UNIVERSITÀ DEGLI STUDI DI MILANO

DIPARTIMENTO DI
SCIENZE FARMACEUTICHE

IRON ACQUISITION AS TARGET FOR NEW ANTITUBERCULAR DRUGS

Matteo Mori

R11797

Supervisor: Prof. Fiorella Meneghetti

Coordinator: Prof. Giancarlo Aldini

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy in Pharmaceutical Sciences (CHIM/08)

Doctoral School of Pharmaceutical Sciences (XXXII cycle)

Department of Pharmaceutical Sciences

University of Milan

Academic year 2018-2019

"The true delight is in the finding out, rather than in the knowing"

Isaac Asimov

Keywords

Abstract

Tuberculosis nowadays ranks among the leading causes of death worldwide. The growing emergence of resistant strains of *Mycobacterium tuberculosis* poses a serious threat to the public's health. Moreover, the inadequacy of the available drugs and the difficult treatment regimens contribute to the spread of the disease. Therefore, the discovery of new antitubercular drugs, acting on new pharmacological targets, is assuming critical importance.

The iron uptake pathways have been identified as a potential option, considering the essentiality of this metal ion for the survival and virulence of the microorganism. In particular, the inhibition of the mycobacterium-specific salicylate synthase MbtI has been recently validated as a new promising approach for the development of novel antitubercular agents.

In this context, the current research project was focused on the identification of new, potent MbtI inhibitors for the development of innovative therapeutic agents. Our investigations started with a structure-based virtual screening, employing the available co-crystal structures of MbtI: we built a pharmacophore model and we used it to screen the Enamine database. Subsequent steps of consensus docking and molecular dynamics allowed us to identify a competitive furan-based inhibitor, endowed with a promising activity on the enzyme. With the aim of understanding the key structural features required for the inhibitory effect, we modified the hit compound to derive preliminary information. Then, we performed an in-depth exploration of the chemical space around this scaffold through a structure-activity relationship study; this investigation led to the identification of the most potent MbtI inhibitors available to date. All the compounds were submitted to a complete biological evaluation. Considering the difficulty in sourcing the natural substrate of the enzyme for the biological analyses, we decided to produce, purify and characterise it, using an in-house optimised process. Apart from an excellent inhibitory activity on the target enzyme, our candidate inhibitors also exhibited an interesting antimycobacterial action, related to the reduced production of siderophores by the pathogen. Notably, we managed to obtain, for the first time,

a solid correlation between an effective inhibition of MbtI and a promising antimycobacterial effect in whole-cell assays.

Parallelly, we submitted the best compounds to co-crystallisation experiments to empirically define the binding mode of this class of inhibitors within the active site of MbtI. The obtainment of high-quality crystals of the protein proved particularly challenging; therefore, we dedicated our efforts to the optimisation of the whole protein production and crystallisation process. These studies allowed us to achieve valuable results, thanks to the obtainment of good crystal samples. The analysis of the structural data allowed to uncover the key protein-ligand interactions at a molecular level, paving the way for the optimisation of the computational model, and for the design and synthesis of improved candidates. Moreover, thanks to these crystallographic investigations, we managed to gather unprecedented experimental evidences regarding the catalytic activity of MbtI.

Finally, we assessed the possibility to design liposome-based nanocarriers for antitubercular inhalation therapy. The promising results obtained from our preliminary experiments, performed on liposomes loaded with one or two first line drugs, laid the foundations for the development of novel delivery systems for the administration of combinations of new and established antitubercular agents.

This multidisciplinary project was developed thanks to the cooperation of scientists from the University of Pavia (Italy), the University of Pisa (Italy), the University of Modena and Reggio Emilia (Italy) and the Institut Pasteur (Paris, France).

Table of Contents

List of Figures

List of Tables

List of Abbreviations

API: Active Pharmaceutical Ingredient
BCG: Bacille Calmette-Guérin
CAS: Chrome Azurol S
CHA: Chorismic acid
Chol: Cholesterol
CUE: Chorismate Utilising Enzyme
DIPEA: *N,N*-Diisopropylethylamine
DLS: Dynamic Light Scattering
DMF: *N,N*-Dimethylformamide
DMSO: Dimethyl sulfoxide
EMB: Ethambutol
ESI: Electrospray Ionisation
ESRF: European Synchrotron Radiation Facility
EtOAc: Ethyl acetate
EtOH: Ethanol
FT-IR: Fourier-Transform Infrared Spectroscopy
GAFF: General Amber Forcefield
HATU: 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide hexafluorophosphate
HPLC: High-Performance Liquid Chromatography
HRMS: High-Resolution Mass Spectrometry
HTS: Hight-Throughput Screening
IC_{50}: Half maximal Inhibitory Concentration
INH: Isoniazid
MAD: Multi-wavelength Anomalous Diffraction
Maf: *Mycobacterium africanum*
MD: Molecular Dynamics
MDR-TB: Multi-drug-resistant tuberculosis
MeOH: Methanol
MES: 2-(*N*-morpholino)ethanesulfonic acid
MIC: Minimum Inhibitory Concentration
MPD: 2-Methyl-2,4-pentanediol
MR: Molecular Replacement
MS: Mass Spectrometry
Mtb: *Mycobacterium tuberculosis*
MTC or MTBC: *Mycobacterium tuberculosis* complex
NanoDSF: Nano-Differential Scanning Fluorimetry
NBS: *N*-Bromosuccinimide
NMR: Nuclear Magnetic Resonance
OD_{600}: Optical Density at 600 nm
PAINS: Pan-Assay Interference Compounds
PC: Phosphatidylcholine
PCS: Photon Correlation Spectroscopy
PDB: Protein Data Bank
PDI: Polydispersity index

PEG: Polyethylene glycol
PEG MME: Polyethylene glycol monomethyl ether
p-TsOH: *p*-Toluenesulfonic acid
PZA: Pyrazinamide
RBT: Rifabutine
REMA: Resazurin Microtiter Assay
REV: Reverse Phase Evaporation
RIF: Rifampicin
RPT: Rifapentine
RR-TB: Rifampicin-resistant tuberculosis
SANS: Small-angle neutron scattering
SLF: Simulated Lung Fluid
SM: Streptomycin
SOLEIL: Source optimisée de lumière d'énergie intermédiaire du LURE
(Laboratoire pour l'utilisation du rayonnement électromagnétique)
TB: Tuberculosis
TEA: Triethylamine
THF: Tetrahydrofuran
TLC: Thin-Layer Chromatography
TMS: Tetramethylsilane
VS: Virtual Screening
WHO: World Health Organisation
XDR-TB: Extensively drug-resistant tuberculosis

Statement of Original Authorship

The work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

Signature:

Date: 09/12/2019

Chapter 1: Introduction

1.1 TUBERCULOSIS

Tuberculosis (TB) is one of the most ancient diseases known to mankind; in its different types and forms, it has affected humans since prehistoric times, causing outbreaks and epidemics which have caused billions of victims over the centuries. *Mycobacterium tuberculosis* (Mtb, Figure 1), the etiological agent of TB, was identified in 1882 by the Nobel Prize winner Robert Koch; since then, many efforts have been devoted to the discovery of antitubercular agents, although the peculiar characteristics of the microorganism have significantly hampered the identification of effective compounds.[1]

Nowadays, TB remains the first determinant of morbidity and mortality from a single infectious agent, still ranking among the leading causes of death worldwide. Despite the undeniable advances in control and prevention, achieved in the last years as a result of the implementation of the measures foreseen in the End TB Strategy (World Health Organisation, 2016-2035), tuberculosis has provoked 1.5 million deaths just in 2018. Most of the 7 million new cases were reported in South-East Asia (44%), in the African region (24%) and in the Western Pacific area (18%). However, the number of estimated cases reached 10 million in the same year.[2]
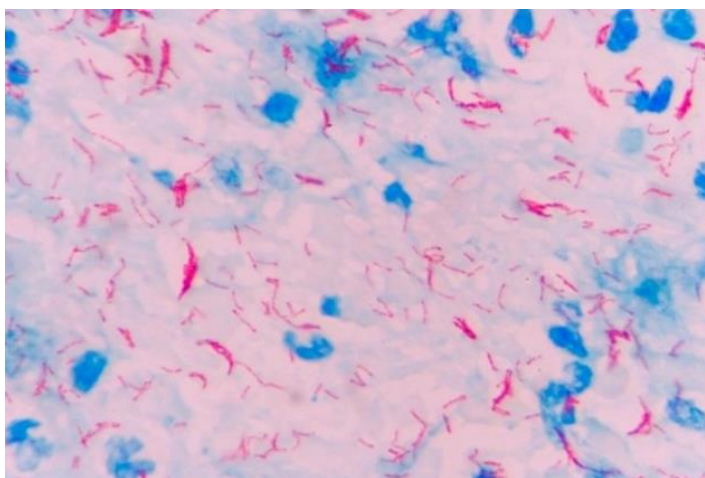
Figure 1. Bright-field microscope image of *Mycobacterium tuberculosis* visualised with the Ziehl–Neelsen stain as red, elongated rods.[3]

Together with *Mycobacterium africanum* (Maf), *Mycobacterium bovis*, *Mycobacterium microti* and *Mycobacterium canetti*, Mtb is part of the so-called *Mycobacterium tuberculosis* complex (MTC or MTBC). Although all of them can cause TB, Mtb and Maf have a specific tropism for humans. Each species shows distinctive phenotypic, genomic, clinical and epidemiological features; anyway, Mtb is by far the most common aetiological agent of TB.[4] It is an obligate intracellular pathogen, aerobic, non-motile, non-encapsulated, acid-fast and non-spore forming. It grows in tissues with high oxygen content, such as the lungs, but it can also endure in hypoxic environments in its quiescent form. The bacilli can be very different in size and shape; typically, they appear as long rods or coccobacilli, depending on the growth conditions and the age of the culture. Their dimensions have been reported to be 1-10 µm in length (usually 3-5 µm) and 0.2-0.6 µm in width. Mtb is considered neither a Gram-positive nor a Gram-negative bacterium, since it is relatively impermeable to basic dyes, unless combined with phenol. Mtb has an extremely slow rate of division compared to other bacteria; it replicates every 15-20 hours, while *Escherichia coli* divides every 20 minutes.[5,6] Another peculiarity of *Mycobacterium tuberculosis* is its cell wall (Figure 2), which is unique among prokaryotes. It is characterised by the presence of strongly hydrophobic molecules called mycolic acids, which form a lipid envelope around the microorganism, affecting the permeability of the cell wall. Hence, mycolic acids play an essential role in the defence against both chemical and enzymatic insults, including those deriving from phagocytic granules and antibiotics. Structurally, the cell wall is composed of an inner cytoplasmic membrane, characterized by a phospholipid bilayer, which contains mainly anionic phospholipids with proteins. Proceeding outward, an intermediate space contains peptidoglycan, covalently linked to arabinogalactan and lipoarabinomannan; finally, the outer membrane consists of mycolic acids in the inner leaflet, and free intercalating glycolipids and waxy components in the outer leaflet.[7]
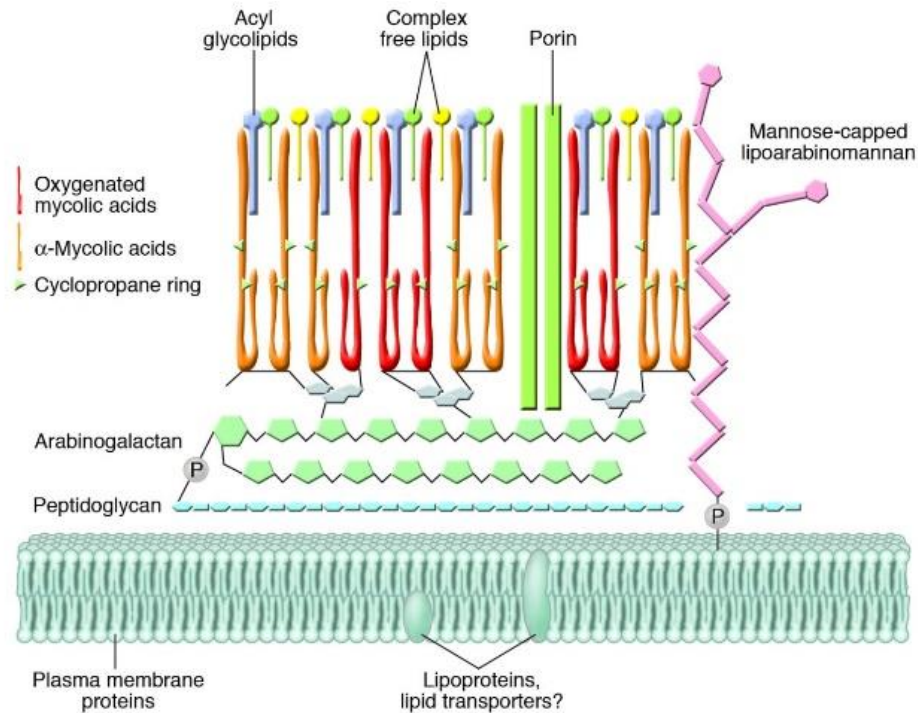
Figure 2. Graphical representation of the cell wall of *Mycobacterium tuberculosis*. The various components are evidenced with different colours and indicated by labels.[8]

Tuberculosis is an infectious disease which spreads through tiny droplets of saliva (0.5-5.0 µm), expelled by affected patients during expiratory activities; these particles are known as "droplet nuclei", and they contain active Mtb bacilli. Due to their small size, these droplets deliver the microorganism directly to pulmonary alveoli, thereby making the lungs the primary site of infection.[9] In fact, although TB affects the respiratory system in more than 70% of all cases, it can also damage other organs, including bones, meninges and lymph nodes. Immunosuppressed people, especially HIV-positive patients, are particularly predisposed to develop an active form of the disease; otherwise, individuals with strong immune defences usually manage to inactivate the pathogen. Despite that, the microorganism may endure in a quiescent form (latent TB) and reactivate later to cause relapse cases in immunocompromised subjects.[10]

The classical clinical symptoms of tuberculosis are rather unspecific: they include chronic cough, loss of weight and appetite, haemoptysis, low fever and night sweats. Since they are common to more frequent pulmonary diseases, the identification of TB is frequently delayed. The most common method for the detection of the disease is the analysis of a sputum sample for the identification

of acid-fast organisms, using the fluorescent complex auramine-rhodamine or acid-fast stains. Another useful diagnostic approach is the Mendel-Mantoux test (also known as Pirquet test or Tuberculine Skin Test), which evaluates the dermal reaction to the subcutaneous injection of purified protein extracts of tubercular bacilli; this technique is particularly valuable because it can detect latent forms of TB. Finally, radiological analyses can be employed to reveal lymphadenopathy and lung lesions.[6,10]

The recommended treatment for drug-susceptible tuberculosis is based on a 6-month regimen with four first-line drugs (Figure 3), namely isoniazid (INH), rifampicin (RIF), ethambutol (EMB) and pyrazinamide (PZA). All of them are used simultaneously in an intensive 2-month phase, followed by a continuation phase, which lasts from 4 to 7 months, with just isoniazid and rifampicin. Other drugs are sometimes employed, usually if resistance to the most common agents is detected or in case of co-morbidities. The classification as first-line drugs is based on the effectiveness, safety and availability of the therapeutic agent; rifabutin (RBT) and rifapentine (RPT) can be used in HIV-positive patients, while streptomycin (SM) is usually adopted in relapse cases. Second-line drugs generally exhibit either weak or unclear bacteriostatic activity, sometimes accompanied by serious side effects; hence, they are usually employed only in case of resistance or intolerance to first-line agents. They include fluoroquinolones (moxifloxacin), aminoglycosides (kanamycin, capreomycin), cycloserine and ethionamide. Finally, third-line agents, including clofazimine, linezolid, amoxicillin/clavulanate, imipenem/cilastatin, clarithromycin, are rarely used as they show an unclear therapeutic profile.[11]
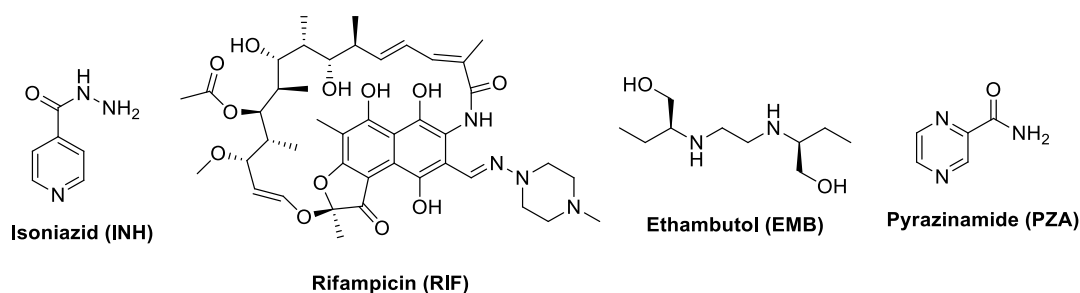
Figure 3. First-line antitubercular drugs.

Nowadays, drug resistance represents the major concern for TB control in many countries. Therapy mismanagement is usually the principal determinant of the emergence of new resistant strains; then, person-to-person transmission contributes to the spread of the infections. The molecular mechanisms behind drug resistance are numerous; they include modifications in the permeability of the cell wall, changes in the pathways involved in drug activation and/or metabolism, mutations in the targets, degradation or inactivation of specific enzymes. These alterations are usually the result of spontaneous mutations in chromosomal genes; no horizontal gene transfer has ever been reported for Mtb.[12] Depending on their drug sensitivity, resistant TB infections can be classified in different categories:

- RR-TB or rifampicin-resistant TB;

- MDR-TB or multidrug-resistant TB, characterised by an insensitivity to both rifampicin and isoniazid;

- XDR-TB or extensively drug-resistant TB, which can be defined as an MDR infection with an additional resistance to a fluoroquinolone and a second-line injectable agent (amikacin, capreomycin, kanamycin).

Although XDR-TB is still not officially recognised by WHO as a sub-type of resistant TB, a total of 10 800 cases belonging to this class were reported by 77 countries in 2017; most cases were registered in the Russian Federation (3661), India (2450), Ukraine (1097), South Africa (747) and Belarus (525).[2]

The only available vaccine against TB was discovered in 1921 by Albert Calmette and Camille Guérin: the use of their attenuated strain of *Mycobacterium bovis* (BCG) can prevent severe forms of extrapulmonary TB in infants, but it is inactive against the more common pulmonary TB at any age. The development of new vaccines has proven a challenging task, with few of them reaching the clinical trials. Three classes are being investigated, namely pre-exposure, preventive post-exposure, and therapeutic vaccines. They belong to three types: subunit vaccines, using antigens typical of active or latent TB, and viable or inactivated whole-cell vaccines. Despite the efforts, the available candidates show a very limited efficacy, especially in the prevention of the disease.[13]

In this context, the research of novel antitubercular agents is of pivotal importance; however, despite many efforts, the identification of valid alternatives to traditional drugs has proven extremely challenging and few significant advances have been made since the 1970s. Two new drugs were approved in 2012 and 2013 (Figure 4), namely bedaquiline (Sirturo®) and delamanid (Deltyba™). The former is a diarylquinoline compound, which acts by inhibiting the bacterial ATP synthase; it was registered after proving effective in phase II clinical trials, but it was later associated with an increased rate of heart attacks, related to its potential to affect cardiac electrophysiology, by prolonging the QTc interval. A similar side effect, although with less severe clinical consequences, can be induced by delamanid, a nitroimidazole derivative approved after phase II clinical trials. Because of their potential toxicity, the limited information regarding their safety profile in combination therapies, and their unavailability in high-burden countries, these new drugs are rarely used. Drug repurposing has led to unsatisfactory results too: clofazimine proved to have a limited bactericidal activity, linezolid induced severe side effects, and the studies on carbapenem efficacy are still fragmentary.[14]

In light of these considerations, research efforts should be enhanced to allow the identification of new chemical entities, acting on hitherto unexplored pharmacological targets.
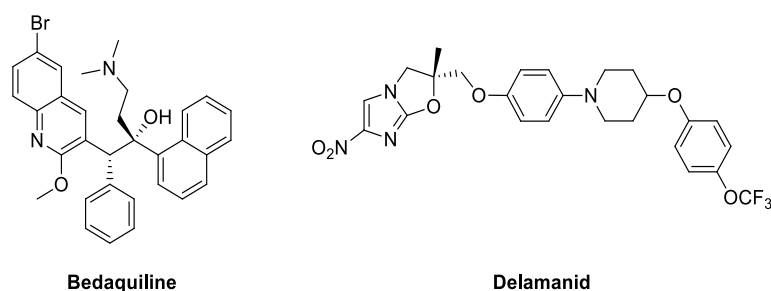
Figure 4. Recently approved antitubercular drugs.

1.2   IRON ACQUISITION IN *Mycobacterium tuberculosis*

Iron is involved in several essential biological processes, and as such it is crucial for the survival of all living beings, including humans and bacteria. In most organisms, the available iron is tightly coordinated to proteins to prevent its toxic

effects. Indeed, uncomplexed iron, which is predominantly ferrous iron (Fe^{2+}), catalyses the formation of reactive oxygen species through the Fenton reaction. Moreover, as a consequence of this process, Fe^{2+} is oxidised to the much less soluble Fe^{3+}, determining the formation of ferric iron precipitate.[15]

The human body contains *circa* 5 g of iron; most of it is coordinated as Fe^{2+} to the porphyrin ring of protoporphyrin IX (PPIX), making up the so-called haem complex, which constitutes the main component of haemoglobin. The majority of the circulating non-haem iron is transported as Fe^{3+} by transferrin and lactoferrin, while the remaining, non-circulating amount is stored in tissues, coordinated by enzymes or proteins like ferritin.[15]

Given the importance of iron for the survival of many pathogen species, the restriction of the accessibility of this metal has evolved as a fundamental part of the mammalian innate immune system. However, at the same time, bacteria have developed strategies to acquire to counteract this defence mechanism and acquire the amount of iron needed for their survival.[15]

Mycobacterium tuberculosis is mainly an intracellular pathogen, but during active infection it disseminates extracellularly, leaving the primary infection site to reach other areas of the body. While inside the macrophage phagosome, Mtb takes advantage of the rather high concentration of iron in the cytosol by producing siderophores which facilitate the permeation of the metal across the phagosomal membrane. Moreover, because macrophages also degrade senescent erythrocytes, haem iron becomes an additional source for the intracellular pathogen. When Mtb reaches the bloodstream, it may scavenge iron from transferrin, lactoferrin and ferritin, through the action of its siderophores, or acquire haem iron, thanks to the release of haemolysins, which lyse red blood cells, inducing the release of methaemoglobin.[15]
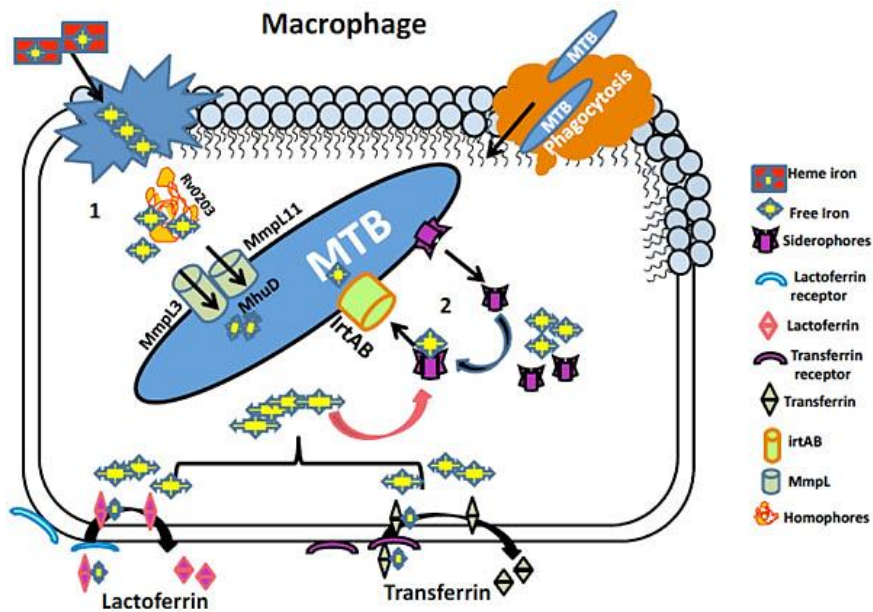
Figure 5. Graphical representation of the main iron acquisition pathways of *Mycobacterium tuberculosis*. Both haem and non-haem iron sources are exploited by the microorganism to ensure the maintenance of constant levels of this essential metal co-factor.[16]

As previously mentioned, a fundamental role in the mycobacterial acquisition of iron is played by siderophores, small iron-chelating molecules endowed with a high affinity for Fe^{3+} ions. These compounds can form octahedral complexes with the metal and are classified according to their iron-binding moiety. Mtb synthesises two mixed-type siderophores, bearing phenolate and hydroxamate groups, known as mycobactins and carboxymycobactins: these molecules share the same core, but they differ in the chemical properties, which in turn influence their localisation (Figure 5).[15]
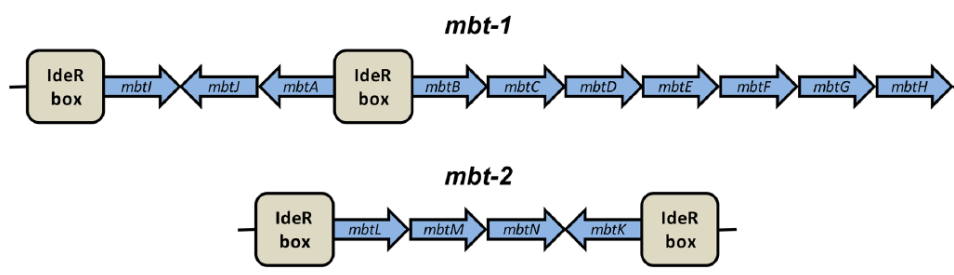
Figure 6. Organisation of the *mbt-1* and *mbt-2* gene clusters, encoding for the proteins involved in the siderophore-mediated iron acquisition pathway of Mtb.[15]

The enzymes that catalyse the formation of these siderophores are encoded by two gene clusters, namely *mbt-1* (*mbtA-mbtJ*) and *mbt-2* (*mbtK-mbtN*); the former is responsible for the formation of the mycobactin core, while the latter

allows the assembly of the aliphatic chain of mycobactin T (Figure 6). The first reaction of the biosynthetic process is mediated by MbtI, a Mg^{2+}-dependent enzyme that catalyses the double-step conversion of chorismic acid to salicylic acid, with the release of pyruvic acid. Then, MbtA activates salicylic acid, promoting its transfer to the thiolate domain of MbtB, the first component of a large complex, comprising three non-ribosomal peptide synthetases (NRPS, MbtB, MbtE, MbtF) and two polyketide synthases (MbtC, MbtD), which allows the building of the core. Additionally, MbtG catalyses the N-hydroxylation of the L-Lys portion, and MbtH assists the folding of the NRPSs; the function of the last protein, MbtJ, is still unknown. The formation of the aliphatic chain of mycobactin T is mediated by MbtL, MbtM and MbtN, while its fusion to the core scaffold is catalysed by MbtK (Figure 7).[15]
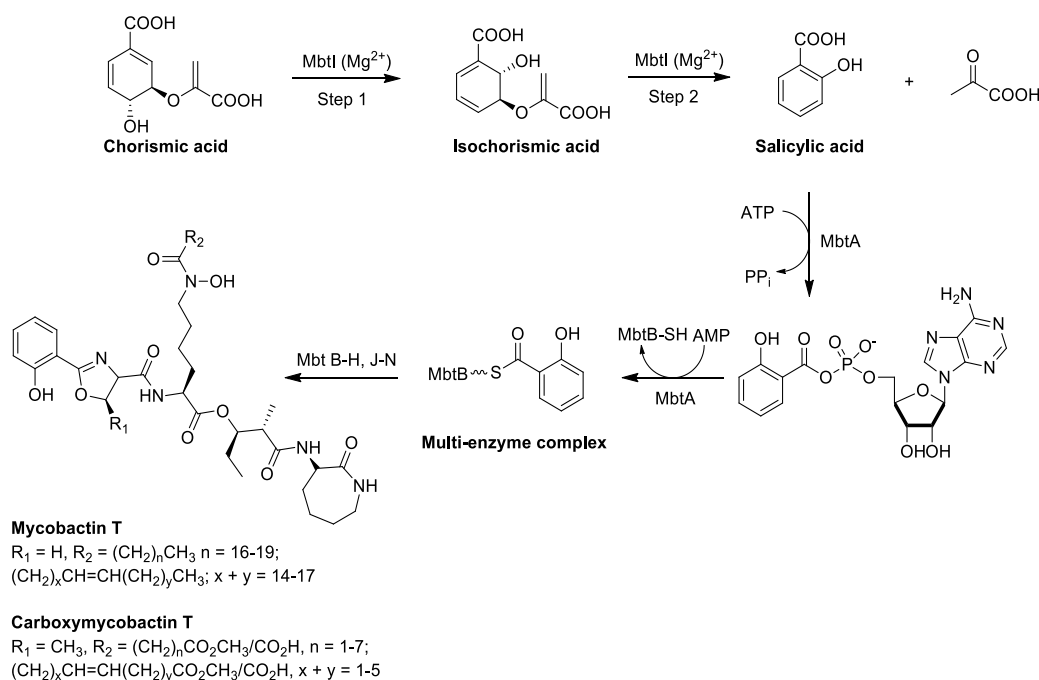
Figure 7. Biosynthetic process leading to the production of mycobacterial siderophores (mycobactins and carboxymycobactins).

This biosynthetic process is controlled by the iron-dependent repressor IdeR, which blocks the expression of the genes involved in the formation of siderophores: in the presence of high concentrations of iron, the complex IdeR-iron binds the so-called iron-box in the *mbt-1* and *mbt-2* loci, repressing the transcription. At the same time, it also promotes the synthesis the ferritin-type iron-storage proteins BfrA and BfrB. A form of positive regulation has been

observed for the *mbt-1* operon by HupB, a cell wall-associated protein which is produced at low concentrations of iron. HupB binds a site located 10 base pair upstream of the IdeR box, activating the *mbtB* promoter.[15]

As previously mentioned, the different chemical properties of mycobacterial siderophores influence their localisation and their activity. Carboxymycobactins feature a short aliphatic chain and a carboxylic function, which increase its solubility: hence, they are exported out of the mycobacterium into the extracellular environment, where they can scavenge iron from the host's iron-storage proteins. Conversely, mycobactins have a long aliphatic chain, which renders them highly hydrophobic and allows their anchorage to both the outer membrane and the cell wall. The export of these siderophores is mediated by the membrane transporters MmpL4 and MmpL5 (mycobacterial membrane protein large) and their ancillary proteins MmpS4 and MmpS5 (mycobacterial membrane protein small). ESX-3, belonging to the type VII secretion systems (T7SS), has also been reported as essential for the transport of carboxymycobactins and mycobactins, even though its role is still not clearly understood. Similarly, the mechanisms that govern the import of Fe^{3+}-bound siderophores into the mycobacterium are hitherto not well characterised. Apparently, after acquiring iron, carboxymycobactins can either deliver it to mycobactins for the internalisation, or use the dedicated transporter IrtA/IrtB, a heterodimeric protein involved in both the import and the release of iron from the siderophore. Finally, a recycling mechanism allows the re-use of the mycobactins after the delivery of the metal into the cytoplasm; their transport is again dependent on the previously described Mmp system.[15]

Beside siderophore-dependent pathways, other minor systems are involved in the acquisition of non-haem iron. Recent studies have shown that Mtb has the ability to relocate housekeeping proteins to the cell surface to scavenge iron from the host carriers. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) belongs to this class: this glycolytic enzyme can sequester iron from both transferrin, which is mainly located in the plasma, and lactoferrin, which is abundant in lung fluid.[17]

As mentioned above, ESX-3 has been shown to be essential for siderophore-mediated iron acquisition in yet-to-be-determined ways; ESX-3 mutants exhibit a defective ferric-carboxymycobactin uptake and a reduced accumulation of mycobactins on the cell wall. Moreover, ESX-3 may also be involved in the utilisation of haem iron and in the secretion of peptidic substrates (PE5, PPE4) counteracting the iron restriction defence mechanisms, adopted by the host.[18,19]

The acquisition of haem iron in the mycobacterium relies on two different uptake pathways. The main one is mediated by Rv0203, a 10.6 kDa extracellular protein which scavenges free haem iron. Although it cannot bind iron from haemoglobin, it was shown to be essential for the effective haem iron acquisition in Mtb. This observation led some scientists to hypothesise the presence of ancillary proteins that assist Rv0203 in the partial degradation of haemoglobin, in analogy with equivalent systems used by other bacteria, like *Porphyromonas gingivalis*. Bound haem iron is then transferred in a rapid and irreversible fashion to the D1 domain of the transmembrane proteins MmpL3 and MmpL11, *via* a direct transient protein-protein interaction. These proteins then act as transporters, delivering the haem inside the cytoplasm. Despite this hypothetical transfer mechanism seems most likely, some researchers have argued that MmpL3 and MmpL11 may be involved in the haem trafficking in a different way. Considering their well-known role as exporters of small molecules, these scientists have postulated that these proteins may not be bidirectional transporters, and hence may not be involved in the internalisation of haem but rather on its secretion, to control its toxicity. However, as of now, clear experimental evidences pointing in this direction are still missing.[15]

A second haem uptake pathway was discovered in a transposon library screening, aimed at identifying Mtb mutants showing resistance to a toxic haem analogue, gallium protoporphyrin IX (Ga-PPIX). The attenuated expression of three mycobacterial proteins, namely PPE36, PPE62 and FecB2, led to an increased resilience to Ga-PPIX and a reduced growth in haem-rich environments, suggesting an involvement of these candidates in haem uptake. Additional studies on these proteins showed that PPE36 and PPE62 are likely exposed on the cell surface, while FecB2 is located in the periplasmic space and

anchored to the inner membrane. While PPE36 and PPE62 supposedly scavenge haem, stealing it from other transporters, FecB2 may act a chaperone, shuttling haem through the periplasmic space into the cytoplasm. However, some researchers have raised doubts about this interpretation, based on the observation that all these proteins actually have a relatively low affinity for haem.[15]

Independently of the uptake mechanism, the internalised haem group is degraded by a mycobacterial IsdG-type protein (IsdG from *Staphylococcus aureus*), known as MhuD, to allow the release of iron. MhuD works by binding two haem groups per active site and degrading them into chromophores, mycobilin isomers and iron. This process can be split into two steps: in the first, MhuD oxidises the haem group to afford *β/δ-meso*-hydroxyhaem, via a hydroperoxy intermediate. Then, *meso*-hydroxyhaem is converted to mycobilin by a dioxygenase reaction, with the release of Fe^{2+}.[20]

The identification of haem uptake systems in Mtb is relatively recent (2011) and most of them are still poorly understood or characterised; as such, these pathways have not yet been considered for the development of anti-TB agents. Moreover, a direct link between the haem uptake and the virulence of Mtb has never been established. Conversely, mycobactins have been shown to be essential for the mycobacterium to build up the infection, especially in the initial stages. Hence, siderophore biosynthesis has been identified as a promising pathway for the discovery of novel drug candidates. Overall, 14 proteins are involved in this process, but so far only three of them have been investigated, namely MbtI, MbtA and MbtM.[15,21]

In the following paragraphs, the structure and activity of MbtI will be extensively discussed, as well as the efforts made towards the identification of selective enzymatic inhibitors.

1.3   MbtI: THE SALICYLATE SYNTHASE FROM *Mycobacterium tuberculosis*

MbtI is a Mg^{2+}-dependent enzyme, which belongs to the larger family of the Chorismate-Utilising Enzymes (CUE), a non-standard heterogeneous class that

includes proteins capable of catalysing different reactions on a common substrate, chorismate. It has been hypothesised that these enzymes may share the same ancestor, a theory that seems to be confirmed by the high structural similarity of the active sites of the CUEs. This characteristic is extremely interesting, for it allows to theorise a potential secondary multi-target action of selective CUE inhibitors. Chorismate is a key intermediate in the biosynthesis of several aromatic compounds in bacteria, including amino acids (phenylalanine, tyrosine, tryptophan), ubiquinone, menaquinone and salicylate.[22]

MbtI catalyses the first committed step in the biosynthesis of all siderophores of *Mycobacterium tuberculosis*, promoting the conversion of chorismate to salicylate. Hence, the inhibition of MbtI has the potential to impair the main iron acquisition pathway of the mycobacterium, blocking the enzymatic cascade from its beginning. Moreover, the absence of a homologous protein in the host makes it a perfect candidate for the development of novel anti-TB agents.[21]

1.3.1 Structure of MbtI

The structure of MbtI was solved for the first time by Harrison, Lott and co-workers in 2006. They used the multi-wavelength anomalous diffraction (MAD) method, employing datasets collected from the SeMet-substituted protein and then refining them with the native data at 1.8 Å to a final R factor of 0.205 and R_{free} of 0.240. The asymmetric unit contained four MbtI molecules, but the crystal packing indicated a monomeric enzyme, in accordance with the chromatographic profiles obtained during the protein purification (gel-filtration).[22]

MbtI is a 450-residue polypeptide which forms one large single domain, much like the other CUE (Figure 8). It is organised in two α/β subdomains, each composed by a large antiparallel β-sheet with helices packed against it on the outside. Subdomain I has a 10-strand β-sheet with five helices, while subdomain II has an 11-strand β-sheet with six helices. The two β-sheets pack approximately orthogonally across each other in the centre of the monomer, burying hydrophobic side chains between them. Outside the region where the β-sheets intersect, near structural elements from the C-terminal portion of the protein, a deep cleft is formed between the two subdomains.[22] This zone has been identified

as the active site of the enzyme, mainly by comparison to the product-bound structures of other CUEs (for example Irp9 from *Yersinia enterocolitica*, PDB code: 2FN1).[22,23] This cleft takes the form of a long groove, about 12 Å long, 10 Å deep, and 7 Å wide. It is delimited on one side by strand β21 and the following C-terminal helix, α11 (from subdomain I), and on the other by the β16-β17 loop, helix α7, and the β15-α6 loop (from subdomain II), while the β19-β20 and β12-β13 loops form the bottom of the cleft. The loops β16-β17 and, to a lesser extent β14-α6, are mobile, determining changes in the conformation of the molecule; while Harrison obtained the structure in an "open" state, other groups also described a "closed" conformation (PDB code: 3LOG), in which the two loops bend over the active site, possibly preventing the access of other molecules.[22,24] Apparently, the arrangement of the loops is determined by the ligands bound to the enzyme. Because the residues of the active sites of CUEs are well-conserved, a comparison can be made between MbtI and its homologues; however, the absence of both the co-factor (Mg^{2+}) and the product determines shifts in the position of the amino acids involved in the bonding. Glu297 is turned away to form an H-bond with His204, while Thr271 and His334, belonging to the mobile loops, are swung away from the active site. In the vicinity of this region, Ala269 is available to form a covalent intermediate during the catalytic reaction, probably allowing the discrimination between the C-2 and C-4 substituted substrates. In the deeper part of the active site, Harrison also observed a typical electron density, which was attributed to bound pyruvate, the byproduct of the enzymatic reaction (Figure 9). This would reasonably suggest the activity of MbtI in the expression vector, which is extremely likely, considering the presence of the substrate in *E. coli*. Moreover, the presence of the ligand in that position is consistent with the data from other CUEs.[22]
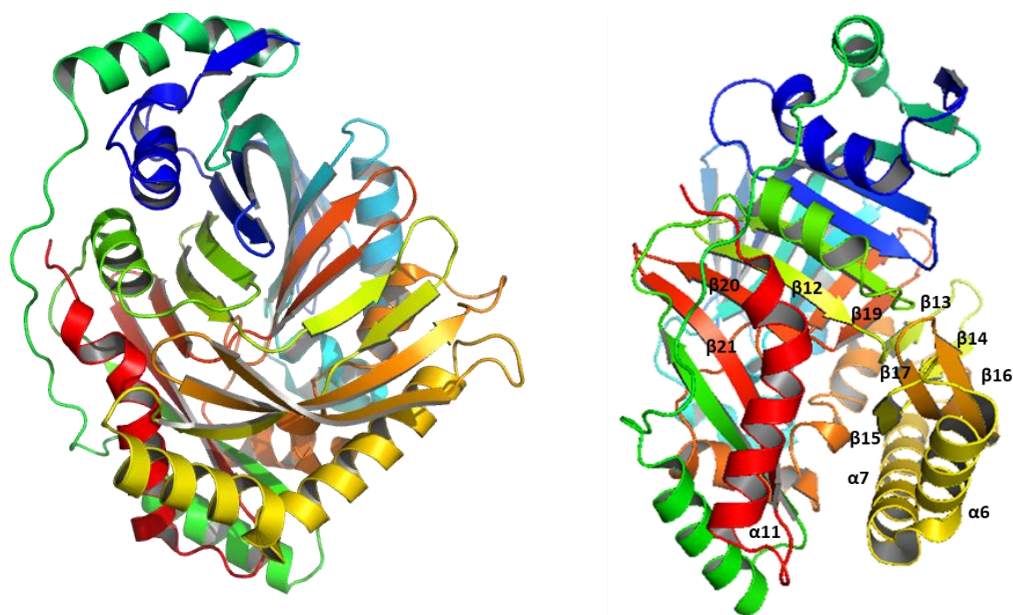
Figure 8. Ribbon diagram of the crystallographic structure of the MbtI monomer (PDB code: 2G5F), coloured blue to red from the N to the C terminus (left). The loops delimiting the active site are labelled on the right panel.
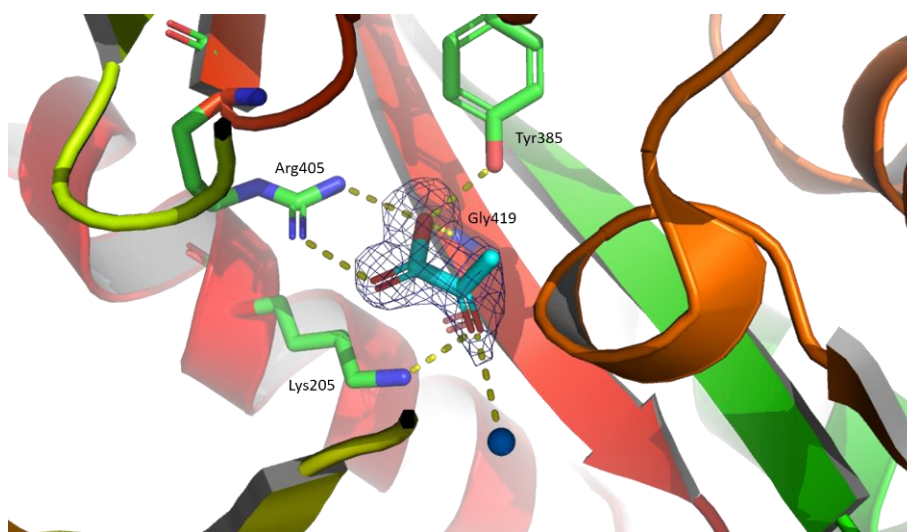
Figure 9. Representation of the active site of MbtI, co-crystallised with a molecule of pyruvate, produced by the conversion of chorismate to salicylate (PDB code: 2G5F). The interaction network of the pyruvate is shown as dashed lines; the side chains and the water molecules are represented as balls and sticks.

