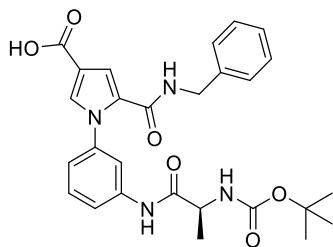
 LiOH (0.053 g, 2.22 mmol) was added to a solution of **110c** (0.200 g, 0.37 mmol) in THF−H$_2$O (4:1, 10 mL). The reaction was stirred at 45 °C for 5 h and then concentrated in vacuo. The residue was dissolved in water (20 mL),

washed with CH$_2$Cl$_2$ (3× 10 mL), and acidified to pH 5 with 1N HCl and extracted with ethyl acetate (3 x 10 mL). The combined organic layers were washed with brine, dried (Na$_2$SO$_4$), filtered and evaporated under reduced pressure. Purification by crystallization (MeOH/H$_2$O) gave the title compound as a white solid (0.131 g, 85%), mp 224 – 227 °C. [α]$_D$ – 16.0 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO) δ 12.26 (bs, 1H, exchangeable with D$_2$O), 10.14 – 10.05 (m, 1H, exchangeable with D$_2$O), 8.96 – 8.82 (m, 1H, exchangeable with D$_2$O), 7.72 – 7.67 (m, 1H), 7.57 – 7.49 (m, 2H), 7.39 – 7.16 (m, 6H), 7.14 – 7.07 (m, 1H,), 7.03 – 6.96 (m, 1H, exchangeable with D$_2$O), 4.32 (d, J = 5.6 Hz, 2H), 4.17 – 4.04 (m, 1H), 1.38 (s, 9H), 1.26 (d, J = 6.9 Hz, 3H). ESI-MS m/z: 529.01 (M + Na)$^+$.

(R)-5-(benzylcarbamoyl)-1-(3-(2-((tert-butoxycarbonyl)amino) propanamido)phenyl)-1H-pyrrole-3-carboxylic acid (*Ent*-110e)
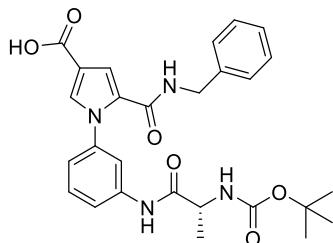
White solid (0.131 g, 85%), mp 224 – 227 °C, [α]$_D$ + 17.7 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO) δ 12.26 (bs, 1H, exchangeable with D$_2$O), 10.14 – 10.05 (m, 1H, exchangeable with D$_2$O), 8.96 – 8.82 (m, 1H, exchangeable with D$_2$O), 7.72 – 7.67 (m, 1H), 7.57 – 7.49 (m, 2H), 7.39 – 7.16 (m, 6H), 7.14 – 7.07 (m, 1H,), 7.03 – 6.96 (m, 1H, exchangeable with D$_2$O), 4.32 (d, J = 5.6 Hz, 2H), 4.17 – 4.04 (m, 1H), 1.38 (s, 9H), 1.26 (d, J = 6.9 Hz, 3H). ESI-MS m/z: 529.01 (M + Na)$^+$.

General procedures for amino derivatives 111 a – f and ent-111a – f.

Example: synthesis of (S)-1-(3-(2-aminopropanamido)phenyl)-N-benzyl-1H-pyrrole-2-carboxamide 2,2,2-trifluoroacetate (111a).
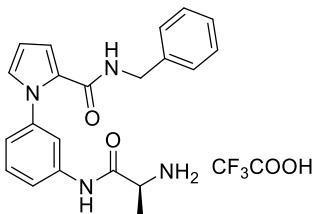
A mixture of TFA and CH_2Cl_2 (1:9, 12 mL) was added to compound **110a** (0.200 g, 0.43 mmol). The reaction was stirred at room temperature for 1h and then concentrated in vacuo to give the title compound as a white solid (0.200 g, 99%), mp 119 - 122 °C. $[\alpha]_D$ + 5.3 (c 0.3, MeOH). ^{1}H NMR (400 MHz, DMSO) δ 10.54 (bs, 1H, exchangeable with D_2O), 8.71 (t, J = 6.1 Hz, 1H, exchangeable with D_2O), 8.16 (s, 3H, exchangeable with D_2O), 7.61 – 7.59 (m, 1H), 7.57 – 7.53 (m, 1H), 7.41 – 7.35 (m, 1H), 7.34 – 7.28 (m, 3H), 7.27 – 7.20 (m, 2H), 7.05 (dd, J = 2.6, 1.8 Hz, 1H), 7.03 – 6.99 (m, 1H), 6.93 (dd, J = 3.8, 1.8 Hz, 1H), 6.26 (dd, J = 3.8, 2.6 Hz, 1H), 4.33 (d, J = 6.1 Hz, 2H), 4.04 – 3.97 (m, 1H). ESI-MS m/z: 363.01 $(M + 1)^+$.

(R)-1-(3-(2-aminopropanamido)phenyl)-N-benzyl-1H-pyrrole-2-carboxamide 2,2,2-trifluoroacetate *ent*-111a*:*
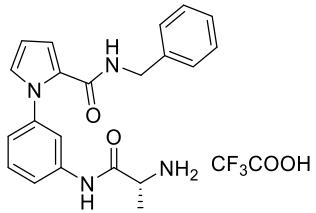
White solid (yield 99%), mp 124 – 125 °C. $[\alpha]_D$ – 6.9 (c 0.3, MeOH). ^{1}H NMR (400 MHz, DMSO) δ 8.91 (t, J = 5.9 Hz, 1H, exchangeable with D_2O), 8.72 (t, J = 6.1 Hz, 1H, exchangeable with D_2O), 8.05 (s, 3H, exchangeable with D_2O), 7.40 – 7.29 (m, 3H), 7.29 – 7.19 (m, 5H), 7.17 – 7.12 (m, J = 7.9 Hz, 1H), 7.04 (dd, J = 2.5, 1.8 Hz, 1H), 6.93 (dd, J = 3.8, 1.8 Hz, 1H), 6.26 (dd, J = 3.8, 2.5 Hz, 1H), 4.38 (d, J = 5.9 Hz, 2H), 4.33 (d, J = 6.1 Hz, 2H), 3.91 – 3.83 (m, J = 6.9 Hz, 1H), 1.38 (d, J = 7.0 Hz, 3H). ESI-MS m/z: 363.01 $(M + 1)^+$.

(S)-1-(3-((2-aminopropanamido)methyl)phenyl)-N-benzyl-1H-pyrrole-2-carboxamide 2,2,2-trifluoroacetate 111b
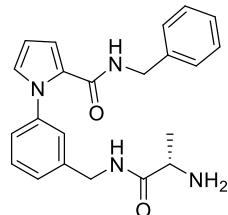

White solid (99%), mp 136 – 138 °C. $[\alpha]_D$ + 6.7 (c 0.3, MeOH). ^{1}H NMR (400 MHz, DMSO) δ 8.91 (t, *J* = 5.9 Hz, 1H, exchangeable with D$_2$O), 8.72 (t, *J* = 6.1 Hz, 1H, exchangeable with D$_2$O), 8.05 (s, 3H, exchangeable with D$_2$O), 7.40 – 7.29 (m, 3H), 7.29 – 7.19 (m, 5H), 7.17 – 7.12 (m, *J* = 7.9 Hz, 1H), 7.04 (dd, *J* = 2.5, 1.8 Hz, 1H), 6.93 (dd, *J* = 3.8, 1.8 Hz, 1H), 6.26 (dd, *J* = 3.8, 2.5 Hz, 1H), 4.38 (d, *J* = 5.9 Hz, 2H), 4.33 (d, *J* = 6.1 Hz, 2H), 3.91 – 3.83 (m, *J* = 6.9 Hz, 1H), 1.38 (d, *J* = 7.0 Hz, 3H). ESI-MS m/z: 377.11 (M + 1)$^+$

(R)-1-(3-((2-aminopropanamido)methyl)phenyl)-N-benzyl-1H-pyrrole-2-carboxamide 2,2,2-trifluoroacetate *ent*-111b
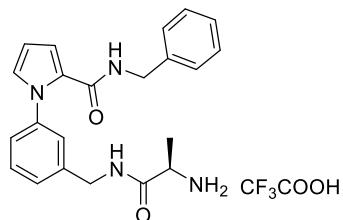
White solid (99%), mp 136 – 138 °C. $[\alpha]_D$ – 6.2 (c 0.3, MeOH). ^{1}H NMR (400 MHz, DMSO) δ 8.91 (t, *J* = 5.9 Hz, 1H), 8.72 (t, *J* = 6.1 Hz, 1H), 8.05 (s, 1H), 7.40 – 7.29 (m, 2H), 7.29 – 7.19 (m, 4H), 7.17 – 7.12 (m, *J* = 7.9 Hz, 1H), 7.04 (dd, *J* = 2.5, 1.8 Hz, 1H), 6.93 (dd, *J* = 3.8, 1.8 Hz, 1H), 6.26 (dd, *J* = 3.8, 2.5 Hz, 1H), 4.38 (d, *J* = 5.9 Hz, 2H), 4.33 (d, *J* = 6.1 Hz, 2H), 3.91 – 3.83 (m, *J* = 6.9 Hz, 1H), 1.38 (d, *J* = 7.0 Hz, 3H). ESI-MS m/z: 377.11 (M + 1)$^+$

(S)-ethyl 1-(3-(2-aminopropanamido)phenyl)-5-(benzylcarbamoyl)-1H-pyrrole-3-carboxylate 2,2,2-trifluoroacetate 111c
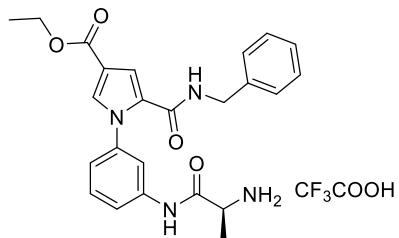
White solid (99%), mp 199 – 200 °C. $[\alpha]_D$ + 5.9 (c 0.3, MeOH). ^{1}H NMR (400 MHz, DMSO) δ 10.13 (s, 1H, exchangeable with D_2O), 8.51 (t, J = 6.1 Hz, 1H, exchangeable with D_2O), 7.73 (s, 3H, exchangeable with D_2O), 7.20 – 7.17 (m, 2H), 7.17 – 7.12 (m, 1H), 7.01 – 6.94 (m, 1H), 6.90 – 6.84 (m, 3H), 6.82 – 6.75 (m, 3H), 6.67 – 6.58 (m, 1H), 3.87 (d, J = 6.0 Hz, 2H), 3.78 (q, J = 7.1 Hz, 2H), 3.60 – 3.53 (m, 1H), 1.01 (d, J = 7.0 Hz, 3H), 0.83 (t, J = 7.1 Hz, 3H). ESI-MS m/z: 435.06 (M + 1)$^+$.

(R)-ethyl 1-(3-(2-aminopropanamido)phenyl)-5-(benzylcarbamoyl)-1H-pyrrole-3-carboxylate 2,2,2-trifluoroacetate *ent*-111c
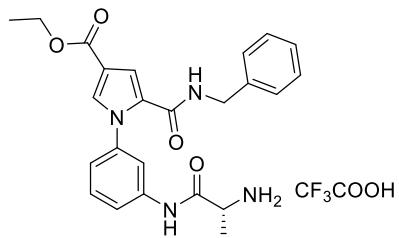
White solid (99%), mp 199 – 200 °C. $[\alpha]_D$ – 3.33 (c 0.3, MeOH). ^{1}H NMR (400 MHz, DMSO) δ 10.13 (s, 1H, exchangeable with D_2O), 8.51 (t, J = 6.1 Hz, 1H, exchangeable with D_2O), 7.73 (s, 3H, exchangeable with D_2O), 7.20 – 7.17 (m, 2H), 7.17 – 7.12 (m, 1H), 7.01 – 6.94 (m, 1H), 6.90 – 6.84 (m, 3H), 6.82 – 6.75 (m, 3H), 6.67 – 6.58 (m, 1H), 3.87 (d, J = 6.0 Hz, 2H), 3.78 (q, J =

7.1 Hz, 2H), 3.60 – 3.53 (m, 1H), 1.01 (d, *J* = 7.0 Hz, 3H), 0.83 (t, *J* = 7.1 Hz, 3H). ESI-MS m/z: 435.06 (M + 1)$^+$.

Synthesis of (S)-ethyl 1-(3-((2-aminopropanamido)methyl)phenyl)-5-(benzylcarbamoyl)-1H-pyrrole-3-carboxylate 2,2,2-trifluoroacetate 111d
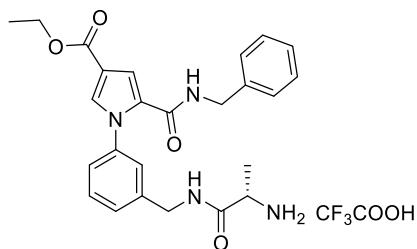
White solid (yield 99%), mp 123 – 126 °C. [α]$_D$ + 12.3 (c 0.3, MeOH). ^{1}H NMR (400 MHz, DMSO) δ 8.97 (t, *J* = 5.9 Hz, 1H, exchangeable with D$_2$O), 8.90 (t, *J* = 5.7 Hz, 1H, exchangeable with D$_2$O), 8.06 (s, 3H, exchangeable with D$_2$O), 7.64 – 7.57 (m, 1H), 7.43 – 7.37 (m, 1H), 7.36 – 7.27 (m, 9H), 4.47 – 4.37 (m, 2H), 4.32 (d, *J* = 5.9 Hz, 2H), 4.24 (q, *J* = 7.1 Hz, 2H), 3.92 – 3.85 (m, *J* = 6.8 Hz, 1H), 1.39 (d, *J* = 6.9 Hz, 2H), 1.28 (t, *J* = 7.1 Hz, 2H). ESI-MS m/z: 449.02 (M + 1)$^+$

Synthesis of (R)-ethyl 1-(3-((2-aminopropanamido)methyl)phenyl)-5-(benzylcarbamoyl)-1H-pyrrole-3-carboxylate 2,2,2-trifluoroacetate *ent-111d*

White solid (yield 99%), mp 123 – 126 °C. [α]$_D$ + 12.3 (c 0.3, MeOH). ^{1}H NMR (400 MHz, DMSO) δ 8.97 (t, *J* = 5.9 Hz, 1H, exchangeable with D$_2$O), 8.90 (t, *J* = 5.7 Hz, 1H, exchangeable with D$_2$O), 8.06 (s, 3H, exchangeable with D$_2$O), 7.64 – 7.57 (m, 1H), 7.43 – 7.37 (m, 1H), 7.36 – 7.27 (m, 9H), 4.47

– 4.37 (m, 2H), 4.32 (d, *J* = 5.9 Hz, 2H), 4.24 (q, *J* = 7.1 Hz, 2H), 3.92 – 3.85 (m, *J* = 6.8 Hz, 1H), 1.39 (d, *J* = 6.9 Hz, 2H), 1.28 (t, *J* = 7.1 Hz, 2H). ESI-MS m/z: 449.02 (M + 1)$^+$.

Synthesis of (S)-1-((3-(2-(benzylcarbamoyl)-4-carboxy-1H-pyrrol-1-yl)phenyl)amino)- 1- oxopropan-2-aminium 2,2,2-trifluoroacetate 111e.

White solid (0.051 g, 99%), 219 – 220 °C. [α]$_D$ + 3.7 (c 0.3, MeOH). ^{1}H NMR (400 MHz, DMSO) δ 10.57 (s, 1H, exchangeable with D$_2$O), 8.93 (t, J = 6.1 Hz, 1H, exchangeable with D$_2$O), 8.40 (s, 3H, exchangeable with D$_2$O), 7.65 – 7.61 (m, 1H), 7.62 – 7.54 (m, 2H), 7.45 – 7.38 (m, 1H), 7.35 – 7.27 (m, 2H), 7.28 – 7.20 (m, 3H), 7.11 – 7.05 (m, 2H), 4.32 (d, J = 6.0 Hz, 2H), 4.04 – 3.97 (m, J = 7.0 Hz, 1H), 1.46 (d, J = 7.0 Hz, 3H). ESI-MS m/z: 407.05 (M + 1)$^+$.

Synthesis of (R)-1-((3-(2-(benzylcarbamoyl)-4-carboxy-1H-pyrrol-1-yl)phenyl)amino)- 1- oxopropan-2-aminium 2,2,2-trifluoroacetate *ent-111e*.
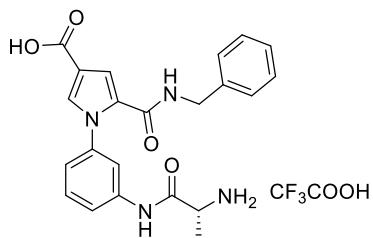
White solid (99%), 219 – 220 °C. [α]$_D$ - 1.7 (c 0.3, MeOH). ^{1}H NMR (400 MHz, DMSO) δ 10.57 (s, 1H, exchangeable with D$_2$O), 8.93 (t, J = 6.1 Hz, 1H, exchangeable with D$_2$O), 8.40 (s, 3H, exchangeable with D$_2$O), 7.65 –

7.61 (m, 1H), 7.62 – 7.54 (m, 2H), 7.45 – 7.38 (m, 1H), 7.35 – 7.27 (m, 2H), 7.28 – 7.20 (m, 3H), 7.11 – 7.05 (m, 2H), 4.32 (d, J = 6.0 Hz, 2H), 4.04 – 3.97 (m, J = 7.0 Hz, 1H), 1.46 (d, J = 7.0 Hz, 3H). ESI-MS m/z: 407.05 (M + 1)⁺.

Synthesis of (S)-1-((3-(2-(benzylcarbamoyl)-4-carboxy-1H-pyrrol-1-yl)benzyl)amino)-1-oxopropan-2-aminium 2,2,2-trifluoroacetate 111f.
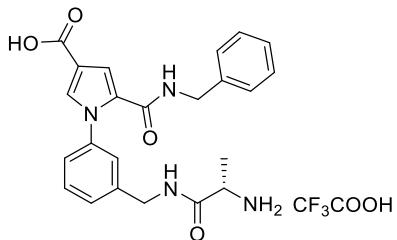
White solid (0.069 g, 99%), 203 – 205 °C. [α]$_D$ + 13.00 (c 0.3, MeOH). ¹H NMR (400 MHz, DMSO) δ 10.57 (s, 1H, exchangeable with D₂O), 8.94 (t, *J* = 6.1 Hz, 1H), 8.89 (t, *J* = 5.9 Hz, 1H), 8.76 (s, 3H), 7.56 – 7.53 (m, 1H), 7.42 – 7.35 (m, 1H), 7.35 – 7.29 (m, 2H), 7.29 – 7.25 (m, 5H), 7.24 – 7.18 (m, 2H), 4.45 – 4.36 (m, 2H), 4.33 (d, *J* = 6.0 Hz, 2H), 3.91 – 3.82 (m, 1H), 1.38 (d, *J* = 7.0 Hz, 3H). ESI-MS m/z: 421.08 (M + 1)⁺.

Synthesis of (R)-1-((3-(2-(benzylcarbamoyl)-4-carboxy-1H-pyrrol-1-yl)benzyl)amino)-1-oxopropan-2-aminium 2,2,2-trifluoroacetate *ent*-111f.
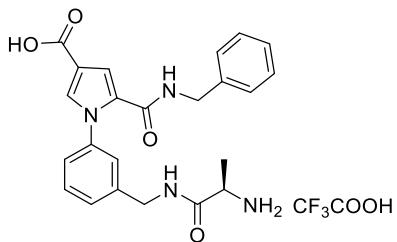
White solid (99%), mp 200 – 202 °C. [α]$_D$ – 12.7 (c 0.3, MeOH). ¹H NMR (400 MHz, DMSO) δ 10.57 (s, 1H, exchangeable with D₂O), 8.94 (t, *J* = 6.1 Hz, 1H), 8.89 (t, *J* = 5.9 Hz, 1H), 8.76 (s, 3H), 7.56 – 7.53 (m, 1H), 7.42 – 7.35 (m, 1H), 7.35 – 7.29 (m, 2H), 7.29 – 7.25 (m, 5H), 7.24 – 7.18 (m, 2H),

4.45 – 4.36 (m, 2H), 4.33 (d, J = 6.0 Hz, 2H), 3.91 – 3.82 (m, 1H), 1.38 (d, J = 7.0 Hz, 3H). ESI-MS m/z: 421.08 (M + 1)$^+$. ESI-MS m/z: 421.08 (M + 1)$^+$

Synthesis of ethyl 2-((3-nitrobenzyl)amino)acetate (113a) [167]**.**
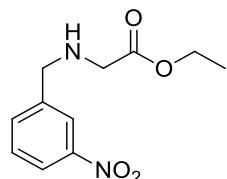
To a solution of glycine ethyl ester hydrochloride (5.00 g, 35.82 mmol) and TEA (3.6 g, 35.82 mmol) was added 3-nitrobenzaldehyde (3.60 g, 23.88 mmol). The mixture was stirred overnight at room temperature and then was added NaBH$_4$ (1.80 g, 47.76 mmol) after cooling at 0 °C. The reaction was sterred at 0 °C for additional 1 h and quenched with water (30 mL). The reaction was concentrated in vacuo and the residue was dissolved in satured solution of NaHCO$_3$ (150 mL). The aqueous layer was washed with diethyl ether (3x 50 mL), acidified to pH 1 with 6 N HCl and extracted with ethyl acetate (3 x 50 mL). The combined organic layers were washed with Brine, dried (Na$_2$SO$_4$), filtered and evaporated under vacuum to give the title compound as a yellow oil (4.62 g, 81 %). ^{1}H NMR (300 MHz, CDCl$_3$): δ 8.27 – 8.20 (m, 1H), 8.15 – 8.10 (m, 1H), 7.74 – 7.66 (m, 1H), 7.56 – 7.46 (m, 1H), 4.21 (q, J = 7.1 Hz, 2H), 3.92 (s, 2H), 3.42 (s, 2H), 1.80 (s, 1H, exchangeable with D$_2$O), 1.29 (t, J = 7.1 Hz, 3H). ESI-MS m/z: 239.10 (M + 1)$^+$

Synthesis of ethyl 2-((3-cyanobenzyl)amino)acetate (113b). [168]
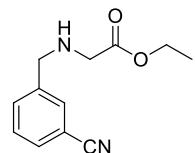
Compound **113b** was obtained as a yellow oil (3.44 g, 66%) from 3-cianobenzaldehyde (3.0 g, 22.8 mmol), according to the procedure used for

**113a.** ^{1}H NMR (300 MHz, CDCl$_3$): δ 7.70 – 7.64 (m, 1H), 7.62 – 7.53 (m, 2H), 7.47 – 7.39 (m, 1H), 4.20 (q, *J* = 7.1 Hz, 2H), 3.84 (s, 2H), 3.39 (s, 2H), 1.75 (s, 1H, scambiato con D$_2$O), 1.28 (t, *J* = 7.1 Hz, 3H). ESI-MS m/z:219.07 (M + 1)$^+$.

Synthesis of ethyl 2-(N-(3-nitrobenzyl)formamido)acetate (114a).
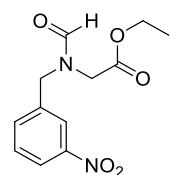
To a mixture of **113a** (4.00 g, 16.97 mmol) and formic acid (9.06 g, 200 mmol) was added dropwise over about 15 minutes acetic anhydride (10.60 g, 100 mmol) while cooling and stirring. The mixture was continuosly stirred for one hour at 0 °C and two hours at room temperature. The reaction mixture was concentrated under vacuum and the residue was dissolved in ethyl acetate (200 mL). The organic layer was washed with a saturated aqueous solution of NaHCO$_3$ (3 x 50 mL), HCl 1 N (3 x 50 mL) and Brine. The organic phase was dried (Na$_2$SO$_4$), filtered and concentrated in vacuo to obtain the title compound (**11a**) as a yellow oil (3.57 g, 80%). ^{1}H NMR (300 MHz, CDCl$_3$): δ 8.39 (s, 1H, exchangeable with D$_2$O), 8.26 – 8.05 (m, 2H), 7.64 – 7.45 (m, 2H), 4.71, 4.65 (2s, 2H), 4.18 (q, *J* = 7.1 Hz, 2H), 3.97, 3.91 (2s, 2H), 1.26 (t, *J* = 7.1 Hz, 3H). ESI-MS m/z: 239.10 (M + 1)$^+$.

Synthesis of ethyl 2-(N-(3-cyanobenzyl)formamido)acetate (114b).
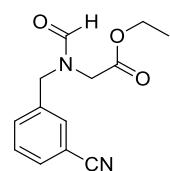
Compound **114b** was obtained as a yellow oil (3.01 g, 91%) from compound **113b** (3.00 g, 13.7 mmol), according to the procedure used for **114a.** ^{1}H NMR

(300 MHz, CDCl$_3$) δ 8.36, 8.21 (2s, 1H), 7.69 – 7.57 (m, 1H), 7.57 – 7.40 (m, 3H), 4.64, 4.58 (2s, 2H), 4.18 (q, J = 7.1 Hz, 2H), 3.95, 3.87 (2s, 2H), 1.26 (t, J = 7.1 Hz, 3H). ESI-MS m/z: 267.91 (M + 1)$^+$.

Synthesis of 2-(N-(3-nitrobenzyl)formamido)acetic acid (115a).
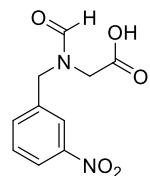
An aqueous solution of sodium hydroxide 2N (16mL) was added to a solution of **11a** (3.00 g, 11.27 mmol) in ethanol (12 mL) . The reaction was stirred at room temperature for 30 minutes,and water (20 ml) was added. The aqueous phase was washed with chloroform (3 x 50 mL) and acidified to pH 1 with 6 N HCl and extracted with ethyl acetate (3 x 50 mL). The combined organic layers were washed with Brine, dried (Na$_2$SO$_4$), filtered and evaporated under vacuum to give the title compound as a yellow solid (1.88 g, 70%); mp 138 – 140 °C. ^{1}H NMR (300 MHz, DMSO): δ 12.82 (s, 1H, exchangeable with D$_2$O), 8.42 (s, 1H, exchangeable with D$_2$O), 8.24 – 8.17 (m, 1H), 8.17 – 8.09 (m, 1H), 7.82 – 7.76 (m, 1H), 7.75 – 7.58 (m, 1H), 4.68, 4.58 (2s, 2H), 4.09, 3.83 (2s, 2H). ESI-MS m/z: 444.50 (M x 2)$^-$

Synthesis of 2-(N-(3-cyanobenzyl)formamido)acetic acid (115b).
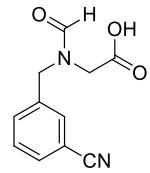
Compound **12b** was obtained as a white solid (2.18 g, 82%) from compound **11b** (3.0 g, 12.18 mmol), according to the procedure used for **12a,** m. p. 121 – 123 °C. ^{1}H NMR (300 MHz, DMSO):δ 12.82 (s, 1H, exchangeable with D$_2$O),

8.37, 8.17 (s, 1H, exchangeable with D$_2$O), 7.86 – 7.63 (m, 2H), 7.63 – 7.46 (m, 2H), 4.58, 5.50 (2s, 2H), 4.06, 3.80 (s, 2H). ESI-MS m/z: 239.03 (M + 1)$^+$.

