UNIVERSITA' DEGLI STUDI DI PARMA

Dottorato di ricerca in Scienze del Farmaco, delle Biomolecole e dei Prodotti per la Salute

Ciclo XXIX

NEW STRATEGIES TO ERADICATE MICROBIAL RESISTANCE

Coordinatore:
Chiar.mo Prof. Marco Mor

Tutor:
Chiar.mo Prof. Gabriele Costantino

Dottoranda:
Giannamaria Annunziato

INDEX

PREFACE

Antibiotics have always been considered one of the most relevant discoveries of the 20th century and their widespread availability has revolutionised healthcare since their introduction.[1] The dawn of the modern "*antibiotic era*" is associated to the names of Paul Ehrlich and Alexander Fleming. Paul Ehrlich is credited with the concept of the so-called "magic bullet", based on target selectivity; the Ehrlich's idea was that a given molecule selectively targets only disease-causing agents and not the host. Ehrlich asserted that it was possible to chemically prepare compounds that would "be able to exert their full action exclusively on the parasite harboured within the organism". On the other hand, Fleming has to be credited for the serendipitous discovery of Penicillin, later to be claimed as the miracle drug of the 20th century.

Although the spread of antibiotics formally starts with the "antibiotic era", however in the pre-antibiotic era numerous remedies belonging to the traditional medicine have constantly put microorganism under the pressure of antibacterial. Antimicrobial activity seems to be present in a number of herbs and plants and the discovery of active components from natural sources not only has contributed to build the antibacterial arsenal, but they still keep fuelling the pipeline of antimicrobials used by the mainstream medicine. In the past 200 years, empirical science and serendipity have come together to bring over the current state of knowledge about antimicrobial drugs.[2]

Unfortunately, the dawn of the antibiotic era has sadly corresponded to the raise of the phenomenon of antimicrobial resistance.

If it is widely accepted that the use of antimicrobial agents has led to the control and even eradication of diseases, at the same time their misuse has also led to the development of resistant strains. Based on his early

observations, in 1945, Fleming had predicted that the inconsiderate use of this discovery could lead to the selection and propagation of mutant forms of bacteria resistant to antibiotics. Indeed, alarm signs of developing resistance were observed just after few years of the golden age of antimicrobials.[3,4]

Antimicrobial resistance (AMR) is a natural process whereby microbes evolve in such a way to withstand the action of drugs, making them apparently ineffective. The pressure that antimicrobials put on populations of microbes is responsible for the selection of resistant strains.

Although AMR is a naturally occurring process, and while in the past decades it appeared to be under control, today it is considered a threat for global health, and the expectations for the future are not encouraging as well. This may be due to several reasons: (i) the use of antimicrobials has strongly increased in the last few decades; (ii) in most cases the patients are not able to carry out properly the therapy; (iii) in some classes of antimicrobials there are very limited numbers of new drugs under development to replace those ineffective by rising drug resistance.[5]

Today much effort is being required not only to develop new drugs but also to understand the basics of resistance.[6] Today, it is generally accepted that there are different mechanisms that contribute to the development and the spread of resistance in the case of antimicrobials (Figure 1) :[7]

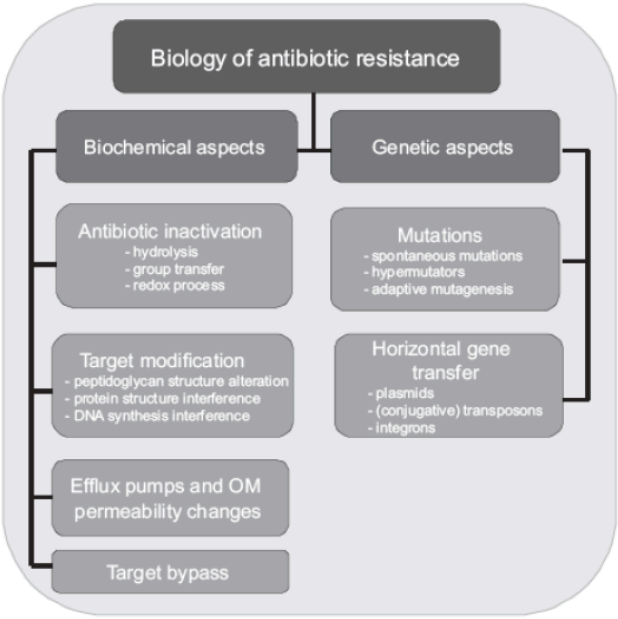


Figure 1. In this graph is reported the different mechanisms that contribute to the development and the spread of resistance. Picture adapted from ref. [8]

(i) the overexpression of the drug target, thus swamping the antimicrobial agent;

(ii) the reverse of overexpression, namely reduction in the concentration of the drug target, thus eliminating it as a site of action;

(iii) changes in molecule biosynthesis;

(iv) the efflux pumps that actively throw out the antimicrobials from the cell.[8]

Drug-resistant infections are responsible for more than half a million deaths globally each year, and the number of new cases, especially as nosocomial infections, is constantly increasing, posing clinic and economic issue. Costs due to infections by multidrug-resistant bacteria in the EU result in extra healthcare costs and productivity losses each year, estimated to be at least 1.5 billion euros. The annual additional

cost of treating hospital-acquired infections from just six species of antibiotic-resistant bacteria was estimated to be at least 1.3 billion in 1992 euros (1.87 billion in 2006 euros). Should the control of resistance fail, early research suggests that deaths will exceed 10 million each year by 2050 and there will be over 100 trillion USD in lost output. Therefore, resistant infections have a deep socioeconomic impact, which affects physicians, patients, health-care administrators, pharmaceutical producers and the public.

For all of these reasons, the concerted action of governmental agencies and healthcare operators is needed to define a number of measures to minimize the detrimental effects of resistance.[5]

The World Health Organization (WHO) promotes a number of different actions to keep new resistance from developing and to prevent the resistance that already exists from spreading. First of all, it promotes sustainable investments that take into account the needs for all countries; increasing investments in new medicines, diagnostic tools, vaccines and other therapeutic interventions are pursued; finally, policies to encourage development of new antibiotics, including tax incentives for research spending, *patent* extensions for other drugs in a pharmaceutical company's portfolio, liability protection from the adverse side effects of new antibiotics are strongly promoted. They also include policies to reduce FDA approval times; thereby allowing companies to obtain a return on their investments earlier, and tax breaks to defray the cost of the FDA approval process.[9] For what concerns the field of basic research and drug discovery to fight antimicrobial resistance, several approaches have been pursued throughout the years, especially considering the increasing number of health-menacing events. One of this approaches involve the identification of new pharmacological targets

that regulate bacterial life cycle, and the synthesis of small molecules able to interact with these targets. The idea underlying such approach is that a bacterial target never hit by the members of the currently available antibiotic arsenal is likely less prone to mutate, and may lead to bacterial death if effectively targeted. Beside the aforementioned approach, where the main targets are represented by those proteins regulating the life cycle of the bacteria, new idea are emerging. Other interesting targets are represented by those proteins, genes and other bio-macromolecules, responsible for bacterial virulence and fitness, whose inhibition would likely weakness the bacteria, making them more susceptible to the attack of the immune system. Indeed, it has been proposed that such targets, being not critical for survival *per se* (bacteria are eventually killed by the immune system), are less susceptible to generate mutations. On the other hand, the main hurdle is related to the fact that, historically, target-based drug design is not recommended for antibacterial, as it does not take into account issues such as permeability, which is a peculiar feature to enter the bacterial cell and so as to reach the target of interest. Therefore, a fine-tuned medicinal chemistry approach is needed in order to deliver mature lead compounds. Another approach that, as in the past as now, may result in the discovery of novel antibiotics is the identification of natural products with antimicrobial activity, and their chemical manipulation in order to ameliorate activity and physicochemical properties. Nature has always represented an endless source of therapeutic agents, especially by virtue of the high chemical specificity of the molecules produced.

Of the many possible approaches, the above reported are those that I have chosen in this PhD project in order to deliver novel tools to study and possibly tackle down bacterial infections.

Indeed, the research activity during my PhD course was mainly focused on:

(i) synthesis of potential inhibitors of microbial targets, which are involved in microbial population growth during infection;

(ii) modifications of the scaffold of natural products endowed with antimicrobial activity.

In this contest the role of a fine-tuned medicinal chemistry approach is crucial, since it can pave the way for the design of new chemical entities able to target unexplored or less explored targets, or modify already existing drugs so as to improve their pharmacological profiles.

.

1. O-ACETYLSERINE SULFHYDRYLASE AS ANTIBACTERIAL DRUG TARGET

1.1 Sulfur Metabolism as a Source of Drug Target

Sulfur metabolism may represent a potential reservoir of targets in the search for novel antibacterial chemotherapeutics.[10] For instance, there are evidences that suppression or reduction of cysteine biosynthesis leads to a decrease of bacterial adaptation and to a reduction of infectivity.[11] Some pathogens spend part of their life-cycle in extremely challenging environments, where survival and proliferation require powerful adaptation mechanisms, involving peculiar metabolic pathways.[12] In such conditions, any interference with the pathogen adaptation strategies, such as interfering with amino acid biosynthesis, may lead to a reduction of their virulence or persistence or increased susceptibility to antibiotics.

It has been observed that the importance of cysteine biosynthetic enzymes varies within the life-cycle of pathogens: their activity can be dispensable during growth *in vitro* or acute infection and becomes indispensable during persistence phase.[13] Molecules targeting biosynthetic pathways like cysteine biosynthesis may have the potential advantage of fighting infections that persist inside the host (i.e. during clinical latency) more effectively than traditional antibiotics, thus contributing to prevent resistance development.[14]

As a matter of fact, of great interest is the observation that both vegetative and actively swarming *S. typhimurium*, knockout for cysteine biosynthesis enzymes, displayed a significant decrease in antibiotic resistance. Thus providing with initial experimental evidences that an

impaired oxidative stress response, due to inhibition of cysteine biosynthesis, can be an interesting strategy to overcome bacterial resistance.[15] Indeed, antibiotic-induced oxidative stress, which is widely recognized as a general mechanism of action of many antibacterials, could explain the reduced resistance rate observed in bacteria with impaired cysteine biosynthesis.[16] Interestingly, the increased rate of antibiotic resistance observed in swarm cells can be reproduced on vegetative populations by induction of the cysteine regulon, further confirming the role played by cysteine in conferring a resistant phenotype. These findings suggest that inhibitors of cysteine biosynthesis may represent a selective and innovative approach to deal with the issue of antibiotic resistance, and enhance the efficacy of antibiotic treatment.[14]

1.1.1 Sulfur Assimilation Pathway

Sulfur became available to mammalians because bacteria and plants incorporate inorganic sulfur into cysteine,[17] which, instead, is made available to mammalians through reverse trans-sulfuration from methionine. Cysteine is the amino acid precursor of all sulfur-containing biomolecules, including methionine, CoA, biotin, Fe-S clusters, penicillin and glutathione, that are crucial for the survival and virulence of the majority of living organisms.[18]

Most bacteria and plants carry out cysteine biosynthesis by reductive sulfate assimilation pathway (RSAP), a multistep reduction of sulfate that culminates in the incorporation of bisulfide into cysteine using an activated form of serine (Figure 2).[19]

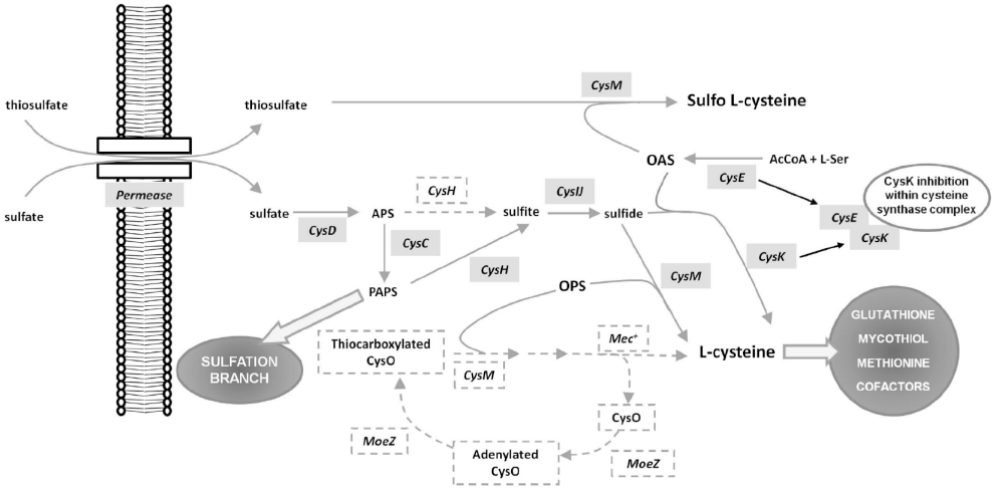


Figure 2. RSAP: Reductive Sulfur Assimilation Pathway. Picture adapted from ref.[11]

RSAP begins with the transport of sulfate inside the cell, followed by his reduction to bisulfide. This process is highly energy-consuming and is tuned on cellular needs. Bacteria find sulfur in the form of sulfate in the environment and actively transport it through the plasma membrane.[20]

After reduction of sulfate to bisulfide, a toxic compound whose concentration inside the cell is kept between 20 and 160 μM, the latter is incorporated into cysteine by a member of a large enzyme family known as cysteine-synthase (CS).[21] The CS complex is composed by the enzymes Serine Acetyltransferase (SAT) and O-acetylserine sulfydrylase (OASS), which catalyse the last step of cysteine biosynthesis.[17,18]

1.1.2 Cysteine synthase complex: SAT and OASS enzymes

Control of sulfur metabolism in plants and bacteria is linked to cysteine synthase complex (CS). The complex consists of serine acetyltransferase (SAT) and *O*-acetyl-*L*-serine sulfydrylase (OASS), which catalyse the final step in cysteine biosynthesis.[22]

SAT is catalytically active in the complex and produces *O*-acetyl-L-serine (OAS, the activated from of Ser), a cysteine precursor where the β-hydroxyl group of the amino acid is acetylated or phosphorylated to generate a better living group in the β-elimination reaction (Figure 3).

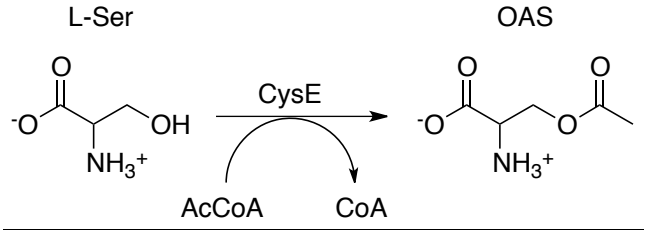


Figure 3. Reaction catalysed by serine acetyltransferase

Most bacteria use OAS as preferred substrate for cysteine biosynthesis.[13] OAS is unstable and spontaneously converts to *N*-acetylserine (NAS), the natural inducer of the cysteine regulon signalling. SAT catalyses the reaction by which an acetyl group from acetyl-CoA (AcCoA) is transferred to the hydroxyl one of *L*-Serine to form OAS and CoA.[23]

It has been hypothesized that SAT is a dimer of trimers, with a flexible C-terminal portion bound to an unusual left-handed parallel β-helix structural domain. The C-terminus tail of SAT is fundamental for function and regulation, being responsible of an intrasteric inhibition in the

presence of cysteine by binding to OASS-A for the formation of cysteine synthase complex.[24]

Depending on the organism and growth conditions, the last step of cysteine biosynthesis is catalysed by different sulphydrilases, sharing high homology but also showing some functional and structural differences. Based on the preferred substrate, these enzymes can be divided into two main groups: *O*-acetylserine sulfhydrylase (OASS-A) and *O*-phosphoserine sulfhydrylase (OASS-B).[18]

Initially, OASS-A and OASS-B were indentified in *S. typhymurium*, and the two *O*-acetylserine sulfhydrylase isoforms are expressed under different growth conditions (aerobic and anaerobic respectively, Figure 4).

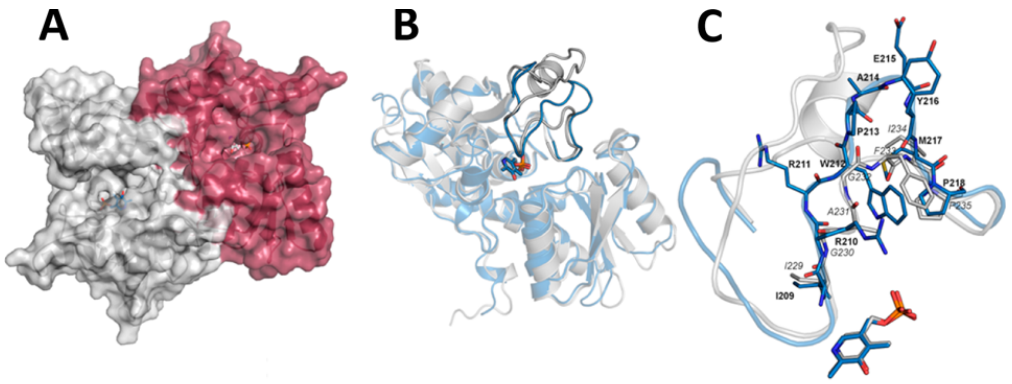


Figure 4. (A) Homodimer of OASS-A taken from 1OAS crystal structure; the two monomers are highlighted with different colors. (B) Superposition of OASS-A (1OAS chain A, in light gray) and OASS-B (2JC3 chain A, in blue). The major structural difference concerning the loops above the binding site is highlighted; the sequences of the loops of the two isoforms, aligned as in the superposition of the two crystal structures, are reported at the bottom of the figure. (C) Zoom on the loops highlighted in panel B: OASS-A in light gray with residues labeled in italics; OASS-B in blue with residues labeled in bold. The structure-based amino acid sequence alignment of the loops of the two isoforms is reported at the bottom of the figure. Picture adapted from ref.[25]

In addition, OASS-B was demonstrated to be a more promiscuous enzyme, capable of accepting also thiosulfate as a sulphur donor. Furthermore, OASS-B accepts a wider variety of substrates and bulkier ligands, suggesting an active site volume greater than OASS-A. The two isozymes share a 43% identity and exhibit very similar three-dimensional structures.[26]

OASS-A and OASS-B are dimeric pyridoxal 5'-phosphate (PLP) dependent enzymes and share a similar Bi-Bi ping-pong reaction mechanism (Figure 5).

The first half-reaction is the β-elimination on the β-substituted L-serine, leading to the accumulation of the α-aminoacrylate intermediate, that, in the second half-reaction, is attacked by sulfide, or other nucleophilic substrates.[27]

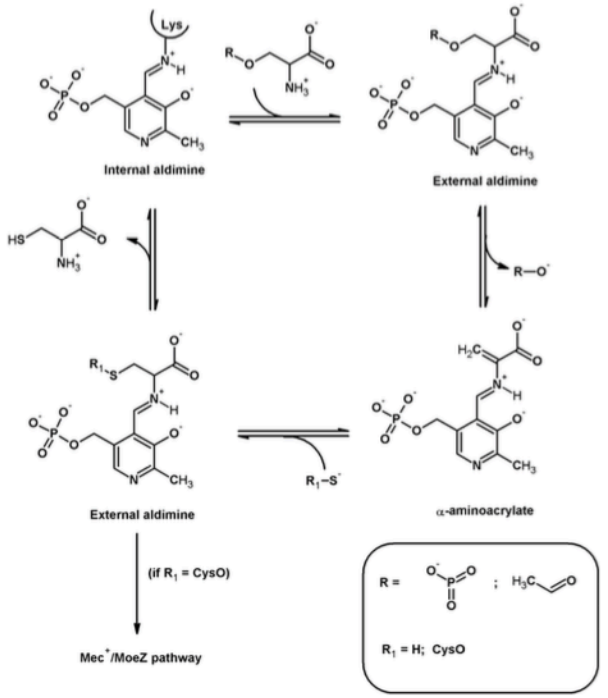


Figure 5. Catalytic cycle of O-acetylserine sulfidrilase. Picture adapted from ref.[11]

The three-dimensional (3-D) structures of OASS-A/OASS-B invariably show a homodimer (Figure 4), with each subunit organized into two domains and active sites, each carrying a molecule of PLP, facing the solvent. The two domains are flexible and close up upon the substrate binding, thus allowing diffusion into the active site only of small molecules and protecting the highly reactive α-aminoacrylate intermediate from attack by water or other nucleophiles. Of the two isoforms, only OASS-A is capable of binding SAT, originating the CS complex (Figure 6).[25]

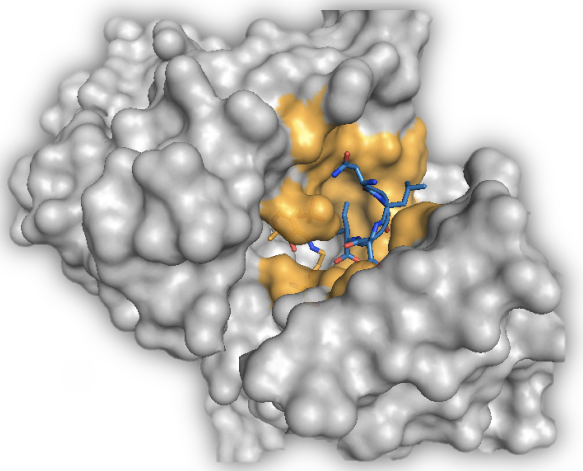


Figure 6. Three dimensional structure of OASS-A (grey and orange surface, where orange surface highlights the OASS active site), in complex with the last 5 amino acid of the C-ter tail of SAT (blue sticks).

Of great interest is the observation that, actively swarming *S. typhimurium*, knockout for *cysK,* the gene coding for OASS-A, displays a 500-fold decrease in antibiotic resistance. Thus providing with initial experimental evidences that blocking OASS-A-catalysed production of cysteine could represent a novel strategy to weaken microorganism defences against oxidative stress and so to counteract their resistance

to antibiotics.[28] Disclosure of OASS-A and SAT 3D structures from various bacteria, such as *S. thyphimurium, H. influenza* and *E. coli,* in addition to extensive biochemical studies, has shed light on the modes of interaction between the two enzymes to assemble the CS complex.[29] The interaction of OASS with serine acetyltransferase (SAT), the preceding enzyme in the anabolic pathway, has been characterized determining the kinetic mechanism and binding affinity. In detail, OASS-A forms a fitted complex with SAT, with a K_d in nanomolar range, whereas OASS-B does not show interaction with SAT. SAT binds OASS active site via its terminal C-terminal peptide, resulting in a competitive inhibition of OASS. OASS-A forms a fast face complex with SAT, followed by a slow conformational change.[30] The structure of OASS-A from *H. influenzae* was determined with the C-terminal decapeptide of SAT bound in the active site. Only the last four amino acids (NLNI) were detected, suggesting that they have a specific role in the interaction. This conclusion is supported by extensive mutational and computational analysis, also showing the relevance of the C-terminal amino acid isoleucine for OASS-SAT formation.[14]

1.2 State of the Art

A previous attempt to identify OASS inhibitors was carried out by exploiting a small library of pentapeptides. Such pentapeptides were designed as analogues of the C-ter tail of SAT and showed affinities for OASS-A active site in the middle-low micromolar range. Even if of foremost importance, such pentapeptides are characterized by practical limitation, such as low stability and low bioavailability. Therefore, there is the need to circumvent the chemical liabilities of the aforementioned peptides, while maintaining the enzyme inhibition capability. In order to

identify new inhibitor of OASS enzymes, with improved pharmacological profile the following observations are relevant: (*i*) the last amino acid residue of most C-ter tails of different pathogens is usually an Ile, (*ii*) the Ile residue is crucial for OASS-SAT interaction, and deletion of such residue completely abrogates binding of SAT to OASS-A. (Figure 7A). On the basis of the aforementioned experimental observation and taking advantages of the structural information obtained by the analysis of the 3-D complex of OASS-A in complex with the NLNI peptide (Figure 7A), we reasoned that the C-ter Ile can be mimicked in search for new inhibitors. Therefore, we speculated that the acid moiety of the isoleucine is crucial for the activity, since it establishes a dense network of hydrogen bonds with the backbone of several residues belonging to the catalytic pocket, namely, Asn72, Thr73, Thr69, and Gln143.[31]

On the other hand, an alkyl moiety mimicking the side chain of isoleucine can be introduced beside the carboxylic acid moiety through a suitable spacer, so as to maintain the two anchoring arms into favourable conformations, and to properly mimic the interaction observed for the C-ter Ile. In Figure 7 is summarized the strategy used to rationally design new OASS inhibitors, together with the structure of the first non-peptidic OASS inhibitors identified (Figure 7B).

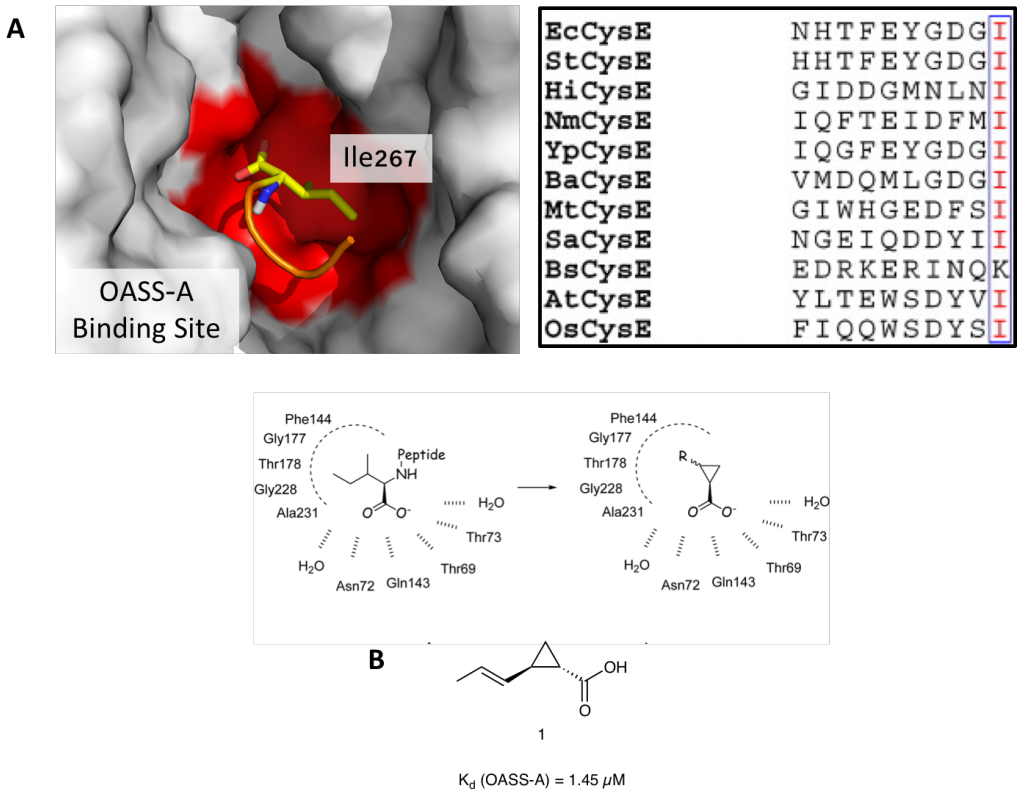


Figure 7. (A) Rational design of cyclopropane carboxylic acids. (B) First OASS-A inhibitor.

Based on the preliminary SAR studies reported by Amori et al,[32] the general structure of these inhibitors relied on the presence of a carboxylic acid moiety linked through a cyclopropane linker to a side chain mimicking that of isoleucine, in a trans-configuration. Analogues of these β-substituted cyclopropane carboxylic acids were prepared, and the most active compound trans -2-(prop-1-enyl)-cyclopropanecarboxylic acid (**1**), showed an encouraging K_d of 1.45 μM toward *Hi*OASS-A (Figure 7). However, this derivative suffered from several issues that make it particularly unsuitable for further investigation: harsh to synthesize, highly volatile, prone to decomposition in the stock solution after a few days, hampering sometimes the reproducibility of

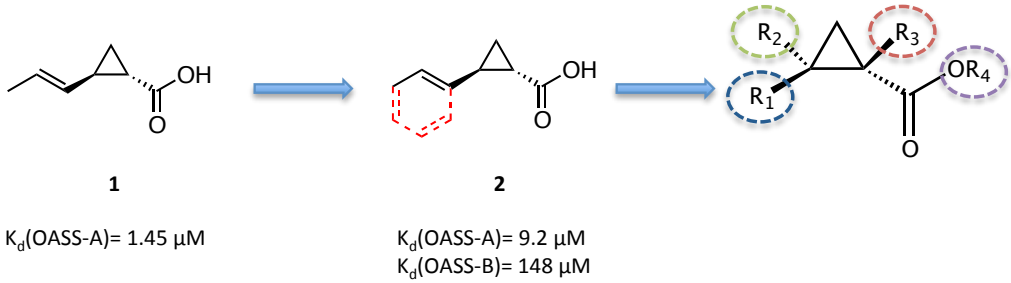
experiments. In addition, since both isoforms of OASS (A and B) must be inhibited to obtain cysteine auxotroph, the lack of activity toward *Hi*OASS-B was considered fatherly detrimental for the development of these compounds.[32]

Since the encouraging activity, and taking into consideration the drawbacks reported in the background, aim of this PhD work was to ameliorate both the activity profile and the drug-like characteristics of these hit compounds.

1.3 Aim of the Project

Taking into account these premises, I have focused my attention to the rational modifications of derivative **1**. In Figure 8 a comprehensive description of the proposed modifications is reported. The modifications were aimed to improve (*i*) the compound feasibility from a synthetic point of view, (*ii*) the stability of the compound, and (*iii*) to explore the chemical space of modifications in position R_1, R_2, R_3 and R_4 so as to identify active compounds also on OASS-B isoform, and therefore define a preliminary SAR.

It must be noticed that the medicinal chemistry work planned was driven by the iterative feedback coming from the biochemical assays, coupled to the fruitful interplay of computational and spectroscopic methods. We reasoned that the first obvious modification would have been the inclusion of the vinyl moiety into a phenyl ring, maintaining the trans stereo-relationship with the carboxylic moiety (Figure 8), enhancing the synthetic feasibility and the stability. Moreover, the insertion of a phenyl ring significantly expands the scope for further decoration and functionalization aimed at better modulating both activity and drug-likeness.



1

K_d(OASS-A)= 1.45 µM

2

K_d(OASS-A)= 9.2 µM
K_d(OASS-B)= 148 µM

Figure 8. Schematic representation of the strategy used to improve the activity profile of 1

When tested against OASS-A and OASS-B from *S. typhimurium*, a well characterized facultative intracellular pathogen for which both the enzyme isoforms have been thoroughly characterized, this modification proved to be effective, as compound trans-2-phenylcyclopropanecarboxylic acid (**2**) showed moderate affinity toward *St*OASS-A, with a K_d of 9.2 µM and resulted also active against *St*OASS-B (K_d = 148 µM) although the binding affinity was still not satisfactory. Thus, with the aim of understanding which structural modification can lead to an increased affinity toward OASS-B while retaining good potency at OASS-A, compounds **1** and **2** were docked into the active site of the two enzyme isoforms (Figure 9).

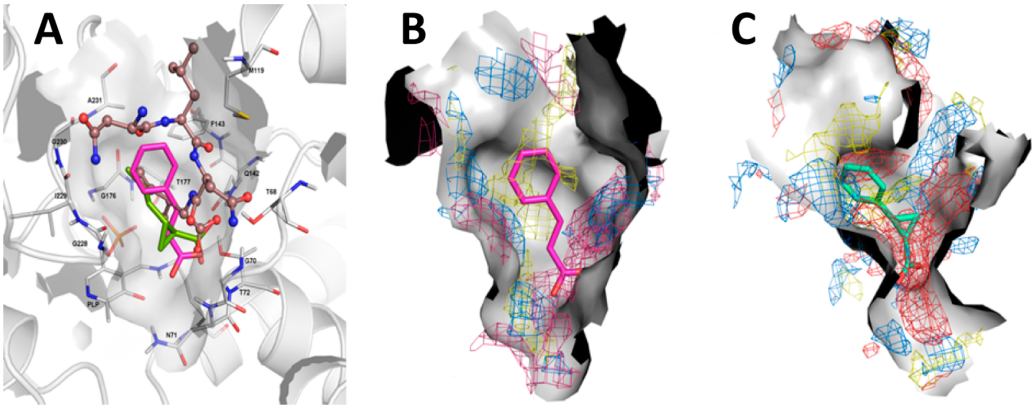


Figure 9. Binding of 2 to StOASS-A. (A) binding mode of **12** in StOASS-A binding site (1OAS crystal structure, in magenta), superposed to the previously identified compound **1** (in green) and to the C-terminal portion of CysE as solved in 1Y7L crystal structure from H. influenzae. (B) Map of the features of StOASS-A binding site with compound **2** as a reference; hydrogen bond donor areas are reported in blue meshes, the hydrogen bond acceptor in red and the hydrophobic ones in yellow. A second small hydrophobic area can be seen at the bottom of the picture, surrounded by polar areas. (C) Map of the features of StOASS-B binding site with compound **2** as a reference; colors are the same as in panel B). Picture adapted from ref.[25]

As expected, we observed that the hydrogen bond network established by the isoleucine of the carboxy-terminal portion of SAT (NLNI) in the crystallographic complex with *Hi*OASS-A (PDB code 1Y7L) is maintained by the carboxylic group of **1** and **2** in the docking poses obtained by using the crystal structure 1OAS from *S.typhymurium* (Figure 9A). In addition, the alkyl portion of **1** and the aromatic ring of **2** are located in the small lipophilic pocket occupied by the hydrophobic portion of isoleucine. As can be appreciated in Figure 9B, a considerable portion of the binding pocket is empty. This region is characterized by the presence of a small lipophilic area surrounded by mildly polar residues. Therefore, we reasoned that the functionalization of the α-carbon of the cyclopropane ring might exploit this portion of the binding site, thus establishing new interactions and increasing the binding affinity. Moreover, also in the OASS-B isoform there is space to be filled by substituents in this direction, even if the cavity is slightly

more polar (Figure 9C). The docking studies reported in Figure 9 also revealed that substituents at the β position of the cyclopropane ring or on the phenyl ring can be less tolerated since the corresponding regions of the enzyme active site are less prone to host bulkier substituents, in particular considering the shape of the binding pocket of the B isoform. Therefore, I have continued the synthesis of novel substituted cyclopropanecarboxylic acids, with the aim improving their stability and their potency toward both the OASS-A and OASS-B isoforms.[12]

The novel compounds (Figure 10) could serve as biological tools for the study and further characterization of OASS enzymes.

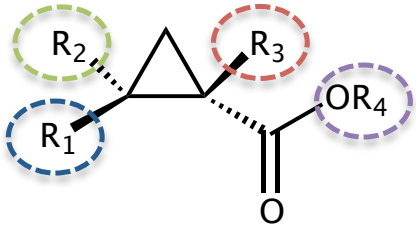


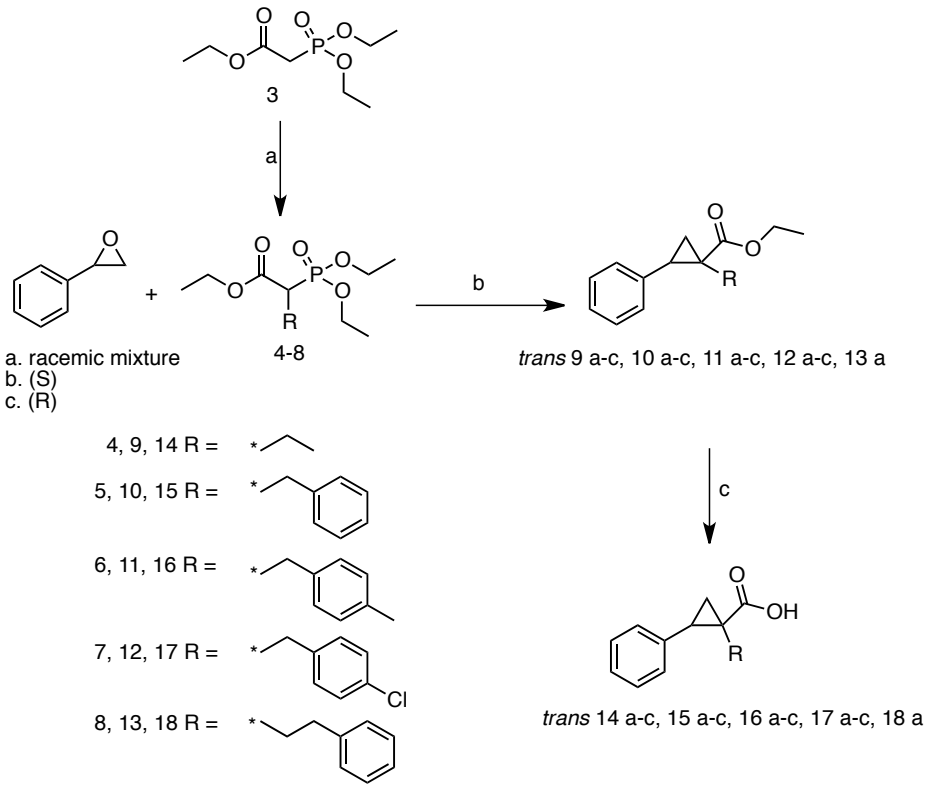
Figure 10. General representation of the structure for the newly proposed OASS inhibitors. Highlighted with coloured circles the site of possible modifications.

Compound	Stereochemistry	R_1	R_2	R_3	R_4
14a	*trans*	-Ph	-H	-Et	H
14b	1R,2S	-Ph	-H	-Et	-H
14c	1S,2R	-Ph	-H	-Et	-H
10a	*trans*	-Ph	-H	-Bn	-OEt
15a	*trans*	-Ph	-H	-Bn	-H
15b	1S,2S	-Ph	-H	-Bn	-H
15c	1R,2R	-Ph	-H	-Bn	-H
16a	*trans*	-Ph	-H	4-CH_3-Bn	-H
16b	1S,2S	-Ph	-H	4-CH_3-Bn	-H
16c	1R,2R	-Ph	-H	4-CH_3-Bn	-H
17a	*trans*	-Ph	-H	4-Cl-Bn	-H
17b	1S,2S	-Ph	-H	4-Cl-Bn	-H
17c	1R,2R	-Ph	-H	4-Cl-Bn	-H
18a	*trans*	-Ph	-H	-PhEt	-H
20	*cis*	-H	-COOH	-H	-H
21	*cis*	-H	-COOH	-H	-OEt
23	*trans*	-H	-COOH	-H	-H
24	*trans*	-H	-COOH	-H	-OEt
38	*rac*	-CH_3	-COOH	-H	-H
39	*rac*	-H	-COOEt	-H	-H
40	*rac*	-Ph	-COOH	-H	-H
41	*rac*	-CH_3	-COOH	-CH_3	-H
42	*rac*	-Ph	-COOH	-CH_3	-H
43	*rac*	-Bn	-COOH	-CH_3	-H
44	*rac*	-Ph	-COOEt	-H	-H
47	*rac*	-CH_2OH	-H	-Ph	-H

Table 1. Synthesized compounds as potential inhibitors of OASS.

1.4 Chemistry

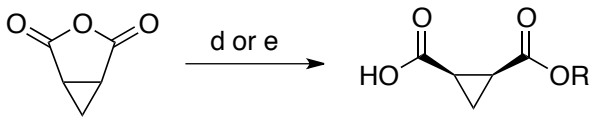
The 2-phenylcyclopropane carboxylic acids (**14 a−c**, **15 a−c**, **16 a−c**, **17a−c** and **18 a**) were prepared through a straightforward protocol already reported,[33,34] reacting the appropriate styrene oxides **a−c** with the suitable phosphonates in dimethoxyethane at 130 °C using butyllithium as deprotonating agent. This procedure has the advantage to direct in the trans position the carboxylic moiety and the phenyl ring, with retention of configuration. This allows synthesizing the enantiopure 2-phenylcyclopropane carboxylic acids simply starting from the commercially available enantiopure styrene oxides. Subsequent basic hydrolysis with $LiOH \cdot H_2O$ under MW irradiation afforded the title compounds in good overall yields (Scheme 1). When not commercially available, phosphonates were obtained by reacting triethyl phosphonoacetate with the proper alkyl bromide, in the presence of NaH as the base, from 0 ° C to room temperature.[35,36] Compound **2** was commercially available.



Scheme 1. Reagents and conditions: (a) R-Br, NaH, DME, 60 °C, 4 h, 54−75%; (b) n-BuLi, DME, 90 °C, 18 h, 67−76%; (c) LiOH·H2O, THF/ MeOH/H2O (3:1:1); 10 min, MW, 86−98%.

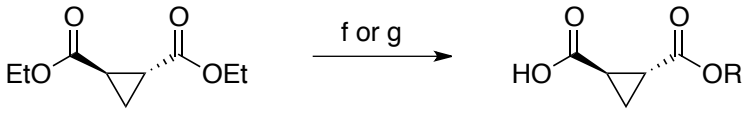
The synthesis of cis-cyclopropane-1,2-carboxylic acid derivatives (**20**, **21**) started from commercially available 3- oxabicyclo[3.1.0]hexane-2,4-dione (**19**), which was hydrolyzed with either water or ethanol to give, respectively, compounds **20** and **21** in good overall yields (Scheme 2). To obtain the trans-cyclopropane- 1,2-carboxylic acid derivatives, diethyl (1R,2R)-cyclopropane- 1,2-dicarboxylate was hydrolyzed in the presence of stoichiometric amount of potassium hydroxide, to give derivatives **23** and **24** in quantitative yields. For the synthesis of substituted cyclopropane-1,2-dicarboxylic acids (Scheme 3), key intermediates **32**–**37** were prepared according to a protocol described by McCoy,[37] since the procedures already set by us proved to be not

efficient.[34] Therefore, the suitable acrylates and a-halo esters were reacted in the presence of NaH, and the diethyl ciscyclopropane- 1,2-dicarboxylates were obtained. The subsequent hydrolysis in basic aqueous media afforded the title products **38**–**44** in good overall yields (Scheme 4). Of note, during the cyclopropanation, the solvent plays a key role in the stereochemistry of the resulting molecules, with toluene being the best in conferring the desired *cis* configuration. Reacting benzyl cyanide and epichlorohydrin in the presence of sodium amide, at 90 °C in benzene, the corresponding cyclopropyl alcohol was obtained, that on turn underwent basic hydrolysis to give the title compound 28 in 44% overall yields (Scheme 5). Attempts to synthesize the ethyl ester of **47** according to other reported procedures[34] failed to give the desired compound.



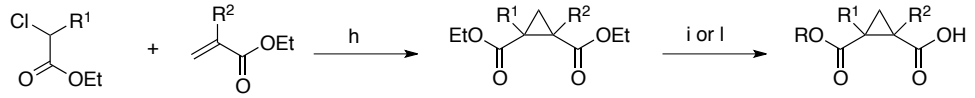
19 **20**, R=H
21, R=Et

Scheme 2. Reagents and conditions. (d) H2O, reflux, 2 h, 98%; (e) EtOH, pyr, reflux, 12 h, 96%.



22 **23**, R=H
24, R=Et

Scheme 3. Reagents and conditions. (f) 1. 6N NaOH, reflux, 12 h, 2. aq. HCl, pH 2-3, 97%; (g) 14N NaOH (1 equiv.), EtOH, reflux, 15 min, 98%.



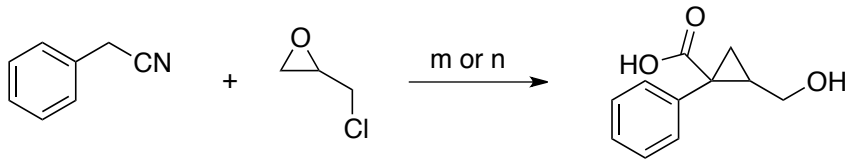
25, R^1=H
26, R^1=CH_3
27, R^1=CH_2CH_3
28, R^1=Ph
29, R^1=Bn

30, R^2=H
31, R^2=CH_3

32, R^1=CH_3 R^2=H
33, R^1=CH_2CH_3 R^2=H
34, R^1=Ph R^2=H
35, R^1=CH_3 R^2=CH_3
36, R^1=Ph R^2=CH_3
37, R^1=Bn R^2=CH_3

38, R=H R^1=CH_3 R^2=H
39, R=H R^1=CH_2CH_3 R^2=H
40, R=H R^1=Ph R^2=H
41, R=H R^1=CH_3 R^2=CH_3
42, R=H R^1=Ph R^2=CH_3
43, R=H R^1=Bn R^2=CH_3
condition c
44, R=Et R^1=Ph R^2=H

Scheme 4. Reagents and conditions. (h) NaH, dry toluene, 20–40 _C, 36–72 h, 65–78%; (i) 1N KOH, THF/H2O (1:2), 100 °C, 24 h, 62–75%; (l) 14N NaOH (1 equiv.), EtOH, reflux, 15 min, 64%.



45 **46** **47**

Scheme 5. Reagents and conditions. m) $NaNH_2$, dry benzene, rt, 16 h; (n) 1N NaOH, reflux, 24 h, 35%.