1.3.2 Activity of MbtI

The discovery of bound pyruvate in the protein structure provided the first evidence that MbtI acts as a salicylate synthase.[22] In the presence of Mg^{2+}, the enzyme catalyses two partial reactions: first, it converts chorismate to its C2-

substituted isomer, isochorismate (isochorismate synthase activity), and then it transforms this intermediate into salicylate with the release of pyruvate (isochorismate-pyruvate lyase activity). It has been hypothesised that there is an initial S_N2'' reaction, in which Lys205 (or Glu297) activates a water molecule for the attack on chorismate, with Glu252 polarising the C4-hydroxyl leaving group; this first step leads to the formation of a diol carbanion intermediate, which then converts to isochorismate with the release of water. An intramolecular [3,3]-sigmatropic rearrangement follows, leading to the production of salicylic acid and pyruvic acid (Figure 10).[21]
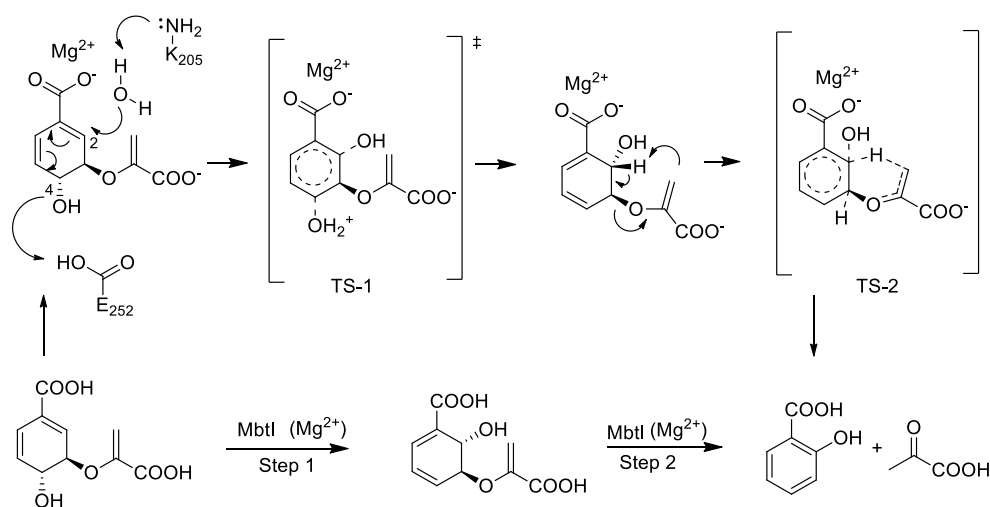
Figure 10. Mechanism for the MbtI-catalysed conversion of chorismate to salicylate, the building block for the formation of all mycobacterial siderophores.

It is widely known that the presence of Mg^{2+} is fundamental for these reactions to occur.[21] By employing a computational approach, Ferrer, Bertrán and co-workers postulated that, in physiological conditions, the ion is coordinated by two water molecules, two glutamate residues (Glu294 and Glu434) and two oxygen atoms of the C1 carboxylate group of the substrate. The presence of the ion deeply affects the overall catalytic reaction, since it produces changes in:

- conformation of the active site: the charge of the Mg^{2+} cation displaces the positively charged Lys293 and drags the negatively charged Glu297 towards the active site. It also allows the approach of the water molecule, that will be activated by Lys 205 (or Glu297);

- polarisation state of amino acid residues: calculations performed with the PropKa software have shown significant shifts in the pka values of the active

site residues, with or without the cation. Recent findings have suggested that Lys205, which was thought to be responsible for the activation of the water molecule, may actually be unable to perform this task, because it is protonated in the presence of Mg^{2+}, at neutral (and even slightly basic) pH, thus being incapable of abstracting the water proton. Glu297 has been proposed as a better candidate for the activation reaction, since it is unprotonated in both presence and absence of Mg^{2+}. Moreover, Glu252 is protonated at physiological pH only when the cation is present, thus being able to donate a proton to the C4 hydroxyl group;

- geometry of the substrate: thanks to the Mg^{2+}, the C1 carboxylate group of chorismate remains coplanar with the ring, favouring the delocalisation of the π system and reducing the electrical density on C2, thus facilitating the nucleophilic attack on this atom.[25]

MbtI is considered to be a highly promiscuous enzyme, since it can allegedly perform different reactions, depending on the surrounding environment. While the conversion of chorismate into salicylate is the favoured pathway in the presence of Mg^{2+}, the absence of the cofactor prevents all the above-mentioned structural modification to occur; in this condition, it has been hypothesised that MbtI may behave as a chorismate mutase, catalysing the conversion of chorismate to prephenate (Figure 11).[26] However, this conjecture remains highly debated, since some researchers observed the loss of the chorismate mutase activity after double purification of the protein, suggesting the possible presence of a contaminant enzyme in the previous studies.[27]
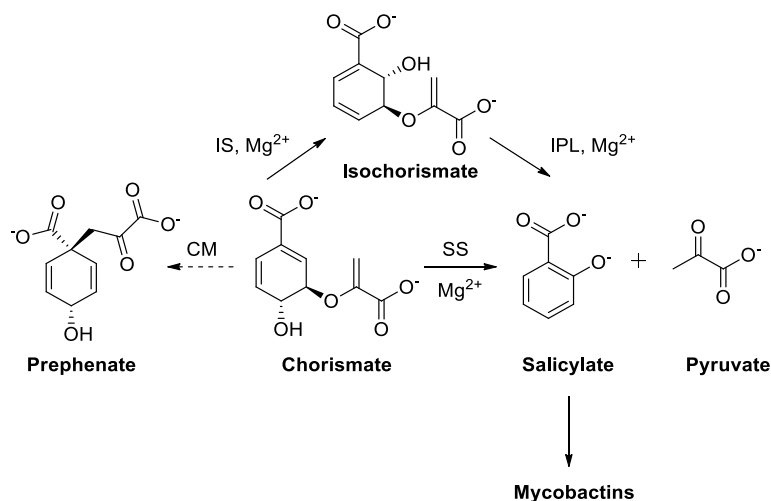
Figure 11. Hypothesised promiscuous activity of MbtI. According to this theory, the enzyme could act either as a salicylate synthase or a chorismate mutase, depending on the availability of the Mg^{2+} ion.

The activity of MbtI has been proven to be pH-dependent: isochorismate tends to accumulate at pH values below 7.5, while at pH 8 the main product is salicylate. This phenomenon is probably ascribable to modifications in the ionisation states of key amino acids.[25]

1.3.3 Inhibition of MbtI

Considering the high potential of MbtI as an innovative molecular target for the development of novel antitubercular agents, its inhibition has been the subject of several studies, yielding various levels of success. However, these inhibitors were never tested *in vivo* in infected animals.

In the following paragraphs, a classification of the literature inhibitors based on their structure will be discussed, to provide an overview of the state of the art in MbtI inhibition (Figure 12).
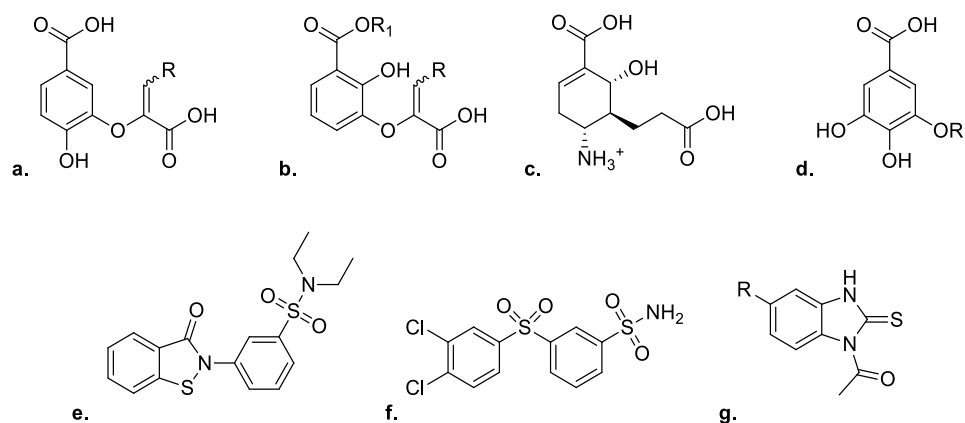
Figure 12. Small-molecule inhibitors of MbtI: **a**. chorismate analogues; **b**. isochorismate analogues; **c**. transition state analogues; **d**. gallate derivatives; **e**. benzisothiazolone derivatives; **f**. diarylsulfone derivatives; **g**. benzimidazole-2-thione derivatives.

1.3.3.1 *Chorismate analogues*

Compounds bearing the 3,4-dihydroxybenzoic acid moiety represent the first attempts to develop MbtI inhibitors, mimicking the structure of the natural substrate; they show a moderate to poor activity, with K_i values in the range 290-1300 µM. Surprisingly, most of the molecules of this series are two-fold less potent than the simplified derivative 3-hydroxybenzoic acid (K_i = 500 µM). The only exception to this trend is represented by compounds bearing small hydrophobic substituents (e.g. methyl group, with a K_i = 290 µM, and isopropyl group, with a K_i = 310 µM) linked to the double bond of the enolpyruvyl side chain, which probably interact with lipophilic residues of the binding pocket. This modification also enhances the selectivity of the inhibitor towards MbtI, since studies have shown that the same groups are ineffective in increasing the potency of the compounds on other salicylate synthases.[21]

1.3.3.2 *Isochorismate analogues*

Considering the mediocre activity of chorismate-related compounds, the intermediate of the catalytic reaction was next taken into consideration for the development of substrate-mimetic inhibitors. Interestingly, compounds possessing the isochorismate-like 2,3-dihydroxybenzoate scaffold show a significantly higher inhibitory potency, with K_i values in the low-micromolar range (≈11-14 µM). Again, small hydrophobic substituents on the enolpyruvyl side chain enhance the inhibitory activity, while the insertion of a phenyl ring determines a slight reduction of potency. The configuration of the double bond

does not seem to play a significant role, even though *Z* isomers are by far the most common in co-crystallised structures.[21]

Binding assays have confirmed the hypothesis regarding the high plasticity of the active site, which allows the accommodation of a wide variety of substituents on the side chain. Detailed studies have shown that this series of ligands can bind MbtI with two different binding modes: the first one, corresponding to the interaction mode of the natural substrate, is shown by the least potent inhibitor, AMT. The second conformation has been observed for the remaining inhibitors, some of which show a better affinity for MbtI compared to the substrates. These conformational studies have been carried out based on crystal structures of the enzyme in complex with the inhibitors. However, the metal cofactor is absent in the crystals; hence, the role it may play in determining the binding mode is still to be elucidated. In point of fact, Chi, Bulloch and co-workers reported the crystal structure of a Mg^{2+}-bound MbtI molecule; however, the interaction mode of the ion seems artificial and unnatural, considering it is coordinated only by water molecules and it does not make direct contacts with any of the amino acids of the active site. Therefore, it is hard to make deductions from this observation, which has been defined as a probable artefact by the author themselves.[28]

Although efficient in enzyme inhibition tests, isochorismate analogues have shown disappointing results in whole cell assays; this is probably due to the high hydrophilicity of the inhibitors, which prevents them from crossing the mycobacterial cell wall. In order to increase the lipophilicity of the compounds, two approaches have been pursued:

- esterification of carboxylic moieties: methyl ester derivatives have been synthesised and tested, showing good potency, and thus suggesting the capability of esters to cross the cell wall and undergo hydrolysis to the corresponding acid by means of bacterial esterases;

- esterification of carboxylic moieties and increase of lipophilicity of the C3 substituent: this aim has been achieved through the insertion of aliphatic residues and aromatic moieties on the side chain. The most potent compound (MIC^{50} = 25 μM) possesses an *o*-trifluoromethylphenyl substituent on the double

bond; notably, this molecule is not as efficient in inhibiting MbtI, as could be expected (K_i = 125 µM). The relatively high bactericidal activity can be explained considering its high degree of hydrophobicity and hypothesising an additional inhibitory activity on another enzyme, possibly a CUE.[21]

1.3.3.3 *Transition state analogues*

Transition state analogues have been proposed as alternative MbtI inhibitors by Zhang, Aldrich and co-workers; however, they do not display any significant activity. Inspired by the work of Bartlett and Kozolowski, who had discovered a potent transition state inhibitor of a homologue of MbtI from *E. coli* (EntC),[29] Aldrich's group proposed to apply a similar approach on the mycobacterial enzyme. They decided to mimic the first transition state (TS1) of the MbtI-catalysed reaction, since it is the easiest to reproduce. In 2015, they proposed (1*R*,5*R*,6*S*)-4-carboxy-6-(2-carboxyethyl)-5-hydroxycyclohex-3-enaminium as a possible mimetic; this compound possesses an amine moiety at position 4, which mimics the charged leaving group, and a propionate fragment, which replaces the labile pyruvic side chain. The compound shows a < 10% inhibition at 100 µM; the reason for this mediocre result can be ascribed to the absence of the C3 oxygen atom, which is essential to form a H bond, and to the presence of a charged NH_3^+ group, which probably forms repulsive electrostatic interactions with an amino acid residue of the active site. Additionally, their lack of activity also suggests the existence of structural differences between MbtI and EntC.[30]

In 2017, they proposed a wider library of compounds, characterised by an uncharged scaffold; they explored different modifications of their previous molecule, inverting the positions of the hydroxyl and amino groups, re-introducing the natural enolpyruvyl side chain or adding lipophilic moieties. Again, the candidates showed no activity against MbtI.[31]

1.3.3.4 *Other MbtI inhibitors*

A high throughput screening (HTS) approach has been employed to identify additional MbtI inhibitors, with poor success.[32] The screening of a library of one hundred thousand compounds has led to the identification of four classes of potential inhibitors:

- gallate derivatives: these compounds display a poor activity on MbtI, probably because of the hydroxyl group in position 5, which increases the steric hindrance of the inhibitor without establishing any interaction with the amino acid residues of the active site.

- benzisothiazolones: they exhibit an extremely high potency, with IC_{50} values near the enzyme concentration; they are thought to be irreversible inhibitors, that covalently label with their isothiazolone moiety thiol nucleophiles to form mixed disulfide bonds with four Cys residues, namely Cys50, Cys211, Cys380 and Cys436. In fact, these molecules are probably Pan-Assay-Interference compounds (PAINS), meaning they always give high scores in biological tests. Therefore, they are unsuitable for further development, since collected data about their inhibitory activity are unreliable;

- diarylsulfones: they are derivatives of the anti-leprotic drug dapsone. The most potent compound shows an interesting IC_{50} of 35 μM; however, it possesses an extremely tight SAR that prevents the implementation of any modification to its structure. All changes in the substituents of the phenyl rings result in an abolishment of the inhibitory activity;

- benzimidazole-2-thiones: the reference compound acts as a reversible non-competitive inhibitor of MbtI, with an IC_{50} of 9.2 μM. SAR studies have determined that:

  - the aromatic ring, the 2-thione and the NH group are all fundamental for the activity;

  - the acetyl group is important for the activity; its deletion or replacement with an *N*-methyl substituent results in a six- to nine-fold decrease in potency;

  - substitution of the aryl ring does not have significant effects, the insertion of a methyl group on C5 allows a slight increase in potency (IC_{50} = 7.6 μM).[21,32]

Chapter 2: Objectives

In the context of the pressing need of novel drugs for the treatment of TB, the aim of this research project was the identification of new, potent inhibitors of the salicylate synthase (MbtI) from *Mycobacterium tuberculosis*. MbtI is an essential mycobacterium-specific enzyme, which does not have any counterpart in human cells; as such, it can be considered an attractive target for the development of new antibiotics. This innovative and validated target is involved in the biosynthesis of all mycobacterial siderophores, namely mycobactins and carboxymycobactins; these high-affinity iron-chelating molecules are responsible for the acquisition of Fe^{2+}, an essential metal cofactor mediating several biological processes, implicated in the survival and pathogenicity of the microorganism.

To this end, we performed a structure-based virtual screening, employing the available co-crystal structures of MbtI: we built a pharmacophore model, and we used it to screen the Enamine database. Subsequent steps of consensus docking and molecular dynamics allowed us to identify a competitive furan-based inhibitor, endowed with a promising activity on the enzyme. With the aim of understanding the key structural features required for the inhibitory effect, we modified the hit compound to derive preliminary information. Then, we performed an in-depth exploration of the chemical space around this scaffold through a structure-activity relationship study.

Finally, we preliminarily assessed the possibility to produce liposome-based nanocarriers for the design of an innovative antitubercular inhalation therapy, employing a combination of new and established antitubercular agents.

Chapter 3: Results and Discussion

3.1 COMPUTATIONAL STUDIES

In order to identify better MbtI inhibitors, a Virtual Screening (VS) study was developed. A pharmacophore model was generated, based on the analysis of the co-crystal structure of MbtI with 3-(1-carboxyprop-1-enyloxy)-2-hydroxybenzoic acid (Methyl-AMT, PDB code: 3VEH).[28,33] Figure 13 depicts the interactions of Methyl-AMT inside the binding site. The carboxylic group of the 2-hydroxy benzoic acid forms an H-bond with the side-chain of Tyr385; the analysis of different MbtI crystal structures suggests that the position of this residue is highly conserved, and it is important for the interaction of the carboxylic function of the reported inhibitors. The aromatic ring of the 2-hydroxy benzoic acid shows lipophilic interactions with Ile207, Ala362, Leu404 and a cation-π interaction with Lys438. The methyl fragment of the 1-carboxyprop-1-enyloxy chain is inserted into a lipophilic pocket mainly delimited by Pro251, Leu268, His334 and Thr361, whereas the second carboxylic group shows an ionic interaction involving Lys438 and compatible with the interaction with the Mg^{2+} ion.[28,33]
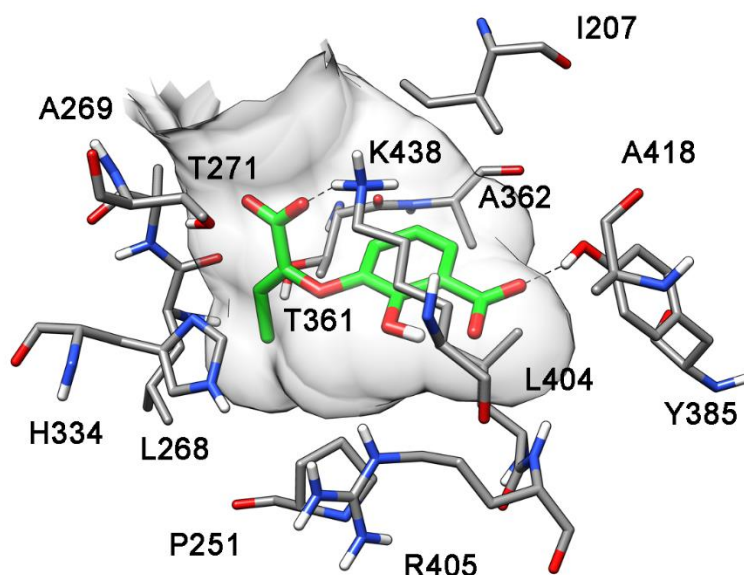
Figure 13. Graphical representation of Methyl-AMT docked in the active site of MbtI. The main interactions with the amino acid residues of the enzyme are evidenced by dashed lines.

Based on these considerations, we identified four features in the pharmacophore, as depicted in Figure 14. The model was generated using the software Ligandscout,[34] and comprised (1) an H-bond acceptor feature, representing the oxygen atom of the carboxylic fragment interacting with Y385, (2, 3) two hydrophobic features, representing the aromatic and methyl substituents, and (4) a negative feature, representing the carboxylic group of the 1-carboxypropenyl-1-enyloxy fragment. In addition, the pharmacophore model was refined by adding excluded volume spheres, mimicking the steric hindrance represented by the MbtI binding site.[33]
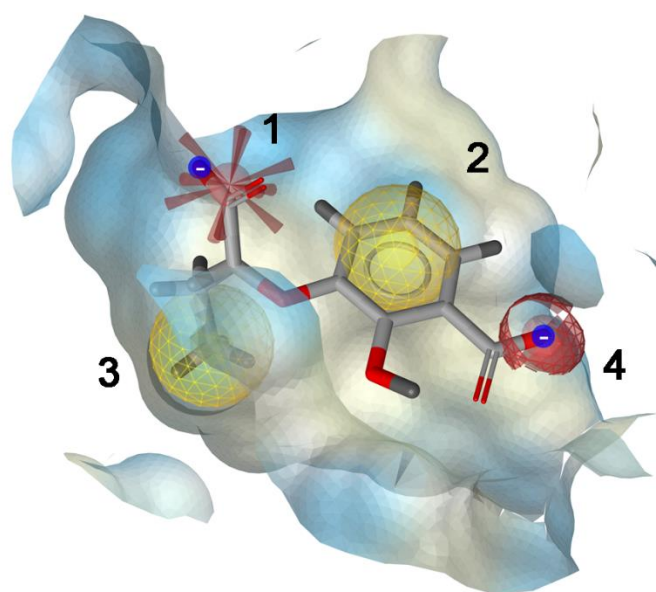
Figure 14. Pharmacophore model obtained by the computational analysis of the available crystal structures of MbtI, in complex with inhibitors. The four features are indicated by numbers 1-4.
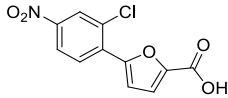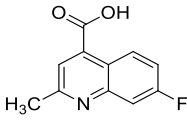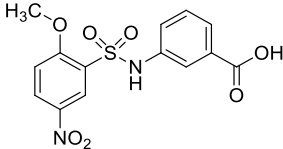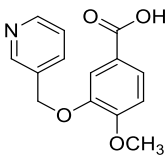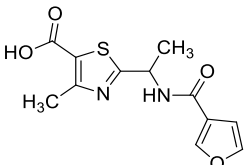
This model was used to screen the Enamine database, comprising about 1500000 commercially available compounds, retrieving only entries which matched all the four pharmacophore features and respected the volume constraints, given by the binding site. By applying this strict filter, only 2050 compounds were selected for the docking studies. Recently, Tuccinardi and co-workers have reported an evaluation of the consensus docking approach.[35] Through this approach, one ligand is docked into the target protein by means of multiple docking methods. Then, among the best-ranked poses (originated by the different docking procedures), the binding pose predicted by the largest number of docking methods is considered as the best docking pose. From a qualitative

point of view, previous results highlighted that consensus docking was able to predict ligand binding poses better than the single docking evaluations. Furthermore, concerning the VS studies, the results suggested that this approach performed as well as the best available methods found in the literature, and it was also able to experimentally identify new active molecules. Hence, the 2050 compounds, previously selected through the pharmacophore filter, were subjected to a consensus docking protocol, calibrated for the docking of MbtI inhibitors. The MbtI-tuned procedure included 5 different docking methods (i.e. Plants and Gold, with the four different fitness scoring functions), which were found to be the most reliable among a total of 11 different methods, based on self-docking studies performed on six MbtI-ligand co-crystal structures. For each of the 2050 potential ligands, the five docking poses, predicted by the five docking procedures, were clustered together to search for common binding modes. The analysis of the results revealed that the five protocols predicted the same binding pose only for 74 compounds, which thus reached the maximum consensus level. Therefore, based on the consensus docking filter, only 3.6% of the docked compounds were considered as potentially active inhibitors. The analysis of the binding mode predicted for the 74 potential MbtI inhibitors revealed that only five of them still matched the receptor-based pharmacophore model in their docking pose. On these bases, the five compounds were purchased at the highest purity level available from the commercial source (≥ 95%) and subjected to biological assays to evaluate their inhibitory activity on the target enzyme.[33]

The effect of the candidates on MbtI activity was tested at subsaturating concentrations of the substrate (chorismic acid, CHA, 50 μM), at a final concentration of 100 μM.[33] As the acquisition of chorismic acid from commercial sources was, unfortunately, demanding, we started its production from an engineered strain of *E. coli* (KA12). After a single-step procedure employing chromatographic techniques, chorismic acid was obtained at high purity (95%) and used for the biological tests. The culturing, isolation, extraction and purification conditions were refined in order to set up an improved protocol for its production in higher yields.

The most potent MbtI inhibitors reported in the literature belong to the AMT series; these derivatives are structurally related to isochorismate and show IC_{50} values of about 10 mM. In particular, Methyl-AMT, which emerged as the most powerful candidate of this class, was used as a positive control, showing an IC_{50} value of 11.6 µM, in agreement with the previously reported values.[24] As shown in Table 1, one compound was found to be provided with a remarkable activity (**VS1**, IC_{50} = 21.1 µM), another displayed a modest activity (compound III, IC_{50} = 77.6 µM), while the remaining three compounds were completely inactive.[33]

Table 1. Biological evaluation of the compounds identified by the VS study (**VS1**-**5**). The inhibitory effect is expressed as percentage of residual enzymatic activity (at 100 µM ligand concentration) for all compounds, and half-maximal inhibitory concentration (IC_{50}, µM) for the active candidates.

Code	Structure	% Residual activity	IC_{50}
VS1		21.3 ± 4.3	21.1 ± 4.1
VS2		101.3 ± 3.6	-
VS3		38.4 ± 11.4	77.6 ± 5.7
VS4		104.0 ± 13.2	-
VS5		108.8 ± 12.0	-

Hence, compounds **VS1** and **VS3** were subjected to molecular dynamics (MD) simulations, with explicit water, in order to further characterise their predicted binding mode. The magnesium ion was also included in MbtI active site in order to simulate its interaction with the ligands. As shown in Figure 15, the

Mg^{2+} ion presents an octahedral coordination geometry: it is bound by a water molecule, one oxygen of the carboxylate groups of Glu294, Glu297 and Glu434 and the two oxygens of the carboxylate of compound I. The furan ring shows an H-bond with Lys438, whereas the benzene ring, as for Methyl-AMT, exhibits lipophilic interactions with Ile207, Leu404 and a cation-π interaction with Lys438. Finally, the nitro group strongly interacts with Tyr385 and forms a second H-bond with the nitrogen backbone of Arg405. Figure 15 shows the average minimized structure of the MtbI-**VS3** complex. The two oxygens of the carboxyl group of the compound interact with the Mg^{2+} ion, the benzoic acid ring shows lipophilic interactions with Ile207, whereas the 2-methoxy-5-nitrophenyl fragment loses the H-bond with the hydroxyl group of Tyr385 during the initial steps of the MD. However, the loss of the H-bond with Tyr385 could be partially compensated with the presence of H-bonds between the sulfonamide group and the backbone oxygen of Thr361 and the side chain of Lys438.[33]
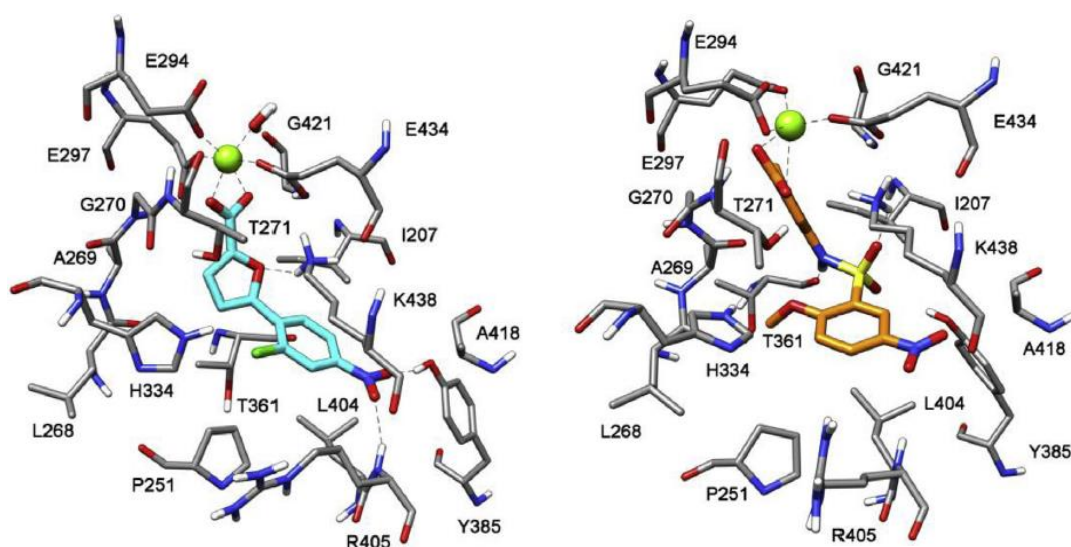
Figure 15. The binding modes of **VS1** (left) and **VS3** (right), as predicted by the computational simulations. The Mg^{2+} ion is represented as a green sphere, and the interacting amino acid residues are shown in sticks.

In order to further assess the solidity of the VS workflow, a retrospective evaluation was carried out. Simpler ligand-based approaches such as fingerprint similarity searches or 3D ligand similarity screenings showed a lower performance in retrieving the most active identified compound (**VS1**) from the initial commercial database. Moreover, the application of the consensus docking

approach was found to outperform the single docking methods in identifying compound I from the 2050 ligands selected through the pharmacophore filter.[33]

3.2   STRUCTURE-ACTIVITY RELATIONSHIP STUDIES

The VS study, described in detail in the previous section, allowed the identification of a novel hit compound (**VS1**), endowed with inhibitory activity against MbtI. **VS1** exhibited a promising IC_{50} value of 21.1 μM, corresponding to a 78.7% inhibition of the enzymatic activity. In order to validate the biological results and effectively compare them with literature data, we synthesised Methyl-AMT, following the procedure published by Manos-Turvey, Payne and co-workers.[24] The biological analysis of the compound confirmed the previously reported results, with an IC_{50} = 11.6 ± 2.4.[33]

Therefore, this hit was used as a starting point for an in-depth investigation into the potentiality of this interesting scaffold for the obtainment of more potent analogues. In the pursuance of this objective, several derivatives were designed and synthesised, to explore the chemical space and to define the structural and functional requirements to achieve the inhibition of MbtI.

Because the computational studies had identified the carboxylic acid as a fundamental moiety for the coordination of the metal cofactor, we first analysed the substitution pattern on the phenyl ring. By alternatively removing the chlorine atom and the nitro group, we determined the essential role of the p-NO$_2$ function. Hence, considering the potential toxicity of this moiety, we attempted its substitution with other groups, with little success. Therefore, we decided to re-introduce a substituent in the *ortho* position; first, we maintained the nitro group, and then we replaced it with the groups that had yielded the best results in the previous series, namely the trifluoromethyl group and the fluorine atom. Among CF$_3$-substituted derivatives, some candidates showed a remarkable inhibitory activity, effectively replacing the p-NO$_2$ derivative. In the hope of finding an even better inhibitor, and to complete the chemical investigation of the phenyl group, we moved to the hitherto unexplored *meta* position: among this series, we discovered some of the best inhibitors, endowed with a similar/better potency compared to our first lead, but lacking the nitro function. Furthermore, we explored the role of the central heterocycle by testing the effects of its replacement with the phenyl ring and of other moieties like a thiophene, oxazole, thiazole and the imidazole, decorated with the best groups, previously identified.
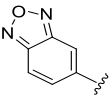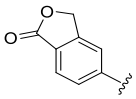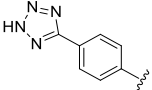
Although all of them retained a certain activity, the furan still performed better than the other cyclic moieties.

3.2.1 *para*-Substituted derivatives

As mentioned above, the synthesis of *p*-substituted derivatives was inspired by the observation that the removal of the chlorine atom of **VS1** led to a more potent derivative (**1**), while the elimination of the *p*-NO_2 group resulted in an inactive compound (**2**). This data suggested the importance of the substituent in the *para* position and, at the same time, hinted the negligible role of the *o*-Cl. Moreover, when both these substituents were removed (**3**), a complete loss of activity was detected. The replacement of the phenyl ring with pyridine (**4**) afforded a partially active compound. The isomeric derivatives of **1**, showing the nitro substituent in *meta* (**5**) and in *ortho* (**6**) positions, determined a loss of potency, with an MbtI inhibitory activity of 50%, at a concentration of 100 µM. These observations indicated the importance of both the NO_2 group and its position to ensure the enzymatic inhibition. However, the potential toxicity of this moiety prompted us to try to find suitable alternatives, with a better safety profile. The *p*-hydroxy derivative **7** proved to be inactive, probably owing to its inability to simultaneously accept two hydrogen bonds in the suitable geometry; likewise, the substitution of the nitro group with a chlorine (**8**), an amine (**9**) or an amide (**10**) led to inactive analogues. Interestingly, the introduction of a CF_3 group (**11**) or a fluorine atom (**12**) led to moderately active inhibitors (66% and 44% residual enzymatic activity at 100 µM, respectively). As a further attempt, the *para*-nitro group of **1** was replaced with the methylsulfonyl (**13**) and sulphonamide (**14**) substituents. With respect to the nitro moiety, they are larger fragments; however, they could potentially interact through two H-bonds with the protein by means of the two oxygens, mimicking the simulated interaction of the nitro moiety. In the first case, the replacement resulted in complete loss of activity, whereas compound **14** retained a very poor activity (84% residual activity). The substitution of the *para*-nitro group with other bioisosteres led to weak (**15,16**) or inactive (**17,18,19**) compounds (Table 2).[33]

Table 2. *p*-Substituted derivatives of **VS1** (**1-19**). The inhibitory effect is expressed as percentage of residual enzymatic activity (at 100 μM ligand concentration) for all compounds, and half-maximal inhibitory concentration (IC$_{50}$, μM) for the most active candidates (residual activity ≤ 25%).
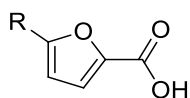
Code	R	% Residual activity	IC$_{50}$
1	4-NO$_2$-Ph	18.2 ± 5.1	7.6 ± 1.6
2	2-Cl-Ph	92.4 ± 13.8	-
3	Ph	102.0 ± 8.5	-
4	pyr	45.0 ± 5.1	-
5	3-NO$_2$-Ph	50.2 ± 13.0	-
6	2-NO$_2$-Ph	55.1 ± 22.7	-
7	4-OH-Ph	96.3 ± 19.2	-
8	4-Cl-Ph	130.3 ± 36.2	-
9	4-NH$_2$-Ph	100.3 ± 11.5	-
10	4-CONH$_2$-Ph	77.2 ± 11.6	-
11	4-CF$_3$-Ph	66.3 ± 9.7	-
12	4-F-Ph	44.1 ± 9.1	-
13	4-SO$_2$CH$_3$-Ph	98.5 ± 14.4	-
14	4-SO$_2$NH$_2$-Ph	84.6 ± 16.2	-
15		46.8 ± 8.3	-
16		58.0 ± 8.9	-
17	4-CN-Ph	125.6 ± 17.9	-
18	4-COOH-Ph	92.4 ± 11.8	-
19		104.4 ± 7.8	

Among this series, we evidenced that compound **1** shows an improved potency with respect to the initial hit, **VS1**, with an IC$_{50}$ value of 7.6 ± 1.6 μM. Moreover, the MbtI kinetic analysis in the presence of **1** proved that this compound is a potent competitive inhibitor towards the chorismate binding site, displaying a K$_i$ value of 5.3 ± 0.6 μM. Additional tests were performed to verify that the candidate was not a PAIN compound: the addition of BSA and Triton X-100 did not influence its IC$_{50}$, proving that it did not form aggregates with the target. Similarly, the addition of DTT did not impact on the activity, showing that the ligand did not interact with the cysteine residues of the protein. In order to

assess its antimycobacterial potency, **1** was tested against *M. tuberculosis* H37Rv. Its methyl ester derivative (71% residual activity) was also assayed, on the basis of the theoretical higher activity of esters with respect to carboxylic acids in whole-cell experiments on Mtb. Such a behaviour was also detected by Manos-Turvey, Payne and co-workers with their isochorismate analogues, and it is conceivably related to a better ability of the esters to cross the mycobacterial cell wall, due to their higher lipophilicity, and be hydrolysed by bacterial esterases.[24] Surprisingly, while compound **1** showed a MIC[99] of 156 μM, its ester did not allow for an improvement of the cellular potency; this counterintuitive behaviour is probably due to the low solubility of our esterified compound in the assay medium. Nevertheless, it is noteworthy that **1** can be considered as the first lead compound, combining a remarkable inhibitory activity on MbtI with a moderate antimycobacterial activity. To assess if the antitubercular activity was indeed related to the inhibition of iron uptake, we used the non-pathogenic *M. bovis* BCG, whose siderophores are closely akin to the mycobactins from *M. tuberculosis*.[36] The compound, when assayed against cells grown in the low-iron-containing Chelated Sauton's medium, showed a 3-fold lower MIC value compared to the one obtained against mycobacteria grown in the Middlebrook 7H9 medium (80 μM vs 240 μM), suggesting an involvement of iron uptake in the mechanism of action of **1**. Thus, in order to support the correlation of the antitubercular activity of this compound with MbtI inhibition, the effects of the compound on siderophore production were assessed in *M. bovis* BCG cells, grown in iron-depleted medium and in the presence of different sublethal concentrations of **1**. To this purpose, the levels of siderophore production were measured using the Universal CAS liquid assay and through the isolation and quantification of the mycobactins.[37] As shown in Figure 16, the CAS assay performed on the medium of the cells grown in different concentrations of **1**, showed that the removal of iron was inversely related to the concentration of the compound. Similarly, the concentration of mycobactins, determined in the same cells, decreased at higher concentrations of **1**, confirming that the inhibitory effect towards Mtb growth could be due to the inhibition of mycobactin biosynthesis (Figure 16).[33]
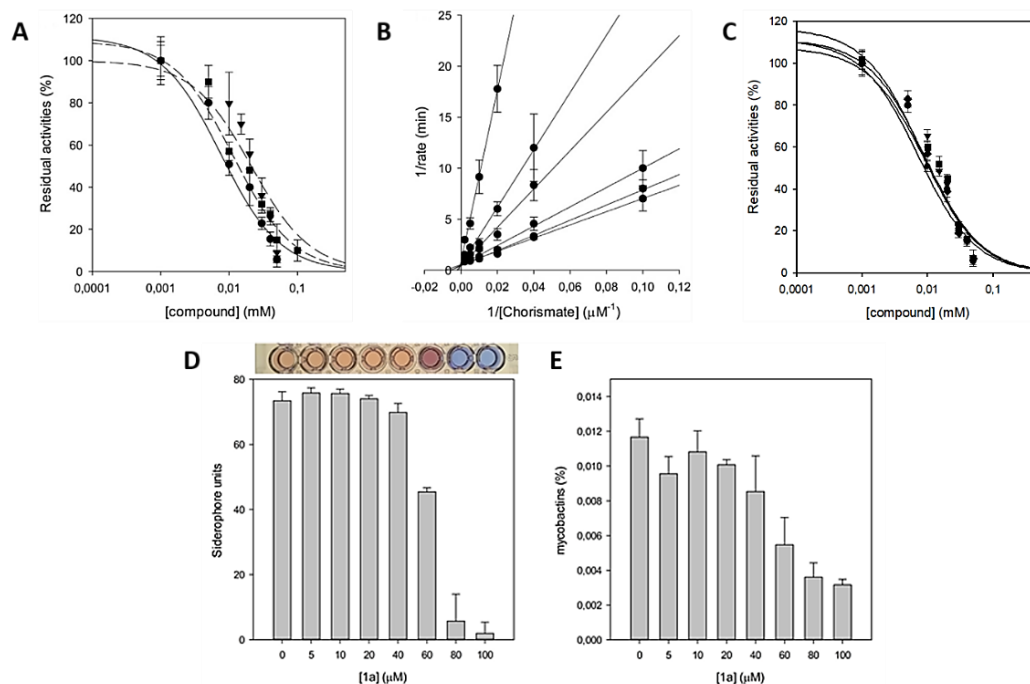
Figure 16. Biological evaluation of **1**. **A.** Graph comparing the IC_{50} values of **VS1** (▼) and Methyl-AMT (■) to that of **1** (●). **B.** Global reciprocal plot of data from MbtI steady-state kinetics analysis towards chorismic acid, in the presence of different concentrations of **1**. **C.** IC_{50} plot of **1** in the presence of BSA (■), Triton X-100 (▼) and DTT (◆). **D.** Universal CAS assay performed on *M. bovis* BCG, grown in the presence of different concentrations of **1**. **E.** Determination of the mycobactins in the abovementioned cells.

3.2.2 Disubstituted derivatives

Because the nitro group is a cause for concern in medicinal chemistry, we further developed our library of furan-based compounds, taking into account the previously obtained results. Specifically, we decided to replace this functional group with the trifluoromethyl moiety and the fluorine atom, as they allowed the retention of a certain activity against MbtI in the previous monosubstituted series, exploring at the same time the effects of different groups in the *ortho* position of the phenyl ring.

We began to build our library of disubstituted analogues of **VS1** with the synthesis of **20**, characterised by the presence of the *p*-trifluoromethyl group and the *o*-Cl. Because this compound displayed encouraging inhibitory properties (% residual enzymatic activity at 100 μM: 6.1 ± 2.5, IC_{50} 28.5 ± 2.6 μM), we decided to adopt it as a new hit and focus on the analysis of the effects of a variety of substituents in the *ortho* position of the phenyl ring, maintaining the *p*-CF₃ group.