Synthesis of ethyl 1-(3-nitrobenzyl)-1H-pyrrole-3-carboxylate (116a).
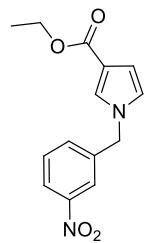

A solution of **115a** (1.40 g, 5.88 mmol) and ethyl propiolate (1.73 g, 17.64 mmol) in acetic acid (7 mL) was stirred at 130 °C for 5 h. The reaction was concentrated under vacuum. The residue was dissolved with ethyl acetate (150 mL) and washed with satured aqueous solution of NaHCO$_3$ (3 x 50 mL), Brine, dried (Na$_2$SO$_4$), filtered and evaporated under vacuum. The residue was purified by silica gel chromatography (ethyl acetate : petroleum ether, 3:7) to give the title compound as a yellow solid (0.790 g, 73 %); m. p. 76 – 78 °C. ^{1}H NMR (300 MHz, CDCl$_3$): δ 8.18 (d, *J* = 8.3 Hz, 1H), 8.04 – 8.01 (m, 1H), 7.58 – 7.49 (m, 1H), 7.43 – 7.37 (m, 1H), 7.36 – 7.30 (m, 1H), 6.73 – 6.65 (m, 2H), 5.17 (s, 2H), 4.27 (q, *J* = 7.1 Hz, 2H), 1.33 (t, *J* = 7.1 Hz, 3H). ESI-MS m/z: 255.62 (M+ 1)$^+$

**Synthesis of ethyl 1-(3-cyanobenzyl)-1H-pyrrole-3-carboxylate (116b).**
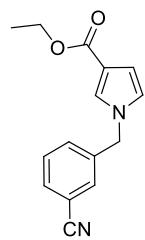
Compound **116b** was obtained as a white solid (1.40 g, 60%) from compound **115b** (2.0 g, 9.16 mmol), according to the procedure used for **116a**, m. p. 87 – 89 °C. ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.63 – 7.57 (m, 1H), 7.50 – 7.42 (m, 1H), 7.39 (s, 1H), 7.35 – 7.28 (m, 2H), 6.70 – 6.57 (m, 1H), 5.10 (s, 2H), 4.27 (q, J = 7.0 Hz, 2H), 1.33 (t, J = 7.0 Hz, 3H). ESI-MS m/z: 275.07 (M + 1)$^+$

Synthesis of ethyl 1-(3-nitrobenzyl)-5-(2,2,2-trichloroacetyl)-1H-pyrrole-3-carboxylate (117a).
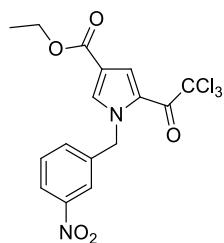
Compound **117a** was obtained as a yellow solid (0.34 g, 88%) from compound **116a** (0.300 g, 1.03 mmol), according to the procedure used for **104a**, m. p. 84 – 86 °C. ^{1}H NMR (300 MHz, CDCl$_3$)δ 8.20 – 8.13 (m, 1H), 7.98 (d, J = 1.1 Hz, 1H), 7.97 – 7.94 (m, 1H), 7.69 (d, J = 1.1 Hz, 1H), 7.57 – 7.49 (m, 1H), 7.46 – 7.40 (m, 1H), 5.66 (s, 2H), 4.34 (q, J = 7.1 Hz, 2H), 1.37 (t, J = 7.1 Hz, 3H). ESI-MS m/z: 418.07 (M + 1)$^+$

Synthesis of ethyl 1-(3-cyanobenzyl)-5-(2,2,2-trichloroacetyl)-1H-pyrrole-3-carboxylate (117b).
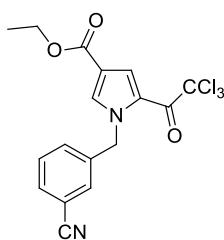
Compound **117b** was obtained as a white solid (1.27 g, 81%) from compound **116b** (1.0 g, 3.90 mmol), according to the procedure used for **104a**, m. p. 109

– 110 °C. ^{1}H NMR (300 MHz, CDCl$_3$): δ 7.97 (d, *J* = 1.1 Hz, 1H), 7.66 (d, *J* = 1.1 Hz, 1H), 7.63 – 7.56 (m, 1H), 7.50 – 7.42 (m, 1H), 7.38 – 7.31 (m, 2H), 5.63 (s, 2H), 4.34 (q, *J* = 7.1 Hz, 2H), 1.37 (t, *J* = 7.1 Hz, 3H).ESI-MS m/z: 399.01 (M + 1)$^+$.

Synthesis of   ethyl 5-(benzylcarbamoyl)-1-(3-nitrobenzyl)-1H-pyrrole-3-carboxylate (118a).
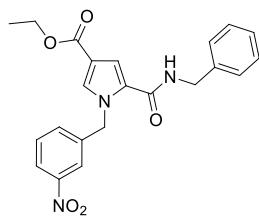
Compound **118a** was obtained as a yellow solid (0.43 g, 88%) from compound **117a** (0.500 g, 1.19 mmol), according to the procedure used for **108a**, m. p. 136 – 137 °C. ^{1}H NMR (300 MHz, CDCl$_3$): δ 8.20 – 8.09 (m, 1H), 8.05 – 7.95 (m, 1H), 7.54 – 7.41 (m, 3H), 7.37 – 7.16 (m, 5H), 7.09 – 7.04 (m, 1H), 6.39 – 6.21 (m, 1H, exchangeable with D$_2$O), 5.71 (s, 2H), 4.50 (d, *J* = 4.6 Hz, 2H), 4.26 (q, *J* = 6.9 Hz, 2H), 1.32 (t, *J* = 6.9 Hz, 3H). ESI-MS m/z: 408.11 (M + 1)$^+$.

Synthesis of   ethyl 5-(benzylcarbamoyl)-1-(3-cyanobenzyl)-1H-pyrrole-3-carboxylate (118b).
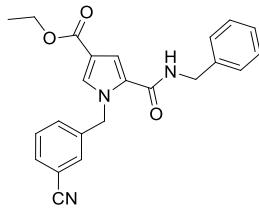
Compound **118b** was obtained as a yellow solid (0.51 g, 45%) from compound **117b** (1.15 g, 2.9 mmol), according to the procedure used for **108a**, m. p. 150 – 153 °C. ^{1}H NMR (300 MHz, CDCl$_3$): δ 7.62 – 7.53 (m, 1H), 7.49 – 7.37 (m, 4H), 7.37 – 7.19 (m, 5H), 7.08 – 7.01 (m, 1H), 6.23 (t, *J* = 5.2 Hz , 1H,

exchangeable with D₂O), 5.65 (s, 2H), 4.51 (d, J = 5.2 Hz, 2H), 4.27 (q, J = 7.0 Hz, 2H), 1.33 (t, J = 7.0 Hz, 3H). ESI-MS m/z: 388.15 (M + 1)$^+$

Synthesis of ethyl 1-(3-aminobenzyl)-5-(2,2,2-trichloroacetyl)-1H-pyrrole-3-carboxylate (119a).
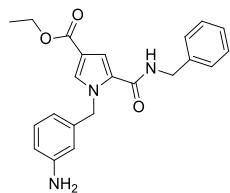
Compound **118a** was obtained as a yellow solid (0.132 g, 73%) from compound **117a** (0.200 g, 0.50 mmol), according to the procedure used for **109a**. ^{1}H NMR (300 MHz, DMSO): δ 8.78 (t, J = 5.7 Hz, 1H, exchangeable with deuterium oxide), 7.71 – 7.55 (m, 1H), 7.38 – 7.12 (m, 6H), 7.00 – 6.81 (m, 1H), 6.48 – 6.38 (m, 1H), 6.35 – 6.29 (m, 1H), 6.28 – 6.22 (m, 1H), 5.48 (s, 2H), 5.03 (bs, 2H, exchangeable with deuterium oxide), 4.36 (d, J = 5.7 Hz, 2H), 4.19 (d, J = 7.0 Hz, 2H), 1.25 (t, J = 7.0 Hz, 3H). ESI-MS m/z: 744.88 (M x 2)$^+$.

Synthesis of ethyl 1-(3-(aminomethyl)benzyl)-5-(benzylcarbamoyl)-1H-pyrrole-3-carboxylate (119b).
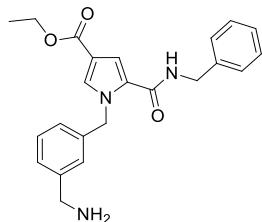
Compound **119b** was obtained as a slurry (0.270 g, 90%) from compound **118b** (0.300 g, 0.77 mmol), according to the procedure used for **109b**. ^{1}H

NMR (300 MHz, DMSO): δ 8.83 (t, *J* = 5.1 Hz, 1H, exchangeable with deuterium oxide), 8.35 – 8.14 (m, 2H), 7.85 – 7.64 (m, 1H), 7.46 – 7.15 (m, 9H), 7.16 – 7.01 (m, 1H), 5.66 (s, 2H, exchangeable with deuterium oxide), 4.35 (d, *J* = 5.1 Hz, 2H), 4.20 (q, *J* = 6.5 Hz, 2H), 4.02 – 3.90 (m, 2H), 1.24 (t, *J* = 6.5 Hz, 3H). ESI-MS m/z: 329.10 (M + 1)[+].

Synthesis of (S)-ethyl 5-(benzylcarbamoyl)-1-(3-(2-((tert-butoxycarbonyl)amino)propanamido)benzyl)-1H-pyrrole-3-carboxylate (120a).
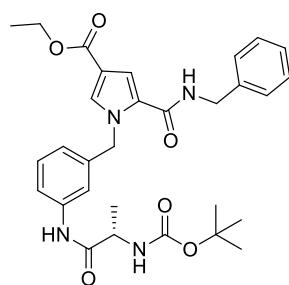

Compound **120a** was obtained as a white solid (0.121, 82%) from compound **119a** (0.121 g, 0.27 mmol), according to the procedure used for **110a**, mp 139 – 142 °C. [α]$_D$ -31.0 (c 0.3, MeOH).^{1}H NMR (300 MHz, DMSO):δ 9.89 (s, 1H, exchangeable with deuterium oxide), 8.78 (t, *J* = 5.5 Hz, 1H, exchangeable with deuterium oxide), 7.74 (s, 1H), 7.62 – 7.51 (m, 1H), 7.31 – 7.26 (m, 3H), 7.25 – 7.23 (m, 1H), 7.23 – 7.20 (m, 1H), 7.20 – 7.14 (m, 3H), 7.02 (d, *J* = 6.8 Hz, 1H, exchangeable with deuterium oxide), 6.80 – 6.73 (m, 1H), 5.61 (s, 2H), 4.33 (d, *J* = 5.5 Hz, 2H), 4.19 (q, *J* = 7.0 Hz, 2H), 4.12 – 4.01 (m, 1H), 1.36 (s, 9H), 1.30 – 1.16 (m, 6H). ESI-MS m/z: 548.84 (M + 1)[+]

Synthesis of (R)-ethyl 5-(benzylcarbamoyl)-1-(3-(2-((tert-butoxycarbonyl)amino)propanamido)benzyl)-1H-pyrrole-3-carboxylate (ent-120a).


Compound *ent-***120a** was obtained as a white solid (0.121, 82%) from compound **119a** (0.121 g, 0.27 mmol), according to the procedure used for *ent-***110a**, mp 136 – 138 °C. [α]$_D$ +27.0 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO) δ 9.89 (s, 1H, exchangeable with deuterium oxide), 8.78 (t, *J* = 5.5 Hz, 1H, exchangeable with deuterium oxide), 7.74 (s, 1H), 7.62 – 7.51 (m, 1H), 7.31 – 7.26 (m, 3H), 7.25 – 7.23 (m, 1H), 7.23 – 7.20 (m, 1H), 7.20 – 7.14 (m, 3H), 7.02 (d, *J* = 6.8 Hz, 1H, exchangeable with deuterium oxide), 6.80 – 6.73 (m, 1H), 5.61 (s, 2H), 4.33 (d, *J* = 5.5 Hz, 2H), 4.19 (q, *J* = 7.0 Hz, 2H), 4.12 – 4.01 (m, 1H), 1.36 (s, 9H), 1.30 – 1.16 (m, 6H). ESI-MS m/z: 548.84 (M + 1)$^+$.

Synthesis of (S)-ethyl 5-(benzylcarbamoyl)-1-(3-((2-((tert-butoxycarbonyl)amino)propanamido)methyl)benzyl)-1H-pyrrole-3-carboxylate (120b).
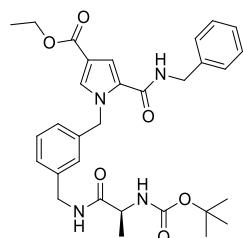
Compound **120b** was obtained as a white solid (0.207 g, 52%) from compound **119b** (0.400 g, 0.71 mmol), according to the procedure used for **110a**, 141 – 144 °C. [α]$_D$ - 7.4 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO): δ 8.81 (t, J = 5.0 Hz, 1H, exchangeable with deuterum oxide), 8.26 (t, J = 5.6 Hz, 1H, exchangeable with deuterum oxide), 7.72 (s, 1H), 7.35 – 7.26 (m, 3H), 7.27 – 7.18 (m, 4H), 7.17 – 7.10 (m, 1H), 7.11 - 7.07 (m, 1H), 7.01 – 6.94 (m, 1H), 6.88 (d, J = 5.6 Hz, 1H, exchangeable with deuterum oxide), 5.61 (s, 2H), 4.36 (d, J = 5.0 Hz, 2H), 4.27 – 4.13 (m, 4H), 3.96 (s, 1H), 1.39 (s, 9H), 1.25 (t, J = 6.6 Hz, 3H), 1.17 (d, J = 6.6 Hz, 3H). ESI-MS m/z: 585.16 (M + Na)$^+$

Synthesis of (R)-ethyl 5-(benzylcarbamoyl)-1-(3-((2-((tert-butoxycarbonyl)amino)propanamido)methyl)benzyl)-1H-pyrrole-3-carboxylate (ent-120b).


Compound ***ent*-120b** was obtained as a white solid (0.207 g, 52%) from compound **119b** (0.400 g, 0.71 mmol), according to the procedure used for ***ent*-110a**, mp 140 – 143 °C. [α]$_D$ + 9.8 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO): δ 8.81 (t, J = 5.0 Hz, 1H, exchangeable with deuterum oxide), 8.26 (t, J = 5.6 Hz, 1H, exchangeable with deuterum oxide), 7.72 (s, 1H), 7.35 – 7.26 (m, 3H), 7.27 – 7.18 (m, 4H), 7.17 – 7.10 (m, 1H), 7.11 – 7.07 (m, 1H), 7.01 – 6.94 (m, 1H), 6.88 (d, J = 5.6 Hz, 1H, exchangeable with deuterum oxide), 5.61 (s, 2H), 4.36 (d, J = 5.0 Hz, 2H), 4.27 – 4.13 (m, 4H), 3.96 (s, 1H), 1.39 (s, 9H), 1.25 (t, J = 6.6 Hz, 3H), 1.17 (d, J = 6.6 Hz, 3H). ESI-MS m/z: 585.16 (M + Na)$^+$.

Synthesis of (S)-5-(benzylcarbamoyl)-1-(3-(2-((tert- butoxycarbonyl) amino) propanamido)benzyl)-1H-pyrrole-3-carboxylic acid (120c).


Compound **120c** was obtained as a white solid (0.160 g, 85%) from compound **120a** (0.200 g, 0.36 mmol), according to the procedure used for **110e**, mp 232 – 234 °C. [α]$_D$ -16.3 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO) δ**:** δ 12.05 (bs, 1H, exchangeable with deuterum oxide), 9.99 – 9.80 (m, 1H, exchangeable with deuterium oxide), 8.75 (t, *J* = 5.8 Hz, 1H, exchangeable with deuterium oxide), 7.69 – 7.63 (m, 1H), 7.61 – 7.54 (m, 1H), 7.34 – 7.28 (m, 2H), 7.28 – 7.20 (m, 4H), 7.20 – 7.14 (m, 2H), 7.02 (d, *J* = 7.2 Hz, 1H, exchangeable with deuterium oxide), 6.80 – 6.74 (m, 1H), 5.59 (s, 2H), 4.34 (d, *J* = 5.8 Hz, 2H), 4.15 – 4.04 (m, 1H), 1.37 (s, 9H), 1.22 (d, *J* = 7.0 Hz, 3H). ESI-MS m/z: 1040.90 (M x 2)$^+$.

Synthesis of (R)-5-(benzylcarbamoyl)-1-(3-(2-((tert-butoxycarbonyl)amino)propanamido)benzyl)-1H-pyrrole-3-carboxylic acid (ent-120c).
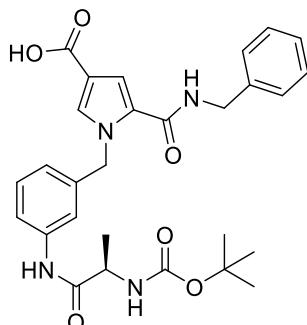
Compound ***ent*-120c** was obtained as a white solid (0.160 g, 85%) from compound ***ent*-119a** (0.200 g, 0.36 mmol), according to the procedure used for **110e,** mp 163 – 166 °C. $[\alpha]_D$ +14.2 (c 0.3, MeOH).^{1}H NMR (300 MHz, DMSO): δ 12.05 (bs, 1H, exchangeable with deuterum oxide), 9.99 – 9.80 (m, 1H, exchangeable with deuterium oxide), 8.75 (t, *J* = 5.8 Hz, 1H, exchangeable with deuterium oxide), 7.69 – 7.63 (m, 1H), 7.61 – 7.54 (m, 1H), 7.34 – 7.28 (m, 2H), 7.28 – 7.20 (m, 4H), 7.20 – 7.14 (m, 2H), 7.02 (d, *J* = 7.2 Hz, 1H, exchangeable with deuterium oxide), 6.80 – 6.74 (m, 1H), 5.59 (s, 2H), 4.34 (d, *J* = 5.8 Hz, 2H), 4.15 – 4.04 (m, 1H), 1.37 (s, 9H), 1.22 (d, *J* = 7.0 Hz, 3H). ESI-MS m/z: 1040.90 (M x 2)$^+$.

Synthesis of (S)-5-(benzylcarbamoyl)-1-(3-((2-((tert-butoxycarbonyl)amino)propanamido)methyl)benzyl)-1H-pyrrole-3-carboxylic acid (120d).

Compound **120d** was obtained as a white solid (0.160 g, 85%) from compound **120b** (0.200 g, 0.36 mmol), according to the procedure used for **110e**, mp 181 – 184 °C. $[\alpha]_D$ -9.7 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO): δ 12.06 (bs, 1H, exchangeable with deuterium oxide), 8.77 (t, *J* = 5.5 Hz, 1H, exchangeable with deuterium oxide), 8.27 (t, *J* = 5.3 Hz, 1H, exchangeable with deuterium oxide), 7.68 – 7.59 (m, 1H), 7.35 – 7.18 (m, 7H), 7.17 – 7.11 (m, 1H), 7.10 – 7.05 (m, 1H), 7.01 – 6.92 (m, 1H), 6.88 (d, *J* = 7.0 Hz, 1H, exchangeable with deuterium oxide), 5.60 (s, 2H), 4.36 (d, *J* = 5.5 Hz, 2H),

4.26 – 4.13 (m, 2H), 4.02 – 3.92 (m, 1H), 1.37 (s, 9H), 1.17 (d, J = 6.9 Hz, 3H). ESI-MS m/z: 534.87 (M + 1)+

Synthesis of (R)-5-(benzylcarbamoyl)-1-(3-((2-((tert-butoxycarbonyl)amino)propanamido)methyl)benzyl)-1H-pyrrole-3-carboxylic acid (*ent*-120d).

Compound ***ent*-120d** was obtained as a white solid (0.160 g, 85%) from compound ***ent*-120b** (0.200 g, 0.36 mmol), according to the procedure used for **110e**, 176 – 179 °C. [α]$_D$ + 12.9 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO) δ : δ 12.06 (bs, 1H, exchangeable with deuterium oxide), 8.77 (t, *J* = 5.5 Hz, 1H, exchangeable with deuterium oxide), 8.27 (t, *J* = 5.3 Hz, 1H, exchangeable with deuterium oxide), 7.68 – 7.59 (m, 1H), 7.35 – 7.18 (m, 7H), 7.17 – 7.11 (m, 1H), 7.10 – 7.05 (m, 1H), 7.01 – 6.92 (m, 1H), 6.88 (d, *J* = 7.0 Hz, 1H, exchangeable with deuterium oxide), 5.60 (s, 2H), 4.36 (d, *J* = 5.5 Hz, 2H), 4.26 – 4.13 (m, 2H), 4.02 – 3.92 (m, 1H), 1.37 (s, 9H), 1.17 (d, *J* = 6.9 Hz, 3H). ESI-MS m/z: 534.87 (M + 1)+

Synthesis of (S)-1-((3-((2-(benzylcarbamoyl)-4-(ethoxycarbonyl)-1H-pyrrol-1-yl)methyl)phenyl)amino)-1-oxopropan-2-aminium 2,2,2-trifluoroacetate (121a).
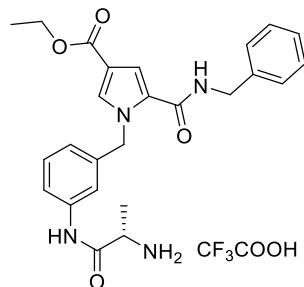

Compound **121a** was obtained as a white solid (0.051 g, 99%) from compound **120a** (0.050 g, 0.09 mmol), according to the procedure used for **110a**, mp 232 – 234 °C. [α]$_D$ + 5.3 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO): δ 10.34 (s, 1H, exchangeable with deuterium oxide), 8.77 (t, *J* = 5.4 Hz, 1H, exchangeable with deuterium oxide), 8.19 – 8.06 (m, 3H, exchangeable with deuterium oxide), 7.77 (s, 1H), 7.60 – 7.52 (m, 1H), 7.34 – 7.31 (m, 1H), 7.31 – 7.24 (m, 4H), 7.24 – 7.22 (m, 1H), 7.21 – 7.14 (m, 2H), 6.92 – 6.84 (m, 1H), 5.63 (s, 2H), 4.34 (d, *J* = 5.4 Hz, 2H), 4.20 (q, *J* = 7.1 Hz, 2H), 3.97 – 3.90 (m, 1H), 1.41 (d, *J* = 6.8 Hz, 3H), 1.26 (t, *J* = 7.1 Hz, 3H). ESI-MS m/z: 449.03 (M + 1)$^+$.

Synthesis of (R)-1-((3-((2-(benzylcarbamoyl)-4-(ethoxycarbonyl)-1H-pyrrol-1-yl)methyl)phenyl)amino)-1-oxopropan-2-aminium 2,2,2-trifluoroacetate (ent-121a).

Compound *ent*-**121a** was obtained as a white solid (0.051 g, 99%) from compound *ent*-**120a** (0.050 g, 0.09 mmol), according to the procedure used for **110a**, mp 228 – 230. [α]$_D$ - 3.0 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO): δ 10.34 (s, 1H, exchangeable with deuterium oxide), 8.77 (t, *J* = 5.4 Hz, 1H, exchangeable with deuterium oxide), 8.19 – 8.06 (m, 3H, exchangeable with deuterium oxide), 7.77 (s, 1H), 7.60 – 7.52 (m, 1H), 7.34 – 7.31 (m, 1H), 7.31 – 7.24 (m, 4H), 7.24 – 7.22 (m, 1H), 7.21 – 7.14 (m, 2H), 6.92 – 6.84 (m, 1H), 5.63 (s, 2H), 4.34 (d, *J* = 5.4 Hz, 2H), 4.20 (q, *J* = 7.1 Hz, 2H), 3.97 – 3.90 (m, 1H), 1.41 (d, *J* = 6.8 Hz, 3H), 1.26 (t, *J* = 7.1 Hz, 3H). ESI-MS m/z: 449.03 (M + 1)$^+$.

Synthesis of (S)-1-((3-((2-(benzylcarbamoyl)-4-(ethoxycarbonyl)-1H-pyrrol-1-yl)methyl)benzyl)amino)-1-oxopropan-2-aminium 2,2,2-trifluoroacetate (121b).

Compound **121b** was obtained as a white solid (0.057 g, 99%) from compound **120b** (0.050 g, 0.10 mmol), according to the procedure used for **110**, 155 – 158 °C. [α]$_D$ +7.3 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO): δ 8.88 – 8.75 (m, 2H, exchangeable with deuterium oxide), 8.10 – 7.96 (m, 3H, exchangeable with deuterium oxide), 7.73 (s, 1H), 7.35 – 7.25 (m, 4H), 7.27 – 7.14 (m, 4H), 7.12 – 7.07 (m, 1H), 7.08 – 6.96 (m, 1H), 5.63 (s, 2H), 4.35 (d, *J* = 5.5 Hz, 2H), 4.29 – 4.14 (m, 4H), 3.86 – 3.77 (m, 1H), 1.35 (d, *J* = 6.5 Hz, 3H), 1.26 (t, *J* = 6.9 Hz, 3H). ESI-MS m/z: 462.54 (M + 1)$^+$.

Synthesis of (R)-1-((3-((2-(benzylcarbamoyl)-4-(ethoxycarbonyl)-1H-pyrrol-1-yl)methyl)benzyl)amino)-1-oxopropan-2-aminium 2,2,2-trifluoroacetate (ent-121b).
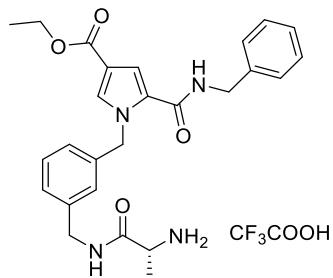

Compound *ent*-**121b** was obtained as a white solid (0.057 g, 99%) from compound *ent*-**120b** (0.050 g, 0.10 mmol), according to the procedure used for **110a**, 154 – 157 °C. $[\alpha]_D$ - 6.3 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO): δ 8.88 – 8.75 (m, 2H, exchangeable with deuterium oxide), 8.10 – 7.96 (m, 3H, exchangeable with deuterium oxide), 7.73 (s, 1H), 7.35 – 7.25 (m, 4H), 7.27 – 7.14 (m, 4H), 7.12 – 7.07 (m, 1H), 7.08 – 6.96 (m, 1H), 5.63 (s, 2H), 4.35 (d, J = 5.5 Hz, 2H), 4.29 – 4.14 (m, 4H), 3.86 – 3.77 (m, 1H), 1.35 (d, J = 6.5 Hz, 3H), 1.26 (t, J = 6.9 Hz, 3H). ESI-MS m/z: 462.54 (M + 1)$^+$.