1.5 Results and Discussion

Based on the observations reported in the previous sections of this chapter, we designed compound **14a**, bearing an ethyl moiety at the α-carbon, which was initially prepared as racemic mixture. The affinity to *St*OASS-A did not change significantly with respect to that of **2**, whereas the affinity toward the B isoform was considerably lower (K_d, OASS-A = 15 μM; K_d, OASS-B = 720 μM).

Compound	OASS A	OASS B	SI[b]
	K_d (μM)		
1	1.45±0.3	>1000	na
2	9.2±0.9	148±16	16
10a	no bind[e]	nd[f]	na
14a	15±1	720±100	48
14b	12.1±0.5	860±90	71
14c	1200±300	>2000	na
15a[d]	0.20±0.02	4.1±0.2	21
15b[g]	0.067±0.007	1.66±0.07	25
15c	50±8	197±27	4
16a	0.077±0.011	1.2±0.2	16
16b	0.028±0.005	0.49±0.05	18
16c	81±17	121±12	1.5
17a	0.156±0.007	0.76±0.10	5
17b	0.054±0.008	0.42±0.06	8
17c	84±10	152±19	1.8
18a	14±3	27±2	1.9
20	215±9	900±90	4.2
21	>1000	>1000	na
23	>1000	640±10	<0.6
24	245±14	675±9	2.8
38	49±1	320±20	6.5
39	8±1	173±8	21
40	7±1	55±3	7.4
41	42±3	350±32	8.3
42	9±1	40±2	4.4
43	48±2	368±47	7.7
44	83±2	>1000	>12
47	168±35	nd	na

Table 2. [a] All the compounds are in trans configuration; [b] SI: Selectivity index is intended as the ratio $K_{d,OASS-A}/K_{d,OASS-B}$; [c]not attainable; [d]stoichiometry calculated for this compound: 2.36; [e]no binding at concentrations < 20 μM. The compound is not soluble in the assay buffer with 1% DMSO at concentrations > 20 μM; [f] not determined; [g] stoichiometry calculated for this compound: 1.05.

Since for optically active compounds the stereochemistry is usually crucial for the interaction with the enzymes, we deemed of interest to synthesize as reported in Scheme 1 the two enantiomers of **14a**. Interestingly, we found that while the 1R,2S enantiomer **14b** was

barely as active as the racemic mixture, the 1S,2R enantiomer **14c** was inactive. The orientation of **14c** in the binding site, as obtained by docking of the compound into St OASS-A active site, is slightly turned compared to the active enantiomer **14b**. This leads to a less efficient interaction of the carboxylic acid with the surrounding residues and to some steric clashes in the hydrophobic pocket occupied by the phenyl ring. Therefore, we can assume that this enantiomer is not properly accommodated inside the binding site, partially explaining the huge difference measured in the K_ds for the two compounds. With regard to **14b**, given the small size of the ligand and the big cavity of the binding site, docking studies found two equally possible binding modes. In both cases, the carboxylic group is involved in a hydrogen bond network and van der Waals interactions with the phenyl ring and the ethyl group.

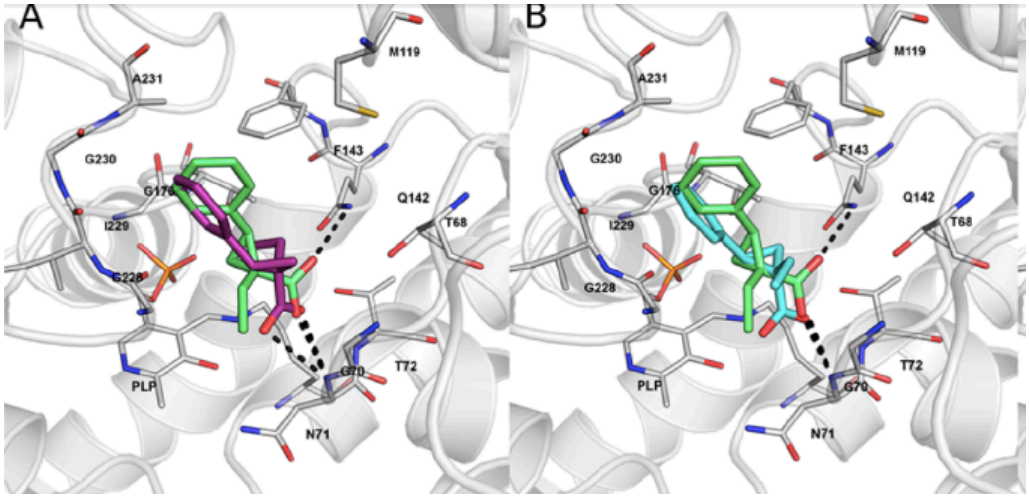


Figure 11. Stereochemical preferences for the binding of cyclopropane carboxylic acid derivatives to *St*OASS-A. (A) Two possible binding modes for compound **14b** (1R,2S); the one represented in green seems to be the preferred one (see text). (B) Docking pose of the inactive compound **14c** (1S,2R, in cyan) compared to the active enantiomer **14b**: the carboxylic group establishes fewer hydrogen bonds and the position of the aromatic ring causes some steric clashes with the residues in proximity to the compound. Picture adapted from ref.[25]

Although it can be speculated that the first interaction is the most likely, as it resembles that of the parent pentapeptide, we took advantage of the fairly high binding constant of **14b** to carry out saturation transfer difference (STD) measurements. STD is a NMR spectroscopic method allowing transfer of transient energy from a protein to its ligand. The distribution of saturation transferred to the ligand protons reveals the relative spatial proximity of moieties of the ligand to the protein. This information is used to determine which part of the ligand molecule is responsible for binding, assuming that the higher the STD effect, the closer the ligand proton is to the receptor surface, and, therefore, the stronger the ligand−protein interaction.[38,39] To simplify the STD data analysis, the saturation received by the different ligand protons is expressed as group epitope mapping (GEM), i.e., percentage of STD signals normalized with respect to the most saturated signal. The higher the value, the more intimate the recognition of the ligand portion is by the protein-binding pocket.[40] From experimental GEM values for compound **14b** (Figure **12 A**), it can be seen that the β-carbon of the cyclopropane ring represents the portion of the ligand that is more deeply buried inside the cavity and the ring hydrogens exhibit the highest saturation. Ethyl and phenyl groups seem to be only partially shielded from the solvent inside the binding pocket. CORCEMA-ST (complete relaxation and conformational exchange matrix analysis of saturation transfer)[41] theory was used to assess the agreement between the docking poses selected for compound **14b** and the corresponding STD spectrum. CORCEMA-ST calculates the predicted STD-NMR intensities for a proposed molecular model of a ligand− protein complex. The predicted STD intensities are expressed as GEM values and compared with the experimental data (Figure 12 B).

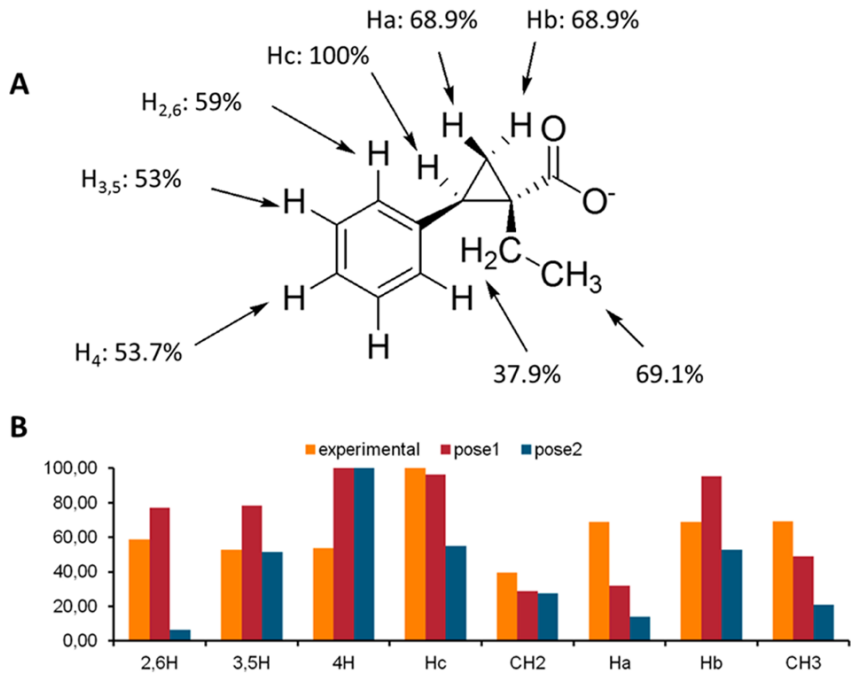


Figure 12. (A) GEM values for compound **14b** (1R,2S) from the experimental STD experiments. (B) The histogram compares the experimental STD values of **14b** (1R,2S) (in orange) to those predicted for the docking poses (in red and blue). Picture adapted from ref.[25]

The first proposed binding mode (Figure 11 A, in green) shows that predicted GEM values are more consistent (75% of the analysed protons, Pose 1, Figure 12 A) with the experimentally measured GEM values with respect to the second binding mode (Pose 2, Figure 12 A). First of all, the GEM value obtained for the hydrogen atom bound to the β-carbon of the cyclopropane ring (Hc) reveals that the docking pose is fully in agreement with the experimental value, predicting this region as being very close to the binding site surface with stable van der Waals interactions. Moreover, similar intermediate GEM values are predicted for the ethyl moiety. Small differences are found for the aromatic portion and the hydrogen atoms in γ position of the cyclopropane ring. In particular, concerning the aromatic ring, docking results predict a more

buried pose compared to the experimental findings, in particular for the *para* position. It can be hypothesized that these differences are due to the proximity of the two portions of the ligand to the flexible loop delimiting the binding pocket. This loop can adopt different conformations in solution, and the observed STD signals are realistically an average of the different conformations, whereas in the docking experiments the conformation of this loop is fixed. To take into account accurately the behaviour of this loop and the influence of protein dynamics on the predicted STD values is beyond the aim of this work and will be tackled separately. The predicted GEM values for the second possible binding mode (Figure 11 A, purple) are substantially different from the experimental ones, with a GEM value of only 55% for the hydrogen atom of the β -carbon of the cyclopropane ring, suggesting that this binding mode is probably not the one adopted in the formation of the protein− ligand complex.

With the aim of gaining further insight into the correct position of the 2-phenylcyclopropane carboxylic acid backbone and to further exploit the available volume inside the binding site to increase the binding affinity, we synthesized a compound bearing a benzyl group at the α -carbon. This substituent might stabilize the binding mode in the most favourable pose, occupying the portion of the binding site that is left empty by smaller ligands. We reasoned that an aromatic group at the α -carbon could be well allocated in this region characterized by a small hydrophobic area surrounded by polar residues, defining mildly polar areas. Moreover, the aromatic group leaves space for easily adding further substituents and maximizing the complementarity with the binding site. Therefore, compounds **15a** and its enantiomeric pure analogues **15b** (1S,2S) and **15c** (1R,2R) were synthesized and tested.

The racemic mixture exhibited a very good activity compared to the ethyl derivative and, of the two enantiomers, **15b** was found to show high affinity to both *St*OASS-A and *St*OASS-B with K_ds in the nanomolar and low micromolar ranges, respectively (Table 2). On the other hand, **15c**, as expected, was found to be several fold (700-fold) less potent than the corresponding enantiomer for *St*OASS-A. The enantiomeric preferences were further confirmed by calculation of the binding stoichiometry to *St*OASS-A. In the case of the racemic mixture the binding stoichiometry is about 2 and decreases to about 1 when compound **15b** is used for the titration. In fact, one molecule of **15b** binds for each active site of *St*OASS-A, whereas 2 molecules of the racemic mixture **15a** are needed, since it is a 50:50 mixture of **15b** and **15c**. The docking of **15b** into the binding pocket of *St*OASS-A gave us other useful information (Figure 13A). First, it allowed identifying the binding mode of the 2-phenylcyclopropane carboxylic acid backbone within the binding site. This is helpful in order to carry out further ligand optimization via decoration of the 2-phenyl ring with various functional groups. Second, the analysis of the pose of **15b** led us to speculate that there is further room for additional small functional groups in the *para* position of the α-benzyl ring.

Therefore, compounds bearing at the *para* position of the benzyl ring either an electron donor group, such as the methyl, or an electron-withdrawing group, such as chlorine, were synthesized and tested. First, we synthesized the 1S,2S enantiomers **16b** and **17b**, assuming that also in this case this stereochemistry would have been the preferred one. However, we deemed it of interest to also prepare the racemic derivatives **16a** and **17a** and the enantiomers 1S,2S **16c** and **17c**, in order to further corroborate, with additional data, the most favourable

stereochemistry (Table 2, Figure 13B).

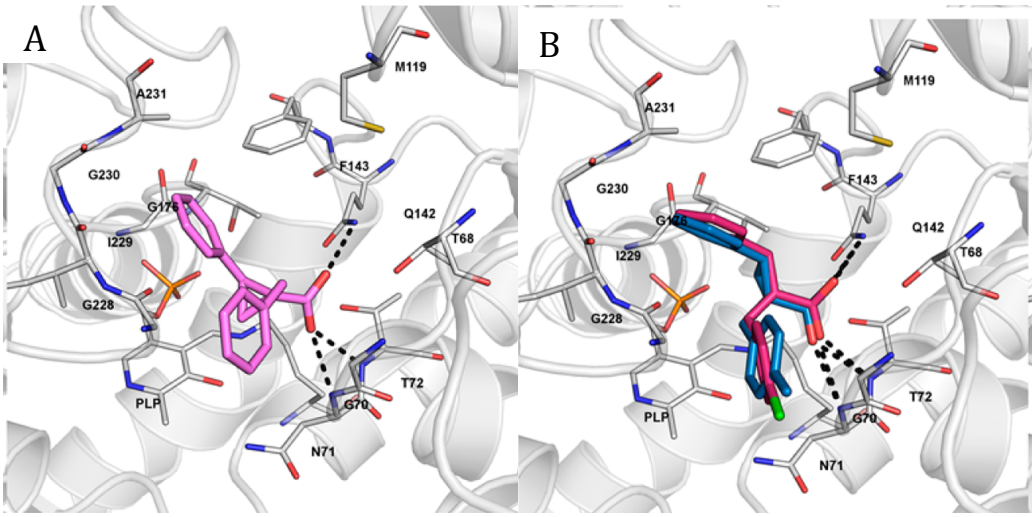


Figure 13. (A) Docking pose for compound **15b** (1S,2S); the additional aromatic moiety is located in the small lipophilic region. (B) Docking poses of derivatives **16b** (1S,2S, blue) and **17b** (1S,2S pink). Picture adapted from ref.[25]

Moreover, since the pocket accommodating the benzyl ring seemed to be large, we also synthesized the α -phenylethyl derivative **18a**, in which the aromatic ring is connected to the α -carbon of the cyclopropane through an ethyl spacer. We found that, compared to **15b**, the affinities of **16b** and **17b** toward St OASS-A were almost the same. However, both compounds showed a more than 3-fold improved affinity toward *St*OASS-B, with K_d s in the low nanomolar range. To our knowledge, this represents the strongest binder reported so far for the B isoform of *St*OASS. Indeed, a previously identified small molecule compound active on *St*OASS-B exhibited dissociation constants for the B isoform of about 30 μM, about 70-fold higher than the best compound here identified. As expected, for what concerned *St*OASS-A, the activity can be ranked based on the absolute configuration as follows: 1S,2S > rac ≫ 1R,2R. Compound **18a** was only marginally active. It can be

speculated that the phenylethyl substituent affects the interaction of the whole molecule since the ethyl linker is too long to allow the accommodation of the phenyl ring in the small lipophilic area occupied by the benzyl group. Therefore, for both the enantiomers of **18a** the phenylethyl moiety points toward the other side of the binding site, where the C-terminal residues of SAT in the 3IQG crystal structure from *Hi*OASS-A bind (Figure 14). The importance of the acid group and its interactions in the pocket of the binding site is further corroborated by the lack of activity of the ester derivative **10a**.

Finally, as predicted by our preliminary computational studies (Figure 9) substitutions of the β position of the cyclopropane ring are less tolerated. Indeed, regardless of the nature and size of the substituents introduced in such position the activity profile of the compounds did not improved, on the contrary the ligand affinities decrease (compounds **20, 21, 23, 24, 38, 39**, **40, 41, 42, 43, 44, 47**). Such experimental evidences highlight that the introduction of substituents on the β position have detrimental effects on the activity of our compounds. Therefore, discussion and further exploration of modification on this position are not any more considered.

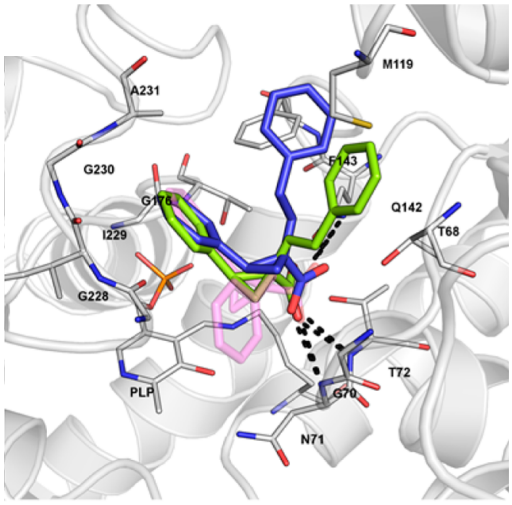


Figure 14. Docking poses of one of the enantiomers of **18a** (in slate blue and lime green), superposed to 15b (in pink, transparent): the ethyl spacer is too long to allow a perfect matching of the aromatic ring with the small hydrophobic area; therefore, the derivative seems to adopt a different conformation inside the binding site, in agreement with the loss of affinity observed for this compound. Picture adapted from ref.[25]

1.5.1 Preliminary Structure Activity Relationships (SAR).

Although the set of compounds synthesized is relatively small, some preliminary hints of SAR can be described. The vinyl moiety of **1** can be embodied in a phenyl ring and kept in a trans conformation with respect to the carboxylic acid moiety. Trapping the acid group into an ester leads to a severe decrease or even loss of the activity. Substitution at the α-carbon of the cyclopropane not only is well tolerated, as in the case of the compound **14a**, but it also remarkably enhances the affinity toward both *St*OASS-A and *St*OASS-B, as in the case of the α- benzyl substitution. Elongating the spacer between the phenyl ring and the α-carbon of cyclopropane from one to two methylene units leads to a loss of activity (**18a**, K_d OASS-A = 14 μM). The same effect is recorded with the deletion of the aromatic ring (**14a**, K_d OASS-A = 15 μM; **14b**, K_d OASS-A = 12.1 μM). The benzyl ring, along with positioning of the 2-phenylcyclopropane carboxylic acid into its most favourable pose for

interaction, is probably responsible for the π− π interaction. Moreover, the benzyl ring can be further decorated with small functional groups, such as methyl and chlorine, regardless of their electron-donor or electron-withdrawing properties. In this first round of modifications, only the *para* position was investigated. However, a medicinal chemistry campaign aimed at refining the SAR for these derivatives is currently underway in our laboratory. Stereochemistry was found to be pivotal for ensuring high affinity. In the case of the α-ethyl substituted derivatives, the levorotatory enantiomer 1R,2S was found to be more than 1000-fold less active than the counterpart. In the case of the α-benzyl substitution, the dextrorotatory enantiomer 1S,2S shows consistently an affinity higher than that of other isomers, according to the following trend: 1S,2S > rac ≫ 1R,2R. Therefore, the most active compounds are those showing a 2S stereochemistry in their structure. This is an important finding to consider in the planning of further derivatives. Indeed, since the synthetic method of cyclopropanation maintains the configuration of the starting epoxides, it can be speculated that (S)-styrene oxides will yield the active derivatives.

1.5.2 Selectivity of the Compounds toward the Two *St*OASS Isoforms.

As described above, inhibition of both isoforms of *St*OASS is important for granting a complete inhibition of cysteine biosynthesis in bacteria. In this work, we have obtained an excellent activity toward *St*OASS-B, especially with compounds **17b** and **16b**. However, considering that the ultimate target of our work is a compound showing the same activity toward both isoforms, we have rationalized the information collected in this work on the basis of the ratio between *St*OASS-B and *St*OASS-A affinities (Table 2). First, it must be noticed that, in general, the affinity

toward St OASS-B is proportional to that for St OASS-A, although always weaker. Therefore, we focused on the increase in the activity toward the A isoform. When the α -carbon is not substituted, the ratio between the activity to *St*OASS-B and *St*OASS-A is around 16 (see compound 2). Substitution with an ethyl group, although maintaining the activity toward the A isoform, led to a sharp decrease in the affinity toward the B isoform (see **14a** and **14b**, with an affinity ratio of around **50** and **70**, respectively). Substitution with the α -benzyl group reduces the activity ratio compared to the α -ethyl, but such a ratio still remains around 20 in the case of the racemic mixture and 25 for the more active enantiomer. What seems to remarkably improve the affinity ratio is the insertion of a lipophilic substituent at the para position of the benzyl ring. Indeed, for those compounds bearing a para-chloro substitution, that is **17a** and **17b**, the K_d for *St*OASS-B is only 5- and 8-fold higher than that for *St*OASS-A, respectively. Although a higher number of derivatives will shed more light into the SAR of these inhibitors, we can conclude that the α - benzyl substitution leads not only to the strongest compounds toward *St*OASS-A and *St*OASS-B so far identified, but also, if properly substituted, to the loss of selectivity between the two isoforms.

1.5.3 Inhibitory Activity of Compounds **17b** and **17c** toward the Two *St*OASS Isoforms.

After characterizing the compounds with regard to their SAR, we investigated whether there was a direct correlation between the binding affinities and the enzyme inhibitory activity. Indeed, it is expected that for compounds that act as competitive inhibitors K_d coincide with K_i. Compound **17b** was selected for the inhibition assay because of its high binding activity and its low selectivity toward the two enzyme isoforms

*St*OASS-A and B.

Compound	OASS-A		OASS-B	
	K_d (μM)	IC_{50} (μM)	K_d (μM)	IC_{50} (μM)
17b	0.054±0.008	0.099±0.004	0.42±0.06	0.50±0.03
17c	84±10	14±1	152±19	48±2

Table 3. IC50 s of Representative Inhibitors of *St*OASS-A and *St*OASS-B.

As expected for a compound that binds to the enzyme active site and competes with the substrate OAS, compound **17b** was found to inhibit both isoforms of *St*OASS with an IC_{50} comparable to the dissociation constants calculated by the direct fluorimetric titration (K_d *St*OASS-A = 0.054 μM ± 0.008 vs IC_{50} *St*OASS-A = 0.099 μM ± 0.004; K_d *St*OASS-B = 0.42 μM ± 0.06 vs IC_{50} *St*OASS-B = 0.50 μM ±0.03). A similar conclusion can be drawn for the 1R,2R isomer, which exhibits a lower affinity for the enzyme (Table 3). Therefore, we have demonstrated that there is a direct correlation between ligand binding and inhibition for both the isoforms.

1.6 Conclusions

Inhibition of cysteine biosynthesis may strongly affect the life cycle of many unicellular microorganisms and plants. O Acetylserine sulfhydrylase catalyzes the last step of cysteine biosynthesis, allowing the transfer of sulfide to acetylserine. Since mammals are not equipped with this enzyme, supplying cysteine through the reverse transsulfuration pathway from methionine, inhibitors of OASS would be highly selective and safe. Taking into account the data from our previous efforts,[14] and combining computational and spectroscopic approaches such as STD NMR, we have rationally designed and synthesized a

series of 2-phenylcyclopropane carboxylic acids to be tested against *St*OASS-A and *St*OASS-B. Overall, this medicinal chemistry campaign led to derivatives about 4000- fold and 7000-fold more active on *St*OASS-A and *St*OASS-B, respectively, than the peptidic inhibitor MNLNI, and to a reduction of the selectivity toward the two isoforms. Moreover, we have demonstrated that compounds that bind to the enzyme active site are inhibitors of both the isoforms efficiently competing with the substrate. Given the various range of K_ds, a preliminary hint of structure−activity relationships could be outlined. The carboxylic acid moiety is crucial for the activity, as its substitution with an ester led to loss of activity. The cyclopropane ring, that maintains in the suitable trans configuration the carboxylic acid moiety and the phenyl ring, proved to be a valuable structural solution, confirming the results of our previous investigations. Substitution at the α -carbon of the cyclopropane strongly affects both the potency of the compounds and their selectivity toward the two isoforms *St*OASS-A and *St*OASS-B. Bulky groups such as the para substituted and unsubstituted benzyl moieties allow to obtain potent inhibitors. However, a substituent such as chlorine at the *para* position of the benzyl ring allows a strong affinity to be coupled with a reduction of selectivity toward the two isoforms of *St*OASS, with the K_d toward OASS-B only 8 times higher than that toward OASS-A. Interestingly, the stereochemistry is extremely important in conferring high potency, and the enantiomer 1S,2S is more potent than the 1R,2R counterpart by an average 1000-fold factor. Finally, to the best of our knowledge, **16b** and **17b** are the most active compounds described so far toward *St*OASS-B. Altogether, these findings provide a solid base to further investigate this series of 2-phenylcyclopropane carboxylic acids as a valuable tool to tune the

events associated with the biosynthesis of cysteine in unicellular organisms, such as bacterial virulence and drug resistance.

1.7. Experimental Section

General Procedure for the Synthesis of Phosphonates (4–8)

Ethyl 2-(diethoxyphosphoryl)acetate (1 equiv) was added dropwise to a cooled suspension of NaH (1.1 equiv) in dry DME (2 mL/mmol). After stirring at 25 °C for 2 h the appropriate halide (1.1 equiv) was added, and the mixture was heated at 60 °C for 2 h. The reaction mixture was poured into ice water and extracted with ethyl acetate and the combined organic layers were washed with H_2O and brine, dried over Na_2SO_4, and evaporated under reduced pressure. The crude material was purified through flash chromatography eluting with dichloromethane/diethyl ether 95:5 to give the desired compounds as colorless oil in yields ranging from 54% to 68%.

General Procedure for the Cyclopropanation

To a solution of the appropriate phosphonate (2 equiv) in dry DME (5 mL/mmol) at 25 °C, *n*-buthyllithium (2.5 M in hexane, 2 equiv) was added dropwise over 5 min. After stirring at the same temperature for 30 min, the appropriate styrene oxide (1 equiv) was added portionwise and the mixture heated at 90 °C for 18 h. After cooling, saturated aqueous solution of NH_4Cl was added and the organic layers were extracted with ethyl acetate (3 × 10 mL), washed with brine, dried over anhydrous Na_2SO_4, and evaporated under reduced pressure. The oil obtained was purified through flash column chromatography eluting with petroleum ether/ethyl acetate (95:5), to give the desired product as yellowish oil in

yields ranging from 65% to 78%.

General Procedure for the Ester Hydrolysis: Synthesis of Compounds 14a–c, 15a–c, 16a–c, 17a–c, and 18a. Esters **9a**–**c**, **10a**–**c**, **11a**–**c**, **12a**–**c**, and **13a.**
(1 equiv) and $LiOH \cdot H_2O$ (4 equiv) were dissolved in a solution of THF/MeOH/H_2O (3/1/1, 1 mL/mmol) and heated under stirring in a microwave oven at 100 °C for 7 min. The reaction mixture is then evaporated *in vacuum*, and the crude is taken up with H_2O, acidified with HCl 1 N, and extracted with ethyl acetate, that is in turn washed with brine and dried over anhydrous Na_2SO_4. After the evaporation of the solvent the crude material was purified through flash column chromatography eluting dichloromethane/methanol (95:5), to give the desired product as a white solid in yields ranging from 56% to 67%.

General procedure for the synthesis of esters 32–37
To a stirred suspension of sodium hydride (60% suspension in mineral oil, 1.1 equiv) in anhydrous toluene (2 mL/mmol), the suitably substituted ethyl 2-halocarboxylate (1 equiv) and ethyl 2- acrylate (1 equiv) were added under nitrogen atmosphere at room temperature. The reaction mixture is allowed to stir at a temperature maintained between 20 and 40 _C until consumption of the starting reagents according to TLC (usually 36–72 h). The reaction is quenched by cautious addition of a small amount of methanol (1 mL) and then poured into ice-water and extracted with Et_2O. The combined organic layers are then washed with brine, dried over $MgSO_4$, and concentrated under reduced pressure. The crude residue was purified by flash chromatography eluting with a mixture of petroleum ether/ethyl acetate

in variable proportion, to give the desired product as a white solid in yields ranging from 65% to 78%.

General procedure for the synthesis of derivatives 38–43

To a solution of the substituted ethyl cyclopropanedicarboxylic ester (1 equiv), dissolved in a mixture of THF/water (1:2) at 0 °C, a solution of aqueous potassium hydroxide 1N (1 mL/mmol) was added dropwise. The reaction mixture was stirred at 50 °C until consumption of the starting material according to the TLC, and then acidified to pH 3 with 1N HCl. The aqueous phase was extracted with Et_2O, and the combined organic layers were washed with brine, dried over $MgSO_4$, and concentrated under reduced pressure, to afford the title compound as a white solid in yields ranging from 62% to 75%. The products underwent biological assays without any further purification. Following a similar procedure, but using equimolar amount of KOH and EtOH as the solvent, title compound **44** was obtained from **34** in 64% yield.

2-(hydroxymethyl)-1-phenylcyclopropane-1-carboxylic acid 45

A solution of benzyl cyanide (2.0 g, 17.1 mmol) was added dropwise over 30 min to a suspension of NaNH2 (1.54 g, 39.3 mmol) in benzene at 0 °C. After stirring for 3 h at room temperature, a solution of epichloroydrin (1.53 g, 16.6 mmol) was added to the reaction mixture over 45 min, using an ice-bath to keep the temperature in a range between 20 and 40 °C. After consumption of the limiting reagent, monitored by TLC, 1N NaOH (16.6 mL) is cautiously added dropwise to the reaction mixture that is allowed to react at 90 °C overnight. After cooling, the benzene was decanted, and the acqueous phases was extracted with dichloromethane, acidified with 2N HCl to pH 2 and extracted again with ethyl acetate. The combined organic layers were

then washed with brine, dried over $MgSO_4$ and concentrated under reduced pressure. The crude residue was purified by flash chromatography eluting with petroleum ether/ ethyl acetate (90:10), to give the desired product as a white solid in 35% overall yield.

Ethyl 2-(diethoxyphosphoryl)butanoate 4

^{1}H NMR (300 MHz, $CDCl_3$) δ: 4.17-4.03 (m, 6H); 3.81 (s, 3H); 3.23-3.09 (m, 3H); 2.25 (s, 3H); 1.30 (t, J= 7.05, 6H).

Ethyl 2-(diethoxyphosphoryl)-3-phenylpropanoate 5

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.07-7.02 (m, 5H); 4.18-4.00 (m, 6H); 3.23-3.09 (m, 3H); 2.25 (s, 3H); 1.30 (t, J= 7.05, 6H).

Ethyl 2-(diethoxyphosphoryl)-3-(4-methylphenyl)propanoate 6

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.07-7.02 (m, 4H); 4.18-4.00 (m, 6H); 3.23-3.09 (m, 3H); 2.25 (s, 3H); 1.30 (t, J= 7.05, 6H); 1.13 (t, J= 5.28, 3H).

Ethyl 2-(diethoxyphosphoryl)-3-(4-chlorophenyl)propanoate 7

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.24-7.11 (m, 4H); 4.21-4.06 (m, 6H); 3.19-3.12 (m, 3H); 1.34 (t, J= 7.08, 6H); 1.15 (t, J= 7.14, 3H).

Ethyl 2-(diethoxyphosphoryl)-3-(naphthalen-2-yl)propanoate 8

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.23-7.04 (m, 7H); 4.17-4.03 (m, 6H); 3.81 (s, 3H); 3.23-3.09 (m, 3H); 1.30 (t, J= 7.05, 6H).

(trans)-Ethyl 1-ethyl-2-phenylcyclopropanecarboxylate (9a)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.43-7.36 (m, 5H); 4.26-4.16 (m, 2H); 3.20

(d, J= 15.72, 1H); 3.02 (t, J= 8.34, 1H); 2.13-1.97 (m, 2H); 1.51-1.47 (m, 1H); 1.22 (t, J= 5.97, 3H); 0.80 (t, J= 7.32 Hz, 3H).

(1R,2S)-ethyl 1-ethyl-2-phenylcyclopropanecarboxylate (9b)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.43-7.36 (m, 5H); 4.26-4.16 (m, 2H); 3.20 (d, J= 15.72, 1H); 3.02 (t, J= 8.34, 1H); 2.13-1.97 (m, 2H); 1.51-1.47 (m, 1H); 1.22 (t, J= 5.97, 3H); 0.80 (t, J= 7.32 Hz, 3H).

(1S,2R)-ethyl 1-ethyl-2-phenylcyclopropanecarboxylate (9c)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.43-7.36 (m, 5H); 4.26-4.16 (m, 2H); 3.20 (d, J= 15.72, 1H); 3.02 (t, J= 8.34, 1H); 2.13-1.97 (m, 2H); 1.51-1.47 (m, 1H); 1.22 (t, J= 5.97, 3H); 0.80 (t, J= 7.32 Hz, 3H).

trans-Ethyl 1-benzyl-2-phenylcyclopropanecarboxylate (10a)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39-7.20 (m, 10H); 4.26-4.16 (m, 2H); 3.20 (d, J= 15.72, 1H); 3.02 (t, J= 8.34, 1H); 2.00-1.93 (m, 2H); 1.51-1.47 (m, 1H); 1.22 (t, J= 5.97, 3H).

(1S,2S)-ethyl 1-benzyl-2-phenylcyclopropanecarboxylate (10b)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39-7.20 (m, 10H); 4.26-4.16 (m, 2H); 3.20 (d, J= 15.72, 1H); 3.02 (t, J= 8.34, 1H); 2.00-1.93 (m, 2H); 1.51-1.47 (m, 1H); 1.22 (t, J= 5.97, 3H).

(1R,2R)-ethyl 1-benzyl-2-phenylcyclopropanecarboxylate (10c)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39-7.20 (m, 10H); 4.26-4.16 (m, 2H); 3.20 (d, J= 15.72, 1H); 3.02 (t, J= 8.34, 1H); 2.00-1.93 (m, 2H); 1.51-1.47 (m, 1H); 1.22 (t, J= 5.97, 3H).

trans-Ethyl 1-(4-methylbenzyl)-2-phenylcyclopropanecarboxylate (11a)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39-7.20 (m, 5H); 7.11 (d, J= 8.43, 4H); 4.26-4.16 (m, 2H); 3.20 (d, J= 15.72, 1H); 3.02 (t, J= 8.34, 1H); 2.25 (s, 3H); 2.00-1.93 (m, 2H); 1.51-1.47 (m, 1H); 1.22 (t, J= 5.97, 3H).

(1S,2S)-ethyl 1-(4-methylbenzyl)-2-phenylcyclopropanecarboxylate (11b)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39-7.20 (m, 5H); 7.11 (d, J= 8.43, 4H); 4.26-4.16 (m, 2H); 3.20 (d, J= 15.72, 1H); 3.02 (t, J= 8.34, 1H); 2.25 (s, 3H); 2.00-1.93 (m, 2H); 1.51-1.47 (m, 1H); 1.22 (t, J= 5.97, 3H).

(1R,2R)-ethyl 1-(4-methylbenzyl)-2-phenylcyclopropanecarboxylate (11c)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39-7.20 (m, 5H); 7.11 (d, J= 8.43, 4H); 4.26-4.16 (m, 2H); 3.20 (d, J= 15.72, 1H); 3.02 (t, J= 8.34, 1H); 2.25 (s, 3H); 2.00-1.93 (m, 2H); 1.51-1.47 (m, 1H); 1.22 (t, J= 5.97, 3H).

trans-Ethyl 1-(4-chlorobenzyl)-2-phenylcyclopropanecarboxylate (12a)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39-7.11 (m, 9H); 4.26-4.16 (m, 2H); 3.20 (d, J= 15.72, 1H); 3.02 (t, J= 8.34, 1H); 2.00-1.93 (m, 2H); 1.51-1.47 (m, 1H); 1.22 (t, J= 5.97, 3H).

(1S,2S)-ethyl 1-(4-chlorobenzyl)-2-phenylcyclopropanecarboxylate (12b)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39-7.11 (m, 9H); 4.26-4.16 (m, 2H); 3.20 (d, J= 15.72, 1H); 3.02 (t, J= 8.34, 1H); 2.00-1.93 (m, 2H); 1.51-1.47 (m, 1H); 1.22 (t, J= 5.97, 3H).

(1R,2R)-ethyl 1-(4-chlorobenzyl)-2-phenylcyclopropanecarboxylate (12c)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39-7.11 (m, 9H); 4.26-4.16 (m, 2H); 3.20 (d, J= 15.72, 1H); 3.02 (t, J= 8.34, 1H); 2.00-1.93 (m, 2H); 1.51-1.47 (m, 1H); 1.22 (t, J= 5.97, 3H).

trans-Ethyl 1-phenethyl-2-phenylcyclopropanecarboxylate (13a)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39-7.20 (m, 10H); 4.25-4.13 (m, 2H); 3.02 (t, J= 8.34, 1H); 3.01-2.89 (m, 1H); 2.78-2.68 (m, 1H); 2.63-2.53 (m, 1H); 2.37-2.21 (m, 1H); 2.20-2.10 (m, 1H); 1.51-1.47 (m, 1H); 0.80 (t, J= 7.32 Hz, 3H).

***trans*-1-Ethyl-2-phenylcyclopropanecarboxylic acid (14a)**

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.43–7.36 (m, 5H); 3.20 (d, *J* = 15.72, 1H); 3.02 (t, *J* = 8.34, 1H); 2.13–1.97 (m, 2H); 1.51–1.47 (m, 1H); 0.80 (t, *J* = 7.32 Hz, 3H).

^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 129.29; 128.21; 126.87; 33.32; 31.17; 21.48; 18.49; 11.59.

HRMS (ESI): m/z found $[M-H]^-$ 190.10.

(*1R,2S*)-1-Ethyl-2-phenylcyclopropanecarboxylic acid (14b)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.43–7.36 (m, 5H); 3.20 (d, *J* = 15.72, 1H); 3.02 (t, *J* = 8.34, 1H); 2.13–1.97 (m, 2H); 1.51–1.47 (m, 1H); 0.80 (t, *J* = 7.32 Hz, 3H).

^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 129.29; 128.21; 126.87; 33.32; 31.17; 21.48; 18.49; 11.59.

$[\alpha]^{20}_D$ – 100.77 (c 1, $CHCl_3$).

HRMS (ESI): m/z found $[M-H]^-$ 190.09.

(*1S,2R*)-1-Ethyl-2-phenylcyclopropanecarboxylic acid (14c)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.43–7.36 (m, 5H); 3.20 (d, *J* = 15.72, 1H); 3.02 (t, *J* = 8.34, 1H); 2.13–1.97 (m, 2H); 1.51–1.47 (m, 1H); 0.80 (t, *J* = 7.32 Hz, 3H).

^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 129.29; 128.21; 126.87; 33.32; 31.17; 21.48; 18.49; 11.59.

$[\alpha]^{20}_D$ + 100.77 (c 1, $CHCl_3$).

HRMS (ESI): m/z found $[M-H]^-$ 190.10.

***trans*-1-Benzyl-2-phenylcyclopropanecarboxylic acid (15a)**

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39–7.20 (m, 10H); 3.20 (d, *J* = 15.72, 1H); 3.02 (t, *J* = 8.34, 1H); 2.00–1.93 (m, 2H); 1.51–1.47 (m, 1H).

^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 181.64; 139.98; 136.22; 129.39; 128.73; 128.49; 128.21; 127.26; 126.04; 34.09; 32.93; 30.73; 18.49.

HRMS (ESI): m/z found $[M-H]^-$ 252.32.

(*1S,2S*)-1-Benzyl-2-phenylcyclopropanecarboxylic acid (15b)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39–7.20 (m, 10H); 3.20 (d, *J* = 15.72, 1H); 3.02 (t, *J* = 8.34, 1H); 2.00–1.93 (m, 2H); 1.51–1.47 (m, 1H).

^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 181.64; 139.98; 136.22; 129.39; 128.73; 128.49; 128.21; 127.26; 126.04; 34.09; 32.93; 30.73; 18.49.

$[\alpha]^{20}_D$ – 25.88 (c 1, $CHCl_3$).

HRMS (ESI): m/z found $[M-H]^-$ 252.31.

(*1R,2R*)-1-Benzyl-2-phenylcyclopropanecarboxylic acid (15c)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39–7.20 (m, 10H); 3.20 (d, *J* = 15.72, 1H); 3.02 (t, *J* = 8.34, 1H); 2.00–1.93 (m, 2H); 1.51–1.47 (m, 1H).

^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 181.64; 139.98; 136.22; 129.39; 128.73; 128.49; 128.21; 127.26; 126.04; 34.09; 32.93; 30.73; 18.49.

$[\alpha]^{20}_{D}$ + 25.88 (c 1, $CHCl_3$).

HRMS (ESI): m/z found $[M-H]^-$ 252.32.

***trans*-1-(4-Methylbenzyl)-2-phenylcyclopropanecarboxylic acid (16a)**

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39–7.20 (m, 5H); 7.11 (d, *J* = 8.43, 4H); 3.20 (d, *J* = 15.72, 1H); 3.02 (t, *J* = 8.34, 1H); 2.25 (s, 3H); 2.00–1.93 (m, 2H); 1.51–1.47 (m, 1H).

^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 181.64; 139.98; 136.22; 129.39; 128.73; 128.49; 128.21; 127.26; 126.04; 34.09; 32.93; 30.73; 18.49; 11.59.

HRMS (ESI): m/z found $[M-H]^-$ 266.32.

(*1S,2S*)-1-(4-Methylbenzyl)-2-phenylcyclopropanecarboxylic acid (16b)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39–7.20 (m, 5H); 7.11 (d, *J* = 8.43, 4H); 3.20 (d, *J* = 15.72, 1H); 3.02 (t, *J* = 8.34, 1H); 2.25 (s, 3H); 2.00–1.93 (m, 2H); 1.51–1.47 (m, 1H).

^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 181.64; 139.98; 136.22; 129.39; 128.73; 128.49; 128.21; 127.26; 126.04; 34.09; 32.93; 30.73; 18.49; 11.59.

$[\alpha]^{20}_{D}$ – 110.9 (c 1, $CHCl_3$).

HRMS (ESI): m/z found $[M-H]^-$ 266.40.

(*1R,2R*)-1-(4-Methylbenzyl)-2-phenylcyclopropanecarboxylic acid (16c)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39–7.20 (m, 5H); 7.11 (d, *J* = 8.43, 4H); 3.20 (d, *J* = 15.72, 1H); 3.02 (t, *J* = 8.34, 1H); 2.25 (s, 3H); 2.00–1.93 (m, 2H); 1.51–1.47 (m, 1H).

^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 181.64; 139.98; 136.22; 129.39; 128.73; 128.49; 128.21; 127.26; 126.04; 34.09; 32.93; 30.73; 18.49; 11.59.

$[\alpha]^{20}_{D}$ + 110.9 (c 1, $CHCl_3$).

HRMS (ESI): m/z found $[M-H]^-$ 266.32.

***trans*-1-(4-Chlorobenzyl)-2-phenylcyclopropanecarboxylic acid (17a)**

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39–7.11 (m, 9H); 3.20 (d, *J* = 15.72, 1H); 3.02 (t, *J* = 8.34, 1H); 2.00–1.93 (m, 2H); 1.51–1.47 (m, 1H).

^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 181.64; 139.98; 136.22; 129.39; 128.73; 128.49; 128.21; 127.26; 126.04; 34.09; 32.93; 30.73; 18.49;

HRMS (ESI): m/z found $[M-H]^-$ 286.80.

(*1S,2S*)-1-(4-Chlorobenzyl)-2-phenylcyclopropanecarboxylic acid (17b)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39–7.11 (m, 9H); 3.20 (d, *J* = 15.72, 1H); 3.02 (t, *J* = 8.34, 1H); 2.00–1.93 (m, 2H); 1.51–1.47 (m, 1H).

^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 181.64; 139.98; 136.22; 129.39; 128.73; 128.49; 128.21; 127.26; 126.04; 34.09; 32.93; 30.73; 18.49.

$[\alpha]^{20}_{D}$ – 40.77 (c 1, $CHCl_3$).

HRMS (ESI): m/z found $[M-H]^-$ 286.76.

(*1R,2R*)-1-(4-Chlorobenzyl)-2-phenylcyclopropanecarboxylic acid (17c)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39–7.11 (m, 9H); 3.20 (d, *J* = 15.72, 1H); 3.02 (t, *J* = 8.34, 1H); 2.00–1.93 (m, 2H); 1.51–1.47 (m, 1H).

^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 181.64; 139.98; 136.22; 129.39; 128.73; 128.49; 128.21; 127.26; 126.04; 34.09; 32.93; 30.73; 18.49.

$[\alpha]^{20}_{D}$ + 40.77 (c 1, $CHCl_3$).

HRMS (ESI): m/z found $[M-H]^-$ 286.76.