The effects of the previously analysed substituents were evaluated in two more series of derivatives, one bearing the *p*-F group and the other displaying the original nitro moiety, to verify that the newly introduced functions could effectively replace (and possibly improve) the *p*-NO$_2$.[38]

The substitution of the chlorine atom with the fluorine in **21** did not affect the activity, while the introduction of the bromine atom in **22** diminished the inhibitory effect of the compound (% RA 38.0 ± 3.7), probably due to the low capability of bromine to act as a hydrogen bond acceptor together with its higher atomic radius, which determines a negative steric constraint. The presence of the electron donating groups in **23**, **24**, and **25** did not lead to an improvement of the biological effects. When an electron withdrawing moiety was introduced in the *ortho* position, as in **26**, we detected an increased activity. This outcome was confirmed by the introduction of an *o*-CF$_3$ moiety, which led to the disclosure of **27**, showing an IC$_{50}$ value comparable to that of the parent compound (**VS1**). Its positional isomer **28** was then prepared and tested to verify the importance of the *ortho* position in determining the activity: moving the CF$_3$ to the *meta* position resulted in a significant decrease of the inhibitory activity. However, quite interestingly, the positional isomer bearing the two CF$_3$ groups in position 3 and 5 (**29**) showed a remarkable activity, with a % residual activity of 8.9 ± 1.6.[38]

Interestingly, *p*-F (**30-37**) derivatives generally showed a lower activity; the only exception was **36**, bearing an electron-withdrawing cyano group in the *ortho* position, which showed an IC$_{50}$ value of 16.0 ± 1.7 µM. Conversely, the *p*-NO$_2$ (**38-44**) derivatives exhibited an inhibitory profile similar to that of the *p*-CF$_3$ analogues. These outcomes confirm the ability of the CF$_3$ to act as a bioisostere of the nitro moiety (Table 3).

Table 3. Disubstituted derivatives of **VS1** (**20**-**44**). The inhibitory effect is expressed as percentage of residual enzymatic activity (at 100 μM ligand concentration) for all compounds, and Half-Maximal Inhibitory Concentrations (IC$_{50}$, μM) for the most active candidates (residual activity ≤ 25%).
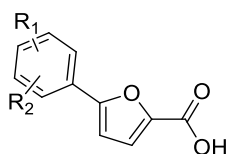
Code	R$_1$	R$_2$	% Residual activity	IC$_{50}$
20	4-CF$_3$	2-Cl	6.1 ± 2.5	28.5 ± 2.6
21	4-CF$_3$	2-F	6.1 ± 1.1	27.6 ±7.3
22	4-CF$_3$	2-Br	38.0 ± 3.7	-
23	4-CF$_3$	2-OH	16.8 ± 2.7	35.9 ± 10.3
24	4-CF$_3$	2-CH$_3$	19.3 ± 2.2	31.5 ± 9.6
25	4-CF$_3$	2-NH$_2$	13.1 ± 2.4	34.6 ± 10.9
26	4-CF$_3$	2-CN	5.7 ± 1.5	18.5 ± 3.2
27	4-CF$_3$	2-CF$_3$	3.9 ± 1.7	13.1 ± 2.0
28	4-CF$_3$	3-CF$_3$	64.3 ± 4.5	-
29	3-CF$_3$	5-CF$_3$	8.9 ± 1.6	18.8 ± 6.8
30	4-F	2-Cl	38.3 ± 7.4	-
31	4-F	2-F	24.0 ± 1.8	31.6 ± 2.8
32	4-F	2-Br	33.2 ± 6.0	-
33	4-F	2-OH	19.4 ± 1.6	26.3 ± 1.2
34	4-F	2-CH$_3$	23.7 ± 5.0	37.5 ± 4.2
35	4-F	2-NH$_2$	25.6 ± 3.4	-
36	4-F	2-CN	3.00 ± 2.7	16.0 ± 1.7
37	4-F	2-CF$_3$	19.4 ± 3.9	26.3 ± 1.2
38	4-NO$_2$	2-F	26.4 ± 4.8	-
39	4-NO$_2$	2-Br	32.1 ± 2.4	-
40	4-NO$_2$	2-OH	13.9 ± 4.0	29.8 ± 4.2
41	4-NO$_2$	2-CH$_3$	13.2 ± 5.8	34.9 ± 3.8
42	4-NO$_2$	2-NH$_2$	12.4 ± 1.1	23.9 ± 5.2
43	4-NO$_2$	2-CN	5.0 ± 1.9	24.4 ± 5.9
44	4-NO$_2$	2-CF$_3$	19.6 ± 3.9	40.9 ± 7.5

For the most active compound, **27**, an accurate biological evaluation was performed in an analogous way to that of **1**. A kinetic analysis allowed the determination of the K$_i$ value (8.8 ± 0.7 μM) and the classification of this compound as competitive inhibitor of MbtI. Then, additional tests were performed to verify that **27** was not a PAIN compound: the addition of BSA and Triton X-100 did not influence its IC$_{50}$, proving that the compound did not form aggregates with the target. Similarly, the addition of DTT did not impact on its

activity, showing that the ligand did not interact with the cysteine residues of the protein. To evaluate the antimycobacterial activity of **27** we used the non-pathogenic *M. bovis* BCG. The minimal inhibitory concentration value, determined in iron-limiting conditions using the chelated Sauton's medium, was slightly lower than that of the previous lead (MIC^{99} = 250 μM). Moreover, to assess if the antitubercular activity was related to iron uptake inhibition, the effects of the compound on siderophore production were evaluated by means of the Universal CAS liquid assay.[37] The test revealed that the removal of iron was inversely related to the concentration of the compound. Similarly, the quantification of mycobactins, isolated from the cell pellet, evidenced that the concentration of the siderophores decreased at higher concentrations of the compound (Figure 17). Finally, **27** was selected along with the parent compound **VS1** and **1** to be screened against human MRC-5 fibroblasts to evaluate their cytotoxicity. The biological results showed that the compounds did not affect the growth of these healthy cells (MRC-5, IC$_{50}$ > 100 μM), thus indicating the potentially low level of toxicity of this class of compounds.[38]

The same biological assays will also be performed on **36**, which provided similar results to **27** in the MbtI inhibition tests.
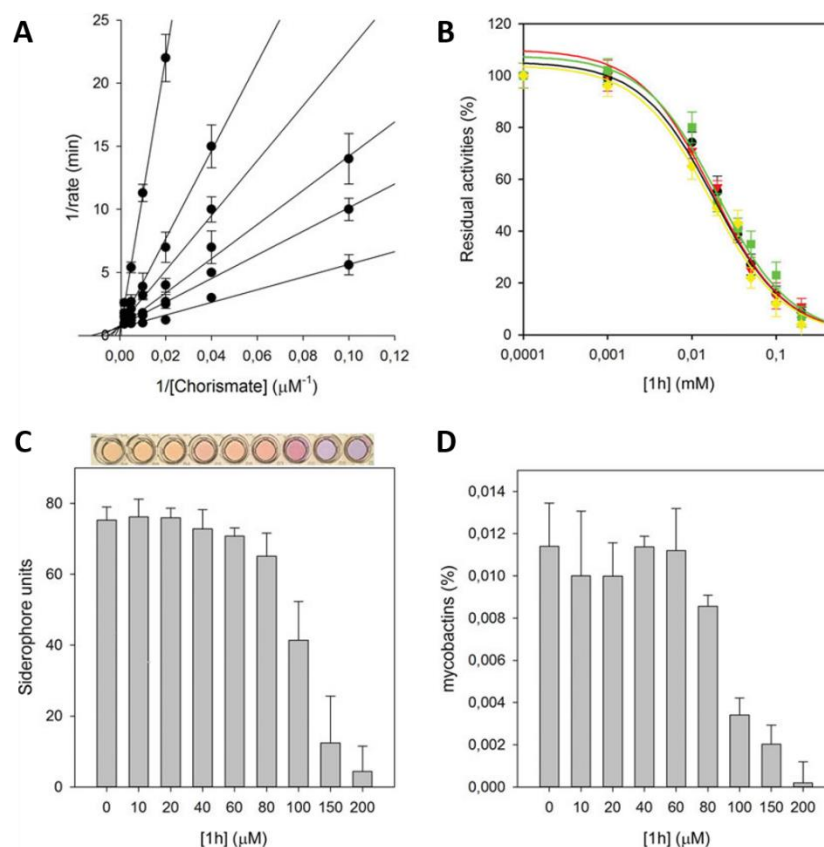
Figure 17. Biological evaluation of **27**. **A.** Global reciprocal plot of data from MbtI steady-state kinetics analysis towards chorismic acid, in the presence of different concentrations of **27**. **B.** IC_{50} plot of **27** in the presence of BSA (■), Triton X-100 (▼) and DTT (♦). **D.** Universal CAS assay performed on *M. bovis* BCG, grown in the presence of different concentrations of **27**. **E.** Determination of the mycobactins in the abovementioned cells.

3.2.3 *meta*-Substituted derivatives

Inspired by the encouraging activity of **29**, which had two trifluoromethyl groups at position 3 and 5 of the phenyl ring, we decided to investigate the hitherto unexplored *meta* position.

We began by assessing the inhibitory effect of **45**, bearing only one CF_3 at position 3; this compound showed a promising, activity, so we continued by testing the effect of other substituents at the *meta* position. The replacement with a halogen (**46**) or with electron-donating functions like the hydroxyl (**47**), methyl (**48**) or amino (**49**) groups abolished or significantly weakened the activity. The insertion of a strong electron-withdrawing group capable of forming a localised negative charge, like the NO_2 (**5**) function, did not lead to an improvement of the activity. Conversely, the amide group of **50** allowed for a significant enhancement

of the inhibition (% residual activity at 100 μM 20.9 ± 4.3). While the methylation of the amide in **51** led to a strong decrement of the activity, the substitution of the amide with classical bioisosteres allowed for a retention (sulphonamide, **52**; carboxylic acid, **53**) of the inhibitory effect. Interestingly, the substitution with a nitrile in **54** afforded the best compound of the series, characterised by a % RA of 3.1 ± 1.0 (Table 4).

Table 4. *m*-Substituted derivatives of **VS1** (**45-56**). The inhibitory effect is expressed as percentage of residual enzymatic activity (at 100 μM ligand concentration) for all compounds, and Half-Maximal Inhibitory Concentrations (IC$_{50}$, μM) for the most active candidates (residual activity ≤ 25%).
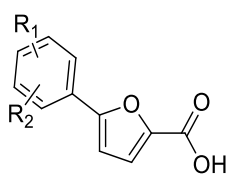
Code	R$_1$	R$_2$	% Residual activity	IC$_{50}$
45	3-CF$_3$	-	42.0 ± 6.3	-
46	3-Cl	-	101.6 ± 17.8	-
47	3-OH	-	70.4 ± 21.8	-
48	3-CH$_3$	-	103.9 ± 4.8	-
49	3-NH$_2$	-	65.8 ± 9.6	-
50	3-CONH$_2$	-	20.9 ± 4.3	31.4 ± 10.3
51	3-CONHCH$_3$	-	84.0 ± 9.1	-
52	3-SO$_2$NH$_2$	-	28.6 ± 6.8	-
53	3-COOH	-	27.2 ± 4.5	-
54	3-CN	-	3.1 ± 1.0	10.5 ± 2.3
55	3-CN	5-OCH$_3$	10.5 ± 3.9	14.5 ± 2.1
56	3-CN	5-OPh	10.5 ± 4.1	23.1 ± 2.6

As previously done with the other lead compounds, **54** underwent an in-depth biological evaluation. The IC$_{50}$ was calculated as 10.5 ± 2.3 μM, proving to be comparable or slightly better than that of the previous candidates. Additional tests were performed to ensure that the compound was not a PAIN, following the protocol described in the previous paragraphs. To evaluate the antimycobacterial activity of **54** we used the non-pathogenic *M. bovis* BCG. The minimal inhibitory concentration (MIC[99]) value was determined in iron-limiting conditions using the chelated Sauton's medium and proved to be similar to that of **27** (250 μM). To assess if the antitubercular activity was related to iron uptake inhibition, the

effects of the compound on siderophore production were evaluated by means of the Universal CAS liquid assay. The removal of iron proved to be inversely related to the concentration of the compound, and the quantification of the mycobactins, isolated from the cell pellet, evidenced that their concentration decreased at higher concentrations of the compound (Figure 18). Finally, **54** was screened against human MRC-5 fibroblasts to evaluate its cytotoxicity, revealing an $IC_{50} >$ 100 µM.

Despite the excellent inhibitory activity on the enzyme, the MIC^{99} value of **54** showed room for improvement. In the hope of enhancing its antimycobacterial activity, we decided to introduce lipophilic groups at position 5 of the phenyl ring. In order to test the effect of substituents endowed with different steric hindrance, we chose a methyl (**55**) and a phenyl (**56**) group, linked to the main phenylfuran scaffold through an oxygen bridge. Both the compounds retained the inhibitory activity on MbtI, showing substantially unvaried values with respect to the parent compound (**54**). The evaluation of their antimycobacterial effect was performed in the previously described conditions. Interestingly, while **55** did not show an improvement of the MIC^{99} value, with **56** we effectively managed to halve the previous result, obtaining an $MIC^{99} = 125$ µM.
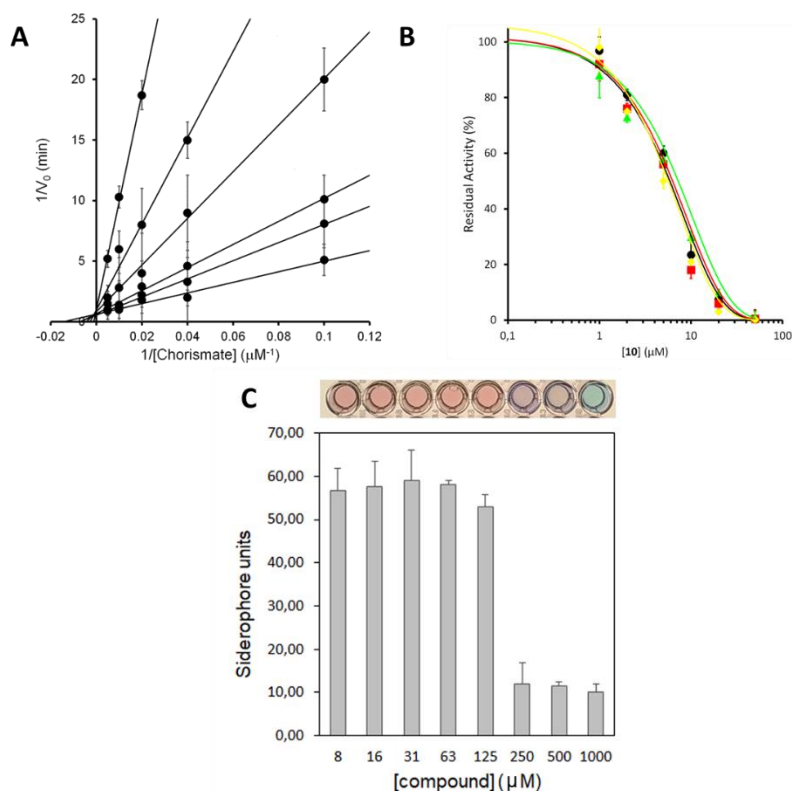
Figure 18. Biological evaluation of **54**. **A.** Global reciprocal plot of data from MbtI steady-state kinetics analysis towards chorismic acid, in the presence of different concentrations of **54**. **B.** IC$_{50}$ plot of **54** in the presence of BSA (■), Triton X-100 (▼) and DTT (◆). **D.** Universal CAS assay performed on *M. bovis* BCG, grown in the presence of different concentrations of **54**.

3.2.4 Heterocyclic derivatives

The chemical exploration of the scaffold of **VS1** was completed with the analysis of the role of the furan. The importance of the central heterocycle was confirmed by the lack of activity displayed by **57**, an analogue of **VS1** bearing the carboxylic function directly attached to the benzene. The substitution of the furan with a phenyl led to synthesis of the biphenyl conger of **1**, **58**, which exhibited a significant drop in the inhibitory effect, suggesting the importance of the furan in determining the activity. Therefore, we decided to test the effect of different five-membered heterocycles, using as substituents on the phenyl ring the best groups of the previous series, namely the *p*-NO$_2$ and *m*-CN moieties. We chose the thiophene (**59,60**) as a classical isostere of the furan, and then we investigated the introduction of a nitrogen at position 3 with the oxazole (**61,62**), the thiazole (**63,64**) and the imidazole (**65,66**). Despite all of the compounds retained a certain inhibitory activity on the enzyme, the introduction of the nitrogen atoms

did not improve the effect, probably because they did not establish any significant interaction in the active site. Similarly, the substitution of the oxygen at position 1 did not yield good results (Table 5).

In conclusion, none of the compounds belonging to this series allowed for an improvement of the inhibitory activity on MbtI compared to the respective parent compounds.

Table 5. Heterocyclic derivatives of **VS1** (**57**-**66**). The inhibitory effect is expressed as percentage of residual enzymatic activity (at 100 µM ligand concentration) for all compounds, and Half-Maximal Inhibitory Concentrations (IC$_{50}$, µM) for the most active candidates (residual activity ≤ 25%).
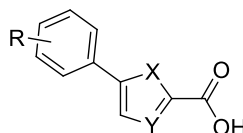
Code	X	Y	R	% Residual activity	IC$_{50}$
57	-	-	-	89.3 ± 16.1	-
58	-	-	-	36.8 ± 7.6	-
59	S	-	4-NO$_2$	23.0 ± 7.8	21.1 ± 2.7
60	S	-	3-CN	23.7 ± 5.0	35.5 ± 1.9
61	O	N	4-NO$_2$	22.5 ± 10.8	18.6 ± 1.7
62	O	N	3-CN	38.0 ± 7.6	-
63	S	N	4-NO$_2$	32.8 ± 2.3	-
64	S	N	3-CN	25.8 ± 1.8	-
65	NH	N	4-NO$_2$	25.6 ± 9.3	-
66	NH	N	3-CN	25.3 ± 1.2	-

3.3   SYNTHETIC CHEMISTRY

All the compounds presented in the previous paragraphs were synthesised according to one of the following general procedures. Compound **49** was synthesised with a different approach, reported in Material and Methods. Compounds **VS1-5**, **1**, **5**, **45**, **57** were acquired from Enamine and Sigma Aldrich at the highest purity level available (≥ 95%).

3.3.1 Synthesis of 2-4,6-8,10-14,17,18,20,27-29,46,47,50,54,58-60
(General Scheme A)
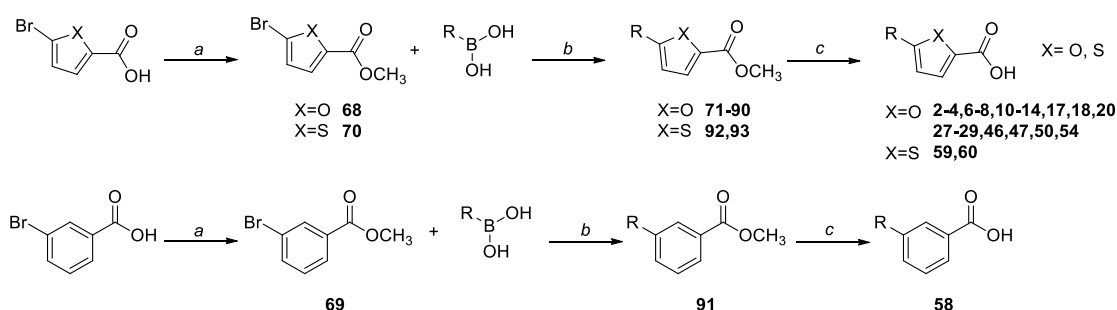
Reagents and conditions: *a)* conc. H_2SO_4, MeOH, reflux, 24 h; *b)* Pd(PPh$_3$)$_2$Cl$_2$, 2M Na$_2$CO$_3$, 1,4-dioxane, 90°C, overnight, N$_2$; *c)* NaOH, EtOH/THF 1:1, reflux, 5 h *or* LiOH·H$_2$O, THF/H$_2$O 1:1, 20°C, 2 h.

The starting carboxylic acids were converted to the corresponding methyl esters through a Fischer-Speier esterification. The so-obtained intermediates were reacted with the suitable boronic acids in a palladium-catalysed Suzuki-Miyaura coupling, using bis(triphenylphosphine)palladium dichloride and sodium carbonate, to afford the corresponding bicyclic derivatives. The resulting compounds were hydrolysed to the corresponding acids under basic conditions.[33,38]

3.3.2 Synthesis of 9,15,16,19,21,23-26,30,31,33-37,40-42,44,48,51-53,55,56 (General Scheme B)
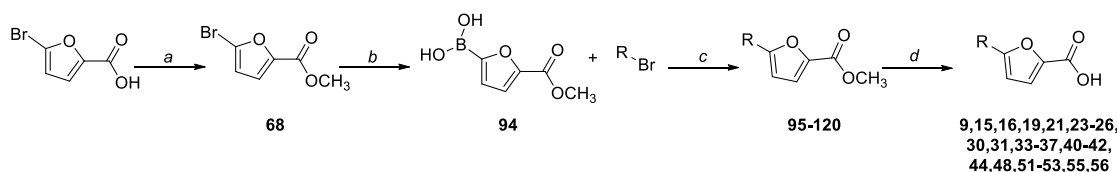
Reagents and conditions: *a)* conc. H_2SO_4, MeOH, reflux, 24 h; *b)* 1. bis[2-(*N,N*-dimethylamino)ethyl] ether, 2M *i*-PrMgCl, THF, 20 min. 10-15°C – 30 min. r.t., N_2; 2. $B(OCH_3)_3$, 0°C, 10 min., N_2; *c)* Pd(PPh$_3$)$_2$Cl$_2$, 2M Na$_2$CO$_3$, 1,4-dioxane, 60°C, 90 min., MW, N_2; *d)* NaOH, EtOH/THF 1:1, reflux, 5 h *or* LiOH·H$_2$O, THF/H$_2$O 1:1, 20°C, 2 h.

Methyl 5-bromofuran-2-carboxylate was obtained through a Fischer-Speier esterification of the commercially available 5-bromofuran-2-carboxylic acid. The corresponding boronic acid was synthesized *via* a multi-step reaction, involving isopropylmagnesium chloride, bis[2-(*N,N*-dimethylamino)ethyl]ether and trimethyl borate. The so-obtained intermediate was reacted with the suitable bromo derivative in a microwave-assisted Suzuki-Miyaura reaction, using bis(triphenylphosphine)palladium dichloride and sodium carbonate. The ester group of the resulting intermediates was hydrolysed to the corresponding acid under basic condition.[33,38]

3.3.3 Synthesis of 22,32,38,39,43 (General Scheme C)
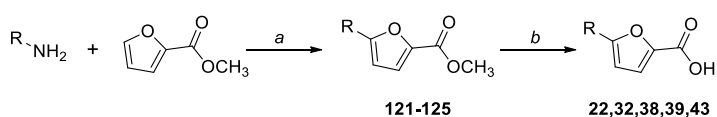
Reagents and conditions: *a)* 1. NaNO$_2$, conc. HCl, 30 min, 0°C; *2*. CuCl$_2$, CH$_3$COCH$_3$/H$_2$O, 2-3 h; *b)* NaOH, EtOH/THF 1:1, 5 h, reflux.

The key ester intermediate was obtained through a Sandmeyer reaction, involving the commercially available methyl furan-2-carboxylate and the suitable aniline derivative. First, the diazotization was conducted, using sodium nitrite in a hydrogen chloride solution; then, the coupling with the furan was performed in the presence of copper (II) chloride in an acetone-water mixture. The obtained

bicyclic compound was then hydrolysed to the corresponding acid, under basic conditions.[38]

3.3.4 Synthesis of 61,62 (General Scheme D)

126,127   **61,62**
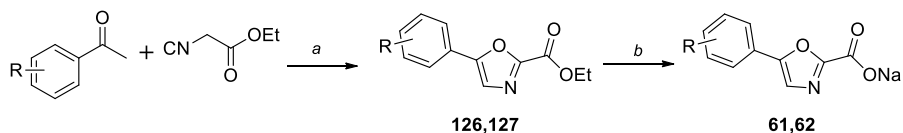
Reagents and conditions: *a*) I$_2$, DMSO, 3 h, 130°C; *b*) NaOH, THF/H$_2$O 1:1, 1.5 h, r.t.

The key ester intermediates were obtained through a one-step mechanism based on an iodine-promoted formal [3+2] cycloaddition of the suitable acetophenone with ethyl isocyanoacetate. The so-obtained 2,5-disubstituted oxazole esters were then hydrolysed to the corresponding acids under basic conditions and isolated as sodium salts for stability reasons.[39]

3.3.5 Synthesis of 63,64 (General Scheme E)
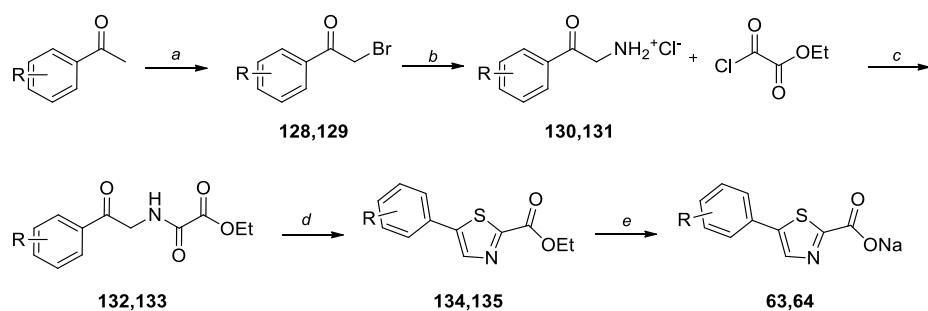
128,129   **130,131**

132,133   **134,135**   **63,64**

Reagents and conditions: *a*) NBS, *p*-TsOH, CH$_2$Cl$_2$, overnight, r.t, N$_2$ atm.; *b*) *1*. hexamine, CH$_2$Cl$_2$, 8 h, r.t.; *2*. conc. HCl, EtOH, overnight, r.t.; *c*) TEA, EtOAc, 3 h, reflux; *d*) Lawesson's reagent, 1,4-dioxane, 2 h, reflux; *v*) NaOH, THF/H$_2$O 1:1, 1.5 h, r.t.

The monobrominated derivatives of the commercially available acetophenones were obtained using NBS and *p*-toluenesulfonic acid. Then, the corresponding amine hydrochloride salts were synthesised through the Delépine reaction, which comprises an amination with hexamine, generating the α-aminoacetophenone, followed by the addition of concentrated hydrochloric acid.

Ethyl-chlorooxoacetate was added to a solution of the hydrochloride salts in ethyl acetate to afford the corresponding amides. Then, the Lawesson's reagent was used for the ring closure reaction, to yield the corresponding ethyl 5-phenylthiazole-2-carboxylates, which were hydrolysed under basic conditions and isolated as sodium salts.[40]

3.3.6 Synthesis of 65,66 (General Scheme F)
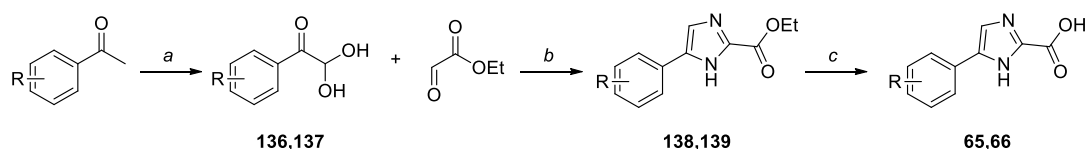
Reagents and conditions: *a*) SeO$_2$, 1,4-dioxane, H$_2$O, 7 h, reflux, N$_2$; *b*) NH$_4$OAc, CH$_3$CN, H$_2$O, 4 h, r.t.; *c*) NaOH, H$_2$O/EtOH 1:1, 5 h, reflux.

The commercially available acetophenones were treated with selenium oxide in 1,4-dioxane and water to afford the corresponding geminal diols, which were reacted with ethyl glyoxylate and ammonium acetate to form the imidazole esters. These key intermediates were hydrolysed in basic conditions to the corresponding acids.[41,42]

3.4 X-RAY CRYSTAL STRUCTURES

With the aim to validate the docking hypothesis and to shed light on the binding mode of selected compounds with MbtI, we designed co-crystallisation experiments, as a support for the development of more active candidates.

In medicinal chemistry, co-crystallisation is generally applied to the characterisation of complexes between macromolecules and ligands, providing a "snapshot" of their real interaction pattern and thus allowing for an in-depth analysis of their binding. As such, crystallisation can be effectively integrated in a drug discovery pipeline, allowing for a rational structure-based design of improved ligands. Once the best candidates have been identified, the careful study of their binding mode provides invaluable information to optimise computational models, and thus drive further synthetic efforts.

With this aim in mind, we approached the crystallisation of MbtI. The obtainment of co-crystal structures was of pivotal importance, considering both the limited size of the compounds and the high plasticity of the active site, as reported by Chi, Bulloch and co-workers.[28]

In the literature, a total of nine structures of the protein are available: three of them are crystals of the enzyme in its apo form (PDB codes: 2I6Y, 2G5F, 3LOG), and six are complexes with derivatives of the isochorismate analogue Methyl-AMT, differently substituted at the double bond of the enolpyruvyl side chain (PDB codes: 3ST6, 3VEH, 3RV6, 3RV7, 3RV8, 3RV9).

The molecules for the co-crystallisation experiments were selected within the available pool of MbtI ligands, to meet specific purposes. First, we decided to reproduce one of the co-crystal structures available in the literature, in order to validate both the protein production process and the crystallisation experiments. Among the available compounds, we selected Methyl-AMT, which we had already synthesised according to the reported procedure, to act as a reference compounds in the biological tests.[24] Then, we attempted to obtain a structure with chorismate, the natural substrate of the enzyme, despite we knew from preliminary stability tests, carried out with ^{1}H-NMR (data not shown), that the compound was significantly unstable in solution even at low temperature (4°C). Finally, we selected among our compounds **1**, **27** and **54**, considering both their

strong inhibitory potency and the fact that they belonged to structurally different chemical series.

Preliminary crystallisation experiments were carried out in 24-well plates, using the vapour diffusion technique, applied to both the hanging-drop and sitting-drop methods. Different parameters were varied, including the protein concentration (5-20 mg/mL), the temperature (4° or 18°C) and the reservoir composition. The tested crystallisation conditions were taken from the literature, and carefully modified to explore the effect of different variations on the crystallisation process.[22,28] A detailed account of the assayed conditions is reported in the Materials and Methods Section.

Although these experiments did not lead to the expected results, they were instrumental to perform an exploratory assessment of the crystallisation of MbtI. We managed to conclude that the protein tended to form crystals at a concentration of 20 mg/mL and only at 4°C. As for the reservoir composition, a deeper optimisation process proved to be necessary, because of the low quality of the obtained crystals. The few hits achieved with these experiments were needle-shaped and diffracted poorly, with resolutions in the range 2.8-3.8 Å; the best sample showed an orthorhombic system with a $P2_12_12_1$ space group (cell dimensions: a: 51.8, b: 164.2, c: 195.3 Å), which was analogous to that obtained by Harrison, Lott $et\ al.$ (PDB code: 2G5F).[22] No bound ligand was detected in any of the tested crystals.

In light of the limited success attained in these experiments, we decided to design an optimisation process aimed to obtain high-quality crystals with the embedded ligands. These new investigations were carried out at the Pasteur Institute in Paris, in the Structural Microbiology Unit directed by Prof. Pedro Alzari, under the supervision of Dr. Marco Bellinzoni.

Considering that the amount of the protein had been a limiting factor in the preliminary crystallisation studies, we decided to attempt to improve the expression and purification processes. To this end, we used a new codon-optimised clone, obtained from GenScript (GenScript, Piscataway, NJ, USA). Codon optimisation is a novel technique, designed to enhance the translational efficiency of a target gene in an expression vector; it is based on the concept of

codon redundancy, according to which the same amino acid can be encoded by a different set of three genes. Considering that different organisms exhibit bias towards using certain codons over others, this procedure allows to design a clone, specifically optimised for the recombinant expression in a particular organism, without modifying the amino acid sequence of the desired protein. With this approach, we obtained a new pET-28a-MbtI plasmid for the transformation of *E. coli* BL21 cells. Moreover, we also decided to adopt the autoinduction method over the traditional IPTG-mediated expression; this approach has been shown to be more versatile than the classical one, allowing for superior OD_{600} values and increased protein yields.[43] Finally, we decided to perform a buffer optimisation study, in order to improve the stability of the protein in its storage conditions. The implementation of these modifications in the protein production and purification procedures allowed for a significant improvement of the yield, thereby increasing the number of possible crystallisation trials.

In order to find the best crystallisation conditions, a *de novo* full screening was performed, using commercial 96-well-plate kits with the sitting drop vapour diffusion method. Overall, we tested 672 different reservoir solutions, with different protein samples. The concentration of the protein solution was kept around 20 mg/mL and the experiments were performed at 4°C, because these conditions had been identified as the best in the previous studies. The same ligands, formerly selected for the preliminary studies, were used in these full screenings, in the presence or absence of magnesium chloride.

With these new experiments, we managed to obtain high-quality crystals; the best candidates were fished and analysed at the ESRF or SOLEIL synchrotron facilities. The diffraction data were solved with the Molecular Replacement technique, using as models the most suitable MbtI structures from the PDB. Overall, 40 datasets were acquired: the best ones were selected for further processing, based on the resolution and on the presence of a bound ligand.

The structures of some of the best samples will be discussed in the following sections. A detailed account of all obtained datasets is provided in the Materials and Methods Section.

3.4.1 Structure of MbtI with Methyl-AMT

As previously stated, the obtainment of a co-crystal with Methyl-AMT had the aim to validate the protein production process and the crystallisation experiments. In this context, we succeeded in achieving a new structure, characterised by a different crystalline form, also improving the quality of the data. With its 1.65 Å, it became the highest-resolution structure of MbtI available in the literature.
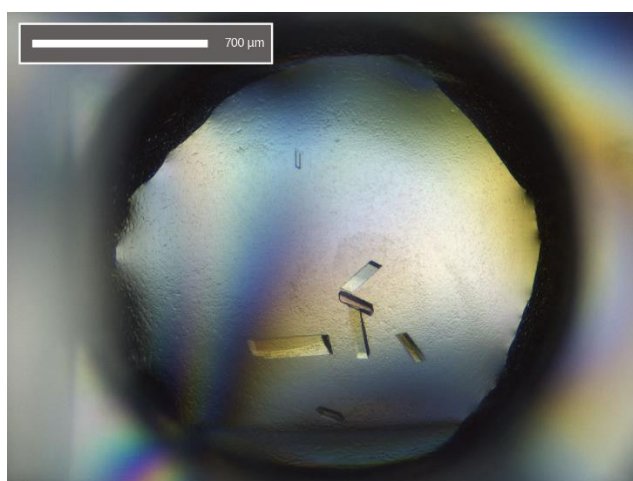
Figure 19. Image of the drop containing the crystals of MbtI in complex with Methyl-AMT (16% PEG 3350, 0.1 M sodium citrate pH 5.6, 2% tacsimate).

The prism-shaped crystals grew in the presence of 16% PEG 3350, 0.1 M sodium citrate pH 5.6, 2% tacsimate, and were cryoprotected using a solution of 25% glycerol (Figure 19). The diffraction data, acquired at the SOLEIL Synchrotron (Saint-Aubin, France) on the PROXIMA-2A beamline, were indexed to the monoclinic $P2_1$ system, with the following cell parameters: a: 95.607 Å, b: 83.064 Å, c: 126.252 Å, β: 111.670°. The structure was solved with the Molecular Replacement method, using as model the structure deposited by Chi, Bulloch and co-workers (PDB code: 3VEH).[28] The refinement was carried out with Buster,[44] until an R = 0.189 and R_{free} = 0.213. The summary of the crystal data and the refinement statistics are reported in Materials and Methods (Table 7).

The crystal presents four molecules in the asymmetric unit; chains A, C and D show a lack of electron density between residues 273-285 and 327-332, corresponding to the mobile loops of the protein. Interestingly, chain B displays a complete electron density map in those same regions. Although the amino acid sequence of chains A, C and D could not be completed, all four molecules seem to assume an open conformation. As described in the introduction section, MbtI is

known to exist in two conformations, depending on the position of its mobile loops. The movement of these regions is known to affect the accessibility to the active site, but the causative factor influencing this shift is yet to be identified. The corresponding structure obtained by Chi, Bulloch and co-workers (PDB code: 3VEH) showed a closed conformation, suggesting that the ligand does not play a significant role in this regard.[28] This conclusion seems to be confirmed by the fact that two of the three available structures of the apo protein show a different conformational state (2G5F: open; 3LOG: closed). Despite the authors of the previously mentioned structure observed an increase of flexibility with sterically hindered ligands, it is also clear that the inhibitor cannot determine the position of the loops (Figure 20).
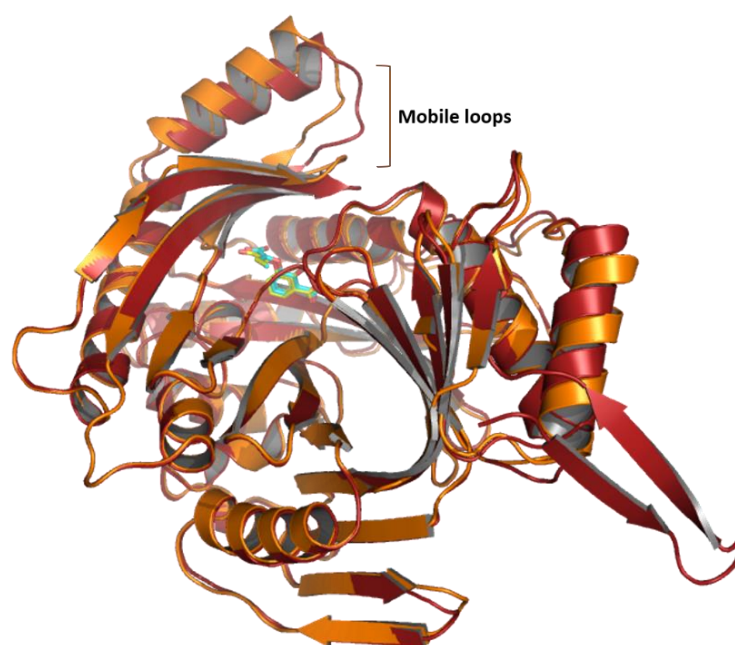
Figure 20. Superimposition of our crystal structure (orange, open) with 3VEH (red, closed). The two mobile loops (273-285 at the top, 327-332 at the bottom) determine the conformation of the protein. Inside the active site, Methyl-AMT is represented as yellow sticks for 3VEH and cyan sticks for our structure.

The electron density for the ligand is strong in all four chains, and its interaction mode inside the active site is consistent with the reported data. The carboxylic group linked to the phenyl ring of Methyl-AMT forms H-bonds with Tyr385, the backbone amide of Gly419, and an ordered water molecule. Another water molecule mediates the hydrogen bonding of Arg405 and His334 with the 2-hydroxyl and enol oxygens of the ligand, respectively. The phenyl ring is

sandwiched between the hydrophobic side chains of Leu404 and Lys438. A cation-π interaction occurs with the amine group of Lys438, which also establishes a hydrogen bond with the enolpyruvyl carboxylate of the molecule; the latter forms a further hydrogen bond with the hydroxyl group of Thr271 (Figure 21).[28]
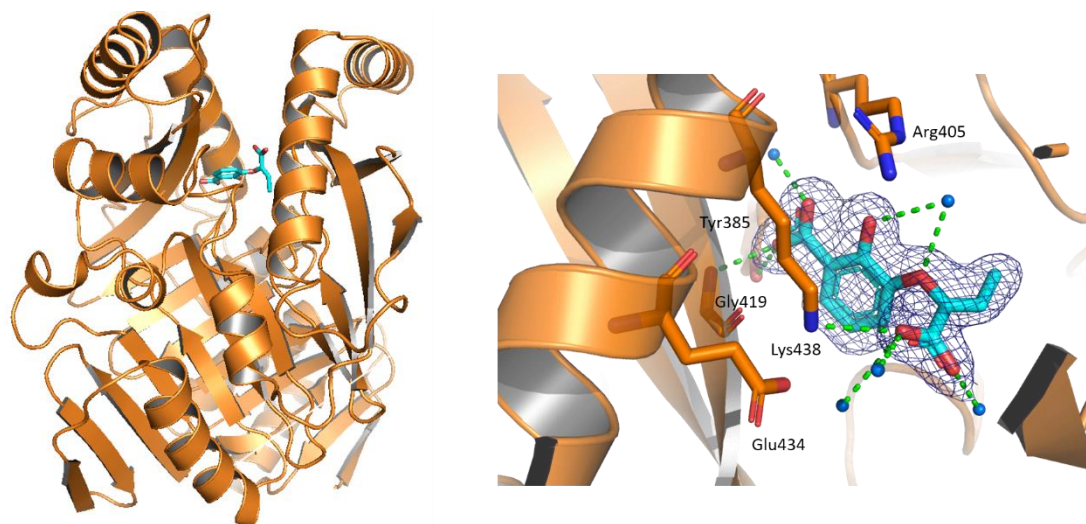
Figure 21. Ribbon diagram of the MbtI-Methyl-AMT structure. On the left, a whole view of the protein, showing the ligand in the active site. On the right, the interactions (green, dashed lines) of Methyl-AMT with the side chains (in sticks) and the water molecules (blue spheres). The blue mesh represents the electron density around the ligand (scaled at 1σ).

3.4.2 Structure of MbtI with 54

In the context of our attempts to gain a co-crystal structure of MbtI with active compounds bearing the 5-phenyfuran-2-carboxylic acid scaffold, we managed to obtain reproducible high-quality crystals with **54**, one of the most potent inhibitors in our library, bearing a cyano moiety at the *meta* position.
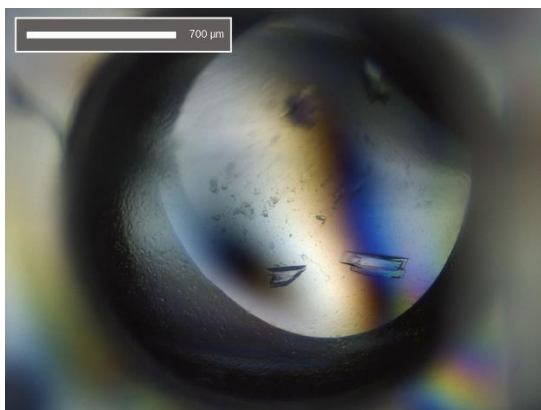
Figure 22. Image of the drop containing the crystals of MbtI in complex with **54** (20% PEG 3350, 0.2 M sodium tartrate).

The crystals grew in the presence of 20% PEG 3350 and 0.2 M solutions of different sodium salts; the best hit was obtained with sodium tartrate, and the cryoprotectant was parathon oil (Figure 22). The diffraction data, acquired at the SOLEIL Synchrotron (Saint-Aubin, France) on the PROXIMA-1 beamline, were indexed to the monoclinic P2$_1$ system, with the following cell parameters: a: 88.336 Å, b: 111.687 Å, c: 94.998 Å, β: 92.670°. The structure was solved with the Molecular Replacement method, using as model the structure having PDB code 3RV7.[28] The refinement was carried out with Buster,[44] reaching an R = 0.210 and R$_{free}$ = 0.238, using data up to 2.09 Å. The summary of the crystal data and the refinement statistics are reported in Materials and Methods (Table 7).