Synthesis of (S)-1-((3-((2-(benzylcarbamoyl)-4-carboxy-1H-pyrrol-1-yl)methyl)phenyl)amino)-1-oxopropan-2-aminium 2,2,2-trifluoroacetate (121c).
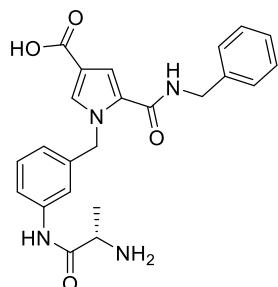
Compound **121c** was obtained as a white solid (0.071 g, 99%) from compound **120c** (0.070 g, 0.13 mmol), according to the procedure used for

110a, 215 – 217 °C. $[\alpha]_D$ + 3.0 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO): δ 12.10 (s, 1H, exchangeable with deuterium oxide), 10.34 (s, 1H, exchangeable with deuterium oxide), 8.74 (t, J = 5.5 Hz, 1H, exchangeable with deuterium oxide), 8.19 – 8.04 (m, 3H, exchangeable with deuterium oxide), 7.63 – 7.65 (m, 1H), 7.60 – 7.50 (m, 1H), 7.33 – 7.24 (m, 5H), 7.24 – 7.22 (m, 1H), 7.21 – 7.15 (m, 2H), 6.95 – 6.84 (m, 1H), 5.62 (s, 2H), 4.34 (d, J = 5.5 Hz, 2H), 4.01 – 3.87 (m, 1H), 1.41 (d, J = 6.9 Hz, 3H). ESI-MS m/z: 421.05 (M + 1)$^+$

Synthesis of (R)-1-((3-((2-(benzylcarbamoyl)-4-carboxy-1H-pyrrol-1-yl)methyl)phenyl)amino)-1-oxopropan-2-aminium 2,2,2-trifluoroacetate (ent-121c).
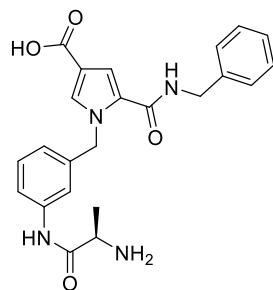
Compound **ent-121c** was obtained as a white solid (0.071 g, 99%) from compound **ent-120c** (0.070 g, 0.13 mmol), according to the procedure used for **110a**, mp 210 – 212 °C. $[\alpha]_D$ -1.8 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO): δ 12.10 (s, 1H, exchangeable with deuterium oxide), 10.34 (s, 1H, exchangeable with deuterium oxide), 8.74 (t, *J* = 5.5 Hz, 1H, exchangeable with deuterium oxide), 8.19 – 8.04 (m, 3H, exchangeable with deuterium oxide), 7.63 – 7.65 (m, 1H), 7.60 – 7.50 (m, 1H), 7.33 – 7.24 (m, 5H), 7.24 – 7.22 (m, 1H), 7.21 – 7.15 (m, 2H), 6.95 – 6.84 (m, 1H), 5.62 (s, 2H), 4.34 (d, *J* = 5.5 Hz, 2H), 4.01 – 3.87 (m, 1H), 1.41 (d, *J* = 6.9 Hz, 3H). ESI-MS m/z: 421.04 (M + 1)$^+$

Synthesis of (S)-1-((3-((2-(benzylcarbamoyl)-4-carboxy-1H-pyrrol-1-yl)methyl)benzyl)amino)-1-oxopropan-2-aminium 2,2,2-trifluoroacetate (121d).
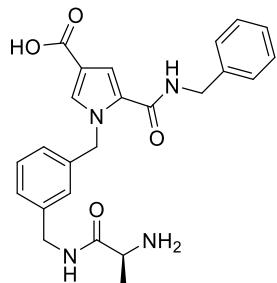

Compound **121d** was obtained as a white solid (0.070 g, 99%) from compound **120d** (0.070 g, 0.13 mmol), according to the procedure used for 110**a,** mp 204 – 206 °C. [α]$_D$ +2.8 (c 0.3, MeOH). ^{1}H NMR (300 MHz, DMSO):δ 11.95 (s, 1H, scambiato con D$_2$O), 8.87 – 8.59 (m, 2H, scambiato con D$_2$O),8.23 – 8.19 (m, 2H, scambiato con D$_2$O), 7.65 (s, 1H),7.33 – 7.14 (m, 8H),7.04 – 6.98 (m, 1H), 5.62 (s, 2H),4.36 (d, J = 5.7 Hz, 2H), 4.32 – 4.20 (m, 2H), 3.86 – 3.73 (m, 1H), 1.34 (d, J = 6.5 Hz, 3H). ESI-MS m/z:435.14 (M + 1)$^+$.

Synthesis of (S)-1-((3-((2-(benzylcarbamoyl)-4-carboxy-1H-pyrrol-1-yl)methyl)benzyl)amino)-1-oxopropan-2-aminium 2,2,2-trifluoroacetate (ent-121d).
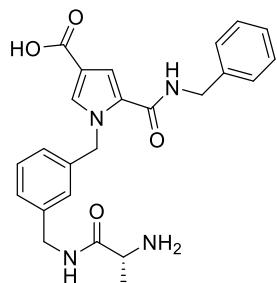
Compound ***ent*-121d** was obtained as a white solid (0.070 g, 99%) from compound ***ent*-120d** (0.070 g, 0.13 mmol), according to the procedure used for **110a**, 204 – 206 °C. [α]$_D$ -3.1 (c 0.3, MeOH). ^{1}H NMR (300 MHz,

DMSO):δ11.95 (s, 1H, scambiato con D_2O),8.87 – 8.59 (m, 2H, scambiato con D_2O),8.23 – 8.19 (m, 2H, scambiato con D_2O), 7.65 (s, 1H),7.33 – 7.14 (m, 8H),7.04 – 6.98 (m, 1H), 5.62 (s, 2H),4.36 (d, J = 5.7 Hz, 2H), 4.32 – 4.20 (m, 2H), 3.86 – 3.73 (m, 1H), 1.34 (d, J = 6.5 Hz, 3H). ESI-MS m/z:435.14 (M + 1)$^+$.

7.3. Bisubstrate inhibitors.

General procedures for peptides synthesis: 123, 126, 139 – 144.

Peptides **123**, **126**, **139 – 144** were prepared following standard Fmoc SPPS protocols. The required N-modified L-arginine building blocks were prepared from a common thiourea precursor as previously described.[161] Peptides were assembled on the 2-chlorotrityl resin working at 0.1 mmol scale, peptide couplings were performed using 4.0 equiv of protected Fmoc amino acid, 4.0 equiv of BOP reagent, and 8.0 equiv of DIPEA in a total volume of 10 mL of DMF at RT for 1 h. Alternatively, incorporation of the N-modified L-arginine building residues was performed using 2.0 equiv of the N-modified L-arginine building blocks, 2.0 equiv of BOP reagent, and 4.0 equiv of DIPEA in a total volume of 10 mL of DMF at RT overnight. Peptide couplings were verified using the Kaiser and bromophenol blue tests. Upon completion of SPPS, peptides were cleaved from the resin and deprotected using a mixture of 95:2.5:2.5 TFA/TIS/H_2O followed by Et_2O precipitation to yield the crude peptides. Each peptide was purified to homogeneity using RP-HPLC, employing a Prosphere C18 column (250 22 mm, 300 Å, 10 m) with a gradient of 5 to 95% acetonitrile (0.1% TFA) in 90 min at a flow rate of 11.5 mL min^{-1}. Peptide identity was confirmed by MALDI-MS analysis, in each case providing the expected mass.

**General procedures for removal of Aloc protecting group:124 and 127**

For removal of the Aloc protecting group of peptides **123** and **126**, the resin was first washed with CH_2Cl_2 (2 × 10 mL) under argon after which $PhSiH_3$ (0.74 mL, 6.0 mmol) in CH_2Cl_2 (4 mL) and Pd- $(PPh_3)_4$ (74 mg, 0.06 mmol) in CH_2Cl_2 (12 mL) were added and the mixture swirled under argon for 1 hour. The reaction mixture was drained and the procedure was repeated. To remove residual palladium catalyst, the resin was then washed with CH_2Cl_2 (5 × 10 mL), a 0.5% solution of diethyldithiocarbamic acid trihydrate sodium salt in DMF (5 × 10 mL), and DMF (5 × 10 mL). After Aloc removal, peptides 124 and 127 not cleaved from the resin were used for coupling with adenosine fragments.

Synthesis of 2-{(3a*R*,4*R*,6*R*,6a*R*)-6-(6-Amino-9*H*-purin-9-yl)-2,2-dimethylperhydrofuro [3,4-*d*][1,3]dioxol-4-yl}ethanenitrile (129)[163]
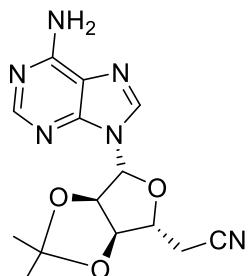
To a well stirred suspension of ((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methanol **(128)** (758 mg, 2.47 mmol) and PPh_3 (1.617 g, 6.166 mmol) in THF (12.4 cm^3), acetone cyanohydrins (0.57 cm^3, 6.2 mmol) was added. Within 5 min, DEAD (0.98 cm^3, 97%, 6.1 mmol) was added at 0 °C. The solution was stirred at 0 °C for 50 min and further at 20 °C for 16 h. Evaporation *in vacuo* and column chromatography (SiO$_2$; EtOAc– MeOH (96:4)) afforded **129** (698 mg, 89%) as a brown oil; δ**H**(200 MHz; CDCl$_3$) 1.42 (3 H, s), 1.64 (3 H, s), 2.84–3.08 (2 H,

m),4.51–4.60 (1 H, m), 5.16 (1 H, dd, *J* 6.2, 3.3), 5.50 (1 H, dd, *J* 6.2, 2.1), 5.71 (2 H, br s), 6.12 (1 H, d, *J* 2.1); 7.93 (1 H, s); 8.38 (s, 1 H).

Synthesis of (2*R*,3*S*,4*R*,5*R*)-2-(2-Aminoethyl)-5-(6-amino-9*H*-purin-9-yl)-tetrahydrofuran-3,4-diol (130)[163]
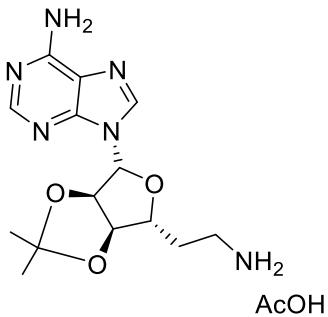
To a 0.1 M solution of **129** in THF (11 cm^3, 1.1 mmol), glacial acetic acid (20 cm^3) and PtO$_2$ (200 mg) were added. The suspension was hydrogenated (4 bar H$_2$) for 19 h. The catalyst was removed by filtration and the solvent evaporated *in vacuo*. The product was used in the next step without further purification.

Synthesis of Ethyl (*E*)-3-{(3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethylperhydrofuro[3,4-*d*][1,3]dioxol-4-yl}prop-2-enoate (132) [163]
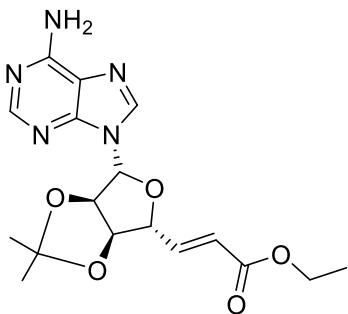
To 128 (12.3 g, 40 mmol) in Me$_2$SO (100 cm^3), Ph$_3$P=CHCO$_2$Et (42.8 g, 100 mmol) and *o*-iodoxybenzoic acid (IBX) (27.8 g,100 mmol) were added. The mixture was stirred at 20 °C for 72 h. Water (500 cm^3) was added and the

mixture extracted with EtOAc (2 × 500 cm^3). The combined organic phases were dried (Na$_2$SO$_4$) and concentrated to a residue which was purified by column chromatography (SiO$_2$; EtOAc–MeOH (96:4)) to give **132** (10.5 g, 70%) as colorless crystals, mp 102 °C (EtOH) (lit., 104–106 _C); δ**H**(300 MHz; CDCl$_3$) 1.22 (3 H, t, *J* 7.2), 1.40 (3 H, s), 1.63 (3 H, s), 4.11 (2 H, q, *J* 7.2), 4.79–4.83 (1 H, m), 5.14 (1 H, dd, *J* 6.2, 3.4), 5.56 (1 H, dd, *J* 6.2, 1.9), 5.81 (1 H, dd, *J* 15.9, 1.7), 6.13 (1 H, d, *J* 1.9), 6.96 (1 H, dd, *J* 15.9, 5.6), 7.87 (1 H, s), 8.33 (1 H, s).

Synthesis of (*E*)-3-{(3a*R*,4*R*,6*R*,6a*R*)-6-(6-Amino-9*H*-purin-9-yl)-2,2-dimethylperhydrofuro[3,4-*d*][1,3]dioxol-4-yl}prop-2-en-1-ol (133) [163].
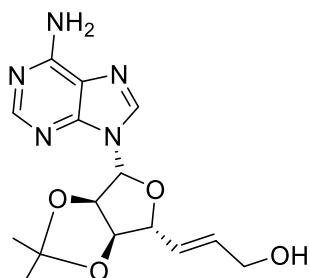
To **132** (375 mg, 1 mmol) in CH$_2$Cl$_2$ (3 cm$_3$), a 1 M solution of DIBAL-H in hexane (8 cm**3**, 8 mmol) was added dropwise. The mixture was stirred at - 78 °C for 2 h and then quenched with MeOH (5 cm**3**). A saturated aqueous solution of potassium sodium tartrate monohydrate (Rochelle salt, 50 cm**3**) was added, and the resulting suspension was stirred vigorously for 16 h at 20 °C, then extracted with EtOAc (3 × 50 cm**3**). The combined organic phases were dried (Na$_2$SO$_4$) and evaporated to dryness *in vacuo* to afford **4** (328 mg, 98%) as a colorless foam; $[\alpha]^{20}_{D}$ 17.8 (*c.* 1 in CHCl$_3$); δ**H**(300 MHz; CDCl**3**) 1.40 (3 H, s), 1.63 (3 H, s), 4.07 (2 H, s), 4.72 (1 H, dd, *J* 4.1, 2.5), 5.02 (1 H, dd, *J* 6.2, 4.1), 5.54 (1 H, dd, *J* 6.2, 2.2), 5.56 (2 H, s), 5.85–5.87 (2 H, m), 6.10 (1 H, d, *J* 2.2), 7.88 (1 H, s), 8.37 (1 H, s).

**Synthesis of 2-(*E*)-3-{(3a*R*,4*R*,6*R*,6a*R*)-6-(6-Amino-9*H*-purin-9-yl)-2,2-dimethylperhydrofuro [3,4-*d*][1,3]dioxol-4-l}prop-2-enyl-2,3-dihydro-1*H*-isoindole-1,3-dione (134)** [163]
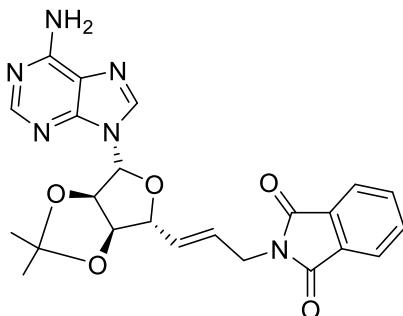
DEAD (0.234 cm³, 97%, 1.46 mmol) was added dropwise to a stirred suspension of **133**(490 mg, 1.46 mmol), phthalimide (215 mg, 1.46 mmol) and Ph₃P (383 mg, 1.46 mmol) in THF (7 cm³). After stirring for 1.5 h at 20 °C, a colorless solid started to precipitate. Stirring was continued for 1 h, after which the mixture was cooled to 0 °C and filtered. The residue was washed with Et₂O and dried *in vacuo* to give **134** (461 mg, 68%) as colorless crystals, mp 215 °C (MeOH); $[\alpha]^{20}_D$ 9.0 (*c.* 1 in CHCl₃); δ**H**(300 MHz;CDCl**3**) 1.36 (3 H, s), 1.59 (3 H, s), 4.19 (2 H, d, *J* 5.3), 4.69(1 H, dd, *J* 6.5, 3.1), 4.98 (1 H, dd, *J* 6.2, 3.1), 5.49 (1 H,dd, *J* 6.2, 2.0), 5.59 (2 H, br s), 5.73 (1 H, dt, *J* 15.6, 5.3), 5.84(1 H, dd, *J* 15.6, 6.5), 6.07 (1 H, d, *J* 2.0), 7.71–7.76 (2 H, m),7.80–7.86 (3 H, m), 8.24 (1 H, s).

Synthesis of 9-{(3a*R*,4*R*,6*R*,6a*R*)-6-[(*E*)-3-Aminoprop-1-enyl]-2,2-dimethylperhydrofuro[3,4-*d*][1,3]dioxol-4-yl}-9*H*-purin-6-amine (135) [163]
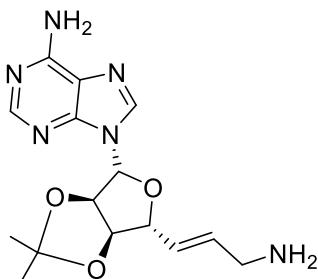
To phthalimide **134** (885 mg, 1.91 mmol), a solution of 33% MeNH$_2$ in EtOH (30 cm^3) was added and the mixture was stirred at 20 °C for 16 h. After evaporation *in vacuo*, the residue was dissolved in CHCl$_3$ (25 cm^3) and extracted with 10% aqueous AcOH (30 cm^3). The aqueous phase was washed with CHCl$_3$ (3 × 25 cm^3), adjusted to pH > 12 with 2 N NaOH and then extracted again with CHCl$_3$ (4 × 25 cm^3). The combined organic phases were dried (Na$_2$SO$_4$) and evaporated to give **7** (605 mg, 95%) as a colorless foam; $[\alpha]^{20}_D$ 30.8 (*c.* 1 in CHCl$_3$); δ**H** (300 MHz; CDCl$_3$) 1.39 (3 H, s), 1.62 (3 H, s), 3.22 (2 H, d, *J* 5.1), 4.68 (1 H, dd, *J* 7.3, 3.4), 4.99 (1 H, dd, *J* 6.2, 3.4), 5.53 (1 H, dd, *J* 6.2, 2.2), 5.71 (1 H, dd, *J* 15.6, 7.3), 5.83 (1 H, dt, *J* 15.6, 5.1), 5.87 (2 H, br s), 6.08 (1 H, d, *J* 2.2), 7.87 (1 H, s), 8.35 (1 H, s);

(2R,3S,4R,5R)-2-(3-Aminopropyl)-5-(6-amino-9H-purin-9-yl)-tetrahydrofuran-3,4-diol (136)
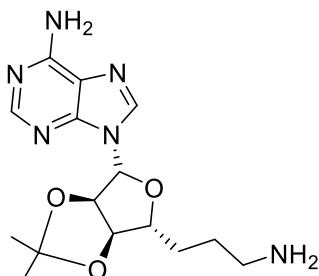
The allylic amine **135** (125 mg, 0.376 mmol) in EtOH (10 cm3) was hydrogenated (4 bar H$_2$) in the presence of 10% Pd/C (100 mg) for 18 h. The catalyst was removed by filtration and the solvent evaporated in vacuo. TFA–H2O (5:2, 30 cm3) was added and the mixture stirred for 1 h at 20 C, then evaporated to dryness to give **136** (94 mg, 84%) as a yellow foam; [α] 20 D δH(300 MHz; CD3OD) 1.75–1.87 (4 H, m), 2.97 (2 H, t, J 7.5), 4.00–4.05 (1 H, m), 4.21 (1 H, t, J 5.0), 4.78 (1 H, t, J 5.0), 5.97 (1 H, d, J 5.0), 8.19 (1 H, s), 8.24 (1 H, s).

General procedures for thiourea building blocks 131, 137 and 138. Example: synthesis of N-((2-((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)ethyl)carbamothioyl)-2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-sulfonamide 131
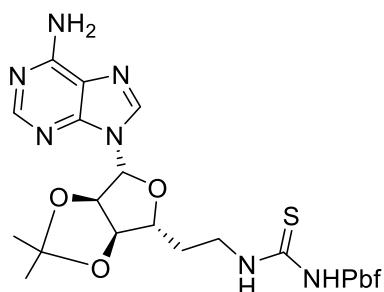
A solution of **130** (0.31 mmol, 0.100 g) in DCM (1 mL) was treated with a 0.1 M solution of Pbf-NCS in DCM (0.62 mmol, 6.2 mL) followed by addition of TEA (0.93 mmol, 0.2 mL). After stirring 1 h at ambient temperature TLC analysis revealed complete consumption of the PbfNCS with formation of the desired product. The mixture was concentrated and applied directly to a silica column, eluting with 9:1 DCM/MeOH. Following solvent removal the title compound was obtained as a yellow solid (0.097 g, 50%) and directly used in the next step given the high instability at room temperature. ESI-MS m/z: 632.5 (M + 1)$^+$.

N-(((E)-3-((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)allyl)carbamothioyl)-2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-sulfonamide 137
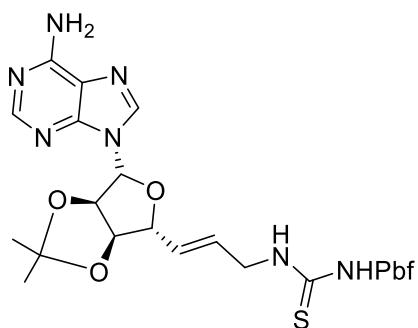
Yellow solid (0.151, 78%). ESI-MS m/z: 644.8 (M + 1)⁺.

N-((3-((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)propyl)carbamothioyl)-2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-sulfonamide 138
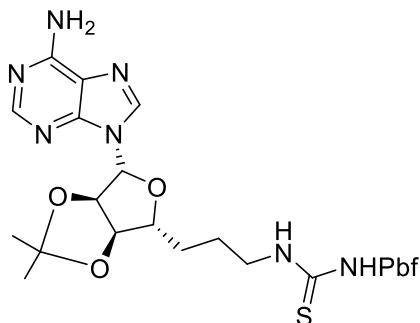

Yellow solid (0.155, 80%). ESI-MS m/z: 645.4 (M + 1)⁺.

General procedure for preparation of final compounds 80 – 85

The resin bearing peptide **124 or 127** (0.1 mmol) was first washed with DCM (2 × 10 mL) under N_2 after which a solution of thiourea precursor (0.3 mmol of 131, 137 or 138) and EDCI (0.3 mmol) in DCM (10 mL) was added and the mixture swirled under nitrogen for 2 hours. The reaction mixture was drained and the resin was washed with DCM (3 x 10 mL), DMF (3 x 10 mL) and again with DCM (3 x 10 mL). The protected peptide was deprotected and cleaved using a mixture of 95:2.5:2.5 TFA/TIS/H_2O followed by Et_2O precipitation to yield the crude peptides. Each peptide was purified to homogeneity using RP-HPLC, employing a Prosphere C18 column (250 22 mm, 300 Å, 10 m) with a gradient of 5 to 95% acetonitrile (0.1% TFA) in 90 min at a flow rate of 11.5 mL min⁻¹ . Peptide identity was confirmed by MALDI-MS analysis, in each case providing the expected mass.