***trans*-1-Phenethyl-2-phenylcyclopropanecarboxylic acid (18a)**

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.39–7.20 (m, 10H); 3.02 (t, *J* = 8.34, 1H); 3.01–2.89 (m, 1H); 2.78–2.68 (m, 1H); 2.63–2.53 (m, 1H); 2.37–2.21 (m, 1H); 2.20–2.10 (m, 1H); 1.51–1.47 (m, 1H).

^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 181.64; 139.98; 136.22; 129.39; 128.73; 128.49; 128.21; 127.26; 126.04; 34.09; 32.93; 30.73; 18.49; 15.23.

HRMS (ESI): m/z found $[M-H]^-$ 266.22.

Cis (±)-diethyl 1-methylcyclopropane-1,2-dicarboxylate 32

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 1.15 (s, 3H); 1.21–1.29 (m, 6H); 1.32–1.36 (m, 1H); 1.63–1.66 (m, 1H); 2.00–2.12 (m, 1H); 4.12–4.19 (m, 4H);

HRMS (ESI) m/z found $[M-H]^-$ 200.23.

Cis (±)-diethyl 1-ethylcyclopropane-1,2-dicarboxylate 33

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 1.11 (t, J.8.34, 3H); 1.21–1.29 (m, 6H); 1.34–1.38 (m, 1H); 1.97–2.06 (m, 2H); 2.15–2.21 (m, 2H); 4.12–4.19 (m, 4H);
HRMS (ESI) m/z found $[M-H]^-$ 214.16.

Cis (±)-diethyl 1-phenylcyclopropane-1,2-dicarboxylate 34

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 1.23–1.30 (m, 6H); 1.41–1.45 (m, 1H); 2.24–2.27 (m,
2H); 4.23–4.27 (m, 4H); 7.36_7.43 (m, 5H);
HRMS (ESI) m/z found $[M-H]^-$ 262.30.

Cis (±)-diethyl 1,2-dimethylcyclopropane-1,2-dicarboxylate 35

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 1.15 (s, 6H); 1.31–1.36 (m, 6H);
2.21–2.25 (m, 2H);
4.21–4.26 (m, 4H);
HRMS (ESI) m/z found $[M-H]^-$ 214.16.

Cis (±)-diethyl 1-methyl-2-phenylcyclopropane-1,2- dicarboxylate 36

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 1.17 (s, 3H); 1.30–1.35 (m, 6H), 2.17–2.21 (m, 2H); 4.31 4.36 (m, 4H); 7.32–7.38 (m, 5H).
HRMS (ESI) m/z found $[M-H]^-$ 276.33.

Cis (±)-diethyl 1-benzyl-2-methylcyclopropane-1,2-dicarboxylate 37

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 1.15 (s, 3H), 1.34–1.38 (m, 6H), 1.93–2.00 (m, 2H); 2.19 2.23 (m, 2H); 4.35–4.39 (m, 4H); 7.48–7.56 (m, 5H);
HRMS (ESI) m/z found $[M-H]^-$ 290.30.

Cis (±)-1-methylcyclopropane-1,2-dicarboxylic acid 38

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 1.13 (s, 3H); 1.23–1.26 (m, 1H); 1.51–1.56 (m, 1H); 2.13–2.17 (m, 2H).

^{13}C-NMR (100.6 MHz, $CDCl_3$) δ: 33.21; 31.41; 23.54; 16.54.

HRMS (ESI) m/z found $[M-H]^-$ 144.13.

Cis (±)-1-ethylcyclopropane-1,2-dicarboxylic acid 39

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 1.11 (t, J.8.34, 3H); 1.26–1.31 (m, 1H), 1.97–2.06 (m, 2H), 2.15–2.21 (m, 2H).

^{13}C-NMR (100.6 MHz, $CDCl_3$) δ: 33.32; 31.17; 21.48; 18.49; 11.59.

HRMS (ESI) m/z found $[M-H]^-$ 158.15.

Cis (±)-1-phenylcyclopropane-1,2-dicarboxylic acid 40

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 1.40–1.44 (m, 1H); 2.24– 2.27 (m, 2H), 7.33_7.39 (m, 5H). ^{13}C-NMR (100.6 MHz, $CDCl_3$) δ: 129.29; 128.21; 126.87; 33.32; 31.17; 21.48.

HRMS (ESI) m/z found $[M-H]^-$ 206.19.

Cis (±)-1,2-dimethylcyclopropane-1,2-dicarboxylic acid 41

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 1.12 (s, 6H); 2.21–2.25 (m, 2H).

^{13}C-NMR (100.6 MHz, $CDCl^3$) δ: 33.32; 31.17; 21.48; 17.58.

HRMS (ESI) m/z found $[M-H]^-$ 158.12.

Cis (±)-1-methyl-2-phenylcyclopropane-1,2-dicarboxylic acid 42

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 1.21 (s, 3H); 2.12–2.15 (m, 2H); 7.31–7.37 (m, 5H).

^{13}C-NMR (100.6 MHz, CDCl3) δ: 129.18; 128.33; 126.67; 31.17; 17.58.

HRMS (ESI) m/z found $[M-H]^-$ 220.20.

Cis (±)-1-benzyl-2-methylcyclopropane-1,2-dicarboxylic acid 43

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 1.21 (s, 3H); 1.94–2.03 (m, 2H); 2.21–2.25 (m, 2H); 7.41–7.49 (m, 5H).

^{13}C-NMR (100.6 MHz, $CDCl_3$) δ: 131.123; 129.56; 126.32; 33.54; 17.34; 11.23.

HRMS (ESI) m/z found $[M-H]^-$ 234.11.

Cis (±)-2-(ethoxycarbonyl)-2-phenylcyclopropanecarboxylic acid 44

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 1.21–1.25 (t, J.7.54, 3H); 1.65–1.67 (m, 1H); 2.33–2.35 (m, 2H); 4.11–4.16 (m, 2H); 7.33–7.37 (m, 5H).

^{13}C-NMR (100.6 MHz, $CDCl_3$) δ: 130.12; 129.67; 127.44; 33.54.

HRMS (ESI) m/z found $[MH]^-$ 234.21.

2-(hydroxymethyl)-1-phenylcyclopropane-1-carboxylic acid 47

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 1.74–1.78 (m, 1H); 1.88–1.94 (m, 2H); 3.98–4.03 (m, 2H); 7.44–7.58 (m, 5H).

^{13}C-NMR (100.6 MHz, CDCl3): 130.24; 129.86; 128.21; 33.61; 32.84; 15.16.

HRMS (ESI) m/z found $[M-H]^-$ 192.20.

2. CARBONIC ANHYDRASE AS ANTIBACTERIAL DRUG TARGET

2.1 Introduction

Resistance to antibiotics belonging to several different classes is escalating and represents a worldwide problem,[42] as both Gram-negative and Gram-positive bacteria (such as among others *Staphylococcus aureus*, *Mycobacterium tuberculosis*, *Helicobacter pylori*, *Brucella suis, Streptococcus pneumoniae*, etc.,) no longer respond to many such drugs.[43] Cloning of the genomes of many bacterial pathogens offers however the possibility to explore alternative pathways for inhibiting virulence factors or proteins essential for their life cycle.[44] Among others, new possible drug targets recently explored, are represented by a class of enzymes that catalyse a simple process: the hydration of carbon dioxide to bicarbonate and protons.[45] The family of enzymes, catalysing such process, is known as carbonic anhydrases (CAs, EC 4.2.1.1), and they are all metalloenzymes.

2.2 Carbonic Anhydrase

Carbonic anhydrases (CAs; also known as carbonate dehydratases EC 4.2.1.1) are metalloenzymes ubiquitously expressed in almost all living organisms.[46] They are encoded by seven evolutionarily unrelated gene families. These are the α-CAs (present in vertebrates, bacteria, algae and cytoplasm of green plants); the β-CAs (predominantly in bacteria, algae and chloroplasts of monodicotyledons and dicotyledons); the γ-CAs (mainly in archaea and some bacteria); the δ-CAs (present in some marine diatoms); ε, ζ and η. [47–49,]

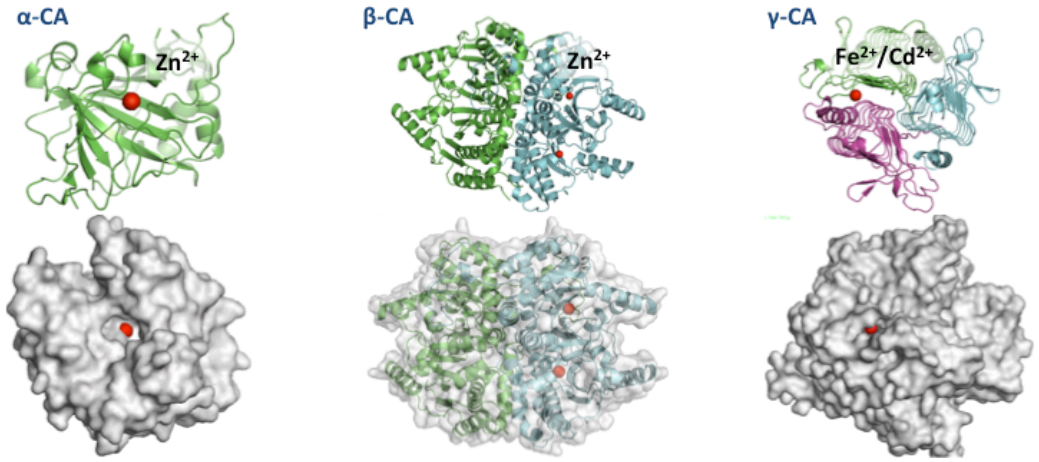


Figure 15. X-ray structures of three different families of CA.

In mammals, 16 α-CA isozymes or CA-related proteins have been described, with different catalytic activity, subcellular localization and tissue distribution.[50] On the basis of aforementioned differences, mammal α-CAs can be classified in different groups:

1. cytosolic isoforms: CA I, CA II, CA III, CA VII and CA XIII
2. membrane-bound isoforms: CA IV, CA IX, CA XII, CA XIV and CA XV
3. mithocondrial isoforms: CA VA and CA VB
4. secreted CAs: CA V

A schematic representation of the reaction catalyses by all these CA is depicted in reaction 1, where CO_2 is converted into bicarbonate ion and protons. The reaction is aided by the presence of zinc cation (Zn^{2+}). Indeed, most of the CA active sites contain Zn^{2+} which is fundamental for catalysis. This simple reaction is implicated in several physiological and pathological processes, such as respiration; electrolyte secretion in

various tissues and organs; gluconeogenesis, lipogenesis and ureagenesis, bone resorption, calcification and in cancer.[51-53]

2.2.1 Carbonic Anhydrase Reaction Mechanism

Carbonic anhydrases (CAs) catalyse the following reaction:

Reaction 1. Reaction catalysed by Carbonic Anhydrase.

The metal ion (which is a Zn^{2+} cation in all α-CAs investigated during my PhD course) is essential for catalysis.[43,44] The analysis of the X-ray crystallographic structure so far disclosed, revealed that is located at the bottom of a 15 Å deep active-site cleft (shown as a pink sphere in the figure 16A). The Zn^{2+} is coordinated by three histidine residues (His94, His96 and His119; shown in green in the figure 16A) and a water molecule/hydroxide ion.[45] Mechanicistic studies provided several clues about the reaction mechanism, where a cluster of histidine (His64, His4, His3, His17, His15 and His10) are directly involved in the protons shuttling process between the enzyme active site and the surrounding environments (Figure 16A). Gln-92 e Phe-131 which are involved in the binding of many sulphonamide/sulphamate inhibitors are highlighted in yellow.[44]

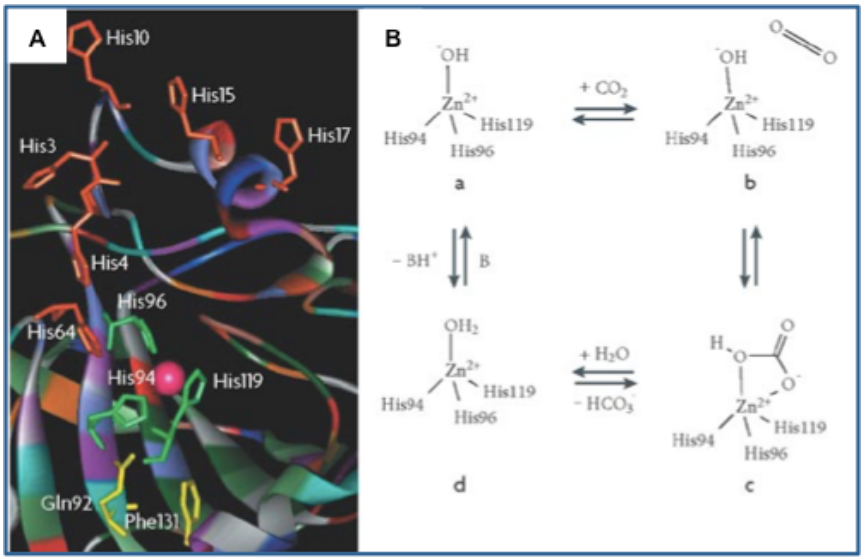
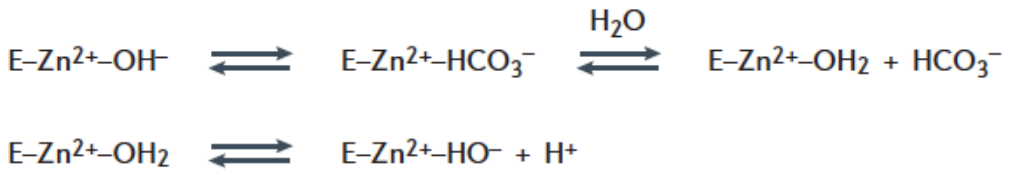


Figure 16. (A) Active site and (B) catalytic cycle of Carbonic Anhydrase (CA). Picture adapted from ref.[53]

The zinc-bound water is also engaged in hydrogen-bond interactions with the hydroxyl moiety of Thr199, which in turn is bridged to the carboxylate moiety of Glu106; these interactions enhance the nucleophilicity of the zinc-bound water molecule, and orient the substrate (CO_2) in a favourable location for nucleophilic attack[54] (Figure 16B). The active form of the enzyme is the basic one, with hydroxide bound to Zn^{2+} (inset a, in Figure 16B). This strong nucleophile attacks the CO_2 molecule that is bound in a hydrophobic pocket in its neighbourhood (the substrate-binding site comprises residues Val121, Val143 and Leu198 in the case of the human isozyme CA II) (inset b, in Figure 16B), leading to the formation of bicarbonate coordinated to Zn^{2+}(inset c, in Figure 16B). The bicarbonate ion is then displaced by a water molecule and liberated into solution, leading to the acid form of the enzyme, with water coordinated to Zn^{2+} (inset d, in Figure 16B), which is catalytically inactive.[51] To regenerate the basic form (inset a, in Figure 16B), a proton transfer reaction from the active site to the

environment takes place, which may be assisted either by active-site residues (such as His64, the proton shuttle in isozymes I, II, IV, VI, VII, IX and XII–XIV among others) or by buffers present in the medium. The process may be schematically represented by the following reactions:



The rate-limiting step in catalysis is reaction 3, that is, the proton transfer that regenerates the zinc-hydroxide species of the enzyme.[44] In the CA isozymes, catalytically very active (such as CA II, IV, VI, VII, IX, XII, XIII and XI), the process is assisted by a histidine residue placed at the entrance of the active site (His64), or by a cluster of histidines (Figure 16A), which protrudes from the rim of the active site to the surface of the enzyme, thus assuring efficient proton-transfer pathways.[43] This may explain why CA II is one of the most active enzymes known (with a kcat/Km of 1.5 x 108 M^{-1} s^{-1}), approaching the limit of diffusion-controlled processes.[42]

2.2.2 Carbonic Anhydrase Inhibition

As mentioned above, the conversion of CO_2 into bicarbonate and protons is implicated in different physiological and pathological processes. It is, therefore, not surprising if CAs represent one of the most important therapeutic targets so far explored to treat several human diseases, including glaucoma, obesity, cancer, epilepsy and osteoporosis.

The inhibition and activation of CAs are well characterized processes, with most classes of inhibitors binding to the metal cation, hampering the substrate entrance and its processing. On the contrary, activators bind at the entrance of the active site cavity, participating to the catalytic mechanism, such as assisting the proton shuttling processes between the metal ion, the bound water molecule and the environment.[55] This leads to the enhanced formation of the metal hydroxide, which is the catalytically active species of the enzyme. A general inhibition mechanism is depicted in Figure 17, where it can be appreciated how inhibitors generally bind the metal ion in their deprotonated state (as anions), leading to the formation of tetrahedral complex with the cation (E, in Figure 17) or alternatively chelating the Zn^{2+} in a trigonal bipyramidal geometry (F, in Figure 17). In the latter case a water molecule coordinates the metal centre, in addition to the inhibitor.[56] Although Figure 17 represents a general catalytic/inhibition mechanism for the α-CA family, it can be generalized also to the other CA genetic families having other metal ions into the active site cavity, such as Cd(II) or Fe(II), or for CAs characterized by different coordination patterns of the metal ion (i.e. two Cys and one His residues), as for the β - and ζ - class CAs.[57,53]

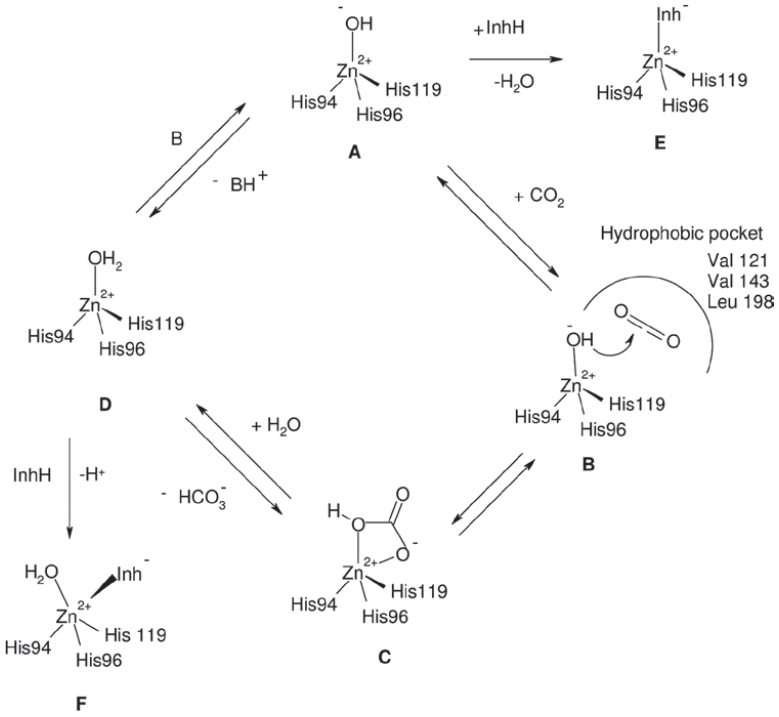
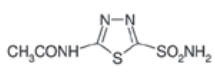


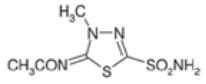
Figure 17. Catalytic and inhibition mechanisms (with zinc ion binders) of α-carbonic anhydrases (CAs) (hCA I amino acid numbering of the zinc ligands). A similar catalytic/inhibition mechanism is valid also for CAs from other classes (β-, γ- and ζ-CAs) but either the metal ion is coordinated by other amino acid residues or a Cd(II) ion is present instead of zinc at the active site. Picture adapted from ref.[53]

It is worth to mention that recently other inhibition mechanisms were described in literature, where the binding to the metal centre does not occur directly. This is the case of (*i*) polyamines, which bind to the enzyme by anchoring to the zinc coordinated water/hydroxide ion, (*ii*) coumarins act as prodrugs and bind at the entrance of the active site cavity, rather far away from the metal ion.[58] Among the well-known CAIs, so far reported, the most important class is represented by the sulphonamides one, with several compounds such as acetazolamide (**48**), methazolamide (**49**), ethoxzolamide (**50**), sulthiame (**51**), dichlorophenamide (**52**), dorzolamide (**53**), brinzolamide (**54**), sulpiride (**55**) and zonisamide (**56**) in clinical use for different human diseases

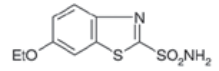
(Figure 18). Sulphamates, such as topiramate (**57**), EMATE (**58**) and irosustat (**59**) although developed independently of their potential CAI properties, are also potent CAIs and are clinically used as anti-epileptics/anti-obesity agents (topiramate) or are in advanced clinical trials as dual, steroid sulphatase inhibitors/CAIs with anti-cancer effects[59,58,53] (Figure 18). Indeed, irosustat (**59**) is the first-in-class steroid sulphatase inhibitor to be used clinically in patients with advanced hormone-dependent cancers,[59,60] but it is also a potent CAI.[60] Other compounds possessing potent CAI activity and developed for other uses are celecoxib (**60)** and valdecoxib (**61)**, originally described as cyclooxygenase 2 inhibitors.[60] Saccharin (**62**), widely used sweetener, is also a CAI, as recently reported by Supuran and Klebe's groups.[61]



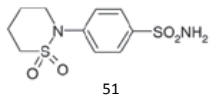
48



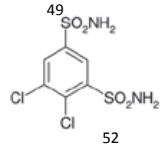
49



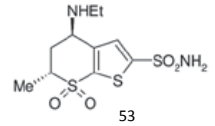
50



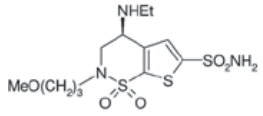
51



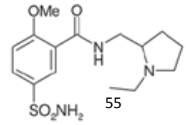
52



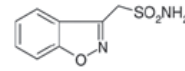
53



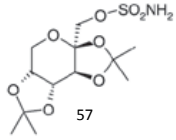
54



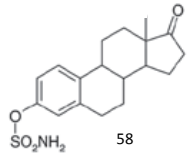
55



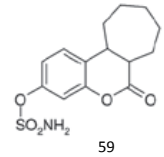
56



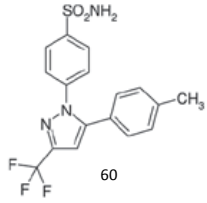
57



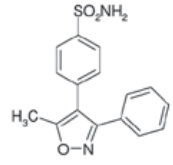
58



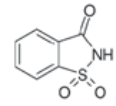
59



60



61



62

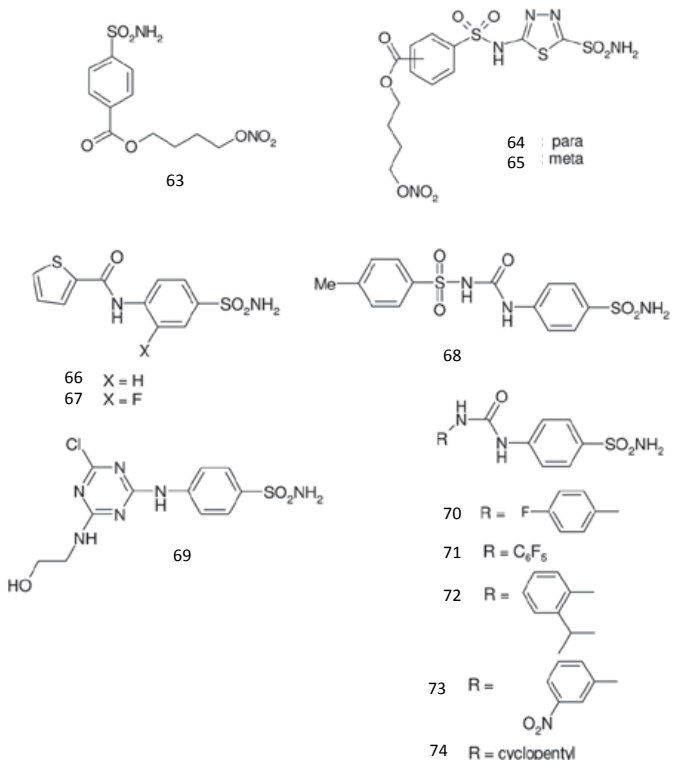


Figure 18. Clinically used/preclinical sulphonamide and sulphamate CAIs 48–74 and compounds developed in the last period.

Generally, sulphonamides and sulphamates are strong inhibitors of different CA genetic families, not only the α–class enzymes.[62,63] Their ability to inhibits several CA families is mainly due to the presence of groups which strongly interact with the metal cation of the enzyme active site (i.e. R-SO_2NH^-), generating a complex with a tetrahedral geometry (E in Figure 18).

Indeed, most of the CAIs so far reported have a sulfonamide group playing the principal role in the binding to the CA active sites. As a result of the prominent role played by the sulfonamide in the interaction with CA active sites, modifications at the core structure have only decreased effects in terms of affinity/selectivity.

Compounds **48-74** may be considered as first generation CAIs, which

indiscriminately inhibit most of the human isoforms known till date, and therefore highlighting selectivity related issues.

Therefore, most of the side effects of CAIs (of the type **48–74)** are mainly due to the non-selective inhibition of other CA isoforms, which can lead to severe adverse events such as metabolic acidosis, kidney stones, bone loss and so on.[64]

It is, therefore, not surprising if in the last years the main goal of drug design activities into the CA field is the attempt to obtain isoform-selective CAIs for the different families/isoforms involved in different human diseases. This is, however, not an easy task, considering that most isoforms possess active site architecture quite similar among each other.[65,66]

2.2.3 Bacterial Carbonic Anhydrase

Recently, it has been demonstrated that in many bacteria, CAs are essential for the life cycle of the organisms.[46] Therefore, CAs have been started to be investigated as new potential antibacterial targets, in the search for new antibiotics with novel mechanism of actions.

As mentioned above, Bacteria encode CAs belonging to the α-, β- and γ-CAs (Figure 19). However, how these enzyme classes are distributed in various bacteria is much less well understood. Indeed, some of them encode CAs belonging only to one family, others from two to three different genetic families, whereas quite a few species do not contain CAs at all,[67] as the Gram-negative bacteria belonging to the genera Buchnera and Rickettsia. As shown in Table 4, α-CAs has been detected only in the genomes of Gram-negative bacteria whereas β-and and γ-CAs have been identified in both Gram-negative and Gram-positive bacteria. Intriguingly, among the Gram-negative bacteria shown

in the Table 4, only six microorganisms (Neisseria *gonorrhoeae*, Helicobacter *pylori*, Vibrio *cholerae*, Sulfurihydrogenibium *yellowstonense*, Sulfurihydrogenibium *azorense* and Ralstonia *eutropha*) expressed CAs belonging to the α-class. Moreover, N. *gonorrhoeae* is a Gram-negative bacterium encoding for only one class of CAs, the α-CA, whereas the genome of Staphylococcus *aureus* and Enterococcus *faecalis* (Gram-positive bacteria) encode only for γ-CAs.

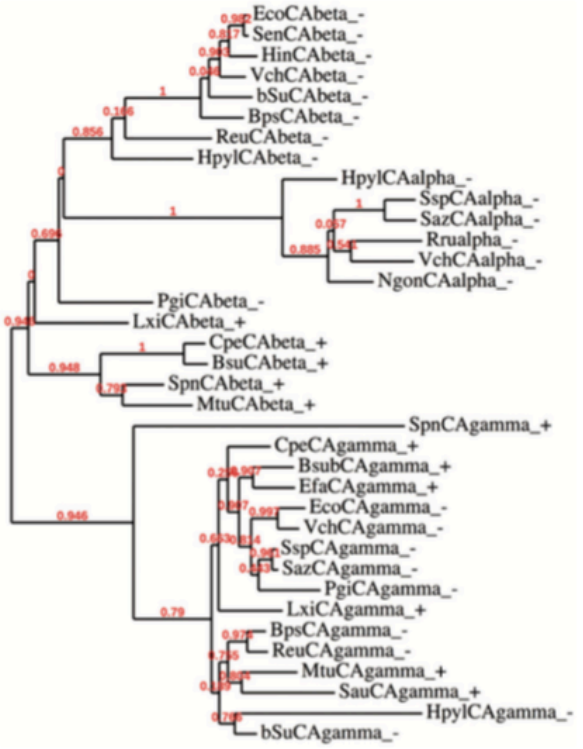


Figure 19. Phylogenetic analysis of the α-, β- and γ-CAs of the Gram- negative and -positive bacteria indicated in Table 3. The tree was constructed using the program PhyML 3.0. Branch support values are reported at branch point.

Gram negative	CA family		
Neisseria gonorrhoeae	α	–	–
Helicobacter pylori	α	β	γ
Escherichia coli	–	β	γ
Haemophilus influenzae	–	β	–
Brucella suis	–	β	γ
Salmonella enterica	–	β	–
Vibrio cholerae	α	β	γ
Sulfurihydrogenibium yellowstonense	α	–	γ
Sulfurihydrogenibium azorense	α	–	γ
Porphyromonas gingivalis	–	β	γ
Ralstonia eutropha	α	β	γ
Burkholderia pseudomallei	–	β	γ
Gram positive	CA family		
Mycobacterium tuberculosis	–	β	γ
Clostridium perfringens	–	β	γ
Streptococcus pneumoniae	–	β	γ
Bacillus subtilis	–	β	γ
Leifsonia xyli	–	β	γ
Staphylococcus aureus	–	–	γ
Enterococcus faecalis	–	–	γ

Table 4. Distribution of CA families in Gram-negative/positive bacteria (for which the genome was cloned).

2.4 Aim of the Project

Starting form preliminary works conducted by the computational group of our lab we embarked on a project aimed at the repurposing of the compound libraries available in house to look for new potential applications for our compounds. In such work, we proceeded with the analysis of the fragments and chemotypes present in our libraries by applying the maximum common substructure (MCS) decomposition approach. [68] The analysis of the data available in the literature for similar classes of chemical structures allowed us to identify the carbonic anhydrase (CA, EC 4.2.1.1) metalloenzyme family as a potential target of some of our compound series. On the basis of these results, a docking protocol was set up and thoroughly validated against all of the CA classes and isoforms crystallized so far. Such a method allowed us to identify eleven compounds as potential CA inhibitors (CAIs). Prompted by these results, several compounds were synthesized so as to (*i*) verify their ability to inhibits *in vitro* specific CA isoforms; (*ii*)

perform preliminary SAR studies; and (*iii*) pursue preliminary hit-to-lead optimization campaigns.

2.5 Preliminary Identification of Potential CAIs Among Our Compound Collection

The knowledge of the main chemotypes/fragments present in a chemical library can direct the identification of new potential targets for compounds in that library. Therefore, we extrapolated and analyzed the main fragments present in our in-house chemical library, composed of 12 main fragments/chemotypes, by applying the MCS approach. The analysis revealed that, among others, the most abundant fragments are 2-aminothiazole (**75**), 2-phenylcyclopropane-1-amino (**76**), 2 hydroxypyridinium (**77**), pyridine-N-oxide (**78**), cyclopropanecarboxylic acid (**79**), and benzoic acid (**80**) (Figure 20).

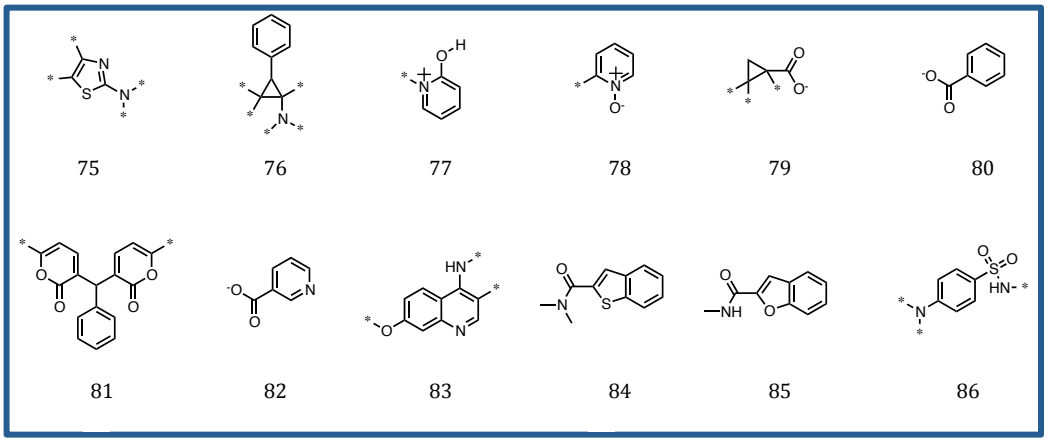


Figure 20. Structures of the main fragments identified in our compound collection library. The asterisks highlight the main points of substitution for the reported scaffolds.

Taking into account that most of the compounds available in our lab were designed as antimicrobial agents[69,70,71] or inhibitors of metalloenzymes,[72] we decided to focus our attention on all of those targets that satisfied the following criteria: 1) they are metalloenzymes

ubiquitously expressed in most living organisms, and 2) there are reported inhibitors that are characterized by fragments similar to those present in our library. An extensive knowledge-based literature search allowed us to identify the carbonic anhydrase superfamily as an eligible protein class.[73] Indeed, sulfonamides are the most common CAIs (similar to fragment **86**, in Figure 20), but recently other CAI chemotypes have been disclosed, such as 2-aminothiazole (similar to **75**), phenols[74] (similar to **77**), hydroxypyridin(thi)ones[75,76] (similar to **78**), benzoic acid (similar to **79**, **80**), coumarins[77] (similar to **81**), and polyamines[78] (not present in our database).

Encouraged by these findings, we set up a computational protocol aimed at the identification of potential CAIs among our compound collection. To this aim, we collected all the CA X-ray crystal structures available in the Protein Data Bank from different organisms, families, and isoforms (about 750 structures). A docking protocol was designed and optimized in a stepwise pipeline, with the aim of 1) identifying the best docking parameters; 2) identifying the CA X-ray crystal structures most suitable for our docking purposes, and finally 3) reliably screening the in-house collection library. Thus, all CA X-ray crystal structures with co-crystallized molecules bound within the active site were discharged from further analysis. Once we had identified the best docking parameters and the proper X-ray dataset, we screened the compound collection. When possible, we used more than one X-ray crystal structure per CA isoform, in order to include side chains and protein flexibilities in our calculations. In Figure 21, the heat map scores for both human isoforms (Figure 21A) and microbial families (Figure 21B) are depicted. From the analysis of Figure 21A, it can be noticed that 1)

several compounds have good Glide scores (>7.5_0.1 kcalmol_1, dark blue) against all but hCA III isoforms (hCA III inset in Figure 21), so hCa III is apparently not druggable by our chemotypes, and 2) some compounds display preferential interactions with specific conformations of the same hCA isoform. For example, compound C1 (inset in Figure 21) shows good scores (>7.5_0.1 kcalmol_1, dark blue) for only a decreased number of hCA II conformations, which supports the observation that small conformational changes in the CA active site can have important effects on the docking results.

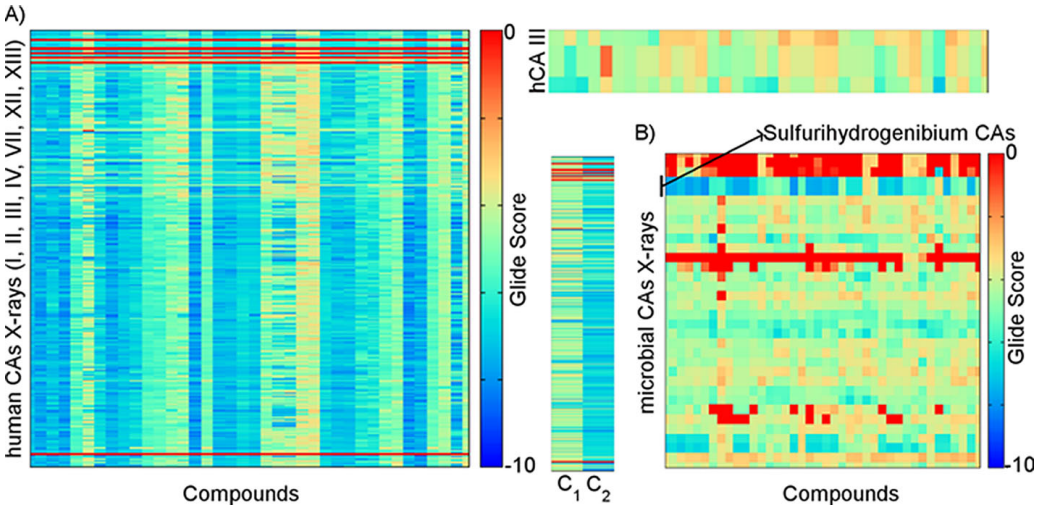


Figure 21. A) Heat map scores for the human (*h*) CAs. The red lines indicate separation between the different *h*CAs in the following order: *h*CA I, II, III, IV, VII, XII, and XIII. Highlighted in the insets are the scores for the *h*CA III isoform and for two different compounds (C1 and C2). B) Heat map scores for microbial CAs. Scores are reported as Glide scores in $kcalmol^{-1}$ (indicated by the colored bar beside the heat map). Picture adapted from ref.[79]

If docked on the microbial CAs (Figure 21B), our compounds generally displayed low Glide scores. This can be explained by the fact that most of these CAs belong to the β family, which is characterized by a narrower and more solvent-exposed active site (Figure 22.) Exceptions are represented by those microbial CAs characterized by the α-family fold type (that is, the Sulfurihydrogenibium bacterial CAs, highlighted by

the black bar and arrow in Figure 21 B), which show higher Glide scores than the others.

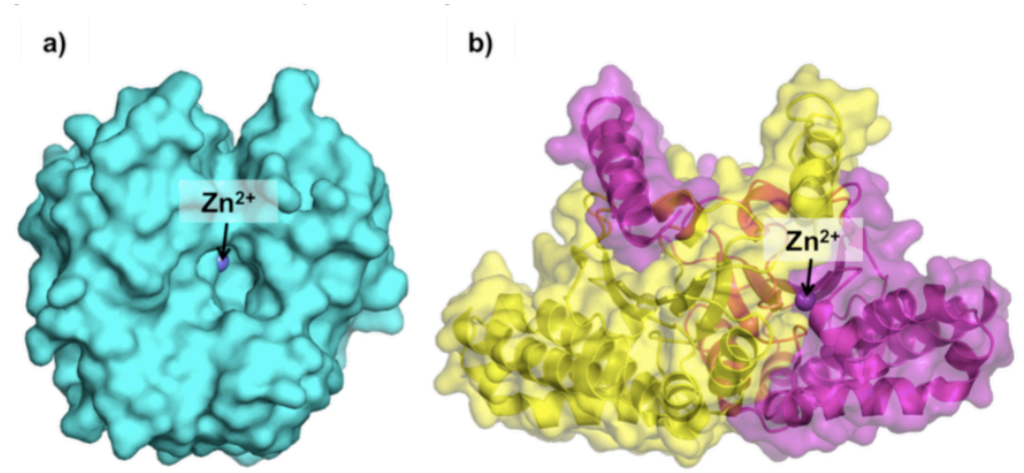


Figure 22. a) Surface representation of the α-family *h*CA II (PDB: 1CA2) and **b)** surface representation of the β-family *Ec*CA. It can be noticed that for the *h*CA II the catalytic Zn is exposed to the solvent, while for the *Ec*CA the active site is shaped by two monomers (purple and yellow surfaces and cartoons) and the catalytic Zn is at the interface between the two monomers. Picture adapted from ref.[79]

On the basis of the preliminary computational observations, we selected a small set of ligands (Figure 23) to be tested in vitro against three *h*CA isoforms (I, II, and III) and against the Sulfurihydrogenibium yellowstonensis (Ssp) CA.

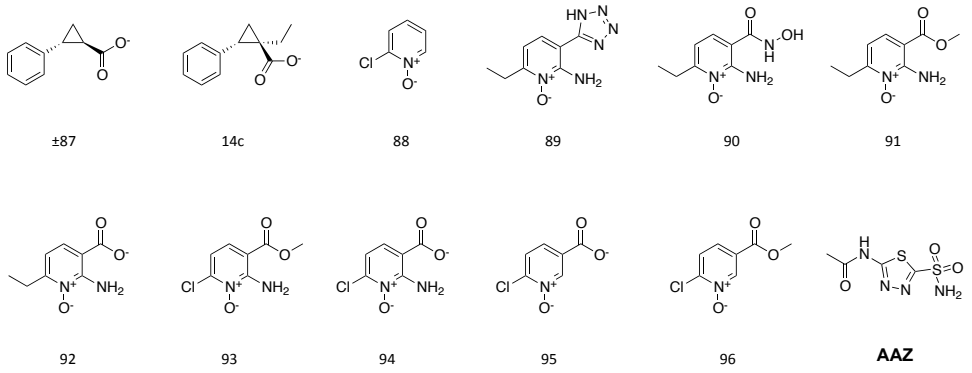
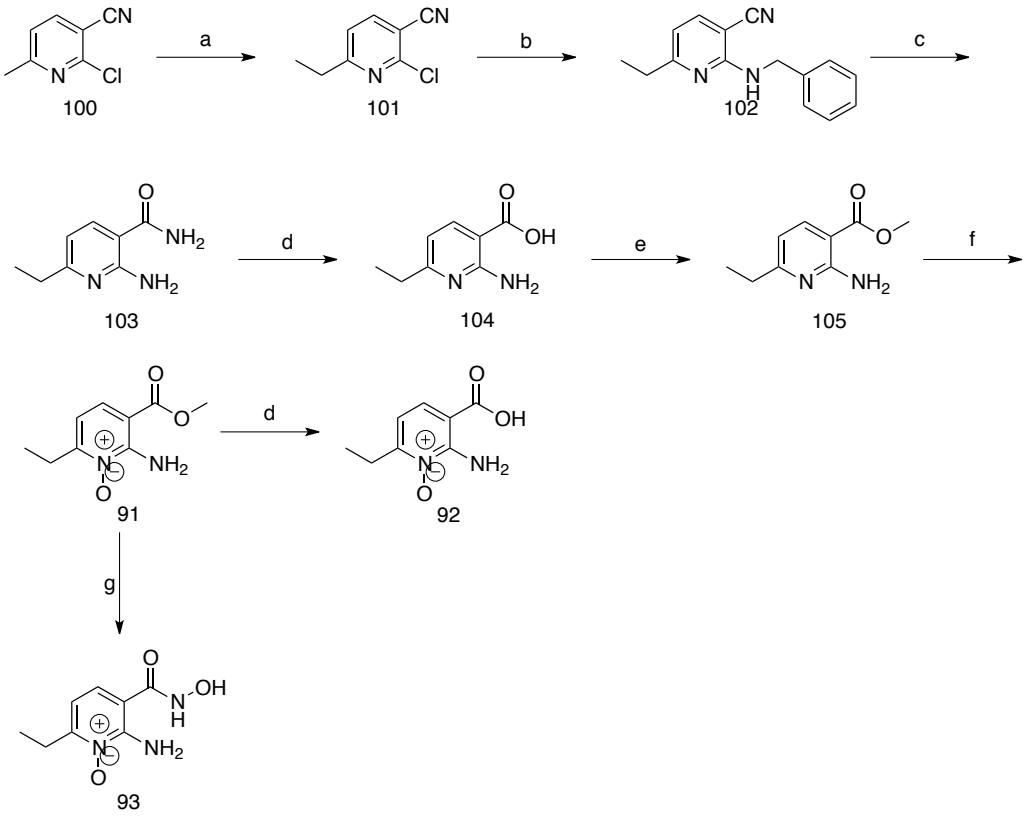


Figure 23. Structures of compounds selected to be tested in vitro.

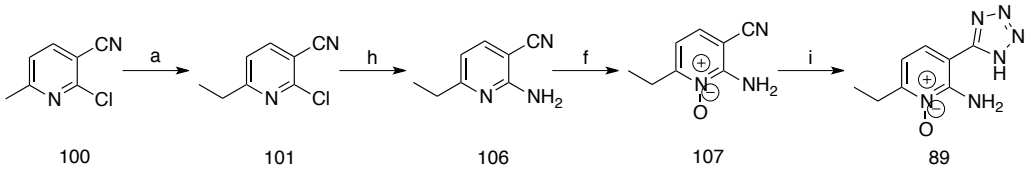
2.6 Chemistry

Compounds **90**, **91**, and **92** were obtained by starting from the commercially available 2-chloro-6-methylnicotinonitrile (**100**), which reacted with methyl iodide in the presence of NaH as a base to yield compound **101**.[72] The chlorine atom in position 2 was substituted by an amino group through the reaction of compound **101** with benzylamine, conducted in a microwave oven, which led to compound **102**. This compound was debenzylated with 98% H_2SO_4 (**103**). Subsequently, a basic hydrolysis in the presence of 10% KOH at reflux temperature[72] gave the acid **104**, which was methylated by an esterification reaction in the presence of trimethylsilyldiazomethane (TMS-diazomethane), to give compound **105**.[72] The desired compound **91** was obtained by oxidation of the pyridine nitrogen atom with methyltrioxorhenium(- VII) (MTO) in a catalytic amount; subsequent basic hydrolysis of the ester gave desired compound **92**. Finally, compound **91** through a reaction with hydroxylamine and sodium methoxide as a base[80] gave compound **93** with a high yield (Scheme 6).