The crystal structure presents four molecules in the asymmetric unit; as in the previously described structure, the electron density between residues 270-292 and 327-335 is weak or sometimes absent. The position of the loop 327-335 appears to be more stable in chains C and D, where the interpretable density map allows for the fitting of the region. The conformation of the protein seems to be open in all four molecules, being superimposable with that of the apo form obtained by Harrison *et al.* (2G5F).[22] The N-terminal region of chain B is extremely well-defined with a clear electron density for all residues.

The electron density for the ligand is detectable in all the four chains, with different intensity: the map is well-defined in chains D and B, while in A and C the density of the phenyl ring is quite weak, even though it is clearer around the cyano moiety. Regarding the interactions, the carboxylic group forms H-bonds with Tyr385, Arg405 and an ordered water molecule; the oxygen of the furan interacts with Arg405, while the phenyl ring forms a cation-π interaction with

Lys438 and a Van der Waals contact with Thr361. Finally, the cyano group forms a contact with Lys205, a key amino acid involved in the first step of the catalytic reaction of MbtI (Figure 23).
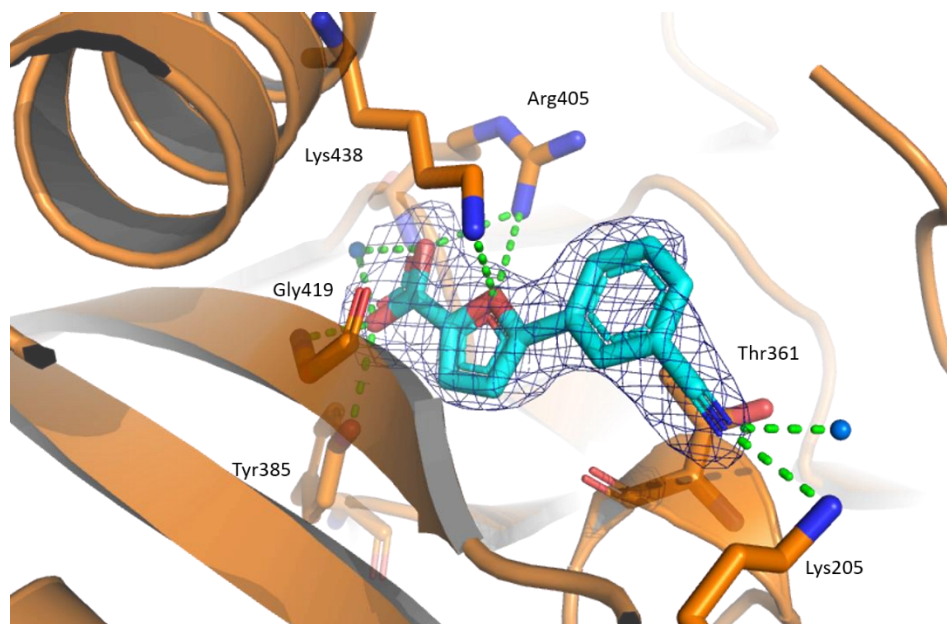
Figure 23. Ribbon diagram of the MbtI-**54** structure: the interactions of the ligand with the side chains (in sticks) and the water molecules (blue spheres) are represented as green, dashed lines. The blue mesh represents the electron density around the ligand (scaled at 1σ).

The analysis of the binding mode of the compound revealed a different spatial disposition with respect to the pose predicted by the computational simulations for this class of compounds. The superimposition of the model and the structure shows the two ligand molecules in a flipped relative position, with two cycles lying on roughly perpendicular planes, and the others in a parallel disposition (Figure 24). The reason of this discrepancy is tightly related to the role of the Mg^{2+} ion. The computational simulations were run based on the assumption that the chelation of the co-factor was essential to ensure the inhibitory activity; consequently, the proposed binding between the carboxylic acid of the molecule and the metal ion becomes the main determinant of the overall orientation of the ligand. The absence of the ion in the structure, despite the presence of 5 mM $MgCl_2$ in the crystallisation condition, seems to affect the binding mode of the ligand, leading the adoption of a different orientation with respect to that indicated by the computational studies. As a consequence, the

carboxylic function forms H-bonds with two amino acid residues of the active site and an ordered water molecule.
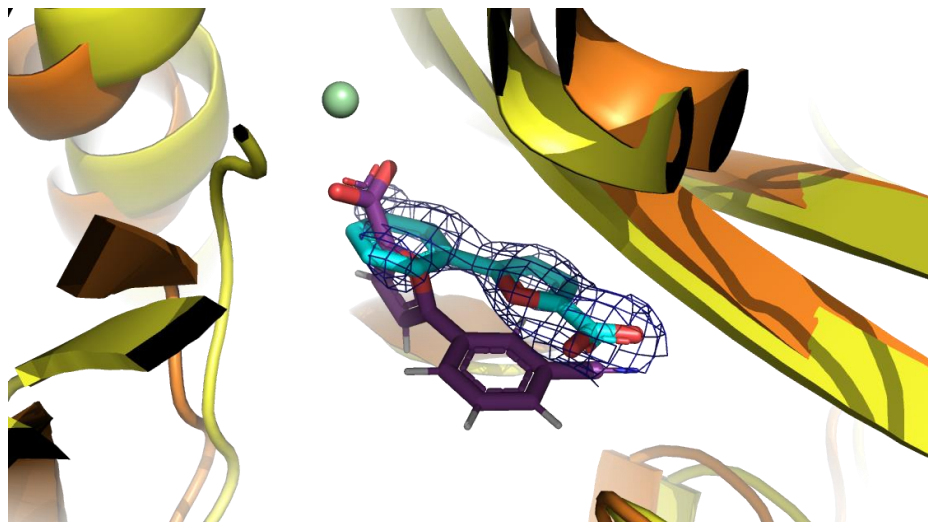
Figure 24. Superimposition of the crystallographic structure of MbtI-**54** (ligand in cyan) *versus* the computational pose (ligand in purple). The green sphere represents the Mg^{2+} ion, while the blue mesh shows the electron density around the ligand (scaled at 1σ).

A similar dichotomy was observed by Chi, Bulloch and co-workers, who obtained two binding modes for their compounds. The first one, shown only by their weakest inhibitor, is hypothetically that of the natural substrate and reproduces the pose predicted by our computational simulations. A second pose was detected for all the other molecules and it is similar to that of our crystal structure.[28]

Therefore, the difference between the computational model and the crystal structure prompted us to deeply investigate the role of the Mg^{2+} ion in the catalytic activity of the enzyme in order to rationalise our results.

Although the involvement of this metal co-factor in the reactions catalysed by MbtI has been the subject of a number of publications,[21,24] its role remains elusive. MbtI is known to be a Mg^{2+}-dependent enzyme, but Ferrer *et al.* also reported on its ability to act as a chorismate mutase in the absence of the metal, leading to the synthesis of prephenate.[26] Ziebart and Toney ascribed this promiscuous activity to an incomplete purification of MbtI,[27] but Ferrer and co-workers further supported their original hypothesis with computational studies, showing the plausibility of their interpretation.[25]

Moreover, the obtainment of a crystal structure of the protein with the co-factor is a challenging task. The only structure of MbtI showing a Mg^{2+} ion in the active site was obtained by Chi, Bulloch and co-workers (3RV6); however, as stated by the authors, the binding is likely the result of an artefact, considering that the ion is coordinated only by water molecules, which is hardly a physiological condition.[28] Furthermore, the inherent difficulty in the obtainment of a crystal structure with the metal would seem counterintuitive, considering that the Mg^{2+} clearly plays such a pivotal role in the catalytic activity of MbtI. Despite our numerous attempts in this direction, we encountered the same issues: different magnesium salts, used at typical binding concentrations (2-5 mM), did not lead to the expected results.

3.4.3 Investigations around the Mg^{2+} ion

In the attempt to shed light on the role of the magnesium and to rationalise our results, we decided to expand our literature research to other Chorismate Utilising Enzymes (CUEs), because the available data regarding MbtI were limited and inconclusive.

In this context, we examined the work of Meneely, Lamb *et al.*; they proposed a very interesting interpretation of the role of the Mg^{2+} ion in this class of enzymes, based on crystallographic and biological evidences.[45] They suggested a model in which, in the presence of normal iron concentrations, ferrous ions bind to the catalytic metal site of the enzyme with high affinity, thereby occluding the access to the active site in a negative-feedback fashion. In iron limiting conditions, the absence of the ferrous ions determines an increased affinity of the enzyme for the natural substrate (chorismate) and the inhibitors. In this situation, the affinity for the metal ion is rather low, so it is conceivable that the ligand is the first driving element of the catalytic reaction. Once the magnesium binds to the active site, a conformational shift determines the occlusion of the binding region. Interestingly, after the conversion of chorismate to isochorismate, a sudden change in the affinity for the metal occurs: the enzyme tightly binds magnesium, allowing the reaction to quickly evolve towards the formation of salicylate.[45]

This theory seemed to provide valid answers to most of the unexplained observations regarding MbtI. For example, the low affinity for the metal could be the cause for the absence of magnesium inside the crystal structures of MbtI. Moreover, it also provided a reasonable interpretation of the favoured binding mode of our compounds. In light of these data, and bearing in mind the similar results obtained by Chi, Bulloch and co-workers,[28] we hypothesised that the magnesium cannot determine the binding mode of the ligand: therefore, the pose observed in the crystallographic structure should actually be the physiological one. As a consequence, the simulated binding mode, in which the co-factor is the determinant of the molecule orientation, must be an artefact.

In order to acquire more data and support the applicability of this hypothesis to MbtI, we designed a series of experiments. We evaluated the activity of the enzyme and its inhibition at different concentrations of magnesium, and we made additional attempts to obtain a crystal structure with the metal ion. Finally, we designed a compound unable to chelate the magnesium, to verify if it could still inhibit the enzyme, thus verifying the crystallographic binding mode.

3.4.3.1 Experiments at different Mg^{2+} concentrations

The activity of MbtI was tested at different concentrations of the metal ion. Our attempts to probe the enzymatic activity in the absence of magnesium revealed that MbtI is completely inactive without its metal co-factor. This confirms that the ion is essential for the enzyme to exert its activity, and, at the same time, corroborates the hypothesis that the supposed chorismate mutase effect is likely due to a faulty purification of the enzyme.

Moreover, in the presence of the inhibitor, the performance of the compound was slightly hindered by very high concentrations of the metal. This observation seemed to support our previous hypothesis: at moderate concentrations of magnesium, the ligand preferentially binds (and inhibits) MbtI and the metal has a low affinity for the enzyme. In particular, when the concentration of the metal reaches very high levels, part of the ions binds to the enzyme, thereby inducing the shift of the mobile loops and thus occluding the

access of the inhibitor to the active site. These data might also indicate that the compound could prevent the binding of the magnesium in a competitive fashion, becoming slightly less effective when the levels of the metal reach higher values (Figure 25).

[Mg^{2+}] (mM)	IC$_{50}$ (μM)
1	9.8 ± 1.3
5	11.4 ± 2.0
25	14.2 ± 1.1
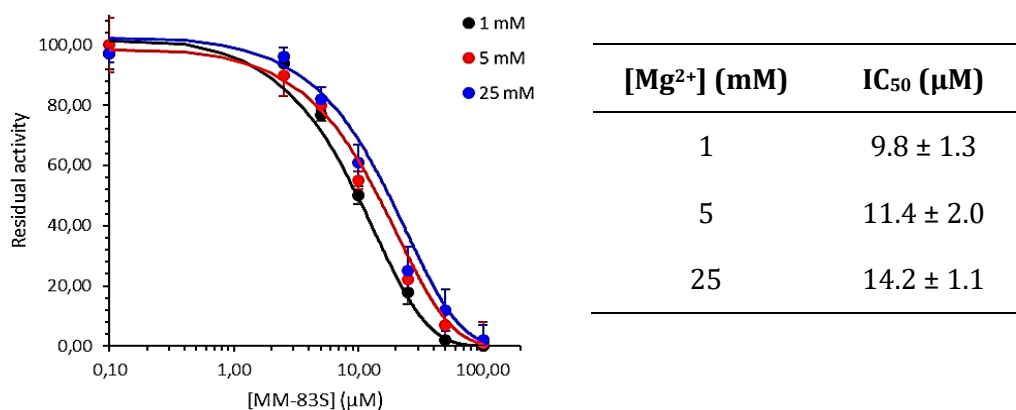
Figure 25. On the left, a graph comparing the profiles of the IC$_{50}$ of **54**, tested at different concentrations of Mg^{2+}. On the right, a table showing the IC$_{50}$ values at the assayed concentrations.

3.4.3.2 Structure of MbtI with the Mg^{2+} ion

In light of the previously described mechanism of action of the enzyme, we decided to adjust our crystallisation trials in order to attempt to obtain a co-crystal structure of MbtI with the magnesium ion. To this end, we tested very high concentrations of the co-factor, using MgCl$_2$ and MgSO$_4$ between 1 and 1.75 M. With this expedient, we aimed at counteracting the low affinity of the enzyme for the magnesium.
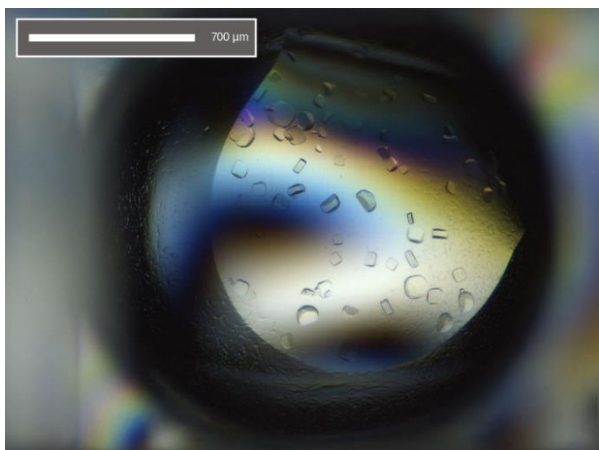
Figure 26. Image of the drop containing the crystals of MbtI in complex with the Mg^{2+} cofactor (1.75 M $MgSO_4$, 0.1 M MES pH 6.5).

Interestingly, we managed to obtain some well-shaped crystals in the presence of 1.75 M $MgSO_4$ and 0.1 M MES pH 6.5. The samples were cryoprotected using parathon oil (Figure 26). The diffraction data, acquired at the SOLEIL Synchrotron (Saint-Aubin, France) on the PROXIMA-2A beamline, were indexed to the tetragonal I422 system, with the following cell parameters: $a=b$: 193.970 Å, c: 257.030 Å. The structure was solved with the Molecular Replacement method, and the refinement was carried out with Buster,[44] until an R = 0.184 and R_{free} = 0.208, using data up to 2.11 Å. The summary of the crystal data and the refinement statistics are reported in Materials and Methods (Table 7).

Being a completely new crystalline form, an intensive refitting and rebuilding process was necessary to solve the structure. The crystal presents four molecules in the asymmetric unit, and, differently from the previous crystals, the electron density is extremely well-defined in all regions; this is probably due to the high stability of the mobile loops, that are tightened to produce a closed conformation. The analysis of the active site evidenced the clear presence of an electron density for the magnesium ion, though its coordination is slightly different in the various chains. In chain A, the magnesium interacts with the oxygen of Gly421, with four ordered water molecules, and with a sulphate anion. Through two of the water molecules, it also contacts Glu431 and Glu434. Chain C does not show the presence of the ion; in its absence, Glu431 binds a water molecule, and there is no clear electron density for the side chain of Glu434. The coordination pattern of magnesium in chains B and D is analogous: the ion directly interacts with Glu297 and Glu434 and two ordered water molecules, contacting through them Glu294 and Glu431. In addition to that, the metal forms

a strong interaction with another molecule, surrounded by a strong electron density. By analogy with the crystal of Irp9 from *Yersinia enterocolitica* (PDB code: 2FN1) the molecule was unambiguously identified as a salicylate.[23] The electron density map of the ligand is perfectly superimposable to that of the Irp9 structure, including the weak negative density corresponding to the hydroxyl group. The density is particularly strong in chain D, while in chain B it is slightly weaker, despite remaining clear. The salicylate coordinates the magnesium ion with its carboxylic moiety, which could also form additional H-bonds with the peptide backbone through Gly270 and Gly421, and with the side chain of Thr271. The hydroxyl group and the phenyl ring do not seem to form significant interactions (Figure 27).

The presence of bound salicylate in the structure was a remarkable result, especially considering that the natural substrate was not added to the crystallisation solution. This clearly suggests that the enzyme is active in the recombinant organism (*E. coli*) and that either chorismate or salicylate is co-purified with the protein. The activity of MbtI in the expression vector is not surprising, considering that chorismate is a fundamental intermediate for the synthesis of all aromatic compounds in bacteria, and it surely is abundant in *E. coli* too. Moreover, it was not the first time that a product of the catalytic reaction was identified in a crystal structure of MbtI. In 2006, Harrison, Lott and co-workers obtained the first structure of the enzyme, observing the presence of pyruvate in the active site; this finding prompted the authors to classify MbtI as a salicylate synthase, as pyruvate is the by-product of the transformation of chorismate.[22] Unsurprisingly, the position that was occupied by the pyruvate in the literature structure is taken by a sulphate ion, which can probably form tighter interactions with the surrounding amino acids. Notably, our structure represents the first evidence of the native binding mode of the ligand inside the active site of MbtI. Thanks to these results, the previously hypothesised conformation is now confirmed by our crystallographic data, which showed the product of the catalytic reaction, bound to the metal co-factor. We managed to observe the presence of the salicylate only thanks to the high concentration of magnesium, which stabilised the bound product. The closed conformation,

determined by the presence of magnesium, contributed to prevent the loss of the salicylate from the active site.

We confirmed that the magnesium induces the closed conformation of the protein by the obtainment of another structure, showing an embedded barium ion. This divalent ion occupies the same position of the magnesium and forms a similar pattern of interactions with the surrounding amino acids, with slight modifications due to the different steric hindrance. The presence of a metal binding site in that region of space was also confirmed by a third structure, containing a calcium ion. In this case, the coordination of the atom was different; however, the conformation of the protein was still "closed", a further evidence that the binding of metal ions is responsible for such an arrangement of the mobile loops.

Moreover, our attempts to obtain a crystal structure with both the inhibitor and the magnesium bound to the active site failed, further validating our hypothesis. In this case, the excess of the co-factor probably makes the lower affinity for the ion insignificant, effectively displacing the inhibitor in the active site. The obtainment of such a structure is substantially in agreement with the biological experiments, in which the inhibitory activity of the compound starts to decrease when the concentration of the magnesium reaches high levels.

In conclusion, our data indicate that the binding mode of our ligands is not influenced by the magnesium ion.
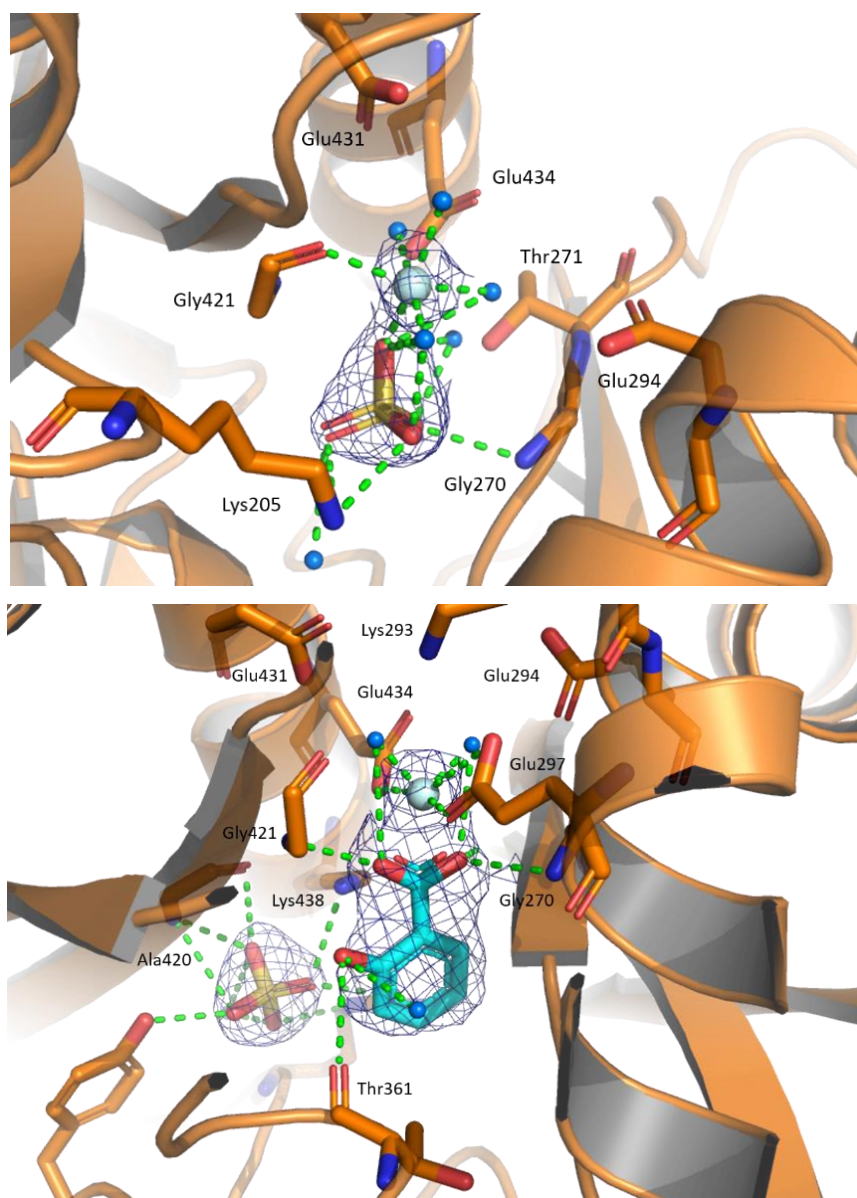
Figure 27. Ribbon diagram of the MbtI-Mg^{2+} structure: above, the active site of chain A, below the same view for chain D. The interactions of the ligands with the side chains (in sticks), the Mg^{2+} ion (light blue sphere) and the water molecules (blue spheres) are represented as green, dashed lines. The blue mesh represents the electron density around the ligand (scaled at 1σ).

3.4.3.3 Design and synthesis of a probe

In order to further confirm our hypothesis, and to ascertain the theory of the non-essentiality of the magnesium in determining the binding mode of the ligand, we decided to design and synthesise a probe molecule, incapable of chelating the metal ion. Therefore, we introduced a modification on **54**: we converted its carboxylic function to an amide, which, despite being incapable of interacting with the magnesium, can still form H-bonds with amino acids and water molecules.

The compound was synthesised by directly converting the carboxylic acid of **54** into an amide function, through a HATU-catalysed reaction.

67
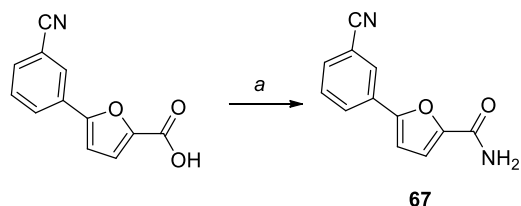
Reagents and conditions: *a*) *1*. HATU, DIPEA, DMF, r.t., 30 min.; *2*. NH$_4$Cl, r.t., 2 h.

The so-obtained compound was tested on MbtI, and it revealed a very good activity (% residual enzymatic activity: 12.3 ± 1.6; IC$_{50}$ = 17.3 ± 1.9), providing yet another demonstration that the previous theory can be effectively applied to MbtI, and that the binding mode of this class of compounds was uncovered by the crystal structure.

In conclusion, the definition of the binding mode of this class of MbtI inhibitors is a fundamental result to continue our studies. The next step will be the adjustment of the computational model, for the design of improved candidates, also considering further substituents and decorations for the improvement of their antimycobacterial properties.

3.5   PRELIMINARY STUDIES ON NANOCARRIERS FOR A NEW INHALED ANTI-TB THERAPY

Since *Mycobacterium tuberculosis* (Mtb) tends to localise in the lungs, the administration of therapeutic agents through the pulmonary route could be a valid strategy to improve the efficacy of the drugs. This approach allows the deposition of the active pharmaceutical ingredient (API) directly at the infection site, avoiding first-pass metabolism and reducing systemic side effects. Recently, liposomes have been proposed as new drug delivery carriers for inhalation therapy, to improve drug targeting and delivery. The success of targeting alveolar macrophages depends critically on the chemico-physical characteristics of the nanocarriers.[46]

Therefore, we decided to analyse the effect of the encapsulation of two established antitubercular drugs, like insoniazid (INH) and rifampicin (RIF), in order to explore the possibility to produce nanoparticles, embedded with our candidate compounds (Figure 28).[47] These studies were performed in collaboration with Prof. Eliana Leo at the University of Modena and Reggio Emilia (Italy).

Liposomes were prepared by Reverse Phase Evaporation (REV), using cholesterol (Chol) and phosphatidylcholine (PC) with a fixed molar ratio (1:1). For RIF-loaded liposomes, RIF:lipid (w/w) ratios of 3:100, 6:100, 12:100 (for RIF 3%, 6%, 12% samples, respectively) were used; as for INH, 15:100, 30:100, 60:100 drug:lipid (w/w) ratios were employed (for INH 15%, 30%, 60% samples, respectively). For the preparation of co-loaded liposomes, a RIF:INH:lipid (w/w) ratio of 12:15:100 was chosen. The size and polydispersity index (PDI) of the nanoparticles were determined by Photon Correlation Spectroscopy (PCS), a technique based on the Dynamic Light Scattering (DLS) of a laser beam of a certain wavelength by particles or in a liquid medium. The drug loading (DL %) and the encapsulation efficiency (EE %) of loaded-liposomes were evaluated by UV–visible spectroscopy. The *in vitro* release of RIF and INH from liposomes was analysed in Simulated Lung Fluid (SLF) at pH 7.4. For the SANS analyses, liposomes were prepared and purified using D_2O instead of milliQ water. Nanoparticles characterisation was undertaken using the fixed geometry, time-of-flight small angle neutron scattering Sans2d instrument at the ISIS Spallation

source at the Rutherford Appleton Laboratory, U.K. Small-angle neutron scattering (SANS) has recently emerged as a powerful tool to investigate amphiphilic aggregates, such as uni-lamellar and multi-lamellar vesicles. In particular, SANS provides crucial information about the arrangement and behaviour of the drug into the nanoparticle after its encapsulation and allows the analysis of these systems at nanometric level. This technique yields valuable and unique data about steric bilayer thickness, particle dispersion and the finer structural features of the liposomes.[47]
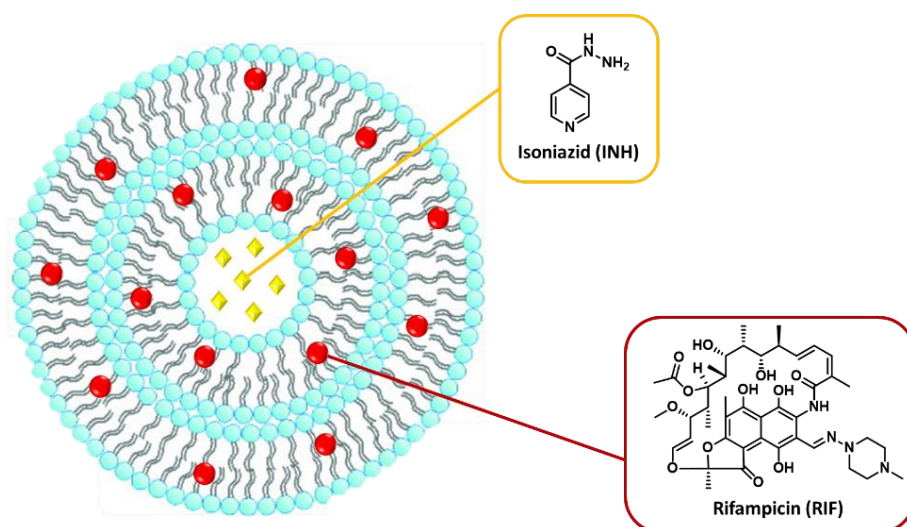
Figure 28. Graphical representation of a liposome, with the localisation of the tested antitubercular drugs. The hydrophilic isoniazid (yellow diamonds) localises in the aqueous core, while the hydrophobic rifampicin (red spheres) is in the lipid bilayer.

The PCS analysis on the liposome samples showed no differences in size, regardless of the type and amount of the embedded drug: all the samples, including the unloaded blanks, displayed an average diameter around 350 nm. Conversely, significant differences in the PDI values were detected. For RIF-loaded liposomes, the PDI increased parallelly to the drug loading, suggesting a reduction of size homogeneity. This effect might be due to changes in the structural organization of the lipid bilayer caused by the hydrophobic drug, which may interact with phospholipid tails of the liposome lamellae. In INH-loaded liposomes, the size homogeneity did not change with the increase of the drug loading, probably because INH is highly soluble in water and its interaction with the phospholipid lamellae is negligible. Finally, PCS analysis of co-loaded liposomes showed an intermediate value of PDI, with respect to the loaded and

unloaded samples, indicating the opposite influence of the two drugs on the homogeneity of the vesicle size. The results of PCS analysis were further supported by SANS data. For RIF-loaded liposomes, a slight decrease in size (too small to be detected by PCS analysis) was detected with respect to the unloaded carrier. This modification is probably due to interactions between the hydrophobic tails of phospholipids and RIF. Moreover, a destabilization of liposomal structure was highlighted. With INH-loaded liposomes, the dimension and multilamellar assembly did not significantly change up to 60% of the drug. The hydrophilicity of this drug leads to its confinement in the aqueous environment of the core and in the inter-bilayer spaces composing the liposome shell. In this respect, no variation in the thickness of the water and lipid layers was observed at low INH concentrations (15% and 30%). Conversely, higher INH concentration induced a slight yet significant decrease in these parameters. Moreover, a change in the inter-bilayer periodical spacing suggested the formation of interactions between the drug and the phospholipid heads, at the water-lipid interface. These interactions are probably hydrogen bonds or van der Waals forces, giving rise to slight changes in the liposome size, which cannot be detected by PCS analysis. This effect is coherent with the observed decrease of the lipid bilayer thickness. Notably, the disorganization of the lipids within the bilayer leads to an increment of the membrane fluidity, promoting the penetration of water molecules. As shown by the PCS studies, the SANS profiles for the co-loaded liposomes highlighted that it is possible to obtain stable liposomes, co-embedded with both drugs. Notably, while the hydrophobic RIF determines a large destabilization of the multi-lamellar vesicular structure, the co-presence of INH and RIF causes an intriguing stabilisation effect on the oligo-lamellar shelled liposome. As confirmed by the drug loading results, this synergistic effect promotes the uptake of both drugs inside the vesicles, allowing for the incorporation of higher amount of INH and RIF. The increased RIF loading might be explained by the structure stabilisation exerted by the co-presence of INH, while the higher INH loading could be attributable to an increase in the lipid bilayer rigidity, induced by the intercalation of the hydrophobic RIF, during liposome formation. Regarding the drug release, this effect did not elicit any significant variation in the *in vitro* behaviour of these systems. Despite the

different localization of the two drugs within the nanocarriers, the presence of RIF and IHN in the same liposome did not modify the release profiles of each drug, which remained substantially unaltered with respect to those of the mono-loaded formulations.[47]

All the experimental details regarding these tests are reported in our referenced publication and will not be further discussed in the Material and Methods section, because, at the present time, they still do not involve MbtI inhibitors. However, our preliminary results show that it is possible to encapsulate two drugs, endowed with different physico-chemical characteristics, in the same liposome nanoparticle. This promising outcome indicates that it would be possible to design an innovative pharmaceutical form for the potential co-delivery of our antitubercular agents with an established drug, in view of a combined therapy for the treatment of TB infections (Figure 29).
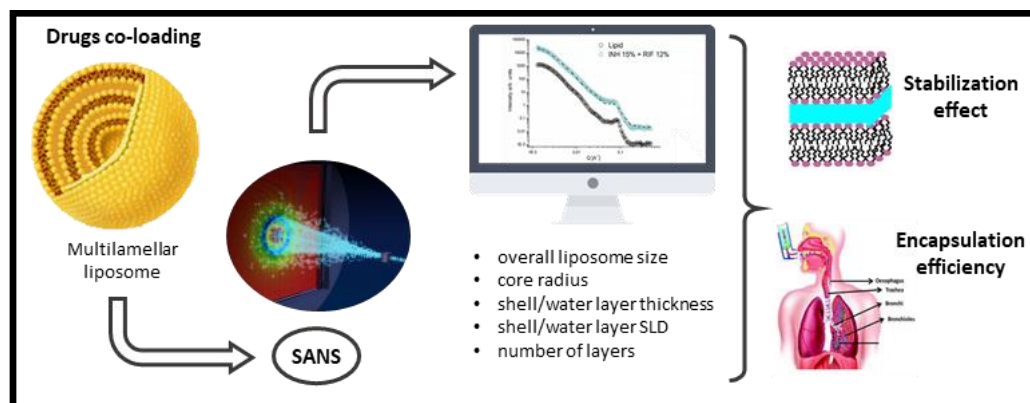
Figure 29. Graphical abstract representing the set-up for the development of anti-TB inhalation therapy.

Chapter 4: Conclusions

This research project, focused on the identification of new, potent MbtI inhibitors for the development of innovative anti-TB therapeutic agents, led to the obtainment of a number of candidate compounds, endowed with an excellent inhibitory activity on the target enzyme, and exhibiting an interesting antimycobacterial action, related to the reduced production of siderophores. Notably, we managed to establish, for the first time, a solid correlation between an effective inhibition of MbtI and a promising antimycobacterial effect in whole-cell assays.

Compounds **1**, **27**, and **54**, the most potent competitive inhibitors identified to date, with K_i values around 10 μM, were submitted to co-crystallisation experiments to empirically define the binding mode of this class of inhibitors within the active site of MbtI. An optimisation of the whole protein production and crystallisation process was necessary to achieve the expected outcomes. The crystal structure of the enzyme in complex with **54** was solved at 2.09 Å, and its analysis suggested the non-essentiality of the magnesium in determining the binding mode of the ligand. These structural data were further supported by biological experiments. According to our observations, and based on literature data, we determined that the ligand binds MbtI independently of the presence of magnesium, because, in this situation, the metal cofactor has a low affinity for the enzyme. Additionally, the inhibitor may also prevent the binding of the magnesium in a competitive fashion; this conclusion was derived by the fact that the compound becomes slightly less effective when the levels of the metal reach very high values.

The crystal structure of the protein in complex with the Mg^{2+} ion and the salicylate (a product of the catalytic reaction), solved at 2.11 Å, showed that the enzyme is active in the recombinant organism (*E. coli*) and that either chorismate or salicylate is co-purified with the protein. This structure represents the first evidence of the native binding mode of the natural ligand inside the MbtI active site. The presence of the magnesium probably contributed to prevent the loss of

the salicylate from the active site, considering that the metal stabilises the closed conformation of the enzyme. The crystal structures of MbtI with calcium and barium also evidenced a closed conformation of the protein, confirming that the binding of metal ions is responsible for such an arrangement of the mobile loops.

Overall, these results were instrumental for our understanding of the determinants of the binding mode of this class of compounds to the enzyme and will be necessary to guide the design of the next generation of inhibitors, for the identification of new anti-TB drugs.

Finally, our preliminary studies on the design of innovative liposome-based nanocarriers for antitubercular inhalation therapy yielded promising results. We proved that it is possible to obtain nanoparticles loaded with one or two drugs, thus laying the foundations for the development of novel delivery systems for the administration of combinations of new and established antitubercular agents.

Chapter 5: Materials and Methods

5.1 COMPUTATIONAL METHODS

The computational studies were performed in collaboration with Prof. Tiziano Tuccinardi at the Department of Pharmacy, University of Pisa in Italy.

5.1.1 Pharmacophore model generation

One of the main drawbacks of the VS consensus docking approach is the required computing time, because, by using this method, a whole dataset of molecules has to be subjected to different docking procedures. Therefore, in order to reduce the number of compounds to be analysed, a pharmacophore screening was used as a prefilter. The pharmacophore hypothesis was built on the basis of the 3D structure of MbtI co-crystallised with Methyl-AMT (PDB code: 3VEH).[28,33] The software LigandScout 4.08 was used for the generation of the receptor-based pharmacophore. First, an exhaustive model including all the possible features was obtained; subsequently, only the desired features were retained, for a total of four features. The excluded volume spheres were defined on the basis of the receptor structure as implemented in the default LigandScout configuration.[33]

5.1.2 Database generation and pharmacophore screening

The Enamine database, comprising about 1500000 commercially available compounds was used as the screening database. The idbgen-gui utility of LigandScout was employed to create the LigandScout 3D database. The iConBest conformational sampling method was applied. The 3D database was then screened using the previously generated pharmacophore model and the receptor excluded volume, imposing that only the compounds matching all the four pharmacophoric features of the model were retrieved.[33]

5.1.3 Docking procedures and pose filtering

Prior to the docking analysis, a self-docking evaluation of the main MbtI-inhibitor X-ray complexes (PDB codes 3RV6, 3RV7, 3RV8, 3RV9, 3ST6 and 3VEH) was carried out using Autodock, Autodock vina, Dock, Fred, Gold with the four fitness functions implemented (i.e. GoldScore, ChemScore, Astex Statistical Potential and ChemPLP), Glide, with the standard and extra precision methods, and Plants. Gold (with the four fitness functions) and Plants resulted to be the most reliable ones, as they showed an average RMSD lower than 2.0 Å. For all the docking procedures, the region of interest for the docking studies was defined in such a manner that it contained all the residues lying within 10 Å from the ligand in the X-ray structure.[33]

By applying the five docking methods, five different binding dispositions were obtained from the docking of each ligand into the protein binding site. The RMSD of these docking poses against the remaining four was evaluated by using the rms_analysis software of the Gold suite. By using an in-house program [18], these results were clustered, so that among the five results, all of the similar docking poses were clustered together. We selected an RMSD clustering threshold of 2.0 Å; therefore, the so-obtained clusters contained the group of poses that were less than 2.0 Å away from all other poses belonging to the same cluster. All of the ligands showing a consensus level of five were taken into account. The filtering of the docking results was carried out by superimposing the docked compounds to the pharmacophore directly from the supplied poses, without changing their coordinates. The retrieval of compounds matching all the four pharmacophore features of the model was imposed in this search.[33]

5.1.4 Molecular dynamics simulations

All simulations were carried out using AMBER14.[48] MD simulations were performed using the ff14 S B forcefield at 300 K. The complex was placed in a rectangular parallelepiped water box. The Mg^{2+} ion was inserted, based on the disposition and interaction of the metal in the crystal structure of Irp9 from *Yersinia enetrocolitica* (PDB code: 2FN1).[23] We used the particle mesh Ewald (PME) compatible Lennard-Jones parameters for divalent metal ions in explicit

solvent, reported by Merz and co-workers.[49] An explicit solvent model for water, TIP3P, was used, and the complexes were solvated with a 20 Å water cap. Sodium ions were added as counterions to neutralize the system. Prior to MD simulations, two steps of minimisation were carried out; in the first stage, we kept the protein fixed with a position restraint of 500 kcal/mol·Å^2 and we solely minimized the positions of the water molecules. In the second stage, we minimized the entire system through 5000 steps of steepest descent followed by conjugate gradient (CG), until a convergence of 0.05 kcal/Å·mol, using the same procedure described above. PME electrostatics and periodic boundary conditions were used in the simulation. The MD trajectory was run using the minimised structure as the starting conformation. The time step of the simulations was 2.0 fs with a cut-off of 10 Å for the non-bonded interaction, and SHAKE was employed to keep all bonds involving hydrogen atoms rigid. Constant-volume periodic boundary MD was carried out for 0.5 ns, during which the temperature was raised from 0 to 300 K. Then 19.5 ns of constant pressure periodic boundary MD was performed at 300 K using the Langevin thermostat to maintain constant the temperature of the system. All the α carbons of the protein were blocked with a harmonic force constant of 10 kcal/mol·Å^2 for the first 3.5 ns, while in the last 16.5 ns no constraints were applied. General Amber forcefield (GAFF) parameters were assigned to the ligands, while partial charges were calculated using the AM1-BCC method as implemented in the Antechamber suite of AMBER 14.[33]

5.2 GENERAL CHEMISTRY

All starting materials, chemicals and solvents were purchased from commercial suppliers (Sigma-Aldrich, St. Louis, MO, USA; FluoroChem, Hadfield, Derbyshire, UK; Carlo Erba, Cornaredo, Italy) and used as received. Anhydrous solvents were utilized without further drying. Aluminium-backed Silica Gel 60 plates (0.2 mm; Merck, Darmstadt, Germany) were used for analytical thin layer chromatography (TLC), to follow the course of the reactions. Microwave-assisted reactions were carried out with a Biotage® Initiator Classic (Biotage, Uppsala, Sweden). Silica gel 60 (40–63 μm; Merck) was used for the purification of intermediates and final compounds, through flash column chromatography. Melting points were determined in open capillary tubes with a Büchi Melting Point B-545 (Büchi, Flawil, Switzerland). All tested compounds were characterized by the means of FT-IR, ^{1}H NMR, ^{13}C NMR, MS. ^{1}H and ^{13}C NMR spectra were acquired at ambient temperature with a Varian-Oxford 300 MHz instrument (Varian, Palo Alto, CA, USA), operating at 300 MHz for ^{1}H and 75 MHz for ^{13}C. Chemical shifts are expressed in ppm (δ) from tetramethylsilane resonance in the indicated solvent (TMS: δ = 0.0 ppm), while J-couplings are given in Hertz. The APT sequence was used when deemed necessary. IR spectra were acquired with a Perkin Elmer Spectrum One FT-IR (Perkin Elmer, Waltham, MA, USA), in a spectral region between 4000 and 450 cm^{-1} and analysed by transmittance technique with 32 scansions and 4 cm^{-1} resolution. Solid samples were mixed in a mortar with KBr (1:100) and pressed to small tablets, using a hydraulic press (14 tons). MS analyses were carried out with a Thermo Finnigan LCQ Advantage LC/MS/MS system (Thermo Finnigan, San Jose, CA, USA), equipped with a quaternary pump, a Diode Array Detector (working wavelength: 254 nm) and a MS spectrometer, with an Electrospray ionization (ESI) source and an Ion Trap mass analyser (ionisation: ESI positive or ESI negative; capillary temperature: 250°C; source voltage: 5.50 kV; source current: 4.00 mA; multipole 1 and 2 offset: 5.50 V and 7.50 V, respectively; intermultipole lens voltage: 16.00 V; trap DC offset voltage: 10.00 V). The purity of the compounds was assessed by means of HPLC (λ: 220 nm; column: EVO C18, Phenomenex, Torrance, CA, USA; eluent system: aqueous formic acid (0.05%)/acetonitrile 50:50; total flow: 1.2 mL/min; injection volume: 30 mL) and was ≥ 95%, unless otherwise stated.