80:
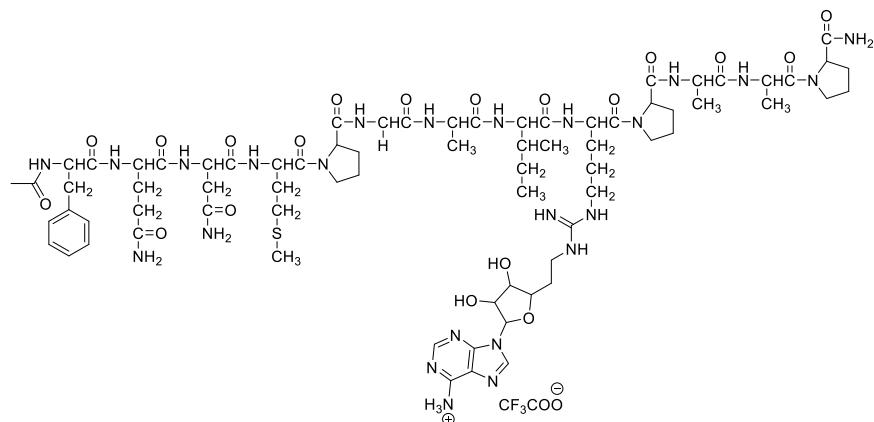
MALDI-MS: 1672.83

81:
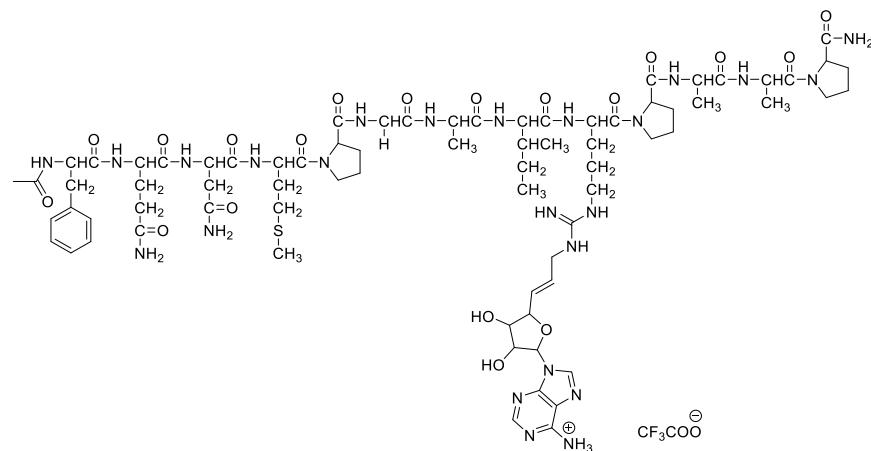
MALDI-MS: 1684.83

82:
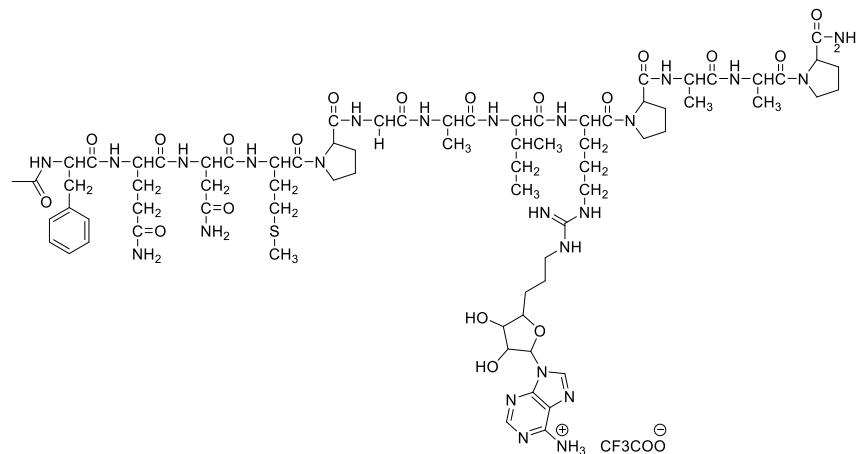
MALDI-MS: 1686.84

83:
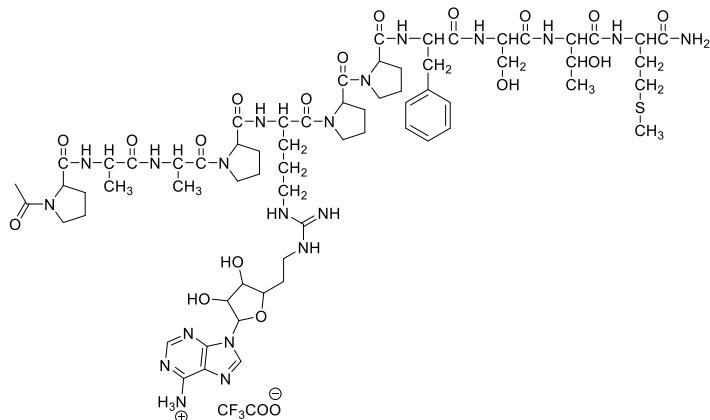
MALDI-MS: 1474.71

84:
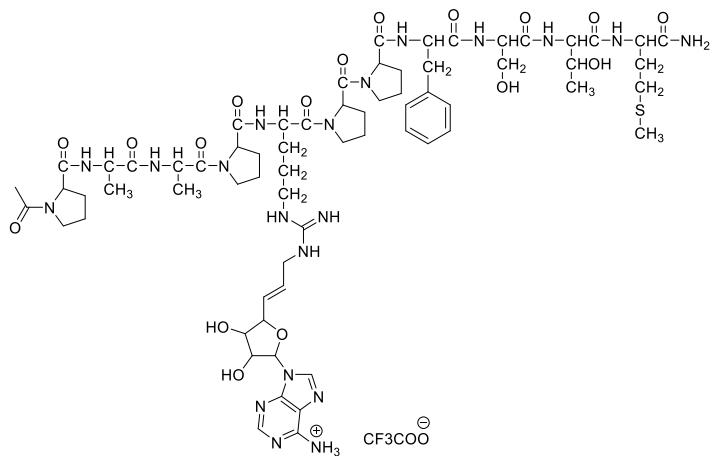
MALDI-MS: 1486.71

85:

MALDI-MS: 1488.73

209
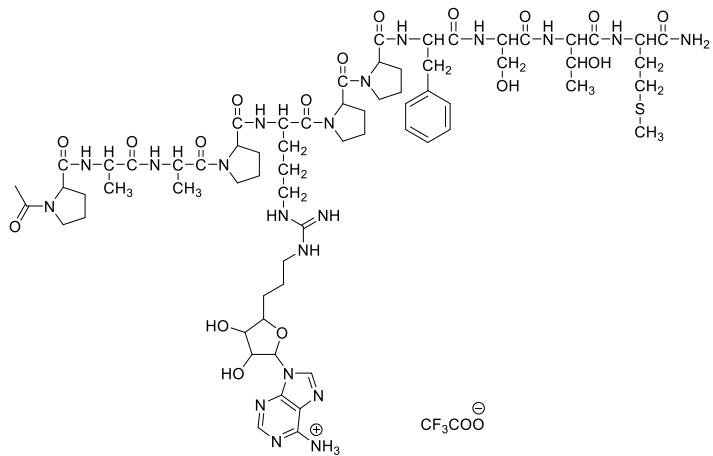
7.4. Indole derivatives

Synthesis of 2-(5-nitro-1H-indol-3-yl)-2-oxoacetyl chloride 150
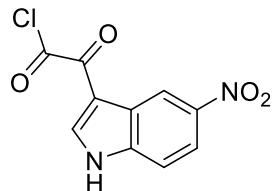
To 5-nitro indole (1.50 g, 9.25 mmol) in EtOH (100 mL) was added dropwise at 0 °C oxalyl chloride (1.50 g, 9.25 mmol). The reaction was left to stirr at room temperature for 48 h. The desired product was obtain as yellow solid after filtration (2.29, 97 %)

Synthesis of 2-(5-nitro-1H-indol-3-yl)-2-oxoacetic acid 151
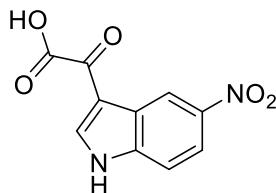
Compound **150** was suspended in 60 mL of H_2O and KOH (2.55, 45.a mmol). The reaction was heated at 120 °C for 24 h. After cooling the reaction was acidified with HCl 12 N until pH 1. The title compound was recovered after filtration as yellow solid (2.10 g, 99%).

Synthesis of 5-nitro-1H-indole-3-carboxylic acid (152)
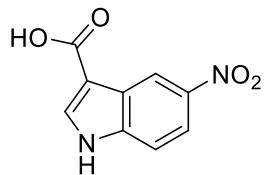
Compound **151** was dissolved in 150 mL of aqueous solution of H_2O_2 and heated at 100 °C for 24 h. Compound**152** was obtained as yellow solid (2.10 g, 9.08 mmol) after filtration, mp 287 – 290 °C. ^{1}H NMR (300 MHz, DMSO): δ

12.47 (br s, 2H, exchangeable with D_2O); 8.91 (d, J = 2 Hz, 1H); 8.29 (s, 1H); 8.11 (dd, J_1 = 9 Hz, J_2 = 2 Hz, 1H); 7.68 (d, J = 9 Hz, 1H). ESI m/z: 207 [M + H]$^+$.

Synthesis of ethyl 5-nitro-1H-indole-3-carboxylate (149b)

To a solution of **152** (1.60 g, 7.73 mmol) in EtOH absolute (95 mL) was added H_2SO_4 (1 mL) and heated at reflux. The mixture was concentrated under reduced pressure, the residue was dissolved with AcOEt (100 mL) and washed with NaOH 2N (3 x 30 mL) and Brine. The organic layer was dried with Na_2SO_4, filtered and evaporated under reduced pressure. Compound **149** was obtained as yellow solid (1.59 g, 88%), mp 229 – 230 °C. ^{1}H NMR (300 MHz, DMSO-d_6): δ 11.90 (s, 1H exchangeable with D_2O); 9.18 (d, J = 1.4 Hz, 1H); 8.46 (s 1H), 8.25 (dd, J_1 = 7.5 Hz, J_2 = 1.15 Hz, 1H); 7.59 (d, J = 7.5 Hz, 1H); 4.29 (q, J = 5.8 Hz, 2H); 1.30 (t, J = 5.9 Hz, 3H). ESI m/z: 234.21 [M + H]$^+$.

Synthesis of ethyl (E)-2-(2-(4-nitrophenyl)hydrazono)propanoate 153
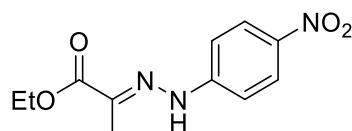
A solution of 4-nitro-phenylhydrazine (4.20 g, 27.4 mmol) in glacial AcOH (52 mL) was heated at 55 °C for 1 h and then was added ethyl piruvate (3 mL, 27.4 mmol). After 5 h were added 100 mL of saturated solution of Ca_2CO_3. The aqueous solution was extracted with AcOEt (3 x 200 mL), then the combined organic layers were washed with Brine, dried with Na_2SO_4, filtered and evaporated inder reduced pressure. Compound **153** was obtained as brown

solid (6.58 g, 95 %) after trituration with hexane, mp 156.8 - 157.2 °C. ^{1}H NMR (300 MHz, CDCl$_3$): δ 8.21 (d, J = 9 Hz, 2H); 7.98 (s, 1H, exchangeable with D$_2$O); 7.27 (d, J = 9 Hz, 2H); 4.34 (q, J = 14 Hz, 2H); 2.17 (s, 3H); 1.39 (t, J = 14 Hz, 3H). ESI m/z: 252 [M + H]$^+$.

Synthesis of ethyl 5-nitro-1H-indole-2-carboxylate 149c
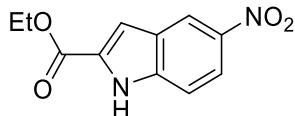
To a solution of **153** (2.00 g, 7.96 mmol) in tolene 200 mL was added PPA (12.5 g) and the reaction wa left to stir at 115 °C for 12 h and 100 mL of saturated aqueous solution of NaHCO$_3$ and 100 mL of NaOH 6N were added. The basic aqueous solution was extracted with toluene (5 x 100 mL) and the combined organic layers were washed with brine dried, filtered and evaporated under reduced pressure. Compound **149c** was obtained as yellow solid (1.40 g, 75 %) after purification by flash chromatography (AcOEt : hexane 1:1); mp 209 – 212 °C. ^{1}H NMR (300 MHz, CDCl$_3$): δ 9.24 (s, 1H exchangeable with D$_2$O); 8.69 (d, J = 2 Hz, 1H); 8.23 (dd, J$_1$ = 9 Hz, J$_2$ = 2 Hz, 1H); 7.49 (d, J = 9 Hz, 1H); 7.39 – 7.38 (m, 1H); 4.45 (q, J = 14 Hz, 2H); 1.44 (t, J = 14 Hz, 3H). ESI m/z: 235 [M + H]$^+$.

Continuous-flow processing for the reduction of 155a – c using the H-Cube Pro™ hydrogenator. Example: synthesis of 1H-indol-5-amine (155a)
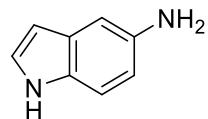
A 124 mL stock solution of **149a** 0.03 M concentration in EtOH/AcOEt 1:1 was prepared. The reaction parameters (full-H$_2$ mode, 30 °C, 10 bar and 1 mL min_1 floow rate) were selected on the H-Cube-Pro™ hydrogenator. The instrument was setted with a 30 mm Pd/C 10% CatCart and the processing was

started, whereby initially only pure solvent was pumped through the system until the instrument had achieved the desired reaction parameters and stable processing. At that point, the sample inlet line was switched to the vial containing the substrate. A total reaction volume of 124 mL was collected and the cartridge subsequently washed with pure solvent for 5 min to remove any substrate/product still adsorbed on the catalyst. Evaporation of the solvent afforded the desired product **155a** (0.810 g, 99%), mp 131 – 133 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.56 (s, 1H exchangeable with D_2O), 7.11 (t, J = 2.7 Hz, 1H), 7.07 (d, J = 8.5 Hz, 1H), 6.72 – 6.63 (m, 1H), 6.47 (dd, J = 8.5, 2.0 Hz, 1H), 6.14 – 6.08 (m, 1H), 4.62 (s, 2H exchangeable with D_2O). ESI-MS m/z: 133 [M+H]$^+$.

Ethyl 5-amino-1H-indole-3-carboxylate 155b
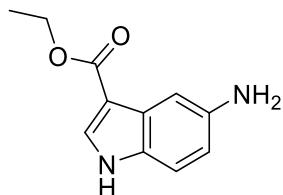
Yellow solid (0.872 g, resa 99%), mp 120 – 121 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.78 (s, 1H exchangeable with D_2O), 8.35 (s, 1H), 7.33 (d, J = 1.5 Hz, 1H), 7.11 (d, J = 7.4 Hz, 1H), 4.33 – 4.23 (m, 4H exchangeable with D_2O), 1.30 (t, J = 5.9 Hz, 3H). ESI m/z: 205 [M + H]$^+$

Ethyl 5-amino-1H-indole-2-carboxylate 155c
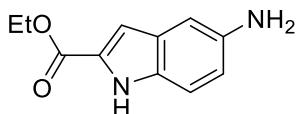
Yellow solid (0.872 g, 99%), 113.6 – 114.9 °C. ^{1}H NMR (300 MHz, DMSO): δ 11.39 (s, 1H, exchangeable with D_2O); 7.16 (d, J = 9 Hz, 1H); 6.84 (s, 1H);

6.71 - 6.68 (m, 2H); 4.65 (s, 2H, exchangeable with D_2O); 4.30 (q, J = 14 Hz, 2H); 1.32 (t, J = 14 Hz, 3H). ESI m/z: 205 $[M + H]^+$.

Synthesis of Series I derivatives

General procedures for coupling reaction (156a – c). Example: synthesis of (9H-fluoren-9-yl)methyl (2-((1H-indol-5-yl)amino)-2-oxoethyl) carbamate (156a)
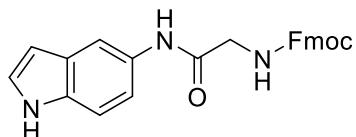
To a solution of HOBt (1.36 g, 8.93 mmol), HBTU (3.38 g, 8.93 mmol), DIPEA (10.93 mL, 17.85 mmol) and Fmoc-Gly-OH (1.66 g, 5.58 mmol) in THF/DMF 4:1 (41 mL) was added a solution of 155a (0.491g, 3.72 mmol)) in THF/DMF 4:1 (12 mL). The reaction was left to stir at room temperature for 12 h and then was concentrated under reduced pressure. The residue was dissolved with AcOEt (100 mL) and washed with) s.s. of $NaHCO_3$ (3 x 10 mL), NaCl and dried over Na_2SO_4. Filtration, evaporation under vacuum follone by chromatography purification (AcOEt/DCM, 1:9) afford the desired product **156a** as white solid (0.760 g, 50%), mp 105-110 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.99 (s, 1H exchangeable with D_2O), 9.73 (s, 1H exchangeable with D_2O), 7.91 – 7.85 (m, 3H), 7.75 – 7.73 (m, 2H), 7.64 – 7.60 (m, 1H), 7.45 – 7.40 (m, 2H), 7.36 – 7.29 (m, 4H), 7.21 – 7.17 (m, 1H), 6.37 (t, *J* = 5.5 Hz, 1H, exchangeable with D_2O), 4.– 4.25 (m, 3H), 3.80 (d, *J* = 5.5 Hz, 2H). ESI m/z: 412 $[M + H]^+$.

Ethyl 5-(2-((tert-butoxycarbonyl)amino)acetamido)-1H-indole-3-carboxylate 156b
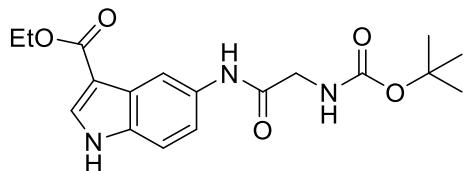
Compound **156b** was obtained as white solid (0.955 g, 50%) starting from **155b** (1.24 g, 5.29 mmol) and Boc-Gly-OH (1.85 g, 10.58 mmol) following the synthetic procedures used for **156a**; mp 134 – 138 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.84 (s, 1H exchangeable with D_2O), 9.85 (s, 1H exchangeable with D_2O), 8.26 (s, 1H), 8.01 (d, $J = 2.6$ Hz, 1H), 7.45 – 7.37 (m, 2H), 7.03 (t, $J = 6.2$ Hz, 1H exchangeable with D_2O), 4.28 (q, $J = 7.1$ Hz, 2H), 3.73 (d, $J = 6.2$ Hz, 2H), 1.42 (s, 9H), 1.35 (t, $J = 7.1$ Hz, 3H). ESI m/z: 362 [M + H]$^+$.

Ethyl 5-(2-((tert-butoxycarbonyl)amino)acetamido)-1H-indole-2-carboxylate 156c
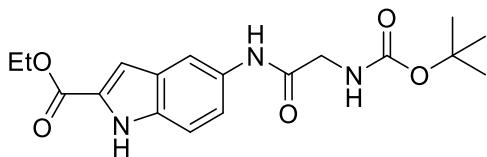
Compound **156c** was obtained as white solid (0.955 g, 50%) starting from **155c** (1.24 g, 5.29 mmol) and Boc-Gly-OH (1.85 g, 10.58 mmol) following the synthetic procedures used for **156a;** 193 – 194 °C ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.79 (s, 1H exchangeable with D_2O), 9.78 (s, 1H exchangeable with D_2O), 7.99 (s, 1H), 7.39-7.31 (m, 2H), 7.058 (d, $J = 2.1$ Hz, 1H), 6.99 (t, $J = 5.2$ Hz, 1H exchangeable with D_2O), 4.31 (q, $J = 6.9$ Hz, 2H), 3.70 (d, $J = 5.2$ Hz, 2H), 1.38 (s, 9H), 1.32 (t, $J = 6.9$ Hz, 3H). ESI m/z: 362 [M + H]$^+$.

Synthesis of 2-amino-N-(1H-indol-5-yl)acetamide 157a
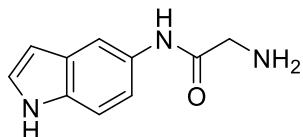
Compound **156a** was dissolved in 50 mL of DCM/piperidine mixture (8:2). The reaction was left to stir at room temperature for 30 min. The mixture was concentrated under reduced pressure, Et$_2$O was added (30 mL) and the product was recovered after filtration as white solid (0.349, 99%), mp 152 – 154 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.05 (s, 1H exchangeable with D$_2$O), 9.71 (s, 1H exchangeable with D$_2$O), 7.95 (d, J = 2.0 Hz, 1H), 7.38 – 7.26 (m, 3H), 6.42 – 6.41 (m, 1H), 2.91 (s, 2H), 2.76 (brs, 2H exchangeable with D$_2$O). ESI m/z: 190 [M + H]$^+$.

Synthesis of ethyl 5-(2-aminoacetamido)-1H-indole-3-carboxylate 2,2,2-trifluoroacetate 157b
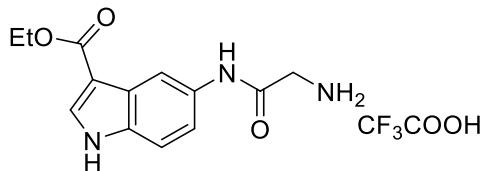
Compound **156 b** was treated with a mixture of DCM/TFA 9:1 (10 mL). The reaction was left to stir at room temperature for 30 min. Evaporation of solvent under vacuum gave the desired product **157b** as white solid (0.517 g, 99%); mp 140 – 142 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.93 (s, 1H exchangeable with D$_2$O), 10.34 (s, 1H exchangeable with D$_2$O), 8.31 (d, J = 1.9 Hz, 1H), 8.20 – 7.99 (m, 4H, 2H exchangeable with D$_2$O), 7.49 – 7.41 (m, 2H), 4.27 (q, *J* = 7.0 Hz, 2H), 3.78 (s, 2H), 1.33 (t, *J* = 7.0 Hz, 3H). ESI m/z: 262 [M + H]$^+$.

Synthesis of ethyl 5-(2-aminoacetamido)-1H-indole-2-carboxylate 2,2,2-trifluoroacetate (157c)
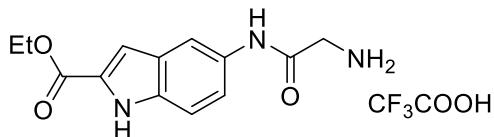
Compound **157c** was obtained as white solid (0.600 g, 99%) starting from **156c** (0.583 g, 1.61 mmol) according to the procedure used for **157b**; mp 206 – 208 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.87 (s, 1H exchangeable with D_2O), 10.31 (s, 1H exchangeable with D_2O), 8.18 – 8.03 (m, 3H, 2H exchangeable with D_2O), 8.00 (s, d, J = 1.3 Hz, 1H), 7.43 (d, J = 7.9 Hz, 1H), 7.34 (7.34 (dd, J = 7.9, 1.3 Hz, 1H), 7.14 (s, 1H), 4.33 (q, J = 7.0 Hz, 2H), 3.77 (s, 2H), 1.33 (t, J = 7.0 Hz, 3H). ESI m/z: 262 [M + H]$^+$

General procedures for preparation of 4methyl-phenyl substituted ureidic derivatives 159a – c.

Example: synthesis of N-(1H-indol-5-yl)-2-(3-(p-tolyl)ureido)acetamide 159a
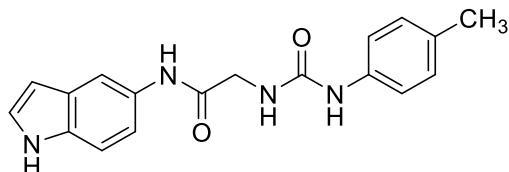
To a suspension of 157a (0.169 g, 0.894 mmol) and TEA (0.31 mL, 2.23 mmol) in THF dry (25 mL) was added dropwise a solution of *p*-tolil-isocyanate (0.17 mL, 1.34 mmol) in THF dry (14 ml). The reaction was left to stir at room temperature for 12h. The mixture of reaction was evaporated under vacuum and crystallization with THF afford the desired product **159** (0.144, 50 %) as white solid; mp 203 – 204 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.00 (s, 1H exchangeable with D_2O), 9.82 (s, 1H exchangeable with D_2O), 8.71 (s, 1H), 7.85 (s, 1H), 7.42 – 7.24 (m, 4H), 7.20 (d, J = 8.8 Hz, 1H),

7.12 – 6.96 (7.03 (d, *J* = 8.3 Hz, 2H)), 6.38 – 6.34 (m, 2H), 3.93 (d, *J* = 5.4 Hz, 2H), 2.22 (s, 3H). ESI m/z: 323 [M + H]$^+$.

Ethyl 5-(2-(3-(p-tolyl)ureido)acetamido)-1H-indole-3-carboxylate 159b
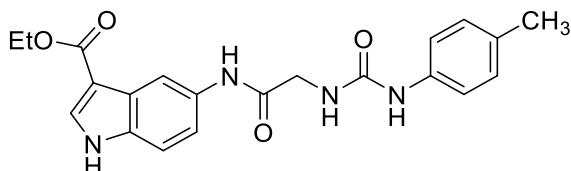
White solid (0.050 g, 60%); 240 – 242 °C. H NMR (300 MHz, DMSO) δ 11.85 (s, 1H, exchangeable with D$_2$O), 9.98 (s, 1H exchangeable with D$_2$O), 8.71 (s, 1H exchangeable with D$_2$O), 8.28 (s, 1H), 8.05 – 7.99 (m, 1H), 7.49 – 7.36 (m, 2H), 7.29 (d, *J* = 8.0 Hz, 2H), 7.03 (d, *J* = 8.0 Hz, 2H), 6.39 (t, *J* = 5.1 Hz, 1H exchangeable with D$_2$O), 4.27 (q, *J* = 7.0 Hz, 2H), 3.94 (d, *J* = 5.1 Hz, 2H), 2.21 (s, 3H), 1.33 (t, *J* = 7.0 Hz, 3H). ESI m/z: 395 [M + H]$^+$.

Ethyl 5-(2-(3-(p-tolyl)ureido)acetamido)-1H-indole-2-carboxylate 159c
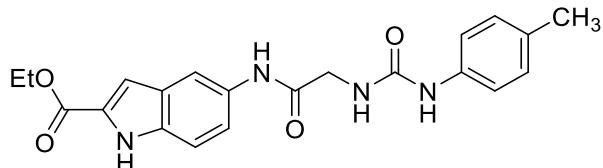
White solid (0.100 g, 60%); 267 – 268 °C. H NMR (300 MHz, DMSO) δ 11.80 (s, 1H exchangeable with D$_2$O), 9.92 (s, 1H exchangeable with D$_2$O), 8.70 (s, 1H exchangeable with D$_2$O), 7.99 (s, 1H),7.40 – 7.33 (m, 2H), 7.28 (d, *J* = 8.1 Hz, 2H), 7.11 – 7.08 (m, 1H), 7.02 (d, *J* = 8.1 Hz, 2H), 6.37 (t, *J* = 5.1 Hz, 1H exchangeable with D$_2$O), 4.36 (q, *J* = 7.0 Hz, 2H), 3.93 (d, *J* = 5.1 Hz, 2H), 2.21 (s, 3H), 1.35 (t, *J* = 7.0 Hz, 3H). ESI m/z: 395 [M + H]$^+$

General procedures for ureidic function formation through CDI (161a – c and 163a – c).

Example: synthesis of N-(1H-indol-5-yl)-2-(3-(2-(piperidin-1-yl)ethyl) ureido)acetamide 161a
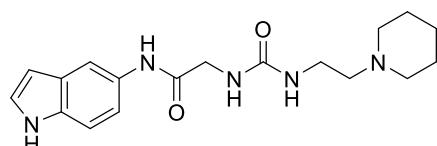
To a cooled solution of N-aminoethyl-piperidine **160** (0.12 ml, 0.868 mmol) and TEA (0.35, 2.55 mmol) in DCM dry (9 mL), was added CDI (0.198 g, 1.22 mmol) and a solution **157a** (0.109g, 0.579 mmol) in THF dry (2.2 mL). The reaction was left to stir at room temperature for 12 h. The mixture of reaction was concentrated under reduced pressure and purification by flash chromatography (DCM/MeOH/TEA, 9:1:0.5) afford the desired product as white product; 234 – 237 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.98 (s, 1H exchangeable with D$_2$O), 9.69 (s, 1H exchangeable with D$_2$O), 7.83 (s, 1H), 7.32 – 7.27 (m, 2H), 7.18 (d, J = 8.4 Hz, 1H), 6.43 – 6.31 (m, 2H exchangeable with D$_2$O), 6.11 (t, J = 4.7 Hz, 1H exchangeable with D$_2$O), 3.82 (d, J = 4.7 Hz, 2H), 3.17 – 3.07 (m, 2H), 2.41 – 2.26 (m, 5H), 1.55 – 1.45 (m, 5H), 1.42 – 1.35 (m, 2H). ESI m/z: 344 [M + H]$^+$.