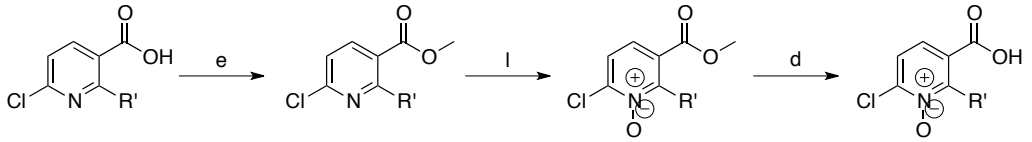
Scheme 6. Reagents and conditions: a) NaH, CH_3I, DMF, 15 min at 0°C, 4 h at RT, 68%; b) benzylamine, 150 W microwaves, 200°C, 15 min, 86%; c) 98% H_2SO_4, RT, 18 h, 98%; d) 10% KOH, 100 °C, 3 h, 98%; e) TMS-diazomethane, toluene/methanol, 0°C, 30 min, 95%; f) MTO, 35% aqueous H_2O_2, EtOH, RT, 3 h, 78%; g) $NH_2OH \cdot HCl$, MeONa, MeOH, RT, 18 h, 76%.

Compound **89** was obtained by starting from the commercially available 2-chloro-6-methylnicotinonitrile (**100**), which reacted with methyl iodide in the presence of NaH as a base to yield compound **101**.[72] The subsequent aromatic amination with ammonia saturated ethanol, led to **106**. The compound **107** was obtained by oxidation of the pyridine nitrogen atom with methyltrioxorhenium(- VII) (MTO) in a catalytic amount and then the last step was conducted in a sealed tube in the presence of sodium azide to obtain the desired compound **89** (Scheme 7).



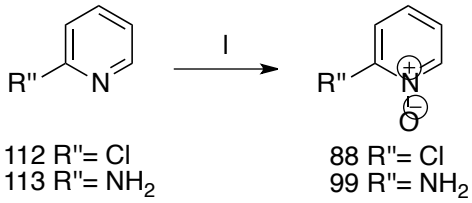
Scheme 7. Reagents and conditions: a) NaH, CH_3I, DMF, 15 min at 0 °C, 4 h at RT, 68%; h) EtOH/NH_3, sealed tube, 24 h, 200 °C, 45 %; f) MTO, 35% aqueous H_2O_2, EtOH, RT, 3 h, 78%; i) NaN_3, NH_4Cl, DMF, 130°C, 3 h, 68%.

Compounds **93**, **94**, **95** and **98** were obtained from commercially available 2 aminonicotinic acid **108** and **109**, via prior conversion into their methyl ester (**110**, **111**) by reaction with (trimethylsilyl)diazomethane in toluene/methanol, followed by methyltrioxorhenium(VII) (MTO)-catalyzed oxidation of the pyridine nitrogen (**93**, **96**) and alkaline ester hydrolysis to obtain the desired compounds **94** and **95** (Scheme 8).



Scheme 8. Reagents and conditions: e) TMS-diazomethane, toluene/methanol, 0 °C, 30 min, 95%; l) meta-chloroperoxybenzoic acid (*m*-CPBA), CH_2Cl_2, 4 h, RT, 67%; d) 10% KOH, 100 °C, 3 h, 98%.

The oxidation of compounds **112** and **113**, which are commercially available, was performed with *m*-CPBA in dichloromethane and gave the desired compounds **88** and **99** (Scheme 9).

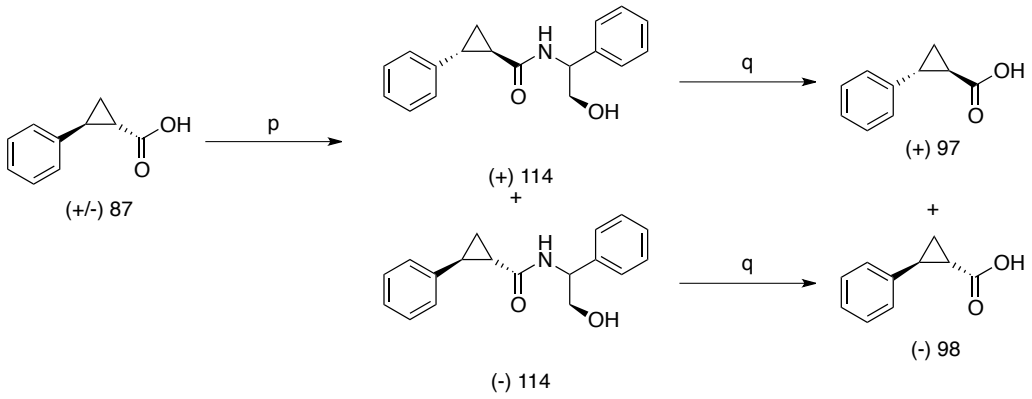


Scheme 9. Reagents and conditions: I) *m*-CPBA, CH_2Cl_2, 4 h, RT, 67%.

Compounds **14b** and **14c** were prepared through a straightforward protocol that has already been reported (Scheme 1).

Compound **87** was commercially available.

The enantiopure compounds **97** and **98** were obtained from compound **87** as a racemic mixture that was treated with (R)-(-)-2-phenylglycinol[81] to obtain the two diastereomers **(+)-114** and **(-)-114**. Compounds **(+)-114** and **(-)-114** were separated through flash chromatography and hydrolyzed in acidic conditions to give compounds **97** and **98** as enantiopure compounds (Scheme 10).



Scheme 10. Reagents and conditions: a) (R)-(-)-2-phenylglycinol, O-(benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium tetrafluoroborate (TBTU), 3-(3-dimethylaminopropyl)-1-ethylcarbodiimide (EDC)·HCl, triethylamine, CH_2Cl_2, 1 h at 0°C, 4 h at RT, 75%; b) 3N H_2SO_4, dioxane, 18 h, 100°C, 98%.

2.7 Results and Discussion

Table 5 lists the obtained inhibition constant (K_i) values for the selected compounds (Figure 23), along with the corresponding Glide scores produced by the docking calculations. The experimental results are in line with our computational predictions, and compounds **87**, **88**, **90** and **91** are endowed with μM affinity against to *h*CA I and II. As predicted by our calculation, none of the compounds was active against *h*CA III (>100 μM). Again, in line with the docking prediction, the compounds reported in Figure 23 showed K_i values in the low μM range for the bacterial α-family *Ssp*CA (Table 5).

Intrigued by the possibility to identify new CAI chemotypes with the ability to preferentially bind microbial CAs over human CAs, we decided to challenge our compounds against another CA family that is relevant for the treatment of human infections. Plasmodium *falciparum*, the causative agent of malaria, expresses a CA (*Pf*CA). The crystal structure of *Pf*CA is not yet available and so was not present in our initial structure dataset. Nevertheless, *Pf*CA represents an already identified target for treating malaria,[82,83] and it belongs to the η family, which is the most similar to the a family with respect to the other CA genetic families.

	CA type *Ki* (μM)			
Compound	***h*CAI**	***h*CAII**	***h*CAIII**	***Ssp*CA**
59	70.6±6.3 (-7.72 ± 0.46)	70.3±5.9 (-7.37±0.42)	>100	2.80±0.21 (-9.01)
14c	100±8.2 (-6.32 ± 0.37)	>100 (-6.41±0.45)	>100	1.71±0.11 (-7.73)
88	71±5.8 (-5.42±0.46)	>100 (-5.29±0.52)	>100	3.45±0.22 (-8.91)
89	100±6.3 (-4.84±0.52)	>100 (-4.55±0.38)	>100	36.30±3.5 (7.71)
90	90.3±8.0 (-8.02±0.43)	>100 (-7.35±0.42)	>100	5.96±0.42 (7.43)
91	91.7±8.3 (-7.55±0.65)	>100 (-4.81±0.54)	>100	25.70±2.4 (-8.94)
92	72±5.9 (-7.53±0.34)	>100 (-7.53±0.70)	>100	8.20±0.63 (-8.00)
93	92.7±7.5 (-7.25±0.24)	>100 (-7.12±0.59)	>100	1.89±0.13 (-9.60)
94	>100 (-7.28±0.30)	>100 (-7.58±0.67)	>100	3.47±0.20 (-7.82)
95	>100 (-7.49±0.37)	>100 (-7.18±0.45)	>100	2.59±0.16 (-7.21)
96	>100 (-5.57±0.34)	>100 (-5.62±0.42)	>100	35.20±1.8 (-8.62)
AAZ	0.25±0.008	0.012±0.001	0.17 ± 0.006	0.005±0.0001

Table 5. Inhibition constant values for *h*CA I, II, and III and *Ssp*CA. Values are the mean standard deviation (SD) of three different assays, n= 3. Glide scores (kcalmol^{-}1) obtained for each ligand are given in parentheses. For *h*CA I and II, the average Glide scores and standard deviations among different conformations are reported, whereas only one X-ray structure is available for *Ssp*CA. Glide scores >7.5_0.1 kcalmol_1 are listed in bold. AAZ: Acetazolamide.

In Table 6, the Ki values for *Pf*CA are reported and, as expected, the compounds show inhibitory activities similar to those observed for *Ssp*CA. Even more interestingly, they all have a significant selectivity over the human CAs. The affinity of our compounds, either pyridine-N-oxides or phenylcyclopropane carboxylates, is much lower than that of reference compounds such as AAZ. Nevertheless, it must be stressed

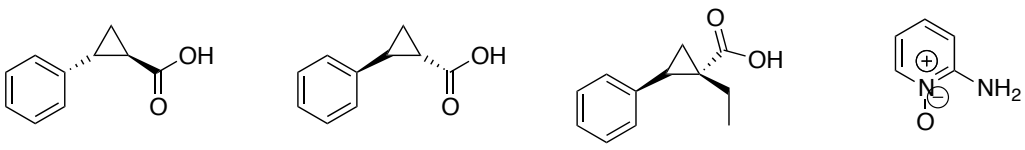
that our compounds come from a recycling approach and that they were optimized against very different targets. For instance, two widely used metrics, namely the ligand efficiency (LE) and the binding efficiency index (BEI), reported in Table 5, indicate that most of our compounds have potential for large improvement in further hit-to-lead optimization cycles.

Compound	K_i (μM)		Selectivity Fold[a]		LE[b]		BEI[c]	
	SspCA	*PfCA*	*SspCA*	*PfCA*	*SspCA*	*PfCA*	*SspCA*	*PfCA*
87	2.80±0.21	5.82±0.23	25.11	12.08	0.648	0.611	0.048	0.045
14c	1.71±0.11	6.73± 0.31	58.48	14.86	0.577	0.517	0.042	0.038
88	3.45±0.22	39.00±2.0	20.58	1.82	0.956	0.772	0.059	0.048
89	36.3±3.5	5.71±0.44	2.76	17.51	0.414	0.489	0.030	0.036
90	5.96±0.42	9.07±0.72	15.15	9.96	0.522	0.504	0.037	0.036
91	25.7±2.4	6.38±0.30	3.57	14.37	0.459	0.520	0.033	0.037
92	8.20± 0.63	6.09±0.51	8.78	11.82	0.548	0.562	0.039	0.040
93	1.89±0.13	37.60±1.84	49.04	2.46	0.616	0.477	0.040	0.031
94	3.47±0.20	39.80±2.16	28.82	2.51	0.637	0.513	0.041	0.033
95	2.59±0.16	41.30±4.01	38.61	2.42	0.711	0.558	0.045	0.035
96	35.0±1.8	8.07±0.62	2.84	12.39	0.520	0.594	0.033	0.038
AAZ	0.005±0.0001	0.17±0.003	2.40	0.07	0.894	0.729	0.052	0.043

Table 6. Inhibition constant values for SspCA and PfCA, with the corresponding fold selectivity, LE, and BEI values. [a] Values are the mean_SD of three different assays, n=3. [b] Calculated as hCA Ki (with the lowest Ki value determined)/(Ssp or Pf)CA Ki. [c] Ligand efficiency (LE)= RTpKi/N, in which N is the number of non-hydrogen atoms. [d] Binding efficiency index (BEI)=RTpKi/Mr, in which Mr is the molecular weight in kDa.

Among the compounds reported in Figure 23, compound **87** was available in our lab as a racemic mixture, whereas compound 14c was a single enantiomer (1R,2S). Therefore, we proceeded with the resolution of the racemic mixture of **87** (**97**(1S,1S) and **98**(1R,1R)) and with the synthesis of the 1S,2R enantiomer of **14c** (**14b**). Moreover, to expand the exploration around the pyridine-N-oxide scaffold, 2-amino-pyridine-

N-oxide was also synthesized (**99**, Figure 24).



(+)97 (-)98 14b 99

Figure 24. Structure of compounds **97**, **98**, **14b** and **99**.

The compounds were tested against all of the CA families and isoforms previously described, and the data are reported in Table 6. Interestingly, the analysis of the data indicates that *h*CA isoforms show an appreciable stereospecificity, because compound **97** is active against *h*CA II, whereas compound **98** is active against *h*CA I, which thus highlights the possibility to obtain isoform selectivity by modulating the stereochemistry of the phenylcyclopropane carboxylate scaffold. The introduction of the ethyl moiety on the aforementioned scaffold has a detrimental effect on the activity against the human isoforms, because both compounds **14c** and **14b** are completely inactive against the human CAs. On the other hand, such a modification of the phenylcyclopropane carboxylate scaffold led to a slight improvement of the activity for *Ssp*CA, because **14c** is more active than **14b**, **97**, and **98** (**14c**>**14b**>**98**>**97**; see Table 6 and Table 7). In contrast, *Pf*CA seems to be unaffected by the stereochemistry and the modification of the main core (phenylcyclopropane carboxylate), because compounds **87**, **14c**, **14b**, **97**, and **98** show similar activity against *Pf*CA (Table 6 and Table 7). Finally, compound **99** shows similar activity against all CAs considered in this study relative to **88**. This observation suggests that the chlorine atom and the amino group in position 2 of the pyridine-N-

oxide ring are well tolerated, whereas substitutions at the other positions of the main core play a crucial role in terms of selectivity and activity.

	CA type *Ki* (μM)				
Compound	***h*CAI**	***h*CAII**	***h*CAIII**	***Ssp*CA**	***Pf*CA**
97	> 100	41.2±2.0	> 100	8.44±0.42	6.87±0.37
98	66.3±4.8	> 100	> 100	6.19±0.25	7.48±0.39
14b	> 100	> 100	> 100	3.66±0.12	8.24±0.50
99	83.5±7.1	> 100	> 100	6.36±0.33	49.8±3.9

Table 7. Inhibition constant values for *h*CA I, II, III, and *Ssp*CA. Values are the mean SD of three different assays, n=3.

2.7.1 Binding Mode and Selectivity Against the Different CA Isoforms and Families.

Among the tested compounds, **91** is quite a promising one, because it shows the best combination of selectivity over *h*CAs, LE and BEI parameters, and, interestingly, was devoid of any activity toward the target for which was originally designed, which thus highly decreases the risk of cross-interaction with other receptors/enzymes.

In Figure 25, the proposed binding modes of **91** into the *h*CA I–III and *Ssp*CA structures are depicted. Compound **91** is nicely accommodated in the *Ssp*CA active site, with the N-oxide moiety chelating the Zn^{2+} ion. The 2-amino and 3-methyl ester groups are accommodated in the hydrophilic half of the active site, defined by the Asn62, Thr65, and Gln67 residues, whereas the 5-ethyl moiety protrudes into the hydrophobic half, which is defined by Val110, Val120, Val173, and Val182. Finally, the pyridine ring engages a π–H-bond interaction with Thr175 (Figure 25A). Through a comparison of the binding modes of the same compound in *h*CA I (Figure 25B) and *h*CA III (Figure 25C), it is

clear that amino acid differences in both the hydrophilic and hydrophobic halves (Figure 25D) of the active site explain: 1) the reduced activity of **91** for *h*CA I and 2) the lack of activity toward *h*CA III. Indeed, **91** in *h*CA I is no longer able to establish the π–H-bond interaction observed in *Ssp*CA. Moreover Asn62 and Val120 are substituted with Val62 and Ala121 in the human isoform, which leads to a less favorable interaction of **91** with the *h*CA I active site. On the other hand, **91** cannot adopt a similar binding mode in *h*CA III as a result of two important amino acid differences: 1) His64 (in *Ssp*CA) is substituted with Lys64 (in *h*CA III), which clashes with the 3-methyl ester group and 2) the *h*CA III active site is further narrowed by the presence of Phe198; these differences hamper the proper accommodation of the 6-ethyl moiety as observed in *Ssp*CA or *h*CA I.

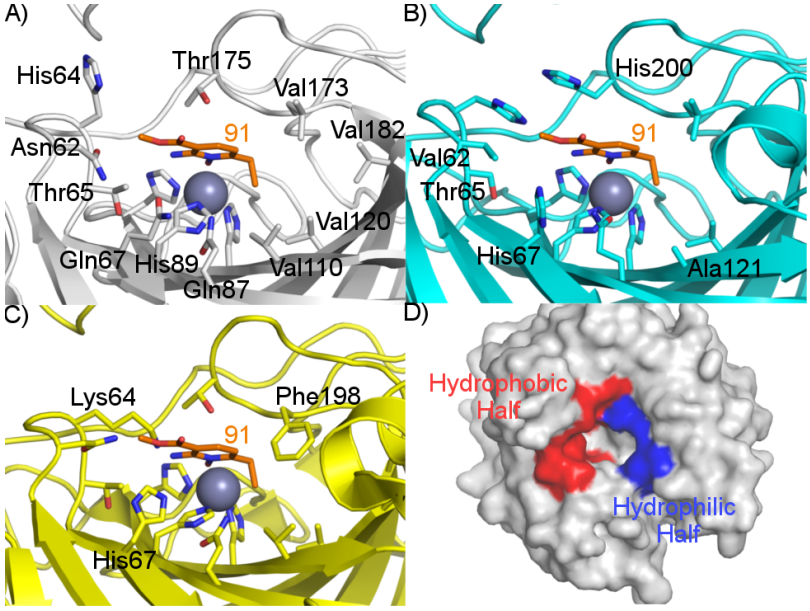


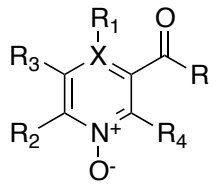
Figure 25. A) Binding mode of compound **91** into the SspCA active site (white sticks and cartoon), B) into the hCA I active site (cyan sticks and cartoon), and C) into the hCA III active site (yellow sticks and cartoon). D) Hydrophilic (blue surface) and hydrophobic (red surface) halves of hCA I (gray surface). Picture adapted from ref.[79]

Compound **91** can be considered as a prototype structure for a new class of zinc-chelating agents as CAIs. Most of the CAIs so far reported have a sulfonamide group playing the principal role in the binding to the CA active sites. As a result of the prominent role played by the sulfonamide in the interaction with CA active sites, modifications at the core structure have only decreased effects in terms of affinity/selectivity. We can hypothesize that, because the N-oxide is a weaker zinc chelating group, compounds bearing this functionality will be more amenable to modification of the organic core with a much higher potential for better control of the affinity and selectivity profile, with respect to the sulfonamide structures. As a matter of fact, the elucidation of the binding mode of **91** shed light on the structural basis underlying its selectivity profile toward the different CAs considered (Figure 25). The engagement of specific interactions with strategic residues inside the CA active site could lead to the design of selective CAIs that preferentially bind microbial CAs over human CAs.

2.8 Hit Expansion and First-Round Optimization

The data previously reported revealed how compound **91** is quite a promising one, because it shows the best combination of selectivity over *h*CAs, LE and BEI parameters. Interestingly, it was devoid of any activity toward the target for which was originally designed, which thus highly decreases the risk of cross-interaction with other receptors/enzymes. Indeed, the activity profile of compound **91** needs to be improved. Therefore, we decided to use such compound as starting point for the expansion of the series and to initiate its optimization in hit-to-lead campaigns. At this regards, the main aim of the first-round optimization activities is to improve/retain the observed affinity profile, while

improving its selectivity toward the human CA isoforms. At this regard, in Figure 26 are identified the main sites of modifications for compound **91**. Indeed, (*i*) the 2-amino N-oxide-pyridine core was retained as main scaffold of most of all the compounds synthesized, (*ii*) the possibility to modify the ester group in position 3 with functional group more stable from a methabolic point of view was considered; (*iii*) the modification of position 6 with alkyl groups of different nature was also considered; (*iv*) introduction of alkyl or aromatic groups in position 4 and 5; and (*v*) the introduction of a nitrogen atom in position 4, since the N^1-oxide pyrazine ring is a scaffold described in several approved drugs (i.e. Minoxidil®, Acipimox®, and so on …).



R = amide

R_1 = alkyl; aromatic

R_2 = alkyl

R_3 = alkyl

R_4 = -NH_2 or -H

X = C or N

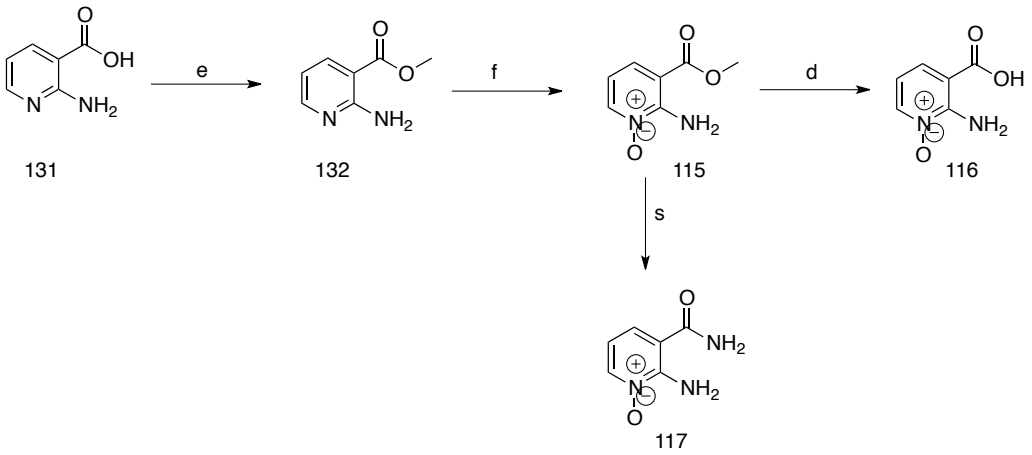
Figure 26: Hit expansion and first round optimization.

Compound	X	R	R_1	R_2	R_3	R_4
115	C	-OMe	-H	-H	-H	$-NH_2$
116	C	-OH	-H	-H	-H	$-NH_2$
117	C	$-NH_2$	-H	-H	-H	$-NH_2$
118	C	$-NH_2$	*i*-Pr	*i*-Pr	-H	$-NH_2$
119	C	$-NH_2$	*i*-Pr	-Et	-H	$-NH_2$
120	C	$-NH_2$	*i*-Pr	-Me	-Me	$-NH_2$
121	C	$-NH_2$	-Et	*i*-Pr	-H	$-NH_2$
122	C	$-NH_2$	-Et	-Et	-H	$-NH_2$
123	C	$-NH_2$	-Et	-Me	-Me	$-NH_2$
124	C	$-NH_2$	-Ph	*i*-Pr	-H	$-NH_2$
125	C	$-NH_2$	-Ph	-Et	-H	$-NH_2$
126	C	$-NH_2$	-Ph	-Me	-Me	$-NH_2$
127	N	-OMe	-	-H	-H	$-NH_2$
128	N	-OMe	-Me	-H	-H	-H
129	N	$-NH_2$	-	-H	-H	$-NH_2$
130	N	$-NH_2$	-Me	-H	-H	-H

Table 8. Hit Expansion And Optimization compounds **115**-**130**.

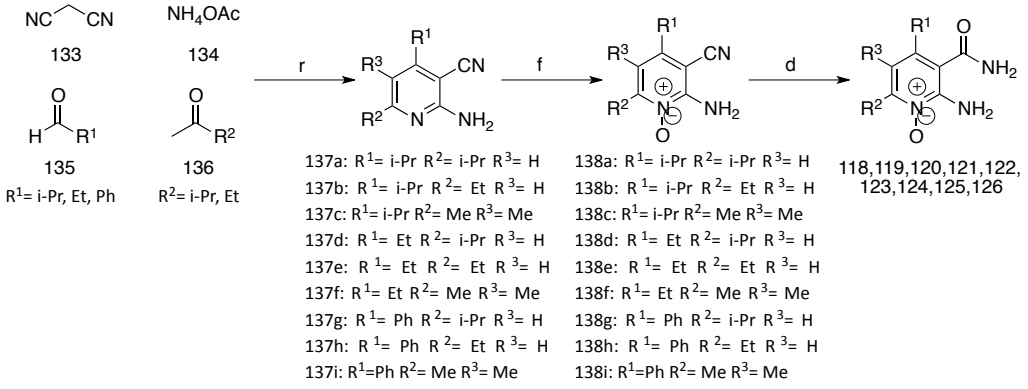
2.9 Chemistry of the Hit Expansion and Optimization Activities

Compounds **115**, **116** and **117** were obtained from commercially available 2 aminonicotinic acid **131**, via prior conversion into their methyl ester **132** by reaction with (trimethylsilyl)diazomethane in toluene/methanol, followed by methyltrioxorhenium(VII) (MTO)-catalyzed oxidation of the pyridine nitrogen (**115**) and alkaline ester hydrolysis to obtain the desired compound **116**. The reaction of **115** with a solution of ammonium hydroxide led to desired compound **117** (Scheme 11).



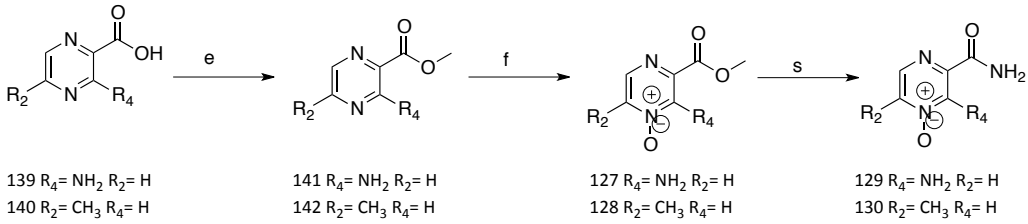
Scheme 11. *Reagents and conditions.* e) TMS-diazomethane, toluene/methanol, 0 °C, 30 min, 95%; f) MTO, 35% aqueous H_2O_2, EtOH, RT, 3 h, 78%; d) 10% KOH, 100 °C, 3 h, 98%; s) NH_4OH, rt, 3h, 98%.

In some cases, due to the lack of useful pyridine-based starting materials from readily available commercial sources, the de novo construction of the pyridine ring, with appropriate substituents, was necessary. A reported microwave-accelerated, one-pot, multicomponent reaction (MCR)[84] was exploited for the synthesis of compounds **118−126**, with slight modifications to the original procedure (Scheme 12). Irradiating a microwave tube containing an appropriate aldehyde (aromatic or aliphatic, **135**), malononitrile (**133**), an enolizable, aliphatic ketone (**136**), and ammonium acetate (**134**) as an ammonia source, with neither solvent nor catalyst, for 15 min at 100 ° C, provided key pyridine products **137a−i** in 37− 65% isolated yield. Oxidation of pyridine nitrogen of MCR-derived intermediates (**138a-i**), followed by hydrolysis of the cyano group, afforded compounds **118−126**.



Scheme 12. *Reagents and conditions:* r) appropriate aldehyde, appropriate ketone, malononitrile, ammonium acetate, μW, 100°C, 10 min, 37-65%; f) MTO, 35% aqueous H_2O_2, EtOH, RT, 3 h, 78%; d) 10% KOH, 100 °C, 3 h, 98%.

Compounds **127**, **128**, **129** and **130** were obtained from commercially available compounds **139** and **140**, via prior conversion into their methyl esters **141** and **142** by reaction with (trimethylsilyl)diazomethane in toluene/methanol, and methyltrioxorhenium(VII) (MTO)-catalyzed oxidation of the pyridine nitrogen to obtain desired compounds **127**, **128**. The reaction of **127** and **128** with a solution of ammonium hydroxide led to desired compounds **129** and **130**, respectively (Scheme 13).



Scheme 13. *Reagents and conditions.* e) TMS-diazomethane, toluene/methanol, 0 °C, 30 min, 95%; f) MTO, 35% aqueous H_2O_2, EtOH, RT, 3 h, 78%; s) NH_4OH, rt, 3h, 98%.

2.10 Conclusions

In the context of the economic downturn and the decreased success rate in drug development programs that is hampering classical drug

discovery activities, medicinal chemists have to identify alternative strategies to overcome productivity problems.[85,86] In academia, one possibility can be represented by the ability to disclose new purposes for compound libraries already available in house. By applying a rational target-focused repurposing approach, we were able to identify a new purpose for some compounds in our whole collection. First of all, we proceeded with the analysis of the fragments and chemotypes present in our libraries by applying the MCS approach.[87] The analysis of the data available in the literature for similar classes of chemical structures allowed us to identify the carbonic anhydrase (EC 4.2.1.1) metalloenzyme family as a potential target of our compound libraries. Modeling studies, together with in vitro assays, allowed us to identify new CAI chemotypes, which are characterized by a low mm affinity for microbial CAs. The obtained modest Ki values are reasonable because the molecules used were not designed to target primarily CAs. Even if the activity profile of the compounds needs to be improved, the identified molecules may represent excellent hits to be further optimized in hit-to-lead campaigns. Indeed, compound **91** seems to be the most promising one, and the analysis of its binding mode in the different CAs allowed us to pave the way for the design of potential selective CAIs that preferentially inhibit microbial CAs. The combination of the herein identified zinc-chelating group with organic cores that are properly decorated may further increase activity and selectivity against CAs of human pathogens characterized by the α- and η-family fold types.

Finally, we started the hit expansion and hit-to-lead campaigns, with the aim to improve the activity profile of compound **91**, while increasing the selectivity profile of the series toward the human CA isoforms. At this regards we optimized several reactions, which allowed us to easily

modify the 2-amino-N-oxide-pyridine core, so as to generate structurally diverse libraries based on compound **91.** Indeed, 16 compounds were synthetized and the biological assays are ongoing in the laboratory of Prof. Supuran.

2.11 Experimental Section

General procedure for the esterification of nicotinic acids: (Trimethylsilyl) diazomethane (2.0 M in diethyl ether; 2.00 equiv) was added dropwise, over a period of 15 min, to a stirred, cooled (0 °C) suspension of substrate (1.00 equiv) in dry toluene/methanol (3:2 v/v; 10 $mLmmol^{-1}$) under a nitrogen atmosphere. The reaction mixture was stirred at this temperature for 30 min. At the end of this time, the reaction mixture became a clear, yellow solution. TLC, by elution with chloroform/methanol (9:1), showed complete consumption of the starting material. The reaction was stopped. The solvents were evaporated under reduced pressure. The crude residue, dissolved in chloroform, was washed with saturated aqueous $NaHCO_3$ and brine. The organic phase was dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure. The desired product was purified by silica flash chromatography, by elution with chloroform, then chloroform/methanol (95:5). Compounds **105**, **110**, **111**, **132**, **141** and **142** were obtained by this method (75–95% yield).

General procedure for the oxidation of the pyridine nitrogen atom: Method A: A 1:2 (v/v) solution of 35% (w/w) aqueous H_2O_2 in absolute ethanol was stirred over anhydrous Na_2SO_4 (5 g/30 mL) at room temperature for 3 h. After filtration, MTO (0.10 equiv) was added to this oxidant solution (4 mLmmol_1), followed by substrate (1.00 equiv). The

reaction mixture was stirred at room temperature for 4 h, after which TLC, by elution with chloroform/methanol (9:1), showed complete consumption of starting material. The reaction was stopped. Brine was added to the reaction mixture, and it was extracted with chloroform. The organic phase was dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure. The desired compounds were purified by silica gel flash chromatography, by elution with chloroform, then chloroform/ methanol (from 99:1 to 95:5). Compounds **91**, **107**, **115**, **127**, **128** and **138a-i** were obtained by this method (68–88% yield).

General procedure for the oxidation of the pyridine nitrogen atom: Method B: The appropriate substrate (1 equiv) was combined with *m*-CPBA (2 equiv) in dichloromethane (4 mLmmol_1), and the mixture was stirred at room temperature for 4 h. After this time, TLC, by elution with dichloromethane/methanol (95:5), showed complete consumption of starting material. The solution was washed with saturated aqueous $NaHCO_3$, and the solvent was evaporated under reduced pressure. The product was purified by silica gel flash chromatography, by elution with dichloromethane/methanol (95:5). Compounds **93**, **96**, **88** and **99** were obtained by this method (56–67% yield).

General procedure for the hydrolysis of methyl esters and amides to the corresponding carboxylic acids: A fine suspension of the substrate in 10% (w/v) aqueous KOH (1 $mLmmol^{-1}$) was heated at reflux and stirred at this temperature for 3 h. TLC after this time showed complete consumption of the starting material. The reaction was stopped. After cooling at room temperature, the clear, colourless solution so obtained was cooled to 0 8C and carefully acidified to pH 4–

5, with 3N aqueous HCl. A white precipitate immediately formed. The desired product was collected by filtration in vacuum in quantitative yield. Compounds **92**, **94**, **95**, **104** and **106** were obtained by this method.

Synthesis of 2-chloro-6-ethylnicotinonitrile (101): Sodium hydride (60% dispersion in mineral oil; 1.50 equiv) was carefully added, under a nitrogen atmosphere, to a stirred, cooled (0 °C) solution of 2-chloro-6-methylnicotinonitrile (**100**; 1.00 equiv) in dry N,N'-dimethylformamide (2.5 $mLmmol^{-1}$). After the reaction mixture had stirred at this temperature at this temperature for 15 min, iodomethane (4.00 equiv) was added. The reaction mixture was allowed to warm to room temperature and stirred for 4 h under nitrogen. TLC after this time, by elution with petroleum ether/ethyl acetate (7:3), showed almost complete consumption of the starting material. The reaction was stopped. The reaction mixture was cooled to 0 °C, then water was carefully added. The mixture was extracted with diethyl ether. The organic phase was washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure. The crude residue was purified by silica gel flash chromatography, by elution with petroleum ether/ethyl acetate (from 9:1 to 8:2) to afford the target product (68%) as a colourless oil.

Synthesis of 2-(benzylamino)-6-ethylnicotinonitrile (102): A 10 mL microwave tube was charged with 2-chloro-6-ethylnicotinonitrile (**101**, 1.00 equiv) and benzylamine (2.00 equiv). The tube was placed in a microwave and irradiated at 200 °C for 15 min (maximum power input, 150 W; maximum pressure, 160 PSI; power max, ON; stirring, ON). TLC

after this time, by elution with petroleum ether/ ethyl acetate (9:1), showed almost complete consumption of the starting material. The reaction was stopped. The reaction mixture was cooled to 0 °C, then 1 N HCl aqueous solution was carefully added. The mixture was extracted with chloroform. The organic phase was washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure. The crude residue was purified by silica gel flash chromatography, by elution with petroleum ether/ethyl acetate (from 98:2 to 9:1), to afford the target product (86%) as a colourless oil.

Synthesis of 2-amino-6-ethylnicotinamide (103): H_2SO_4 (95%) was carefully added to compound **102**, and the dark mixture was stirred at room temperature for 18 h. TLC after this time, by elution with petroleum ether/ethyl acetate (9:1), showed almost complete consumption of the starting material. The reaction was stopped. The reaction mixture was cooled at 0 °C, then 1M NaOH aqueous solution was carefully added. The mixture was extracted with chloroform. The organic phase was washed with brine, dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude residue was used for the next step without further purification.

Synthesis of 2-amino-6-ethyl-3-(hydroxycarbamoyl)pyridine 1-oxide (93): Compound **91** (1.00 equiv) was added carefully, dropwise through dropping funnel, to a suspension of $NH_2OH{\cdot}HCl$ (1.5 equiv) and MeONa (1.5 equiv) in MeOH at 0 °C. The reaction was stirred at room temperature for 18 h. TLC after this time, by elution with dichloromethane/Methanol (85:15), showed complete consumption of the starting material. The reaction was stopped. The reaction mixture

was cooled to 0 °C and 1N HCl aqueous solution was carefully added. The mixture was extracted with chloroform. The organic phase was washed with brine, dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude residue was purified by silica gel flash chromatography, by elution with chloroform/methanol/formic acid (8.7:1:0.3), to afford the target product (76%) as a white solid.

Synthesis of 2-amino-6-ethylnicotinonitrile (106): A solution of the substrate in ammonia-saturated ethanol (1 $mLmmol^{-1}$) was placed in a sealed tube, heated to 200 °C, and maintained at this temperature for 24 h. After the mixture had cooled to room temperature, the solvent was removed by evaporation in vacuum, and the desired product was purified by silica gel flash chromatography, by elution with petroleum ether/ethyl acetate (from 9:1 to 6:4), to afford the target product (45 %).

Synthesis of 6-ethyl-3-(1 H-tetrazol-5-yl)pyridin-2-amine (89): A sealed tube was charged with 2-amino-3-cyano-6-ethylpyridine 1- oxide (**107**; 1.00 equiv), NaN_3 (3.00 equiv), NH_4Cl (3.00 equiv), and dry DMF (1 $mLmmol^{-1}$). The reaction was stirred at 130 °C for 3 h. TLC after this time, by elution with dichloromethane/Methanol (85:15), showed complete consumption of the starting material. The reaction was stopped. The reaction mixture was cooled at 0 °C, and water was added. The mixture was extracted with chloroform, and the organic phase was washed with brine, dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude residue was purified by silica gel flash chromatography, by elution with chloroform/methanol/formic acid (8.7:1:0.3), to afford the target product (68%) as a white solid.

Synthesis of N-(2-hydroxy-1-phenylethyl)-2phenylcyclopropane carboxamide ((+/-)114): (R)-(-)-2-Phenylglycinol (1.5 equiv), TBTU (1 equiv), EDC·HCl (1.5 equiv), and triethylamine (1 equiv) were added to a solution of compound **87** (1 equiv) in dichloromethane (1 $mLmmol^{-1}$). The reaction mixture was stirred for 1 h at 0 °C and then for 4 h at room temperature. TLC after this time, by elution with dichloromethane/methanol (9:1), showed almost complete consumption of the starting material. The reaction was stopped. An aqueous solution of NH_4Cl was carefully added to the reaction mixture, which was then extracted with chloroform. The organic phase was washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure. The crude residue was purified by silica gel flash chromatography, by elution with dichloromethane/methanol (from 99:1 to 95:5) to afford the two diastereomers **(+)-114** and **(-)-114** as white solids in 43 and 31% yields, respectively.

General procedure for the hydrolysis of N-(2-hydroxy-1-phenylethyl)- 2-phenylcyclopropanecarboxamides: A solution of substrate in 3N H_2SO_4 and dioxane (1:1; 1 $mLmmol^{-1}$) was stirred at 100 °C for 18 h. TLC after this time, by elution with dichloromethane/ methanol (9:1), showed almost complete consumption of the starting material. The reaction was stopped. Water was carefully added to the reaction mixture, which was then extracted with chloroform. The organic phase was washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure. The crude residue was purified by silica gel flash chromatography, by elution with dichloromethane/methanol (from 99:1 to 95:5), to afford the target

compounds as white solids (98 %). Compounds **97** and **98** were obtained by this method.

General Procedure for the One-Pot Synthesis of 2-Amino-3-cyanopyridines. A 10 mL microwave tube was charged with malononitrile (**133**, 1.00 equiv), appropriate aldehyde (**135**, 1.00 equiv), appropriate ketone (**136**, 1.00 equiv), and ammonium acetate (**134**, 1.50 equiv). The tube was placed in a microwave oven and irradiated at 100 °C for 15 min (maximum power input, 150 W; maximum pressure, 160 PSI; power max, ON; stirring, ON). The resulting crude was purified by silica gel flash chromatography, eluting with appropriate petroleum ether/ethyl acetate mixtures, affording the target products in 37−65% yield. Compounds **137a-i** were obtained by this method.

General Procedure for the Hydrolysis of Nitriles to the Corresponding amide. A fine suspension of the substrate in 10% (w/v) aqueous KOH (1 mL/mmol) was heated to reflux and stirred at this temperature for 3 h. TLC after this time, eluting with chloroform/ethanol/formic acid, 8.7/1.0/0.3, showed complete consumption of the starting material. The reaction was stopped. After cooling to room temperature, the clear, colourless solution so obtained was cooled to 0 °C and carefully acidified to pH 4−5, with 3 N aqueous HCl. A white precipitate immediately formed. The desired product was collected by filtration in vacuum in quantitative yield. Compounds **118**, **119**, **120**, **121**, **122**, **123**, **124**, **125** and **126** were obtained by this method.

General Procedure for the Conversion of Methyl Esters to the Corresponding amide. A solution of the substrate in NH_4OH solution (1 mL/mmol) was stirred at room temperature for 3 h. TLC after this time, eluting with chloroform/ethanol/formic acid, 8.7/1.0/0.3, showed complete consumption of the starting material. The reaction was stopped. A white precipitate immediately formed. The desired product was collected by filtration in vacuum in quantitative yield. Compounds **117**, **129** and **130** were obtained by this method.

2-chloro-6-ethylnicotinonitrile (101)

^{1}H NMR ($CDCl_3$, 400 MHz) δ: 7.90-7.92 (d, 1H); 7.24-7.26 (d, 1H); 2.85-2.90 (q, 2H); 1.29-1.33 (t, 3H).

^{13}C NMR ($CDCl_3$, 100.6 MHz) δ: 169.01; 151.97; 142.58; 120.77; 115.03; 107.65; 31.49; 13.12.

HRMS (ESI): found m/z $[M + H]^+$ 167.3.

2-(benzylamino)-6-ethylnicotinonitrile (102)

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.58 (d, J=7.8 Hz, 1 H); 7.41–7.30 (m, 5 H); 6.51 (d, J= 7.8 Hz, 1H); 5.45 (br s, 1H); 4.75 (d, J=5.6 Hz, 2 H); 2.76–2.68 (q, 2 H); 1.30–1.25 (t, 3H).

HRMS (ESI): m/z found $[M+H]^+$ 238.30.

2-amino-6-ethylnicotinamide (103)

^{1}H NMR (300 MHz, DMSO) δ: 7.86 (d, J=7.7 Hz, 1 H); 7.80 (br s, 2H); 7.16 (br s, 2 H); 6.43 (d, J= 7.7 Hz, 1 H); 2.55–2.50 (q, 2H); 1.19–1.14 (t, 3H).

HRMS (ESI): m/z found $[M+H]^+$ 166.19.

2-amino-6-ethylnicotinic acid (104)

^{1}H NMR (300 MHz, DMSO) δ: 13.48 (br s, 1 H); 7.94 (d, J=7.7 Hz, 1 H); 7.18 (br s, 2 H); 6.46 (d, J= 7.7 Hz, 1 H); 2.57–2.50 (q, 2H); 1.19–1.14 (t, 3H).
HRMS (ESI): m/z found $[M+H]^{+}$ 167.18.

Methyl 2-amino-6-ethylnicotinate (105)

^{1}H NMR (300 MHz, DMSO) δ: 7.96 (d, J=8.0 Hz, 1 H); 7.11 (br s, 2 H); 6.51 (d, J=8.0 Hz, 2 H); 3.79 (s, 3 H); 2.61–2.54 (q, 2H); 1.20–1.15 (t, 3H).
HRMS (ESI): m/z found $[M+H]^{+}$ 181.20.

Methyl 2-amino-6-ethylnicotinate 1-oxide (91)

^{1}H NMR (300 MHz, DMSO) δ: 7.73 (br s, 2 H); 7.68 (d, J=8.4 Hz, 1H); 6.71 (d, J= 8.4 Hz, 1H); 3.86 (s, 3H); 2.88–2.80 (q, 2 H); 1.24–1.89 (t, 3H).
^{13}C NMR (75.5 MHz, DMSO) δ: 168.3; 154.7; 152.2; 128.2; 109.6; 106.7; 52.8; 24.7; 10.8.
HRMS (ESI): m/z found $[M+H]^{+}$ 197.20.

2-Amino-6-ethyl-N-hydroxynicotinamide 1-oxide (93)

^{1}H NMR (300 MHz, DMSO) δ: 11.20 (br s, 1 H); 9.34 (br s, 1 H); 7.73 (br s, 2 H);
7.62 (d, J=8.4 Hz, 1H); 6.47 (d, J=8.4 Hz, 1 H); 2.79–2.73 (q, 2H); 1.21–1.17 (t, 3H).
^{13}C NMR (75.5 MHz, DMSO) δ: 161.2; 155.7; 152.0; 127.7; 110.0; 106.7; 24.7; 10.8.
HRMS (ESI): m/z found $[M+H]^{+}$ 198.19.

2-Amino-6-ethylnicotinic Acid 1-Oxide (92)

^{1}H NMR (400 MHz, DMSO) δ: 13.49 (br s, 1H); 7.75 (br s, 2H); 7.66-7.68 (d, 1H); 6.66-6.69 (d, 1H); 2.82-2.85 (q, 2H); 1.19-1.23 (t, 3H).

^{13}C NMR (100.6 MHz, DMSO) δ: 168.22; 154.74; 152.12; 128. 19; 109.56; 106.75; 24.63; 10.87.