Alternatively, the purity was confirmed by means of high-resolution mass spectrometry, using a Q Exactive™ Hybrid Quadrupole-Orbitrap™ Mass Spectrometer (Thermo Fischer, Waltham, MA, USA).

The five compounds selected by the computational study (**VS1**-**5**) were acquired from Enamine (Kyiv, Ukraine) at the highest purity level available (≥ 95%). Compounds **1**, **5**, **45**, **57** were purchased from Sigma Aldrich.

5.2.1 Synthetic procedures

5.2.1.1 Synthesis of 2-4,6-8,10-14,17,18,20,27-29,46,47,50,54,58-60 (General Scheme A)
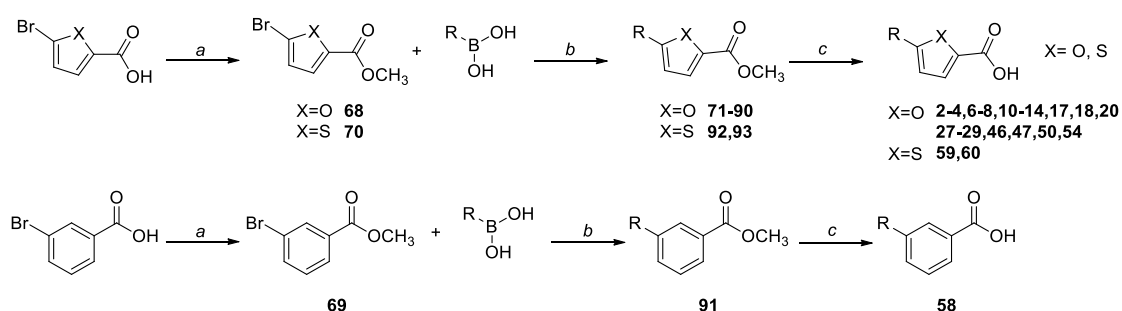
Reagents and conditions: *a)* conc. H_2SO_4, MeOH, reflux, 24 h; *b)* Pd(PPh$_3$)$_2$Cl$_2$, 2M Na$_2$CO$_3$, 1,4-dioxane, 90°C, overnight, N$_2$; *c)* NaOH, EtOH/THF 1:1, reflux, 5 h *or* LiOH·H$_2$O, THF/H$_2$O 1:1, 20°C, 2 h.

General procedure for the esterification of aromatic carboxylic acids (*a*). The appropriate aromatic carboxylic acid (1 mmol) was dissolved in methanol (4.8 mL) before concentrated sulfuric acid (0.7 mL) was added dropwise while stirring. The reaction mixture was refluxed for 24 h. After completion, the solution was cooled to room temperature and the methanol was removed *in vacuo*; the aqueous phase was treated with sodium bicarbonate to ensure neutral-basic pH and then extracted with ethyl acetate. The combined organic layers were washed with brine, dried over anhydrous sodium sulphate, filtered and concentrated *in vacuo*.[33,38]

Methyl 5-bromofuran-2-carboxylate (**68**)*.* Starting compound: 5-bromo-2-furoic acid. White solid. Yield: 89%. Mp: 64°C. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.59. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.12 (d, *J* = 3.5 Hz, 1H, ArH), 6.45 (d, *J* = 3.5 Hz, 1H, ArH), 3.89 (s, 3H, CH$_3$).

Methyl 3-bromobenzoate (**69**)*.* Starting compound: 3-bromobenzoic acid. White solid. Yield: 96%. TLC (cyclohexane – ethyl acetate 9:1): R_f = 0.66. Mp: 32°C (lit.). ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm) 8.18 (s, 1H, ArH), 7.97 (d, *J* = 7.8 Hz, 1H, ArH), 7.68 (d, *J* = 7.8 Hz, 1H, ArH), 7.32 (t, *J* = 7.8 Hz, 1H, ArH), 3.92 (s, 3H, CH$_3$).

Methyl 5-bromothiophene-2-carboxylate (**70**)*.* Starting compound: 5-bromothiophene-2-carboxylic acid. Pale yellow solid. Yield: 97%. TLC (cyclohexane – ethyl acetate 9:1): R_f = 0.60. Mp: 60°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.55 (d, *J* = 4.0 Hz, 1H, ArH), 7.07 (d, *J* = 4.0 Hz, 1H, ArH), 3.88 (s, 3H, CH$_3$).

General procedure for the conventional Suzuki-Miyaura coupling (*b*). The suitable aromatic ester (1 mmol), the appropriate phenylboronic acid (1.3 mmol) and bis(triphenylphosphine)palladium(II) dichloride (5% mol) were dissolved in dry 1,4-dioxane (10 mL), under nitrogen atmosphere. A 2M sodium carbonate solution (2 mmol) was then added and the resulting mixture was stirred overnight at 90°C. After completion, the solution was cooled to room temperature and then filtered on a celite pad. The filtrate was diluted with water and extracted with ethyl acetate (3-4 mL). The organic layer was dried over anhydrous sodium sulphate, filtered and concentrated *in vacuo*.[33,38]

Methyl 5-(2-chlorophenyl)furan-2-carboxylate (**71**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (2-chlorophenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) to give the desired product as a greyish solid. Yield: 62%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.45. Mp: 68°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.98 (dd, *J* = 7.8, 1.4 Hz, 1H, ArH), 7.45 (dd, *J* = 7.8, 1.4 Hz, 1H, ArH), 7.35 (dt, *J* = 7.8, 1.4 Hz, 1H, ArH), 7.30 – 7.23 (m, 2H, H$_{10}$, ArH), 7.20 (d, *J* = 3.7 Hz, 1H, ArH), 3.92 (s, 3H, CH$_3$).

Methyl 5-phenylfuran-2-carboxylate (**72**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and phenylboronic acid. The crude was purified by flash column chromatography (cyclohexane – dichloromethane – ethyl acetate 9.5:0.25:0.25) to give the desired product as a white solid. Yield: 80%. TLC (cyclohexane – dichloromethane – ethyl acetate 9.5:0.25:0.25): R_f = 0.15. Mp: 70.1 °C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm) 7.79 (d, *J* = 8.1 Hz, 2H, ArH), 7.45-7.35 (m,

3H, ArH), 7.25 (d, J = 3.7 Hz, 1H, ArH), 6.74 (d, J = 3.7 Hz, 1H, ArH), 3.92 (s, 3H, CH$_3$).

Methyl 5-(pyridin-4-yl)furan-2-carboxylate (**73**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and pyridin-4-ylboronic acid. The compound was obtained according to the general procedure, using a mixture of 1,4-dioxane (9 mL) and dimethylformamide (2 mL). The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 1:1) to give the desired product as a white solid. Yield: 65%. TLC (cyclohexane – ethyl acetate 6:4): R$_f$ = 0.08. Mp: 98°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm) 8.68 (dd, J = 4.8, 1.3 Hz, 2H, ArH), 7.66 (dd, J = 4.8, 1.3 Hz, 2H, ArH), 7.28 (d, J = 3.7 Hz, 1H, ArH), 6.98 (d, J = 3.7 Hz, 1H, ArH), 3.94 (s, 3H, CH$_3$).

Methyl 5-(2-nitrophenyl)furan-2-carboxylate (**74**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (2-nitrophenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – dichloromethane – ethyl acetate 6.5:2.5:1) to give the desired product as a yellow solid. Yield: 13%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.40. Mp: 80°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.84-7.79 (m, 2H, ArH), 7.65 (dt, J = 7.8, 1.3 Hz, 1H, ArH), 7.52 (dt, J = 7.8, 1.3 Hz, 1H, ArH), 7.25 (d, J = 3.7 Hz, 1H, ArH), 6.70 (d, J = 3.7 Hz, 1H, ArH), 3.91 (s, 3H, CH$_3$).

Methyl 5-(4-hydroxyphenyl)furan-2-carboxylate (**75**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (4-hydroxyphenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 7:3) to give the desired product as a light grey solid. Yield: 79%. TLC (cyclohexane – ethyl acetate 6:4): R$_f$ = 0.46. Mp: 188°C. ^{1}H NMR (300 MHz, CDCl$_3$) δ (ppm): 7.69 (dd, J = 8.7, 1.9 Hz, 2H, ArH), 7.23 (d, J = 3.6 Hz, 1H, ArH), 6.91 (dd, J = 8.7, 1.9 Hz, 2H, ArH), 6.60 (d, J = 3.6 Hz, 1H, ArH), 5.20 (bs exch. D$_2$O, 1H, OH), 3.91 (s, 3H, CH$_3$).

Methyl 5-(4-chlorophenyl)furan-2-carboxylate (**76**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (4-chlorophenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) to give the desired product as an ochre solid. Yield: 67%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.43. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.70 (dd, J =

7.8, 0.8 Hz, 2H, ArH), 7.38 (dd, J = 7.8, 0.8 Hz, 2H, ArH), 7.23 (d, J = 3.7 Hz, 1H, ArH), 6.72 (d, J = 3.7 Hz, 1H, ArH), 6.05 (bs exch. D_2O, 1H, NH), 5.60 (bs exch. D_2O, 1H, NH), 3.91 (s, 3H, CH_3).

Methyl 5-(4-carbamoylphenyl)furan-2-carboxylate (**77**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (4-carbamoylphenyl)boronic acid. The crude was purified by flash column chromatography (dichloromethane – ethyl acetate 7:3) and crystallised from acetone/hexane to give the desired product as a white solid. Yield: 64%. TLC (cyclohexane – ethyl acetate 4:6): R_f = 0.20. ^{1}H-NMR (300 MHz, $CDCl_3$) δ (ppm): 7.87 (s, 4H, ArH), 7.27 (d, 1H, partially hidden by solvent peak), 6.85 (d, J = 3.6 Hz, 1H, ArH), 3.93 (s, 3H, CH_3).

Methyl 5-(4-(trifluoromethyl)phenyl)furan-2-carboxylate (**78**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (4-(trifluoromethyl)phenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) to give the desired product as a yellowish solid. Yield: 65%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.47. Mp: 87°C. ^{1}H-NMR (300 MHz, $CDCl_3$) δ (ppm): 7.88 (d, J = 8.1 Hz, 2H, ArH), 7.67 (d, J = 8.1 Hz, 2H, ArH), 7.26 (d, J = 3.7 Hz, 1H, ArH), 6.84 (d, J = 3.7 Hz, 1H, ArH), 3.93 (s, 3H, CH_3).

Methyl 5-(4-fluorophenyl)furan-2-carboxylate (**79**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (4-fluorophenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a white solid. Yield: 68%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.41. ^{1}H-NMR (300 MHz, $CDCl_3$) δ (ppm): 7.76 (dd, J = 8.7, 5.3 Hz, 2H, ArH), 7.24 (d, J = 3.6 Hz, 1H, ArH), 7.11 (t, J = 8.7 Hz, 2H, ArH), 6.68 (d, J = 3.6 Hz, 1H, ArH), 3.91 (s, 3H, CH_3).

Methyl 5-(4-(methylsulfonyl)phenyl)furan-2-carboxylate (**80**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (4-(methylsulfonyl)phenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 6:4) to give the desired product as a sparkling white solid. Yield: 65%. TLC (cyclohexane – ethyl acetate 6:4): R_f = 0.24. Mp: 164°C. ^{1}H-NMR (300 MHz, $CDCl_3$) δ (ppm): 8.10 – 7.85 (m, 4H, ArH),

7.27 (d, J = 3.7 Hz, 1H, ArH), 6.92 (d, J = 3.7 Hz, 1H, ArH), 3.93 (s, 3H, OCH$_3$), 3.07 (s, 3H, SO$_2$CH$_3$).

Methyl 5-(4-sulfamoylphenyl)furan-2-carboxylate (**81**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (4-sulfamoylphenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 5:5) to give the desired product as a white solid. Yield: 30%. TLC (cyclohexane – ethyl acetate 5:5): R$_f$ = 0.35. Mp: 208°C (dec.). ^{1}H-NMR (300 MHz, CD$_3$OD) δ (ppm): 7.97 (s, 4H, ArH), 7.34 (d, J = 3.7 Hz, 1H, ArH), 7.12 (d, J = 3.7 Hz, 1H, ArH), 3.91 (s, 3H, CH$_3$).

Methyl 5-(4-cyanophenyl)furan-2-carboxylate (**82**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (4-cyanophenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – dichloromethane – ethyl acetate 7:1.5:1.5) to give the desired product as a white solid. Yield: 72%. TLC (cyclohexane – dichloromethane – ethyl acetate 7:1.5:1.5): R$_f$ = 0.43. Mp: 156°C. ^{1}H NMR (300 MHz, CDCl$_3$) δ (ppm): 7.87 (dd, J = 8.7, 1.9 Hz, 2H, ArH), 7.71 (dd, J = 8.7, 1.9 Hz, 2H, ArH), 7.27 (d, J = 3.6 Hz, 1H, ArH), 6.89 (d, J = 3.6 Hz, 1H, ArH), 3.93 (s, 3H, CH$_3$).

Methyl 5-(2-chloro-4-trifluoromethyl-2-phenyl)furan-2-carboxylate (**83**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (2-chloro-4-(trifluoromethyl)phenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate - dichloromethane 9:0.5:0.5) to give the desired product as a white solid. Yield: 62%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.65. Mp: 58°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.14 (d, 1H, J = 8.3 Hz, ArH), 7.73 (s, 1H, ArH), 7.60 (d, 1H, J = 8.3 Hz, ArH), 7.33 (d, 1H, J = 3.6 Hz, ArH), 7.26 (d, 1H, J = 3.6 Hz, ArH), 3.94 (s, 3H, CH$_3$).

Methyl 5-(2,4-bis(trifluoromethyl)2-phenyl)furan-2-carboxylate (**84**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (2,4-bis(trifluoromethyl)phenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) to give the desired product as a light brown solid. Yield: 29%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.66. Mp: 149°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.08-7.97 (m, 2H, ArH), 7.88 (d,

1H, *J* = 8.1 Hz, ArH), 7.28 (d, 1H, *J* = 3.7 Hz, ArH), 6.92 (d, 1H, *J* = 3.7 Hz, ArH), 3.93 (s, 3H, CH₃).

Methyl 5-(3,4-bis(trifluoromethyl)2-phenyl)furan-2-carboxylate (**85**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (3,4-bis(trifluoromethyl)phenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a light brown solid. Yield: 33%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.43. Mp: 147°C. ¹H-NMR (300 MHz, CDCl₃) δ (ppm): 8.19 (s, 1H, ArH), 8.06 (d, 1H, *J* = 8.4 Hz, ArH), 7.90 (d, 1H, *J* = 8.4 Hz, ArH), 7.29 (d, 1H, *J* = 3.6 Hz, ArH), 6.95 (d, 1H, *J* = 3.6 Hz, ArH), 3.95 (s, 3H, CH₃).

Methyl 5-(3,5-bis(trifluoromethyl)phenyl)furan-2-carboxylate (**86**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (3,5-bis(trifluoromethyl)phenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) to give the desired product as a white solid. Yield: 58%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.54. Mp: 139°C. ¹H-NMR (300 MHz, CDCl₃) δ (ppm): 8.18 (s, 2H, ArH), 7.83 (s, 1H, ArH), 7.29 (d, *J* = 3.7 Hz, 1H, ArH), 6.94 (d, *J* = 3.7 Hz, 1H, ArH), 3.95 (s, 3H, CH₃).

Methyl 5-(3-chlorophenyl)furan-2-carboxylate (**87**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (3-chlorophenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a white solid. Yield: 68%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.53. Mp: 76°C. ¹H-NMR (300 MHz, CDCl₃) δ (ppm): 7.76 (s, 1H, ArH), 7.64 (d, *J* = 7.3 Hz, 1H, ArH), 7.40 – 7.13 (m, 3H, ArH), 6.75 (d, *J* = 3.6 Hz, 1H, ArH), 3.92 (s, 3H, CH₃).

Methyl 5-(3-hydroxyphenyl)furan-2-carboxylate (**88**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (3-hydroxy)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a white solid. Yield: 31%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.23. Mp: 149°C. ¹H-NMR (300 MHz, CDCl₃) δ (ppm): 7.51 – 7.10 (m, 4H, ArH), 6.85 (d, *J* = 7.6 Hz, 1H, ArH), 6.72 (d, *J* = 3.6 Hz, 1H, ArH), 5.34 (bs, 1H, OH), 3.92 (s, 3H, CH₃).

Methyl 5-(3-carbamoylphenyl)furan-2-carboxylate (**89**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (3-carbamoylphenyl)boronic acid. The crude was purified by flash column chromatography (dichloromethane – ethyl acetate 7:3) to give the desired product as a white solid. Yield: 25%. TLC (cyclohexane – ethyl acetate 4:6): R_f = 0.24. Mp: 192°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.22 (t, *J* = 1.4 Hz, 1H, ArH), 7.94 (dt, *J* = 7.8, 1.4 Hz, 1H, ArH), 7.79 (dt, *J* = 7.8, 1.4 Hz, 1H, ArH), 7.52 (t, *J* = 7.8 Hz, 1H, ArH), 7.26 (d partially hidden by solvent peak, *J* = 3.6 Hz, 1H, ArH), 6.83 (d, *J* = 3.6 Hz, 1H, ArH), 6.20 (bs, 1H, N*H*), 5.66 (bs, 1H, N*H*), 3.93 (s, 3H, CH$_3$).

Methyl 5-(3-cyanophenyl)furan-2-carboxylate (**90**). Starting compounds: methyl 5-bromofuran-2-carboxylate (**68**) and (3-cyanophenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a white solid. Yield: 75%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.33. Mp: 149°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.06 (s, 1H, ArH), 8.00 (d, *J* = 7.8 Hz, 1H, ArH), 7.62 (d, *J* = 7.8 Hz, 1H, ArH), 7.54 (t, *J* = 7.8 Hz, 1H, ArH), 7.26 (d partially hidden by solvent peak, *J* = 3.6 Hz, 1H, ArH), 6.83 (d, *J* = 3.6 Hz, 1H, ArH), 3.94 (s, 3H, CH$_3$).

Methyl 4'-nitro-[1,1'-biphenyl]-3-carboxylate (**91**). Starting compounds: methyl 3-bromobenzoate (**69**) and (4-nitrophenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) to give the desired product as a white solid. Yield: 87%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.53. Mp: 120°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm) 8.45 – 8.24 (m, 3H, ArH), 8.12 (d, *J* = 7.8 Hz, 1H, ArH), 7.87 – 7.74 (m, 3H, ArH), 7.58 (t, *J* = 7.8 Hz, 1H, ArH), 3.97 (s, 3H, CH$_3$).

Methyl 5-(4-nitrophenyl)thiophene-2-carboxylate (**92**). Starting compounds: methyl 5-bromothiophene-2-carboxylate (**70**) and (4-nitrophenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 95:5) and recrystallisation (dichloromethane – methanol) to give the desired product as a yellow solid. Yield: 26%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.50. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.28 (d, *J* = 9.0 Hz, 2H, ArH), 7.86 – 7.73 (m, 3H, ArH), 7.44 (d, *J* = 4.0 Hz, 1H, ArH), 3.93 (s, 3H, CH$_3$).

Methyl 5-(3-cyanophenyl)thiophene-2-carboxylate (**93**). Starting compounds: methyl 5-bromothiophene-2-carboxylate (**70**) and (3-cyanophenyl)boronic acid. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) and recrystallisation (dichloromethane – methanol) to give the desired product as a white solid. Yield: 25%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.33. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.91 (s, 1H, ArH), 7.85 (d, *J* = 7.8 Hz, 1H, ArH), 7.79 (d, *J* = 3.9 Hz, 1H, ArH), 7.64 (d, *J* = 7.8 Hz, 1H, ArH), 7.54 (t, *J* = 7.8 Hz, 1H, ArH), 7.34 (d, *J* = 3.9 Hz, 1H, ArH), 3.92 (s, 3H, CH$_3$).

General procedure for the base-catalysed hydrolysis (*c*). *1.* The appropriate methyl ester derivative (1 mmol) was dissolved in a mixture of tetrahydrofuran/ethanol 1:1 (15 mL) and a 1M solution of sodium hydroxide (2.5 mmol) was added dropwise while stirring. The reaction mixture was heated at reflux for 5 h. After completion, the solvent was evaporated under reduced pressure; the aqueous phase was washed with chloroform (1-5 mL), acidified with hydrochloric acid (3M) and then extracted with ethyl acetate (3-7 mL). The organic layers were washed with brine, dried over anhydrous sodium sulphate and then concentrated in vacuo. The resulting solid was finally washed with cool hexane (3 mL). *2.* The appropriate methyl ester derivative (1 mmol) was treated with lithium hydroxide (3.0 mmol) in a mixture of tetrahydrofuran - water 2:1 (15 mL) at room temperature for 2 h. After completion, the solution was acidified with hydrochloric acid (1M) and then extracted with ethyl acetate (3 x 7 mL). The organic layers were washed with brine, dried over anhydrous sodium sulphate and then concentrated *in vacuo*. Variations to this general procedure were applied to the synthesis of 5-(pyridin-4-yl)furan-2-carboxylic acid (**4**).[33,38]

5-(2-Chlorophenyl)furan-2-carboxylic acid (**2**). Starting compound: methyl 5-(2-chlorophenyl)furan-2-carboxylate (**71**). Procedure 1. White solid. Yield: 75%. TLC (dichloromethane – methanol 8:2 + 3 drops of acetic acid): R_f = 0.71. Mp: 215°C (dec.). ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 7.85 (dd, *J* = 9.4, 1.7 Hz, 1H, ArH), 7.61 (dd, *J* = 7.7, 1.7 Hz, 1H, ArH), 7.50 (dt, *J* = 7.7, 1.7 Hz, 1H, ArH), 7.45 (ddd, *J* = 9.4, 7.7, 1.7 Hz, 1H, ArH), 7.34 (d, *J* = 3.8 Hz, 1H, ArH), 7.22 (d, *J* = 3,8 Hz, 1H, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.9, 153.1, 145.1, 131.6, 131.0,

130.7, 129.4, 128.5, 128.2, 120.0, 113.4. FTIR (KBr) ν cm^{-1}: 3172, 3117, 2916, 2707, 2660, 1689, 1675, 1518, 1469, 1429, 1314, 1026, 761, 751. ESI-MS *m/z*: C$_{11}$H$_7$ClO$_3$ calcd 223, found 222, [M-H]$^-$. HPLC purity: t = 2.30', ≥ 98%.

5-Phenylfuran-2-carboxylic acid (**3**). Starting compound: methyl 5-phenylfuran-2-carboxylate (**72**). Procedure 1. White solid. Yield: 98%. TLC (dichloromethane – methanol 7:3): Rf = 0.19. Mp: 153°C. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.13 (bs exch. D$_2$O, 1H, COOH), 7.81-1.77 (m, 2H, ArH), 7.49-7.44 (m, 2H, ArH), 7.38 (ddt, *J* = 6.3, 1.3 Hz, 1H, ArH), 7.30 (d, *J* = 3,8 Hz, 1H, ArH), 7.13 (d, *J* = 3,8 Hz, 1H, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.7, 156.6, 144.7, 129.8, 129.6, 129.5, 124.8, 120.2, 108.4. FTIR (KBr) ν cm^{-1}: 2962, 2921, 2851, 1676, 1588, 1571, 1521, 1473, 1312, 1263, 1102, 804, 766, 756. ESI-MS *m/z*: C$_{11}$H$_8$O$_3$ calcd 188, found 187[M-H]$^-$, 397 [2M+Na]$^+$. HPLC purity: t = 7.43', ≥ 98%.

5-(Pyridin-4-yl)furan-2-carboxylic acid (**4**). Starting compound: methyl 5-(pyridin-4-yl)furan-2-carboxylate (**73**). Procedure 2. After the completion of the reaction, 2M HCl was slowly dripped until the formation of a white precipitate (pH 4-5); the solid was filtered off, dried *in vacuo* and tested without further purification. White solid. Yield: 89%. TLC (dichloromethane – methanol 7:3): R$_f$ = 0.20. Mp: ~292°C (dec.). ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.2 (bs exch. D$_2$O, 1H, COOH), 8.64 (d, *J* = 4.6 Hz, 2H, ArH), 7.73 (d, *J* = 4.6 Hz, 2H, ArH), 7.42 (d, *J* = 3.7 Hz, 1H, ArH), 7.36 (d, *J* = 3,7 Hz, 1H, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.5, 153.9, 150.9, 146.1, 136.2, 120.0, 118.7, 111.9. FTIR (KBr) ν cm^{-1}: 3422, 3134, 2919, 2850, 1720, 1613, 1603, 1581, 1569, 1549, 1473, 1420, 1375, 1305, 1027, 1015, 836, 803, 770, 759, 707, 544. ESI-MS *m/z*: C$_{10}$H$_7$NO$_3$ calcd 189, found 188 [M-H]$^-$. HPLC purity: t = 1.25', ≥ 98%.

5-(2-Nitrophenyl)furan-2-carboxylic acid (**6**). Starting compound: methyl 5-(2-nitrophenyl)furan-2-carboxylate (**74**). Procedure 1. Yellow solid. Yield: 69%. TLC (dichloromethane – methanol 8:2 + 3 drops of acetic acid): R$_f$ = 0.53. Mp: 218°C (dec.). ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.25 (bs exch. D$_2$O, 1H, COOH), 7.96 (d, *J* = 8.0 Hz, 1H, ArH), 7.87 (dd, *J* = 7.7, 1.3, 1H, ArH), 7.78 (dt, *J* = 7.7, 1.0 Hz, 1H, ArH), 7.67 (dt, *J* = 8.0, 1.3 Hz, 1H, ArH), 7.32 (d, *J* = 3.4 Hz, 1H, ArH), 7.02 (d, *J* = 3,4 Hz, 1H, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.4, 151.8,

147.8, 146.1, 133.3, 131.0, 130.3, 124.7, 122.6, 119.7, 112.0. FTIR (KBr) ν cm^{-1}: 3130, 3080, 2918, 2852, 2704, 2663, 2580, 1698, 1614, 1574, 1515, 1472, 1442, 1316, 1253, 1220. 1170, 1029, 761. ESI-MS *m/z*: C$_{11}$H$_7$NO$_5$ calcd 233, found 257 [M+Na]$^+$. HPLC purity: t = 1.61', ≥ 98%.

5-(4-Hydroxyphenyl)furan-2-carboxylic acid (**7**). Starting compound: methyl 5-(4-hydroxyphenyl)furan-2-carboxylate (**75**). Procedure 1. White solid. Yield: 59%. TLC (dichloromethane – methanol 7:3 + 3 drops of acetic acid): R$_f$ = 0.33. Mp: 265°C (dec.). ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 10.4-10.0 (bs exch. D$_2$O, 1H, COOH), 7.52 (d, *J* = 8.5 Hz, 2H, ArH), 6.80 (m, 3H, ArH), 6.64 (d, *J* = 3,3 Hz, 1H, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 164.1, 158.4, 154.4, 151.1, 126.1, 122.2, 116.4, 115.61, 105.1. FTIR (KBr) ν cm^{-1}: 3369, 3129, 2960, 1614, 1587, 1563, 1533, 1440, 1383, 1269, 1215, 1171, 1022, 962, 837, 811, 780. ESI-MS *m/z*: C$_{11}$H$_8$O$_4$ calcd 204, found 203 [M-H]$^-$. HPLC purity: t = 1.09', ≥ 98%.

5-(4-Chlorophenyl)furan-2-carboxylic acid (**8**). Starting compound: methyl 5-(4-chlorophenyl)furan-2-carboxylate (**76**). Procedure 1. White solid. Yield: 79%. TLC (dichloromethane – methanol 8:2 + 3 drops of acetic acid): R$_f$ = 0.69. Mp: 200° C (dec.). ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 7.82 (d, J=8.5 Hz, 2H, H$_7$, H$_{7'}$), 7.55 (d, J=8.5 Hz, 2H, H$_8$, H$_{8'}$), 7.30 (d, J = 3.4 Hz, 1H, H$_3$), 7.18 (d, J = 3,4 Hz, 1H, H$_4$). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.92, 155.63, 145.28, 134.06, 129.88, 128.73, 126.78, 120.44, 128.16, 109.28. FTIR (KBr): ν = 3435, 3131, 2962, 2919, 2851, 1666, 1586, 1567, 1418, 1411, 1032, 820, 807, 731 cm^{-1}. ESI-MS *m/z* calcd for C$_{11}$H$_7$ClO$_3$ calcd 222, found 221 [M-H]$^-$.

5-(4-Carbamoylphenyl)furan-2-carboxylic acid (**10**). Starting compound: methyl 5-(4-carbamoylphenyl)furan-2-carboxylate (**77**). Procedure 1. White solid. Yield: 84%. TLC (dichloromethane – methanol 7:3): R$_f$ = 0.32. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 8.03 (bs exch. D$_2$O, 1H, NH), 7.96 (d, *J* = 8.1 Hz, 2H, ArH), 7.86 (d, *J* = 8.1 Hz, 2H, ArH), 7.41 (bs exch. D$_2$O, 1H, NH), 7.31 (d, *J* = 3.6 Hz, 1H, ArH), 7.25 (d, *J* = 3.6 Hz, 1H, ArH).

5-(4-(Trifluoromethyl)phenyl)furan-2-carboxylic acid (**11**). Starting compound: methyl 5-(4-(trifluoromethyl)phenyl)furan-2-carboxylate (**78**). Procedure 1. Yellowish solid. Yield: 72%. TLC (dichloromethane – methanol 8:2 + 3 drops of acetic acid): R$_f$ = 0.48. Mp: 209°C. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ

(ppm): 13.3 (bs exch. D_2O, 1H, COOH), 8.00 (d, J = 8.1 Hz, 2H, ArH), 7.82 (d, J = 8.1 Hz, 2H, ArH), 7.35 (d, J = 3.8 Hz, 1H, ArH), 7.33 (d, J = 3.4 Hz, 1H, ArH). ^{13}C NMR (75 MHz, DMSO-d_6) δ (ppm): 159.8, 155,1, 145.8, 133.4, 130.2-126.6-122.9-119.3 (q), 129.9-129.7-129.5-129.1 (q), 126.8-126.7-126.7-126.6 (q), 125.6, 120.5, 110.8. FTIR (KBr) ν cm^{-1}: 3424, 2919, 2853, 1704, 1678, 1620, 1578, 1537, 1429, 1326, 1111, 844, 831, 693. ESI-MS m/z: $C_{12}H_7F_3O_3$ calcd 256, found 255 [M-H]$^-$. HPLC purity: t = 2.75', ≥ 98%.

5-(4-Fluorophenyl)furan-2-carboxylic acid (**12**). Starting compound: methyl 5-(4-fluorophenyl)furan-2-carboxylate (**79**). Procedure 1. White solid. Yield: 79%. TLC (dichloromethane – methanol 7:3): R_f = 0.38. ^{1}H NMR (300 MHz, CD$_3$OD) δ (ppm): 7.85 (dd, J = 8.5, 5.6 Hz, 2H, ArH), 7.27 (d, J = 3.6 Hz, 1H, ArH), 7.19 (t, J = 8.5 Hz, 1H, ArH), 6.90 (d, J = 3.6 Hz, 1H, ArH).

5-(4-(Methylsulfonyl)phenyl)furan-2-carboxylic acid (**13**). Starting compound: methyl 5-(4-(methylsulfonyl)phenyl)furan-2-carboxylate (**80**). Procedure 1. White solid. Yield: 76%. TLC (dichloromethane – methanol 8:2 + 3 drops of acetic acid): R_f = 0.40. Mp: 216°C. ^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm): 13.2 (bs exch. D_2O, 1H, COOH), 8.04 (d, J = 8.1 Hz, 2H, ArH), 7.99 (d, J = 8.1 Hz, 2H, ArH), 7.38 (d, J = 3.8 Hz, 1H, ArH), 7.35 (d, J = 3,8 Hz, 1H, ArH). ^{13}C NMR (75 MHz, DMSO-d_6) δ (ppm): 159.8, 154.9, 146.1, 141.0, 134.1, 128.6, 125.6, 120.5, 111.4, 44.1. FTIR (KBr) ν cm^{-1}: 3433, 3014, 2928, 1690, 1603, 1590, 1570, 1536, 1475, 1423, 1335, 1302, 1283, 1147, 960, 778, 558, 531. ESI-MS m/z: $C_{12}H_{10}O_5S$ calcd 266, found 265 [M-H]$^-$. HPLC purity: t = 1.20', ≥ 98%.

5-(4-Sulfamoylphenyl)furan-2-carboxylic acid (**14**). Starting compound: methyl 5-(4-sulfamoylphenyl)furan-2-carboxylate (**81**). Procedure 1. Beige solid. Yield: 94%. TLC (dichloromethane – methanol 8:2 + 2 drops of acetic acid): R_f = 0.08. Mp > 300°C. ^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm): 13.2 (bs exch. D_2O, 1H, COOH), 7.97 (d, J = 8.0 Hz, 2H, ArH), 7.90 (d, J = 8.0 Hz, 2H, ArH), 7.42 (bs exch. D_2O, 2H, NH$_2$), 7.34 (d, J = 3.4 Hz, 1H, ArH), 7.30 (d, J = 3,4 Hz, 1H, ArH). ^{13}C NMR (75 MHz, DMSO-d_6) δ (ppm): 159.6, 155.1, 145.4, 144.3, 132.4, 127.0, 125.1, 120.3, 110.4. FTIR (KBr) ν cm^{-1}: 3358, 3269, 3119, 2962, 2920, 2852, 1685, 1604, 1567, 1527, 1480, 1427, 1339, 1308, 1272, 1221, 1155, 960, 801, 758, 558, 538.

ESI-MS *m/z*: $C_{11}H_9NO_5S$ calcd 267, found 266 [M-H]⁻. HPLC purity: t = 1.10', ≥ 98%.

5-(4-Cyanophenyl)furan-2-carboxylic acid (**17**). Starting compound: methyl 5-(4-cyanophenyl)furan-2-carboxylate (**82**). Procedure 1. Yellow solid. Yield: 55% TLC (dichloromethane – methanol 7:3): R_f = 0.23. Mp: 250° C (dec.). ¹H NMR (300 MHz, DMSO-d₆) δ (ppm): 13.3 (bs exch. D₂O, 1H, COOH), 7.97 (d, *J* = 6.6 Hz, 2H, ArH); 7.93 (d, *J* = 6.6 Hz, 2H, ArH); 7.38 (d, *J* = 3.6 Hz, 1H, ArH); 7.35 (d, *J* = 3,6 Hz, 1H, ArH). ¹³C NMR (75 MHz, DMSO-d₆) δ (ppm): 159.5, 154.5, 145.9, 133.5, 133.6, 125.3, 120.2, 119.0, 111.4, 111.3. FTIR (KBr) ν cm⁻¹: 2849, 2668, 2566, 2227, 1690, 1674, 1608, 1589, 1535, 1483, 1422, 1373, 1316, 1275, 1221, 1167, 957, 845, 801, 761, 550. ESI-MS *m/z*: $C_{12}H_7NO_3$ calcd 213, found 212 [M-H]⁻. HPLC purity: t = 1.53', ≥ 98%.

5-(4-Carboxyphenyl)furan-2-carboxylic acid (**18**). 5-(4-Cyanophenyl)furan-2-carboxylic acid (**17**, 1 mmol) was further treated with potassium hydroxide (4.5 mmol) in water (12 mL) at reflux for 6 h. After completion, the solution was acidified with hydrochloric acid (10M) and then extracted with ethyl acetate (3-7 mL). The organic layers were washed with brine, dried over anhydrous sodium sulphate and then concentrated *in vacuo*. Light-yellow solid. Yield: 87%. TLC (dichloromethane – methanol 7:3 + 3 drops of acetic acid): R_f = 0.14. Mp > 300° C. ¹H NMR (300 MHz, DMSO-d₆) δ (ppm): 13.2-13.0 (bs exch. D₂O, 2H, COOH), 8.0 (d, *J* = 8.3 Hz, 2H, ArH), 7.9 (d, *J* = 8.3 Hz, 2H, ArH), 7.4 (d, *J* = 3,4 Hz, 1H, ArH), 7.3 (d, *J* = 3,4 Hz, 1H, ArH). ¹³C NMR (75 MHz, DMSO-d₆) δ (ppm): 167.4, 159.8, 155.8, 145.7, 133.5, 131.3, 130.8, 125.0, 120.4, 110.5. FTIR (KBr) ν cm⁻¹: 3431, 2963, 2922, 2852, 2671, 2546, 1675, 1611, 1588, 1539, 1483, 1428, 1306, 1276, 1218, 1167, 1024, 955, 858, 803, 722. ESI-MS *m/z*: $C_{12}H_8O_5$ calcd 232, found 231 [M-H]⁻. HPLC purity: t = 1.12', ≥ 98%.

5-(2-Chloro-4-trifluoromethyl-2-phenyl)furan-2-carboxylic acid (**20**). Starting compound: methyl 5-(2-chloro-4-trifluoromethyl-2-phenyl)furan-2-carboxylate (**83**). Procedure 1. White solid. Yield: 95%. TLC (dichloromethane – methanol 7:3): R_f = 0.56. Mp: 173.7°C. ¹H NMR (300 MHz, DMSO-d₆) δ (ppm): 14.2 (bs exch. D₂O, 1H, COOH), 8.87 (d, 1H, *J* = 8.2 Hz, ArH), 8.78 (s, 1H, ArH), 8.65 (d, 1H, *J* = 8.2 Hz, ArH), 8.20 (d, 1H, *J* = 3.6 Hz, ArH), 8.18 (d, 1H, *J* = 3.6 Hz, ArH). ¹³C

NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.43, 151.77, 146.69, 132.39, 131.98, 130.73-130.24-129.84-129.39 (q), 130.22, 129.39-125.41-121.81-118.19 (q), 128.78-125.17-121.54-117.91 (q), 127.54-127.11-126.68-126.25 (q), 124.39,119.55, 114.16. FTIR (KBr) ν cm^{-1}: 3430, 3180, 3139, 2963, 2918, 2851, 1684, 1628, 1591, 1532, 1490, 1434, 1347, 1267, 1097, 816. HRMS (ESI-QOrbitrap) *m/z*: C$_{12}$H$_6$ClF$_3$O$_3$ calcd 288.9884, found 288.9888 [M-H]$^-$.

5-(2,4-Bis(trifluoromethyl)2-phenyl)furan-2-carboxylic acid (**27**). Starting compound: methyl 5-(2,4-bis(trifluoromethyl)2-phenyl)furan-2-carboxylate (**84**). Procedure 1. Light brown solid. Yield: 56%. TLC (dichloromethane – methanol 7:3): R$_f$ = 0.57. Mp: 155.7°C. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.4-13.2 (bs exch. D$_2$O, 1H, COOH), 8.17-8.15 (m, 2H, ArH), 8.65 (d, 1H, *J* = 8.2 Hz, ArH), 7.36 (d, 1H, *J* = 3.7 Hz, ArH); 7.11 (d, 1H, *J* = 3.7 Hz, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.43, 151.77, 146.69, 132.39, 131.98, 130.73-130.24-129.84-129.39 (q), 130.22, 129.39-125.41-121.81-118.19 (q), 128.78-125.17-121.54-117.91 (q), 127.54-127.11-126.68-126.25(q), 124.39,119.55, 114.16. FTIR (KBr) ν cm^{-1}: 3430, 3180, 3139, 2963, 2918, 2851, 1684, 1628, 1591, 1532, 1490, 1434, 1347, 1267, 1097, 816. HRMS (ESI-QOrbitrap) *m/z*: C$_{13}$H$_6$F$_6$O$_3$ calcd 323.0148, found 323.0151 [M-H]$^-$.

5-(3,4-Bis(trifluoromethyl)2-phenyl)furan-2-carboxylic acid (**28**). Starting compound: methyl 5-(3,4-bis(trifluoromethyl)2-phenyl)furan-2-carboxylate (**85**). Procedure 1. White solid. Yield: 53%. TLC (dichloromethane – methanol 8:2): R$_f$ = 0.26. Mp: 178.0°C. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.5-13.3 (broad s exch D$_2$O, 1H, COOH), 8.20-8.17 (m, 2H, ArH) 8.10 (d, 1H, *J* = 8.4 Hz, ArH); 7.35 (d, 1H, *J* = 3.5 Hz, ArH); 7.11 (d, 1H, *J* = 3,5 Hz, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.84, 153.85, 130.58, 128.45-128.44-128.40-128.38 (q), 127.65-127.58-127.50-127.43 (q), .127.27-126.86-126.46-126.05 (q), 126.23, 122.61, 119.84, 112.76. FTIR (KBr) ν cm^{-1}: 3180, 3139, 2964, 2917, 2851, 2699, 2563, 2473, 1684, 1628, 1591, 1574, 1533, 1490, 1434, 1347, 1305, 1292, 1267, 1036, 816, 763, 749. HRMS (ESI- QOrbitrap) *m/z*: C$_{13}$H$_6$F$_6$O$_3$ calcd 323.0148, found 323.0153 [M-H]$^-$.

5-(3,5-Bis(trifluoromethyl)phenyl)furan-2-carboxylic acid (**29**). Starting compound: methyl 5-(3,5-bis(trifluoromethyl)phenyl)furan-2-carboxylate (**86**).

Procedure 1. White solid. Yield: 86%. TLC (dichloromethane – methanol 7:3): R_f = 0.39. Mp: 189.4°C. ^{1}H NMR (300 MHz, CD$_3$OD) δ (ppm): 8.39 (s, 2H, ArH), 7.93 (s, 1H, ArH), 7.32 (d, J = 3.6 Hz, 1H, ArH), 7.28 (d, J = 3.6 Hz, 1H, ArH).

5-(3-Chlorophenyl)furan-2-carboxylic acid (**46**). Starting compound: methyl 5-(3-chlorophenyl)furan-2-carboxylate (**87**). Procedure 1. Beige solid. Yield: 89%. TLC (dichloromethane – methanol 8:2): R_f = 0.27. Mp: 187°C (dec.). ^{1}H NMR (300 MHz, DMSO-d6) δ (ppm): 7.76 (t, J = 1.7 Hz, 1H, ArH), 7.67 (d, J =7.7, 1H, ArH), 7.44 (t, J = 7.7 Hz, 1H, ArH), 7.33 (d, J =7.7, 1H, ArH), 7.06 (d, J = 3.4 Hz, 1H, ArH), 6.88 (s, 1H, ArH). ^{13}C NMR (75 MHz DMSO-d6) δ:160.72, 160.45. 160.36, 151.98, 134.45, 132.97, 131.56, 128.09, 123.92, 122.95, 109.29. FTIR (KBR): ν cm^{-1}: 3412, 2961, 2920, 2851, 1674, 1660, 1602, 1525, 1461, 1420, 1384, 1305, 1278, 776. HRMS (ESI- QOrbitrap) *m/z*: C$_{11}$H$_7$ClO$_3$ calcd 221.0011, found 221.0022 [M-H]$^-$.