Ethyl 5-(2-(3-(2-(piperidin-1-yl)ethyl)ureido)acetamido)-1H-indole-3-carboxylate 161b
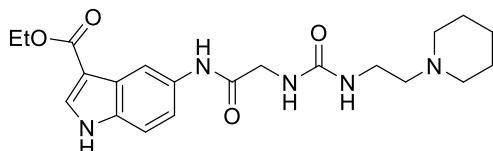
Compound **161b** was obtained as white solid (0.070 g, 64%) starting from **157b** (0.100 g, 0.266 mmol) according to the procedure used for **161a**; mp 168 - 170 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.83 (s, 1H exchangeable with

D$_2$O), 9.85 (s, 1H exchangeable with D$_2$O), 8.26 (s, 1H), 8.01 (s, 1H), 7.49 – 7.30 (m, 2H), 6.37 (s, 1H exchangeable with D$_2$O), 6.10 (t, *J* = 4.8 Hz, 1H exchangeable with D$_2$O), 4.26 (q, *J* = 6.9 Hz, 2H), 3.83 (d, *J* = 5.2 Hz, 2H), 3.11 (q, *J* = 5.7 Hz, 2H), 2.38 – 2.23 (m, 6H), 1.56 – 1.44 (m, 4H), 1.42 – 1.38 (m, 2H), 1.33 (t, *J* = 7.0 Hz, 3H). ESI m/z: 416 [M + H]$^+$

Ethyl 5-(2-(3-(2-(piperidin-1-yl)ethyl)ureido)acetamido)-1H-indole-2-carboxylate 161c
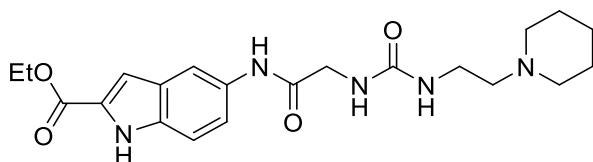
Compound **161c** was obtained as white solid (0.070 g, 64%) starting from **157c** (0.100 g, 0.266 mmol) according to the procedure used for **161a**; mp 182 – 184 °C. ^{1}H NMR (300 MHz, DMSO) δ 11.81 (s, 1H exchangeable with D$_2$O), 9.83 (s, 1H exchangeable with D$_2$O), 7.99 (s, 1H), 7.43 – 7.30 (m, 2H), 7.14 – 7.06 (m, 1H), 6.37 (t, *J* = 5.0 Hz, 1H exchangeable with D$_2$O), 6.10 (t, *J* = 6.0 Hz, 1H exchangeable with D$_2$O), 4.34 (q, *J* = 6.9 Hz, 2H), 3.82 (d, *J* = 5.2 Hz, 2H), 3.16 – 3.06 (m, 2H), 2.39 – 2.24 (m, 6H), 1.56 – 1.44 (m, 4H), 1.43 – 1.29 (m, 5H). ESI m/z: 416 [M + H]$^+$

2-(3-(2-azidoethyl)ureido)-N-(1H-indol-5-yl)acetamide 163a
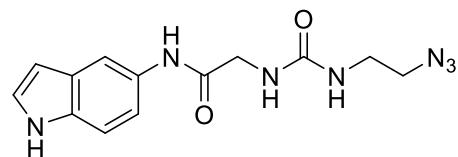
Compound **163a** was obtained as white solid (0.156 g, 40%) starting from **157a** (0.244 g, 1.29 mmol) and 2-azidoethan-1-amine (0.236 g, 1.93 mmol) according to the procedure used for **161a**; mp 154 – 157 °C. ^{1}H NMR (300 MHz, DMSO) δ 10.97 (s, 1H exchangeable with D$_2$O), 9.68 (s, 1H

exchangeable with D$_2$O), 7.84 (s, 1H), 7.37 – 7.24 (m, 2H), 7.18 (dd, *J* = 8.7, 1.5 Hz, 1H), 6.45 (t, *J* = 5.6 Hz, 1H exchangeable with D$_2$O), 6.40 – 6.33 (m, 1H), 6.29 (t, *J* = 5.4 Hz, 1H exchangeable with D$_2$O), 3.84 (d, *J* = 5.5 Hz, 2H), 3.35 (d, *J* = 5.8 Hz, 2H), 3.26 – 3.16 (m, 2H). ESI m/z: 302 [M + H]$^+$.

Ethyl 5-(11,11-dimethyl-4,9-dioxo-10-oxa-3,5,8-triazadodecanamido)-1H-indole-3-carboxylate 163b
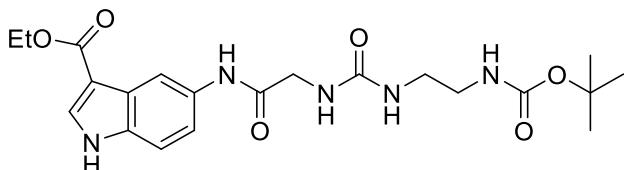
Compound **163b** was obtained as white solid (0.350 g, 75%) starting from **157b** (0.391 g, 1.04 mmol) and tert-butyl (2-aminoethyl)carbamate (0.250 g, 1.56 mmol) according to the procedure used for **161a**; 166 – 170 °C. ^{1}H NMR (300 MHz, DMSO-*d$_6$*) δ 11.83 (s, 1H exchangeable with D$_2$O), 9.85 (s, 1H exchangeable with D$_2$O), 8.26 (s, 1H), 8.01 (s, 1H), 7.47 – 7.34 (m, 2H), 6.83 – 6.73 (m, 1H exchangeable with D$_2$O), 6.33 – 6.18 (m, 2H exchangeable with D$_2$O), 4.26 (q, *J* = 6.9 Hz, 2H), 3.83 (d, *J* = 4.8 Hz, 2H), 3.10 – 3.00 (m, 2H), 3.00 – 2.91 (m, 2H), 1.37 (s, 9H), 1.31 (t, *J* = 6.8 Hz, 3H). ESI m/z: 448 [M + H]$^+$.

Ethyl 5-(11,11-dimethyl-4,9-dioxo-10-oxa-3,5,8-triazadodecanamido)-1H-indole-2-carboxylate 163c
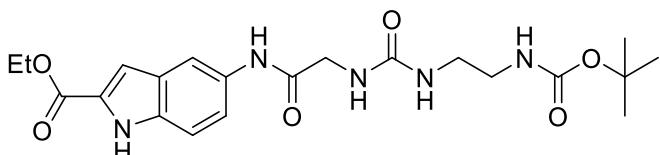
Compound **163c** was obtained as white solid ((0.110 g, 63%) starting from **157b** (0.146 g, 0.389 mmol) and tert-butyl (2-aminoethyl)carbamate (0.250 g,

1.56 mmol) according to the procedure used for **161a;** 194 - 196 °C. ^{1}H NMR (300 MHz, DMSO) δ 11.79 (s, 1H, exchangeable with D$_2$O), 9.80 (s, 1H exchangeable with D$_2$O), 7.99 (s, 1H), 7.42 – 7.31 (m, 2H), 7.12 – 7.04 (m, 1H), 6.78 (t, *J* = 5.2 Hz, 1H exchangeable with D$_2$O), 6.32 – 6.18 (m, 2H exchangeable with D$_2$O), 4.33 (q, *J* = 7.0 Hz, 2H), 3.82 (d, *J* = 5.2 Hz, 2H), 3.10 – 3.00 (m, 2H), 2.99 – 2.90 (m, 2H), 1.38 (s, 9H), 1.32 (t, *J* = 7.0 Hz, 3H). ESI m/z: 448 [M + H]$^+$.

Synthesis of 2-(3-(2-aminoethyl)ureido)-N-(1H-indol-5-yl)acetamide 164a
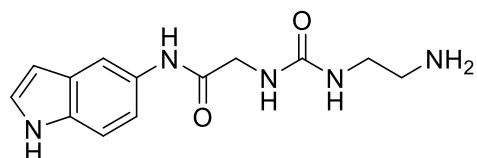
Compound **163a** (0.062 g, 0.205 mmol) was dissolved in MeOH (10 mL). The reaction parameters of H-Cube-Pro™ hydrogenator (full-H$_2$ mode, 30 °C, 10 bar and 1 mL min_1 floow rate) were selected on the H-Cube-Pro™ hydrogenator. The instrument was setted with a 30 mm Pd/C 10% CatCart and the processing was started, whereby initially only pure solvent was pumped through the system until the instrument had achieved the desired reaction parameters and stable processing. At that point, the sample inlet line was switched to the vial containing the substrate. A total reaction volume of 10 mL was collected and the cartridge subsequently washed with pure solvent for 5 min to remove any substrate/product still adsorbed on the catalyst. Evaporation of the solvent afforded the desired product **164a** (0.052 g, 93%); 150 - 152 °C. ^{1}H NMR (300 MHz, DMSO-*d$_6$*) δ 10.98 (s, 1H exchangeable with D$_2$O), 9.69 (s, 1H exchangeable with D$_2$O), 7.84 (s, 1H), 7.35 – 7.26 (m, 2H), 7.22 – 7.14 (m, 1H), 6.41 – 6.33 (m, 1H), 6.32 – 6.14 (m, 2H exchangeable with D$_2$O), 3.82 (d, *J* = 5.6 Hz, 2H), 3.17 (s, 1H), 3.00 (q, *J* = 6.1 Hz, 2H), 2.55 (t, *J* = 6.2 Hz, 2H). ESI m/z: 276 [M + H]$^+$.

Synthesis of ethyl 5-(2-(3-(2-aminoethyl)ureido)acetamido)-1H-indole-3-carboxylate 164b
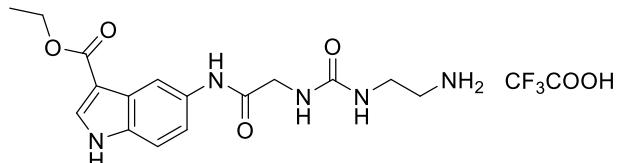
Compound 163b (0.097 g, 0.217 mmol) was treated with a solution of DCM/TFA 9:1 (5 mL) and the reaction was left to stir at room temperature for 30 min. Evaporation under vacuum of reaction solvent afford the desired product **164b** as white solid; 124 – 126 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 12.04 (s, 1H exchangeable with D$_2$O), 11.81 (s, 1H exchangeable with D$_2$O), 10.62 (s, 1H), 9.13 (s, 1H), 8.37 (s, 1H exchangeable with D$_2$O), 8.22 (s, 2H), 8.10 (s, 1H), 7.61 (d, J = 1.1 Hz, 2H), 7.36 (s, 1H), 4.29 (q, J = 5.9 Hz, 2H), 4.10 (s, 2H exchangeable with D$_2$O), 3.22 (t, J = 7.4 Hz, 2H), 2.79 (t, J = 7.4 Hz, 2H), 1.30 (t, J = 5.9 Hz, 3H). ESI m/z: 348 [M + H]$^+$.

Synthesis of ethyl 5-(2-(3-(2-aminoethyl)ureido)acetamido)-1H-indole-3-carboxylate 164c
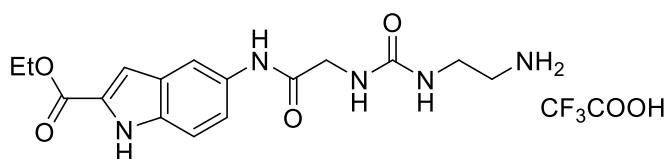
Compound **164c** was obtained as white solid (0.079 g, 77%) starting from **163c** (0.099 g, 0.222 mmol) according to the procedures used for **164b**; 215 – 217 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.81 (s, 1H exchangeable with D$_2$O), 9.84 (s, 1H exchangeable with D$_2$O), 7.99 (s, 1H), 7.72 (s, 2H exchangeable with D$_2$O), 7.44 – 7.28 (m, 2H), 7.10 (s, 1H), 6.56 – 6.39 (m, 2H exchangeable with D$_2$O), 4.33 (q, J = 7.0 Hz, 2H), 3.85 (d, J = 5.0 Hz, 2H),

3.28 – 3.20 (m, 2H), 2.90 – 2.80 (m, 2H), 1.33 (t, J = 7.0 Hz, 3H). ESI m/z: 348 [M + H]$^+$.

Series II derivatives

Synthesis of ethyl ((1H-indol-5-yl)carbamoyl)glycinate 165a
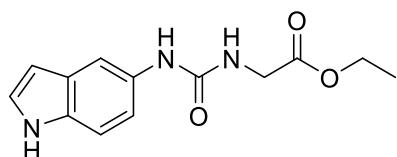
To a cooled solution of ethyl ester glycine chlorohydrate (0.351 g, 2.52 mmol) and TEA (0.35 mL, 7.40 mmol) in DCM dry (40 mL) were added CDI (0.432 g, 2.69 mmol) and a solution of **155a** (0.222 g, 1.68 mmol) in THF dry (5.7 mL). The reaction was left to stir at room temperature for 12h. The mixture of reaction was concentrated under reduced pressure, the residue was dissolved with AcOEt (100 mL) and washed with HCl 1N (3 x 20 mL), s.s. of NaHCO$_3$. The organic phase was dried over Na$_2$SO$_4$, filtered and evaporated under reduced pressure. The desired product **165a** was obtained as white solid (0.307 g, 70%) after crystallization in CHCl$_3$; mp 157 – 160 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.88 (s, 1H exchangeable with D$_2$O), 8.45 (s, 1H exchangeable with D$_2$O), 7.60 (d, J = 2.0 Hz, 1H), 7.33 – 7.21 (m, 2H), 7.01 (dd, J = 8.7, 2.0 Hz, 1H), 6.35 – 6.24 (m, 2H exchangeable with D$_2$O), 4.11 (q, J = 7.1 Hz, 2H), 3.86 (d, J = 5.9 Hz, 2H), 1.21 (t, J = 7.1 Hz, 3H). ESI m/z: 262 [M + H]$^+$.

Synthesis of ethyl 5-(3-(2-(tert-butoxy)-2-oxoethyl)ureido)-1H-indole-3-carboxylate 165b
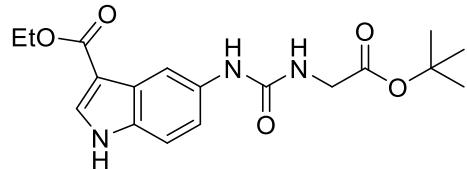
Compound **165b** was obtained white solid (0.720 g, 50%) according to the procedure used for **165a**, starting from **155b** (0.813 g, 3.98 mmol) and tert butyl-glycine ester (0.626 g, 4.77 mmol); mp 145 - 147 °C . ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.74 (s, 1H exchangeable with D_2O), 8.65 (s, 1H exchangeable with D_2O), 8.00 (s, 1H), 7.96 (d, J = 2.8 Hz, 1H), 7.37 − 7.24 (m, 2H), 6.23 (t, J = 6.0 Hz, 1H exchangeable with D_2O), 4.26 (q, J = 7.1 Hz, 2H), 3.76 (d, J = 5.8 Hz, 2H), 1.43 (s, 9H), 1.32 (t, J = 7.1 Hz, 3H). ESI m/z: 362 [M + H]$^+$

Synthesis of ethyl 5-(3-(2-(tert-butoxy)-2-oxoethyl)ureido)-1H-indole-2-carboxylate 165c
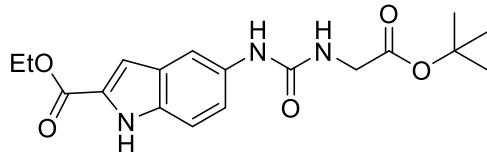
Compound **165c** was obtained white solid (0.720 g, 50%) according to the procedure used for **165a**, starting from **155c** (0.813 g, 3.98 mmol) and tert butyl-glycine ester (0.626 g, 4.77 mmol); mp 135 - 140 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.68 (s, 1H exchangeable with D_2O), 8.58 (s, exchangeable with D_2O), 7.75 (s, 1H), 7.32 (d, J = 9.0 Hz, 1H), 7.23 − 7.14 (m, 1H), 7.08 − 7.01 (m, 1H), 6.26 (t, J = 5.8 Hz, 1H exchangeable with D_2O), 4.32 (q, J = 7.1 Hz, 2H), 3.75 (d, J = 5.6 Hz, 2H), 1.42 (s, 9H), 1.33 (t, J = 6.9 Hz, 3H). ESI m/z: 362 [M + H]$^+$.

Synthesis of ((1H-indol-5-yl)carbamoyl)glycine 166a
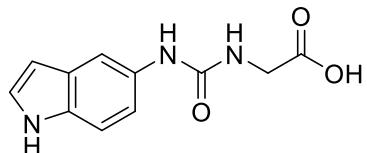
To a solution of 165a (0.100 g, 0.38 mmol) in EtOH (1 mL) was added NaOH 2N (2 eq). The reaction was left to stir for 2h and then was concentrated under reduced pressure. The residue was dissolved with s. s. of $NaHCO_3$ (100 mL) and washed with $CHCl_3$. The aqueous phase was acidified with HCl 0.3 N until pH 1 and extracted with AcOEt (3 x 20 mL). The organic phase was washed with brine, dried with Na_2SO_4, filtered and evaporated under vacuum. The desired product was obtained as white solid (0.070 g, 80 %) and was used in the next step without further purification; mp 205 - 207 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 12.51 (s, 1H exchangeable with D_2O), 10.87 (s, 1H exchangeable with D_2O), 8.43 (s, 1H exchangeable with D_2O), 7.60 (d, J = 2.0 Hz, 1H), 7.31 – 7.21 (m, 2H), 7.01 (dd, J = 8.7, 2.1 Hz, 1H), 6.36 – 6.27 (m, 1H), 6.20 (t, J = 5.8 Hz, 1H exchangeable with D_2O), 3.79 (d, J = 5.7 Hz, 2H). ESI m/z: 234 [M + H]$^+$.

Synthesis of ((3-(ethoxycarbonyl)-1H-indol-5-yl)carbamoyl)glycine 166b
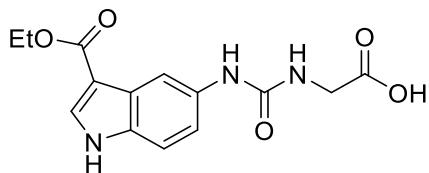
A solution of DCM/TFA (8:2, 13 mL) was added to **165b**. The reaction was left to stir at room temperature for 1h. Evaporation of mixture reaction under vacuum affar the desired product as yellow solid (0.227 g, resa 98%), which was used in the next step without further purification. Mp 250 - 255 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 12.33 (s, 1H exchangeable with D_2O), 11.85 (s,

1H exchangeable with D$_2$O), 9.40 (s, 1H exchangeable with D$_2$O), 9.13 (s, 1H exchangeable with D$_2$O), 8.38 (s, 1H), 7.93 (s, 3H), 4.29 (q, *J* = 5.9 Hz, 2H), 4.00 (s, 2H), 1.30 (t, *J* = 5.9 Hz, 3H). ESI m/z: 306 [M + H]$^+$

Synthesis of ((2-(ethoxycarbonyl)-1H-indol-5-yl)carbamoyl)glycine 166c
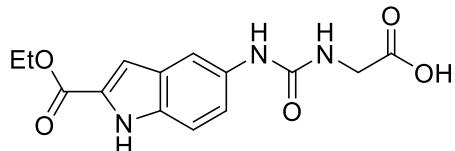
Compound **166c** was obtained as white solid starting from **165c** (0.200 g, 0.553 mmol) according to the procedure described for **166b.** Mp 173 - 176 °C; ^{1}H NMR (300 MHz, DMSO-*d$_6$*) δ 11.68 (s, 1H exchangeable with D$_2$O), 8.60 (s, 1H exchangeable with D$_2$O), 7.76 (s, 1H), 7.32 (d, *J* = 8.9 Hz, 1H), 7.18 (d, *J* = 8.7 Hz, 1H), 7.03 (s, 1H), 6.39 (s, 1H exchangeable with D$_2$O), 6.33 – 6.22 (m, 1H exchangeable with D$_2$O), 4.32 (q, *J* = 7.1 Hz, 2H), 3.70 (s, 1H), 1.33 (t, *J* = 7.1 Hz, 3H). ESI m/z: 306 [M + H]$^+$.

General procedure for coupling reaction (168a – c; 169a – c and 170a – c) Example: synthesis 2-(3-(1H-indol-5-yl)ureido)-N-(p-tolyl)acetamide 168a
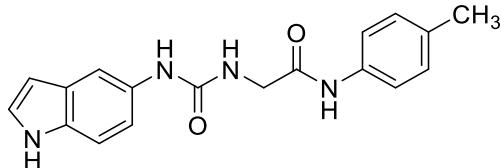
To a solution of **164 a** (0.132 g, 0.565 mmol) and TBTU (0.237 g, 0.738 mmol) in DMF dry (2 mL) was added a solution of *p*-methyl-aniline (0.073 g, 0.678 mmo) and TEA (0.24 mL, 1.71 mmol) in DMF (2 mL). The reaction was left to stir at room temperature for 12 h. The reaction was concentrated under reduced pressure, the residue was dissolved with AcOEt (50 mL) and washed with satuturated aqueous solution of NaHCO$_3$ and brine. The organic phase was dried over Na$_2$SO$_4$, filtered and evaporated under reduced pressure.

Purification by by silica gel chromatography DCM/MeOH, 95:5 gave the desired product **168a** as white solid (0.110 g, resa 60%). ^{1}H NMR (300 MHz, DMSO) δ 10.88 (s, 1H, exchangeable with D_2O), 9.92 (s, 1H exchangeable with D_2O), 8.50 (s, 1H exchangeable with D_2O), 7.62 (s, 1H), 7.60 – 7.39 (m, 2H), 7.33 – 7.20 (m, 2H), 7.11 (d, J = 8.0 Hz, 2H), 7.02 (d, J = 8.4 Hz, 1H), 6.35 – 6.24 (m, 2H exchangeable with D_2O), 3.92 (d, J = 5.1 Hz, 2H), 2.25 (s, 3H). ESI m/z: 323.2 $[M + H]^+$.

Ethyl 5-(3-(2-oxo-2-(p-tolylamino)ethyl)ureido)-1H-indole-3-carboxylate 168b
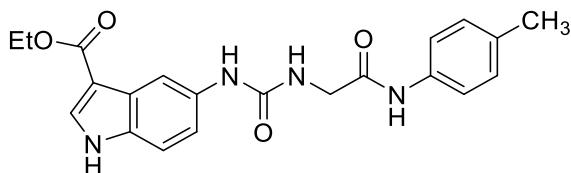
Compound **168b** was obtained as white solid (0.160 g, 50%) starting from **166b** (0.340 g, 0.811 mmol) according to the procedure described for **168a**; mp 214 – 216 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 13.68 (s, 1H exchangeable with D_2O), 11.88 (s, 1H exchangeable with D_2O), 9.38 (s, 1H exchangeable with D_2O), 9.13 (s, 1H exchangeable with D_2O), 8.43 – 8.29 (m, 2H), 8.18 (d, J = 1.4 Hz, 1H), 7.52 – 7.39 (m, 3H), 7.15 – 7.05 (m, 2H), 4.29 (q, J = 5.9 Hz, 2H), 4.10 (s, 2H), 2.32 (d, J = 1.1 Hz, 3H), 1.30 (t, J = 5.9 Hz, 3H). ESI m/z: 395 $[M + H]^+$.

Ethyl 5-(3-(2-oxo-2-(p-tolylamino)ethyl)ureido)-1H-indole-2-carboxylate 168c
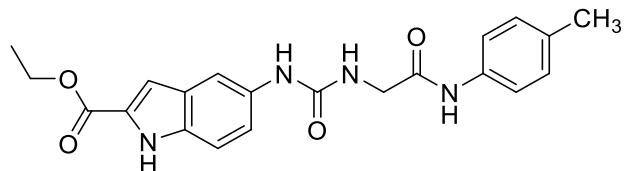
Compound **168c** was obtained as white solid (0.168 g, 60%) starting from **166c** (0.297 g, 0.709 mmol) according to the procedure described for **168a;** mp 233 – 236 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.69 (s, 1H exchangeable with D_2O), 9.94 (s, 1H exchangeable with D_2O), 8.72 (s, 1H exchangeable with D_2O), 7.78 (s, 1H), 7.55 – 7.42 (m, 2H), 7.32 (d, J = 8.8 Hz, 1H), 7.19 (d, J = 8.8 Hz, 1H), 7.15 – 7.07 (m, 2H), 7.05 – 7.01 (m, 1H), 6.41 (t, J = 4.8 Hz, 1H exchangeable with D_2O), 4.32 (q, J = 7.0 Hz, 2H), 3.92 (d, J = 5.2 Hz, 2H), 2.25 (s, 3H), 1.33 (t, J = 7.1 Hz, 3H). ESI m/z: 395 [M + H]$^+$.

2-(3-(1H-indol-5-yl)ureido)-N-(2-(piperidin-1-yl)ethyl)acetamide 169a
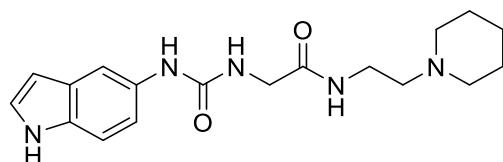
Compound **169a** was obtained as white solid starting from **166a** (0.150 g, 0.643 mmol) and 2-aminoethyl-piperidine (0.263 g, 0.83 mmol) according to the procedure described for **168a;** mp 157 – 159 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.87 (s, 1H exchangeable with D_2O), 8.46 (s, 1H exchangeable with D_2O), 7.84 (s, 1H exchangeable with D_2O), 7.60 (s, 1H), 7.32 – 7.20 (m, 2H), 7.05 – 6.97 (m, 1H), 6.33 – 6.28 (m, 1H), 6.23 (t, J = 4.8 Hz, 1H exchangeable with D_2O), 3.70 (d, J = 4.8 Hz, 2H), 3.26 – 3.15 (m, 2H), 2.47 – 2.33 (m, 5H), 1.55 – 1.43 (m, 5H), 1.42 – 1.32 (m, 2H). ESI m/z: 344 [M + H]$^+$.

Ethyl 5-(3-(2-oxo-2-((2-(piperidin-1-yl)ethyl)amino)ethyl)ureido)-1H-indole-3-carboxylate 169b
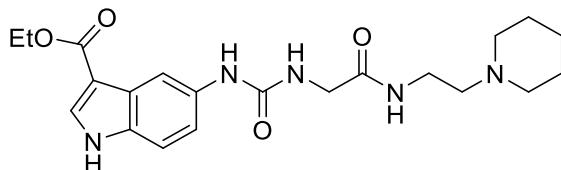
Compound **169b** was obtained as white solid (0.121g, 60%) starting from **166b** (0.200 g, 0.477 mmol) according to the procedure described for **169a;** mp 189 – 190 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.73 (s, 1H exchangeable with D_2O), 8.66 (s, 1H, exchangeable with D_2O), 8.07 – 7.92 (m, 2H), 7.88 – 7.77 (m, 1H exchangeable with D_2O), 7.42 – 7.23 (m, 2H), 6.24 (t, *J* = 5.0 Hz, 1H exchangeable with D_2O), 4.25 (d, *J* = 6.6 Hz, 2H), 3.71 (d, *J* = 4.6 Hz, 2H), 3.24 – 3.14 (m, *J* = 5.9 Hz, 2H), 2.41 – 2.24 (m, 6H), 1.55 – 1.41 (m, 4H), 1.36 – 1.28 (m, 5H). ESI m/z: 416 [M + H]$^+$.

Ethyl 5-(2-(2-oxo-2-((2-(piperidin-1-yl)ethyl)amino)ethyl)ureido)-1H-indole-3-carboxylate 169c
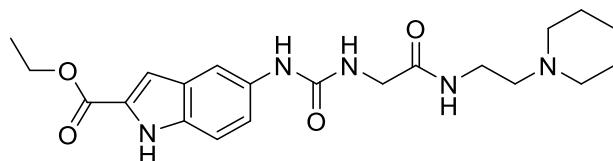
Compound **169c** was obtained as white solid (0.121g, 60%) starting from **166c** (0.200 g, 0.477 mmol) according to the procedure described for **169a;** mp 202 – 204 °C. ^{1}H NMR (300 MHz, DMSO) δ 11.68 (s, 1H, exchangeable with D_2O), 8.60 (s, 1H exchangeable with D_2O), 7.80 (t, *J* = 5.4 Hz, 1H exchangeable with D_2O), 7.76 (s, 1H), 7.32 (d, *J* = 8.8 Hz, 1H), 7.17 (dd, *J* = 8.9, 1.7 Hz, 1H), 7.03 (s, 1H), 6.27 (t, *J* = 5.2 Hz, 1H exchangeable with D_2O), 4.32 (q, *J* = 7.0 Hz, 2H), 3.70 (d, *J* = 5.2 Hz, 2H), 3.23 – 3.12 (m, 2H), 2.40 –

2.22 (m, 6H), 1.51 – 1.41 (m, 4H), 1.38 – 1.29 (m, 5H). ESI m/z: 416 [M + H]$^+$.