HRMS (ESI): found m/z $[M + H]^+$ 183.41.

2-Amino-6-ethylnicotinonitrile (106)

^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.60-7.62 (d, 1H); 6.58-6.60 (d, 1H); 5.30 (br s, 2H); 2.65-2.71 (q, 2H); 1.25-1.28 (t, 3H).

^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 172.13; 167.48; 147.58; 119.85; 104.21; 86.92; 30.61; 12.19.

HRMS (ESI): found m/z $[M + H]^+$ 148.32.

2-Amino-6-ethylnicotinonitrile 1-Oxide (107).

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.32-7.34 (d, 1H); 6.67-6.70 (d, 1H); 6.45 (br s, 2H); 2.98-3.06 (q, 2H); 1.32-1.37 (t, 3H).

^{13}C NMR (75.5 MHz, $CDCl_3$) δ: 157.10; 152.12; 129.13; 114. 95; 111.61; 90.43; 24.52; 10.33.

HRMS (ESI): found m/z $[M + H]^+$ 164.44.

6-Ethyl-3-(1 H-tetrazol-5-yl)pyridin-2-amine 1-oxide (89)

^{1}H NMR (300 MHz, DMSO) δ: 11.72 (br s, 1H); 7.88 (br s, 2H); 7.81 (d, J= 8.2 Hz, 1 H); 6.87 (d, J=8.2 Hz, 1 H); 2.91–2.85 (q, 2H); 1.27– 1.23 (t, 3H).

^{13}C NMR (75.5 MHz, DMSO) δ: 160.2; 157.3; 151.7; 127.0; 110.5; 106.7; 24.7; 10.8.

HRMS (ESI): m/z found $[M+H]^+$ 206.45.

Methyl 6-Chloro-2-aminonicotinate (110)

^1H NMR (400 MHz, DMSO) δ: 8.04-8.06 (d, 1H); 7.54 (br s, 2H); 6.65-6.67 (d, 1H); 3.82 (s, 3H).

^{13}C NMR (100.6 MHz, DMSO) δ: 166.73; 159.75; 154.06; 143.33; 111.61; 104.16; 52.51.

HRMS (ESI): found m/z [M + H]+ 187.33.

Methyl 6-chloronicotinate (111)

^1H NMR (400 MHz, DMSO) δ: 8.04-8.06 (d, 1H); 6.65−6.67 (d, 2 H); 3.82 (s, 3H).

^{13}C NMR (100.6 MHz, DMSO) δ: 166.73; 159.75; 154.06; 143.33; 111.61; 104.16; 52.51.

HRMS (ESI): found m/z $[M + H]^+$ 171.36.

Methyl 6-Chloro-2-aminonicotinate 1-Oxide (93)

^1H NMR (400 MHz, DMSO) δ: 7.95 (br s, 2H); 7.69-7.71 (d, 1H); 7.03-7.05 (d, 1H); 3.88 (s, 3H).

^{13}C NMR (100.6 MHz, DMSO) δ: 166.00; 153.10; 142.93; 128.18; 111.92; 105.71; 53.07.

HRMS (ESI): found m/z $[M + H]^+$ 203.22.

Methyl 6-chloronicotinate 1-Oxide (96)

^1H NMR (300 MHz, DMSO) δ: 8.68 (s, 1H); 7.91-7.93 (d, 1H); 7.71-7.73 (d, 1H); 3.88 (s, 3H).

^{13}C NMR (75.5 MHz, $CDCl_3$) δ: 163.29; 144.90; 140.78; 128.01; 127.92; 125.87; 53.49.

HRMS (ESI): found m/z $[M + H]^+$ 188.31.

6-Chloro-2-aminonicotinic Acid 1-Oxide (94)

^{1}H NMR (300 MHz, DMSO) δ: 7.95 (br s, 2H); 7.67-7.70 (d, 1H); 7.00-7.03 (d, 1H).

^{13}C NMR (75.5 MHz, DMSO) δ: 168.43; 158.98; 153.59; 139.27; 110.28; 102.26.

HPLC/MS: found m/z [M + H]+ 189.57.

6-Chloro-Nicotinic Acid 1-Oxide (95)

^{1}H NMR (400 MHz, DMSO) δ: 14.05 (br s, 1H); 8.68 (s, 1H); 7.90-7.93 (d, 1H); 7.72-7.74 (d, 1H).

^{13}C NMR (100.6 MHz, DMSO) δ: 164.34; 144.33; 140.95; 129.00; 127.80; 126.26.

HRMS (ESI): found m/z $[M + H]^{+}$ 174.08.

2-Chloropyridine 1-oxide (88)

^{1}H NMR (300 MHz, DMSO) δ: 8.44 (d, 1 H); 7.72–7.81 (t, 1H); 7.57–7.53 (t, 1 H); 7.40 (d, 1H).

^{13}C NMR (75.5 MHz, DMSO) δ: 125.31; 111.32; 105.83.

HRMS (ESI): m/z found $[M+H]^{+}$ 114.15.

Pyridin-2-amine 1-oxide (99)

^{1}H NMR (300 MHz, DMSO) δ: 8.38 (d, 1 H); 7.97 (br s, 1 H); 7.62–7.67 (t, 1H); 7.51–7.48 (t, 1H); 7.35 (d, 1H).

^{13}C NMR (75.5 MHz, DMSO) δ: 124.34; 110.35; 104.26.

HRMS (ESI): m/z found $[M+H]^{+}$ 95.11.

(1R,2R)-N-(2-Hydroxy-1-phenylethyl)-2

phenylcyclopropanecarboxamide ((+)-114)

1H NMR (300 MHz, DMSO) δ: 8.57 (d, 1H); 7.12–7.33 (m, 10 H); 4.87–4.91 (m, 1H); 3.52–3.56 (t, 2 H); 2.21–22.26 (m, 1 H); 2.02–2.06 (m, 1H); 1.28–1.32 (m, 1H); 1.18–1.20 (m, 1H).

HRMS: m/z found $[M+H]^+$ 282.35.

(1S,2S)-N-(2-Hydroxy-1-phenylethyl)-2-phenylcyclopropanecarboxamide ((-)-114)

^{1}H NMR (300 MHz, DMSO) δ: 8.57 (d, 1H); 7.12–7.33 (m, 10 H; 4.87–4.91 (m, 1H); 3.52–3.56 (t, 2 H); 2.21–22.2 6(m, 1 H); 2.02–2.06 (m, 1H); 1.28–1.32 (m, 1H); 1.18–1.20 (m, 1H).

HRMS: m/z found $[M+H]^+$ 282.74.

(1R,2R)-2-Phenylcyclopropanecarboxylic acid ((+)97)

^{1}H NMR (300 MHz, DMSO) δ: 11.06 (br s, 1H); 7.22–7.35 (m, 5H); 2.60–2.67 (m, 1H); 1.91–1.96 (m, 1H); 1.68–1.72 (m, 1H); 1.40–1.47 (m, 1H).

^{13}C NMR (100.6 MHz, DMSO) δ: 129.31; 128.22; 126.93; 33.34; 31.25; 21.56.

HRMS: m/z found $[M-H]^-$ 161.19;

[α]D=+47.62.

(1S,2S)-2-Phenylcyclopropanecarboxylic acid ((-)98)

^{1}H NMR (300 MHz, DMSO) δ: 11.06 (br s, 1H); 7.22–7.35 (m, 5H); 2.60–2.67 (m, 1 H); 1.91–1.96 (m, 1H); 1.68–1.72 (m, 1H); 1.40–1.47 (m, 1H).

^{13}C NMR (100.6 MHz, CDCl3) δ: 129.32; 128.26; 126.91; 33.35; 31.21; 21.98.

HRMS: m/z found $[M-H]^-$ 161.19;

[α]D=-47.62.

Methyl 2-aminonicotinate (132)

^{1}H NMR (300 MHz, DMSO) δ: 7.75 (br s, 1 H); 7.55–7.60 (t, 1H); 7.45–7.48 (t, 1H); 7.30 (d, 1H); 3.88 (s, 3 H).

HRMS: m/z found $[M+H]^+$ 152.15.

Methyl 2-aminonicotinate-1-Oxide (115)

^{1}H NMR (300 MHz, DMSO) δ: 8.38 (d, 1 H); 7.97 (br s, 2 H); 7.62–7.67 (t, 1H); 7.48–7.51 (t, 1H); 7.35 (d, 1H); 3.88 (s, 3 H).

^{13}C NMR (75.5 MHz, DMSO) δ: 124.32; 110.35; 104.23; 53.07.

HRMS: m/z found $[M+H]^+$ 168.15.

2-aminonicotinic acid-1-Oxide (116)

^{1}H NMR (300 MHz, DMSO) δ: 11.35 (br s, 1H); 8.38 (d, 1H); 7.97 (br s, 2H); 7.62–7.67 (t, 1H); 7.48–7.51 (t, 1H); 7.35 (d, 1H).

^{13}C NMR (75.5 MHz, DMSO) δ: 124.36; 110.31; 104.27.

HRMS: m/z found $[M+H]^+$ 154.12.

2-aminonicotinamide-1-Oxide (117)

^{1}H NMR (400 MHz, DMSO) δ: 8.54 (d, 1H); 8.22 (br s, 1H); 7.85 (br s, 3H); 7.75 (br s, 1H).

^{13}C NMR (100.6 MHz, DMSO) δ: 167.61; 150.98; 137.91; 130.12; 126.41; 115.22; 111.02.

HRMS (ESI): found m/z $[M + H]^+$ 138.06.

2-amino-4,6-diisopropylnicotinonitrile (137a)

^{1}H NMR (400 MHz, DMSO) δ: 7.49 (br s, 2H); 6.54 (s, 1H); 3.53-3.62 (m, 1H); 2.94-3.03 (m, 1H); 1.16-1.22 (m, 12H).

HRMS (ESI): found m/z $[M + H]^+$ 204.28.

2-amino-4,6-diisopropylnicotinonitrile-1-Oxide (138a)

^{1}H NMR (400 MHz, DMSO) δ: 8.12 (br s, 2H); 7.99 (s, 1H); 3.53-3.62 (m, 1H); 2.94-3.03 (m, 1H); 1.16-1.22 (m, 12H).

HRMS (ESI): found m/z $[M + H]^+$ 220.28.

2-amino-4,6-diisopropylnicotinamide (118)

1H NMR (400 MHz, DMSO) δ: 8.35 (br s, 1H); 8.22 (br s, 1H); 8.12 (br s, 2H); 7.99 (s, 1H); 3.53-3.62 (m, 1H); 2.94-3.03 (m, 1H); 1.16-1.22 (m, 12H).

^{13}C NMR (100.6 MHz, DMSO) δ: 168.21; 163.56; 160.44; 109.51; 28.82; 22.91.

HRMS (ESI): found m/z $[M + H]^+$ 238.30.

2-amino-6-ethyl-4-isopropylnicotinonitrile (137b)

^{1}H NMR (400 MHz, DMSO) δ: 7.69 (br s, 2H); 6.50 (s, 1H); 2.73-2.80 (m, 2H); 2.50-2.51 (m, 1H); 1.68-1.74 (m, 3H); 1.14-1.21 (m, 6H).

HRMS (ESI): found m/z $[M + H]^+$ 190.13.

2-amino-6-ethyl-4-isopropylnicotinonitrile-1-Oxide (138b)

^{1}H NMR (400 MHz, DMSO) δ: 7.69 (br s, 2H); 7.63 (s, 1H); 2.73-2.80 (m, 2H); 2.50-2.51 (m, 1H); 1.68-1.74 (m, 3H); 1.14-1.21 (m,6H).

HRMS (ESI): found m/z $[M + H]^+$ 206.12.

2-amino-6-ethyl-4-isopropylnicotinamide-1-Oxide (119)

^{1}H NMR (400 MHz, DMSO) δ: 7.94 (br s, 1H); 7.84 (br s, 1H); 7.69 (br s, 1H); 7.60 (s, 1H); 2.73-2.80 (m, 2H); 2.50-2.51 (m, 1H); 1.68-1.74 (m, 3H); 1.14-1.21 (m, 6H).

^{13}C NMR (100.6 MHz, DMSO) δ: 168.33; 162.15; 157.96; 111.35; 32.14; 28.73; 22.32; 13.51.

HRMS (ESI): found m/z $[M + H]^+$ 224.11.

2-amino-4-isopropyl-5,6-dimethylnicotinonitrile (137c)

^{1}H NMR (400 MHz, DMSO) δ: 6.23 (br s, 2H); 2.49-2.59 (m, 1H); 2.40 (s, 3H); 2.14 (s, 3H); 1.07 (m, 6H).

HRMS (ESI): found m/z $[M + H]^+$ 190.26.

2-amino-4-isopropyl-5,6-dimethylnicotinonitrile-1-Oxide (138c)

^{1}H NMR (400 MHz, DMSO) δ: 7.02 (br s, 2H); 2.51-2.61 (m, 1H); 2.45 (s, 3H); 2.23 (s, 3H); 1.12 (m, 6H).

HRMS (ESI): found m/z $[M + H]^+$ 206.33.

2-amino-4-isopropyl-5,6-dimethylnicotinamide-1-Oxide (120)

^{1}H NMR (400 MHz, DMSO) δ: 7.88 (br s, 1H); 7.64 (br s, 1H); 7.02 (br s, 2H); 2.51-2.61 (m, 1H); 2.45 (s, 3H); 2.23 (s, 3H); 1.12 (m, 6H).

^{13}C NMR (100.6 MHz, DMSO) δ: 167.41; 161.32; 159.43; 155.71; 121.35; 28.36; 23.25; 19.61; 16.40.

HRMS (ESI): found m/z $[M + H]^+$ 224.63.

2-amino-4-ethyl-6-isopropylnicotinonitrile (137d)

^{1}H NMR (400 MHz, DMSO) δ: 7.01 (s, 1H); 6.50 (br s, 2H); 3.60 (m, 2H); 2.55-2.63 (m, 1H); 1.16 (m, 6H); 1.11 (m, 3H).

HRMS (ESI): found m/z $[M + H]^+$ 190.63.

2-amino-4-ethyl-6-isopropylnicotinonitrile-1-Oxide (138d)

^{1}H NMR (400 MHz, DMSO) δ: 7.54 (br s, 2H); 7.01 (s, 1H); 3.54 (m, 2H); 2.50-2.56 (m, 1H); 1.12 (m, 6H); 1.09 (m, 3H).

HRMS (ESI): found m/z $[M + H]^+$ 206.28.

2-amino-4-ethyl-6-isopropylnicotinamide-1-Oxide (121)

^{1}H NMR (DMSO-d6, 400 MHz) δ: 7.90 (br s, 1H); 7.76 (br s, 1H); 7.60 (br s, 2H); 7.21 (s, 1H); 3.52 (m, 2H); 2.50-2.56 (m, 1H); 1.15 (m, 6H); 1.12 (m, 3H).

^{13}C NMR (100.6 MHz, DMSO) δ: 169.33; 163.21; 161.22; 115.23; 111.74; 37.85; 23.46; 18.57; 13.28.

HRMS (ESI): found m/z $[M + H]^+$ 224.27.

2-amino-4,6-diethylnicotinonitrile (137e)

^{1}H NMR (400 MHz, DMSO) δ: 6.54 (s, 1H); 6.50 (br s, 2H); 2.73-2.80 (m, 4H); 1.07-1.24 (m, 6H).

HRMS (ESI): found m/z $[M + H]^+$ 176.11.

2-amino-4,6-diethylnicotinonitrile-1-Oxide (138e)

^{1}H NMR (400 MHz, DMSO) δ: 7.43 (br s, 2H); 7.09 (s, 1H); 2.73-2.80 (m, 4H); 1.07-1.24 (m, 6H).

HRMS (ESI): found m/z $[M + H]^+$ 192.54.

2-amino-4,6-diethylnicotinamide (122)

^{1}H NMR (400 MHz, DMSO) δ: 7.78 (br s, 1H); 7.59 (br s, 1H); 7.43 (br s, 2H); 7.09 (s, 1H); 2.73-2.80 (m, 4H); 1.07-1.24 (m, 6H).

^{13}C NMR (100.6 MHz, DMSO) δ: 168.21; 162.32; 159.33; 156.44; 111.75; 31.36; 27.47; 14.78.
HRMS (ESI): found m/z $[M + H]^+$ 210.54.

2-amino-4-ethyl-5,6-dimethylnicotinonitrile (137f)

^{1}H NMR (400 MHz, DMSO) δ: 6.50 (br s, 2H); 2.86 (s, 3H); 2.73 (s, 3H); 2.21 -2.29 (m, 2H); 1.07-1.24 (m, 3H).
HRMS (ESI): found m/z $[M + H]^+$ 176.11.

2-amino-4-ethyl-5,6-dimethylnicotinonitrile-1-Oxide (138f)

^{1}H NMR (400 MHz, DMSO) δ: 7.67 (br s, 2H); 2.89 (s, 3H); 2.79 (s, 3H); 2.24 -2.2 31(m, 2H); 1.13-1.25 (m, 3H).
HRMS (ESI): found m/z $[M + H]^+$ 193.33.

2-amino-4-ethyl-5,6-dimethylnicotinamide-1-Oxide (123)

^{1}H NMR (400 MHz, DMSO) δ: 7.92 (br s, 1H); 7.89 (br s, 1H); 7.67 (br s, 2H); 2.89 (s, 3H); 2.79 (s, 3H); 2.24 -2.2 31(m, 2H); 1.13-1.25 (m, 3H).
^{13}C NMR (100.6 MHz, DMSO) δ: 169.31; 157.62; 155.43; 110.24; 22.75; 17.36; 14.57.
HRMS (ESI): found m/z $[M + H]^+$ 210.24.

2-amino-6-isopropyl-4-phenylnicotinonitrile (137g)

^{1}H NMR (400 MHz, DMSO) δ: 7.36-7.41 (m, 5H); 7.21 (s, 1H); 6.40 (br s, 2H); 2.45 (s, 3H); 1.82 (s, 3H); 1.54-1.63 (m, 1H).
HRMS (ESI): found m/z $[M + H]^+$ 239.34.

2-amino-6-isopropyl-4-phenylnicotinonitrile-1-Oxide (138g)

^{1}H NMR (400 MHz, DMSO) δ: 7.54 (br s, 2H); 7.36-7.41 (m, 5H); 7.21 (s, 1H); 62.45 (s, 3H); 1.82 (s, 3H); 1.54-1.63 (m, 1H).
HRMS (ESI): found m/z $[M + H]^+$ 254.28.

2-amino-6-isopropyl-4-phenylnicotinamide (124)

^{1}H NMR (400 MHz, DMSO) δ: 7.93 (br s, 1H); 7.86 (br s, 1H); 7.54 (br s, 2H); 7.36-7.41 (m, 5H); 7.21 (s, 1H); 62.45 (s, 3H); 1.82 (s, 3H); 1.54-1.63 (m, 1H).
^{13}C NMR (100.6 MHz, DMSO) δ: 169.71; 161.12; 153.63; 141.14; 128.45; 122.39; 97.34; 33.72; 22.21.
HRMS (ESI): found m/z $[M + H]^+$ 272.32.

2-amino-6-ethyl-4-phenylnicotinonitrile (137h)

^{1}H NMR (400 MHz, DMSO) δ: 7.38-7.44 (m, 5H); 7.34 (s, 1H); 6.55 (br s, 2H); 2.17-2.25 (m, 2H); 1.03-1.12 (m, 3H).
HRMS (ESI): found m/z $[M + H]^+$ 224.27.

2-amino-6-ethyl-4-phenylnicotinonitrile-1-Oxide (138h)

^{1}H NMR (400 MHz, DMSO) δ: 7.66 (br s, 2H); 7.38-7.44 (m, 5H); 7.34 (s, 1H); 2.17-2.25 (m, 2H); 1.03-1.12 (m, 3H).
HRMS (ESI): found m/z $[M + H]^+$ 240.11.

2-amino-6-ethyl-4-phenylnicotinamide-1-Oxide (125)

^{1}H NMR (400 MHz, DMSO) δ: 7.88 (br s, 1H); 7.65 (br s, 1H); 7.66 (br s, 2H); 7.38-7.44 (m, 5H); 7.34 (s, 1H); 2.17-2.25 (m, 2H); 1.03-1.12 (m, 3H).
^{13}C NMR (100.6 MHz, DMSO) δ: 170.15; 162.32; 159.44; 152.36; 141.39; 129.73; 123.11; 110.25: 98.01; 32.34; 17.47.

HRMS (ESI): found m/z $[M + H]^+$ 258.29.

2-amino-5,6-dimethyl-4-phenylnicotinonitrile (137i)

^{1}H NMR (400 MHz, DMSO) δ: 7.11-7.19 (m, 5H); 6.50 (br s, 2H); 2.21 (s, 3H); 1.10 (s, 3H).

HRMS (ESI): found m/z $[M + H]^+$ 224.23.

2-amino-5,6-dimethyl-4-phenylnicotinonitrile-1-Oxide (138i)

^{1}H NMR (400 MHz, DMSO) δ: 7.61 (br s, 2H); 7.11-7.19 (m, 5H); 2.21 (s, 3H); 1.10 (s, 3H).

HRMS (ESI): found m/z $[M + H]^+$ 240.11.

2-amino-5,6-dimethyl-4-phenylnicotinamide-1-Oxide (126)

^{1}H NMR (400 MHz, DMSO) δ: 7.82 (br s, 1H); 7.74 (br s, 1H); 7.61 (br s, 2H); 7.11-7.19 (m, 5H); 2.21 (s, 3H); 1.10 (s, 3H).

^{13}C NMR (100.6 MHz, DMSO) δ: 168.03; 158.71; 155.93; 151.12; 137.77; 129.41; 127.66; 118.32; 109.16; 19.72; 12.34.

HRMS (ESI): found m/z $[M + H]^+$ 258.29.

Methyl 3-aminopyrazine-2-carboxylate (141)

^{1}H NMR (400 MHz, DMSO) δ: 7.13 (s, 1H); 6.98 (s, 1H); 6.34 (br s, 2H); 3.21 (s, 3H).

HRMS (ESI): found m/z $[M + H]^+$ 154.05.

Methyl 3-aminopyrazine-2-carboxylate-4-Oxide (127)

^{1}H NMR (400 MHz, DMSO) δ: 7.64 (br s, 2H); 6.98 (s, 1H); 6.63 (s, 1H); 3.17 (s, 3H).

^{13}C NMR (100.6 MHz, DMSO) δ: 165.74; 155.31; 148.23; 130.17;

125.72; 61.34.

HRMS (ESI): found m/z $[M + H]^+$ 170.14.

3-aminopyrazine-2-carboxamide-4-Oxide (129)

^{1}H NMR (400 MHz, DMSO) δ: 7.94 (br s, 1H); 7.87 (br s, 1H); 7.78 (br s, 2H); 6.56 (s, 1H); 6.24 (s, 1H).

^{13}C NMR (100.6 MHz, DMSO) δ: 166.15; 154.73; 146.46; 132.34; 123.70.

HRMS (ESI): found m/z $[M + H]^+$ 155.11.

Methyl 5-methylpyrazine-2-carboxylate (142)

^{1}H NMR (400 MHz, DMSO) δ: 7.02 (s, 1H); 6.54 (s, 1H); 3.45 (s, 3H); 1.22 (s, 3H).

HRMS (ESI): found m/z $[M + H]^+$ 152.06.

Methyl 5-methylpyrazine-2-carboxylate-4-Oxide (128)

^{1}H NMR (400 MHz, DMSO) δ: 7.45 (s, 1H); 6.60 (s, 1H); 3.22 (s, 3H); 1.19 (s, 3H).

^{13}C NMR (100.6 MHz, DMSO) δ: 166.62; 154.24; 146.71;, 132.55; 127.74; 15.78.

HRMS (ESI): found m/z $[M + H]^+$ 169.15.

5-methylpyrazine-2-carboxamide (130)

^{1}H NMR (400 MHz, DMSO) δ: 7.90 (br s, 1H); 7.81 (br s, 1H); 7.42 (s, 1H); 6.68 (s, 1H); 3.19 (s, 3H).

^{13}C NMR (100.6 MHz, DMSO) δ: 168.35; 155.12; 147.62; 130.86; 127.79.

HRMS (ESI): found m/z $[M + H]^+$ 154.17.

3. SYNTHETIC STRATEGIES TO FUNCTIONALIZE THE NATURAL PRODUCT BREFELDIN-A

3.1 The Use of Natural Products for Medical Use

Nature is the greatest source of chemical diversity, providing humankind with a vast arsenal of natural products with diverse chemical structures and biological activities. Indeed, the majority, if not all, of the living organisms can produce specialized metabolites endowed with some biological activity. Therefore, it is not surprising that in the last three decades half of the approved small-molecule drugs derived from natural products. [88,89]

The majority of these molecules are in clinical use as (*i*) antibiotics, (*ii*) cancer therapeutics, (*iii*) immunosuppressive agents, whereas, to a lesser extent, they are also commercialized as (*iv*) antivirals, (*v*) anthelmintics, (*vi*) enzyme inhibitors, (*vii*) nutraceuticals (or 'health foods'), (*viii*) polymers, (*ix*) surfactants, (*x*) bioherbicides and (*xi*) vaccines.[90]

Before the medicinal chemistry revolution in the 20th century, crude and semi-pure extracts from plants, animals, microbes and minerals represented the only medications available to treat human and domestic animal illnesses. The receptor theory and the rational explanation of the molecular mechanism of action represented a milestone in the use of drugs.[91] The idea that the effects brought by a drug are mediated by specific interactions of one molecule with one biological macromolecule (such as proteins or nucleic acids in most cases) led scientists to the conclusion that individual chemical compounds in extracts, rather than some mystical "power of life", are the required factors for the biological activity. On turn, this led to the beginning of a totally new era in

pharmacology, as pure, isolated chemicals, instead of extracts, became the standard treatments for diseases.[92]

The chemical structures of many bioactive compounds, responsible for the effects of crude extracts, were elucidated. At the same time, along with the desired biological effects, for most of these natural products undesired effects were identified. A well-known example is salicylic acid: the first FANS reported in human history, but endowed with severe side effects on the gastro-enteric apparatus.[93] Since the raise of side effects cannot be unbound from the biological action, from the moment that both depend on the molecular structure, it appears quite obvious that the rational manipulation of the chemical structure may lead to ameliorate the activity/toxicity balance. In this scenario, medicinal chemistry plays a crucial role in making a given molecule manageable for therapeutic purpose. Unfortunately, the structural modification of natural products is usually not an easy task, since the chemical complexity of the molecules (i.e. taxol).[94] Indeed, the negative cost/benefit balance in producing/using natural products as medication has hampered, for long time, the advances in this field. Nevertheless, technological breakthroughs experienced in different scientific fields, together with enormous difficulties in delivering new drugs in many key therapeutic areas, have renewed the interest in natural product research.

Natural products research continues to furnish a variety of lead structures, which may be used as templates for the development of new drugs by the pharmaceutical industry. There is no doubt that natural products have been, are, and will be, important sources of new pharmaceutical compounds.[91]

3.2 Chemical Diversity of Natural Products

Natural products usually possess complex molecular structures, with cyclic semi-rigid scaffolds, several chiral centers, and they usually do not comply with the Lipinski's rule of five.

As an example, in Figure 27A, anti-malarial artemisinin (**143**) is made of a fused trioxene system with peroxy, lactone, cyclic acetal and ketal moieties. The scaffold of artemisinin maintains both the oxidative potential as well as the chemical stability. The well-known anti-cancer drug paclitaxel (**144**) possesses a 6,8,6,4-tetracycle-fused skeleton linking functional groups at peculiar positions that ensures binding to tubulin. Asperlicin, a non-peptide cholecystokinin (CCK) antagonist (**145**) originated from extracts of *Aspergillus alliaceus*, is a chemically complex molecule with a quinazolinone-fused benzodiazepine core.

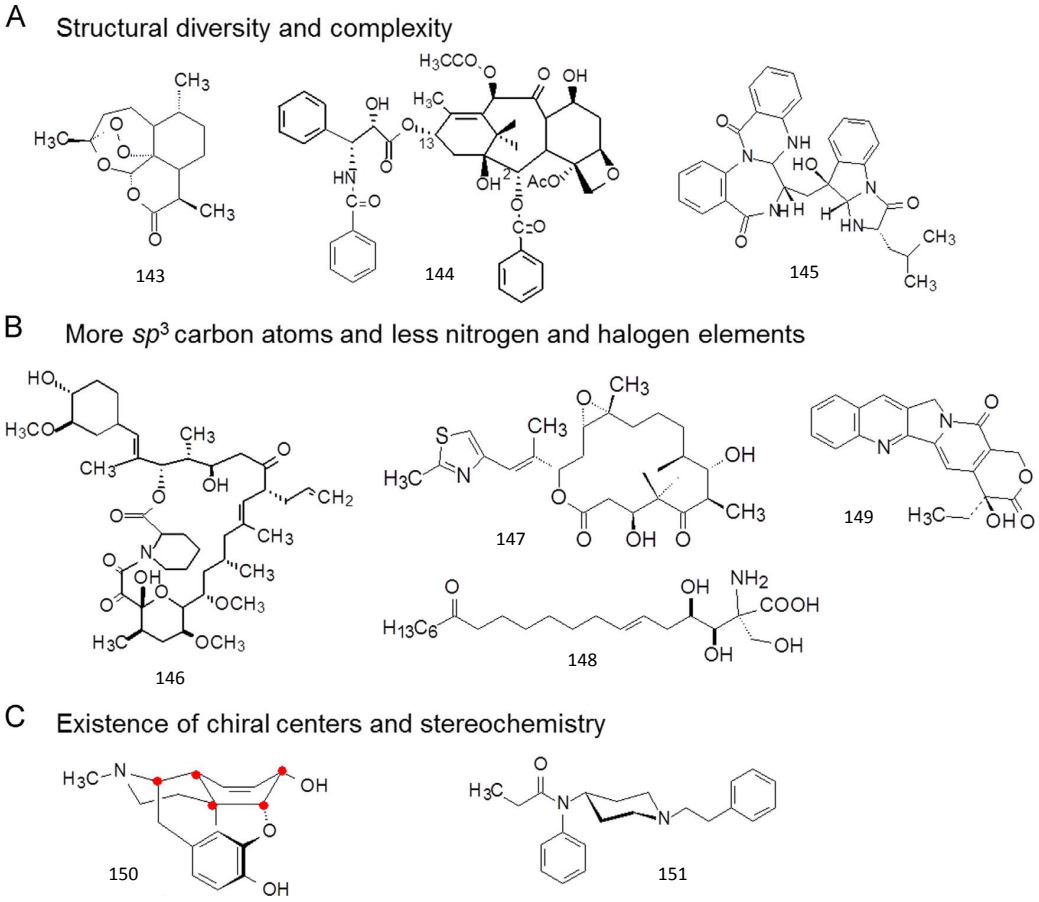


Figure 27. The general features of natural products.

Even if single structures and the presence of numerous functional groups constitute the prerequisite for the strong binding of natural products to the targets, they are also responsible for relevant drawbacks: (*i*) the presence of several functional groups increases the probability of interaction with off-targets, leading to side effects; (*ii*) chemical synthesis difficulties increase, strikingly, with the size and complexity of the molecules. Therefore, it is generally accepted that natural products possess “redundant atoms”, which are not directly involved in the binding to target protein(s), but rather are responsible for the conferring of unpleasant physico-chemical, pharmacokinetic and

biopharmaceutical properties. Therefore, aimed structural optimization and SAR campaigns should be pursued in order to appropriately remove, when feasible, unnecessary moieties/groups/atoms and, in turn, improve the ligand efficiency (LE) of the molecules.[95] LE is a measure of the binding affinity (or also binding free energy) per unit of mass for a compound relative to its molecular target. High values of LE imply fewer unnecessary atoms relative to target binding. This parameter is routinely used as one of the driving factors in the early stages of lead optimization.

3.2.1 More sp^3 Carbon Atoms And Less Nitrogen And Halogen Elements.

Most natural products are composed of more sp^3 hybrid carbons than sp^2 carbons, as a consequence the tetrahedron carbons connected to each other forms flexible chain or cyclic structures. As examples in Figure 27B, both the immunological regulator tacrolimus (**146**) and the anti-cancer compound epothilone B (**147**) are macrolides containing many saturated carbon atoms. Another immune regulator, ISP-1 (**148**), isolated from *Cordyceps sinensis Sacc*, is a linear and flexible molecule. An exception is the anti-cancer drug camptothecin (**149**), a fused and conjugated aromatic system with multiple sp^2 carbons.

Most natural products are characterized by a reduced presence of nitrogen atoms, and a possible explanation of this aspect can be identified in the reduced ability of microorganisms and plants (exception are *Leguminoceae* plants) to fix nitrogen. These features provide various options for introducing nitrogen atoms during modification. Nitrogen is a versatile atom in many aspects, having a stronger nucleophilicity than carbon and oxygen, providing 3 or 5 valence bond,

forming salts for alkaline amines or being a neutral amide, building heterocycles, aromatic and fused compounds, being able to be a terminal group or to serve as a linker to connect other groups. Therefore, introduction or modification of various nitrogen atoms is an important molecular operation in lead compound optimization. Natural products rarely contain halogen atoms, although marine products sometimes incorporate a bromine atom.

3.2.2 Chiral Centres and Stereochemistry

Natural products are synthesized in organisms by enzymatic catalysis, which take place with stereochemical specificity, leading to stereospecific compounds with chiral centers, chiral axes, or cis/trans configurations. These asymmetrical factors usually convey a critical contribution to pharmacological actions (i.e. selectivity). However, those chiral centers not participating in the binding to target protein can be considered useless. In manipulating or simplifying natural products the unnecessary chirality and stereochemistry have to be removed as much as possible, since the presence of several chiral centers can lead to technological issues (i.e. purification) and practical issues (i.e. identification of the active enantiomer is usually not an easy task). Therefore, natural products derivatives lacking or with a reduced number of chiral centers have the advantages to be more chemically accessible and cheaper with respect to the starting natural products. For example, in Figure 27C, the old analgesic morphine (1**50**) contains 21 heavy atoms, 5 fused rings, and 5 chiral carbon atoms. The synthetic analgesic drug fentanyl (**151**), derived from morphine, is a simplified non-chiral molecule.

3.3 Identification of Biologically Active Material

Two main approaches are currently pursued to identify new bioactive molecules from natural sources; (*i*) random collection and screening of material, or (*ii*) knowledge-based approaches, where information coming from *traditional medicine* are used for selection purpose.[89] The first approach is based on the notion that just a small part of globes' biodiversity has been explored for biological activities as yet. In this case, researchers take advantages from peculiar organisms living in a species-rich environment, where they need to evolve defence and competition mechanisms to survive. Thus, collection of plant, animal and microbial samples from rich ecosystems may give rise to isolation of novel biological activities. A well-known example, where such approach led to successful results is represented by the screening activities performed by the National Cancer Institute in USA in 1960s.[92] In that screening campaigns cytostate paclitaxel (taxoid) was identified from Pacific yew tree *Taxus brevifolia*. Paclitaxel showed antitumor activities through a previously unknown mechanism (stabilization of microtubules) and this natural product is now approved for clinical use to treat lung, breast and ovary cancers, as well as for Kapos sarcoma.[92]

Besides random collection, the selection of the starting material may be done by collecting knowledge on the use of plants and other natural products as herbal medicines and thereby identifying potential biological activities. Indeed, a major part of earth's population still relies on nature-derived drugs as their only medication (i.e. traditional Chinese medicine).[96] Traditional medicine may therefore provide invaluable information, as illustrated by the example of artemisinin, an antimalarial agent from sweet worm tree *Artemisiae annua*, used in Chinese

medicine since 200 DC and nowadays in use against multi-resistant malarial protozoa *Plasmodium falciparum*.

3.4 Natural Products as Antimicrobial Drugs: State of the Art

Without any doubts, natural products are the most fruitful source of antimicrobials nowadays and of the several classes of antibacterials, only a few are those derived from chemical synthesis. This is due to the fact that natural compounds such as macro and micro fungi have been part of human life since thousands of years. They were used as food (mushrooms), in preparation of alcoholic beverages (yeasts), medication in traditional medicine and for cultural purposes.[97,98] Undoubtedly the most famous fungus-derived molecule, so far discovered, is penicillin (**152**, Figure 28) obtained from *Penicillium notatum* by Fleming in 1929.[99]

152 was produced in high yields by a counter-current extractive separation approach, allowing its first experimentation *in vivo*. For that discovery, Chain, Florey together with Fleming were honoured with the Nobel Prize in Physiology and Medicine.[100] Between 1942–1944 the results of the first clinical trials on penicillin G (**152**) were published, leading to worldwide efforts aimed at the discovery new antibiotics from microorganisms and bioactive natural products. Such efforts conducted in 1968 to the conclusion that all natural β-lactams had been discovered,[101] but this was not the case. Indeed, in the 70s advancements in the biochemical research, among which the introduction of new screening methods, resulted in the discovery of novel structural classes of antibiotics (i.e. norcardicins, carbapenems and monobactams), leading to the isolation of new antibiotics for clinical

use: norcardicin (**153**), imipenem **154**) and aztreonam (**155**), respectively (Figure 28).[102]

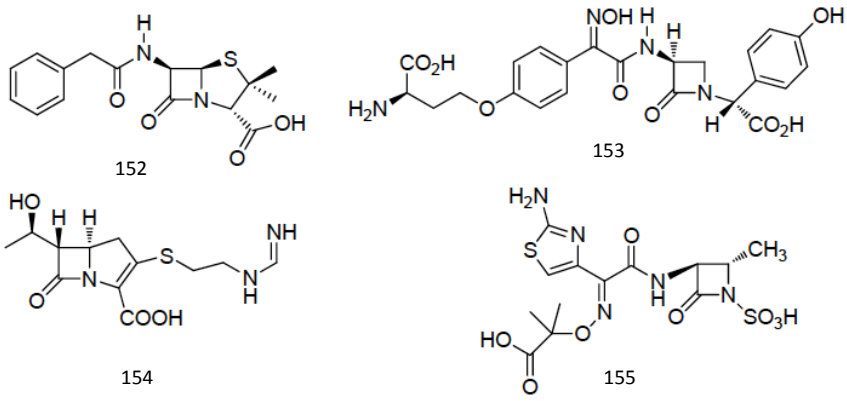
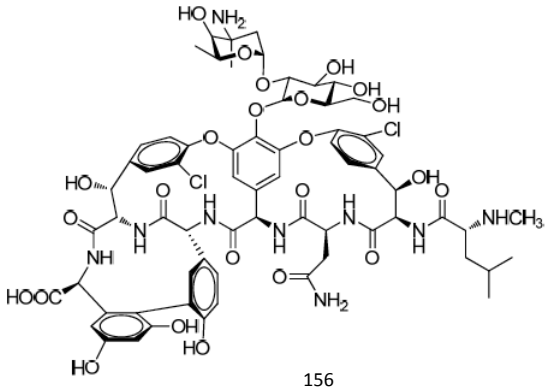


Figure 28. Penicillin (152), Norcardicin (153), Imipenem (154) and Aztreonam (155).

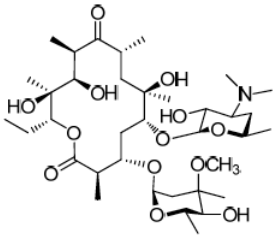
Another source of potential pharmacologically active substances is represented by macro fungi. Indeed, it has been estimated that exist about 25,000 species of basidiomycetes of which about 500 are members of the Aphyllophorales family.[103] Approximately 75% of tested polypore fungi have shown strong antimicrobial activity and may constitute a viable source for the development of novel antibiotics. Many other compounds have also displayed antiviral, cytotoxic, antineoplastic, cardiovascular, anti-inflammatory, immune-stimulating and anticancer properties.[97] Novel bioactive secondary metabolites derived from fungal sources have yielded some of the most important natural products for the pharmaceutical industry.[104]

Vancomycin (**156**) (Figure 29) was discovered in 1953 by Edmund Kornfeld, which isolated a glycopeptide antibiotic produced in cultures of Amycolatopsis orientalis actives against a wide range of gram-positive bacteria (such as Staphylococci and Streptococci) and also against

gram-negative, mycobacteria and fungi. Vancomycin was approved by the FDA in 1958. It is nowadays used for the treatment of severe infections and against susceptible organisms in patients hypersensitive to penicillin (**152**).[105]



156



157

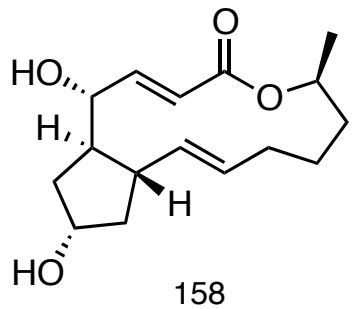
Figure 29. Vancomycin (**156**) and Erythromycin (**157**).

The macrolide erythromycin (**157**) from Saccharopolyspora erythraea is an antibacterial drug, containing a 14-membered macrocycle composed entirely of propionate units (Figure 27). Erythromycin (**157**) has broad-spectrum activities against gram-positive bacteria and is used for mild to moderate, upper and lower respiratory tract infections.[99] Currently there are three semisynthetic ketolide derivatives of erythromycin (**157**), cethromycin (ABT-773, Restanza™), EP-420 (by Enanta Pharmaceuticals) and BAL-19403 (by Basilea), which are in clinical development.

3.5 Brefeldin A

Among the large arsenal of antimicrobial natural products, we decided to focus our attention and efforts on Brefeldin A (BFA, **158**, Figure 30),

which is a secondary metabolite of Ascomycetes species. It was first isolated in 1958 by Singleton et al.[106] from *Penicillium decumbens* and the complete structure was determined by Sigg et al[107] by X-ray analysis.



158

Figure 30. Chemical structure of Brefeldin A (BFA).

Brefeldin A displays a variety of biological properties, including antifungal,[108] antiviral,[109] antinematodal,[108] and antimitotic[110] activities. Structurally, it is a 13-membered macrolide which presents two double bond in *trans*-position, an ester moiety and 5 chiral centres.

BFA is a lactone antibiotic, which exerts its effects by acting on the protein transport from the endoplasmic reticulum to the Golgi apparatus.[111] Unfortunately, BFA is endowed with severe side effects, which make it a toxin rather than a medication. It is therefore of great interest to modify the BFA structure so as to take advantage from the beneficial effects and get rid of the unwanted ones. The first enantioselective synthesis of BFA was reported in 1979 by Corey,[112] and there have been numerous partial and total synthesis described since then.[113,114]

3.6 Aim of The Project

In this contest, we embarked in a project aimed at developing a versatile, practical, and stereocontrolled route that would minimize protecting group manipulations and adapt a platform that leads to a variety of analogues of BFA. In particular, our efforts led to the disclosure of a modified enantioselective synthesis, which allows us to:

(i) control the stereochemistry of the five chiral centres;

(ii) expand the chemical diversity of the substitutions in position C-13 in just one step.

Introduction of substituents in position C13 seems to be a good medicinal chemist strategy since the introduction of functional groups in this position may reduce cytotoxic effects of BFA.[108] As previously mentioned, several partial and total syntheses of BFA are reported in literature. Particularly, *Seehafer et al.*[108] reported the structural modification of BFA at the C13 position. Even if of foremost importance, the aforementioned work has the limitation to require several reaction steps to derivatized position C13. Thus, we took advantage from some of the literature produced through the years in order to design a retrosynthetic route which allows us to identify synthons **159** and **160** as suitable starting points for the production of C-13 substituted analogues: (*i*) synthon **159** allows us to control the stereochemistry of position C-4, C-6, C-4b, C-7b (Figure 31, Scheme 14); (*ii*) alcohol **160** represents a highly versatile substrate, where different **R** alkyl groups can be introduced with the desired stereochemistry (Figure 31, Scheme 16).