5-(3-Hydroxyphenyl)furan-2-carboxylic acid (**47**). Starting compound: methyl 5-(3-hydroxyphenyl)furan-2-carboxylate (**88**). Procedure 1. White solid. Yield: 94%. TLC (dichloromethane – methanol 7:3): R_f = 0.40. Mp: 170°C. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.10-13.00 (bs exch. D$_2$O, 1H, COOH), 9.80-9.60 (bs exch. D$_2$O, 1H, OH), 7.26 (d, J = 3.4 Hz, 1H, ArH), 7.18-7.06 (m, 3H, ArH), 7.04 (d, J = 3.4 Hz, 1H, ArH), 6.77 (dt, J = 7.3, 2.1 Hz, 1H, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.72, 158.30, 156.76, 144.46, 130.77, 130.75, 122.18, 116.55, 115.83, 111.33, 108.19. FTIR (KBr) ν cm^{-1}: 3401, 2919, 1675, 1649, 1582, 1533, 1470, 1418, 1309, 1298, 1284, 1200, 1160, 1025, 954, 852, 799, 785. HRMS (ESI-QOrbitrap) *m/z*: C$_{11}$H$_9$O$_4$ calcd 205.0495, found 205.0499 [M+H]$^+$.

5-(3-Carbamoylphenyl)furan-2-carboxylic acid (**50**). Starting compound: methyl 5-(3-carbamoylphenyl)furan-2-carboxylate (**89**). Procedure 1. White solid. Yield: 81%. TLC (dichloromethane – methanol 7:3): R_f = 0.29. Mp: 250°C. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.10-13.00 (bs exch. D$_2$O, 1H,COOH), 8.27 (s, 1H, ArH), 8.11 (bs exch. D$_2$O, 1H, NH$_2$), 7.92 (d, J = 7.8 Hz, 1H, ArH), 7.86 (d, J = 7.8 Hz, 1H, ArH), 7.55 (t, J = 7.8 Hz, 1H, ArH), 7.49 (bs exch. D$_2$O, 1H, NH$_2$), 7.33 (d, J = 3.6 Hz, 1H, ArH), 7.19 (d, J = 3.6 Hz, 1H, ArH). ^{13}C NMR (75 MHz, DMSO-d6) δ (ppm): 167.68, 159.69, 156.09, 144.93, 135.59, 129.69, 129.56, 128.19, 127.39, 123.90, 120.26, 108.92. FTIR (KBr) ν cm^{-1}: 3451, 3188, 2920, 2649, 2527, 1682,

1598, 1518, 1449, 1403, 1299, 1273, 1160, 1025, 958, 942, 795, 761. HRMS (ESI-QOrbitrap) m/z: $C_{12}H_{10}NO_4$ calcd 232.0604, found 232.0614 [M+H]$^+$.

5-(3-Cyanophenyl)furan-2-carboxylic acid (**54**). Starting compound: methyl 5-(3-cyanophenyl)furan-2-carboxylate (**90**). Procedure 2. White solid. Yield: 87%. TLC (dichloromethane – methanol 8:2): R_f = 0.20. Mp: 260°C (dec.). ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.30 (bs exch D$_2$O, 1H, COOH), 8.28 (s, 1H, ArH), 8.10 (d, J=7.8 Hz, 1H, ArH); 7.80 (d, J=7.8 Hz, 1H, ArH), 7.69 (t, J=7.8 Hz, 1H, ArH), 7.34 (s, 2H, ArH). ^{13}C NMR (75 MHz DMSO-d$_6$) δ (ppm): 159.61, 154.33, 145.60, 132.55, 130.87, 130.75, 129.00, 128.36, 120.12, 118.77, 112.81, 110.22. FTIR (KBr) ν cm^{-1}: 3112, 2915, 2850, 2666, 2576, 2231, 1710, 1686, 1608, 1572, 1519, 1475, 1436, 1318, 1289, 1232, 1174, 1033, 997, 819, 801, 761, 582. HRMS (ESI-QOrbitrap) m/z: $C_{12}H_7NO_3$ calcd 212.0353, found 212.0351 [M-H]$^-$.

4'-Nitro-[1,1'-biphenyl]-3-carboxylic acid (**58**). Starting compound: methyl 4'-nitro-[1,1'-biphenyl]-3-carboxylate (**91**). Procedure 1. White solid. Yield: 92%. TLC (dichloromethane – methanol 7:3): R_f = 0.46. Mp: 232°C. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.2 (bs exch. D$_2$O, 1H, COOH), 8.31 (d, J = 8.9 Hz, 2H, ArH), 8.26 (t, J = 1.6 Hz, 1H, ArH), 8.05-7.97 (m, 4H, ArH), 7.66 (t, J = 7.8 Hz, 1H, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 167.4, 147.5, 146.1, 138.6, 132.3, 132.1, 130.2, 128.6, 128.3, 124.7. FTIR (KBr) ν cm^{-1}: 3072, 2965, 2917, 1707, 1668, 1689, 1606, 1594, 1517, 1449, 1309, 1263, 1109, 856, 741. ESI-MS m/z: $C_{13}H_9NO_4$ calcd 243, found 242 [M-H]$^-$. HPLC purity: t = 8.91', ≥ 98%.

5-(4-Nitrophenyl)thiophene-2-carboxylic acid (**59**). Starting compound: methyl 5-(4-nitrophenyl)thiophene-2-carboxylate (**92**). Procedure 1. Yellow solid. Yield: 98%. TLC (dichloromethane – methanol 9:1): R_f = 0.20. Mp: 190°C. ^{1}H NMR (300 MHz, CD$_3$OD) δ (ppm): 8.30 (d, J = 8.9 Hz, 2H, ArH), 7.95 (d, J = 8.9 Hz, 2H, ArH), 7.79 (d, J = 3.9 Hz, 1H, ArH), 7.64 (d, J = 3.9 Hz, 1H, ArH).

5-(3-Cyanophenyl)thiophene-2-carboxylic acid (**60**). Starting compound: methyl 5-(3-cyanophenyl)thiophene-2-carboxylate (**93**). Procedure 2. White solid. Yield: 90%. TLC (dichloromethane – methanol 9:1): R_f = 0.20. Mp: 207°C. ^{1}H NMR (300 MHz, CD$_3$OD) δ (ppm): 8.08 (s, 1H, ArH), 7.99 (d, J = 7.8 Hz, 1H, ArH), 7.76 (d, J = 3.9 Hz, 1H, ArH), 7.72 (d, J = 7.8 Hz, 1H, ArH), 7.61 (t, J = 7.8 Hz, 1H, ArH), 7.54 (d, J = 3.9 Hz, 1H, ArH).

5.2.1.2 Synthesis of 9,15,16,19,21,23-26,30,31, 33-37,40-42,44,48,51-53,55,56 (General Scheme B)
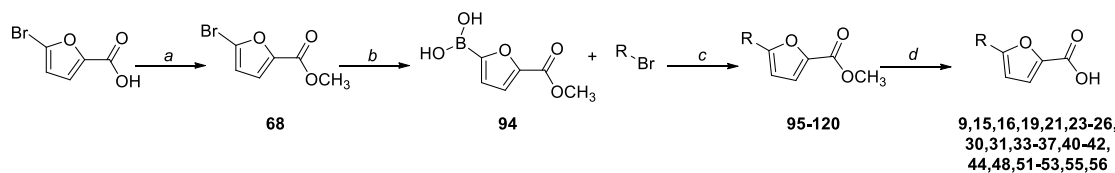
Reagents and conditions: *a)* conc. H_2SO_4, MeOH, reflux, 24 h; *b)* 1. bis[2-(*N*,*N*-dimethylamino)ethyl] ether, 2M *i*-PrMgCl, THF, 20 min. 10-15°C – 30 min. r.t., N_2; 2. $B(OCH_3)_3$, 0°C, 10 min., N_2; *c)* $Pd(PPh_3)_2Cl_2$, 2M Na_2CO_3, 1,4-dioxane, 60°C, 90 min., MW, N_2; *d)* NaOH, EtOH/THF 1:1, reflux, 5 h *or* LiOH·H_2O, THF/H_2O 1:1, 20°C, 2 h.

Synthesis of methyl 5-bromofuran-2-carboxylate (68, *a*). See the corresponding reaction in General Scheme A (*a*).

Synthesis of (5-(methoxycarbonyl)furan-2-yl)boronic acid (94, *b*). Isopropylmagnesium chloride 2M in THF (1.2 mmol) was added to a solution of bis[2-(*N*,*N*-dimethylamino)ethyl] ether (1.2 mmol) in THF (5.1 mL) under nitrogen atmosphere. The resulting mixture was stirred for 20 minutes at 10-15°C, before methyl 5-bromofuran-2-carboxylate (**68**, 1 mmol) was added; the stirring was continued at room temperature for 30 minutes. Then, trimethyl borate (2 mmol) was added at 0°C and the reaction mixture was stirred for 10 minutes. After quenching with diluted HCl (1M), the reaction was extracted with ethyl acetate and the organic layers were washed with brine, dried over anhydrous sodium sulphate and evaporated *in vacuo*. The resulting brown solid was purified *via* crystallisation from hexane and ethyl acetate, to afford the desired product as a beige solid. Yield: 75%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.54. Mp: 128°C (dec.). ^{1}H-NMR (300 MHz, CDCl₃) δ (ppm): 7.19 (d, *J* = 3.5 Hz, 1H, ArH), 7.08 (d, *J* = 3.5 Hz, 1H, ArH), 3.91 (s, 3H, CH₃).[33,38]

General procedure for the microwave-assisted Suzuki-Miyaura coupling (*c*). Methyl 5-boronofuran-2-carboxylate (**94**, 1.3 mmol), the appropriate bromo derivatives (1.0 mmol) and bis(-triphenylphosphine)palladium(II) dichloride (5% mol) were dissolved in dry 1,4-dioxane (10 mL), under nitrogen atmosphere. A 2M sodium carbonate solution (2 mmol) was then added, and the resulting mixture was stirred in a microwave

synthesizer (Biotage®Initiator Classic) for 1 h at 60°C. After completion, the solution was cooled to room temperature and then filtered on a celite pad. The filtrate was diluted with water and extracted with ethyl acetate (3-4 mL). The organic layer was dried over anhydrous sodium sulphate, filtered and concentrated *in vacuo*.[33,38]

Methyl 5-(4-aminophenyl)furan-2-carboxylate (**95**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 4-bromoaniline. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 6:4) to give the desired product as an amber solid. Yield: 61%. TLC (cyclohexane – ethyl acetate 6:4): R_f = 0.33. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.59 (d, J = 8.7 Hz, 2H, ArH), 7.21 (d, J = 3.7 Hz, 1H, ArH), 6.70 (d, J = 8.7 Hz, 2H, ArH), 6.52 (d, J = 3.7 Hz, 1H, ArH), 3.89 (s, 3H, CH$_3$).

Methyl 5-(benzo[c][1,2,5]oxadiazol-5-yl)furan-2-carboxylate (**96**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 5-bromobenzo[*c*][1,2,5]oxadiazole (**141**). The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a brown solid. Yield: 30%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.42. Mp: 177°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm) 8.25 (s, 1H, ArH), 7.91 (dd, J = 9.4, 1.3 Hz, 1H, ArH), 7.80 (dd, J = 9.4, 1.3 Hz, 1H, ArH), 7.30 (d, J = 3.7 Hz, 1H, ArH), 6.97 (d, J = 3.7 Hz, 1H, ArH), 3.95 (s, 3H, CH$_3$).

Methyl 5-(1-oxo-1,3-dihydroisobenzofuran-5-yl)furan-2-carboxylate (**97**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 5-bromoisobenzofuran-1(3*H*)-one (**143**). The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 7:3) to give the desired product as a white solid. Yield: 21%. TLC (cyclohexane – ethyl acetate 7:3): R_f = 0.27. Mp: 218°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm) 8.07 – 7.81 (m, 3H, ArH), 7.28 (d, J = 3.7 Hz, 1H, ArH), 6.93 (d, J = 3.7 Hz, 1H, ArH), 5.36 (s, 2H, CH$_2$), 3.94 (s, 3H, CH$_3$).

Methyl 5-(4-(2H-tetrazol-5-yl)phenyl)furan-2-carboxylate (**98**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 5-(4-bromophenyl)-2*H*-tetrazole (**144**). The reaction work-up was modified as follows: after completion, the reaction was acidified with hydrochloric acid (1M) and then extracted with ethyl acetate (7 x 4 mL); the combined organic layers

were dried over anhydrous sodium sulphate, filtered and concentrated *in vacuo*. The crude was purified by flash column chromatography (dichloromethane – methanol 8:2) to give the desired product as a greyish solid. Yield: 94%. TLC (dichloromethane – methanol 8:2): R_f =0.20. Mp: 220°C (dec.). ^{1}H-NMR (300 MHz, CD$_3$OD) δ (ppm) 8.12 (d, J = 8.6 Hz, 2H, ArH), 7.91 (d, J = 8.6 Hz, 2H, ArH), 7.33 (d, J = 3.7 Hz, 1H, ArH), 7.02 (d, J = 3.7 Hz, 1H, ArH), 3.91 (s, 3H, CH$_3$).

Methyl 5-(2-fluoro-4-(trifluoromethyl)phenyl)furan-2-carboxylate (**99**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 1-bromo-2-fluoro-4-(trifluoromethyl)benzene. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 95:5) to give the desired product as a white solid. Yield: 61%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.60. Mp: 64°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.13 (t, 1H, J = 7.7 Hz, ArH), 7.51 (d, 1H, J = 8.3 Hz, ArH), 7.43 (d, 1H, J = 8.3 Hz, ArH), 7.30 (d, 1H, J = 3.7 Hz, ArH), 7.05 (t, 1H, J = 3.7 Hz, ArH), 3.94 (s, 3H, CH$_3$).

Methyl 5-(2-hydroxy-4-(trifluoromethyl)phenyl)furan-2-carboxylate (**100**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 2-bromo-5-(trifluoromethyl)phenol. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) to give the desired product as a yellow solid. Yield: 48%. TLC (cyclohexane – ethyl acetate 6:4): R_f = 0.66. Mp: 232.0°C. ^{1}H-NMR (300 MHz, D$_3$COD$_3$) δ (ppm): δ 10.13 (bs exch. D$_2$O, 1H, OH), 8.05 (dd, 1H, J = 8.8, 0.9 Hz, ArH), 7.38-732 (m, 3H, ArH), 7.29 (d, 1H, J = 3.6 Hz, ArH), 3.89 (s, 3H, CH$_3$).

Methyl 5-(2-methyl -4-(trifluoromethyl)phenyl)furan-2-carboxylate (**101**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 1-bromo-2-methyl-4-(trifluoromethyl)benzene. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 98:2) to give the desired product as a white solid. Yield: 85%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.58 Mp: 66.0°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.90 (d, 1H, J = 8.7 Hz, ArH), 7.56-7.50 (m, 2H, ArH), 7.30 (d, 1H, J = 3.6 Hz, ArH), 6.74 (d, 1H, J = 3.6 Hz, ArH), 3.93 (s, 3H, OCH$_3$), 2.59 (s, 3H, CH$_3$).

Methyl 5-(2-amino -4-(trifluoromethyl)phenyl)furan-2-carboxylate (**102**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 2-

bromo-5-(trifluoromethyl)aniline. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) to give the desired product as a pale yellow solid. Yield: 24%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.63. Mp: 154°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ 7.60 (d, 1H, J = 7.7 Hz, ArH), 7.28 (d, 1H, J = 3.7 Hz, ArH), 7.02-7.00 (m, 2H, ArH), 6.78 (d, J = 3.7 Hz, 1H, ArH), 4.19 (bs exch. D$_2$O, 2H, NH$_2$), 3.92 (s, 3H, CH$_3$).

Methyl 5-(2-cyano-4-(trifluoromethyl)phenyl)furan-2-carboxylate (**103**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 2-bromo-5-(trifluoromethyl)benzonitrile. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 95:5) to give the desired product as a white solid. Yield: 51%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.49. Mp: 142°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.24 (d, J = 8.3 Hz, 1H, ArH), 8.00 (s, 1H, ArH), 7.92 (d, 1H, J = 8.4 Hz, ArH), 7.58 (d, 1H, J = 3.8 Hz, ArH), 7.33 (d, 1H, J = 3.8 Hz, ArH), 3.96 (s, 3H, CH$_3$).

Methyl 5-(2-chloro-4-fluorophenyl)furan-2-carboxylate (**104**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 1-bromo-2-chloro-4-fluorobenzene. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) to give the desired product as a white solid. Yield: 52%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.38. Mp: 87.2°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.97 (dd, J = 8.9, 6.1 Hz, 1H, ArH), 7.28 (s, 1H, ArH), 7.22 (dd, J = 8.4, 2.6 Hz, 1H, ArH), 7.16 – 7.04 (m, 2H, ArH), 3.92 (s, 3H, CH$_3$).

Methyl 5-(2,4-difluorophenyl)furan-2-carboxylate (**105**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 1-bromo-2,4-difluorobenzene. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 95:5) to give the desired product as a white solid. Yield: 38%. TLC (cyclohexane – ethyl acetate 9:1): R_f = 0.51. Mp: 94.6°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.97 (ddd, J = 9.4, 7.2, 3.6 Hz, 1H, ArH), 7.38 – 7.18 (m, 1H, ArH), 7.13 – 6.72 (m, 3H, ArH), 3.93 (s, 3H, CH$_3$).

Methyl 5-(4-fluoro-2-hydroxyphenyl)furan-2-carboxylate (**106**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 2-bromo-5-fluorophenol. The crude was purified by flash column chromatography

(cyclohexane – ethyl acetate 8:2) to give the desired product as a pale yellow solid. Yield: 29%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.25. Mp: 139.2°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.73 – 7.52 (m, 1H, ArH), 7.29 (d, J = 3.7 Hz, 1H, ArH), 6.87 – 6.57 (m, 3H, ArH), 3.92 (s, 3H, CH$_3$), 1.62 (s, 1H, OH).

Methyl 5-(4-fluoro-2-methylphenyl)furan-2-carboxylate (**107**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 1-bromo-4-fluoro-2-methylbenzene. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) to give the desired product as a pale yellow oil. Yield: 63%. TLC (cyclohexane – ethyl acetate 9:1): R_f = 0.26. Mp: 44.3°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.79 – 7.64 (m, 1H, ArH), 7.26 (d, J = 3.6 Hz, 1H, ArH), 7.05 – 6.89 (m, 2H, ArH), 6.58 (d, J = 3.6 Hz, 1H, ArH), 3.92 (s, 3H, OCH$_3$), 2.51 (s, 3H, CH$_3$).

Methyl 5-(2-amino-4-fluorophenyl)furan-2-carboxylate (**108**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 2-bromo-5-fluoroaniline. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a yellow solid. Yield: 24%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.23. Mp: 123.7°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.55 – 7.42 (m, 1H, ArH), 7.26 (d, J = 4.1 Hz, 1H, ArH), 6.64 (d, J = 3.7 Hz, 1H, ArH), 6.60 – 6.48 (m, 2H, ArH), 4.30 (d, 2H, NH$_2$), 3.89 (s, 3H, CH$_3$).

Methyl 5-(2-cyano-4-fluorophenyl)furan-2-carboxylate (**109**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 2-bromo-5-fluorobenzonitrile. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) to give the desired product as a light yellow solid. Yield: 44%. TLC (cyclohexane – ethyl acetate 9:1): R_f = 0.31. Mp: 189.6°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.11 – 8.03 (m, 1H, ArH), 7.47 – 7.34 (m, 3H, ArH), 7.28 (d, J = 3.7 Hz, 1H, ArH), 3.93 (s, 3H, CH$_3$).

Methyl 5-(4-fluoro-2-(trifluoromethyl)phenyl)furan-2-carboxylate (**110**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 1-bromo-4-fluoro-2-(trifluoromethyl)benzene. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) to give the desired product as a light yellow solid. Yield: 29%. TLC (cyclohexane – ethyl acetate 8:2):

R$_f$ = 0.53. Mp: 47.9°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.82 (dd, J = 8.7, 5.4 Hz, 1H, ArH), 7.57 – 7.42 (m, 1H, ArH), 7.40 – 7.21 (m, 2H, ArH), 6.74 (d, 1H, J = 3.7 Hz, ArH), 3.91 (s, 3H, CH$_3$).

Methyl 5-(2-hydroxy-4-nitrophenyl)furan-2-carboxylate (**111**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 2-bromo-5-nitrophenol. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 6:4) to give the desired product as a yellow solid. Yield: 59%. TLC (cyclohexane – ethyl acetate 6:4): R$_f$ = 0.17. Mp: 238°C. ^{1}H-NMR (300 MHz, CD$_3$COCD$_3$) δ (ppm): 10.45 (bs exch. D$_2$O, 1H, OH), 8.12–8.06 (m, 1H, ArH), 7.93-7.88 (m, 2H, ArH), 7.42-7.36 (m, 2H, ArH), 3.90 (s, 3H, CH$_3$).

Methyl 5-(2-methyl-4-nitrophenyl)furan-2-carboxylate (**112**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 1-bromo-2-methyl-4-nitrobenzene. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) to give the desired product as a yellow solid. Yield: 22%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.54. Mp: 175.2°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.19-8.06 (m, 2H, ArH), 7.97 (d, 1H, J = 8.5 Hz, ArH), 7.31 (d, 1H, J = 3.6 Hz, ArH), 6.84 (d, 1H, J = 3.6 Hz, ArH), 3.94 (s, 3H, OCH$_3$), 2.65 (s, 3H, CH$_3$).

Methyl 5-(2-amino-4-nitrophenyl)furan-2-carboxylate (**113**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 2-bromo-5-nitroaniline. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as red needles. Yield: 28%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.22. Mp: 256.3°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.67 – 7.55 (m, 3H, ArH), 7.31 (d, 1H, J = 3.6 Hz, ArH), 6.85 (d, 1H, J = 3.6 Hz, ArH), 3.93 (s, 3H, CH$_3$).

Methyl 5-(4-nitro-2-(trifluoromethyl)phenyl)furan-2-carboxylate (**114**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 1-bromo-4-nitro-2-(trifluoromethyl)benzene. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a brown solid. Yield: 26%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.33. Mp: 241.2°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.66 (d, 1H, J = 2.3 Hz, ArH),

8.47 (dd, 1H, J = 8.7, 2.3 Hz, ArH), 8.15 (d, 1H, J = 8.7 Hz, ArH), 7.30 (d, 1H, J = 3.7 Hz, ArH), 7.04 (d, 1H, J = 3.7 Hz, ArH), 3.95 (s, 3H, CH$_3$).

Methyl 5-(m-tolyl)furan-2-carboxylate (**115**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 1-bromo-3-methylbenzene. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a white solid. Yield: 52%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.56. Mp: 76.0°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.62 (s, 1H, ArH), 7.57 (d, J = 7.7 Hz, 1H, ArH), 7.30 (t, J = 7.7 Hz, 1H, ArH), 7.24 (d, J = 3.6 Hz, 1H, ArH), 7.16 (d, J = 7.7 Hz, 1H, ArH), 6.72 (d, J = 3.6 Hz, 1H, ArH), 3.92 (s, 3H, OCH$_3$), 2.40 (s, 3H, CH$_3$).

Methyl 5-(3-(methylcarbamoyl)phenyl)furan-2-carboxylate (**116**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 3-bromo-*N*-methylbenzamide (**146**). The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 5:5) to give the desired product as a yellow solid. Yield: 32%. TLC (cyclohexane – ethyl acetate 5:5): R$_f$ = 0.15. Mp: 76.0°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.11 (s, 1H, ArH), 7.86 (d, J = 7.8 Hz, 1H, ArH), 7.73 (d, J = 7.8 Hz, 1H, ArH), 7.45 (t, J = 7.8 Hz, 1H, ArH), 7.23 (d, J = 3.6 Hz, 1H, ArH), 6.78 (d, J = 3.6 Hz, 1H, ArH), 6.45 (bs, 1H, NH), 3.91 (s, 3H, OCH$_3$), 3.03 (d, J = 4.6 Hz, 3H, NHC*H*$_3$).

Methyl 5-(3-sulfamoylphenyl)furan-2-carboxylate (**117**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 3-bromobenzenesulfonamide. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a white solid. Yield: 77%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.41. Mp: 238.0°C. ^{1}H-NMR (300 MHz, D$_3$COD$_3$) δ (ppm): 8.32 (t, J = 1.8 Hz, 1H, ArH), 8.06 (ddd, J = 7.8, 1.8, 1.1 Hz, 1H, ArH), 7.91 (ddd, J = 7.8, 1.8, 1.1 Hz, 1H, ArH), 7.69 (t, J = 7.8 Hz, 1H, ArH), 7.36 (d, J = 3.7 Hz, 1H, ArH), 7.21 (d, J = 3.7 Hz, 1H, ArH), 6.73 (bs, 2H, SO$_2$N*H*$_2$), 3.90 (s, 3H, CH$_3$).

Methyl 5-(3-(methoxycarbonyl)phenyl)furan-2-carboxylate (**118**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and methyl 3-bromobenzoate. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a white solid.

Yield: 48%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.69. Mp: 115.1°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.42 (s, 1H, ArH), 8.04-7.98 (m, 2H, ArH), 7.52 (t, J = 7.8 Hz, 1H, ArH), 7.28 (d, partially hidden by solvent peak, 1H, ArH), 6.84 (d, J = 3.6 Hz, 1H, ArH), 3.96 (s, 3H, CH$_3$), 3.93 (s, 3H, CH$_3$).

Methyl 5-(3-cyano-5-methoxyphenyl)furan-2-carboxylate (**119**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 3-bromo-5-methoxybenzonitrile. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a white solid. Yield: 21%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.24. Mp: 157.9°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.63 (s, 1H, ArH), 7.51 (s, 1H, ArH), 7.27-7.26 (m, 1H, hidden by solvent peak, ArH), 7.12 (s, 1H, ArH), 6.81 (d, J = 3.6 Hz, 1H, ArH), 3.94 (s, 3H, CH$_3$), 3.91 (s, 3H, CH$_3$).

Methyl 5-(3-cyano-5-phenoxyphenyl)furan-2-carboxylate (**120**). Starting compounds: (5-(methoxycarbonyl)furan-2-yl)boronic acid (**94**) and 3-bromo-5-phenoxybenzonitrile (**147**). The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) to give the desired product as a white solid. Yield: 53%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.39. Mp: 130.6°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.76 (s, 1H, ArH), 7.62 (s, 1H, ArH), 7.45 – 7.40 (m, 2H, ArH), 7.29 – 7.18 (m, 2H, ArH), 7.15 – 7.00 (m, 3H, ArH), 6.79 (d, J = 3.6 Hz, 1H, ArH), 3.93 (s, 3H, ArH).

General procedure for the base-catalysed hydrolysis (*d*). See the corresponding reaction in General Scheme A (*c*). Variations to the general procedures were applied to the synthesis of 5-(1-oxo-1,3-dihydroisobenzofuran-5-yl)furan-2-carboxylic acid (**16**).[33]

5-(4-Aminophenyl)furan-2-carboxylic acid (**9**). Starting compound: methyl 5-(4-aminophenyl)furan-2-carboxylate (**95**). Procedure 1. Brownish solid. Yield: 71%. TLC (dichloromethane – methanol 7:3 + 3 drops of acetic acid): R_f = 0.47. Mp: 158° C (dec.). ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 12.8 (broad s. exch. D$_2$O, 1H, COOH), 7.42 (d, J = 8.5 Hz, 2H, ArH), 7.17 (d, J = 3.4 Hz, 1H, ArH), 6.70 (d, J = 3,4 Hz, 1H, ArH) 6.65 (d, J = 8.5 Hz, 2H, ArH), 5.40 (broad s. exch. D$_2$O, 2H, NH$_2$). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 158.54, 150.48, 126.44, 120.55, 117.56, 114.42, 104.51. FTIR (KBr): ν = 3431, 2920, 2851, 1621, 1603, 1587, 1569, 1539,

1488, 1431, 1375, 1302, 1277, 1020, 962, 810, 800, 783, 558, 503 cm^{-1}. ESI-MS *m/z*: C$_{11}$H$_9$NO$_3$ calcd 203, found 204 [M+H]$^+$, 227 [M+Na]$^+$.

5-(Benzo[c][1,2,5]oxadiazol-5-yl)furan-2-carboxylic acid (**15**). Starting compound: methyl 5-(benzo[*c*][1,2,5]oxadiazol-5-yl)furan-2-carboxylate (**96**). Procedure 1. Pale brown solid. Yield: 51%. TLC (dichloromethane – methanol 8:2): R$_f$ = 0.35. Mp: 255°C (dec.). ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.3 (bs exch. D$_2$O, 1H, COOH), 8.29 (s, 1H, ArH), 8.14 (d, *J* = 9.4, 1H, ArH), 8.04 (dd, *J* = 9.4, 1.0 Hz, 1H, ArH); 7.51 (d, *J* = 3,7 Hz, 1H, ArH), 7.38 (d, *J* = 3.7, 1H, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.4, 154.0, 149.5, 148.9, 146.4, 132.5, 130.5, 120.2, 117.7, 112.9, 109.9. FTIR (KBr) ν cm^{-1}: 3435, 3128, 3078, 2962, 2919, 2850, 2371, 1742, 1684, 1623, 1583, 1464, 1377, 1299, 1258, 803, 759. ESI-MS *m/z*: C$_{11}$H$_6$N$_2$O$_4$ calcd 230, found 229 [M-H]$^-$, 459 [2M-H]$^-$. HPLC purity: t = 7.25', ≥ 98%.

5-(1-Oxo-1,3-dihydroisobenzofuran-5-yl)furan-2-carboxylic acid (**16**). Methyl 5-(1-oxo-1,3-dihydroisobenzofuran-5-yl)furan-2-carboxylate (**97**, 1.0 mmol) was dissolved in a mixture of tetrahydrofuran – ethanol 1:1 (15 mL) and a 0.5M solution of sodium hydroxide (2.5 mmol) was added dropwise while stirring. The reaction mixture stirred at room temperature for 30 min. After completion, the solvent was evaporated under reduced pressure; the aqueous phase was washed with chloroform (1 x 5 mL), acidified with hydrochloric acid (3M) and then extracted with ethyl acetate (3 x 7 mL). The organic layers were washed with brine, dried over anhydrous sodium sulphate and then concentrated *in vacuo*. The resulting solid was solubilized in methanol and purified by preparative HPLC (column: EVO C18 Phenomenex; eluent system: aqueous formic acid (0.05%)/acetonitrile 50:50). White solid. Yield: 72%. TLC (dichloromethane – methanol 7:3): R$_f$ = 0.11. Mp: 268°C (dec.). ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.0 (bs exch. D$_2$O, 1H, COOH), 8.12 (s, 1H, ArH), 7.92 (d, *J* = 8.1, 1H, ArH), 7.76 (dd, 1H, *J* = 8.1, 1.7 Hz, ArH), 7.30 (d, *J* = 3.6 Hz, 1H, ArH), 7.20 (d, *J* = 3.6 Hz, 1H, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 168.2, 159.7, 155.9, 146.0, 145.4, 132.4, 131.5, 128.6, 122.7, 122.6, 120.2, 109.9, 61.5. FTIR (KBr) ν cm^{-1}: 3420, 2961, 2918, 2850, 1682, 1610, 1586, 1565, 1524, 1472, 1307, 1260,

1220, 1029, 802, 790, 761. ESI-MS *m/z*: $C_{13}H_8O_5$ calcd 244, found 261 [M-H+H_2O]⁻. HPLC purity: t = 5.51, ≥ 98%.

5-(4-(2H-Tetrazol-5-yl)phenyl)furan-2-carboxylic acid (**19**). Starting compound: methyl 5-(4-(2*H*-tetrazol-5-yl)phenyl)furan-2-carboxylate (**98**). Procedure 1. The crude was purified by preparative TLC (dichloromethane – methanol 8:2 + 10 drops of acetic acid / 100 mL) to give the desired product as a white solid. Yield: 42%. TLC (dichloromethane – methanol 7:3): R_f = 0.10. Mp > 300 °C. ¹H NMR (300 MHz, DMSO-d₆) δ (ppm): 8.07 (d, *J* = 8.1 Hz, 2H, ArH), 7.80 (d, *J* = 8.1 Hz, 2H, ArH), 7.17 (d, *J* = 3,3 Hz, 1H, ArH), 7.07 (d, *J* = 3,3 Hz, 1H, ArH). ¹³C NMR (75 MHz, DMSO-d₆) δ (ppm): 160.6, 160.3, 155.8, 146.7, 132.8, 128.6, 126.7, 124.5, 118.7, 107.7. FTIR (KBr) ν cm⁻¹: 3429, 2955, 2920, 2371, 2355, 2342, 2325, 1617, 1606, 1568, 1489, 1380, 1278, 1175, 782, 754. ESI-MS *m/z*: $C_{12}H_7N_4O_3$ calcd 256, found 255 [M-H]⁻, 397 [2M+Na]⁺. HPLC purity: t = 5.43', ≥ 98%.

5-(2-Fluoro-4-(trifluoromethyl)phenyl)furan-2-carboxylic acid (**21**). Starting compound: methyl 5-(2-fluoro-4-(trifluoromethyl)phenyl)furan-2-carboxylate (**99**). Procedure 1. White solid. Yield: 80%. TLC (dichloromethane – methanol 9:1): R_f = 0.17. Mp: 175°C. ¹H NMR (300 MHz, DMSO-d₆) δ (ppm): 13.6-13.2 (broad s exch D_2O, 1H, COOH), 8.05 (t, 1H, *J* = 7.7 Hz , ArH), 7.87 (dd, 1H, *J* = 11.29, 1.0 Hz, ArH), 7. 71 (dd, 1H, *J* = 8.12, 1.0 Hz, ArH), 7.35 (d, 1H, *J* = 3.7 Hz, ArH), 7.14 (t, 1H, *J* = 3.7 Hz, ArH). ¹³C NMR (75 MHz, DMSO-d₆) δ (ppm): 160.05, 159.57, 156.1, 149.91, 146.0, 130.80-130.69 (d), 130.36-130.25 (d), 129.89, 128.05-128.01 (d), 125.40, 122.6-122.57-122.50 (t), 121.79, 121.52-121.36 (d), 119.84, 114.37-114.68 (d), 114.38, 114.24. FTIR (KBr) ν cm⁻¹: 3429, 2962, 2924, 2853, 1707, 1675, 1628, 1596, 1542, 1490, 1436, 1333, 1170, 808. HRMS (ESI-QOrbitrap) *m/z*: $C_{12}H_6F_4O_3$ calcd 273.0180, found 273.0183 [M-H]⁻.

5-(2-Hydroxy-4-(trifluoromethyl)phenyl)furan-2-carboxylic acid (**23**). Starting compound: methyl 5-(2-hydroxy-4-(trifluoromethyl)furan-2-carboxylate (**100**). Procedure 1. Yellow solid. Yield: 70%. TLC (dichloromethane – methanol 7:3): R_f = 0.28. Mp: 221°C. ¹H NMR (300 MHz, DMSO-d₆) δ (ppm): 13.2-13.0 (bs exch. D_2O, 1H, COOH), 11.2-11.0 (bs exch. D_2O, 1H, OH), 7.92 (dd, 1H, *J* = 8.5, 0.55 Hz, ArH), 7.31 (d, 1H, *J* = 3.6 Hz, ArH), 7.28-7.24 (m, 2H, H8, ArH),

7.20 (d, 1H, J = 3.6 Hz, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.68, 154.96, 152.33, 144.25, 130.39-129.97-129.55-129.12 (q), 129.73-126.12-122.52-118.91 (q), 127.14, 120.30, 120.03, 116.36-116.31, 113.51, 113.04-112.98. FTIR (KBr) ν cm^{-1}: 3392, 3174, 3033, 2857, 1656, 1625, 1594, 1534, 1494, 1438, 1421, 1326, 1308, 1168, 811. HRMS (ESI- QOrbitrap) m/z: C$_{12}$H$_7$F$_3$O$_4$ calcd 271.0224, found 271.0226 [M-H]$^-$.

5-(2-Methyl -4-(trifluoromethyl)phenyl)furan-2-carboxylic acid (**24**). Starting compound: methyl 5-(2-methyl -4-(trifluoromethyl)phenyl)furan-2-carboxylate (**101**). Procedure 1. Orange solid. Yield: Quantitative. TLC (dichloromethane – methanol 7:3): R$_f$ = 0.30. Mp: 131°C. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.4-13.2 (bs exch. D$_2$O, 1H, COOH), 7.92 (d, 1H, J = 8.2, ArH), 7.70 (s, 1H, ArH), 7.65 (d, 1H, J = 8.3 Hz, ArH), 7.35 (d, 1H, J = 3.6, ArH), 7.07 (d, 1H, J = 3.6 Hz, ArH), 2.49 (s, 3H, CH$_3$). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.67, 154.59, 145.28, 132.63, 129.92-126.31-122.69-119.09 (q), 129.70-129.29-128.86-128.45 (q), 128.49-128.45-128.40-128.34 (q), 123.49-123.43-123.38 (q), 119.66, 113.11, 18.93. FTIR (KBr) ν cm^{-1}: 3446, 3163, 2918, 1696, 1674, 1590, 1532, 1491, 1429, 1333, 1311, 1169, 803. HRMS (ESI- QOrbitrap) m/z: C$_{13}$H$_9$F$_3$O$_3$ calcd 269.0431, found 269.0434 [M-H]$^-$.

5-(2-Amino-4-(trifluoromethyl)phenyl)furan-2-carboxylic acid (**25**). Starting compound: methyl 5-(2-amino-4-(trifluoromethyl)phenyl)furan-2-carboxylate (**102**). Procedure 1. Yellow solid. Yield: 70%. TLC (dichloromethane – methanol 7:3): R$_f$ = 0.16. Mp: 175°C. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.2-13.0 (bs exch. D$_2$O, 1H, COOH), 7.68 (d, 1H, J = 8.3 Hz, ArH), 7.55 (d, 1H, J = 3.6 Hz, ArH), 7.33 (d, 1H, J = 3.6 Hz, ArH), 7.09 (d, 1H, J = 1.1 Hz, ArH), 6.92 (dd, 1H, J = 8.2, 1.1 Hz, ArH), 5.03 (bs exch. D$_2$O, 2H, NH$_2$). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.59, 154.91, 145.97, 144.16, 130. 08-130.39-129.97-129.55 (q), 128.73, 127.65-126.41-127.50-127.43 (q), 130.02-126.86-122.79-119.18 (q), 119.74, 116.32, 113.34-113.28, 112.72-112.67, 110.29. FTIR (KBr) ν cm^{-1}: 3513, 3435, 3405, 3384, 2961, 2919, 2850, 1665, 1633, 1591, 1523, 1486, 1448, 1351, 1248, 1167, 805, 795. HRMS (ESI- QOrbitrap) m/z: C$_{12}$H$_8$F$_3$NO$_3$ calcd 270.0385, found 270.0387 [M-H]$^-$.

5-(2-Cyano-4-(trifluoromethyl)phenyl)furan-2-carboxylic acid (**26**). Starting compound: methyl 5-(2-cyano-4-(trifluoromethyl)phenyl)furan-2-carboxylate (**103**). Procedure 2. White solid. Yield: Quantitative. TLC (dichloromethane – methanol 9:1): R_f = 0.28. Mp: 181°C. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.5-13.4 (broad s exch D$_2$O, 1H, COOH), 8.43 (d, 1H, J = 1.7 Hz, ArH), 8.16-8.14 (m, 2H, H10, ArH), 7.50 (d, 1H, J = 3.7 Hz, ArH), 7.42 (d, 1H, J = 3.7 Hz, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.30, 151.14, 146.63, 134.92, 132.29-132.24 (d), 130.88-130.83 (d), 130.23-129.79-129.34-128.89 (q), 128.86-125.24-121.63-118.01 (q), 128.24, 119.85, 117.43, 114.21, 108.29. FTIR (KBr) ν cm^{-1}: 3420, 3060, 2962, 2917, 2850, 1692, 1615, 1586, 1569, 1534, 1484, 1421, 1334, 1316, 1149, 816. HRMS (ESI- QOrbitrap) *m/z*: C$_{13}$H$_6$F$_3$NO$_3$ calcd 280.0227, found 280.0230 [M-H]$^-$.

5-(2-Chloro-4-fluorophenyl)furan-2-carboxylic acid (**30**). Starting compound: methyl 5-(2-chloro-4-fluorophenyl)furan-2-carboxylate (**104**). Procedure 1. White solid. Yield: 95%. TLC (dichloromethane – methanol 8:2): R_f = 0.34. Mp > 300°C (dec.). ^{1}H NMR (300 MHz, CD$_3$OD) δ (ppm): 8.17 (dd, J = 8.9, 6.2 Hz, 1H, ArH), 7.31 (dd, J = 8.7, 2.6 Hz, 1H, ArH), 7.24 – 7.14 (m, 1H, ArH), 7.12 (d, J = 3.5 Hz, 1H, ArH), 7.02 (d, J = 3.5 Hz, 1H, ArH).

5-(2,4-Difluorophenyl)furan-2-carboxylic acid (**31**). Starting compound: methyl 5-(2,4-difluorophenyl)furan-2-carboxylate (**105**). Procedure 1. White solid. Yield: Quantitative. TLC (dichloromethane – methanol 9:1): R_f = 0.23. Mp: 215°C. ^{1}H NMR (300 MHz, CD$_3$OD) δ (ppm): 8.22 – 7.92 (m, 1H, ArH), 7.31 (d, J = 3.7 Hz, 1H, ArH), 7.12 (m, 2H, ArH), 6.92 (t, J = 3.7 Hz, 1H, ArH).

5-(4-Fluoro-2-hydroxyphenyl)furan-2-carboxylic acid (**33**). Starting compound: methyl 5-(4-fluoro-2-hydroxyphenyl)furan-2-carboxylate (**106**). Procedure 2. White solid. Yield: Quantitative. TLC (dichloromethane – methanol 8:2): R_f = 0.17. Mp: 256.4°C. ^{1}H NMR (300 MHz, CD$_3$OD) δ (ppm): 7.89 (dd, J = 8.4, 6.7 Hz, 1H, ArH), 7.27 (d, J = 3.6 Hz, 1H, ArH), 7.05 (d, J = 3.6 Hz, 1H, ArH), 6.77 – 6.59 (m, 2H, ArH).