2-(3-(1H-indol-5-yl)ureido)-N-(2-azidoethyl)acetamide 170a
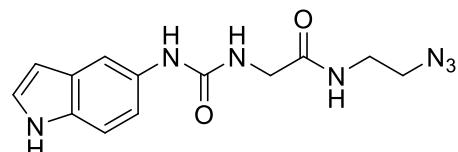
Compound **170a** was obtained as white solid (0.206 g, 53%) starting from **166a** (0.300 g, 1.28 mmol) and 2-azidoethan-1-amine (0.205 g, 1.68 mmol) according to the procedure described for **168a**; mp 165 – 167 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.88 (s, 1H exchangeable with D$_2$O), 8.46 (s, 1H exchangeable with D$_2$O), 8.17 (t, J = 5.4 Hz, 1H exchangeable with D$_2$O), 7.63 – 7.61 (m, 1H), 7.28 – 7.26 (m, 1H), 7.26 – 7.24 (m, 1H), 7.02 (dd, J = 8.6, 2.1 Hz, 1H), 6.34 – 6.30 (m, 1H), 6.22 (t, J = 5.5 Hz, 1H exchangeable with D$_2$O), 3.74 (d, J = 5.4 Hz, 2H), 3.39 (t, J = 5.7 Hz, 2H), 3.29 (t, J = 5.7 Hz, 2H). ESI m/z: 376 [M + H]$^+$.

Ethyl-5-(3-(2-((2-(((tert-butoxycarbonyl)amino)ethyl)amino)-2-oxoethyl) ureido)-1H-indole-3-carboxylate 170b
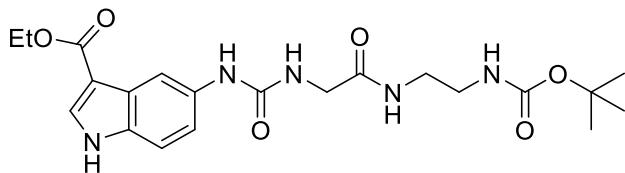
Compound **170b** was obtained as white solid (0.100 g, 76%) starting from **166b** (0.116 g, 0.274 mmol) and tert-butyl (2-aminoethyl)carbamate (0.105 g, 0.602 mmol) according to the procedure described for **168a**; 157 – 160 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.73 (s, 1H exchangeable with D$_2$O), 8.67 (s, 1H exchangeable with D$_2$O), 8.00 (s, 1H), 7.98 – 7.95 (m, 2H

exchangeable with D$_2$O), 7.40 – 7.22 (m, 2H), 6.89 – 6.76 (m, 1H, exchangeable with D$_2$O), 6.31 – 6.13 (m, 1H, exchangeable with D$_2$O), 4.25 (q, *J* = 6.9 Hz, 2H), 3.72 (d, *J* = 4.6 Hz, 2H), 3.17 – 3.06 (m, 2H), 3.05 – 2.93 (m, 2H), 1.37 (s, 9H), 1.32 (t, *J* = 6.9 Hz, 3H). ESI m/z: 478 [M + H]$^+$.

Ethyl5-(3-(2-((2-((tert-butoxycarbonyl)amino)ethyl)amino)-2-oxoethyl)ureido)-1H-indole-2-carboxylate 170c
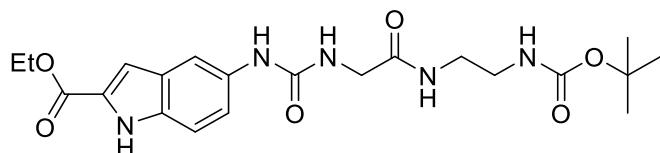
Compound **170c** was obtained as white solid (0.130 g, 60%) starting from **166b** (0.190 g, 0.450 mmol) and tert-butyl (2-aminoethyl)carbamate (0.250 g, 1 mmol) according to the procedure described for **168a;** 210 – 212 °C. ^{1}H NMR (300 MHz, DMSO-*d*$_6$) δ 11.67 (s, 1H exchangeable with D$_2$O), 8.61 (s, 1H exchangeable with D$_2$O), 8.04 – 7.88 (m, 1H exchangeable with D$_2$O), 7.76 (s, 1H), 7.38 – 7.26 (m, 1H), 7.21 – 7.12 (m, 1H), 7.03 (s, 1H), 6.86 – 6.74 (m, 1H exchangeable with D$_2$O), 6.31 – 6.18 (m, 1H exchangeable with D$_2$O), 4.32 (q, *J* = 6.9 Hz, 2H), 3.71 (d, *J* = 4.4 Hz, 2H), 3.18 – 3.04 (m, 2H), 3.03 – 2.92 (m, 2H), 1.37 (s, 9H), 1.32 (t, *J* = 7.1 Hz, 3H). ESI m/z: 478 [M + H]$^+$.

Synthesis of 2-(3-(1H-indol-5-yl)ureido)-N-(2-aminoethyl)acetamide171a
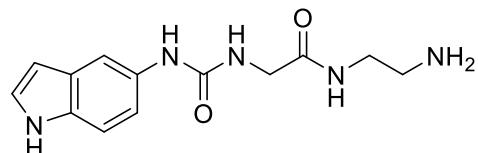
Compound **171a** was obtained as white solid starting from **170a** (0.070 g, 0.230 mmol) according to the procedure described for **164a;** mp 126 – 128 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.88 (s, 1H exchangeable with D$_2$O), 8.48 (s, 1H exchangeable with D$_2$O), 7.90 (t, J = 5.7 Hz, 1H exchangeable with D$_2$O), 7.62 (d, J = 1.9 Hz, 1H), 7.26 (dd, J = 5.9, 3.1 Hz, 2H), 7.03 (dd, J = 8.6, 2.0 Hz, 1H), 6.34 – 6.30 (m, 1H), 6.23 (t, J = 5.4 Hz, 1H exchangeable with D$_2$O), 3.73 (d, J = 5.3 Hz, 2H), 3.19 (s, 2H exchangeable with D$_2$O), 3.11 (q, J = 6.2 Hz, 2H), 2.61 (t, J = 6.4 Hz, 2H). ESI m/z: 276 [M + H]$^+$.

Synthesis of ethyl 5-(3-(2-((2-aminoethyl)amino)-2-oxoethyl)ureido)-1H-indole-3-carboxylate 2,2,2-trifluoroacetate 171b
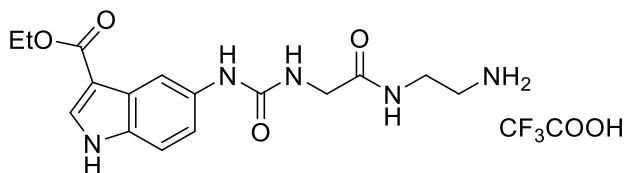
Compound 171 b was obtained as white solid (0.070 g, resa 85%) starting from **170b** (0.080 g, 0.178 mmol) according to the procedure described for **164b**; 188 – 190 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.75 (s, 1H exchangeable with D$_2$O), 8.73 (s, 1H exchangeable with D$_2$O), 8.10 (t, J = 5.1 Hz, 1H exchangeable with D$_2$O), 8.02 (s, 1H), 7.97 (d, J = 2.1 Hz, 1H), 7.69 (s, 2H exchangeable with D$_2$O), 7.33 (d, J = 8.7 Hz, 1H), 7.30 (dd, J = 8.7, 1.2 Hz, 1H), 6.29 (t, J = 5.1 Hz, 1H exchangeable with D$_2$O), 4.25 (q, J = 7.0 Hz, 2H), 3.76 (d, J = 5.2 Hz, 2H), 2.88 (t, J = 6.1 Hz, 2H), 1.32 (t, J = 7.1 Hz, 3H). ESI m/z: 348 [M + H]$^+$.

Synthesis of Ethyl 5-(3-(2-((2-aminoethyl)amino)-2-oxoethyl)ureido)-1H-indole-2-carboxylate 2,2,2-trifluoroacetate 171c
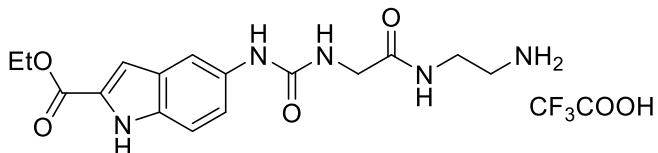
Compound 171c was obtained as white solid (0.050 g, resa 60%) starting from **170c** (0.080 g, 0.180 mmol) according to the procedure described for **164b**; mp 183 – 185 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.69 (s, 1H exchangeable with D$_2$O), 8.68 (s, 1H exchangeable with D$_2$O), 8.11 (s, 1H exchangeable with D$_2$O), 7.88 – 7.64 (m, 3H exchangeable with D$_2$O), 7.33 (d, $J = 8.6$ Hz, 1H), 7.18 (d, $J = 8.6$ Hz, 1H), 7.10 – 6.98 (m, 1H), 6.33 (t, $J = 4.8$ Hz, 1H exchangeable with D$_2$O), 4.32 (q, $J = 7.0$ Hz, 2H), 3.75 (d, $J = 4.4$ Hz, 2H), 2.96 – 2.81 (m, 2H), 1.33 (t, $J = 7.0$ Hz, 3H). ESI m/z: 348 [M + H]$^+$

Synthesis of Series III derivatives

Synthesis of 1-(2-hydroxyethyl)-3-(p-tolyl)urea 173a
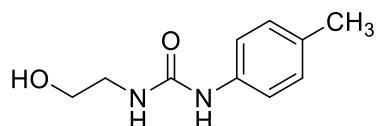
To a cooled solution of *p*-tolilisocyanate (0.94 mL, 7.51 mmol) and TEA (2.06 mL, 15.02 mmol) in THF dry was added dropwise a solution of 2-amino-ethanol (0.90 mL, 15.02 mmol) in THF dry (16 mL). The reaction was left to stir at room temperature for 12 h and then was concentrated under vacuum. The residue was dissolved in AcOEt (150 mL) and washed with HcL 1n (3 x 50mL) and brine.The organic phase was dried with Na$_2$SO$_4$ and evaporated under vacuum. The desired product **173a** was obtained as white solid (1.20 g,

resa 82%) after crystallization with AcOEt; mp 130 -132 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 8.42 (s, 1H, exchangeable with D$_2$O), 7.24 (d, J = 8.0 Hz, 2H), 7.01 (d, J = 8.0 Hz, 2H), 6.11 (t, J = 5.6 Hz, 1H, exchangeable with D$_2$O), 4.74 (t, J = 5.2 Hz, 1H), 3.42 (q, J = 5.5 Hz, 2H), 3.34 (s, 1H, exchangeable with D$_2$O), 2.20 (s, 3H). ESI m/z: 195 [M + H]$^+$

Synthesis of 1-(2-hydroxyethyl)-3-(2-(piperidin-1-yl)ethyl)urea 173b
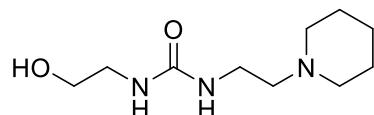
To a solution of N-(2aminoethyl)-piperidine (1.11 mL, 7.80 mmol) in AcOE (16 mL) and TEA (1.19 mL, 8.58 mmol) was added dropwise phenyl-chloroformate (1.2 mL, 8.58) and was left to stir at room temperature for 4h, Water (5 mL) was added, the two phase were separated and the organic phase was washed with water (2 x 5 mL) and Brine. The organic phase was dried, filtered and evaporated under vacuum. Purification by silica gel chromatography (DCM/MEOH 9:1) afford phenyl (2-(piperidin-1-yl)ethyl)carbamate intermediate (**172a**) as yellow oil (1.93 g, 52%). Compound **172a** was dissolved in 1, 2-dichloroethane (10 mL) and then were added TEA (0.56 mL, 4.05 mmol) and 2-amino-ethanol (0.24 mL, 4.05 mmol). The reaction was left to stirr at 120 °C for 2 h. The mixture of reaction was concentrated under reduced pressure and purified by silica gel chromatography (DCM/MeOH/NH$_3$, 9:1:0.1) to afford the desired product **173b** as yellow oil (0.226 g, 78%). ^{1}H NMR (500 MHz, Chloroform-d) δ 6.21 (s, 1H, exchangeable with D$_2$O), 5.24 (s, 1H, exchangeable with D$_2$O), 3.69 (t, J = 4.8 Hz, 2H), 3.33 (q, J = 5.2 Hz, 2H), 3.24 (q, J = 5.5 Hz, 2H), 2.44 (t, J = 5.5 Hz, 2H), 2.40 (m, 4H, exchangeable with D$_2$O), 1.56 (m, J = 5.7 Hz, 5H), 1.48 – 1.40 (m, 2H). ESI m/z: 216 [M + H]$^+$.

Synthesis of 1-(2-azidoethyl)-3-(2-hydroxyethyl)urea 173c
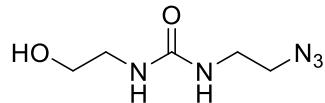
Compound **173c** was obtained as colourless oil (0.600 g, 72%) starting from 2-azidoethan-1-amine (1.00 g, 8.16 mmol) according to the procedures reported for **173b**. ^{1}H NMR (300 MHz, DMSO-d_6) δ 6.16 (t, J = 5.8 Hz, 1H, exchangeable with D_2O), 5.98 (t, J = 5.8 Hz, 1H, exchangeable with D_2O), 4.64 (t, J = 5.2 Hz, 1H, exchangeable with D_2O), 3.39 – 3.34 (m, 2H), 3.29 (d, J = 5.6 Hz, 1H), 3.17 (q, J = 5.8 Hz, 2H), 3.05 (q, J = 5.8 Hz, 2H). ESI m/z: 174 [M + H]$^+$.

Synthesis of 1-(2-methoxyethyl)-3-(p-tolyl)urea 174a
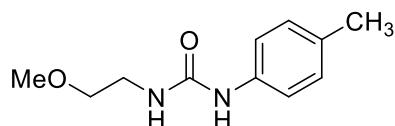
To a cooled solution of **173a** (0.500 g, 2.57 mmol) in THF dry (26 mL) were added dropwise TEA (0.53 mL, 3.86 mmol) and mesyl chloride (0.30 mL, 3.86 mmol). The reaction was left to stir at room temperature for 1 h and was directly used for preparation of **175a – c.**

Synthesis of 1-(2-methoxyethyl)-3-(2-(piperidin-1-yl)ethyl)urea 174b
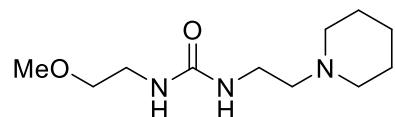
Compound **174b** was obtained starting from **173b** (0.150g, 0.70 mmol) according to the procedure used for **174a.** The mixture of reaction was directly used for the preparation of **175a – c.**

Synthesis of 1-(2-azidoethyl)-3-(2-methoxyethyl)urea 174c
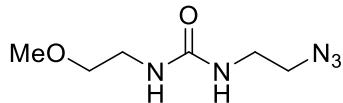
Compound **174c** was obtained starting from **173c** (0.150 g, 0.866 mmol) according to the procedures described for **174a.**

Synthesis of 1-(2-((1H-indol-5-yl)amino)ethyl)-3-(p-tolyl)urea 175a
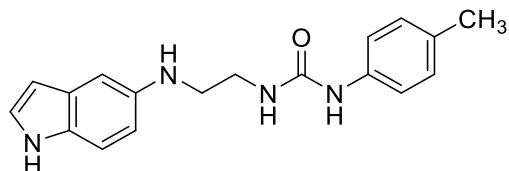
To a mixture of reaction of 174a were added 155a (0.255 g, 1.92 mmol) and TEA (0.26 mL, 1.92 mmol). The reaction was left to stir at 50 °C for 1 h. The reaction was concentrated under vacuum, the residue was dissolved with AcOEt (100 mL) and washed with brine (3 x 35 mL). The organic layer was dried over Na_2SO_4, filtered and evaporated under reduced pressure. Purification by silica gel chromatography (Hex/AcOEt, 1:1) gave the desired product 175a as white solid (0.150 g, 40%); mp 160- 162 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.60 (s, 1H, exchangeable with D_2O), 8.43 (s, 1H, exchangeable with D_2O), 7.27 (d, J = 8.0 Hz, 2H), 7.15 – 7.10 (m, 2H), 7.01 (d, J = 7.9 Hz, 2H), 6.66 (s, 1H), 6.56 – 6.51 (m, 1H), 6.21 (t, J = 5.4 Hz, 1H, exchangeable with D_2O), 6.17 (s, 1H,), 5.02 (t, J = 5.1 Hz, 1H, exchangeable with D_2O), 3.32 – 3.26 (m, 2H), 3.08 (q, J = 6.1 Hz, 2H), 2.21 (s, 3H). ESI m/z: 309 [M + H]$^+$

Synthesis of ethyl 5-((2-(3-(p-tolyl)ureido)ethyl)amino)-1H-indole-3-carboxylate 175b
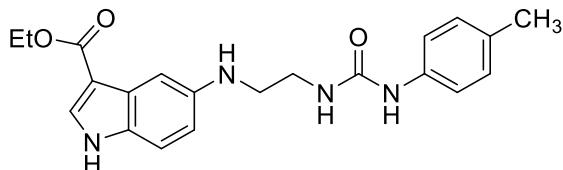
Compound **175b** was obtained as white solid (0.131 g, resa 53%) starting from **155b** (0.200g, 0.98 mmol) and **174a** (0.65 mol) according to the procedure reported for **175a;** 145 – 146 °C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.52 (s, 1H, exchangeable with D_2O), 8.47 (s, 1H, exchangeable with D_2O), 7.82 (s, 1H), 7.27 (d, J = 8.0 Hz, 2H), 7.20 (d, J = 8.7 Hz, 1H), 7.14 (d, J = 2.3 Hz, 1H), 7.01 (d, J = 8.0 Hz, 2H), 6.62 (dd, J = 8.9, 2.3 Hz, 1H), 6.28 – 6.22 (m, 1H, exchangeable with D_2O), 5.40 (t, J = 5.7 Hz, 1H, exchangeable with D_2O), 4.23 (q, J = 7.1 Hz, 2H), 3.31 (m, 2H)3.12 (q, J = 6.2 Hz, 2H), 2.21 (s, 3H), 1.31 (t, J = 7.1 Hz, 3H). ESI m/z: 381 [M + H]$^+$.

Synthesis of ethyl 5-((2-(3-(p-tolyl)ureido)ethyl)amino)-1H-indole-2-carboxylate 175c
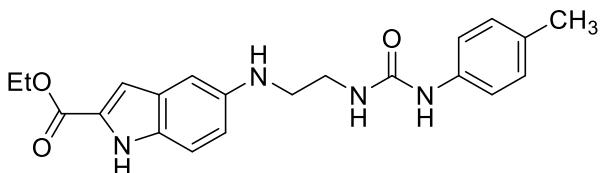
Compound **175b** was obtained as white solid (0.110g, resa 42%) starting from **155c** (0.200g, 0.98 mmol) and **174a** (0.65 mol) according to the procedure reported for **175a**, 130 – 132 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.44 (s, 1H, exchangeable with D_2O), 8.44 (s, 1H, exchangeable with D_2O), 7.27 (d, J = 8.1 Hz, 2H), 7.20 (d, J = 8.8 Hz, 1H), 7.02 (d, J = 8.1 Hz, 2H), 6.87 (s, 1H),

6.75 (dd, J = 8.1, 1.5 Hz, 1H), 6.65 (s, 1H), 6.21 (t, J = 5.4 Hz, 1H, exchangeable with D_2O), 5.33 (t, J = 6.1 Hz, 1H, exchangeable with D_2O), 4.30 (q, J = 7.1 Hz, 2H), 3.54 – 3.39 (m, 2H), 3.16 – 3.05 (m, 2H), 2.21 (s, 3H), 1.32 (t, J = 7.0 Hz, 3H). ESI m/z: 381 [M + H]$^+$.

Synthesis of 1-(2-((1H-indol-5-yl)amino)ethyl)-3-(2-(piperidin-1-yl)ethyl)urea 176a
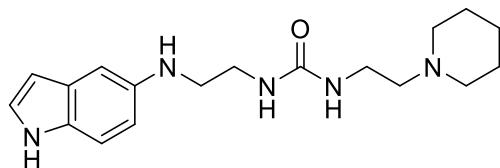
Compound **176a** was obtained as white solid (0.109 g, 40%) starting from **155a** (0.138 g, 1.05 mmol) and **146b** (0.7 mmol) according to the procedure reported for **175a**; 125 – 127 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.53 (s, 1H, exchangeable with D_2O), 7.77 (s, 1H, exchangeable with D_2O), 7.43 (d, J = 7.5 Hz, 1H), 7.34 – 7.22 (m, 2H), 6.69 – 6.51 (m, 3H, exchangeable with D_2O), 5.94 (s, 1H, exchangeable with D_2O), 3.58 (t, J = 4.8 Hz, 2H), 3.51 – 3.36 (m, 4H), 2.52 – 2.37 (m, 6H), 1.55 – 1.33 (m, 6H). ESI m/z: 275 [M + H]$^+$.

Synthesis of ethyl 5-((2-(3-(2-(piperidin-1-yl)ethyl)ureido)ethyl)amino)-1H-indole-3-carboxylate 176b
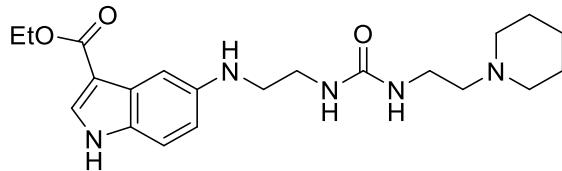
Compound **176b** was obtained as white solid (0.150 g, 54%) starting from **155b** (0.212 g, 1.04 mmol) and **174b** (0.69 mmol) according the procedure described for **175a**; mp 120 – 121 °C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.58

(s, 1H, exchangeable with D$_2$O), 9.89 (s, 1H, exchangeable with D$_2$O), 7.83 (s, 1H, exchangeable with D$_2$O), 7.18 (d, *J* = 28.9 Hz, 2H), 6.63 (s, 1H), 5.22 – 4.99 (m, 1H, exchangeable with D$_2$O), 4.23 (q, *J* = 7.1 Hz, 2H), 3.10 – 3.02 (m, 8H), 1.31 (t, *J* = 7.1 Hz, 3H), 1.21 – 1.14 (m, 10H). ESI m/z: 402 [M + H]$^+$.

Synthesis of ethyl 5-((2-(3-(2-(piperidin-1-yl)ethyl)ureido)ethyl)amino)-1H-indole-2-carboxylate 176c
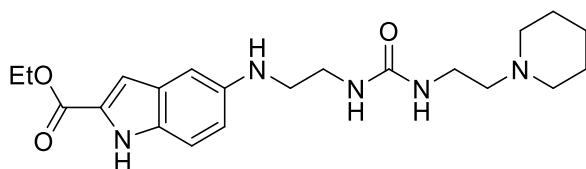
Compound **176c** was obtained as white solid (0.100 g, 57%) starting from **155c** (0.133 g, 0.65 mmol) and **174b** (0.43 mmol) according the procedure described for **175a**; mp 145 – 147 °C. ^{1}H NMR (300 MHz, DMSO-*d$_6$*) δ 11.45 (s, 1H, exchangeable with D$_2$O), 9.27 (s, 1H, exchangeable with D$_2$O), 6.87 (s, 1H, exchangeable with D$_2$O), 6.72 (d, *J* = 8.9 Hz, 1H), 6.64 (s, 1H), 6.43 – 6.31 (m, 2H), 5.33 – 5.22 (m, 1H, exchangeable with D$_2$O), 4.30 (q, *J* = 7.0 Hz, 2H), 3.55 – 3.42 (m, 3H), 3.30 – 3.18 (m, 3H), 3.14 – 3.02 (m, 6H), 2.96 – 2.80 (m, 2H), 1.88 – 1.73 (m, 2H), 1.73 – 1.53 (m, 2H), 1.32 (t, *J* = 7.0 Hz, 3H). ESI m/z: 402 [M + H]$^+$

ACKNOWLEDGMENTS

References

1. Luger, K.; Mader, A. W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J., Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **1997,** *389* (6648), 251-60.

2. Srahl, B. O.; Allis, C. D., The language of covalent histone modifications. *Nature (London)* **2000,** *403* (6765), 41-45.

3. Ruthenburg, A. J.; Allis, C. D.; Wysocka, J., Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol. Cell* **2007,** *25* (1), 15-30.

4. Roth, S. Y.; Denu, J. M.; Allis, C. D., Histone acetyltransferases. *Annu. Rev. Biochem.* **2001,** *70*, 81-120.

5. Nowak, S. J.; Corces, V. G., Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet.* **2004,** *20* (4), 214-220.

6. Gatti, M.; Pinato, S.; Maspero, E.; Soffientini, P.; Polo, S.; Penengo, L., A novel ubiquitin mark at the N-terminal tail of histone H2As targeted by RNF168 ubiquitin ligase. *Cell Cycle* **2012,** *11* (13), 2538-2544.

7. Robertson, K. D., DNA methylation, methyltransferases, and cancer. *Oncogene* **2001,** *20* (24), 3139-3155.

8. de Ruijter, A. J. M.; van Gennip, A. H.; Caron, H. N.; Kemp, S.; van Kuilenburg, A. B. P., Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochem. J.* **2003,** *370* (3), 737-749.

9. Bannister, A. J.; Kouzarides, T., Regulation of chromatin by histone modifications. *Cell Res.* **2011,** *21* (3), 381-395.

10. Honkanen, R. E.; Golden, T., Regulators of serine/threonine protein phosphatases at the dawn of a clinical era? *Curr. Med. Chem.* **2002,** *9* (22), 2055-2075.

11. Shi, Y., Histone lysine demethylases: emerging roles in development, physiology and disease. *Nat. Rev. Genet.* **2007,** *8* (11), 829-833.

12. Musselman, C. A.; Lalonde, M.-E.; Cote, J.; Kutateladze, T. G., Perceiving the epigenetic landscape through histone readers. *Nat. Struct. Mol. Biol.* **2012,** *19* (12), 1218-1227.

13. Martin, C.; Zhang, Y., The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* **2005,** *6* (11), 838-849.

14. Shi, X.; Kachirskaia, I.; Yamaguchi, H.; West, L. E.; Wen, H.; Wang, E. W.; Dutta, S.; Appella, E.; Gozani, O., Modulation of p53 function by SET8-mediated methylation at lysine 382. *Mol. Cell* **2007,** *27* (4), 636-646.