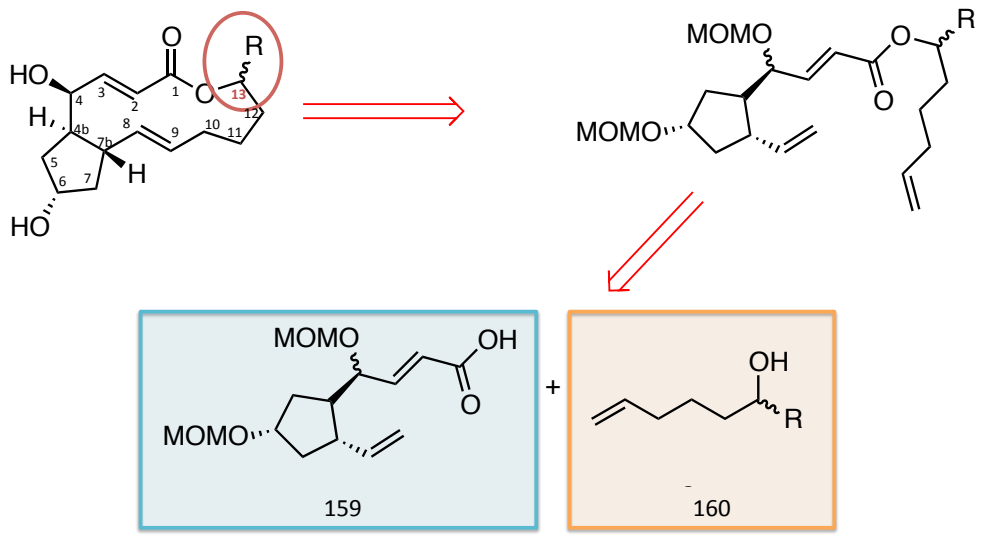
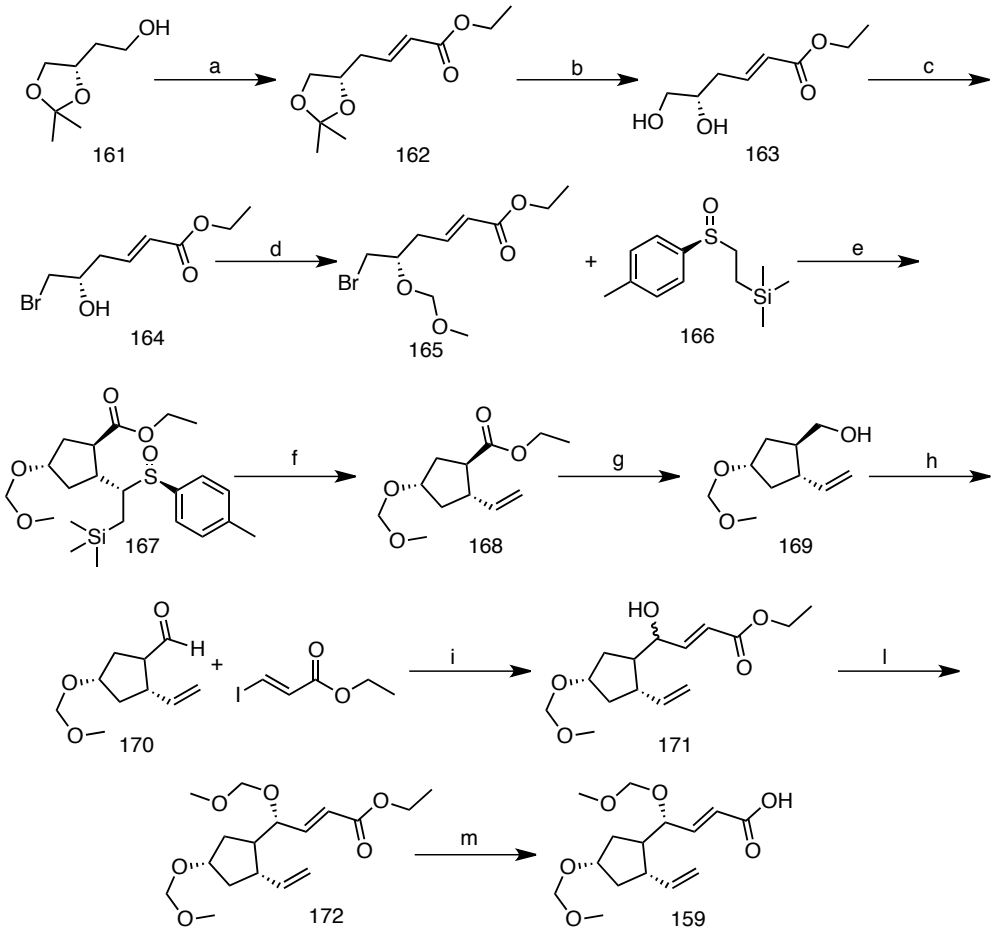


Figure 31. Retrosynthetic analysis of Brefeldin A derivatives.

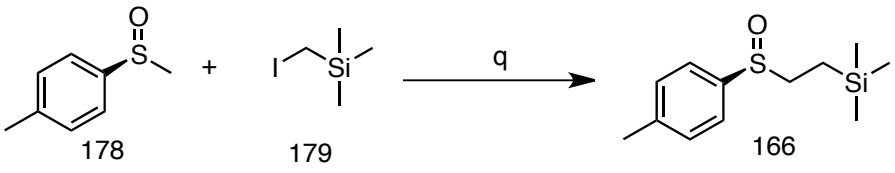
3.7 Chemistry

The α,β-unsatured ester was obtained by one-pot Swern oxidation-Wittig reaction to give **162**, which was in turn deprotected under acidic conditions to afford the unsatured ester **163**.[115] The primary alcohol was selectively brominated using triphenylphosphine and carbon tetrabromide. Reaction between compound **164** and methoxymethyl chloride using *N,N*-diisopropyethylamine as the base provided the electrophile **165**. Formation of compound **167** through the conjugate addition of the α-sulfinyl carbanion derived from a chiral tolylsulfoxide **166** to compound **165** was used. The absolute stereochemistry of the newly formed stereocenters is controlled by the chiral sulfoxide.[115] Treatment with TBAF immediately after purification of **167** was necessary due to rapid decomposition of **167**, presumably through syn-elimination of the arylsulfinic acid.[116] The aldehyde **170** was first prepared from the ester **168** by reduction/oxidation sequence. Subsequent organochromium-mediated addition of ethyl 3-iodoacrylate

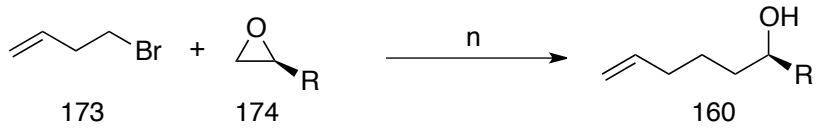
to **169** led to a 1:1 mixture of diastereomeric alcohols which was separated by silica gel to obtain the desiderate distereoisomer **171**. Protection of alcoholic function by methoxymethyl chloride and subsequent saponification led to the key acids **172**. The key intermediate **160** was obtained through a Grignard reaction between and epoxide with appropriate configuration and 4-bromobutene, with retention of configuration. The esters **175** were obtained through Steghlich esterification with the alcohol **160**. Grubb's first generation catalyst was used to prepare the macrolides **176** through a ring closing metathesis (RCM) reaction. Finally the acidic deprotection of methoxymethyl groups provided to the desired derivative of Brefeldin A **177**. It must be noticed that the availability, or, in some cases, the easy preparation of epoxides 174 allow for a wide range of substituents at position C-13, further expanding the series and the SAR exploration.



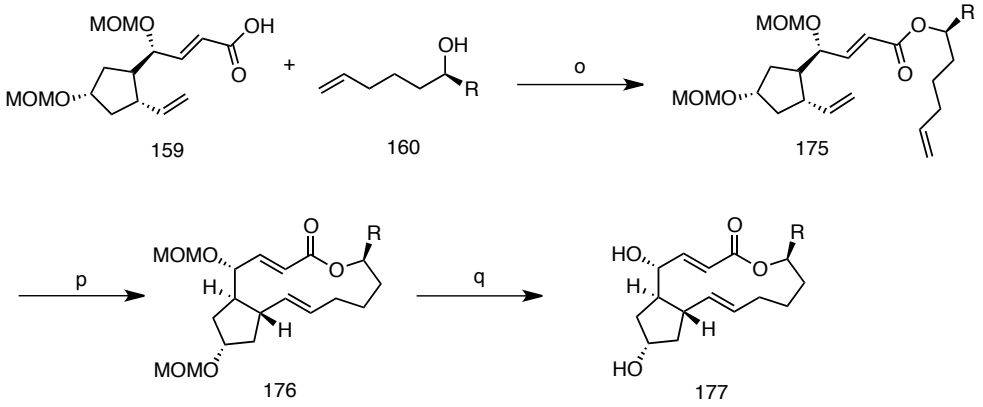
Scheme 14. Reagents and conditions. ***a.*** $(COCl)_2$, DMSO, DIPEA, $(Ph)_3PCHCOOEt$, DCM, -78°C->r.t., 89%; ***b.*** HCl 1N, THF, r.t., 4h, 88%; ***c.*** CBr_4, $(Ph)_3P$, DCM, r.t->0°C->r.t., 6h, 60%; ***d.*** MOMCl, DIPEA, DCM, r.t->0°C->r.t., 18h, 72%; ***e.*** LDA, THF, -78°C->r.t., 1h, 70%; ***f.*** TBAF, THF, r.t., 3h, 70%; ***g.*** $LiAlH_4$, THF, 0°C, 30', 88%; ***h.*** PCC, DCM, r.t., DCM, 75%; ***i.*** $CrCl_2$, $NiCl_2$, DMF, r.t., 6h, 67%; ***l.*** MOMCl, DIPEA, DCM, r.t.->0°C->r.t., 71%; ***m.*** LiOH, THF, H_2O, MeOH, r.t., 98%.



Scheme 15. Reagents and conditions. q*)* LDA, THF, -78°C->r.t., 4 h, 65%.



Scheme 16. Reagents and conditions. n*)* Mg^0, I_2, CuI, THF, r.t., 4 h, 62%. R= alkyl group.



Scheme 17. Reagents and conditions *o.* DCC, DMAP, DCM, r.t.->0°C->r.t., 15%; *p.* Grubb's 1st catalyst, DCM, reflux, 16h, 80%; *q.* HCl 37%, MeOH, r.t. 6h, 70%. R= alkyl group.

3.8 Conclusions and Future Perspectives

Nature has been a source of medicinal products for millennia. Under the point of view of the drug discovery pipeline and the clinical use of such molecules, structural modifications of the main scaffold are usually required, so as to (*i*) disjoint desired and undesired effects and (*ii*) optimize the chemical accessibility. Usually, natural product–based drug discovery starts with the identification of an active compound that necessitates "tailor-made" or individualized manipulation of the structure so as to reach a drug criterion. In this context, we embarked in a project aimed at the structural modification of a secondary metabolite of Fungi, namely Brefeldin A (BFA), which is endowed with interesting biological activities, including antimicrobial effects. In details, we decided to modify

the structure of BFA, in order to improve its antimicrobial effects and selectivity.
Therefore, we designed a modified synthetic route that allows us:

a) to modify position C-13, in just one step;
b) to control the stereochemistry of the five stereocentres;

The final optimized synthesis consists of 12 steps to synthetized intermediate **159** (which is common to all the derivatives reported in this work), while the intermediate **160** is obtained in just one step, where we can control the stereochemistry and the nature of R. The total number of steps is equal to 18 and even if the number is quite high, our synthetic route allows us to obtain the final compounds with the desired stereochemistry for all the 5 stereocenters (Figure 32). This represent a great improvement with respect to the synthetic route reported in the work of *Seehafer et al.*[108]. Indeed, in that work the authors have to optimize different reaction routes depending on the nature of the substituents on C13. In contrast, our synthetic strategy can be apply regardless of the nature of the substituent at the same position.

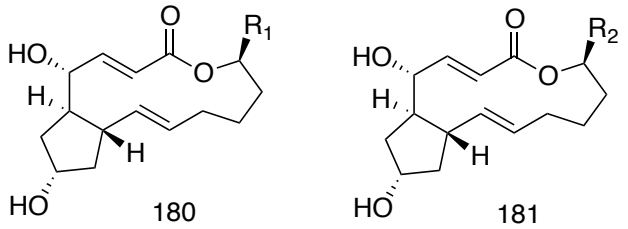


Figure 32. Derivatives compounds of Brefeldin A synthesized. R= alkyl group.

The compounds obtained will be tested as antimicrobial agents so as to perform preliminary SAR studies. Moreover, the compounds will be

evaluated for their toxicity on different human cell lines. Should the compounds give promising results, further derivatives will be synthesized, taking advantages of our new synthetic strategy, pursuing an aimed hit-to-lead optimization campaigns.

As these compounds have not yet been patented, we cannot disclose their structure to the finest detail.

3.9 Experimental Section

Synthesis of (*S*,*E*)-ethyl 4-(2,2-dimethyl-1,3-dioxolan-4-yl)but-2-enoate (162)

A cold (-78°C) stirred solution of oxalyl chloride (3.23 ml; 38 mmol) in 62 ml of dry methylene chloride, under Nitrogen flux, was treated with a solution of DMSO (4.95 ml; 67 mmol) in the same solvent (17 ml). After the resultant mixture was stirred at -78°C for 10 min, a solution of **161** (5g; 34 mmol) in 25 ml of methylene chloride was added. The resultant slurry was stirred at -78°C for 40 min and then treated with diisopropylethylamine (29.61 ml; 170 mmol). The cooling bath was removed, and the reaction mixture was stirred at room temperature for 1 h to afford a yellow solution of *O*-isopropylidene-4-oxo-(S)-butane-1,2-diol.

This solution was cooled to 0°C and treated with (carbethoxymethylene)triphenylphosphorane (29.61 g; 85 mmol) at 0°C for 1 h and at room temperature for 4 h. The resultant solution was diluited with diethyl ether (300 ml), washed with water (3 x 100 ml), 10% aqueous $NaHSO_4$ (100 ml) and brine (3 x 100 ml), and dried over Na_2SO_4. Removal of solvent gave viscous yellow oil. Diethyl ether (300 ml) and hexane (300 ml) were added, and mixture was kept at -10°C for 15 h. Filtration of the white precipitate ($Ph_3P=O$) and removal of solvent

gave a yellow oil, which was purified by flash chromatography. Elution of the column with a petroleum ether-ethyl acetate mixture (9:1) gave 4% of **Z** isomer as colourless oil; further elution of column afforded 6.504 g (89,3%) of the **E** isomer of **162** as colourless oil.

Synthesis of (*S*,*E*)-ethyl 5,6-dihydroxyhex-2-enoate (163)

To a solution of **162** (6.500 g; 30,3 mmol) in 132 ml of tetrahydrofuran was added 1N HCl (90 ml). The reaction mixture was stirred at room temperature for 4 h. NaCl (100 g) and ethyl acetate (400 ml) were added. The organic layer was separated and washed with brine (3 x 100 ml). The aqueous washings were extracted with ethyl acetate (3 x 100 ml), the extracts washed with brine (50 ml), and the organic layers combined. Drying (Na_2SO_4) and removal of solvent gave 4.600 g (88%) of **163** as viscous oil.

Synthesis of (*S*,*E*)-ethyl 6-bromo-5-hydroxyhex-2-enoate (164)

An oven-dried, round-bottomed flask, under nitrogen flux, was charged with methylene chloride (160 ml), compound **163** (4.600 g; 26 mmol, in 30 ml of methylene chloride), and carbon tetrabromide (10.612 g; 32 mmol) in the indicated order. After the solution was cooled to 0°C in the ice bath, triphenylphosphine (8.311 g; 32 mmol) was added in 30 min using syringe pump. After TLC analysis (petroleum ether-ethyl acetate 7:3) indicated that the starting material was consumed, the solution was quenched by cautious addition of satured aqueous ammonium chloride (50 ml). The mixture was extracted with diethyl ether (3 x 50 ml). The combined organic layers were washed brine (3 x 30 ml), and dried over Na_2SO_4. The resulting liquid was filtered and concentrated under reduced pressure. The product was purified by flash chromatography

(petroleum ether-ethyl acetate 8:2) to give 4.300 g (60%) of **164** as colourless oil.

Synthesis of (*S*,*E*)-ethyl 6-bromo-5-(methoxymethoxy)hex-2-enoate (165)

An oven-dried, round-bottomed flask, under nitrogen flux, was charged with methylene chloride (92 ml), compound **164** (2.120 g; 8.95 mmol, in 5 ml of methylene chloride), and MOMCl (3.41 ml; 45 mmol) in the indicated order. After the solution was cooled to 0°C in the ice-bath, diisopropylamine (7.84 ml; 45 mmol) was added in one portion. After overnight reaction at room temperature, diethyl ether (25 ml) was added to the mixture and the solution was washed with satured ammonium chloride solution (7 ml) and dried over Na_2SO_4. The resulting orange solution was filtered and concentrated under reduced pressure. The product was purified by flash chromatography on silica (petroleum ether-ethyl acetate 8:2) to give 1,785 g (72%) of **165** as colourless oil.

Synthesis of (*S*)-trimethyl(2-(*p*-tolylsulfinyl)ethyl)silane (166)

To a solution of diisopropylamine dry (1.768 ml; 13 mmol) in THF dry (7 ml) was added *n*-buthyllitium (2.5 M in hexane; 5.60 ml; 14 mmol) drop wise at 0°C, and the mixture was stirred for 10 min. The reaction mixture was then cooled to -78°C, and a solution of *p*-toluylmethyl sulfoxide (**178**) (1 g; 6.48 mmol) in THF dry (6 ml) was added. After being stirring for 1 h at -78°C, trimethylsililmethylene iodide (**179**) (1.54 ml; 10.4 mmol) was added in one portion. The temperature was rise to 25°C in 2 h, after that the mixture was quenched with ammonium chloride solution (5 ml) and extracted with ethyl acetate (3 x 7 ml). The combined organic extracts were washed with brine (3 x 5 ml), dried over Na_2SO_4 and

concentrated under reduced pressure to give a residue that was purified by coloumn chromatography (petroleum ether-ethyl acetate 8:2) to obtain **166** as yellow solid (0.930 g; 65.12%).

Synthesis of (1*R*,2*S*,4*R*)-ethyl 4-(methoxymethoxy)-2-((*S*)-1-((*R*)-*p*-tolylsulfinyl)-2-(trimethylsilyl)ethyl)cyclopentanecarboxylate (167)

An oven-dried, round-bottomed flask, under Argon flux, was charged with tetrahydrofuran (7 ml) and diisopropylamine (1.73 ml; 12 mmol) in order. After the solution was cooled to 0°C, *n*-buthyllitium (2.5 M in hexane, 5.32 ml; 13 mmol) was added. After being stirred for 10 min, the reaction mixture was cooled to -78°C and compound **166** (0.930 g; 4.22 mmol) was added. After the solution was stirred for 15 min, compound **165** was added in one portion. The solution was stirred 15 min at -78°C and then warmed to 0°C. After 30 min a TLC analysis indicated that the starting material was consumed. The solution was quenched by cautious addition of satured ammonium chloride solution (10 ml). The mixture was extracted with diethyl ether (3x 10 ml) and washed with brine (3 x 5ml). The combined organic layers were dried over Na_2SO_4. The resulting liquid was filtered and concentrated under reduced pressure. The product was purified by flash chromatography (petroleum ether-ethyl acetate 8:2) to afford 1.200 g (70%) of **167** as yellow oil. To prevent decomposition of the product, the next reaction was performed immediately after characterization data was obtained).

Synthesis of (1*R*,2*S*,4*S*)-ethyl 4-(methoxymethoxy)-2-vinylcyclopentanecarboxylate (168)

An oven-dried, round-bottomed flask, under argon flux, was charged with tetrahydrofuran (15 ml) and compound **167** (1.200 g; 2.72 mmol, in

5 ml of tetrahydrofuran). After the solution was cooled to 0°C in the ice bath, TBAF (1.0 M in tetrahydrofuran, 8.17 ml; 8.17 mmol) was added in one portion. After TLC analysis (petroleum ether-ethyl acetate 9:1) indicated that the starting was consumed, the solution was quenched by cautious addition of satured aqueous ammonium chloride (7 ml). The mixture was extracted with diethyl ether (3 x 5 ml). The combined organic layers were washed with brine (3 x 5 ml) and dried over Na_2SO_4. The resulting solution was filtered and concentrated under reduced pressure. The product was purified by flash chromatography (petroleum ether-ethyl acetate 95:5) to afford 0,400 g (70%) of **168** as colourless oil.

Synthesis of ((1*R*,2*S*,4*S*)-4-(methoxymethoxy)-2-vinylcyclopentyl)methanol (169)

An oven-dried, round-bottomed flask, under nitrogen flux, was charged with tetrahydrofuran dry (10 ml) and $LiAlH_4$ (0.256 g; 6.74 mmol). The resultant grey suspension was cooled to 0°C and then a solution of **168** (0.385 g; 1.68 mmol; in 5 ml of tetrahydrofuran) was added drop wise. The mixture was stirred at 0°C, after 30 min a TLC analysis showed that the starting was consumed. The grey suspension was quenched at 0°C by cautious addition of NaOH 1N solution until the suspension became white. The solution was extracted with ethyl acetate (3 x 5 ml), washed with brine (3 x 5 ml) and dried over Na_2SO_4. The resulting solution was filtered and concentrated under reduced pressure. The resulting product was used without further purification in the next step. Colourless oil, 0.276 g, (88%).

Synthesis of (2*S*,4*S*)-4-(methoxymethoxy)-2-vinylcyclopentanecarbaldehyde (170)

An oven-dried, round-bottomed flask, under nitrogen flux, was charged with methylene chloride dry (5 ml), a solution of **169** (0.300 g; 1.61 mmol, in 5 ml of methylene chloride) and PCC (1.042 g; 4.83 mmol), in the indicated order. The mixture was stirred at room temperature. After 3 h, a TLC analysis showed that the starting was consumed. The mixture was quenched with water (5 ml), extracted with methylene chloride (3 x 5 ml), washed with brine and dried over Na_2SO_4. The resulting solution was filtered and concentrated under reduced pressure. The product was purified by flash chromatography (petroleum ether-ethyl acetate 9:1) to afford 0,218 g (75%) of **170** as yellow oil.

Synthesis of (*E*)-ethyl 4-hydroxy-4-((2*S*,4*S*)-4-(methoxymethoxy)-2-vinylcyclopentyl)but-2-enoate (171)

To a solution of $CrCl_2$ (0.667 g; 5.43 mmol) and $NiCl_2$ (0.004 g; 0.032 mmol) in degassed DMF (4 ml) was added a solution of **170** (0.200 g; 1.08 mmol) and (*E*)-ethyl 3-iodoacrylate (0.279 ml; 2.17 mmol) in degassed DMF (4 ml), drop wise at room temperature. The reaction was stirred at room temperature and was monitored by TLC. Upon consumption of **170** the reaction was quenched with satured ammonium chloride solution (5 ml), extracted with ethyl acetate (3 x 5 ml), washed with brine (3 x 5 ml) and dried over Na_2SO_4. The resulting solution was filtered and concentrated under reduced pressure. The product was purified by flash chromatography (petroleum ether-ethyl acetate 7:3) to afford a mixture of two diastereoisomers: **A** (hydroxy group in alpha-position) 0.065 g and **B** (hydroxy group in beta-position) 0,085 g (67%) as colourless oil.

Synthesis of (*E*)-ethyl 4-(methoxymethoxy)-4-((2*S*,4*S*)-4-(methoxymethoxy)-2-vinylcyclopentyl)but-2-enoate (172)

An oven-dried, round-bottomed flask, under nitrogen flux, was charged with methylene chloride (2 ml), compound **171** (0.050 g; 0.176 mmol, in 2 ml of methylene chloride), and MOMCl (0.067 ml; 0.879 mmol) in the indicated order. After the solution was cooled to 0°C in the ice-bath, diisopropylamine (0.153 ml; 0.879 mmol) was added in one portion. After overnight reaction at room temperature, diethyl ether (5 ml) was added to the mixture and the solution was washed with satured ammonium chloride solution (5 ml) and dried over Na_2SO_4. The resulting orange solution was filtered and concentrated under reduced pressure. The product was purified by flash chromatography on silica (petroleum ether-ethyl acetate 8:2) to give 0.041 g (71%) of **172** (diastereoisomers mixture) as colourless oil.

Synthesis of (*E*)-4-(methoxymethoxy)-4-((2*S*,4*S*)-4-(methoxymethoxy)-2-vinylcyclopentyl)but-2-enoic acid (159)

To a solution of **172** (0.040 g; 0.122 mmol) in tetrahydrofuran (4 ml), water (4 ml) and methanol (0.8 ml) was added LiOH (0.012 g; 0.487 mmol) in one portion. The reaction was stirred at room temperature for 2 h. The mixture was quenched with satured ammonium chloride solution (4 ml), extracted with ethyl acetate (3 x 5 ml), washed with brine (3 x 5 ml) and dried over Na_2SO_4. The resulting colourless solution was filtered and concentrated under reduced pressure to afford **159** (diastereoisomers mixture, 0.038 g; 98%) as colourless oil. The product was used in the next step without further purification.

Synthesis of general intermediate 160

An oven-dried, round-bottomed flask, under argon flux, was charged with magnesium turnings (0.832 g; 35 mmol), THF (15 ml), and a crystal of iodine and the mixture was stirred at room temperature until the brown colour of the mixture disappeared. A solution of 4-bromo butane (**173**) (1.83 ml; 0.018 mmol) in THF (5 ml) was added drop wise and the mixture was stirred for an additional 1 h and was brought to -40°C, and CuI (0.066 g; 0.347 mmol) was added. The resulting mixture was stirred for an additional 10 min, at which time the epoxide (**174**) was added in a single portion. The mixture was allowed to warm to -15°C, stirred for an additional 2.5 h, and then quenched with satured ammonium chloride solution (10 ml) and extracted with ethyl acetate (3 x 10 ml), washed with brine (3 x 5 ml) and dried over Na_2SO_4. The resulting solution was filtered and concentrated under reduced pressure. The product was purified by flash chromatography on silica (petroleum ether-ethyl acetate 8:2) to give 0.544 g (62%) of **160** as yellow oil.

Synthesis of general intermediate 175

An oven-dried, round-bottomed flask, under nitrogen flux, was charged with methylene chloride (4 ml), compound **159** (0.048 g; 0.160 mmol), compound **160** (0.082 g; 0.639 mmol) and a catalytic amount of DMAP. The temperature was brought to 0°C and a solution of DCC (0.043 g; 0.192 mmol, in 2 ml of methylene chloride) was added drop wise. After 5 min the temperature was allowed to room temperature and the reaction was stirred for 4 h. The mixture was quenched with satured ammonium chloride solution (5 ml), extracted with methylene chloride (3 x 5 ml), washed with brine and dried over Na_2SO_4. The resulting solution was filtered and concentrated under reduced pressure. The product was purified by flash chromatography on silica (petroleum ether-ethyl acetate

95:5) to give 0.010 g (15%) of **175** (diastereoisomers mixture) as colourless oil.

Synthesis of general intermediate 176

An oven-dried, round bottomed flask, under nitrogen flux, was charged with methylene chloride (4 ml), compound **175** (0.010; 0.0244 mmol) and Grubb's catalyst 1st generation (0.0008 g; 0.00097 mmol). This solution was heated to reflux for 8 h and another portion of catalyst was added (0.0008 g; 0.00097 mmol). Reflux was continued for additional 8 h till the TLC analysis indicated that the starting material was consumed. The solution was allowed to cool and concentrated under reduced pressure. The residue was purified by flash chromatography on silica (methylene chloride-methanol 95:5) to give 0.0016 g (66%) of **176** (diastereisomers mixture).

Synthesis of Brefeldin A derivative 177

To a solution of compound **176** (0.004 g; 0.0105 mmol) in methanol (4 ml), was added at 0°C drop wise HCl 37% (0.240 ml). The reaction was stirred at room temperature until a TLC analysis showed that the starting material was consumed. The mixture was concentrated under reduced pressure and quenched with NaOH 1 N solution (3 ml), extracted with ethyl acetate (3 x 5 ml), washed with brine (3 x 2 ml) and dried over Na_2SO_4 to gave the product (0.002 g; 60%) as white solid.

(*S*,*E*)-ethyl 4-(2,2-dimethyl-1,3-dioxolan-4-yl)but-2-enoate (162)

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 6.94-6.87 (m, 1H); 5.91-5.87 (m, 1H); 4.24-4.14 (m, 3H); 4.06-4.03 (m, 1H); 3.58-3.55 (m, 1H); 2.54-2.40 (m, 2H); 1.40 (s, 3H); 1.33 (s, 3H); 1.27 (t, J= 7.02, 3H).

(*S*,*E*)-ethyl 5,6-dihydroxyhex-2-enoate (163)

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 7.04-6.94 (m, 1H); 5.98-5.93 (m, 1H); 4.22 (q, J= 8.24, 2H); 3.92-3.90 (m, 1H); 3.72-3.69 (m, 1H); 3.57-3.53 (m, 1H); 2.45-2.40 (m, 2H); 2.23 (d, J=8.72, 1H); 1.91-1.87 (m, 1H); 1.31 (t, J= 8.34, 3H).

(*S*,*E*)-ethyl 6-bromo-5-hydroxyhex-2-enoate (164)

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 6.99-6.89 (m, 1H); 5.97-5.91 (m, 1H); 4.22-4.15 (q, J=9.44, 2H); 3.98-3.92 (m, 1H); 3.54-3.38 (m, 2H); 2.61 (br s, 1H); 2.54-2.48 (m, 2H); 1.29 (t, J=8.34, 3H).

(*S*,*E*)-ethyl 6-bromo-5-(methoxymethoxy)hex-2-enoate (165)

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 6.94-6.87 (m, 1H); 5.95-5.90 (m, 1H); 4.72-4.66 (m, 2H); 4.20-4.14 (q, J=9.44, 2H); 4.20 (m, 1H); 3.57-3.44 (m, 2H); 3.43-3.39 (d, J=6.32, 3H); 1.27 (t, J=8.34, 3H).

(*S*)-trimethyl(2-(*p*-tolylsulfinyl)ethyl)silane (166)

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 7.52-7.50 (m, 2H); 7.35-7.33 (m, 2H); 2.83-2.44 (m, 5H); 2.44 (s, 3H); 0.84-0.78 (m, 2H); 0.01 (s, 9H).

(1*R*,2*S*,4*R*)-ethyl4-(methoxymethoxy)-2-((*S*)-1-((*R*)-*p*-tolylsulfinyl)-2 (trimethylsilyl)ethyl)cyclopentanecarboxylate (167)

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 7.49-7.30 (m, 5H); 4.78-4.64 (m, 2H); 4.26-4.14 (m, 2H); 3.40 (s, 3H); 2.82-2.71 (m, 2H);

(1*R*,2*S*,4*S*)-ethyl 4-(methoxymethoxy)-2-vinylcyclopentanecarboxylate (168)

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 5.87-5.65 (m, 1H); 5.15-4.95 (m, 2H); 4.62 (s, 2H); 4.32-4.17 (m, 3H); 3.38 (s, 3H); 2.82-2.67 (m, 2H); 2.45-2.25 (m, 1H); 2.17-2.12 (m, 2H); 1.68-1.52 (m, 1H); 1.24 (t, J=7.74, 3H).

((1*R*,2*S*,4*S*)-4-(methoxymethoxy)-2-vinylcyclopentyl)methanol (169)

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 5.93-5.81 (m, 1H); 5.06-5.01 (m, 2H); 4.64 (s, 2H); 4.20-4.12 (m, 1H); 3.37 (s, 3H); 2.83-2.74 (m, 2H); 2.32-2.27 (m, 1H); 2.12-1.97 (m, 3H); 1.67-1.61 (m 2H); 1.27-1.22 (m, 1H).

(2*S*,4*S*)-4-(methoxymethoxy)-2-vinylcyclopentanecarbaldehyde (170)

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 9.66 (s, 1H); 5.93-5.81 (m, 1H); 5.06-5.01 (m, 2H); 4.64 (s, 2H); 4.20-4.12 (m, 1H); 3.37 (s, 3H); 2.83-2.74 (m, 2H); 2.32-2.27 (m, 1H); 2.12-1.97 (m, 3H); 1.67-1.61 (m 2H); 1.27-1.22 (m, 1H).

(*E*)-ethyl 4-hydroxy-4-((2*S*,4*S*)-4-(methoxymethoxy)-2-vinylcyclopentyl)but-2-enoate (171)

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 6.94-6.88 (m, 1H); 6.06-6.02 (m, 1H); 5.89-5.80 (m, 1H); 5.07-4.96 (m, 2H); 4.60 (s, 2H); 4.29-4.26 (m, 1H); 4.22-4.18 (m, 2H); 4.16-4.11 (m, 1H); 3.34 (3H); 2.43-2.38 (2H); 2.23-2.12 (m, 2H); 1.87-1.86(m, 1H); 1.69-1.56 (m, 1H); 1.25 (t, J=8.24, 3H).

(*E*)-ethyl4-(methoxymethoxy)-4-((2*S*,4*S*)-4-(methoxymethoxy)-2-vinylcyclopentyl)but-2-enoate (172)

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 6.94-6.88 (m, 1H); 6.06-6.02 (m, 1H); 5.89-5.80 (m, 1H); 5.07-4.96 (m, 2H); 4.72 (s, 2H); 4.60 (s, 2H); 4.29-4.26 (m, 1H); 4.22-4.18 (m, 2H); 4.16-4.11 (m, 1H); 3.40 (s, 2H); 3.34 (3H); 2.43-2.38 (2H); 2.23-2.12 (m, 2H); 1.87-1.86(m, 1H); 1.69-1.56 (m, 1H); 1.25 (t, J=8.24, 3H).

(*E*)-4-(methoxymethoxy)-4-((2*S*,4*S*)-4-(methoxymethoxy)-2-vinylcyclopentyl)but-2-enoic acid (159)

^{1}H-NMR ($CDCl_3$ 300 MHz) δ: 11.03 (s, 1H); 6.94-6.88 (m, 1H); 6.06-6.02 (m, 1H); 5.89-5.80 (m, 1H); 5.07-4.96 (m, 2H); 4.72 (s, 2H); 4.60 (s, 2H); 4.29-4.26 (m, 1H); 4.16-4.11 (m, 1H); 3.40 (s, 2H); 3.34 (3H); 2.43-2.38 (2H); 2.23-2.12 (m, 2H); 1.87-1.86(m, 1H); 1.69-1.56 (m, 1H).

4. HIT VALIDATION IN DNA-ENCODED CHEMICAL LIBRARY

4.1 Technologies for Drug Discovery

The identification of small-molecules able to interact with a specific biological target of pharmaceutical interest is a central problem in chemistry, biology and pharmaceutical sciences.

As a matter of fact, the numbers of potential pharmaceutical targets for which no small-molecule probes are available largely exceeds the number of protein targets with reported tool compounds.[117] The sequencing of the human genome, together with advance in scientific knowledge and technologies breakthrough in several scientific fields (e.g. structural biology, improved analytics) has led to a substantial increase in potential targets. As a consequence, there is a growing need for efficient ligand discovery technologies.

The drug discovery and development pipeline is a complex process. The identification of new chemical entities (NCEs) capable of specific binding to biological target is not *per se* a guarantee that a drug has been discovered. Nevertheless, the identification of specific binders against relevant biological targets typically represents the starting point in the drug discovery pipeline, which can lead to a new drug.[118]

In the past, the drug discovery process started with the identification of bioactive molecules from natural sources, on the basis of random screening assays. Nowadays, drugs are often discovered by screening libraries of small molecules (also called chemical libraries) using suitable bioassay read-out.

In a classical drug discovery program, after the target identification and validation steps, the hit discovery process is generally performed by high-throughput screening[119,120,121] (HTS) of large collections of molecules, called chemical-libraries. HTS represents one of the key technologies for pharmaceutical drug discovery. Nevertheless, this approach has several limitations: (*i*) high false-positive rates; (*ii*) primary hits typically need to be optimized by medicinal chemistry efforts before they can be considered for industrial applications; (*iii*) the synthesis, quality control and management of conventional chemical libraries are associated with high costs, lengthy procedures and complex logistics. Moreover, not all protein targets are suitable for HTS assays or can be produced in sufficient amounts to screen hundreds of thousands of compounds.

Over the last decades, there have been new powerful technologies for chemical library construction and screening. In this context, fragment based discovery approaches,[118,122] virtual drug discovery,[123,124,125] speed screen procedures and DNA-Encoded-Libraries (DECLs) have been considered.

DNA-encoded chemical libraries (which are described below), were inspired by phage display methodologies and by biological combinatorial libraries. In Phage Display Technology,[126,127] billions of polypeptides, displayed on the surface of the filamentous phage, exhibit a physical linkage between their potential binding properties (“phenotype”) and the corresponding genetic information contained inside the phage (“genotype”). Libraries can be “panned” on target antigens immobilized on solid support, and washing cycles allow the isolation of binding molecules.[128] Following a similar logic, ribosome display[129] allows the association of mRNA, coding for a polypeptide, together with its

translation product. Phage display, ribosome display, and conceptually similar technologies (peptides on plasmid; yeast display, puromycin display, covalent DNA display)[130,131] allow the amplification of the genetic information of selected binders after affinity capture, e.g., through bacterial amplification (phage display, Figure 33A). As these technologies rely on transcription and translation, they are only compatible with the display and selection of polypeptides and cannot directly be used for the display and selection of small organic molecules. In light of such limitations it would be useful to devise strategies also for the identification of small organic molecules, capable of binding to target proteins with high affinity and specificity, based on the association of individual chemical compounds to unique DNA-fragments serving as amplifiable identification barcodes (Figure 33B).[132,133]

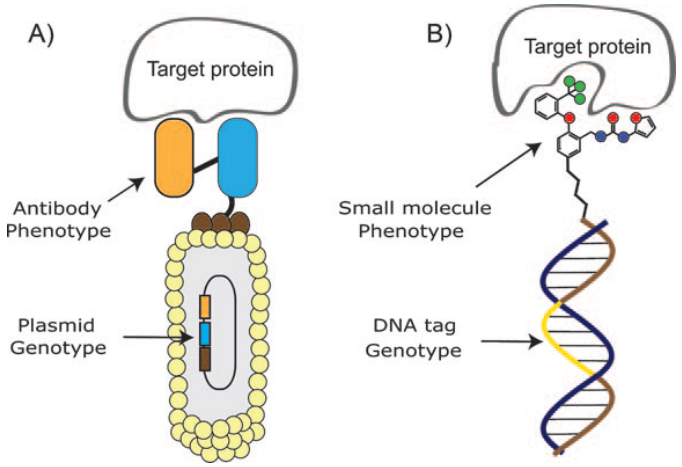


Figure 33. Analogy of antibody phage display technology and DNA-encoded chemical libraries. A) A phage particle displaying a scFv antibody fragment (depicted in orange and blue) as a fusion protein to the minor coat protein pIII on the tip of the phage. The phage particle provides a physical link between the antibody fragment (the "phenotype") and the gene encoding for it ("genotype"). B) A conjugate of a small organic molecule (potentially capable of binding to a target protein) and a DNA oligonucleotide. The DNA tag ("genotype") serves as an identification barcode, unambiguously encoding the molecule ("phenotype") to which it is attached. Picture was readapted from ref.[134]

In 1992, Brenner and Lerner introduced the concept of DNA-Encoded-Libraries (DECLs). The authors proposed the synthesis of peptide libraries on individual beads, containing oligonucleotide barcodes, as amplifiable identifications tags (Figure 34).[132]

In contrast with display technologies, DECLs do not represent a "biosynthetic" encoding method, since nucleic acid does not direct the synthesis of the binding molecule. Indeed, DNA fragments merely serve as "identifiers" to the chemical structure to which they are attached. DNA barcodes can be easily coupled to molecules during library synthesis procedures and can be easily identified by high-throughput DNA sequencing.

During the last years, DNA-encoded chemical procedures have been improved by omitting beads in the library construction step, as described by the research groups of Prof. Dario Neri at ETH Zürich and of Prof. David Liu at Harvard University.[135] The technology does not reqiore complex logistics and large compounds collections can be generated by split and pool methods. Thus, inexpensive DECLs provide an avenue for researchers from academia or small companies to identify hits that bind to targets of pharmaceutical interest. Several examples of binders discovered by DECL technology are described in literature, including: (*i*) ADAMTS-5 inhibitors,[136] (*ii*) tankyrase 1 inhibitors,[137] (*iii*) XIPA inhibitors[138] and (*iv*) integrin lymphocyte function-associated antigen 1 (LFA-1) antagonists.[139]

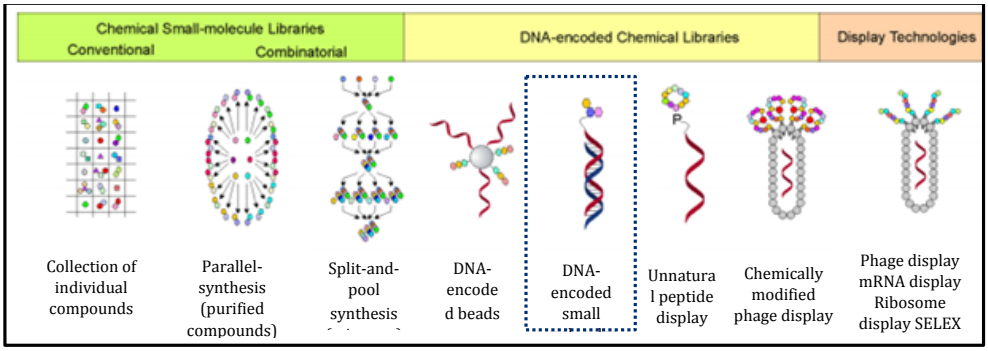


Figure 34. Schematic representation of different types of chemical libraries. DNA-encoded chemical libraries can be considered as an intermediate technology between the screening of compound libraries and the selection of encoded combinatorial libraries of polypeptides (e.g., phage display libraries).Picture adapted from reference[140]

4.2 DNA-Encoded Chemical Libraries (DECLs)

DNA-encoded chemical libraries represent collections of small-molecules coupled to distinctive DNA fragments (barcodes), which allow the identification and relative quantification of molecules in complex mixtures.

In contrast to HTS swelling procedures, DECLs allow for the creation and screening of very large libraries, which can no longer be handled as 'one well, one compound'.

In principle, the use of DECL technology presents distinctive advantages, compared with classical HTS lead discovery: (*i*) a very large library size[141] can be obtained, depending on the number of sets of building blocks used.[142] For example, three sets of building blocks with 1,000 members each will yield a library with 1,0003 = 109 displayed compounds; (*ii*) a variety of DNA-compatible synthesis approaches are available;[143] (*iii*) a DECL may be stored in only one small vessel, as the DNA barcodes allow the unambiguous identification of each library member stored as a pool. This feature represents a distinctive advantage compared to conventional chemical libraries, in which

individual compounds must be stored separately, typically in microtiter plates; (*iv*) structure-activity relationships (SARs) can be obtained from the selection results if structurally related compounds are incorporated into the library.[137]

DNA-encoded chemical library synthesis can be carried out with numerous and structurally diverse chemical building blocks. The most commonly used reactions include: (*i*) reductive amination, (*ii*) amide bond formation, (*iii*) Suzuki coupling and (*iv*) nucleophilic aromatic substitutions.[144]

These libraries are typically synthesized by a stepwise assembly procedure, in which a variety of chemical building blocks (BBs) can be combined in different ways to create the final structures of the library members.

A number of different types of DECLs have been developed. They can be classified as “single-pharmacophore” and “dual-pharmacophore” chemical libraries, depending on whether a single molecule is coupled to a DNA fragment, or whether two molecules are attached at the adjoining extremities of two complementary DNA strands (Figure 35). In single-pharmacophore chemical libraries, individual molecules (regardless their complexity) are attached to individual DNA fragments. By contrast, in dual-pharmacophore libraries (also called ESAC libraries), two sets of compounds are attached to the adjacent complementary strands of DNA, thus allowing for a combinatorial self-assembly of the library.[135]

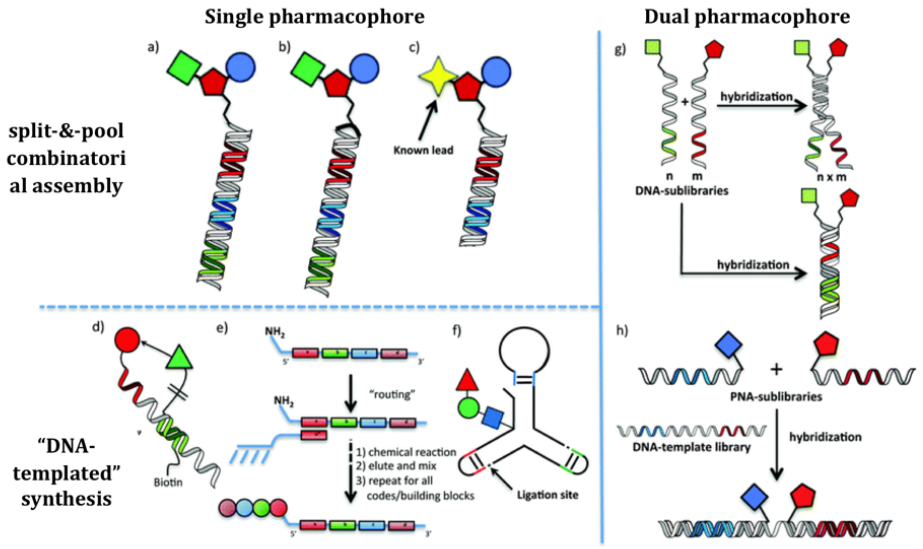


Figure 35. Differences between "single-pharmacophore" and "dual-pharmacophore" chemical libraries. Picture adapted from N. Favalli, master thesis 2015.

4.2.1 Single-pharmacophore DNA-encoded libraries

In single-pharmacophore DNA-encoded chemical libraries[134] individual organic molecules are attached to individual DNA fragments, serving as unambiguous identification barcodes. Different methodologies can be used to generate single-pharmacophore DECLs.