5-(4-Fluoro-2-methylphenyl)furan-2-carboxylic acid (**34**). Starting compound: methyl 5-(4-fluoro-2-methylphenyl)furan-2-carboxylate (**107**). Procedure 1. White solid. Yield: 99%. TLC (dichloromethane – methanol 8:2): R_f

= 0.33. Mp: 181.2°C. ^{1}H NMR (300 MHz, CD$_3$OD) δ (ppm): 7.76 (dd, *J* = 8.6, 5.9 Hz, 1H, ArH), 7.30 (d, *J* = 3.6 Hz, 1H, ArH), 7.14 – 6.95 (m, 2H, ArH), 6.74 (d, *J* = 3.6 Hz, 1H, ArH), 2.52 (s, 3H, CH$_3$).

5-(2-Amino-4-fluorophenyl)furan-2-carboxylic acid (**35**). Starting compound: methyl 5-(2-amino-4-fluorophenyl)furan-2-carboxylate (**108**). Procedure 1. Yellow solid. Yield: 99%. TLC (dichloromethane – methanol 8:2): R$_f$ = 0.23.Mp: 184.8°C. ^{1}H NMR (300 MHz, CD$_3$OD) δ (ppm): 7.52 (dd, *J* = 8.7, 6.4 Hz, 1H, ArH), 7.29 (d, *J* = 3.7 Hz, 1H, ArH), 6.74 (d, *J* = 3.7 Hz, 1H, ArH), 6.56-6.39 (m, 2H, ArH).

5-(2-Cyano-4-fluorophenyl)furan-2-carboxylic acid (**36**). Starting compound: methyl 5-(2-cyano-4-fluorophenyl)furan-2-carboxylate (**109**). Procedure 2. White solid. Yield: 93%. TLC (dichloromethane – methanol 9:1): R$_f$ = 0.19. Mp: 213°C. ^{1}H NMR (300 MHz, CD$_3$OD) δ (ppm): 8.15 – 8.07 (m, 1H, ArH), 7.72 – 7.66 (m, 1H, ArH), 7.62 – 7.51 (m, 1H, ArH), 7.40 – 7.31 (m, 2H, ArH).

5-(4-Fluoro-2-(trifluoromethyl)phenyl)furan-2-carboxylic acid (**37**). Starting compound: methyl 5-(4-fluoro-2-(trifluoromethyl)phenyl)furan-2-carboxylate (**110**). Procedure 1. Light yellow solid. Yield: 70%. TLC (dichloromethane – methanol 8:2): R$_f$ = 0.32. Mp: 146.2°C. ^{1}H NMR (300 MHz, CD$_3$OD) δ (ppm): 7.89 (dd, *J* = 8.8, 5.4 Hz, 1H, ArH), 7.62 (dd, *J* = 9.2, 2.7 Hz, 1H, ArH), 7.51 (td, *J* = 8.3, 2.7 Hz, 1H, ArH), 7.29 (d, *J* = 3.6 Hz, 1H, ArH), 6.81 (d, *J* = 3.6 Hz, 1H, ArH).

5-(2-Hydroxy-4-nitrophenyl)furan-2-carboxylic acid (**40**). Starting compound: methyl 5-(2-hydroxy-4-nitrophenyl)furan-2-carboxylate (**111**). Procedure 1. Yellow solid. Yield: 82%. TLC (dichloromethane – methanol 8:2): R$_f$ = 0.22. Mp: 218°C. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.3-13.2 (bs exch. D$_2$O, 1H, COOH), 11.5-11.4 (bs exch. D$_2$O, 1H, OH), 7.95 (dd, 1H, *J* = 7.7, 1.6 Hz, ArH), 7.8 (m, 2H, H11, ArH), 7.3 (d, 1H, *J* = 3.6 Hz, ArH), 7.2 (d, 1H, *J* = 3.6 Hz, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.61, 154.99, 151.79, 147.61, 144.79, 126.86, 122.67, 120.14, 114.95, 114.77, 111.09. FTIR (KBr) ν cm^{-1}: 3207, 3073, 2844, 1701, 1603, 1574, 1521, 1490, 1428, 1331, 1263, 1149, 1028, 955, 886, 726. HRMS (ESI- QOrbitrap) *m/z*: C$_{11}$H$_6$FNO$_5$ calcd 248.0201, found 248.0201 [M-H]$^-$.

5-(2-Methyl-4-nitrophenyl)furan-2-carboxylic acid (**41**). Starting compound: methyl 5-(2-methyl-4-nitrophenyl)furan-2-carboxylate (**112**). Procedure 1. Light yellow solid. Yield: 74%. TLC (dichloromethane – methanol 7:3): R_f = 0.41. Mp: 216.4°C. ^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm): 13.3-13.2 (bs exch. D_2O, 1H, COOH), 8.19 (d, 1H, J = 2.4 Hz, ArH), 8.15 (dd, 1H, J = 8.6 , 2.4 Hz, ArH), 7.9 (d, 1H, J = 8.6 Hz, ArH), 7.3 (d, 1H, J = 3.8 Hz, ArH), 7.2 (d, 1H, J = 3.8 Hz, ArH), 2.60 (s, 3H, CH$_3$). ^{13}C NMR (75 MHz, DMSO-d_6) δ (ppm): 159.55, 153.91, 147.00, 145.67, 137.19, 134.70, 128.51, 126.58, 121.79, 119.79, 114.37, 21.87. FTIR (KBr) ν cm^{-1}: 3468, 3420, 3082, 2917, 2849, 1696, 1673, 1612, 1582, 1515, 1474, 1424, 1346, 1317, 1303, 1255, 1173, 1035, 902, 803. HRMS (ESI- QOrbitrap) *m/z*: $C_{12}H_9NO_5$ calcd 246.0408, found 246.0408 [M-H]$^-$.

5-(2-Amino-4-nitrophenyl)furan-2-carboxylic acid (**42**). Starting compound: methyl 5-(2-amino-4-nitrophenyl)furan-2-carboxylate (**113**). Procedure 1. Orange solid. Yield: 40%. TLC (dichloromethane – methanol 7:3): R_f = 0.52. Mp: 228.5°C. ^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm): 13.4-13.2 (bs exch. D_2O, 1H, COOH), 7.78-7.73 (m, 2H, H8, ArH), 7.46 (dd, 1H, J = 8.5, 2.1 Hz, ArH), 7.39 (d, 1H, J = 3.6 Hz, ArH), 7.18 (d, 1H, J = 3.6 Hz, ArH), 6.0 (bs exch. D_2O, 2H, NH$_2$). ^{13}C NMR (75 MHz, DMSO-d_6) δ (ppm): 159.99, 154.52, 148.65, 146.78, 144.95, 129.17, 120.27, 118.95, 111.94, 111.58, 111.50. FTIR (KBr) ν cm^{-1}: 3512, 3404, 3120, 2956, 1682, 1633, 1585, 1576, 1508, 1480, 1344, 1314, 1256, 1172, 1040, 1032, 816, 800, 745. HRMS (ESI- QOrbitrap) *m/z*: $C_{11}H_8N_2O_5$ calcd 247.0360, found 247.0361 [M-H]$^-$.

5-(4-Nitro-2-(trifluoromethyl)phenyl)furan-2-carboxylic acid (**44**). Starting compound: methyl 5-(4-nitro-2-(trifluoromethyl)phenyl)furan-2-carboxylate (**114**). Procedure 1. Grey solid. Yield: 87%. TLC (dichloromethane – methanol 7:3): R_f = 0.36. Mp: 167.5°C. ^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm): 14.4-14.0 (bs exch. D_2O, 1H, COOH), 8.57 (dd, 1H, J = 8.6, 2.3 Hz, ArH), 8.55 (d, 1H, J = 2.3 Hz, ArH), 8.18 (d, 1H, J = 8.6 Hz, ArH), 7.40 (d, 1H, J = 3.7 Hz, ArH), 7.23 (d, 1H, J = 3.7 Hz, ArH). ^{13}C NMR (75 MHz, DMSO-d_6) δ (ppm): 159.77, 151.67, 147.89, 147.47, 133.87, 132.99, 128.91-125.28-121.65-118.02 (q), 128.52, 127.58-127.14-126.71-126.27, (q), 123.51-123.43-123.34, 123.26 (q), 120.21, 120.36, 115.72-115.69 (d). FTIR (KBr) ν cm^{-1}: 3435, 3122, 2962, 2917, 2850, 1689, 1639, 1615,

1595, 1571, 1479, 1424, 1352, 1316, 1129, 820. HRMS (ESI- QOrbitrap) *m/z*: $C_{12}H_6F_3NO_5$, calcd 300.0125, found 300.0128 [M-H]⁻.

5-(m-Tolyl)furan-2-carboxylic acid (**48**). Starting compound: methyl 5-(*m*-tolyl)furan-2-carboxylate (**115**). Procedure 1. Brown solid. Yield: 86%. Mp: 147°C. TLC (dichloromethane – methanol 8:2): R_f = 0.35. ^{1}H NMR (300 MHz, DMSO-d6) δ (ppm): 13.07 (bs exch. D_2O, 1H, COOH), 7.62-7.56 (m, 2H, ArH), 7.34 (t, *J* = 7.7 Hz, 1H, ArH), 7.29 (d, *J* = 3.7 Hz, 1H, ArH), 7.19 (d, *J* = 7.7 Hz, 1H, ArH), 7.09 (d, *J* = 3.7 Hz, 1H, ArH), 2.35 (s, 3H, CH_3). ^{13}C NMR (75 MHz, DMSO-d6) δ (ppm): 159.71, 156.85, 144.50, 138.82, 130.08, 129.59, 129.38, 125.23, 122.09, 120.29, 108.24, 21.35. FTIR (KBr) ν cm⁻¹: 3435, 2918, 2666, 2611, 2573, 1674, 1609, 1593, 1573, 1518, 1472, 1421, 1366, 1312, 1275, 1218, 1164, 1024, 788, 760. HRMS (ESI-QOrbitrap) *m/z*: $C_{12}H_{10}O_3$ calcd 203.0703, found 203.0702 [M+H]⁺.

5-(3-(Methylcarbamoyl)phenyl)furan-2-carboxylic acid (**51**). Starting compound: methyl 5-(3-(methylcarbamoyl)phenyl)furan-2-carboxylate (**116**). Procedure 2. Yellow solid. Yield: 87%. TLC (dichloromethane – methanol 7:3): R_f = 0.32. Mp: 159.0°C. ^{1}H NMR (300 MHz, DMSO-d6) δ (ppm): 8.58 (q exch. D_2O, *J* = 4.5 Hz, 1H, NH), 8.22 (t, *J* = 1.7 Hz, 1H, ArH), 7.92 (dt, *J* = 7.8, 1.7 Hz, 1H, ArH), 7.82 (dt, *J* = 7.8, 1.7 Hz, 1H, ArH), 7.55 (t, *J* = 7.8 Hz, 1H, ArH), 7.33 (d, *J* = 3.6 Hz, 1H, ArH), 7.18 (d, *J* = 3.6 Hz, 1H, ArH), 2.80 (d, *J* = 4.5 Hz, 3H, CH_3). ^{13}C NMR (75 MHz, DMSO-d6) δ (ppm): 166.54, 159.67, 156.10,144.87, 135.85, 129.70, 129.62, 127.83, 127.28, 123.42, 120.31, 108.95, 26.73. FTIR (KBr) ν cm⁻¹: 3324, 3116, 3066, 2925, 1718, 1692, 1649, 1584, 1549, 1521, 1481, 1468, 1423, 1407, 1310, 1260, 1220, 1158, 1026, 803, 760. HRMS (ESI-QOrbitrap) *m/z*: $C_{13}H_{11}NO_4$ calcd 246.0761, found 246.0761 [M+H]⁺.

5-(3-Sulfamoylphenyl)furan-2-carboxylic acid (**52**). Starting compound: methyl 5-(3-sulfamoylphenyl)furan-2-carboxylate (**117**). Procedure 1. Yellow solid. Yield: 84%. TLC (dichloromethane – methanol 7:3): R_f = 0.24. Mp: 261.0°C (dec.). ^{1}H NMR (300 MHz, CD3OD) δ (ppm): 8.34 (s, *J* = 1.7 Hz, 1H, ArH), 8.03 (d, *J* = 7.9 Hz, 1H, ArH), 7.88 (d, *J* = 7.9 Hz, 1H, ArH), 7.63 (t, *J* = 7.8 Hz, 1H, ArH), 7.31 (d, *J* = 3.6 Hz, 1H, ArH), 7.07 (d, *J* = 3.6 Hz, 1H, ArH). HRMS (ESI- QOrbitrap) *m/z*: $C_{12}H_8O_5$ calcd 266.0129, found 266.0129 [M-H]⁻.

5-(3-Carboxyphenyl)furan-2-carboxylic acid (**53**). Starting compound: methyl 5-(3-(methoxycarbonyl)phenyl)furan-2-carboxylate (**118**). Procedure 1. White solid. Yield: 88%. TLC (dichloromethane – methanol 8:2): R_f = 0.26. Mp > 300°C. ^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm): 13.20 (2H, bs exch. D_2O, 2H, COOH), 8.30 (s, 1H, ArH), 8.04 (d, J = 7.8, 1H, ArH), 7.92 (d, J = 7.8, 1H, ArH), 7.60 (t, J = 7.8 Hz, 1H, ArH), 7.32 (d, J = 3.6 Hz, 1H, ArH), 7.25 (d, J = 3.6 Hz, 1H, ArH). ^{13}C NMR (75 MHz, DMSO-d_6) δ (ppm): 167.2, 159.7, 155.6, 145.0, 132.2, 129.9, 129.8, 129.1, 125.2, 120.2, 119.2. FTIR (KBr) ν cm^{-1}: 3435, 2966, 2917, 2851, 2661, 2546, 1681, 1612, 1521, 1455, 1420, 1298, 1165, 1030, 804 761. HRMS (ESI-QOrbitrap) *m/z*: $C_{12}H_8O_5$ calcd 231.0299, found 231.0301 [M-H]$^-$.

5-(3-Cyano-5-methoxyphenyl)furan-2-carboxylic acid (**55**). Starting compound: methyl 5-(3-cyano-5-methoxyphenyl)furan-2-carboxylate (**119**). Procedure 2. White solid. Yield: 40%. TLC (dichloromethane – methanol 8:2): R_f = 0.11. Mp: 225.8°C (dec.). ^{1}H NMR (300 MHz, CD_3OD) δ (ppm): 7.77 (t, J = 1.5 Hz, 1H, ArH), 7.68 (dd, 1H, J = 2.4, 1.5 Hz, ArH), 7.33 – 7.23 (m, 2H, ArH), 7.10 (d, J = 3.7 Hz, 1H, ArH), 3.91 (s, 3H, CH_3).

5-(3-Cyano-5-phenoxyphenyl)furan-2-carboxylic acid (**56**). Starting compound: methyl 5-(3-cyano-5-phenoxyphenyl)furan-2-carboxylate (**120**). Procedure 2. White solid. Yield: 73%. TLC (dichloromethane – methanol 8:2): R_f = 0.41. Mp: 227.9°C (dec.). ^{1}H NMR (300 MHz, CD_3OD) δ (ppm): 7.92 (t, J = 1.5 Hz, 1H, ArH), 7.72 (dd, 1H, J = 2.3, 1.5 Hz, ArH), 7.51 – 7.39 (m, 2H, ArH), 7.32 – 7.18 (m, 3H, ArH), 7.15 – 7.05 (m, 3H, ArH).

5.2.1.3 Synthesis of 22,32,38,39,43 (General Scheme C)
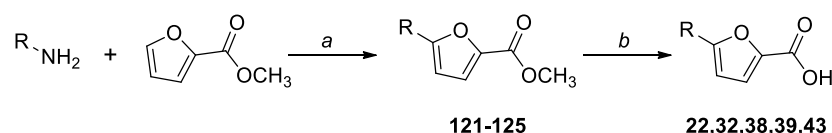
121-125   **22,32,38,39,43**

Reagents and conditions: *a) 1.* NaNO$_2$, conc. HCl, 30 min, 0°C; *2.* CuCl$_2$, CH$_3$COCH$_3$/H$_2$O, 2-3 h; *b)* NaOH, EtOH/THF 1:1, 5 h, reflux.

General procedure for the Sandmeyer reaction (*a*). The appropriate disubstituted aniline (3.2 mmol) was dissolved in 6M HCl (5 mL), and the solution

was cooled to 0°C in an ice bath. NaNO$_2$ (1.12 mL), previously dissolved in H$_2$O, was then added; the temperature was kept between 0-5°C during the addition. To the so-obtained diazonium salt, a solution of methyl furan-2-carboxylate (3.0 mmol) and CuCl$_2$ (0.214 mmol) in acetone (1.88 mL) was added dropwise under stirring; during the addition, the temperature was maintained in the range 20-30°C, so that the rate of evolution of nitrogen was 2-3 bubbles per minute. The reaction was left stirring for ~3 hours. The formed precipitate was filtered off, and then crystallised, to obtain the desired products.[38]

Methyl 5-(2-bromo-4-(trifluoromethyl)phenyl)furan-2-carboxylate (**121**). Starting compounds: methyl furan-2-carboxylate and 2-bromo-4-(trifluoromethyl)aniline. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 95:5) to give the desired product as an orange solid. Yield: 39%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.65. Mp: 62°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.07 (d, 1H, J = 8.3 Hz, ArH), 7.94 (s, 1H, ArH), 7.65 (d, 1H, J = 8.4 Hz, ArH), 7.40 (d, 1H, J = 3.7 Hz, ArH), 7.30 (d, 1H, J = 3.7 Hz, ArH), 3.94 (s, 3H, CH$_3$).

Methyl 5-(2-bromo-4-fluorophenyl)furan-2-carboxylate (**122**). Starting compounds: methyl furan-2-carboxylate and 2-bromo-4-fluoroaniline. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 95:5) to give the desired product as an orange solid. Yield: 10%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.51. Mp: 80.8°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.89 (dd, J = 8.8, 6.1 Hz, 1H, ArH), 7.42 (dd, J = 8.2, 2.5 Hz, 1H, ArH), 7.27 (d, hidden by solvent peak, 1H, ArH), 7.17 (d, J = 3.7 Hz, 1H, ArH), 7.16 – 7.10 (m, 2H, ArH), 3.92 (s, 3H, CH$_3$).

Methyl 5-(2-fluoro-4-nitrophenyl)furan-2-carboxylate (**123**). Starting compounds: methyl furan-2-carboxylate and 2-fluoro-4-nitroaniline. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 9:1) to give the desired product as an orange solid. Yield: 14%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.39. Mp: 196.3°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.20-8.16 (m, 1H, ArH), 8.15 - 8.13 (m, 1H, ArH), 8.11-8.07 (m, 1H, ArH), 7.31 (d, 1H, J = 3.6 Hz, ArH), 7.15 (d, 1H, J = 3.6 Hz, ArH), 3.95 (s, 3H, CH$_3$).

Methyl 5-(2-bromo-4-nitrophenyl)furan-2-carboxylate (**124**). Starting compounds: methyl furan-2-carboxylate and 2-bromo-4-nitroaniline (**145**). The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 95:5) to give the desired product as an orange solid. Yield: 39%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.65. Mp: 62°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.07 (d, 1H, J = 8.3 Hz, ArH), 7.94 (s, 1H, ArH), 7.65 (d, 1H, J = 8.4 Hz, ArH), 7.40 (d, 1H, J = 3.7 Hz, ArH), 7.30 (d, 1H, J = 3.7 Hz, ArH), 3.94 (s, 3H, CH$_3$).

Methyl 5-(2-cyano-4-nitrophenyl)furan-2-carboxylate (**125**). Starting compounds: methyl furan-2-carboxylate and 2-cyano-4-nitroaniline. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a dark yellow solid. Yield: 39%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.46. Mp: 207°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.62 (d, 1H, J = 2.3 Hz, ArH), 8.52 - 8.47 (m, 1H, ArH), 8.33 - 8.28 (m, 1H, ArH), 7.69 (d, 1H, J = 3.7 Hz, ArH), 7.35 (d, 1H, J = 3.7 Hz, ArH), 3.96 (s, 3H, CH$_3$).

General procedure for the base-catalysed hydrolysis (*b*). See the corresponding reaction in General Scheme A (*c*).

5-(2-Bromo-4-(trifluoromethyl)phenyl)furan-2-carboxylic acid (**22**). Starting compound: methyl 5-(2-bromo-4-(trifluoromethyl)phenyl)furan-2-carboxylate (**121**). Procedure 1. White solid. Yield: Quantitative. TLC (dichloromethane – methanol 9:1): R_f = 0.14. Mp: 180°C. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.6-13.2 (bs exch. D$_2$O, 1H, COOH), 8.13 (s, 1H, ArH) 8.02 (d, 1H, J = 8.2 Hz, ArH), 7.87 (d, 1H, J = 8.2 Hz, ArH), 7.42 (d, 1H, J = 3.1 Hz, ArH), 7.37 (d, 1H, J = 3.1 Hz, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 152.55, 133.70, 131.47-131.41-131.66-131.34 (q), 130.58, 131.14-130.66-130.24-129.80 (q), 128.86-125.24-121.63-118.01 (q), 125.53-125.49-125.44-125.40 (q), 120.27, 119.54, 114.55. FTIR (KBr) ν cm^{-1}: 3446, 2965, 2918, 2850, 1703, 1681, 1612, 1584, 1525, 1479, 1420, 1329, 1308, 1116, 804. HRMS (ESI- QOrbitrap) m/z: C$_{12}$H$_6$BrF$_3$O$_3$ calcd 332.9380, found 332.9384 [M-H]$^-$.

5-(2-Bromo-4-fluorophenyl)furan-2-carboxylic acid (**32**). Starting compound: methyl 5-(2-bromo-4-fluorophenyl)furan-2-carboxylate (**122**). Procedure 1. Light yellow solid. Yield: 65%. TLC (dichloromethane – methanol

8:2): R$_f$ = 0.29. Mp: 206.8°C. ^{1}H NMR (300 MHz, CD$_3$OD) δ (ppm): 7.92 (dd, J = 8.9, 6.0 Hz, 1H, ArH), 7.56 (dd, J = 8.5, 2.6 Hz, 1H, ArH), 7.39 – 7.09 (m, 3H, ArH).

5-(2-Fluoro-4-nitrophenyl)furan-2-carboxylic acid (**38**). Starting compound: methyl 5-(2-fluoro-4-nitrophenyl)furan-2-carboxylate (**123**). Procedure 1. Yellow solid. Yield: 92%. TLC (dichloromethane – methanol 8:2): R$_f$ = 0.26. Mp: 248.6°C. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.0-12.8 (bs exch. D$_2$O, 1H, COOH), 8.26 (dd, 1H, J = 11.03, 2.2 Hz, ArH), 8.19 (dd, 1H, J = 8.9, 2.2 Hz, ArH), 8.09 (d, 1H, J = 8.8 Hz, ArH), 7.38 (d, 1H, J = 3.8 Hz, ArH), 7.22 (t, 1 H, J = 3.8, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.63-156.25 (d), 159.39, 148.72-148.68 (d), 147.89-147.77 (d), 146.34, 127.63-127.79 (d), 123.64-123.48 (d), 120.94-120.89 (d), 120.08, 115.56-115.41 (d), 113.07-112.71 (d). FTIR (KBr) ν cm^{-1}: 3435, 3098, 1622, 1584, 1573, 1618, 1509, 1396, 1378, 1339, 1119, 1024, 813, 780. HRMS (ESI- QOrbitrap) *m/z*: C$_{11}$H$_6$FNO$_5$ calcd 250.0157, found 250.0158 [M-H]$^-$.

5-(2-Bromo-4-nitrophenyl)furan-2-carboxylic acid (**39**). Starting compound: methyl 5-(2-bromo-4-nitrophenyl)furan-2-carboxylate (**124**). Procedure 1. Yellow solid. Yield: 98%. TLC (dichloromethane – methanol 8:2): R$_f$ = 0.24. Mp: 255.0°C. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 8.47 (d, 1H, J = 2.2 Hz, ArH), 8.30 (dd, 1H, J = 8.9, 2.2 Hz, ArH), 8.07 (d, 1H, J = 8.9 Hz, ArH), 7.46 (d, 1H, J = 3.5 Hz, ArH), 6.78 (d, 1H, J = 3.5 Hz, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 161.75, 156.14, 147.67, 146.12, 136.60, 129.56, 129.08, 123.41, 118.16, 115.81, 113.48. FTIR (KBr) ν cm^{-1}: 3435, 3098, 1622, 1584, 1573, 1618, 1509, 1396, 1378, 1339, 1119, 1024, 813, 780. HRMS (ESI-QOrbitrap) *m/z*: C$_{11}$H$_6$BrNO$_5$ calcd 309.9357, found 309.9360 [M-H]$^-$.

5-(2-Cyano-4-nitro-phenyl)furan-2-carboxylic acid (**43**). Starting compound: methyl 5-(2-cyano-4-nitrophenyl)furan-2-carboxylate (**125**). Procedure 2. Dark yellow solid. Yield: 87%. TLC (dichloromethane – methanol 8:2): R$_f$ = 0.37. Mp: 246°C. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 13.56-13.46 (bs exch. D$_2$O, 1H, COOH), 8.81 (d, 1H, J = 2.4 Hz, ArH), 8.57 (dd, 1H, J = 8.9, 2.4 Hz, ArH), 8.20 (d, 1H, J = 8.9 Hz, ArH), 7.58 (d, 1H, J = 3.8 Hz, ArH), 7.44 (d, 1H, J = 3.8 Hz, ArH). ^{13}C NMR (75 MHz, DMSO-d$_6$) δ (ppm): 159.24, 150.75, 147.18, 146.96, 136.43, 130.53, 128.91, 128.52, 120.08, 117.02, 115.36, 108.21. FTIR (KBr) ν cm^{-1}: 3538, 3469,

3418, 3068, 2232,1682, 1697, 1639, 1607, 1579, 1530, 1513, 1420, 1367, 1292, 1275, 1173, 1041, 810, 801, 747. HRMS (ESI-QOrbitrap) *m/z*: $C_{12}H_6N_2O_5$ calcd 257.0204, found 257.0205 [M-H]$^-$.

5.2.1.4 Synthesis of 61,62 (General Scheme D)
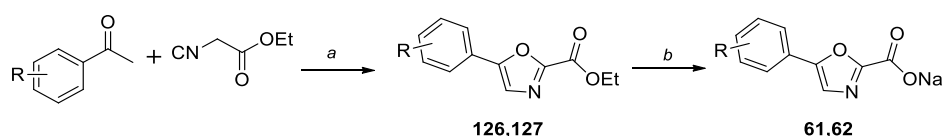
Reagents and conditions: *a*) I$_2$, DMSO, 3 h, 130°C; *b*) NaOH, THF/H$_2$O 1:1, 1.5 h, r.t.

General procedure for the coupling reaction (*a*). To a solution of the suitable acetophenone (1.0 mmol) and iodine (1.6 mmol) in DMSO (3 mL) was added ethyl 2-isocyanoacetate (2.0 mmol). The resulting mixture was stirred at 130°C for 3 hours; then, the reaction was quenched with water (50 mL) and extracted with ethyl acetate (3 × 50 mL). The combined organic layers were washed with brine, dried over anhydrous sodium sulphate and concentrated under reduced pressure.[39]

Ethyl 5-(4-nitrophenyl)oxazole-2-carboxylate (**126**). Starting compound: 1-(4-nitrophenyl)ethenone. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a yellow solid. Yield: 7%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.15. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.35 (d, *J* = 8.8 Hz, 2H, ArH), 7.95 (d, *J* = 8.8 Hz, 2H, ArH), 7.72 (s, 1H, ArH), 4.53 (q, *J* = 7.1 Hz, 2H, CH$_2$), 1.48 (t, *J* = 7.1 Hz, 3H, CH$_3$).

Ethyl 5-(3-cyanophenyl)oxazole-2-carboxylate (**127**). Starting compound: 3-acetylbenzonitrile. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a light brown solid. Yield: 6%. TLC (cyclohexane – ethyl acetate 6:4): R$_f$ = 0.46. Mp: 154.3°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.06 (s, 1H, ArH), 7.99 (d, *J* = 7.8 Hz, 1H, ArH), 7.70 (d, *J* = 7.8 Hz, 1H, ArH), 7.66 – 7.55 (m, 2H, ArH), 4.52 (q, *J* = 7.1 Hz, 2H, CH$_2$), 1.47 (t, *J* = 7.1 Hz, 3H, CH$_3$).

General procedure for the base-catalysed hydrolysis (*b*). Pellets of sodium hydroxide (0.11 mmol) were dissolved in water (2 mL), and the resulting solution was added to a solution of the suitable esters (0.11 mmol) in tetrahydrofuran (2 mL). The reaction mixture was stirred at room temperature for 1.5 h; then, the solvent was removed under reduced pressure, and the aqueous layer was freeze-dried to afford the final products.[39]

Sodium 5-(4-nitrophenyl)oxazole-2-carboxylate (**61**). Starting compound: ethyl 5-(4-nitrophenyl)oxazole-2-carboxylate (**126**). Pale yellow solid. Yield: 80%. TLC (dichloromethane – methanol 8:2): R_f = 0.10. Mp > 300°C (dec.). ^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm): 8.32 (d, *J* = 9.0 Hz, 2H, ArH), 7.97 (d, *J* = 9.0 Hz, 2H, ArH), 7.90 (s, 1H, ArH).

Sodium 5-(3-cyanophenyl)oxazole-2-carboxylate (**62**). Starting compound: ethyl 5-(3-cyanophenyl)oxazole-2-carboxylate (**127**). Grey solid. Yield: 89%. TLC (dichloromethane – methanol 8:2): R_f = 0.10. Mp > 300°C (dec.). ^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm): 8.18 (td, *J* = 1.7, 0.6 Hz, 1H, ArH), 8.00 (dt, *J* = 7.8, 1.7, 1H, ArH), 7.80 (dt, *J* = 7.8, 1.7 Hz, 1H, ArH), 7.74 (s, 1H, ArH), 7.67 (td, *J* = 7.8, 0.6 Hz, 1H, ArH).

5.2.1.5 Synthesis of 63,64 (General Scheme E)
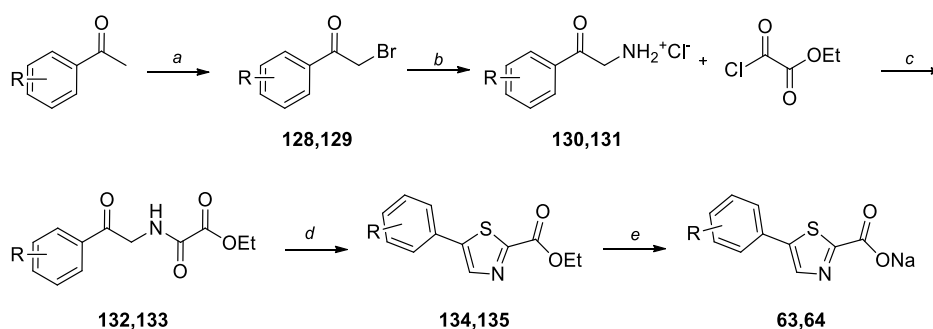
Reagents and conditions: *a*) NBS, *p*-TsOH, CH$_2$Cl$_2$, overnight, r.t, N$_2$ atm.; *b*) *1*. hexamine, CH$_2$Cl$_2$, 8 h, r.t.; *2*. conc. HCl, EtOH, overnight, r.t.; *c*) TEA, EtOAc, 3 h, reflux; *d*) Lawesson's reagent, 1,4-dioxane, 2 h, reflux; *v*) NaOH, THF/H$_2$O 1:1, 1.5 h, r.t.

General procedure for the bromination (*a*). A solution of the suitable acetophenone (3.03 mmol) in dichloromethane (0.61 mL) was added dropwise to a solution of *N*-bromosuccinimide (3.64 mmol) and *p*-toluenesulfonic acid (0.3

mmol,) in dichloromethane (2.42 mL) at 0°C. The reaction mixture was then stirred under nitrogen atmosphere at room temperature for 16 hours. After the addition of water (3 mL), the organic layer was separated, and the aqueous layer was extracted with dichloromethane (2 × 15 mL). The combined organic layers were washed with saturated aqueous sodium bicarbonate (10 mL) and brine (10 mL), dried over anhydrous sodium sulphate, and concentrated under reduced pressure.

2-Bromo-1-(4-nitrophenyl)ethenone (**128**). Starting compound: 1-(4-nitrophenyl)ethenone. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a white solid. Yield: 45%. Mp: 105°C. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.38. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.36 (dd, J = 9.0, 2.1 Hz, 2H, ArH), 8.17 (dd, J = 9.0, 2.1 Hz, 2H, ArH), 4.47 (s, 2H, CH$_2$).

3-(2-Bromoacetyl)benzonitrile (**129**). Starting compound: 3-acetylbenzonitrile. The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to give the desired product as a white solid. Yield: 20%. Mp: 105.4°C. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.30. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.27 (s, 1H, ArH), 8.22 (d, J = 7.7 Hz, 1H, ArH), 7.89 (d, J = 7.7 Hz, 1H, ArH), 7.66 (t, J = 7.7 Hz, 1H, ArH), 4.43 (s, 2H, CH$_2$).

General procedure for the Delépine reaction (*b*). Hexamine (0.6 mmol) was added to a stirred solution of the monobrominated compound (0.5 mmol) in dichloromethane (1.8 mL), and the reaction mixture was stirred at room temperature for 8 hours. After completion, the reaction was filtered, washed with dichloromethane, and dried under reduced pressure to afford an off-white solid that was dissolved in ethanol (1.8 mL) and concentrated hydrochloric acid (0.1 mL). The reaction mixture was stirred at room temperature for 16 hours. The obtained solid was filtered, washed with cold ethanol, and dried to give an off-white solid.[40]

2-(4-Nitrophenyl)-2-oxoethanaminium chloride (**130**). Starting compound: 2-bromo-1-(4-nitrophenyl)ethenone (**128**). White solid. Yield: 54%. Mp: 228.6 (dec.). TLC (dichloromethane – methanol 9:1): R_f = 0.17. ^{1}H-NMR (300 MHz,

DMSO-d$_6$) δ (ppm): 8.48 (bs exch. D$_2$O, 3H, NH$_3^+$), 8.38 (d, J = 8.8 Hz, 2H, ArH), 8.26 (d, J = 8.8 Hz, 2H, ArH), 4.69 (s, 2H, CH$_2$).

2-(3-Cyanophenyl)-2-oxoethanaminium chloride (**131**). Starting compound: 3-(2-bromoacetyl)benzonitrile (**129**). White solid. Yield: 65%. TLC (dichloromethane – methanol 9:1): R$_f$ = 0.24. ^{1}H-NMR (300 MHz, DMSO-d$_6$) δ (ppm): 8.52 (s, 1H, ArH), 8.42 (bs exch. D$_2$O, 3H, NH$_3^+$), 8.30 (d, J = 7.8 Hz, 1H, ArH), 8.21 (d, J = 7.8 Hz, 1H, ArH), 7.82 (t, J = 7.8 Hz, 1H, ArH), 4.65 (s, 2H, CH$_2$).

General procedure for the amide formation (*c*). To a solution of the suitable hydrochloride salt (1.34 mmol) in ethyl acetate (5.83 mL), triethylamine (0.22 mL, 1.61 mmol) was added, and reaction mixture was cooled to 5°C. Then, ethyl oxalylchloride (0.18 mL, 1.61 mmol,) was added dropwise, and the resulting mixture was refluxed for 3 hours. After completion, the reaction was cooled and quenched with water; the organic layer was separated, dried over sodium sulphate and evaporated under reduced pressure to afford the crude product.[40]

Ethyl 2-((2-(4-nitrophenyl)-2-oxoethyl)amino)-2-oxoacetate (**132**). Starting compound: 2-(4-nitrophenyl)-2-oxoethanaminium chloride (**130**). The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 5:5) to give the desired product as a yellow solid. Yield: 42%. Mp: 152°C. TLC (cyclohexane – ethyl acetate 6:4): R$_f$ = 0.15. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.38 (d, J = 9.0 Hz, 2H, ArH), 8.17 (d, J = 9.0 Hz, 2H, ArH), 8.01 (bs exch. D$_2$O, 1H, NH), 4.89 (d, J = 4.9 Hz, 2H, CH$_2$), 4.40 (q, J = 7.1 Hz, 2H, OCH$_2$), 1.41 (t, J = 7.1 Hz, 3H, CH$_3$).

Ethyl 2-((2-(3-cyanophenyl)-2-oxoethyl)amino)-2-oxoacetate (**133**). Starting compound: 2-(3-cyanophenyl)-2-oxoethanaminium chloride (**131**). The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 5:5) to give the desired product as a pale yellow solid. Yield: 61%. TLC (cyclohexane – ethyl acetate 6:4): R$_f$ = 0.12. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.29 (td, J = 1.7, 0.6 Hz, 1H, ArH), 8.21 (dt, J = 7.8, 1.7, 1H, ArH), 8.00 (s, 1H, ArH), 7.93 (dt, J = 7.8, 1.7, 1H, ArH), 7.69 (td, J = 7.8, 0.6 Hz, 1H, ArH), 4.85 (d, J = 4.9 Hz, 2H, CH$_2$), 4.41 (q, J = 7.1 Hz, 2H, OCH$_2$), 1.42 (t, J = 7.1 Hz, 3H, CH$_3$).

General procedure for the ring closure (*d*). To a solution of the proper amide (0.49 mmol) in 1,4-dioxane (2.8 mL), the Lawesson's reagent (0.54 mmol)

was added, and the mixture was refluxed for 2 hours. The reaction was subsequently cooled, added to water, and neutralised with a saturated solution of sodium carbonate. The product was separated from the aqueous phase using ethyl acetate. The organic layer was dried over sodium sulphate, filtered, and concentrated under reduced pressure to give a dark brown residue.[40]

Ethyl 5-(4-nitrophenyl)thiazole-2-carboxylate (**134**). Starting compound: ethyl 2-((2-(4-nitrophenyl)-2-oxoethyl)amino)-2-oxoacetate (**132**). The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 7:3) to give the desired product as a pale yellow solid. Yield: 88%. Mp: 173.3°C. TLC (cyclohexane – ethyl acetate 6:4): R_f = 0.45. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.32 (dd, J = 9.1, 2.2 Hz, 1H, ArH), 8.28 (s, 1H, ArH), 7.80 (dd, J = 9.1, 2.2 Hz, 2H, ArH), 4.51 (q, J = 7.1 Hz, 2H, CH$_2$), 1.47 (t, J = 7.1 Hz, 3H, CH$_3$).

Ethyl 5-(3-cyanophenyl)thiazole-2-carboxylate (**135**). Starting compound: ethyl 2-((2-(3-cyanophenyl)-2-oxoethyl)amino)-2-oxoacetate (**133**). The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 7:3) to give the desired product as a white solid. Yield: 60%. TLC (cyclohexane – ethyl acetate 6:4): R_f = 0.43. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.19 (s, 1H, ArH), 7.89 (td, J = 1.7, 0.6 Hz, 1H, ArH), 7.84 (dt, J = 7.8, 1.7, 1H, ArH), 7.69 (dt, J = 7.8, 1.7 Hz, 1H, ArH), 7.58 (td, J = 7.8, 0.6 Hz, 1H, ArH), 4.50 (q, J = 7.1 Hz, 2H, CH$_2$), 1.46 (t, J = 7.1 Hz, 3H, CH$_3$).

General procedure for the base-catalysed hydrolysis (*e*). See the corresponding reaction in General Scheme D (*b*).

Sodium 5-(4-nitrophenyl)thiazole-2-carboxylate (**63**). Starting compound: ethyl 5-(4-nitrophenyl)thiazole-2-carboxylate (**134**). Dark green solid. Yield: 85%. TLC (dichloromethane – methanol 8:2): R_f = 0.10. Mp > 300°C (dec.). ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 8.31 (d, J = 8.9 Hz, 2H, ArH), 8.28 (s, 1H, ArH), 7.93 (d, J = 8.9 Hz, 2H, ArH).

Sodium 5-(3-cyanophenyl)thiazole-2-carboxylate (**64**). Starting compound: ethyl 5-(3-cyanophenyl)thiazole-2-carboxylate (**135**). White solid. Yield: 86%. TLC (dichloromethane – methanol 8:2): R_f = 0.10. Mp > 300°C (dec.). ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 8.19 (s, 1H, ArH), 8.09 (t, J = 1.6 Hz, 1H, ArH), 7.98 (dt, J

= 7.8, 1.6 Hz, 1H, ArH), 7.72 (dt, *J* = 7.8, 1.6 Hz, 1H, ArH), 7.62 (t, *J* = 7.8 Hz, 1H, ArH).

5.2.1.6 Synthesis of 65,66 (General Scheme F)
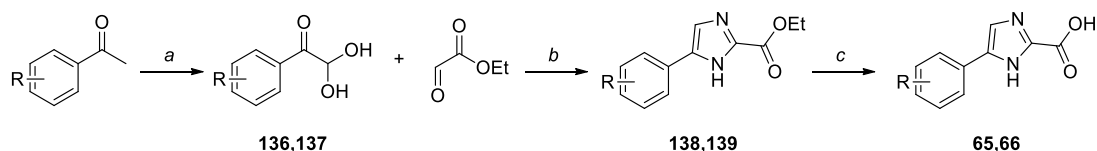
Reagents and conditions: *a*) SeO$_2$, 1,4-dioxane, H$_2$O, 7 h, reflux, N$_2$; *b*) NH$_4$OAc, CH$_3$CN, H$_2$O, 4 h, r.t.; *c*) NaOH, H$_2$O/EtOH 1:1, 5 h, reflux.

General procedure for the diol formation (*a*). The suitable acetophenone (1 mmol) was added to a solution of selenium dioxide (2 mmol) in 1,4-dioxane (1 mL) and water (2 drops); the resulting mixture was stirred and refluxed for 7 hours. After completion, the reaction was cooled to room temperature and filtered on a celite pad, after the addition of 10 mL of dichloromethane. The combined organic layers were concentrated under reduced pressure to afford a liquid, which was heated at reflux with 2 mL of water for 10 minutes. The resulting solution was cooled in an ice bath, until the formation of a precipitate that was retrieved by filtration and used without further purification.[41]

2,2-Dihydroxy-1-(4-nitrophenyl)ethenone (**136**). Starting compound: 1-(4-nitrophenyl)ethenone. Yellow solid. Yield: 36%. ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 8.33 (d, *J* = 9.2 Hz, 2H, ArH), 8.27 (d, *J* = 9.2 Hz, 2H, ArH), 7.02 (bs exch. D$_2$O, 2H, OH), 5.65 (s, 1H, CH).