15. Chuikov, S.; Kurash, J. K.; Wilson, J. R.; Xiao, B.; Justin, N.; Ivanov, G. S.; McKinney, K.; Tempst, P.; Prives, C.; Gamblin, S. J.; Barlev, N. A.; Reinberg, D., Regulation of p53 activity through lysine methylation. *Nature (London, U. K.)* **2004,** *432* (7015), 353-360.

16. Yang, Y.; Bedford, M. T., Protein arginine methyltransferases and cancer. *Nat. Rev. Cancer* **2013,** *13* (1), 37-50.

17.   Copeland, R. A.; Solomon, M. E.; Richon, V. M., Protein methyltransferases as a target class for drug discovery. *Nat. Rev. Drug Discovery* **2009,** *8* (9), 724-732.

18.   Min, J.; Feng, Q.; Li, Z.; Zhang, Y.; Xu, R.-M., Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. *Cell* **2003,** *112* (5), 711-23.

19.   Anglin, J. L.; Song, Y., A Medicinal Chemistry Perspective for Targeting Histone H3 Lysine-79 Methyltransferase DOT1L. *J. Med. Chem.* **2013,** *56* (22), 8972-8983.

20.   Dillon, S. C.; Zhang, X.; Trievel, R. C.; Cheng, X., The SET-domain protein superfamily: Protein lysine methyltransferases. *GenomeBiology* **2005,** *6* (8), No pp. given.

21.   Min, J.; Zhang, X.; Cheng, X.; Grewal, S. I. S.; Xu, R.-M., Structure of the SET domain histone lysine methyltransferase Clr4. *Nat. Struct. Biol.* **2002,** *9* (11), 828-832.

22.   Schapira, M., Structural chemistry of human SET domain protein methyltransferases. *Curr. Chem. Genomics* **2011,** *5*, 85-94.

23.   Collins, R. E.; Tachibana, M.; Tamaru, H.; Smith, K. M.; Jia, D.; Zhang, X.; Selker, E. U.; Shinkai, Y.; Cheng, X., In Vitro and in Vivo Analyses of a Phe/Tyr Switch Controlling Product Specificity of Histone Lysine Methyltransferases. *J. Biol. Chem.* **2005,** *280* (7), 5563-5570.

24.   Rathert, P.; Dhayalan, A.; Murakami, M.; Zhang, X.; Tamas, R.; Jurkowska, R.; Komatsu, Y.; Shinkai, Y.; Cheng, X.; Jeltsch, A., Protein lysine methyltransferase G9a acts on non-histone targets. *Nat. Chem. Biol.* **2008,** *4* (6), 344-346.

25.   Van Duyne, R.; Easley, R.; Wu, W.; Berro, R.; Pedati, C.; Klase, Z.; Kehn-Hall, K.; Flynn, E. K.; Symer, D. E.; Kashanchi, F., Lysine methylation of HIV-1 Tat regulates transcriptional activity of the viral LTR. *Retrovirology* **2008,** *5*, No pp. given.

26.   Pagans, S.; Kauder, S. E.; Kaehlcke, K.; Sakane, N.; Schroeder, S.; Dormeyer, W.; Trievel, R. C.; Verdin, E.; Schnolzer, M.; Ott, M., The cellular lysine methyltransferase Set7/9-KMT7 binds HIV-1 TAR RNA, monomethylates the viral transactivator Tat, and enhances HIV transcription. *Cell Host Microbe* **2010,** *7* (3), 234-244.

27.   Okada, Y.; Jiang, Q.; Lemieux, M.; Jeannotte, L.; Su, L.; Zhang, Y., Leukaemic transformation by CALM-AF10 involves upregulation of Hoxa5 by hDOT1L. [Erratum to document cited in CA145:332773]. *Nat. Cell Biol.* **2006,** *8* (10), 1178.

28.   Huang, J.; Dorsey, J.; Chuikov, S.; Zhang, X.; Jenuwein, T.; Reinberg, D.; Berger, S. L., G9a and Glp Methylate Lysine 373 in the Tumor Suppressor p53. *J. Biol. Chem.* **2010,** *285* (13), 9636-9641.

29.   Mosammaparast, N.; Shi, Y., Reversal of histone methylation: biochemical and molecular mechanisms of histone demethylases. *Annu. Rev. Biochem.* **2010,** *79*, 155-179.

30.   Arrowsmith, C. H.; Bountra, C.; Fish, P. V.; Lee, K.; Schapira, M., Epigenetic protein families: a new frontier for drug discovery. *Nat. Rev. Drug Discovery* **2012,** *11* (5), 384-400.

31.   Krivtsov, A. V.; Armstrong, S. A., MLL translocations, histone modifications and leukaemia stem-cell development. *Nat. Rev. Cancer* **2007,** *7* (11), 823-833.

32.   Ayton, P. M.; Cleary, M. L., Transformation of myeloid progenitors by MLL oncoproteins is dependent on Hoxa7 and Hoxa9. *Genes Dev.* **2003,** *17* (18), 2298-2307.

33.   Li, Z.; Luo, R. T.; Mi, S.; Sun, M.; Chen, P.; Bao, J.; Neilly, M. B.; Jayathilaka, N.; Johnson, D. S.; Wang, L.; Lavau, C.; Zhang, Y.; Tseng, C.; Zhang, X.; Wang, J.; Yu, J.; Yang, H.; Wang, S. M.; Rowley, J. D.; Chen, J.; Thirman, M. J., Consistent Deregulation of Gene Expression between Human and Murine MLL Rearrangement Leukemias. *Cancer Res.* **2009,** *69* (3), 1109-1116.

34.   Okada, Y.; Feng, Q.; Lin, Y.; Jiang, Q.; Li, Y.; Coffield, V. M.; Su, L.; Xu, G.; Zhang, Y., hDOT1L links histone methylation to leukemogenesis. *Cell (Cambridge, MA, U. S.)* **2005,** *121* (2), 167-178.

35.   Zhang, X.; Cheng, X., Structure of the Predominant Protein Arginine Methyltransferase PRMT1 and Analysis of Its Binding to Substrate Peptides. *Structure (Cambridge, MA, U. S.)* **2003,** *11* (5), 509-520.

36.   Chen, X.; El Gazzar, M.; Yoza, B. K.; McCall, C. E., The NF-κB Factor RelB and Histone H3 Lysine Methyltransferase G9a Directly Interact to Generate Epigenetic Silencing in Endotoxin Tolerance. *J. Biol. Chem.* **2009,** *284* (41), 27857-27865.

37.   Meissner, A., Epigenetic modifications in pluripotent and differentiated cells. *Nat. Biotechnol.* **2010,** *28* (10), 1079-1088.

38.   Shi, Y.; Do, J. T.; Desponts, C.; Hahm, H. S.; Scholer, H. R.; Ding, S., A combined chemical and genetic approach for the generation of induced pluripotent stem cells. *Cell Stem Cell* **2008,** *2* (6), 525-528.

39.   Chen, D.; Han, M.; Hong, H.; Koh, S. S.; Huang, S.-M.; Schurter, B. T.; Aswad, D. W.; Stallcup, M. R., Regulation of transcription by a protein methyltransferase. *Science (Washington, D. C.)* **1999,** *284* (5423), 2174-2177.

40.   Bedford, M. T.; Clarke, S. G., Protein arginine methylation in mammals: who, what, and why. *Mol. Cell* **2009,** *33* (1), 1-13.

41.   Wysocka, J.; Allis, C. D.; Coonrod, S., Histone arginine methylation and its dynamic regulation. *Front. Biosci.* **2006,** *11* (1), 344-355.

42.   Boffa, L. C.; Karn, J.; Vidali, G.; Allfrey, V. G., Distribution of NG,NG-dimethylarginine in nuclear protein fractions. *Biochem. Biophys. Res. Commun.* **1977,** *74* (3), 969-76.

43.   Gary, J. D.; Clarke, S., RNA and protein interactions modulated by protein arginine methylation. *Prog. Nucleic Acid Res. Mol. Biol.* **1998,** *61*, 65-131.

44.   Zurita-Lopez, C. I.; Sandberg, T.; Kelly, R.; Clarke, S. G., Human Protein Arginine Methyltransferase 7 (PRMT7) Is a Type III Enzyme Forming ω-NG-Monomethylated Arginine Residues. *J. Biol. Chem.* **2012,** *287* (11), 7859-7870.

45.   Yang, Y.; Hadjikyriacou, A.; Xia, Z.; Gayatri, S.; Kim, D.; Zurita-Lopez, C.; Kelly, R.; Guo, A.; Li, W.; Clarke, S. G.; Bedford, M. T., PRMT9 is a Type II methyltransferase that methylates the splicing factor SAP145. *Nat. Commun.* **2015,** *6*, 6428.

46.   Zhang, X.; Bolt, M.; Guertin, M. J.; Chen, W.; Zhang, S.; Cherrington, B. D.; Slade, D. J.; Dreyton, C. J.; Subramanian, V.; Bicker, K. L.; Thompson, P. R.; Mancini, M. A.; Lis, J. T.; Coonrod, S. A., Peptidylarginine deiminase 2-catalyzed histone H3 arginine 26 citrullination facilitates estrogen receptor α target gene activation. *Proc. Natl. Acad. Sci. U. S. A.* **2012,** *109* (33), 13331-13336, S13331/1-S13331/10.

47.   Weiss, V. H.; McBride, A. E.; Soriano, M. A.; Filman, D. J.; Silver, P. A.; Hogle, J. M., The structure and oligomerization of the yeast arginine methyltransferase, Hmt1. *Nat. Struct. Biol.* **2000,** *7* (12), 1165-1171.

48.   Yue, W. W.; Hassler, M.; Roe, S. M.; Thompson-Vale, V.; Pearl, L. H., Insights into histone code syntax from structural and biochemical studies of CARM1 methyltransferase. *EMBO J.* **2007,** *26* (20), 4402-4412.

49.   Troffer-Charlier, N.; Cura, V.; Hassenboehler, P.; Moras, D.; Cavarelli, J., Functional insights from structures of coactivator-associated arginine methyltransferase 1 domains. *EMBO J.* **2007,** *26* (20), 4391-4401.

50.   Schubert, H. L.; Blumenthal, R. M.; Cheng, X., Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem. Sci.* **2003,** *28* (6), 329-335.

51.   Cheng, X.; Collins, R. E.; Zhang, X., Structural and sequence motifs of protein (histone) methylation enzymes. *Annu. Rev. Biophys. Biomol. Struct.* **2005,** *34*, 267-294.

52.   Cura, V.; Troffer-Charlier, N.; Wurtz, J.-M.; Bonnefond, L.; Cavarelli, J., Structural insight into arginine methylation by the mouse protein arginine methyltransferase 7: a zinc finger freezes the mimic of the dimeric state into a single active site. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2014,** *70* (9), 2401-2412.

53.   Antonysamy, S.; Bonday, Z.; Campbell, R. M.; Doyle, B.; Druzina, Z.; Gheyi, T.; Han, B.; Jungheim, L. N.; Qian, Y.; Rauch, C.; Russell, M.; Sauder, J. M.; Wasserman, S. R.; Wichert, K.; Willard, F. S.; Zhang, A.; Emtage, S., Crystal structure of the human PRMT5:MEP50 complex. *Proc. Natl. Acad. Sci. U. S. A.* **2012,** *109* (44), 17960-17965, S17960/1-S17960/9.

54.   Sun, L.; Wang, M.; Lv, Z.; Yang, N.; Liu, Y.; Bao, S.; Gong, W.; Xu, R.-M., Structural insights into protein arginine symmetric dimethylation by PRMT5. *Proc. Natl. Acad. Sci. U. S. A.* **2011,** *108* (51), 20538-20543, S20538/1-S20538/3.

55.   Wang, H.; Huang, Z.-Q.; Xiap, L.; Feng, Q.; Erdjument-Bromage, H.; Strahl, B. D.; Briggs, S. D.; Allis, C. D.; Wong, J.; Tempst, P.; Zhang, Y., Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science (Washington, DC, U. S.)* **2001,** *293* (5531), 853-856.

56.   Lin, W.-J.; Gary, J. D.; Yang, M. C.; Clarke, S.; Herschman, H. R., The mammalian immediate-early TIS21 protein and the leukemia-associated BTG1 protein interact with a protein-arginine N-methyltransferase. *J. Biol. Chem.* **1996,** *271* (25), 15034-15044.

57.   Tang, J.; Frankel, A.; Cook, R. J.; Kim, S.; Paik, W. K.; Williams, K. R.; Clarke, S.; Herschman, H. R., PRMT1 is the predominant type I protein arginine methyltransferase in mammalian cells. *J. Biol. Chem.* **2000,** *275* (11), 7723-7730.

58.   Pawlak, M. R.; Scherer, C. A.; Chen, J.; Roshon, M. J.; Ruley, H. E., Arginine N-methyltransferase 1 is required for early postimplantation mouse development, but cells deficient in the enzyme are viable. *Mol. Cell. Biol.* **2000,** *20* (13), 4859-4869.

59.   Goulet, I.; Gauvin, G.; Boisvenue, S.; Cote, J., Alternative Splicing Yields Protein Arginine Methyltransferase 1 Isoforms with Distinct Activity, Substrate Specificity, and Subcellular Localization. *J. Biol. Chem.* **2007,** *282* (45), 33009-33021.

60.   Pahlich, S.; Zakaryan, R. P.; Gehring, H., Protein arginine methylation: Cellular functions and methods of analysis. *Biochim. Biophys. Acta, Proteins Proteomics* **2006,** *1764* (12), 1890-1903.

61.   Yu, Z.; Chen, T.; Hebert, J.; Li, E.; Richard, S., A mouse PRMT1 null allele defines an essential role for arginine methylation in genome maintenance and cell proliferation. *Mol. Cell. Biol.* **2009,** *29* (11), 2982-2996.

62.   Seligson, D. B.; Horvath, S.; Shi, T.; Yu, H.; Tze, S.; Grunstein, M.; Kurdistani, S. K., Global histone modification patterns predict risk of prostate cancer recurrence. *Nature (London, U. K.)* **2005,** *435* (7046), 1262-1266.

63.   Yang, Y.-Z.; Lu, Y.; Espejo, A.; Wu, J.-C.; Xu, W.; Liang, S.-D.; Bedford, M. T., TDRD3 is an Effector Molecule for Arginine-Methylated Histone Marks. *Mol. Cell* **2010,** *40* (6), 1016-1023.

64.   Nagahata, T.; Onda, M.; Emi, M.; Nagai, H.; Tsumagari, K.; Fujimoto, T.; Hirano, A.; Sato, T.; Nishikawa, K.; Akiyama, F.; Sakamoto, G.; Kasumi, F.; Miki, Y.; Tanaka, T.; Tsunoda, T., Expression profiling to predict postoperative prognosis for estrogen receptor-negative breast cancers by analysis of 25,344 genes on a cDNA microarray. *Cancer Sci.* **2004,** *95* (3), 218-225.

65. Yu, Z.; Vogel, G.; Coulombe, Y.; Dubeau, D.; Spehalski, E.; Hebert, J.; Ferguson, D. O.; Masson, J. Y.; Richard, S., The MRE11 GAR motif regulates DNA double-strand break processing and ATR activation. *Cell Res.* **2012,** *22* (2), 305-320.

66. Boisvert, F.-M.; Rhie, A.; Richard, S.; Doherty, A. J., The GAR motif of 53BP1 is arginine methylated by PRMT1 and is necessary for 53BP1 DNA binding activity. *Cell Cycle* **2005,** *4* (12), 1834-1841.

67. Le Romancer, M.; Treilleux, I.; Leconte, N.; Robin-Lespinasse, Y.; Sentis, S.; Bouchekioua-Bouzaghou, K.; Goddard, S.; Gobert-Gosse, S.; Corbo, L., Regulation of estrogen rapid signaling through arginine methylation by PRMT1. *Mol. Cell* **2008,** *31* (2), 212-221.

68. Cheung, N.; Chan, L. C.; Thompson, A.; Cleary, M. L.; So, C. W. E., Protein arginine-methyltransferase-dependent oncogenesis. *Nat. Cell Biol.* **2007,** *9* (10), 1208-1215.

69. Meyer, R.; Wolf, S. S.; Obendorf, M., PRMT2, a member of the protein arginine methyltransferase family, is a coactivator of the androgen receptor. *J. Steroid Biochem. Mol. Biol.* **2007,** *107* (1-2), 1-14.

70. Wolf, S. S., The protein arginine methyltransferase family: An update about function, new perspectives and the physiological role in humans. *Cell. Mol. Life Sci.* **2009,** *66* (13), 2109-2121.

71. Lakowski, T. M.; Frankel, A., A Kinetic Study of Human Protein Arginine N-Methyltransferase 6 Reveals a Distributive Mechanism. *J. Biol. Chem.* **2008,** *283* (15), 10015-10025.

72. Sayegh, J.; Webb, K.; Cheng, D.; Bedford, M. T.; Clarke, S. G., Regulation of Protein Arginine Methyltransferase 8 (PRMT8) Activity by Its N-terminal Domain. *J. Biol. Chem.* **2007,** *282* (50), 36444-36453.

73. Iwasaki, H.; Kovacic, J. C.; Olive, M.; Beers, J. K.; Yoshimoto, T.; Crook, M. F.; Tonelli, L. H.; Nabel, E. G., Disruption of Protein Arginine N-Methyltransferase 2 Regulates Leptin Signaling and Produces Leanness In Vivo Through Loss of STAT3 Methylation. *Circ. Res.* **2010,** *107* (8), 992-1001.

74. Qi, C.; Chang, J.; Zhu, Y.; Yeldandi, A. V.; Rao, S. M.; Zhu, Y.-J., Identification of protein arginine methyltransferase 2 as a coactivator for estrogen receptor α. *J. Biol. Chem.* **2002,** *277* (32), 28624-28630.

75. Yoshimoto, T.; Boehm, M.; Olive, M.; Crook, M. F.; San, H.; Langenickel, T.; Nabel, E. G., The arginine methyltransferase PRMT2 binds RB and regulates E2F function. *Exp. Cell Res.* **2006,** *312* (11), 2040-2053.

76. Kzhyshkowska, J.; Schutt, H.; Liss, M.; Kremmer, E.; Stauber, R.; Wolf, H.; Dobner, T., Heterogeneous nuclear ribonucleoprotein E1B-AP5 is methylated in its Arg-Gly-Gly (RGG) box and interacts with human arginine methyltransferase HRMT1L1. *Biochem. J.* **2001,** *358* (2), 305-314.

77.   Swiercz, R.; Cheng, D.; Kim, D.; Bedford, M. T., Ribosomal Protein rpS2 Is Hypomethylated in PRMT3-deficient Mice. *J. Biol. Chem.* **2007,** *282* (23), 16917-16923.

78.   Lai, Y.; Song, M.; Hakala, K.; Weintraub, S. T.; Shiio, Y., Proteomic Dissection of the von Hippel-Lindau (VHL) Interactome. *J. Proteome Res.* **2011,** *10* (11), 5175-5182.

79.   Singh, V.; Miranda, T. B.; Jiang, W.; Frankel, A.; Roemer, M. E.; Robb, V. A.; Gutmann, D. H.; Herschman, H. R.; Clarke, S.; Newsham, I. F., DAL-1/4.1B tumor suppressor interacts with protein arginine N-methyltransferase 3 (PRMT3) and inhibits its ability to methylate substrates in vitro and in vivo. *Oncogene* **2004,** *23* (47), 7761-7771.

80.   Alexiou, G. A.; Markoula, S.; Gogou, P.; Kyritsis, A. P., Genetic and molecular alterations in meningiomas. *Clin Neurol Neurosurg* **2011,** *113* (4), 261-7.

81.   Fronz, K.; Otto, S.; Koelbel, K.; Kuehn, U.; Friedrich, H.; Schierhorn, A.; Beck-Sickinger, A. G.; Ostareck-Lederer, A.; Wahle, E., Promiscuous Modification of the Nuclear Poly(A)-binding Protein by Multiple Protein-arginine Methyltransferases Does Not Affect the Aggregation Behavior. *J. Biol. Chem.* **2008,** *283* (29), 20408-20420.

82.   Brais, B.; Bouchard, J.-P.; Xie, Y.-G.; Rochefort, D. L.; Chretien, N.; Tome, F. M. S.; Lafreniere, R. G.; Rommens, J. M.; Uyama, E.; Nohira, O.; Blumen, S.; Korczyn, A. D.; Heutink, P.; Mathieu, J.; Duranceau, A.; Codere, F.; Fardeau, M.; Rouleau, G. A., Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. [Erratum to document cited in CA128:176812]. *Nat. Genet.* **1998,** *19* (4), 404.

83.   Chen, X.; Niroomand, F.; Liu, Z.; Zankl, A.; Katus, H. A.; Jahn, L.; Tiefenbacher, C. P., Expression of nitric oxide related enzymes in coronary heart disease. *Basic Res. Cardiol.* **2006,** *101* (4), 346-353.

84.   Cheng, D.; Cote, J.; Shaaban, S.; Bedford, M. T., The arginine methyltransferase CARM1 regulates the coupling of transcription and mRNA processing. *Mol. Cell* **2007,** *25* (1), 71-83.

85.   Ohkura, N.; Takahashi, M.; Yaguchi, H.; Nagamura, Y.; Tsukada, T., Coactivator-associated Arginine Methyltransferase 1, CARM1, Affects Pre-mRNA Splicing in an Isoform-specific Manner. *J. Biol. Chem.* **2005,** *280* (32), 28927-28935.

86.   El Messaoudi, S.; Fabbrizio, E.; Rodriguez, C.; Chuchana, P.; Fauquier, L.; Cheng, D.; Theillet, C.; Vandel, L.; Bedford, M. T.; Sardet, C., Coactivator-associated arginine methyltransferase 1 (CARM1) is a positive regulator of the cyclin E1 gene. *Proc. Natl. Acad. Sci. U. S. A.* **2006,** *103* (36), 13351-13356.

87.   Lee, Y.-H.; Stallcup, M. R., Roles of protein arginine methylation in DNA damage signaling pathways: is CARM1 a life-or-death decision point? *Cell Cycle* **2011,** *10* (9), 1343-1344.

88.   Lupien, M.; Eeckhoute, J.; Meyer, C. A.; Krum, S. A.; Rhodes, D. R.; Liu, X. S.; Brown, M., Coactivator function defines the active estrogen receptor alpha cistrome. *Mol. Cell. Biol.* **2009,** *29* (12), 3413-3423.

89.   Zhao, H.-Y.; Zhang, Y.-J.; Dai, H.; Zhang, Y.; Shen, Y.-F., CARM1 mediates modulation of Sox2. *PLoS One* **2011,** *6* (10), e27026.

90.   Ceschin, D. G.; Walia, M.; Wenk, S. S.; Duboe, C.; Gaudon, C.; Xiao, Y.; Fauquier, L.; Sankar, M.; Vandel, L.; Gronemeyer, H., Methylation specifies distinct estrogen-induced binding site repertoires of CBP to chromatin. *Genes Dev.* **2011,** *25* (11), 1132-1146.

91.   Naeem, H.; Cheng, D.; Zhao, Q.; Underhill, C.; Tini, M.; Bedford, M. T.; Torchia, J., The activity and stability of the transcriptional coactivator p/CIP/SRC-3 are regulated by CARM1-dependent methylation. *Mol. Cell. Biol.* **2007,** *27* (1), 120-134.

92.   Feng, Q.; He, B.; Jung, S.-Y.; Song, Y.; Qin, J.; Tsai, S. Y.; Tsai, M.-J.; O'Malley, B. W., Biochemical Control of CARM1 Enzymatic Activity by Phosphorylation. *J. Biol. Chem.* **2009,** *284* (52), 36167-36174.

93.   Chevillard-Briet, M.; Trouche, D.; Vandel, L., Control of CBP co-activating activity by arginine methylation. *EMBO J.* **2002,** *21* (20), 5457-5466.

94.   Sims, R. J., III; Rojas, L. A.; Beck, D.; Bonasio, R.; Schueller, R.; Drury, W. J., III; Eick, D.; Reinberg, D., The C-Terminal Domain of RNA Polymerase II Is Modified by Site-Specific Methylation. *Science (Washington, DC, U. S.)* **2011,** *332* (6025), 99-103.

95.   Hong, H.; Kao, C.; Jeng, M.-H.; Eble, J. N.; Koch, M. O.; Gardner, T. A.; Zhang, S.; Li, L.; Pan, C.-X.; Hu, Z.; MacLennan, G. T.; Cheng, L., Aberrant expression of CARM1, a transcriptional coactivator of androgen receptor, in the development of prostate carcinoma and androgen-independent status. *Cancer (N. Y., NY, U. S.)* **2004,** *101* (1), 83-89.

96.   Majumder, S.; Liu, Y.; Ford, O. H., III; Mohler, J. L.; Whang, Y. E., Involvement of arginine methyltransferase CARM1 in androgen receptor function and prostate cancer cell viability. *Prostate (Hoboken, NJ, U. S.)* **2006,** *66* (12), 1292-1301.

97.   Al-Dhaheri, M.; Wu, J.; Skliris, G. P.; Li, J.; Higashimato, K.; Wang, Y.; White, K. P.; Lambert, P.; Zhu, Y.; Murphy, L.; Xu, W., CARM1 Is an Important Determinant of ERα-Dependent Breast Cancer Cell Differentiation and Proliferation in Breast Cancer Cells. *Cancer Res.* **2011,** *71* (6), 2118-2128.

98.   Lahusen, T.; Henke, R. T.; Kagan, B. L.; Wellstein, A.; Riegel, A. T., The role and regulation of the nuclear receptor co-activator AIB1 in breast cancer. *Breast Cancer Res. Treat.* **2009,** *116* (2), 225-237.

99.   Frietze, S.; Lupien, M.; Silver, P. A.; Brown, M., CARM1 regulates estrogen-stimulated breast cancer growth through up-regulation of E2F1. *Cancer Res.* **2008,** *68* (1), 301-306.

100. Xu, W.; Chen, H.; Du, K.; Asahara, H.; Tini, M.; Emerson, B. M.; Montminy, M.; Evans, R. M., A transcriptional switch mediated by cofactor methylation. *Science (Washington, DC, U. S.)* **2001,** *294* (5551), 2507-2511.

101. Ou, C.-Y.; LaBonte, M. J.; Manegold, P. C.; So, A. Y.-L.; Ianculescu, I.; Gerke, D. S.; Yamamoto, K. R.; Ladner, R. D.; Kahn, M.; Kim, J. H.; Stallcup, M. R., A Coactivator Role of CARM1 in the Dysregulation of β-Catenin Activity in Colorectal Cancer Cell Growth and Gene Expression. *Mol. Cancer Res.* **2011,** *9* (5), 660-670.