The iterative assembly of sets of BBs in a "split-and-pool" methodology combined with the stepwise DNA-tagging possibly represents the most effective strategy used for the preparation of DECLs. Starting from just few hundred small molecules and oligonucleotides this strategy speed up the preparation of DECLs containing millions or even billions of compounds. The procedure mainly consists in the following sequential steps (Figure 36):[145]

a) *coupling* reactions of a first set of chemical moieties (*w* building blocks of type “A”) to amino-tagged oligonucleotides carrying code 1 (i.e. through amide bond formation);

b) oligonucleotide conjugates *pooling* in a single mixture, as any single conjugates contain a DNA sequence necessary for the unambiguous identification of the corresponding compound, and then splitting into different reaction vessels;

c) incorporation of the second group of small molecules (*x*, type “B”) and *encoding* of the second building block “B” through hybridization and fill-in reaction mediated by Klenow DNA polymerase with an oligonucleotide fragment carrying code 2;

d) *pooling* of all reactions and incorporation of the third set of building blocks (*y*, type “C”) resulting in a *n* x *m* x *y* member DECL.

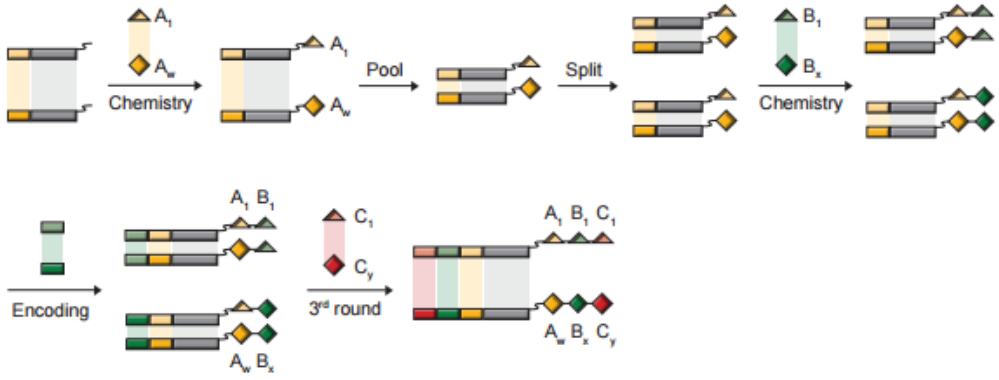


Figure 36. Schematic synthesis of a three building block DNA-encoded chemical library. Picture adapted from M Leimbacher, phd thesis 2012.

These steps can be repeated *n* times and, at the end of this synthesis procedure; libraries of very large size (usually is between 10^5 and 10^9 encoded small molecules) can be created. However, it is worth to notice that the efficacy, purity and quality of the library may decrease as the library size grows. As an example, this is the case when the chemical

transformations for the coupling of individual small molecules do not proceed with quantitative yields and when the molecular weight of individual pharmacophore exceed 600 Da (i.e., using more than 3 sets of building blocks), thus violating the Lipinski's rule of five.[140] Other important parameters to consider for the design of DECLs, are (*i*) the number of diversity points (including the sets of BBs and synthetic cycles) and (*ii*) the assembly geometry, which directly influences the chemical properties of the encoded BBs.

An alternative strategy for the construction of single-pharmacophore DECLs is represented by the method called "DNA-routing".[140] Harbury and co-workers used a solid phase containing complementary DNA-codons and purification-elution steps for the sequential capture and modification of growing molecular structure (Figure 35).[146] Similarly, Hansen *and co.* described three-way DNA-hairpin-looped junctions able to assist the library synthesis through the transfer of appropriate donor chemical moieties onto an acceptor site (Figure 35).[147]

4.2.2 Dual-pharmacophore DNA-encoded libraries

Dual-pharmacophore DECLs, also known as "Encoded Self-Assembling Chemical libraries" (ESAC library technology), are composed by two molecules coupled at the extremities (5' and 3' ends) of two complementary DNA strands. The main advantage of dual-pharmacophore DNA-encoded libraries with respect to single-pharmacophore DNA-encoded libraries refers to the fact that the linkers that connect two adjacent BBs to the individual DNA-tags are quite flexible. This "bidentate" binding mode may produces an increase in binding affinity due to the simultaneous engagement of two distinct non-overlapping binding sites on the target.[140] The group of Prof. Neri has

pioneered in the field of generating dual-pharmacophore DNA-encoded.[148] The synthesis of ESAC libraries is based on combinatorial assembly of reciprocal sub-libraries (often called A and B) through DNA hybridization. Typically, sub-library A is composed of small chemical moieties connected to the 5'-extremity of single stranded encoded oligonucleotides. Similarly, a sub-library B displays BBs attached at the 3'-extremity of specifically tagged oligonucleotides. The combination of two sub-libraries (A x B) with 1000 members give rise to a 1,000,000 member ESAC library (Figure 37). Complementary DNA sequences allows the combinatorial assembly of the two sublibraries. The oligonucleotide conjugates are individually synthesized by coupling of chemical entities, purified by high-performance liquid chromatography (HPLC) and characterized by liquid chromatography-mass spectrometry (LC-MS). Moreover, the DNA tags in addition to their role as identification barcodes confer enhanced solubility to small organic molecules in aqueous phase. It is worth to mention that other strategies for ESAC synthesis are emerging, such as that propose by Winssinger *and co.,* in which PNA-conjugate molecules are combinatorial self-assembled on a DNA template.

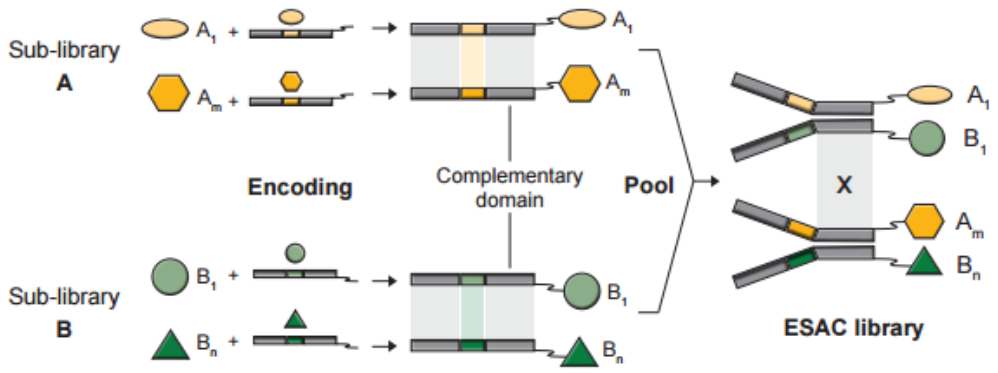


Figure 37. General scheme of ESAC library construction. Picture adapted from M Leimbacher, phd thesis 2012.

4.2.3 Affinity selection test.

DNA-encoded chemical libraries can be selected by affinity capture on virtually any target protein of interest immobilised on a solid support (Figure 38),[149,150] in analogy to what happens in an antibody phage display library selections.

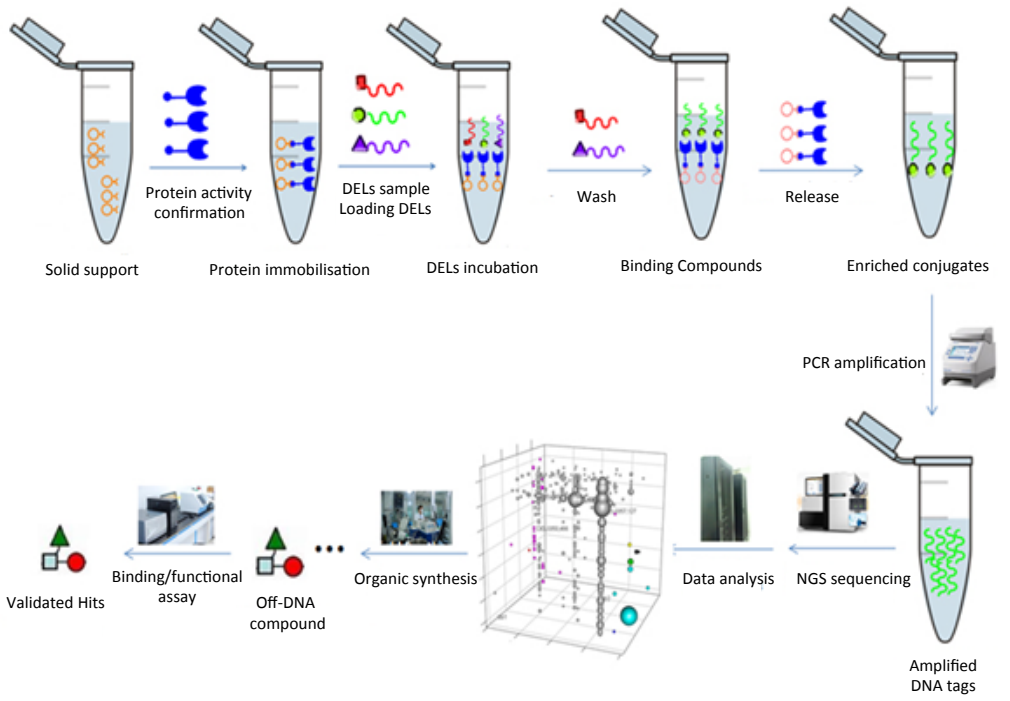


Figure 38. "Panning" (i.e., affinity selection procedure) for the identification of binding compounds from DNA-encoded chemical libraries.
(Picture adapted from [http://www.hitgen.com/?page_id=6864])

The library is typically incubated at low concentration (fM to pM concentration of individual library members) with target protein of interest immobilised on a solid support. Upon separation of the "bound" and "unbound" fractions, the DNA tags at the enriched compounds are PCR-amplified (Figure 38). The resulting amplicon mixture must then be deconvoluted, by quantifying the relative concentration of all library members (e.g., by high-throughput sequencing) and counting of the corresponding tags. Subsequently, identified "hits" have been characterized biologically or biochemically, in the presence or absence of linked DNA (Figure 38).[151,152]

During affinity selection procedures, the target protein of interest is typically immobilized on magnetic beads and subsequently incubated

with the encoded chemical library. If suitable small organic happens to be present in the DECL, they will be retained on the magnetic beads because of a binding interaction with the immobilized target protein (Figure 38). Several washing steps remove nonbinding library members, whereas binding library molecules remain associated with the solid support and can be recovered for PCR amplification procedures.

The selection buffer used for the affinity selection depends on the target protein. Most proteins can readily be stored and screened in PBS. If PBS may not be used as selection buffer, e.g., for the screening of phosphatases, alternative buffers such as HEPES may be used.

Additional precautions have to be considered. For example, Tween 20 is usually added to the selection solution in order to prevent the coagulation of magnetic beads and the sticking of beads to the plastic plate and tip comb. In some case, excessive amounts of Tween 20 may lead to suboptimal selection or to the dissociation of multi-subunit proteins.

The number of washing steps may be adjusted for the individual selection procedures. If a high-affinity ligand is expected, the stringency (e.g., number and duration) of the washing steps can be increased. By contrast, if a lower binding affinity is expected, fewer washing steps of shorter duration should be considered. A good compromise is found in the use of five washing steps with duration of 30 seconds each (Figure 39).[153]

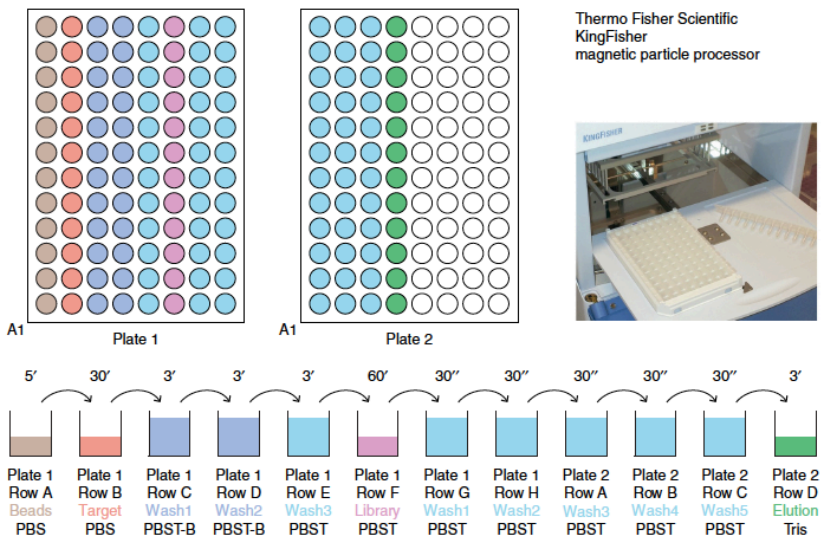


Figure 39. Plate loading scheme. Top, KingFisher 200-µl plates shown from above. As the magnetic particle processor transfers all magnetic beads contained in a row (e.g., A1 to A12) during each step, the wells are loaded in a row-wise manner (e.g., target proteins in row B of plate 1). The respective solutions are filled into the wells as depicted at the bottom; 200 µl per well for all washing steps and 100 µl of washed beads, target protein, library and Tris buffer. The numbers above the arrows indicate the incubation time of the beads at each step. In this setup, each column allows the performance of an independent selection. Although the handling by the magnetic particle processor is identical for all plate columns, individual selection parameters may be varied in terms of target protein, DECL type and general buffer composition. Picture adapted from reference[153]

The DNA-tagged binders are eventually eluted and the DNA-codes can be amplified by PCR in order to obtain sufficient quantities of double-stranded DNA to sequence. The resulting amplicon mixture must then be deconvoluted, and identified “hits” can be characterized biologically or biochemically, in the presence or absence of linked DNA.

4.2.4 High throughput DNA sequencing technologies.

The technology used for DNA sequencing mainly involves the high-throughput sequencing (HTP) developed by Illumina/Solexa. The set of data, derived from the sequence counts of individual compounds, is submitted to statistical evaluation and represented by graphical display.

The HTP results are visualized by using three-dimensional matrices where the code combinations are showed in the *xy* plane and the corresponding sequence counts for individual library components in the *z* axis. The HTP data resulting from a library composed by three sets of BBs would require a four-dimensional space. In practice, the code combinations are visualized in *xyz* space, while the sequence count is represented by spheres of different colours and size (Figure 40). It is reasonable to assume that compounds with high binding affinity for the target protein are preferentially detected preferentially with HTP, thus yielding "spikers" in the fingerprints of library selections.[137]

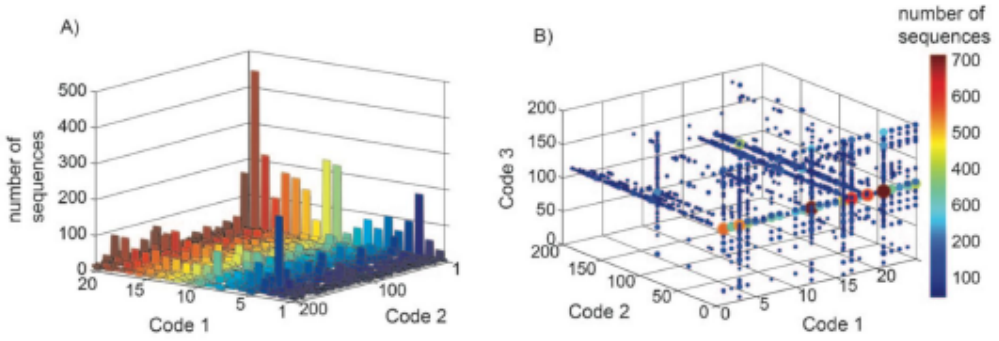


Figure 40. Evaluation of DECLs selections after HTS decoding. A) Sequence plot of a two BBs DECL. B) Sequence plot of a three BBs DECL. (Picture adapted from N. Favalli's master thesis, 2015)

Other HTP DNA sequencing technologies can be used, such as the "Roche 454 technology",[154] which uses emulsion-based PCR with the aim to amplify a single DNA strand on individual beads. After the PCR step, the beads are divided on microscopic wells on a picotiter plate and analyzed by Pyrosequencing followed by luminescence detection.

Compared with Roche 454 technology, Illumina produces a much higher amount of sequence data (about 20 GB) per run. On the other hand, the

Illumina platform generates shorter read lengths (currently up to 100 bp) and it needs a longer run time (several days) to obtain the final data. The libraries synthetized within the laboratory of Prof. Neri, based on two or three sets of building blocks, are encoded with DNA fragments shorter than 80 bases and for these reasons are compatible both with 454 and with Illumina sequencing.

Another relevant sequencing method is represented by the “applied Biosystem’s SOLiD technology” which depends on a sequencing-by-ligation chemistry. Similar to the 454 platform, DNA fragments immobilized on beads are amplified by emulsion PCR. However, the beads are then blocked on a slide and the DNA fragments are interrogated through ligation steps with labeled oligonucleotide probes. The result of the process is that each base is "sequenced" twice, allowing increased accuracy. The sequence output of SOLiD is comparable to that observed in Illumina sequencing. At the moment, about 15-30 GB of 50 bp reads can be achieved in a week.[155]

4.2.5 Hit validation in DECLs.

After selection of binding molecules from a DNA-encoded chemical library, preferential binders are typically resynthesized without DNA tag in order to confirm protein interaction and to determine affinity constants.[156] Good correlations between the sequencing counts after the selection experiment and the experimental affinity of the compounds without DNA tag have been reported. However, large DNA-encoded libraries may identify hundreds of enriched compounds, making conventional hit-validation processes (e.g., affinity or inhibition measurements, fluorescence polarization measurements of fluorescein conjugated compounds) more cumbersome. In this case, additional

selections with an increased selection pressure might define the number of hits to validate.[156]

Different hit-validation strategies for DNA-conjugates in biological assays have been recently explored. They include: (*i*) the measurement of a solution-based reporter assay (e.g. inhibition assay or competition assay) using a DNA-tagged small molecule and the target protein (Figure 41 A); (*ii*) the electromobility shift assay that measures the affinity of the DNA-conjugates with a protein (Figure 41 B); and (*iii*) the affinity chromatography of radiolabeled DNA-conjugates able to rank the hits according to their retention time (Figure 41 C). Moreover, the target protein can be immobilized on a solid support (i.e., ELISA plate), and the oligonucleotide can be employed for detection (Figure 41 D). Another possible strategy features a reverse setup, in which a DNA-conjugate is immobilised on a solid support and the interacting protein is detected using affinity reagents (e.g. antibodies, Figure 41 E).[157]

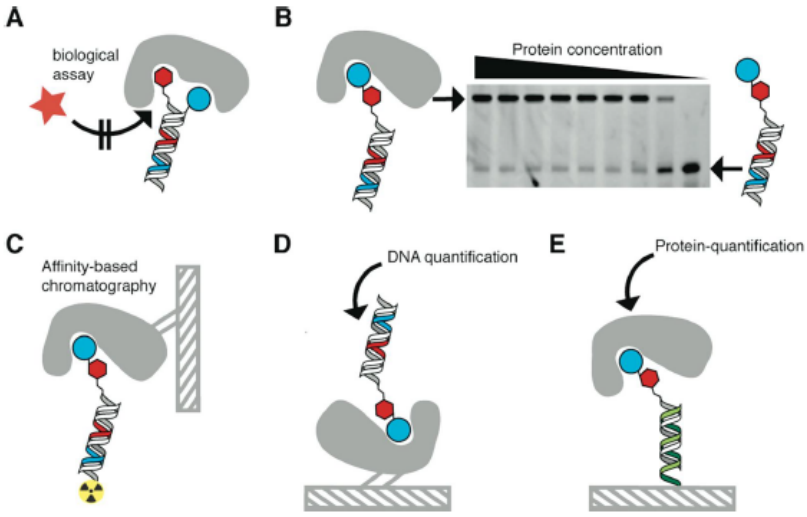


Figure 41. Hit-validation strategies using DNA-conjugated small molecules. (A) Scheme for the measurement of a reporter assay (e.g., inhibition assay, competition assay) with a small molecule conjugated to a DNA tag and a protein of interest. (B) Affinity-measurement of a DNA-conjugated small molecule with a protein in an electromobility shift assay: Complex formation between the protein and the DNA-conjugated ligand leads to band shift in a polyacrylamide gel electrophoresis compared to the free DNA-conjugated ligand. (C) Affinity chromatography of radiolabeled DNA-small molecule conjugates allows ranking of binding structures according to their retention time on the resin during washing cycles. (D) In analogy to the hit-validation of polypeptide display technologies, the target protein can be immobilized on a solid support and the DNA-tag used for quantification (e.g., using PCR, biotin-streptavidin interaction, or surface plasmon resonance measurements). (E) The reverse setup consists of DNA-conjugate immobilization and protein quantification (e.g., using antibodies, biotin-streptavidin interaction, or surface plasmon resonance measurements).Picture adapted from reference[157]

4.3 Aim of The Project

During the course of my PhD, I have worked, for several months, at the Institute of Pharmaceutical Sciences – ETH Zürich, in the laboratory of Prof. Dario Neri.

During this time, I have mainly focused on the hit validation process for a large DNA-encoded chemical library, containing over 30 millions of macrocycle derivatives.

The macrocycle DECL on which I have worked consisted of a cyclic-peptide scaffold, functionalized in three positions with carboxylic acids of different structure (building blocks). The library, comprising 283 x 386 x

324 = 35393112 members, had previously been synthesized by Dr Yizhou Li. In Figure 42, the results of an affinity selection against a splice variant of fibronectin are depicted. The analysis of the plot reported in Figure 42 revealed that the building block 238, from Code3, is preferentially enriched, being present in almost all the compounds analysed (Figure 42 A). The 2-D projection of the results can be useful in understanding which other building blocks (R_1 or BB_1 and R_2 or BB_2) are enriched and are plus likely interact with the protein under investigation (Figure 42 B).

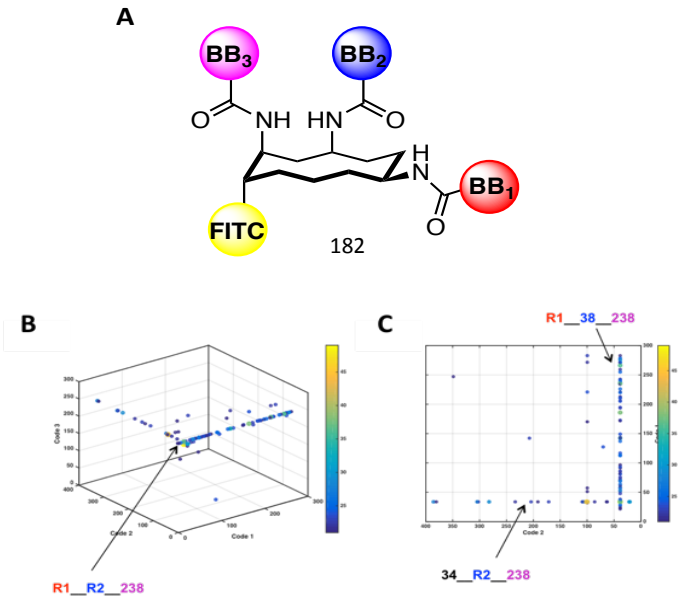


Figure 42. Analysis of Affinity selection. A) Schematic representation of the general structure of the compounds. B) Sequence plot of a three BBs DECL; C) 2-D projection (x= code 2; y= code 1) of the Sequence plot.

Among the compounds identified in Figure 42B, we decided to synthesized the 3 most enriched cyclic peptide derivatives, bearing the BB 238 in position 3 and other different building blocks in position R_1 and R_2.

Each of the selected compounds was synthesized, labelled with a fluorescine tag and purified by HPLC (Figure 43). Finally, a sandwich ELISA test was performed *in vitro*, in order to obtain preliminarily confirm, the action of target-ligand interaction.

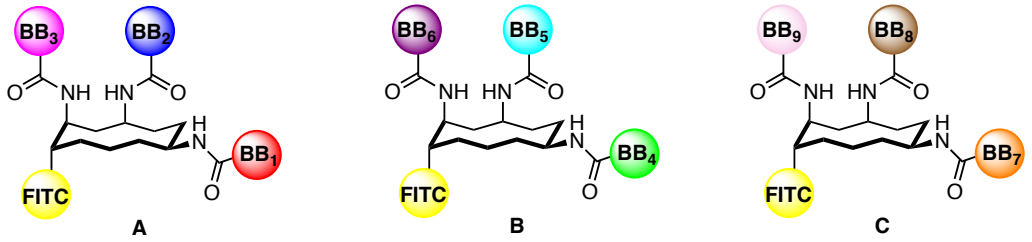


Figure 43. Synthetized compounds.

4.4 Macrocycle Peptide Scaffold

As explained in the previous paragraph, the macrocyclic core structure of the library consists of a cyclic peptide.

This macrocycle has been designed with distinctive steric and structural requirements. In particular, Dr Yizhou Li aimed at directing three sets of building blocks with the same spatial orientation toward the target.

The macrocyclic compound **182** is synthesized from a linear macromolecule **184**, which contains four side chains carrying primarily amino groups. These functionalities are suitably protected with selectively removable protecting groups (PG), which are stable during the macrocycle synthesis. The building block are added one by one during the construction of the peptide. After ring closing reaction, the Boc protecting group (tert-butyl carbamate) has to be removed from macrocyclic compound to allow the addition of the fluorescent probe. The other protecting groups are then removed one by one during the synthesis. In the figure below (Figure 43), I present a retrosynthetic

approach for the synthesis of macrocyclic compound. The linear macromolecule **184**, is closed using a condensation reaction. The obtained macrocycle **183** is then selectively deprotected (to remove the Boc protecting group). Finally, fluorescent probe is added to the scaffold yielding compound **182**.

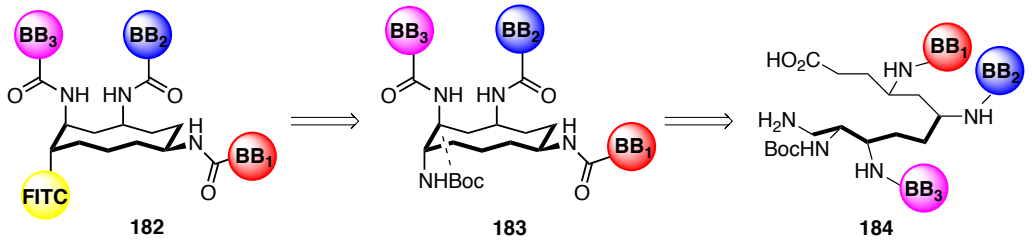


Figure 43. A possible retrosynthetic approach to obtain the macrocyclic peptide scaffold **182**.

As these compounds have not yet been patented, we cannot disclose their structure to the finest detail.

4.5 Chemistry

4.5.1 Synthesis Of The Linear Macromolecule

The desired macrocyclic peptides were obtained through a solid phase peptide synthesis (SPPS). The linear macromolecule **184** was synthesized starting from resin **186** containing the first monomer (bound to the resin by ester bond) featuring one free reactive nucleophilic group. The chain was elongated by adding pre-activated monomer **187** (to increase the electrophilicity of X, Figure 44). However, the nucleophilic group of **187** had to be suitably protected to avoid polymerization reactions or other side reactions. After the coupling, the protective group of the nucleophilic group of **187** was removed and the process was iterated several times, yielding the macromolecule **185**.

The building blocks consisting of previously activated carboxylic acids, were added one by one during the construction of the peptide.

In a final step, the resin was cleaved in acidic conditions (trifluoroethanol and acetic acid in dichloromethane solution) and the product **184** was recovered in good yields (Figure 44). The choice of the monomers was carefully planned to obtain the linear macromolecule with the desired properties. A preferential conformation in solution (strongly dependent from the chosen monomers) promoted the ring closing reaction, rather than a polymerization reaction. Moreover, the choice of monomers has to allow the ramification of the macromolecule in the right position with four functionalizable (pre-protected) amino groups (Scheme 18).

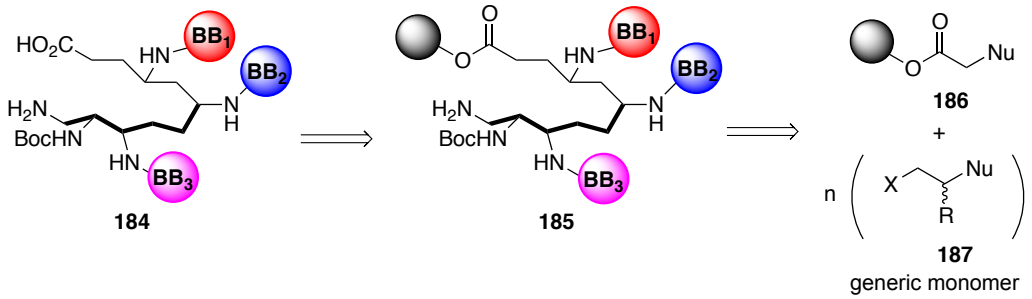
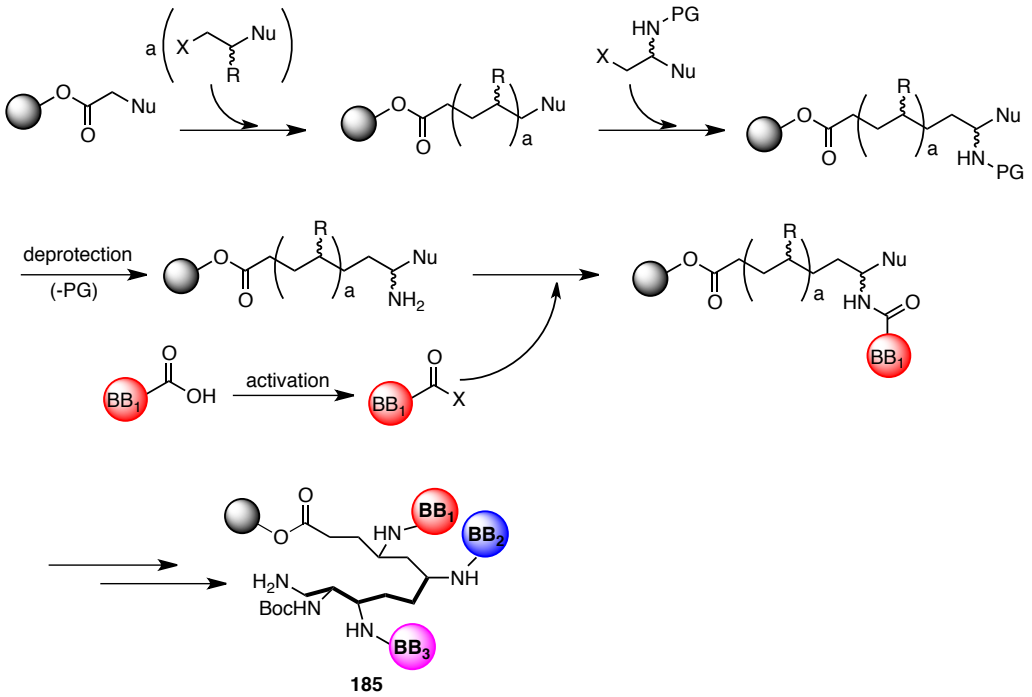


Figure 44. Retrosynthetic scheme for the preparation of linear macromolecule **3**.



Scheme 18. Solid phase synthesis of linear macromolecule **185**.

4.5.2 Cycloaddition Reaction

Compound **184** was closed by intermolecular reaction between the amino group and the carboxylic group (condensation reaction, Figure 45) present in the same molecule. The carboxylic acid was activated using PyBOP ((Benzotriazol-1-yloxy) tripyrrolidino phosphoniumhexafluoro phosphate, Figure 46) and DIPEA (Ethyldiisopropylamine) as a base (Figure 47). The labile intermediate reacts immediately with the amino group yielding a stable amide. In order to promote the cyclization reaction (intramolecular reaction) compared to polymerization reaction (intermolecular reaction), we worked in high dilution conditions (0.5 mM solution of linear compound **184** in dry DMF).

After the cyclization reaction, the product **183** was purified by HPLC.

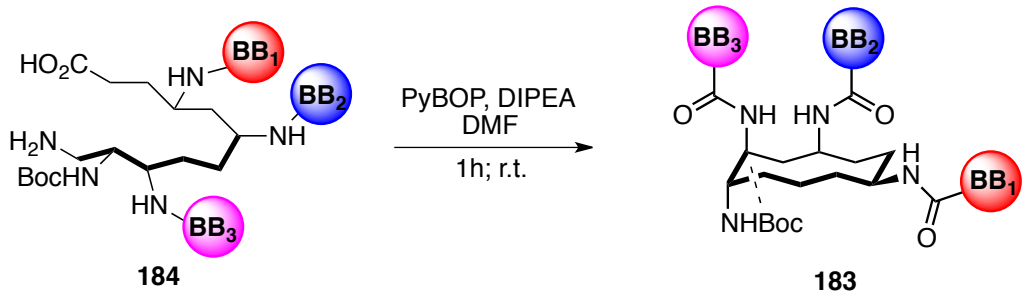


Figure 45. Cycloaddition reaction. In this cycloaddition reaction, the carboxylic acid 4 is activated by PyBOP and (in high dilution conditions) the activated carboxylic acid reacts intramolecularly with the amino group present in the same molecule.

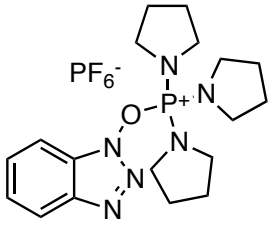


Figure 46. Chemical structure of PyBOP

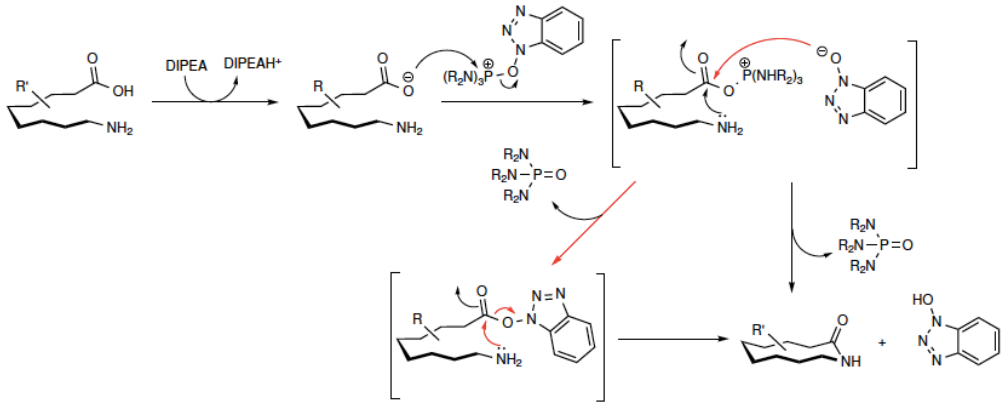


Figure 47. Mechanism of the cycloaddition reaction activated by PyBOP.

4.5.3 Boc Deprotection Reaction

The Boc amino-protected macrocycle **183** was selectively deprotected in strong acidic conditions (Figure 48) obtained by dissolving the compound 3 in water and adding a large amount of 2,2,2-trifluoroacetic acid (TFA:water = 20:1). The resulting solution was stirred for 1 hour at room temperature. Finally, the solvent was removed and the crude product **188** was isolated in quantitative yields. The reaction was followed by LC-MS and HPLC analysis.

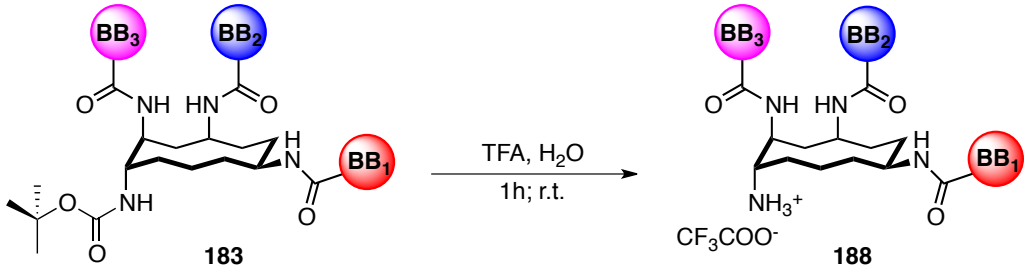


Figure 48. Boc deprotection reaction.

4.5.4 Addition of The Fluorescent Probe

Fluorescein isothiocyanate is a convenient reagent for covalently attaching the fluorescein chromophore to amino groups of peptides.

The reaction with fluorescein isothiocyanate was obtained by dissolving compound **188** in a mixture of pyridine, N,N-dimethyl formamide and dichloromethane (12:7:5 ratio), then adding fluorescein isothiocyanate (1.5 equivalents) and 0.1% of DIPEA. The reaction was conducted at room temperature for 20 minutes (Figure 49). After the analysis of reaction by UPLC revealed that all the starting material had been converted in the desired product. The crude of reaction was purified by HPLC, leading to final desired cyclopeptide **182** with high purity, ready to be tested in affinity measurements.

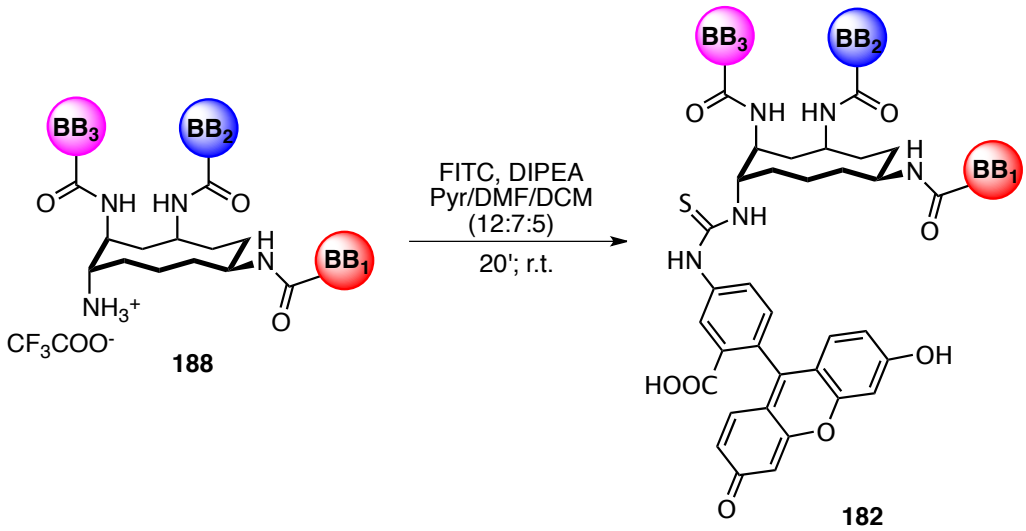


Figure 49. Reaction with Fluoresceine isothiocyanate.

4.6 RESULTS

Preliminary sandwich ELISA tests were performed on compound **A**, **B** and **C**. The experimental scheme for this assay is depicted in Figure 50.

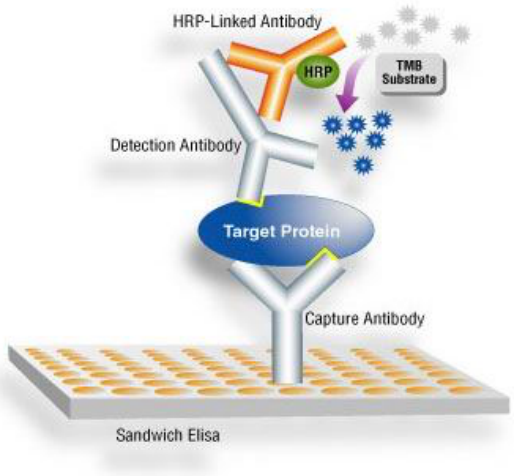


Figure 50. Sandwich ELISA assay. Plate is coated with a capture antibody; sample is added, and any antigen present binds to capture antibody; detecting antibody is added, and binds to antigen; enzyme-linked secondary antibody is added, and binds to detecting antibody; substrate is added, and is converted by enzyme to detectable form. Picture adapted from (http://www.iaszoology.com/elisa/).

Only compound **B** showed a specific interaction with the cognate target (EDA-coated fibronectin) (Figure 51). From the analysis of the plot we could observe how the signal increases proportionally with the binder concentration. By contrast, the un-modified cyclic peptide scaffold did not exhibit any detectable binding interaction. These are preliminary results and further tests will be performed, so as to confirm the specific interaction of compound **B** with the target of interest and to better characterized such compound.

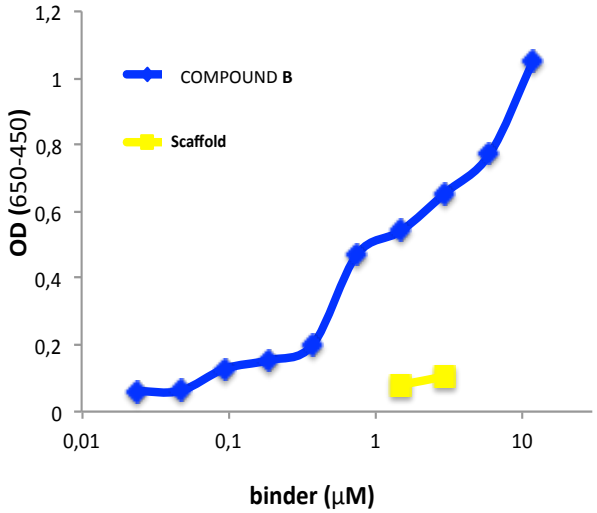


Figure 51. Results of an ELISA assay for compound **B**.

4.7 Conclusions

In the context of the work conducted at ETH Zurich, three cyclic-peptides (compound A, B, and C) carrying different building blocks at three different positions, were synthesized. All the compounds were labelled with a fluorescein tag, thus facilitating their characterization in an ELISA test. All the aforementioned compounds were purified by

HPLC, and the analysis of the corresponding chromatograms confirmed their purity (Figure 20-22). Among the compounds tested, compound B exhibited the most promising binding profile (Figure 19). Further assays will be required in order to quantitatively assess the dissociation constant at this molecule towards splice isoform at fibronectin, counting the EDA domain.

4.8 HPLC Analysis of Purified Compounds

4.8.1 Chromatogram of compound A

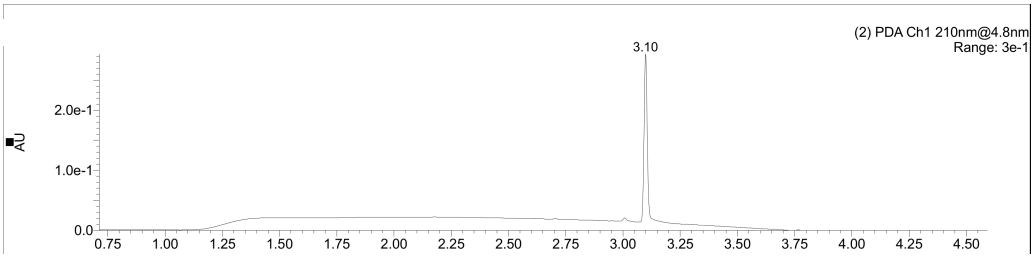


Figure 52. HPLC chromatogram of compound 4, registered at 260 nm. Elution gradient: CH3CN:Buffer A (0.1% TFA in m.Q. water) from 5:95 to 60:40 in 22 minutes. In the x-axis the time (min) is reported while in the y-axis the absorbance of the compound

4.8.2 Chromatogram of compound B

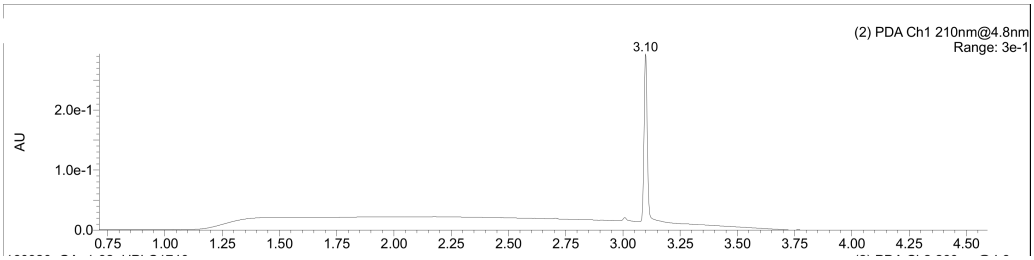


Figure 53. HPLC chromatogram of compound 4, registered at 260 nm. Elution gradient: CH3CN:Buffer A (0.1% TFA in m.Q. water) from 5:95 to 60:40 in 22 minutes. In the x-axis the time (min) is reported while in the y-axis the absorbance of the compound.

4.8.3 Chromatogram of compound C

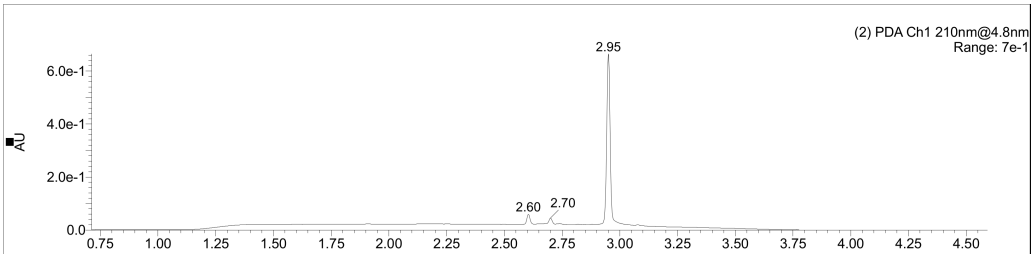


Figure 54. HPLC chromatogram of compound 4, registered at 260 nm. Elution gradient: CH3CN:Buffer A (0.1% TFA in m.Q. water) from 5:95 to 60:40 in 22 minutes. In the x-axis the time (min) is reported while in the y-axis the absorbance of the compound.