102. Kim, J.; Lee, J.; Yadav, N.; Wu, Q.; Carter, C.; Richard, S.; Richie, E.; Bedford, M. T., Loss of CARM1 Results in Hypomethylation of Thymocyte Cyclic AMP-regulated Phosphoprotein and Deregulated Early T Cell Development. *J. Biol. Chem.* **2004,** *279* (24), 25339-25344.

103. Zhou, Z.; Sun, X.; Zou, Z.; Sun, L.; Zhang, T.; Guo, S.; Wen, Y.; Liu, L.; Wang, Y.; Qin, J.; Li, L.; Gong, W.; Bao, S., PRMT5 regulates Golgi apparatus structure through methylation of the golgin GM130. *Cell Res.* **2010,** *20* (9), 1023-1033.

104. Friesen, W. J.; Paushkin, S.; Wyce, A.; Massenet, S.; Pesiridis, G. S.; Van Duyne, G.; Rappsilber, J.; Mann, M.; Dreyfuss, G., The methylosome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. *Mol. Cell. Biol.* **2001,** *21* (24), 8289-8300.

105. Karkhanis, V.; Hu, Y.-J.; Baiocchi, R. A.; Imbalzano, A. N.; Sif, S., Versatility of PRMT5-induced methylation in growth control and development. *Trends Biochem. Sci.* **2011,** *36* (12), 633-641.

106. Pal, S.; Baiocchi, R. A.; Byrd, J. C.; Grever, M. R.; Jacob, S. T.; Sif, S., Low levels of miR-92b/96 induce PRMT5 translation and H3R8/H4R3 methylation in mantle cell lymphoma. *EMBO J.* **2007,** *26* (15), 3558-3569.

107. Hou, Z.; Peng, H.; Ayyanathan, K.; Yan, K.-P.; Langer, E. M.; Longmore, G. D.; Rauscher, F. J., III, The LIM protein AJUBA recruits protein arginine methyltransferase 5 to mediate SNAIL-dependent transcriptional repression. *Mol. Cell. Biol.* **2008,** *28* (10), 3198-3207.

108. Lankat-Buttgereit, B.; Goeke, R., The tumour suppressor Pdcd4: recent advances in the elucidation of function and regulation. *Biol. Cell* **2009,** *101* (6), 309-317.

109. Powers, M. A.; Fay, M. M.; Factor, R. E.; Welm, A. L.; Ullman, K. S., Protein arginine methyltransferase 5 accelerates tumor growth by arginine methylation of the tumor suppressor programmed cell death 4. *Cancer Res.* **2011,** *71* (16), 5579-5587.

110. Abramovich, C.; Yakobson, B.; Chebath, J.; Revel, M., A protein-arginine methyltransferase binds to the intracytoplasmic domain of the IFNAR1 chain in the type I interferon receptor. *EMBO J.* **1997,** *16* (2), 260-266.

111. Liu, F.; Zhao, X.; Perna, F.; Wang, L.; Koppikar, P.; Abdel-Wahab, O.; Harr, M. W.; Levine, R. L.; Xu, H.; Tefferi, A.; Deblasio, A.; Hatlen, M.;

Menendez, S.; Nimer, S. D., JAK2V617F-Mediated Phosphorylation of PRMT5 Downregulates Its Methyltransferase Activity and Promotes Myeloproliferation. *Cancer Cell* **2011,** *19* (2), 283-294.

112.   Aggarwal, P.; Vaites, L. P.; Kim, J. K.; Mellert, H.; Gurung, B.; Nakagawa, H.; Herlyn, M.; Hua, X.; Rustgi, A. K.; McMahon, S. B.; Diehl, J. A., Nuclear Cyclin D1/CDK4 Kinase Regulates CUL4 Expression and Triggers Neoplastic Growth via Activation of the PRMT5 Methyltransferase. *Cancer Cell* **2010,** *18* (4), 329-340.

113.   Stein, C.; Riedl, S.; Ruethnick, D.; Noetzold, R. R.; Bauer, U.-M., The arginine methyltransferase PRMT6 regulates cell proliferation and senescence through transcriptional repression of tumor suppressor genes. *Nucleic Acids Res.* **2012,** *40* (19), 9522-9533.

114.   Hyllus, D.; Stein, C.; Schnabel, K.; Schiltz, E.; Imhof, A.; Dou, Y.; Hsieh, J.; Bauer, U.-M., PRMT6-mediated methylation of R2 in histone H3 antagonizes H3 K4 trimethylation. *Genes Dev.* **2007,** *21* (24), 3369-3380.

115.   Guccione, E.; Bassi, C.; Casadio, F.; Martinato, F.; Cesaroni, M.; Schuchlautz, H.; Luescher, B.; Amati, B., Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. *Nature (London, U. K.)* **2007,** *449* (7164), 933-937.

116.   El-Andaloussi, N.; Valovka, T.; Toueille, M.; Steinacher, R.; Focke, F.; Gehrig, P.; Covic, M.; Hassa, P. O.; Schar, P.; Hubscher, U.; Hottiger, M. O., Arginine methylation regulates DNA polymerase β. *Mol. Cell* **2006,** *22* (1), 51-62.

117.   Michaud-Levesque, J.; Richard, S., Thrombospondin-1 Is a Transcriptional Repression Target of PRMT6. *J. Biol. Chem.* **2009,** *284* (32), 21338-21346.

118.   Bleibel, W. K.; Duan, S.; Huang, R. S.; Kistner, E. O.; Shukla, S. J.; Wu, X.; Badner, J. A.; Dolan, M. E., Identification of genomic regions contributing to etoposide-induced cytotoxicity. *Hum. Genet.* **2009,** *125* (2), 173-180.

119.   Buhr, N.; Carapito, C.; Schaeffer, C.; Kieffer, E.; Van Dorsselaer, A.; Viville, S., Nuclear proteome analysis of undifferentiated mouse embryonic stem and germ cells. *Electrophoresis* **2008,** *29* (11), 2381-2390.

120.   Karkhanis, V.; Wang, L.; Tae, S.; Hu, Y.-J.; Imbalzano, A. N.; Sif, S., Protein arginine methyltransferase 7 regulates cellular response to DNA damage by methylating promoter histones H2A and H4 of the polymerase δ catalytic subunit gene, POLD1. *J. Biol. Chem.* **2012,** *287* (35), 29801-29814.

121.   Cheng, D.; Yadav, N.; King, R. W.; Swanson, M. S.; Weinstein, E. J.; Bedford, M. T., Small Molecule Regulators of Protein Arginine Methyltransferases. *J. Biol. Chem.* **2004,** *279* (23), 23892-23899.

122.   Viviano, M.; Milite, C.; Rescigno, D.; Castellano, S.; Sbardella, G., A continuous-flow synthesis of 1,4-benzodiazepin-5-ones, privileged scaffolds for drug discovery. *RSC Adv.* **2015,** *5* (2), 1268-1273.

123.   Castellano, S.; Milite, C.; Feoli, A.; Viviano, M.; Mai, A.; Novellino, E.; Tosco, A.; Sbardella, G., Identification of Structural Features of 2-Alkylidene-1,3-Dicarbonyl Derivatives that Induce Inhibition and/or Activation of Histone Acetyltransferases KAT3B/p300 and KAT2B/PCAF. *ChemMedChem* **2015,** *10* (1), 144-157.

124.   Lenoci, A.; Tomassi, S.; Conte, M.; Benedetti, R.; Rodriguez, V.; Carradori, S.; Secci, D.; Castellano, S.; Sbardella, G.; Filetici, P.; Novellino, E.; Altucci, L.; Rotili, D.; Mai, A., Quinoline-Based p300 Histone Acetyltransferase Inhibitors with Pro-apoptotic Activity in Human Leukemia U937 Cells. *ChemMedChem* **2014,** *9* (3), 542-548.

125.   Valente, S.; Lepore, I.; Dell'Aversana, C.; Tardugno, M.; Castellano, S.; Sbardella, G.; Tomassi, S.; Di Maro, S.; Novellino, E.; Di Santo, R.; Costi, R.; Altucci, L.; Mai, A., Identification of PR-SET7 and EZH2 selective inhibitors inducing cell death in human leukemia U937 cells. *Biochimie* **2012,** *94* (11), 2308-2313.

126.   Castellano, S.; Spannhoff, A.; Milite, C.; Dal Piaz, F.; Cheng, D.; Tosco, A.; Viviano, M.; Yamani, A.; Cianciulli, A.; Sala, M.; Cura, V.; Cavarelli, J.; Novellino, E.; Mai, A.; Bedford, M. T.; Sbardella, G., Identification of Small-Molecule Enhancers of Arginine Methylation Catalyzed by Coactivator-Associated Arginine Methyltransferase 1. *J. Med. Chem.* **2012,** *55* (22), 9875-9890.

127.   Milite, C.; Castellano, S.; Benedetti, R.; Tosco, A.; Ciliberti, C.; Vicidomini, C.; Boully, L.; Franci, G.; Altucci, L.; Mai, A.; Sbardella, G., Modulation of the activity of histone acetyltransferases by long chain alkylidenemalonates (LoCAMs). *Bioorg. Med. Chem.* **2011,** *19* (12), 3690-3701.

128.   Dal Piaz, F.; Vassallo, A.; Rubio, O. C.; Castellano, S.; Sbardella, G.; De Tommasi, N., Chemical biology of Histone acetyltransferase natural compounds modulators. *Mol. Diversity* **2011,** *15* (2), 401-416.

129.   Cheng, D.; Valente, S.; Castellano, S.; Sbardella, G.; Di Santo, R.; Costi, R.; Bedford, M. T.; Mai, A., Novel 3,5-Bis(bromohydroxybenzylidene)piperidin-4-ones as Coactivator-Associated Arginine Methyltransferase 1 Inhibitors: Enzyme Selectivity and Cellular Activity. *J. Med. Chem.* **2011,** *54* (13), 4928-4932.

130.   Castellano, S.; Kuck, D.; Viviano, M.; Yoo, J.; Lopez-Vallejo, F.; Conti, P.; Tamborini, L.; Pinto, A.; Medina-Franco, J. L.; Sbardella, G., Synthesis and Biochemical Evaluation of Δ2-Isoxazoline Derivatives as DNA Methyltransferase 1 Inhibitors. *J. Med. Chem.* **2011,** *54* (21), 7663-7677.

131.   (a) Medina-Franco, J. L.; Singh, N.; Lopez-Vallejo, F.; Kuck, D.; Lyko, F.; Duenas-Gonzalez, A.; Sbardella, G. In *DNA methyltransferase inhibitors: Molecular modeling and virtual screening*, American Chemical Society: 2010; pp MEDI-68; (b) Dal Piaz, F.; Tosco, A.; Eletto, D.; Piccinelli, A. L.; Moltedo, O.; Franceschelli, S.; Sbardella, G.; Remondelli, P.; Rastrelli,

L.; Vesci, L.; Pisano, C.; De Tommasi, N., The Identification of a Novel Natural Activator of p300 Histone Acetyltransferase Provides New Insights into the Modulation Mechanism of this Enzyme. *ChemBioChem* **2010,** *11* (6), 818-827.

132.   Mai, A.; Rotili, D.; Tarantino, D.; Nebbioso, A.; Castellano, S.; Sbardella, G.; Tini, M.; Altucci, L., Identification of 4-hydroxyquinolines inhibitors of p300/CBP histone acetyltransferases. *Bioorg. Med. Chem. Lett.* **2009,** *19* (4), 1132-1135.

133.   Sbardella, G.; Castellano, S.; Vicidomini, C.; Rotili, D.; Nebbioso, A.; Miceli, M.; Altucci, L.; Mai, A., Identification of long chain alkylidenemalonates as novel small molecule modulators of histone acetyltransferases. *Bioorg. Med. Chem. Lett.* **2008,** *18* (9), 2788-2792.

134.   Mai, A.; Cheng, D.; Bedford, M. T.; Valente, S.; Nebbioso, A.; Perrone, A.; Brosch, G.; Sbardella, G.; De Bellis, F.; Miceli, M.; Altucci, L., Epigenetic Multiple Ligands: Mixed Histone/Protein Methyltransferase, Acetyltransferase, and Class III Deacetylase (Sirtuin) Inhibitors. *J. Med. Chem.* **2008,** *51* (7), 2279-2290.

135.   Castellano, S.; Kuck, D.; Sala, M.; Novellino, E.; Lyko, F.; Sbardella, G., Constrained Analogues of Procaine as Novel Small Molecule Inhibitors of DNA Methyltransferase-1. *J. Med. Chem.* **2008,** *51* (7), 2321-2325.

136.   Mai, A.; Rotili, D.; Tarantino, D.; Ornaghi, P.; Tosi, F.; Vicidomini, C.; Sbardella, G.; Nebbioso, A.; Miceli, M.; Altucci, L.; Filetici, P., Small-Molecule Inhibitors of Histone Acetyltransferase Activity: Identification and Biological Properties. *J. Med. Chem.* **2006,** *49* (23), 6897-6907.

137.   Castellano, S.; Milite, C.; Ragno, R.; Simeoni, S.; Mai, A.; Limongelli, V.; Novellino, E.; Bauer, I.; Brosch, G.; Spannhoff, A.; Cheng, D.; Bedford, M. T.; Sbardella, G., Design, synthesis and biological evaluation of carboxy analogues of arginine methyltransferase inhibitor 1 (AMI-1). *ChemMedChem* **2010,** *5* (3), 398-414.

138.   Spannhoff, A.; Heinke, R.; Bauer, I.; Trojer, P.; Metzger, E.; Gust, R.; Schuele, R.; Brosch, G.; Sippl, W.; Jung, M., Target-Based Approach to Inhibitors of Histone Arginine Methyltransferases. *J. Med. Chem.* **2007,** *50* (10), 2319-2325.

139.   Bissinger, E.-M.; Heinke, R.; Spannhoff, A.; Eberlin, A.; Metzger, E.; Cura, V.; Hassenboehler, P.; Cavarelli, J.; Schuele, R.; Bedford, M. T.; Sippl, W.; Jung, M., Acyl derivatives of p-aminosulfonamides and dapsone as new inhibitors of the arginine methyltransferase hPRMT1. *Bioorg. Med. Chem.* **2011,** *19* (12), 3717-3731.

140.   t Hart, P.; Lakowski, T. M.; Thomas, D.; Frankel, A.; Martin, N. I., Peptidic Partial Bisubstrates as Inhibitors of the Protein Arginine N-Methyltransferases. *ChemBioChem* **2011,** *12* (9), 1427-1432.

141.   Dowden, J.; Hong, W.; Parry, R. V.; Pike, R. A.; Ward, S. G., Toward the development of potent and selective bisubstrate inhibitors of protein

arginine methyltransferases. *Bioorg. Med. Chem. Lett.* **2010,** *20* (7), 2103-2105.

142.   van Haren, M.; van Ufford, L. Q.; Moret, E. E.; Martin, N. I., Synthesis and evaluation of protein arginine N-methyltransferase inhibitors designed to simultaneously occupy both substrate binding sites. *Org. Biomol. Chem.* **2015,** *13* (2), 549-560.

143.   Yan, L.; Yan, C.; Qian, K.; Su, H.; Kofsky-Wofford, S. A.; Lee, W.-C.; Zhao, X.; Ho, M.-C.; Ivanov, I.; Zheng, Y. G., Diamidine Compounds for Selective Inhibition of Protein Arginine Methyltransferase 1. *J. Med. Chem.* **2014,** *57* (6), 2611-2622.

144.   (a) Eram, M. S.; Shen, Y.; Szewczyk, M. M.; Wu, H.; Senisterra, G.; Li, F.; Butler, K. V.; Kaniskan, H. U.; Speed, B. A.; dela Sena, C.; Dong, A.; Zeng, H.; Schapira, M.; Brown, P. J.; Arrowsmith, C. H.; Barsyte-Lovejoy, D.; Liu, J.; Vedadi, M.; Jin, J., A Potent, Selective, and Cell-Active Inhibitor of Human Type I Protein Arginine Methyltransferases. *ACS Chem. Biol.* **2015**, Ahead of Print; (b) Siarheyeva, A.; Senisterra, G.; Allali-Hassani, A.; Dong, A.; Dobrovetsky, E.; Wasney, G. A.; Chau, I.; Marcellus, R.; Hajian, T.; Liu, F.; Korboukh, I.; Smil, D.; Bolshan, Y.; Min, J.; Wu, H.; Zeng, H.; Loppnau, P.; Poda, G.; Griffin, C.; Aman, A.; Brown, P. J.; Jin, J.; Al-awar, R.; Arrowsmith, C. H.; Schapira, M.; Vedadi, M., An Allosteric Inhibitor of Protein Arginine Methyltransferase 3. *Structure (Oxford, U. K.)* **2012,** *20* (8), 1425-1435.

145.   Liu, F.; Li, F.; Ma, A.; Dobrovetsky, E.; Dong, A.; Gao, C.; Korboukh, I.; Liu, J.; Smil, D.; Brown, P. J.; Frye, S. V.; Arrowsmith, C. H.; Schapira, M.; Vedadi, M.; Jin, J., Exploiting an Allosteric Binding Site of PRMT3 Yields Potent and Selective Inhibitors. *J. Med. Chem.* **2013,** *56* (5), 2110-2124.

146.   Kaniskan, H. U.; Szewczyk, M. M.; Yu, Z.; Eram, M. S.; Yang, X.; Schmidt, K.; Luo, X.; Dai, M.; He, F.; Zang, I.; Lin, Y.; Kennedy, S.; Li, F.; Dobrovetsky, E.; Dong, A.; Smil, D.; Min, S.-J.; Landon, M.; Lin-Jones, J.; Huang, X.-P.; Roth, B. L.; Schapira, M.; Atadja, P.; Barsyte-Lovejoy, D.; Arrowsmith, C. H.; Brown, P. J.; Zhao, K.; Jin, J.; Vedadi, M., A Potent, Selective and Cell-Active Allosteric Inhibitor of Protein Arginine Methyltransferase 3 (PRMT3). *Angew. Chem., Int. Ed.* **2015,** *54* (17), 5166-5170.

147.   Purandare, A. V.; Chen, Z.; Huynh, T.; Pang, S.; Geng, J.; Vaccaro, W.; Poss, M. A.; Oconnell, J.; Nowak, K.; Jayaraman, L., Pyrazole inhibitors of coactivator associated arginine methyltransferase 1 (CARM1). *Bioorg. Med. Chem. Lett.* **2008,** *18* (15), 4438-4441.

148.   Wan, H.; Huynh, T.; Pang, S.; Geng, J.; Vaccaro, W.; Poss, M. A.; Trainor, G. L.; Lorenzi, M. V.; Gottardis, M.; Jayaraman, L.; Purandare, A. V., Benzo[d]imidazole inhibitors of Coactivator Associated Arginine

Methyltransferase 1 (CARM1)-Hit to Lead studies. *Bioorg. Med. Chem. Lett.* **2009,** *19* (17), 5063-5066.

149.   Therrien, E.; Larouche, G.; Manku, S.; Allan, M.; Nguyen, N.; Styhler, S.; Robert, M.-F.; Goulet, A.-C.; Besterman, J. M.; Nguyen, H.; Wahhab, A., 1,2-Diamines as inhibitors of co-activator associated arginine methyltransferase 1 (CARM1). *Bioorg. Med. Chem. Lett.* **2009,** *19* (23), 6725-6732.

150.   Allan, M.; Manku, S.; Therrien, E.; Nguyen, N.; Styhler, S.; Robert, M.-F.; Goulet, A.-C.; Petschner, A. J.; Rahil, G.; Robert MacLeod, A.; Deziel, R.; Besterman, J. M.; Nguyen, H.; Wahhab, A., N-Benzyl-1-heteroaryl-3-(trifluoromethyl)-1H-pyrazole-5-carboxamides as inhibitors of co-activator associated arginine methyltransferase 1 (CARM1). *Bioorg. Med. Chem. Lett.* **2009,** *19* (4), 1218-1223.

151.   Huynh, T.; Chen, Z.; Pang, S.; Geng, J.; Bandiera, T.; Bindi, S.; Vianello, P.; Roletto, F.; Thieffine, S.; Galvani, A.; Vaccaro, W.; Poss, M. A.; Trainor, G. L.; Lorenzi, M. V.; Gottardis, M.; Jayaraman, L.; Purandare, A. V., Optimization of pyrazole inhibitors of coactivator associated arginine methyltransferase 1 (CARM1). *Bioorg. Med. Chem. Lett.* **2009,** *19* (11), 2924-2927.

152.   Sack, J. S.; Thieffine, S.; Bandiera, T.; Fasolini, M.; Duke, G. J.; Jayaraman, L.; Kish, K. F.; Klei, H. E.; Purandare, A. V.; Rosettani, P.; Troiani, S.; Xie, D.; Bertrand, J. A., Structural basis for CARM1 inhibition by indole and pyrazole inhibitors. *Biochem. J.* **2011,** *436* (2), 331-339.

153.   Schurter, B. T.; Koh, S. S.; Chen, D.; Bunick, G. J.; Harp, J. M.; Hanson, B. L.; Henschen-Edman, A.; Mackay, D. R.; Stallcup, M. R.; Aswad, D. W., Methylation of Histone H3 by Coactivator-Associated Arginine Methyltransferase 1. *Biochemistry* **2001,** *40* (19), 5747-5756.

154.   Lee, J.; Bedford, M. T., PABP1 identified as an arginine methyltransferase substrate using high-density protein arrays. *EMBO Rep.* **2002,** *3* (3), 268-273.

155.   Boriack-Sjodin, P. A.; Jin, L.; Jacques, S. L.; Drew, A.; Sneeringer, C.; Scott, M. P.; Moyer, M. P.; Ribich, S.; Moradei, O.; Copeland, R. A., Structural insights into ternary complex formation of human CARM1 with various substrates. *ACS Chem. Biol.* **2015**, Ahead of Print.

156.   Shaitanova, E. N.; Gerus, I. I.; Kukhar, V. P., A new synthetic route to 3-polyfluoroalkyl-containing pyrroles. *Tetrahedron Lett.* **2008,** *49* (7), 1184-1187.

157.   Van Leusen, A. M.; Siderius, H.; Hoogenboom, B. E.; Van Leusen, D., Chemistry of sulfonylmethyl isocyanides. 6. New and simple synthesis of the pyrrole ring system from Michael acceptors and (p-tolylsulfonyl)methyl isocyanide. *Tetrahedron Lett.* **1972,** (52), 5337-40.

158.   Padwa, A.; Burgess, E. M.; Gingrich, H. L.; Roush, D. M., On the problem of regioselectivity in the 1,3-dipolar cycloaddition reaction of

munchnones and sydnones with acetylenic dipolarophiles. *J. Org. Chem.* **1982,** *47* (5), 786-91.

159.   Antilla, J. C.; Baskin, J. M.; Barder, T. E.; Buchwald, S. L., Copper-diamine-catalyzed N-arylation of pyrroles, pyrazoles, indazoles, imidazoles, and triazoles. [Erratum to document cited in CA141:207140]. *J. Org. Chem.* **2004,** *69* (19), 6514.

160.   Gourlay, B. S.; Molesworth, P. P.; Ryan, J. H.; Smith, J. A., A new and high yielding synthesis of unstable pyrroles via a modified Clauson-Kaas reaction. *Tetrahedron Lett.* **2006,** *47* (5), 799-801.

161.   Martin, N. I.; Liskamp, R. M. J., Preparation of NG-substituted L-arginine analogs suitable for solid phase peptide synthesis. *J. Org. Chem.* **2008,** *73* (19), 7849-7851.

162.   t Hart, P.; Kleijn, L. H. J.; de Bruin, G.; Oppedijk, S. F.; Kemmink, J.; Martin, N. I., A combined solid- and solution-phase approach provides convenient access to analogues of the calcium-dependent lipopeptide antibiotics. *Org. Biomol. Chem.* **2014,** *12* (6), 913-918.

163.   Lerner, C.; Masjost, B.; Ruf, A.; Gramlich, V.; Jakob-Roetne, R.; Zurcher, G.; Borroni, E.; Diederich, F., Bisubstrate inhibitors for the enzyme catechol-O-methyltransferase (COMT): influence of inhibitor preorganisation and linker length between the two substrate moieties on binding affinity. *Org Biomol Chem* **2003,** *1* (1), 42-9.

164.   Narayana, B.; Ashalatha, B. V.; Vijaya Raj, K. K.; Fernandes, J.; Sarojini, B. K., Synthesis of some new biologically active 1,3,4-oxadiazolyl nitroindoles and a modified Fischer indole synthesis of ethyl nitro indole-2-carboxylates. *Bioorg. Med. Chem.* **2005,** *13* (15), 4638-4644.

165.   La Regina, G.; Bai, R.; Coluccia, A.; Famiglini, V.; Pelliccia, S.; Passacantilli, S.; Mazzoccoli, C.; Ruggieri, V.; Sisinni, L.; Bolognesi, A.; Rensen, W. M.; Miele, A.; Nalli, M.; Alfonsi, R.; Di Marcotullio, L.; Gulino, A.; Brancale, A.; Novellino, E.; Dondio, G.; Vultaggio, S.; Varasi, M.; Mercurio, C.; Hamel, E.; Lavia, P.; Silvestri, R., New Pyrrole Derivatives with Potent Tubulin Polymerization Inhibiting Activity As Anticancer Agents Including Hedgehog-Dependent Cancer. *J. Med. Chem.* **2014,** *57* (15), 6531-6552.

166.   Pinto, D. J. P.; Orwat, M. J.; Wang, S.; Fevig, J. M.; Quan, M. L.; Amparo, E.; Cacciola, J.; Rossi, K. A.; Alexander, R. S.; Smallwood, A. M.; Luettgen, J. M.; Liang, L.; Aungst, B. J.; Wright, M. R.; Knabb, R. M.; Wong, P. C.; Wexler, R. R.; Lam, P. Y. S., Discovery of 1-[3-(Aminomethyl)phenyl]-N-[3-fluoro-2'-(methylsulfonyl)-[1,1'-biphenyl]-4-yl]-3-(trifluoromethyl)-1H-pyrazole-5-carboxamide (DPC423), a Highly Potent, Selective, and Orally Bioavailable Inhibitor of Blood Coagulation Factor Xa. *J. Med. Chem.* **2001,** *44* (4), 566-578.

167.   Wurtz, N. R.; Priestley, E. S.; Cheney, D. L.; Glunz, P. W.; Zhang, X.; Ladziata, V.; Parkhurst, B.; Mueller, L. Preparation of macrocyclic amino acid

derivatives as factor VIIa inhibitors useful as anticoagulants. WO2008079836A2, 2008.

168.   Roethle, P. A.; McFadden, R. M.; Yang, H.; Hrvatin, P.; Hui, H.; Graupe, M.; Gallagher, B.; Chao, J.; Hesselgesser, J.; Duatschek, P.; Zheng, J.; Lu, B.; Tumas, D. B.; Perry, J.; Halcomb, R. L., Identification and Optimization of Pteridinone Toll-like Receptor 7 (TLR7) Agonists for the Oral Treatment of Viral Hepatitis. *J. Med. Chem.* **2013,** *56* (18), 7324-7333.