5. MATERIALS AND METHODS

All the reagents were purchased from Sigma-Aldrich, Alfa-Aesar, and Enamine at reagent purity and, unless otherwise noted, were used without any further purification. Dry solvents used in the reactions were obtained by distillation of technical grade materials over appropriate dehydrating agents. MCRs were performed using CEM Microwave Synthesizer-Discover model. Reactions were monitored by thin layer chromatography on silica gel-coated aluminum foils (silica gel on Al foils, SUPELCO Analytical, Sigma-Aldrich) at 254 and 365 nm. Where indicated, intermediates and final products were purified by silica gel flash chromatography (silica gel, 0.040−0.063 mm), using appropriate solvent mixtures. 1H NMR and 13C NMR spectra were recorded on a BRUKER AVANCE spectrometer at 400 and 100 MHz, respectively, with TMS as internal standard. For STD experiments please refer to the
1H NMR spectra are reported in this order: multiplicity and number of protons. Standard abbreviation indicating the multiplicity were used as follows: s = singlet, d = doublet, dd = doublet of doublets, t = triplet, q = quadruplet, m = multiplet, and br = broad signal. HPLC/MS experiments were performed with an Agilent 1100 series HPLC apparatus, equipped with a Waters Symmetry C18, 3.5 µm, 4.6 mm Å~ 75 mm column and an MS: Applied Biosystem/ MDS SCIEX instrument, with API 150EX ion surce. HRMS experiments were performed with an LTQ ORBITRAP XL THERMO apparatus.
All compounds were tested as 95% purity or higher (by HRMS (ESI)).

6. BIBLIOGRAFY

(1) Davies, J.; Davies, D. Origins and Evolution of Antibiotic Resistance. *Microbiol. Mol. Biol. Rev. MMBR* **2010**, *74* (3), 417–433.

(2) Aminov, R. I. A Brief History of the Antibiotic Era: Lessons Learned and Challenges for the Future. *Front. Microbiol.* **2010**, *1*, 134.

(3) Spellberg, B.; Gilbert, D. N. The Future of Antibiotics and Resistance: A Tribute to a Career of Leadership by John Bartlett. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* **2014**, *59 Suppl 2*, S71-75.

(4) Hartman, F. C.; LaMuraglia, G. M.; Tomozawa, Y.; Wolfenden, R. The Influence of pH on the Interaction of Inhibitors with Triosephosphate Isomerase and Determination of the pKa of the Active-Site Carboxyl Group. *Biochemistry (Mosc.)* **1975**, *14* (24), 5274–5279.

(5) World Health Organization, WHO. Antimicrobial Resistance.

(6) Antibacterial and Antifungal Drug Discovery. *Nat. Biotechnol.* **2000**, *18 Suppl*, IT24-26.

(7) Lister, P. D.; Wolter, D. J.; Hanson, N. D. Antibacterial-Resistant Pseudomonas Aeruginosa: Clinical Impact and Complex Regulation of Chromosomally Encoded Resistance Mechanisms. *Clin. Microbiol. Rev.* **2009**, *22* (4), 582–610.

(8) Fernández, L.; Hancock, R. E. W. Adaptive and Mutational Resistance: Role of Porins and Efflux Pumps in Drug Resistance. *Clin. Microbiol. Rev.* **2012**, *25* (4), 661–681.

(9) European Antimicrobial Resistance Surveillance System (EARSS). RIVM (2005).

(10) Sekowska, A.; Kung, H. F.; Danchin, A. Sulfur Metabolism in Escherichia Coli and Related Bacteria: Facts and Fiction. *J. Mol. Microbiol. Biotechnol.* **2000**, *2* (2), 145–177.

(11) Campanini, B.; Pieroni, M.; Raboni, S.; Bettati, S.; Benoni, R.; Pecchini, C.; Costantino, G.; Mozzarelli, A. Inhibitors of the Sulfur Assimilation Pathway in Bacterial Pathogens as Enhancers of Antibiotic Therapy. *Curr. Med. Chem.* **2015**, *22* (2), 187–213.

(12) Kredich, N. M. Regulation of L-Cysteine Biosynthesis in Salmonella Typhimurium. I. Effects of Growth of Varying Sulfur Sources and O-Acetyl-L-Serine on Gene Expression. *J. Biol. Chem.* **1971**, *246* (11), 3474–3484.

(13) Hulanicka, M. D.; Hallquist, S. G.; Kredich, N. M.; Mojica-A, T. Regulation of O-Acetylserine Sulfhydrylase B by L-Cysteine in Salmonella Typhimurium. *J. Bacteriol.* **1979**, *140* (1), 141–146.

(14) Spyrakis, F.; Singh, R.; Cozzini, P.; Campanini, B.; Salsi, E.; Felici, P.; Raboni, S.; Benedetti, P.; Cruciani, G.; Kellogg, G. E.; Cook, P. F.; Mozzarelli, A. Isozyme-Specific Ligands for O-Acetylserine Sulfhydrylase, a Novel Antibiotic Target. *PloS One* **2013**, *8* (10), e77558.

(15) Bonner, E. R.; Cahoon, R. E.; Knapke, S. M.; Jez, J. M. Molecular Basis of Cysteine Biosynthesis in Plants: Structural and Functional Analysis of O-Acetylserine Sulfhydrylase from Arabidopsis Thaliana. *J. Biol. Chem.* **2005**, *280* (46), 38803–38813.

(16) Agarwal, S. M.; Jain, R.; Bhattacharya, A.; Azam, A. Inhibitors of Escherichia Coli Serine Acetyltransferase Block Proliferation of Entamoeba Histolytica Trophozoites. *Int. J. Parasitol.* **2008**, *38* (2), 137–141.

(17) Beinert, H. A Tribute to Sulfur. *Eur. J. Biochem.* **2000**, *267* (18), 5657–5664.

(18) Mihara, H.; Esaki, N. Bacterial Cysteine Desulfurases: Their Function and Mechanisms. *Appl. Microbiol. Biotechnol.* **2002**, *60* (1–2),

12–23.

(19) Kessler, D. Enzymatic Activation of Sulfur for Incorporation into Biomolecules in Prokaryotes. *FEMS Microbiol. Rev.* **2006**, *30* (6), 825–840.

(20) Liu, C.; Wang, R.; Varlamova, O.; Leyh, T. S. Regulating Energy Transfer in the ATP Sulfurylase-GTPase System. *Biochemistry (Mosc.)* **1998**, *37* (11), 3886–3892.

(21) Schelle, M. W.; Bertozzi, C. R. Sulfate Metabolism in Mycobacteria. *Chembiochem Eur. J. Chem. Biol.* **2006**, *7* (10), 1516–1524.

(22) Campanini, B.; Speroni, F.; Salsi, E.; Cook, P. F.; Roderick, S. L.; Huang, B.; Bettati, S.; Mozzarelli, A. Interaction of Serine Acetyltransferase with O-Acetylserine Sulfhydrylase Active Site: Evidence from Fluorescence Spectroscopy. *Protein Sci. Publ. Protein Soc.* **2005**, *14* (8), 2115–2124.

(23) Kredich, N. M. The Molecular Basis for Positive Regulation of Cys Promoters in Salmonella Typhimurium and Escherichia Coli. *Mol. Microbiol.* **1992**, *6* (19), 2747–2753.

(24) Salsi, E.; Campanini, B.; Bettati, S.; Raboni, S.; Roderick, S. L.; Cook, P. F.; Mozzarelli, A. A Two-Step Process Controls the Formation of the Bienzyme Cysteine Synthase Complex. *J. Biol. Chem.* **2010**, *285* (17), 12813–12822.

(25) Pieroni, M.; Annunziato, G.; Beato, C.; Wouters, R.; Benoni, R.; Campanini, B.; Pertinhez, T. A.; Bettati, S.; Mozzarelli, A.; Costantino, G. Rational Design, Synthesis, and Preliminary Structure-Activity Relationships of α-Substituted-2-Phenylcyclopropane Carboxylic Acids as Inhibitors of Salmonella Typhimurium O-Acetylserine Sulfhydrylase. *J. Med. Chem.* **2016**, *59* (6), 2567–2578.

(26) Turnbull, A. L.; Surette, M. G. Cysteine Biosynthesis, Oxidative Stress and Antibiotic Resistance in Salmonella Typhimurium. *Res. Microbiol.* **2010**, *161* (8), 643–650.

(27) Chattopadhyay, A.; Meier, M.; Ivaninskii, S.; Burkhard, P.; Speroni, F.; Campanini, B.; Bettati, S.; Mozzarelli, A.; Rabeh, W. M.; Li, L.; Cook, P. F. Structure, Mechanism, and Conformational Dynamics of O-Acetylserine Sulfhydrylase from Salmonella Typhimurium: Comparison of A and B Isozymes. *Biochemistry (Mosc.)* **2007**, *46* (28), 8315–8330.

(28) Turnbull, A. L.; Surette, M. G. L-Cysteine Is Required for Induced Antibiotic Resistance in Actively Swarming Salmonella Enterica Serovar Typhimurium. *Microbiol. Read. Engl.* **2008**, *154* (Pt 11), 3410–3419.

(29) Cook, P. F.; Wedding, R. T. Overall Mechanism and Rate Equation for O-Acetylserine Sulfhydrylase. *J. Biol. Chem.* **1977**, *252* (10), 3459.

(30) Cook, P. F.; Wedding, R. T. A Reaction Mechanism from Steady State Kinetic Studies for O-Acetylserine Sulfhydrylase from Salmonella Typhimurium LT-2. *J. Biol. Chem.* **1976**, *251* (7), 2023–2029.

(31) Bruno, A.; Amori, L.; Costantino, G. Computational Insights into the Mechanism of Inhibition of OASS-A by a Small Molecule Inhibitor. *Mol. Inform.* **2013**, *32* (5–6), 447–457.

(32) Amori, L.; Katkevica, S.; Bruno, A.; Campanini, B.; Felici, P.; Mozzarelli, A.; Costantino, G. Design and Synthesis of Trans-2-Substituted-Cyclopropane-1-Carboxylic Acids as the First Non-Natural Small Molecule Inhibitors of O-Acetylserine Sulfhydrylase. *MedChemComm* **2012**, *3* (9), 1111.

(33) Bray, C. D.; Minicone, F. Stereocontrolled Synthesis of Quaternary Cyclopropyl Esters. *Chem. Commun. Camb. Engl.* **2010**, *46*

(32), 5867–5869.

(34) Pieroni, M.; Annunziato, G.; Azzali, E.; Dessanti, P.; Mercurio, C.; Meroni, G.; Trifiró, P.; Vianello, P.; Villa, M.; Beato, C.; Varasi, M.; Costantino, G. Further Insights into the SAR of α-Substituted Cyclopropylamine Derivatives as Inhibitors of Histone Demethylase KDM1A. *Eur. J. Med. Chem.* **2015**, *92*, 377–386.

(35) Hackelöer, K.; Schnakenburg, G.; Waldvogel, S. R. Oxidative Coupling Reactions of 1,3-Diarylpropene Derivatives to Dibenzo[a,c]cycloheptenes by PIFA. *Eur. J. Org. Chem.* **2011**, *2011* (31), 6314–6319.

(36) Özer, G.; Saraçoğlu, N.; Balci, M. Synthesis and Chemistry of Unusual Bicyclic Endoperoxides Containing the Pyridazine Ring. *J. Org. Chem.* **2003**, *68* (18), 7009–7015.

(37) McCoy, L. L. Three-Membered Rings. The Preparation of Some 1,2-Cyclopropanedicarboxylic Acids. *J. Am. Chem. Soc.* **1958**, *80* (24), 6568–6572.

(38) Mayer, M.; Meyer, B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angew. Chem. Int. Ed.* **1999**, *38* (12), 1784–1788.

(39) Angulo, J.; Nieto, P. M. STD-NMR: Application to Transient Interactions between Biomolecules-a Quantitative Approach. *Eur. Biophys. J. EBJ* **2011**, *40* (12), 1357–1369.

(40) Mayer, M.; Meyer, B. Group Epitope Mapping by Saturation Transfer Difference NMR To Identify Segments of a Ligand in Direct Contact with a Protein Receptor. *J. Am. Chem. Soc.* **2001**, *123* (25), 6108–6117.

(41) Jayalakshmi, V.; Krishna, N. R. Complete Relaxation and Conformational Exchange Matrix (CORCEMA) Analysis of

Intermolecular Saturation Transfer Effects in Reversibly Forming Ligand-Receptor Complexes. *J. Magn. Reson. San Diego Calif 1997* **2002**, *155* (1), 106–118.

(42) Ginsberg, A. M. Emerging Drugs for Active Tuberculosis. *Semin. Respir. Crit. Care Med.* **2008**, *29* (5), 552–559.

(43) Cloeckaert, A.; Schwarz, S. Molecular Characterization, Spread and Evolution of Multidrug Resistance in Salmonella Enterica Typhimurium DT104. *Vet. Res.* **2001**, *32* (3–4), 301–310.

(44) Graham, D. Y.; Fischbach, L. Helicobacter Pylori Infection. *N. Engl. J. Med.* **2010**, *363* (6), 595–596; author reply 596.

(45) Smith, K. S.; Jakubzick, C.; Whittam, T. S.; Ferry, J. G. Carbonic Anhydrase Is an Ancient Enzyme Widespread in Prokaryotes. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (26), 15184–15189.

(46) Supuran, C. T. Bacterial Carbonic Anhydrases as Drug Targets: Toward Novel Antibiotics? *Front. Pharmacol.* **2011**, *2*, 34.

(47) McKenna, R. Carbonic Anhydrase: Its Inhibitors and Activators. CRC Enzyme Inhibitors Series, Volume 1 Edited by Claudiu T. Supuran, Andrea Scozzafava (Universita Degli Studi, Firenze), and Janet Conway (Pfizer Inc., New York). CRC Press LLC: Boca Raton, FL. 2004. Xii + 364 Pp. $139.95. ISBN 0-415-30673-6. *J. Am. Chem. Soc.* **2005**, *127* (10), 3643–3643.

(48) Scozzafava, A.; Mastrolorenzo, A.; Supuran, C. T. Carbonic Anhydrase Inhibitors and Activators and Their Use in Therapy. *Expert Opin. Ther. Pat.* **2006**, *16* (12), 1627–1664.

(49) Supuran, C. T.; Scozzafava, A.; Casini, A. Carbonic Anhydrase Inhibitors. *Med. Res. Rev.* **2003**, *23* (2), 146–189.

(50) Alterio, V.; Vitale, R. M.; Monti, S. M.; Pedone, C.; Scozzafava, A.; Cecchi, A.; De Simone, G.; Supuran, C. T. Carbonic Anhydrase

Inhibitors: X-Ray and Molecular Modeling Study for the Interaction of a Fluorescent Antitumor Sulfonamide with Isozyme II and IX. *J. Am. Chem. Soc.* **2006**, *128* (25), 8329–8335.

(51) Nishimori, I.; Minakuchi, T.; Onishi, S.; Vullo, D.; Cecchi, A.; Scozzafava, A.; Supuran, C. T. Carbonic Anhydrase Inhibitors: Cloning, Characterization, and Inhibition Studies of the Cytosolic Isozyme III with Sulfonamides. *Bioorg. Med. Chem.* **2007**, *15* (23), 7229–7236.

(52) Vullo, D.; Franchi, M.; Gallori, E.; Antel, J.; Scozzafava, A.; Supuran, C. T. Carbonic Anhydrase Inhibitors. Inhibition of Mitochondrial Isozyme V with Aromatic and Heterocyclic Sulfonamides. *J. Med. Chem.* **2004**, *47* (5), 1272–1279.

(53) Supuran, C. T. Carbonic Anhydrase Inhibitors and Activators for Novel Therapeutic Applications. *Future Med. Chem.* **2011**, *3* (9), 1165–1180.

(54) Pastorekova, S.; Parkkila, S.; Pastorek, J.; Supuran, C. T. Carbonic Anhydrases: Current State of the Art, Therapeutic Applications and Future Prospects. *J. Enzyme Inhib. Med. Chem.* **2004**, *19* (3), 199–229.

(55) Briganti, F.; Mangani, S.; Orioli, P.; Scozzafava, A.; Vernaglione, G.; Supuran, C. T. Carbonic Anhydrase Activators: X-Ray Crystallographic and Spectroscopic Investigations for the Interaction of Isozymes I and II with Histamine. *Biochemistry (Mosc.)* **1997**, *36* (34), 10384–10392.

(56) Maresca, A.; Temperini, C.; Vu, H.; Pham, N. B.; Poulsen, S.-A.; Scozzafava, A.; Quinn, R. J.; Supuran, C. T. Non-Zinc Mediated Inhibition of Carbonic Anhydrases: Coumarins Are a New Class of Suicide Inhibitors [#]. *J. Am. Chem. Soc.* **2009**, *131* (8), 3057–3062.

(57) Xu, Y.; Feng, L.; Jeffrey, P. D.; Shi, Y.; Morel, F. M. M. Structure

and Metal Exchange in the Cadmium Carbonic Anhydrase of Marine Diatoms. *Nature* **2008**, *452* (7183), 56–61.

(58) Touisni, N.; Maresca, A.; McDonald, P. C.; Lou, Y.; Scozzafava, A.; Dedhar, S.; Winum, J.-Y.; Supuran, C. T. Glycosyl Coumarin Carbonic Anhydrase IX and XII Inhibitors Strongly Attenuate the Growth of Primary Breast Tumors. *J. Med. Chem.* **2011**, *54* (24), 8271–8277.

(59) Woo, L. W. L.; Ganeshapillai, D.; Thomas, M. P.; Sutcliffe, O. B.; Malini, B.; Mahon, M. F.; Purohit, A.; Potter, B. V. L. Structure-Activity Relationship for the First-in-Class Clinical Steroid Sulfatase Inhibitor Irosustat (STX64, BN83495). *ChemMedChem* **2011**, *6* (11), 2019–2034.

(60) Weber, A.; Casini, A.; Heine, A.; Kuhn, D.; Supuran, C. T.; Scozzafava, A.; Klebe, G. Unexpected Nanomolar Inhibition of Carbonic Anhydrase by COX-2-Selective Celecoxib: New Pharmacological Opportunities due to Related Binding Site Recognition. *J. Med. Chem.* **2004**, *47* (3), 550–557.

(61) Köhler, K.; Hillebrecht, A.; Schulze Wischeler, J.; Innocenti, A.; Heine, A.; Supuran, C. T.; Klebe, G. Saccharin Inhibits Carbonic Anhydrases: Possible Explanation for Its Unpleasant Metallic Aftertaste. *Angew. Chem. Int. Ed.* **2007**, *46* (40), 7697–7699.

(62) Hilvo, M.; Innocenti, A.; Monti, S. M.; De Simone, G.; Supuran, C. T.; Parkkila, S. Recent Advances in Research on the Most Novel Carbonic Anhydrases, CA XIII and XV. *Curr. Pharm. Des.* **2008**, *14* (7), 672–678.

(63) Di Fiore, A.; Monti, S. M.; Hilvo, M.; Parkkila, S.; Romano, V.; Scaloni, A.; Pedone, C.; Scozzafava, A.; Supuran, C. T.; De Simone, G. Crystal Structure of Human Carbonic Anhydrase XIII and Its Complex with the Inhibitor Acetazolamide. *Proteins* **2009**, *74* (1), 164–175.

(64) Supuran, C. T. Carbonic Anhydrases: Novel Therapeutic

Applications for Inhibitors and Activators. *Nat. Rev. Drug Discov.* **2008**, *7* (2), 168–181.

(65) Whittington, D. A.; Waheed, A.; Ulmasov, B.; Shah, G. N.; Grubb, J. H.; Sly, W. S.; Christianson, D. W. Crystal Structure of the Dimeric Extracellular Domain of Human Carbonic Anhydrase XII, a Bitopic Membrane Protein Overexpressed in Certain Cancer Tumor Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (17), 9545–9550.

(66) Whittington, D. A.; Grubb, J. H.; Waheed, A.; Shah, G. N.; Sly, W. S.; Christianson, D. W. Expression, Assay, and Structure of the Extracellular Domain of Murine Carbonic Anhydrase XIV: Implications for Selective Inhibition of Membrane-Associated Isozymes. *J. Biol. Chem.* **2004**, *279* (8), 7223–7228.

(67) Ueda, K.; Nishida, H.; Beppu, T. Dispensabilities of Carbonic Anhydrase in Proteobacteria. *Int. J. Evol. Biol.* **2012**, *2012*, 1–5.

(68) Kawabata, T. Build-Up Algorithm for Atomic Correspondence between Chemical Structures. *J. Chem. Inf. Model.* **2011**, *51* (8), 1775–1787.

(69) Pieroni, M.; Wan, B.; Cho, S.; Franzblau, S. G.; Costantino, G. Design, Synthesis and Investigation on the Structure-Activity Relationships of N-Substituted 2-Aminothiazole Derivatives as Antitubercular Agents. *Eur. J. Med. Chem.* **2014**, *72*, 26–34.

(70) Vincetti, P.; Caporuscio, F.; Kaptein, S.; Gioiello, A.; Mancino, V.; Suzuki, Y.; Yamamoto, N.; Crespan, E.; Lossani, A.; Maga, G.; Rastelli, G.; Castagnolo, D.; Neyts, J.; Leyssen, P.; Costantino, G.; Radi, M. Discovery of Multitarget Antivirals Acting on Both the Dengue Virus NS5-NS3 Interaction and the Host Src/Fyn Kinases. *J. Med. Chem.* **2015**, *58* (12), 4964–4975.

(71) Pieroni, M.; Machado, D.; Azzali, E.; Santos Costa, S.; Couto, I.;

Costantino, G.; Viveiros, M. Rational Design and Synthesis of Thioridazine Analogues as Enhancers of the Antituberculosis Therapy. *J. Med. Chem.* **2015**, *58* (15), 5842–5853.

(72) Vallerini, G. P.; Amori, L.; Beato, C.; Tararina, M.; Wang, X.-D.; Schwarcz, R.; Costantino, G. 2-Aminonicotinic Acid 1-Oxides Are Chemically Stable Inhibitors of Quinolinic Acid Synthesis in the Mammalian Brain: A Step toward New Antiexcitotoxic Agents. *J. Med. Chem.* **2013**, *56* (23), 9482–9495.

(73) Alterio, V.; Di Fiore, A.; D'Ambrosio, K.; Supuran, C. T.; De Simone, G. Multiple Binding Modes of Inhibitors to Carbonic Anhydrases: How to Design Specific Drugs Targeting 15 Different Isoforms? *Chem. Rev.* **2012**, *112* (8), 4421–4468.

(74) Martin, D. P.; Cohen, S. M. Nucleophile Recognition as an Alternative Inhibition Mode for Benzoic Acid Based Carbonic Anhydrase Inhibitors. *Chem. Commun.* **2012**, *48* (43), 5259.

(75) Martin, D. P.; Blachly, P. G.; McCammon, J. A.; Cohen, S. M. Exploring the Influence of the Protein Environment on Metal-Binding Pharmacophores. *J. Med. Chem.* **2014**, *57* (16), 7126–7135.

(76) Schulze Wischeler, J.; Innocenti, A.; Vullo, D.; Agrawal, A.; Cohen, S. M.; Heine, A.; Supuran, C. T.; Klebe, G. Bidentate Zinc Chelators for α-Carbonic Anhydrases That Produce a Trigonal Bipyramidal Coordination Geometry. *ChemMedChem* **2010**, *5* (9), 1609–1615.

(77) Vu, H.; Pham, N. B.; Quinn, R. J. Direct Screening of Natural Product Extracts Using Mass Spectrometry. *J. Biomol. Screen.* **2008**, *13* (4), 265–275.

(78) Carta, F.; Aggarwal, M.; Maresca, A.; Scozzafava, A.; McKenna, R.; Masini, E.; Supuran, C. T. Dithiocarbamates Strongly Inhibit

Carbonic Anhydrases and Show Antiglaucoma Action in Vivo. *J. Med. Chem.* **2012**, *55* (4), 1721–1730.

(79) Annunziato, G.; Angeli, A.; D'Alba, F.; Bruno, A.; Pieroni, M.; Vullo, D.; De Luca, V.; Capasso, C.; Supuran, C. T.; Costantino, G. Discovery of New Potential Anti-Infective Compounds Based on Carbonic Anhydrase Inhibitors by Rational Target-Focused Repurposing Approaches. *ChemMedChem* **2016**, *11* (17), 1904–1914.

(80) Chaiyaveij, D.; Batsanov, A. S.; Fox, M. A.; Marder, T. B.; Whiting, A. An Experimental and Computational Approach to Understanding the Reactions of Acyl Nitroso Compounds in [4 + 2] Cycloadditions. *J. Org. Chem.* **2015**, *80* (19), 9518–9534.

(81) Cho, S. J.; Jensen, N. H.; Kurome, T.; Kadari, S.; Manzano, M. L.; Malberg, J. E.; Caldarone, B.; Roth, B. L.; Kozikowski, A. P. Selective 5-Hydroxytryptamine 2C Receptor Agonists Derived from the Lead Compound Tranylcypromine: Identification of Drugs with Antidepressant-like Action. *J. Med. Chem.* **2009**, *52* (7), 1885–1902.

(82) Krungkrai, S. R.; Krungkrai, J. Malaria Parasite Carbonic Anhydrase: Inhibition of Aromatic/Heterocyclic Sulfonamides and Its Therapeutic Potential. *Asian Pac. J. Trop. Biomed.* **2011**, *1* (3), 233–242.

(83) Järvisalo, J.; Saris, N. E. Action of Propranolol on Mitochondrial Functions--Effects on Energized Ion Fluxes in the Presence of Valinomycin. *Biochem. Pharmacol.* **1975**, *24* (18), 1701–1705.

(84) One-Pot Synthesis of 2-Amino-3-Cyanopyridine Derivatives under Microwave Irradiation without Solvent. *Arkivoc* **2005**, *2005* (1), 137.

(85) Ashburn, T. T.; Thor, K. B. Drug Repositioning: Identifying and Developing New Uses for Existing Drugs. *Nat. Rev. Drug Discov.* **2004**, *3* (8), 673–683.

(86) Oprea, T. I.; Bauman, J. E.; Bologa, C. G.; Buranda, T.; Chigaev, A.; Edwards, B. S.; Jarvik, J. W.; Gresham, H. D.; Haynes, M. K.; Hjelle, B.; Hromas, R.; Hudson, L.; Mackenzie, D. A.; Muller, C. Y.; Reed, J. C.; Simons, P. C.; Smagley, Y.; Strouse, J.; Surviladze, Z.; Thompson, T.; Ursu, O.; Waller, A.; Wandinger-Ness, A.; Winter, S. S.; Wu, Y.; Young, S. M.; Larson, R. S.; Willman, C.; Sklar, L. A. Drug Repurposing from an Academic Perspective. *Drug Discov. Today Ther. Strateg.* **2011**, *8* (3–4), 61–69.

(87) Kawabata, T. Build-Up Algorithm for Atomic Correspondence between Chemical Structures. *J. Chem. Inf. Model.* **2011**, *51* (8), 1775–1787.

(88) Smanski, M. J.; Zhou, H.; Claesen, J.; Shen, B.; Fischbach, M. A.; Voigt, C. A. Synthetic Biology to Access and Expand Nature's Chemical Diversity. *Nat. Rev. Microbiol.* **2016**, *14* (3), 135–149.

(89) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs over the 30 Years from 1981 to 2010. *J. Nat. Prod.* **2012**, *75* (3), 311–335.

(90) Demain, A. L. Importance of Microbial Natural Products and the Need to Revitalize Their Discovery. *J. Ind. Microbiol. Biotechnol.* **2014**, *41* (2), 185–201.

(91) Lahlou, M. The Success of Natural Products in Drug Discovery. *Pharmacol. Amp Pharm.* **2013**, *4* (3), 17–31.

(92) Cragg, G. M.; Newman, D. J. Natural Products: A Continuing Source of Novel Drug Leads. *Biochim. Biophys. Acta* **2013**, *1830* (6), 3670–3695.

(93) Anastasio, G. D.; Robinson, M. D.; Little, J. M.; Leitch, B. B.; Pettice, Y. L.; Norton, H. J. A Comparison of the Gastrointestinal Side Effects of Two Forms of Erythromycin. *J. Fam. Pract.* **1992**, *35* (5), 517–

523.

(94) Nicolaou, K. C.; Yang, Z.; Liu, J. J.; Ueno, H.; Nantermet, P. G.; Guy, R. K.; Claiborne, C. F.; Renaud, J.; Couladouros, E. A.; Paulvannan, K.; Sorensen, E. J. Total Synthesis of Taxol. *Nature* **1994**, *367* (6464), 630–634.

(95) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand Efficiency: A Useful Metric for Lead Selection. *Drug Discov. Today* **2004**, *9* (10), 430–431.

(96) Leonti, M.; Casu, L. Traditional Medicines and Globalization: Current and Future Perspectives in Ethnopharmacology. *Front. Pharmacol.* **2013**, *4*.

(97) Dias, D. A.; Urban, S.; Roessner, U. A Historical Overview of Natural Products in Drug Discovery. *Metabolites* **2012**, *2* (4), 303–336.

(98) Stadler, M.; Hoffmeister, D. Fungal Natural Productsâ€"the Mushroom Perspective. *Front. Microbiol.* **2015**, *6*.

(99) Dias, D. A.; Urban, S.; Roessner, U. A Historical Overview of Natural Products in Drug Discovery. *Metabolites* **2012**, *2* (4), 303–336.

(100) Abraham, E. P.; Chain, E.; Fletcher, C. M.; Florey, H. W.; Gardner, A. D.; Heatley, N. G.; Jennings, M. A. Further Observations on Penicillin. 1941. *Eur. J. Clin. Pharmacol.* **1992**, *42* (1), 3–9.

(101) Gallant, C. V.; Daniels, C.; Leung, J. M.; Ghosh, A. S.; Young, K. D.; Kotra, L. P.; Burrows, L. L. Common β-Lactamases Inhibit Bacterial Biofilm Formation: β-Lactamases Impair Biofilm Formation. *Mol. Microbiol.* **2005**, *58* (4), 1012–1024.

(102) Fabbretti, A.; Gualerzi, C. O.; Brandi, L. How to Cope with the Quest for New Antibiotics. *FEBS Lett.* **2011**, *585* (11), 1673–1681.

(103) Zjawiony, J. K. Biologically Active Compounds from Aphyllophorales (Polypore) Fungi. *J. Nat. Prod.* **2004**, *67* (2), 300–310.

(104) Cragg, G. M.; Newman, D. J. Biodiversity: A Continuing Source of

Novel Drug Leads. *Pure Appl. Chem.* **2005**, *77* (1).

(105) Butler, M. S. The Role of Natural Product Chemistry in Drug Discovery. *J. Nat. Prod.* **2004**, *67* (12), 2141–2153.

(106) Singleton, V. L.; Bohonos, N.; Ullstrup, A. J. Decumbin, a New Compound from a Species of Penicillium. *Nature* **1958**, *181* (4615), 1072–1073.

(107) Baudouy, R.; Crabbé, P.; Greene, A. E.; Le Drian, C.; Orr, A. F. A Synthesis of Brefeldin A. *Tetrahedron Lett.* **1977**, *18* (34), 2973–2976.

(108) Seehafer, K.; Rominger, F.; Helmchen, G.; Langhans, M.; Robinson, D. G.; Özata, B.; Brügger, B.; Strating, J. R. P. M.; van Kuppeveld, F. J. M.; Klein, C. D. Synthesis and Biological Properties of Novel Brefeldin A Analogues. *J. Med. Chem.* **2013**, *56* (14), 5872–5884.

(109) Tamura, G.; Ando, K.; Suzuki, S.; Takatsuki, A.; Arima, K. Antiviral Activity of Brefeldin A and Verrucarin A. *J. Antibiot. (Tokyo)* **1968**, *21* (2), 160–161.

(110) Cagliero, C.; Zhou, Y. N.; Jin, D. J. Spatial Organization of Transcription Machinery and Its Segregation from the Replisome in Fast-Growing Bacterial Cells. *Nucleic Acids Res.* **2014**, *42* (22), 13696–13705.

(111) Grose, C.; Klionsky, D. J. Alternative Autophagy, Brefeldin A and Viral Trafficking Pathways. *Autophagy* **2016**, *12* (9), 1429–1430.

(112) Corey, E. J.; Nicolaou, K. C.; Melvin, L. S. Synthesis of Brefeldin A, Carpaine, Vertaline, and Erythronolide B from Nonmacrocyclic Precursors. *J. Am. Chem. Soc.* **1975**, *97* (3), 654–655.

(113) Greene, A. E.; Le Drian, C.; Crabbe, P. An Efficient Total Synthesis of (.+-.)-Brefeldin-A. *J. Am. Chem. Soc.* **1980**, *102* (25), 7583–7584.

(114) Bartlett, P. A.; Green, F. R. Total Synthesis of Brefeldin A. *J. Am.*

Chem. Soc. **1978**, *100* (15), 4858–4865.

(115) Nakamura, null; Watanabe, null; Toru, null. Extremely Efficient Chiral Induction in Conjugate Additions of P-Tolyl Alpha-Lithio-Beta-(Trimethylsilyl)ethyl Sulfoxide and Subsequent Electrophilic Trapping Reactions. *J. Org. Chem.* **2000**, *65* (6), 1758–1766.

(116) Kusuda, S.; Ueno, Y.; Toru, T. Vinyl Anion Equivalent V. Asymmetric Synthesis of Allylic Alcohols Using Chiral 2-(Trialkylsilyl)ethyl Sulfoxides. *Tetrahedron* **1994**, *50* (4), 1045–1062.

(117) Edwards, A. M.; Isserlin, R.; Bader, G. D.; Frye, S. V.; Willson, T. M.; Yu, F. H. Too Many Roads Not Taken. *Nature* **2011**, *470* (7333), 163–165.

(118) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science* **1996**, *274* (5292), 1531–1534.

(119) Mayr, L. M.; Bojanic, D. Novel Trends in High-Throughput Screening. *Curr. Opin. Pharmacol.* **2009**, *9* (5), 580–588.

(120) Wigglesworth, M. J.; Murray, D. C.; Blackett, C. J.; Kossenjans, M.; Nissink, J. W. M. Increasing the Delivery of next Generation Therapeutics from High Throughput Screening Libraries. *Curr. Opin. Chem. Biol.* **2015**, *26*, 104–110.

(121) Macarron, R.; Banks, M. N.; Bojanic, D.; Burns, D. J.; Cirovic, D. A.; Garyantes, T.; Green, D. V. S.; Hertzberg, R. P.; Janzen, W. P.; Paslay, J. W.; Schopfer, U.; Sittampalam, G. S. Impact of High-Throughput Screening in Biomedical Research. *Nat. Rev. Drug Discov.* **2011**, *10* (3), 188–195.

(122) Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R. Fragment-Based Lead Discovery. *Nat. Rev. Drug Discov.* **2004**, *3* (8), 660–672.

(123) Jorgensen, W. L. The Many Roles of Computation in Drug

Discovery. *Science* **2004**, *303* (5665), 1813–1818.

(124) Reker, D.; Perna, A. M.; Rodrigues, T.; Schneider, P.; Reutlinger, M.; Mönch, B.; Koeberle, A.; Lamers, C.; Gabler, M.; Steinmetz, H.; Müller, R.; Schubert-Zsilavecz, M.; Werz, O.; Schneider, G. Revealing the Macromolecular Targets of Complex Natural Products. *Nat. Chem.* **2014**, *6* (12), 1072–1078.

(125) Reker, D.; Schneider, G. Active-Learning Strategies in Computer-Assisted Drug Discovery. *Drug Discov. Today* **2015**, *20* (4), 458–465.

(126) Smith, G. P. Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science* **1985**, *228* (4705), 1315–1317.

(127) McCafferty, J.; Griffiths, A. D.; Winter, G.; Chiswell, D. J. Phage Antibodies: Filamentous Phage Displaying Antibody Variable Domains. *Nature* **1990**, *348* (6301), 552–554.

(128) Winter, G.; Griffiths, A. D.; Hawkins, R. E.; Hoogenboom, H. R. Making Antibodies by Phage Display Technology. *Annu. Rev. Immunol.* **1994**, *12*, 433–455.

(129) Mattheakis, L. C.; Bhatt, R. R.; Dower, W. J. An in Vitro Polysome Display System for Identifying Ligands from Very Large Peptide Libraries. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91* (19), 9022–9026.

(130) Heinis, C.; Rutherford, T.; Freund, S.; Winter, G. Phage-Encoded Combinatorial Chemical Libraries Based on Bicyclic Peptides. *Nat. Chem. Biol.* **2009**, *5* (7), 502–507.

(131) Roberts, R. W. Totally in Vitro Protein Selection Using mRNA-Protein Fusions and Ribosome Display. *Curr. Opin. Chem. Biol.* **1999**, *3* (3), 268–273.

(132) Brenner, S.; Lerner, R. A. Encoded Combinatorial Chemistry. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (12), 5381–5383.

(133) Melkko, S.; Dumelin, C. E.; Scheuermann, J.; Neri, D. Lead Discovery by DNA-Encoded Chemical Libraries. *Drug Discov. Today* **2007**, *12* (11–12), 465–471.

(134) Scheuermann, J.; Neri, D. DNA-Encoded Chemical Libraries: A Tool for Drug Discovery and for Chemical Biology. *ChemBioChem* **2010**, *11* (7), 931–937.

(135) Kleiner, R. E.; Dumelin, C. E.; Liu, D. R. Small-Molecule Discovery from DNA-Encoded Chemical Libraries. *Chem. Soc. Rev.* **2011**, *40* (12), 5707–5717.

(136) Deng, H.; Grunder, S.; Cordova, K. E.; Valente, C.; Furukawa, H.; Hmadeh, M.; Gándara, F.; Whalley, A. C.; Liu, Z.; Asahina, S.; Kazumori, H.; O'Keeffe, M.; Terasaki, O.; Stoddart, J. F.; Yaghi, O. M. Large-Pore Apertures in a Series of Metal-Organic Frameworks. *Science* **2012**, *336* (6084), 1018–1023.

(137) Franzini, R. M.; Ekblad, T.; Zhong, N.; Wichert, M.; Decurtins, W.; Nauer, A.; Zimmermann, M.; Samain, F.; Scheuermann, J.; Brown, P. J.; Hall, J.; Gräslund, S.; Schüler, H.; Neri, D. Identification of Structure-Activity Relationships from Screening a Structurally Compact DNA-Encoded Chemical Library. *Angew. Chem. Int. Ed Engl.* **2015**, *54* (13), 3927–3931.

(138) Seigal, B. A.; Connors, W. H.; Fraley, A.; Borzilleri, R. M.; Carter, P. H.; Emanuel, S. L.; Fargnoli, J.; Kim, K.; Lei, M.; Naglich, J. G.; Pokross, M. E.; Posy, S. L.; Shen, H.; Surti, N.; Talbott, R.; Zhang, Y.; Terrett, N. K. The Discovery of Macrocyclic XIAP Antagonists from a DNA-Programmed Chemistry Library, and Their Optimization to Give Lead Compounds with in Vivo Antitumor Activity. *J. Med. Chem.* **2015**, *58* (6), 2855–2861.

(139) Kollmann, C. S.; Bai, X.; Tsai, C.-H.; Yang, H.; Lind, K. E.;

Skinner, S. R.; Zhu, Z.; Israel, D. I.; Cuozzo, J. W.; Morgan, B. A.; Yuki, K.; Xie, C.; Springer, T. A.; Shimaoka, M.; Evindar, G. Application of Encoded Library Technology (ELT) to a Protein-Protein Interaction Target: Discovery of a Potent Class of Integrin Lymphocyte Function-Associated Antigen 1 (LFA-1) Antagonists. *Bioorg. Med. Chem.* **2014**, *22* (7), 2353–2365.

(140) Franzini, R. M.; Samain, F.; Abd Elrahman, M.; Mikutis, G.; Nauer, A.; Zimmermann, M.; Scheuermann, J.; Hall, J.; Neri, D. Systematic Evaluation and Optimization of Modification Reactions of Oligonucleotides with Amines and Carboxylic Acids for the Synthesis of DNA-Encoded Chemical Libraries. *Bioconjug. Chem.* **2014**, *25* (8), 1453–1461.

(141) Litovchick, A.; Dumelin, C. E.; Habeshian, S.; Gikunju, D.; Guié, M.-A.; Centrella, P.; Zhang, Y.; Sigel, E. A.; Cuozzo, J. W.; Keefe, A. D.; Clark, M. A. Encoded Library Synthesis Using Chemical Ligation and the Discovery of sEH Inhibitors from a 334-Million Member Library. *Sci. Rep.* **2015**, *5*, 10916.

(142) Franzini, R. M.; Neri, D.; Scheuermann, J. DNA-Encoded Chemical Libraries: Advancing beyond Conventional Small-Molecule Libraries. *Acc. Chem. Res.* **2014**, *47* (4), 1247–1255.

(143) Satz, A. L.; Cai, J.; Chen, Y.; Goodnow, R.; Gruber, F.; Kowalczyk, A.; Petersen, A.; Naderi-Oboodi, G.; Orzechowski, L.; Strebel, Q. DNA Compatible Multistep Synthesis and Applications to DNA Encoded Libraries. *Bioconjug. Chem.* **2015**, *26* (8), 1623–1632.

(144) Gartner, Z. J. DNA-Templated Organic Synthesis and Selection of a Library of Macrocycles. *Science* **2004**, *305* (5690), 1601–1605.

(145) Mannocci, L.; Leimbacher, M.; Wichert, M.; Scheuermann, J.; Neri, D. 20 Years of DNA-Encoded Chemical Libraries. *Chem. Commun.*

2011, *47* (48), 12747.

(146) Wrenn, S. J.; Weisinger, R. M.; Halpin, D. R.; Harbury, P. B. Synthetic Ligands Discovered by in Vitro Selection. *J. Am. Chem. Soc.* **2007**, *129* (43), 13137–13143.

(147) Hansen, M. E.; Bentin, T.; Nielsen, P. E. High-Affinity Triplex Targeting of Double Stranded DNA Using Chemically Modified Peptide Nucleic Acid Oligomers. *Nucleic Acids Res.* **2009**, *37* (13), 4498–4507.

(148) Melkko, S.; Scheuermann, J.; Dumelin, C. E.; Neri, D. Encoded Self-Assembling Chemical Libraries. *Nat. Biotechnol.* **2004**, *22* (5), 568–574.

(149) Melkko, S.; Dumelin, C. E.; Scheuermann, J.; Neri, D. Lead Discovery by DNA-Encoded Chemical Libraries. *Drug Discov. Today* **2007**, *12* (11–12), 465–471.

(150) Scheuermann, J.; Dumelin, C. E.; Melkko, S.; Neri, D. DNA-Encoded Chemical Libraries. *J. Biotechnol.* **2006**, *126* (4), 568–581.

(151) Mannocci, L.; Zhang, Y.; Scheuermann, J.; Leimbacher, M.; De Bellis, G.; Rizzi, E.; Dumelin, C.; Melkko, S.; Neri, D. High-Throughput Sequencing Allows the Identification of Binding Molecules Isolated from DNA-Encoded Chemical Libraries. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (46), 17670–17675.

(152) Buller, F.; Zhang, Y.; Scheuermann, J.; Schäfer, J.; Bühlmann, P.; Neri, D. Discovery of TNF Inhibitors from a DNA-Encoded Chemical Library Based on Diels-Alder Cycloaddition. *Chem. Biol.* **2009**, *16* (10), 1075–1086.

(153) Decurtins, W.; Wichert, M.; Franzini, R. M.; Buller, F.; Stravs, M. A.; Zhang, Y.; Neri, D.; Scheuermann, J. Automated Screening for Small Organic Ligands Using DNA-Encoded Chemical Libraries. *Nat. Protoc.* **2016**, *11* (4), 764–780.

(154) Schuster, S. C. Next-Generation Sequencing Transforms Today's Biology. *Nat. Methods* **2008**, *5* (1), 16–18.

(155) Myllykangas, S.; Buenrostro, J. D.; Natsoulis, G.; Bell, J. M.; Ji, H. P. Efficient Targeted Resequencing of Human Germline and Cancer Genomes by Oligonucleotide-Selective Sequencing. *Nat. Biotechnol.* **2011**, *29* (11), 1024–1027.

(156) Melkko, S.; Mannocci, L.; Dumelin, C. E.; Villa, A.; Sommavilla, R.; Zhang, Y.; Grütter, M. G.; Keller, N.; Jermutus, L.; Jackson, R. H.; Scheuermann, J.; Neri, D. Isolation of a Small-Molecule Inhibitor of the Antiapoptotic Protein Bcl-xL from a DNA-Encoded Chemical Library. *ChemMedChem* **2010**, *5* (4), 584–590.

(157) Buller, F.; Mannocci, L.; Scheuermann, J.; Neri, D. Drug Discovery with DNA-Encoded Chemical Libraries. *Bioconjug. Chem.* **2010**, *21* (9), 1571–1580.