3-(2,2-Dihydroxyacetyl)benzonitrile (**137**). Starting compound: 3-acetylbenzonitrile. Orange oil. Yield: 85%. TLC (cyclohexane – ethyl acetate 5:5): R$_f$ = 0.50. Not isolated.

General procedure for the coupling (*b*). Ethyl glyoxylate polymer in toluene (3.0 mmol) was depolymerised by heating at 60°C for 15 minutes. The resulting liquid was added dropwise to a solution of ammonium acetate (3.0 mmol) in water (0.7 mL) and acetonitrile (1.5 mL) at 0°C, followed by the addition of the suitable diol (1.0 mmol). The mixture was stirred at 0°C for 30 minutes, and

then at room temperature for 1.5 hours. After the removal of acetonitrile, the mixture was partitioned between ethyl acetate and water; the organic layer was washed with brine and dried over anhydrous sodium sulphate, filtered and concentrated *in vacuo*.[42]

Ethyl 5-(4-nitrophenyl)-1H-imidazole-2-carboxylate (**138**). Starting compound: 2,2-dihydroxy-1-(4-nitrophenyl)ethenone (**136**). The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 6:4) and recrystallisation (ethyl acetate – hexane) to give the desired product as a yellow solid. Yield: 31%. TLC (cyclohexane – ethyl acetate 6:4): R_f = 0.39. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 8.27 (d, J = 8.7 Hz, 2H, ArH), 8.00 (d, J = 8.7 Hz, 2H, ArH), 7.63 (s, 1H, ArH), 4.52 (q, J = 7.1 Hz, 2H, CH$_2$), 1.47 (t, J = 7.1 Hz, 3H, CH$_3$).

Ethyl 5-(3-cyanophenyl)-1H-imidazole-2-carboxylate (**139**). Starting compound: 3-(2,2-dihydroxyacetyl)benzonitrile (**137**). The crude was purified by flash column chromatography (cyclohexane – ethyl acetate 7:3) and recrystallisation (ethyl acetate – hexane) to give the desired product as a white solid. Yield: 40%. TLC (cyclohexane – ethyl acetate 5:5): R_f = 0.52. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 10.72 (bs exch. D$_2$O, 1H, NH), 8.14 (s, 1H, ArH), 8.07 (dd, J = 7.9, 1.6 Hz, 1H, ArH), 7.69 – 7.42 (m, 3H, ArH), 4.49 (q, J = 7.2 Hz, 2H, CH$_2$), 1.44 (t, J = 7.2 Hz, 3H, CH$_3$).

General procedure for the base-catalysed hydrolysis (*c*). *1.* The thiazole ester (1.0 mmol) was suspended in THF (5 mL), and an aqueous solution of NaOH (5.0 mmol in 5 mL) was added. The solution was stirred at reflux for 6 hours, and then at room temperature overnight. The precipitate was filtered off and treated with 3M HCl until pH 4; the resulting precipitate was filtered to afford the free acid. *2.* The thiazole ester (1.0 mmol) was suspended in a 2:1 THF-water mixture (10 mL), and lithium hydroxide (5.0 mmol) was added on an ice bath. The reaction was stirred at room temperature overnight. After completion, the solvent was evaporated, and the aqueous phase was first washed with dichloromethane (3 mL) and then acidified with 3M HCl to pH 4. The resulting precipitate was collected by filtration.[42]

5-(4-Nitrophenyl)-1H-imidazole-2-carboxylic acid (**65**). Starting compound: ethyl 5-(4-nitrophenyl)-1*H*-imidazole-2-carboxylate (**138**). Procedure 1. Yellow

solid. Yield: 70%. Mp: 137°C. TLC (dichloromethane – methanol 9:1): R_f = 0.11. ^{1}H-NMR (300 MHz, CD$_3$OD) δ (ppm): 8.28 (d, J = 8.8 Hz, 2H, ArH), 8.06 (d, J = 8.8 Hz, 2H, ArH), 7.89 (s, 1H, ArH).

5-(3-Cyanophenyl)-1H-imidazole-2-carboxylic acid (**66**). Starting compound: ethyl 5-(3-cyanophenyl)-1*H*-imidazole-2-carboxylate (**139**). Procedure 2. White solid. Yield: 54%. Mp: 165°C. TLC (dichloromethane – methanol 9:1): R_f = 0.10. ^{1}H-NMR (300 MHz, CD$_3$OD) δ (ppm): 8.22 (s, 1H, ArH), 8.13 (d, J = 7.8 Hz, 1H, ArH), 8.06 (s, 1H, ArH), 7.80 (d, J = 7.8 Hz, 1H, ArH), 7.69 (t, J = 7.8 Hz, 1H, ArH).

5.2.1.7 Procedure for the synthesis of 5-(3-aminophenyl)furan-2-carboxylic acid (49)
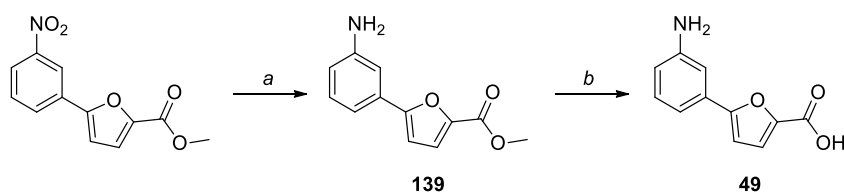
Reagents and conditions: *a*) SnCl$_2$, EtOAc, 5 h, reflux; *b*) NaOH, EtOH/THF 1:1, 5 h, reflux.

Methyl 5-(3-aminophenyl)furan-2-carboxylate (139). To a solution of methyl 5-(3-nitrophenyl)furan-2-carboxylate (1 mmol) in ethyl acetate (4 mL), tin(II) chloride (0.310 mmol) was added and the mixture was refluxed for 5 hours. After quenching by addition of a saturated aqueous sodium bicarbonate until pH= 7-8, the precipitated tin salts were eliminated by filtration, and the aqueous phase was extracted with ethyl acetate. The organic layer was dried over sodium sulphate, filtered and evaporated under vacuum. The crude residue was purified by flash chromatography (cyclohexane – ethyl acetate 8:2) to provide the desired compound as a white solid. Yield: 76%. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.45. Mp: 215°C (dec.). ^{1}H NMR (300 MHz, DMSO-d$_6$) δ (ppm): 7.36 (d, J = 3.7 Hz, 1H, ArH), 7.09 (t, J = 7.8 Hz, 1H, ArH), 7.00 (t, J = 2.0 Hz, 1H, ArH), 6.97 (d, J = 3.7 Hz, 1H, ArH), 6.94 (d, J = 7.8 Hz, 1H, ArH), 6.57 (dd, J = 7.8, 2.0 Hz, 1H, ArH), 5.31 (bs, 2H, NH$_2$), 3.81 (s, 3H, CH$_3$).[50]

5-(3-Aminophenyl)furan-2-carboxylic acid (49). See the corresponding reaction in General Scheme A (*c*). Procedure 1. Starting compound: methyl 5-(3-aminophenyl)furan-2-carboxylate (**139**). Yellow solid. Yield: 89%. Mp: 225°C (dec.). TLC (dichloromethane – methanol 7:3): R_f = 0.33. ^{1}H-NMR (300 MHz, DMSO-d6) δ (ppm): 7.25 (d, *J* = 3.7 Hz, 1H, ArH), 7.10 (t, *J* = 7.9 Hz, 1H, ArH), 6.99 (t, *J* = 2.0 Hz, 1H, ArH), 6.92 (m, 1H, ArH), 6.91 (d, *J* = 3.7 Hz, 1H, ArH), 6.56 (ddd, *J* = 7.9, 2.0, 1.0 Hz, 1H, ArH). ^{13}C NMR (75 MHz, DMSO-d6) δ (ppm): 159.79, 157.61, 149.65, 144.21, 130.12, 129.95, 120.11, 115.09, 112.75, 109.69, 107.49. FTIR (KBr) ν cm^{-1}: 3439,3350, 3241, 3122, 3056, 2961, 2920, 2850, 1682, 1621, 1604, 1573, 1525, 1490,1471, 1376, 1257, 959, 949, 803, 781. HRMS (ESI-QOrbitrap) *m/z*: $C_{11}H_9NO_3$ calcd 202.0510, found 202.0512 [M-H]$^-$.

5.2.1.8 Procedure for the synthesis of 5-(3-cyanophenyl)furan-2-carboxamide (67)
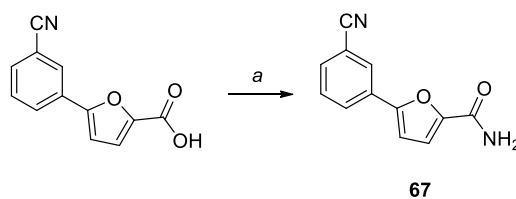
Reagents and conditions: *a*) *1.* HATU, DIPEA, DMF, r.t., 30 min.; *2.* NH$_4$Cl, r.t., 2 h.

HATU (1.0 mmol) and DIPEA (0.70 mL) were added to a solution of 5-(3-cyanophenyl)furan-2-carboxylic acid (**54**, 1.0 mmol) in *N,N*-dimethylformamide (3.70 mL), and the resulting mixture was stirred for 30 minutes at room temperature. Then, ammonium chloride (3.0 mmol) was added, and the stirring was continued for 2 more hours. After completion, the reaction was neutralised with 1M HCl and partitioned between ethyl acetate and water. The organic layer was washed three times with cold water, dried over anhydrous sodium sulphate and concentrated *in vacuo*. The crude product was purified by crystallisation from dichloromethane/hexane to afford an off-white solid. Yield: 32%. TLC (dichloromethane – methanol 95:5): R_f = 0.27. Mp: 202.0°C. ^{1}H-NMR (300 MHz, DMSO-d6) δ (ppm): 8.45 (t, *J* = 1.7 Hz, 1H, ArH), 8.21 (d, *J* = 7.8 Hz, 1H, ArH), 8.08 (bs exch. D$_2$O, 1H, NH$_2$), 7.80 (d, *J* = 7.8, 1H, ArH), 7.65 (t, *J* = 7.8 Hz, 1H, ArH), 7.52 (bs exch. D$_2$O, 1H, NH$_2$), 7.25 (d, *J* = 3.6 Hz, 1H, ArH), 7.15 (d, *J* = 3.6 Hz, 1H, ArH).

^{13}C NMR (75 MHz DMSO-d6) δ (ppm): 159.51, 152.57, 148.48, 132.29, 131.06, 130.62, 129.01, 128.15, 118.96, 116.18, 112.66, 109.92. FTIR (KBr) ν cm^{-1}: 3474, 3167, 2962, 2923, 2852, 2225, 1697, 1614, 1535, 1518,1470, 1425, 1395, 1261, 1099, 1037, 958, 903, 893, 798. HRMS (ESI-QOrbitrap) *m/z*: $C_{12}H_8N_2O_2$ calcd 213.0659, found 213.0659 [M+H]$^+$.[51]

5.2.1.9 Procedure for the synthesis of 5-bromobenzo[*c*][1,2,5]oxadiazole (141)

140   **141**
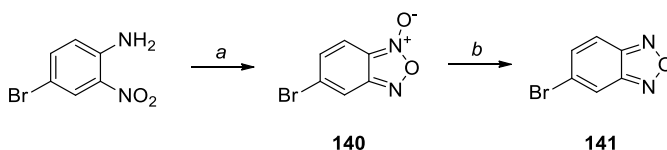
Reagents and conditions: *a*) *1.* KOH, EtOH, 47°C → 65°C, 2 h; *2.* 15% NaClO, 2-5°C, 90 min; *3.* 20°C, 18 h; *b*) PPh$_3$, toluene, 110°C, 4 h, Ar.

6-Bromobenzo[*c*][1,2,5]oxadiazole 1-oxide (140). A stirred solution of potassium hydroxide (1.6 mmol) in ethanol (0.95 mL) at 47°C was treated portion-wise with 4-bromo-2-nitroaniline (1 mmol), ensuring the maintenance of a constant temperature; then, the solution was heated at 65°C for 2 hours. The reaction mixture was cooled to 2°C, treated with aqueous sodium hypochlorite (15%, 2.1 mL), and stirred for 90 minutes, keeping the temperature below 5°C. Then, the reaction was gradually warmed to room temperature and stirred for an additional 18 hours. After completion, the precipitate was filtered off, washed with water and dried *in vacuo*, to afford the desired product as a red solid. The product was used without further purification.[33]

5-Bromobenzo[*c*][1,2,5]oxadiazole (141). Triphenylphosphine (1.1 mmol) was dissolved in dry toluene (3 mL) and stirred at 110°C under argon atmosphere; then, 6-bromobenzo[*c*][1,2,5]oxadiazole 1-oxide (**140**, 1 mmol) in toluene (0.5 mL) was added dropwise over 1 hour and the resulting mixture was stirred for 3 hours. After completion, the solvent was evaporated *in vacuo*, and the crude was purified through flash column chromatography (cyclohexane – diethyl ether 98:2), to afford the desired product as a pale pink solid. Yield: 60%. TLC (cyclohexane – diethyl ether 98:2): R$_f$ = 0.50. Mp: 74°C. ^{1}H-NMR (300 MHz,

DMSO-d$_6$) δ (ppm): 8.49 (s, 1H, ArH), 8.04 (d, J = 9.4 Hz, 1H, ArH), 7.69 (dd, J = 9.4, 1.6 Hz, 1H, ArH).[33]

5.2.1.10 Procedure for the synthesis of 5-bromoisobenzofuran-1(3H)-one (143)
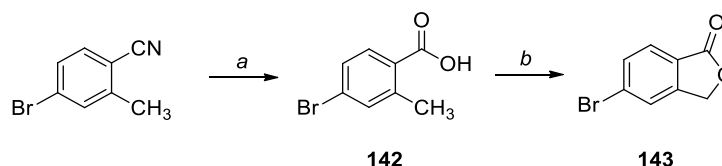
Reagents and conditions: *a*) KOH, H$_2$O, reflux, overnight; *b*) KBrO$_3$, NaHSO$_3$, EtOAc/H$_2$O, 20°C, 40 h.

4-Bromo-2-methylbenzoic acid (142). 4-Bromo-2-methylbenzonitrile (1 mmol) was suspended in water (49 mL) and potassium hydroxide was added (50 mmol); the reaction mixture was stirred at reflux overnight. After cooling, the basic solution was washed with ethyl acetate (3 x 15 mL), acidified with hydrochloric acid (6M) and extracted with ethyl acetate (3 x 15 mL). The organic phase deriving from the acidic extraction was dried over anhydrous sodium sulphate and concentrated *in vacuo*, yielding the desired product as a white solid. Yield: 60%. TLC (dichloromethane – methanol 9:1): R$_f$ = 0.46. Mp: 182°C. ^{1}H-NMR (300 MHz, CD$_3$OD) δ (ppm): 7.80 (d, J = 8.4 Hz, 1H, ArH), 7.48 (s, 1H, ArH), 7.42 (d, J = 8.4 Hz, 1H, ArH), 2.56 (s, 3H, CH$_3$).[33]

5-Bromoisobenzofuran-1(3*H*)-one (143). Potassium bromate (3 mmol) in water (1.1 mL) was added to a solution of 4-bromo-2-methylbenzoic acid (**142**, 1 mmol) in ethyl acetate (3 mL); then, a solution of sodium bisulfite (3 mmol) in water (3 mL) was slowly dripped. The resulting two-phase mixture was stirred at 20°C for 40 hours. After completion, the organic phase was separated, and the aqueous layer extracted with ethyl acetate (3 x 3 mL); the unified organic phases were dried over anhydrous sodium sulphate and concentrated *in vacuo*, affording the desired product as a white solid. Yield: 59%. TLC (cyclohexane – ethyl acetate 8:2): R$_f$ = 0.38. Mp: 155°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.79 (d, J = 8.5 Hz, 1H, ArH), 7.73 – 7.64 (m, 2H, ArH), 5.30 (s, 2H, CH$_2$).[33]

5.2.1.11   Procedure for the synthesis of 5-(4-bromophenyl)-2H-tetrazole (144)

144
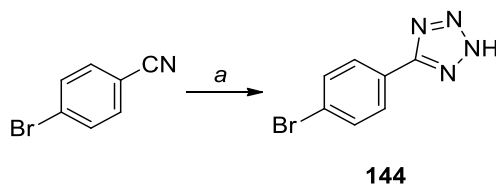
Reagents and conditions: *a*) NH$_4$Cl, NaN$_3$, LiCl, DMF, reflux, 20.5 h, N$_2$.

5-(4-Bromophenyl)-2*H*-tetrazole (144). A mixture of 4-bromobenzonitrile (1 mmol), ammonium chloride (0.1 mmol), sodium azide (1.1 mmol) and lithium chloride (1 mmol) in dry DMF (0.5 mL) was heated at reflux for 20.5 hours under nitrogen atmosphere. After completion, the reaction was diluted with water, and hydrochloric acid (2M) was added to obtain the precipitation of the product. The resulting solid was collected by filtration, dried *in vacuo* and purified through an acid-base extraction. Beige solid. Yield: 75%. TLC (dichloromethane – methanol 8:2): R$_f$ = 0.17. Mp: 271°C (dec.). ^{1}H-NMR (300 MHz, DMSO-d$_6$) δ (ppm): 7.96 (d, *J* = 8.6 Hz, 2H, ArH), 7.82 (d, *J* = 8.6 Hz, 2H, ArH).[33]

5.2.1.12   Procedure for the synthesis of 2-bromo-4-nitroaniline (145)

145
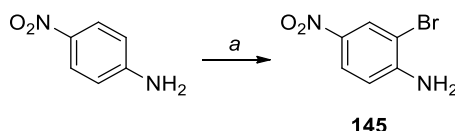
Reagents and conditions: *a*) NBS, PEG 400, r.t., 20 min.

2-Bromo-4-nitroaniline (145). NBS (1.05 mmol) was slowly added to a suspension of 4-nitroaniline (1 mmol) in PEG-400 (2 g). The mixture was stirred at room temperature for 20 minutes. After completion, the mixture was poured into water and extracted with ethyl acetate. The combined extracts were concentrated, and the residue was purified by flash column chromatography (cyclohexane – ethyl acetate 8:2) to obtain the pure monobrominated product as a yellow solid. Yield: 61%. TLC (cyclohexane – ethyl acetate 7:3): R$_f$ = 0.50. ^{1}H-

NMR (300 MHz, CDCl$_3$) δ (ppm): 8.38 (d, J = 2.1 Hz, 1H, ArH), 8.03 (d, J = 8.9 Hz, 1H, ArH), 6.74 (d, J = 8.9 Hz, 1H, ArH).[38]

5.2.1.13 Procedure for the synthesis of 3-bromo-N-methylbenzamide (146)

146
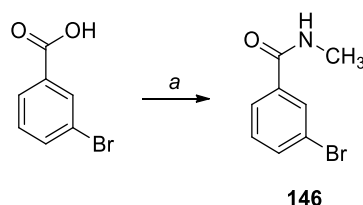
Reagents and conditions: *a*) CH$_3$NH$_2$, HATU, DIPEA, THF, r.t., 18 h.

3-Bromo-*N*-methylbenzamide (146). To a solution of 3-bromobenzoic acid (1 mmol) in THF (4 mL), HATU (1.5 mmol) and DIPEA (2 mmol) were added at 0°C. The reaction mixture was stirred for 30 min at room temperature. Methylamine (2M in THF) (2 mmol) was added at room temperature and the reaction was stirred for 18 hours. After completion, the mixture was extracted with ethyl acetate. The organic phase was dried over sodium sulphate, filtered and concentrated under reduced pressure. The resulting residue was purified by flash column chromatography (dichloromethane – methanol 97:3) to afford a grey solid. Yield: 71%. TLC (dichloromethane – methanol 97:3): R$_f$ = 0.33. Mp: 90.8°C. ^{1}H-NMR (300 MHz, CDCl$_3$) δ (ppm): 7.89 (bs exch. D$_2$O, 1H, NH), 7.66 (d, J = 7.8 Hz, 1H, ArH), 7.56 (d, J = 7.8 Hz, 1H, ArH), 7.23 (t, J = 7.8 Hz, 1H, ArH), 6.80 (s, 1H, ArH), 2.95 (d, J = 4.7 Hz, 3H, CH$_3$).[52]

5.2.1.14 Procedure for the synthesis of 3-bromo-5-phenoxybenzonitrile (147)

147
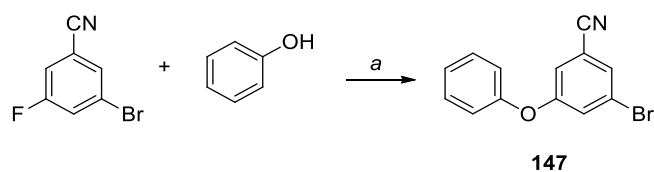
Reagents and conditions: *a*) K$_2$CO$_3$, DMSO, 135°C, 5 min., MW, N$_2$.

3-Bromo-5-fluorobenzonitrile (1.0 mmol), phenol (1.0 mmol) and potassium carbonate (2.0 mmol) were dissolved in dry dimethyl sulfoxide (1.2 mL) under nitrogen atmosphere. The mixture was stirred in a microwave synthesizer for 5 minutes at 135°C. The reaction mixture was diluted with water and extracted with diethyl ether (3 x 3mL). The organic layer was washed with cold water, dried over anhydrous sodium sulphate, filtrated and concentrated *in vacuo*. The product was directly used in the next reaction without further purification. White solid. TLC (cyclohexane – ethyl acetate 8:2): R_f = 0.72. Not isolated.[53]

5.2.1.15   Analytical data for the synthesis of (*E*,*Z*)-3-((1-carboxyprop-1-en-1-yl)oxy)-2-hydroxybenzoic acid (Methyl-AMT)

Methyl-AMT was synthesised according to the reported procedure, to be used as the reference compound in the biological assays.[24] The analytical data for the intermediates and the final compound are listed below.

Methyl 2-diazo-2-(dimethoxyphosphoryl)acetate. ^{1}H NMR (300 MHz, CDCl$_3$) δ (ppm): 3.84 (s, 3H, CH$_3$), 3.80 (s, 3H, CH$_3$), 3.79 (s, 3H, CH$_3$).

Methyl 3-(1-(dimethoxyphosphoryl)-2-methoxy-2-oxoethoxy)-2-hydroxybenzoate. ^{1}H NMR (300 MHz, CDCl$_3$) δ (ppm): 11.03 (bs exch. D$_2$O, 1H, OH), 7.58 (d, *J* = 8.0 Hz, 1H, ArH), 7.27 (d, *J* = 8.0 Hz, 1H, ArH), 6.79 (t, *J* = 8.0 Hz, 1H, ArH), 5.27 (d, *J* = 18.7 Hz, 1H, CH), 3.96 (s, 3H, CH$_3$), 3.94 (s, 6H, CH$_3$), 3.82 (s, 3H, CH$_3$).

(*E*,*Z*)-Methyl 2-hydroxy-3-((1-methoxy-1-oxobut-2-en-2-yl)oxy)benzoate. ^{1}H NMR (300 MHz, CDCl$_3$) δ (ppm): 11.02 (s, 1H, OH), 10.95 (s, 1H, OH), 7.56 (dd, *J* = 8.1, 0.9 Hz, 1H, ArH), 7.51 (dd, *J* = 8.0, 0.9 Hz, 1H, ArH), 7.06 (dd, *J* = 7.9, 1.4 Hz, 1H, ArH), 6.90 (dd, *J* = 7.9, 1.4 Hz, 1H, ArH), 6.82 – 6.68 (m, 3H, ArH+CH$_2$), 5.87 (q, *J* = 7.6 Hz, 1H, CH), 3.96 (s, 6H, OCH$_3$), 3.76 (s, 3H, OCH$_3$), 3.70 (s, 3H, OCH$_3$), 2.09 (d, *J* = 7.6 Hz, 3H, CH$_3$), 1.81 (d, *J* = 7.2 Hz, 3H, CH$_3$).

(*E*,*Z*)-3-((1-Carboxyprop-1-en-1-yl)oxy)-2-hydroxybenzoic acid. ^{1}H NMR (300 MHz, D$_2$O) δ (ppm): 7.47 (dd, *J* = 8.0, 1.5 Hz, 1H, ArH), 7.41 (dd, *J* = 8.0, 1.5 Hz, 1H, ArH), 7.01 (dd, *J* = 8.0, 1.5 Hz, 1H, ArH), 6.90 (dd, *J* = 8.0, 1.5 Hz, 1H,

ArH), 6.80 – 6.68 (m, 3H, ArH+CH$_2$), 5.89 (q, J = 7.6 Hz, 1H,), 1.88 (d, J = 7.6 Hz, 3H, CH$_3$), 1.61 (d, J = 7.2 Hz, 3H, CH$_3$).

5.3 BIOLOGICAL METHODS

The biological evaluation of the compounds was performed in collaboration with Prof. Laurent R. Chiarelli at the Department of Biology and Biotechnology "L. Spallanzani", University of Pavia in Italy.

5.3.1 Production and purification of chorismic acid

Chorismic acid, the natural substrate of MbtI, was produced using the *Escherichia coli* KA12 strain, according to Grisostomi, Hilver *et al.*, and purified by a modified protocol.[54] The identity of the substance was confirmed by ^{1}H-NMR and HRMS. The so-obtained product was used in the enzymatic assays to evaluate the inhibitory effect of the candidate compounds.

One liter of growth medium A (2 g/L casamino acids, 2 g/L yeast extract, 41 mg/L l-tryptophan, 20 mL 50X Vogel & Bonner salts, 0.16% w/v glucose) was inoculated with 5 mL of an overnight pre-culture. The cultures were incubated at 30°C (220 rpm) until OD_{600} 1.9–2.1, then centrifuged at 2000 g for 20 minutes and resuspended in 1 L non-sterile accumulation medium B (12.8 g/L Na_2HPO_4, 1.36 g/L KH_2PO_4, 18 g/L glucose, 2.7 g/L NH_4Cl, 20.3 mg/L $MgCl_2·6H_2O$, 2 mg/L L-tryptophan). This suspension was shaken at 30°C (220 rpm) for 16 hours, allowing the production of chorismate and its secretion into the medium. After removing the cells by centrifugation at 2000 g for 20 min, the supernatant was collected, frozen with liquid nitrogen and stored at -80°C.

The culture medium (1 L) obtained from the previous procedure was treated to isolate chorismate. The volume reduction was achieved using a rotary evaporator at room temperature and low pressure (6 mbar). The concentrated fractions were acidified with diluted HCl (3M) to pH 1.5; the aqueous phase was then extracted with ethyl acetate (3 x 20 mL). The combined organic layers were dried over anhydrous sodium sulphate and then concentrated *in vacuo*.

Chorismic acid purification was performed using a Varian LC-940 HPLC system (Varian, Leinì, Italy) equipped with a binary pump system (pump head volume 10 mL), an autosampler, a fraction collector, a UV-DAD detector operating at λ = 280 nm and an analytical/semi-preparative scale-up module. Analytical and semipreparative separations were carried out using a Kinetex C18 column (100

mm × 3 mm, pore size 100 Å, particle size 5 μm) and AXIA Kinetex C18 column (100 mm × 21.2 mm, pore size 100 Å, particle size 5 μm) (Phenomenex), respectively. Solvent system: 10% acetonitrile in 0.05% formic acid. Flow 2.0 mL/min for the analytical runs and 20.0 mL/min for semi-preparative scale up. Volume of injection: 10 μL for the analytical configuration and 4000 μL for the semi-preparative scale-up. Automatic collection was adjusted each time to compensate variations in the chromatographic peak retention times (RT).

The aqueous fractions, deriving from the chromatographic procedure, were extracted with ethyl acetate, to afford the desired compound as a free acid. The combined organic layers were dried over sodium sulphate and then concentrated *in vacuo*. The resulting white solid was freeze-dried and stored under inert atmosphere (N_2) at -80°C, in order to prevent its degradation.

The NMR spectrum was acquired on a Varian-Oxford 300 MHz instrument (Varian, Palo Alto, CA, USA), operating at 300 MHz for ^1H. The high-resolution mass spectrometry analysis was performed on a Q-TOF Synapt G2-Si (Waters, Milford, MA, USA).

White solid. Yield: 200 mg. TLC (dichloromethane–methanol-acetic acid 9:1:0.1): R_f = 0.11. Mp: 140°C. ^1H-NMR (300 MHz, D_2O) δ (ppm): 6.84 (bs exch. D_2O, 1H, H_1), 6.26 (d, J = 10.0 Hz, 1H, H_5), 5.94 (dd, J = 10.0, 2.7 Hz, 1H, H_4), 5.47 (d, J = 3.5 Hz, 1H, $H_{6/7}$), 4.97 (dd, J = 11.4, 2.7 Hz, 1H, H_2), 4.83 (d, J = 3.5 Hz, 1H, $H_{6/7}$), 4.69-4.60 (m, obscured by water peak, 1H, H_3). HRMS (ESI-QOrbitrap) *m/z*: $C_{10}H_{10}O_6$ Na ($^{+1}$) calcd 249.03696, found 249.03645; $C_{10}H_8O_6$ Na ($^{-1}$) calcd 247.02241, found 247.02260; $C_{10}H_7O_6$ Na_2 ($^{-1}$) calcd 269.00435, found 269.00476, $C_{10}H_9O_6$ Na_2 ($^{+1}$) calcd 271.01890, found 271.01832, $C_{20}H_{19}O_{12}$ Na_2 ($^{+1}$) calcd 497.06664, found 497.06535.

5.3.2 MbtI expression and purification

The Rv2386c gene, encoding the salicylate synthase (MbtI) from *M. tuberculosis*, was amplified by PCR, using the following primers: forward (5'-ATGGGTCGCGGATCCCTGGAAGTTCTGTTCCAGGGGCCCGTGTCCGAGCTCAGCGTC-3') and reverse (5'- TGCGGCCGCAAGCTTTTACTGGCGTGCAACCAGAT- 3'). The

primers were designed according to the In-fusion HD Cloning Kit protocol; the forward primer carried the PreScission Protease cleavage site in order to remove the 6-histidine tag from the protein. The purified PCR fragment was recombined into the *Bam*HI-*Hind*III digested pET-28b following the manufacturer's instructions, to give the pET28b-MbtI vector. The recombinant products were transformed into *E. coli* Stellar™ competent cells and the resulting colonies were checked for the presence of the insert by colony PCR and sequencing.

For protein expression, the *E. coli* BL21 (DE3) strain was transformed with pET28b-MbtI, plated onto LB-kanamycin (50 µg/mL) and a single colony was used as a starter culture. 6x1 L of LB-kanamycin was inoculated with 20 mL of starter culture and grown at 37°C until OD_{600} = 0.6. Protein expression was induced with 0.5 mM IPTG at 25°C for 15 hours.

Cells were harvested by centrifugation, resuspended in 50 mM potassium phosphate buffer pH 8.0, 500 mM KCl, (Buffer A), supplemented with 1 mM phenylmethylsulphonyl fluoride and lysed by sonication. The cell-free extract, obtained by centrifuging the lysate at 50000 g for 20 minutes at 4°C, was loaded on a His-Trap (1 mL, GE Healthcare) equilibrated in buffer, the column washed with 50 mM imidazole in buffer A, and MbtI eluted with 250 mM imidazole in buffer A.

The purified enzyme was dialysed in 50 mM potassium phosphate pH 8.0, 150 mM KCl, 2 mM DTT (Buffer B), and digested with PreScission protease (GE Healthcare, 400 mU/mL). The digested protein was further purified by a second affinity chromatography, followed by size exclusion chromatography on HiLoad Superdex 200 column (GE Healthcare) equilibrated in buffer B. Samples purity was checked by SDS-PAGE and protein concentration evaluated according to Lowry.[55]

5.3.3 MbtI inhibition assays

The residual enzymatic activity of MbtI upon treatment with the *p*-substituted candidates was determined at 37°C by an L-lactic coupled assay.[24] The standard reaction mixture contained 50 mM HEPES pH 8.0, 0.2 mM NADH,

10 mM MgCl$_2$, 3 units of L-lactic dehydrogenase (Sigma-Aldrich, St. Louis, MO, USA), 10-20 μM MbtI, and the reactions were started by the addition of the substrate (chorismate, CHA).[33]

For all other compounds the inhibition tests were performed with a different method, in which the production of salicylic acid was measured by a fluorimetric assay.[32] Briefly, the experiments were performed at 37°C in a final volume of 400 μL of 50 mM Hepes pH 7.5, 5mM MgCl$_2$, using 1–2 μM MbtI and the reactions were started by the addition of chorismic acid and monitored using a Perkin-Elmer LS3 fluorimeter (Ex. λ = 305 nm, Em. λ = 420 nm).[38]

Steady-state kinetic parameters were determined by assaying the enzyme at variable concentrations of chorismate (10-500 μM). The experiments were performed in triplicate, and the kinetic constants K_m and k_{cat} were determined by fitting the data to the Michaelis-Menten equation using the Origin 8 software. Initially, MbtI inhibition was screened for all compounds at 100 μM (dissolved in a solution containing 10% DMSO) and 50 μM chorismic acid. The solubility of the compound in the medium assay was verified up to 100 μM concentration.

For compounds that significantly inhibited the enzyme activity, IC$_{50}$ and K$_i$ values were determined. For IC$_{50}$ determinations, the enzyme activities were measured in the presence of the compound and the values were estimated according to the following equation, where A$_{[I]}$ is the enzyme activity at inhibitor concentration [I] and A$_{[0]}$ is the enzyme activity without inhibitor.[33]

$$A_{[I]} = A_{[0]} \cdot \left(1 - \frac{[I]}{[I] + IC_{50}} \right)$$

The inhibition constant (K$_i$) values were determined by assaying the MbtI enzymatic activity at different substrate and compound concentrations, using an adapted equation for competitive inhibition with Origin 8 software.[33]

$$v = \frac{V_{max} \cdot [S]}{[S] + K_m \left(1 + \frac{[I]}{K_i} \right)}$$

5.3.4 Pan Assay Interference Compounds (PAINS) analysis

As previously reported, there are a number of substructural features which could help identify compounds that appear as PAINS in many biochemical assays.[56] The corresponding filters were included in the Filter-it™ software and the most active compounds were thus analysed by using this program. The results highlighted that the candidates do not possess any of the substructural features shared by the most common PAINS. Moreover, to verify the possible formation of aggregates, the ability of the compounds to inhibit MbtI activity was also tested in the presence of 0.1 mg/mL of bovine serum albumin (BSA) or in the presence of 0.01% (v/v) Triton X-100 as detergent.[57] The IC_{50} values of the compounds in the presence of these additives were not significantly modified, supporting the notion that the ligands do not act by forming aggregates. Finally, to exclude a promiscuous enzyme inhibition due to covalent reaction with cysteines, MbtI inhibition activity was also tested in the presence of 100 μM of 1,4-dithio-dl-threitol (DTT).[58] Again, the IC_{50} was not significantly influenced by DTT, thus excluding an interaction of the compounds with the cysteine residues of MbtI.[33]

5.3.5 Minimal inhibitory concentration determinations

The MIC[99] of **1** (as sodium salt) was determined against *M. tuberculosis* H37Rv in solid medium. A single colony was inoculated in complete Middlebrook 7H9, supplemented with 10% OADC Middlebrook Enrichment, and grown at 37°C until exponential growth phase (~108 CFU/mL). Dilutions to the final concentration of ~106 CFU/mL were performed, and about 1 mL of cell culture was streaked onto plates containing two-fold serial dilutions of the compound. The MIC value was assigned as the lowest drug concentration inhibiting bacterial cell growth. The experiments were performed in triplicate. Additionally, the compound was assayed against *M. bovis* BCG in liquid medium, in high and low iron media, by determining the MIC against cells grown both in 7H9/OADC medium and in low-iron Chelated Sauton's medium.[33]

The MIC[99] of **27** and **54** (as sodium salts) was determined against *M. bovis* BCG, in low-iron Chelated Sauton's medium, using the resazurin reduction assay (REMA).[38]

5.3.6 Siderophore production assay

Siderophore activity in the culture was tested in *M. bovis* BCG through the isolation of mycobactins or using the Universal CAS liquid assay.[37]

To this purpose, *M. bovis* was grown in 7H9 medium, subcultured in chelated Sauton's medium, then diluted 1:1000 in chelated Sauton's containing different concentrations of the most active compounds. After 15 days of incubation at 37°C, cells were harvested by centrifugation. For the isolation of mycobactins, cell pellets were extracted in ethanol overnight. 0.1M $FeCl_3$ in ethanol was added until no further change in colour was observed, and the mixture was incubated at room temperature for 1 hour. Mycobactins were then extracted in chloroform, washed 3 times with water to remove the excess of iron, and evaporated. The residue was dissolved in methanol. The concentration of the mycobactins was determined by measuring the absorbance at 450 nm (1% solution of mycobactins gives an absorbance of 42.8).[33]

For the Universal CAS assay, 100 μL of the supernatant of the cell cultures was mixed with 100 μL of CAS assay liquid solution in a 96-well plate; then, the solution was incubated for 10 minutes at room temperature, and the absorbance was measured at 630 nm. The siderophore units were calculated using the following equation, where A_r is the absorbance of the blank medium with the CAS assay solution at 630 nm, and A_s is the absorbance of the supernatants with the CAS assay solution.[33]

$$\frac{A_r - A_s}{A_r} \cdot 100$$

5.3.7 Cell viability assay

MRC5 human fibroblast lung cells were maintained at 37°C in a humidified atmosphere containing 5% CO_2, according to the supplier's indications. Cells

were plated in 96-well culture plates. The day after seeding, vehicle or compounds were added at different concentrations to the medium. Compounds were added to the cell culture at a concentration ranging from 200 to 0.01 μM. Cell viability was measured after 96 hours according to the supplier's instructions (G7571; Promega, Madison, WI, USA) with a Tecan M1000 PRO instrument (Tecan, Männedorf, Switzerland). IC_{50} values were calculated from logistic dose-response curves.[38]

5.4 CRYSTALLOGRAPHY

5.4.1 Production and purification of MbtI for crystallisation trials

Transformation. *Escherichia coli* BL21 cells were transformed with a codon-optimised pET-28a plasmid (GenScript, Piscataway, NJ, USA), containing the genic sequence of MbtI from *Mycobacterium tuberculosis*, fused to an N-Terminal His-Tag and a TEV cleavage site. 1 μL of the plasmid solution (100 ng/μL in water) was added in sterile conditions to an Eppendorf tube containing the cells, and the mixture was allowed to rest in ice for 20 minutes. After a heat-shock sequence (45 seconds at 42°C), performed using an Eppendorf ThermoMixer® C (Eppendorf, Hamburg, Germany) at 0 rpm, the cells were moved back in ice for 5 minutes. Then, 1 mL of Miller's LB medium was added to the tube, and the cells were heated at 37°C with 1250 rpm shaking for 1 hour in the same thermomixer. The mixture was centrifuged at 3000 rpm for 4 minutes using an Eppendorf 5417C centrifuge, and part of the supernatant was removed. Cells were re-suspended (circa 150 μL) and then plated on an LB-agar Petri dish, containing kanamycin (50 μg/mL) to allow the selective growth of transformed colonies. The culture was stored overnight at 37°C in an incubator (Jouan EB-170, ThermoFisher Scientific, Waltham, MA, USA).

Production by auto-induction. A sterile 100 mL Erlenmeyer flask was prepared with 9.2 mL of 2YT medium pH 7.0, 200 μL of 40% glucose, 100 μL of 1 M $MgSO_4$, 500 μL of 20X NPS buffer, 20 μL of 50 mg/mL kanamycin (Sigma-Aldrich). One or two transformed colonies were added to the so-obtained starter medium using a cell scraper, and the mixture was stirred at 37°C with 180 rpm shaking for 7 hours in an Infors HT shaking incubator (Infors, Bottmingen, Switzerland).

Then, sterile 5 L Erlenmeyer flasks were filled with 1 L of 2YT medium pH 7.0, 50 mL of auto-induction supplement (0.5% glycerol, 0.05% glucose, 0.2% α-lactose, 25 mM $(NH_4)_2SO_4$, 50 mM KH_2PO_4, 50 mM Na_2HPO_4, 1 mM $MgSO_4$) and 2 mL of 50 mg/mL kanamycin; 1 mL of the starter suspension was added to each flask, and the culture was stirred at 180 rpm for 4 hours at 37°C and 12 hours at 25°C in a shaking incubator.

Extraction. The resulting cell culture was collected from the flasks and centrifuged at 4000 rpm for 20 minutes at 4°C in a Sorvall™ Lynx 4000 centrifuge (ThermoFisher Scientific); the cell pellet was collected and used in the next steps or stored at -80°C, while the supernatant was discarded. A 1X PBS solution (20 mL) was used to wash the centrifuge tubes, thereby allowing the retrieval of all the deposited cells.

The cell pellet (10-15 g from 2 L of culture) was re-suspended using 50 mL of IMAC A solution (25 mM Tris·HCl pH 8.5, 300 mM NaCl, 25 mM imidazole), with the addition of 50 μL of Benzonase® Nuclease (Sigma-Aldrich) and a tablet of EDTA-free protease inhibitor complex (cOmplete™, Sigma-Aldrich). The mixture was stirred with a vortex, or overnight on a rotary shaker (10 rpm) at 4°C when using frozen pellets. Then, the cells were lysed with a CF2 cell disruptor (Constant Systems Ltd., Daventry, UK), employing a pressure of 20 kpsi (1.4 kbar). The resulting lysate (circa 100 mL) was centrifuged at 14000 rpm for 2 hours, at 4°C, to remove the membranes; the supernatant containing the protein was collected, while the cell residue was discarded.

Purification. The supernatant was transferred in an Erlenmeyer flask and charged on a 1 mL HisTrap™ High Performance column (GE Healthcare, Chicago, IL, USA) for the first step of the purification process. The affinity chromatography was performed on an Äkta Explorer 3D purifier (GE Healthcare), using a linear gradient obtained with two buffers containing different concentrations of imidazole, namely IMAC A (25 mM Tris·HCl pH 8.5, 300 mM NaCl, 25 mM imidazole) and IMAC B (25 mM Tris·HCl pH 8.5, 300 mM NaCl, 400 mM imidazole). The collection volume was set at 500 μL and the flow was 1 mL/min; the peak corresponding to the protein usually eluted between fractions A7-B1 (96-well plates, serpentine mode), with an average absorbance of 800 mAU at 280 nm. The most significant fractions were chosen to be analysed through electrophoresis on a 12% SDS-PAGE gel, employing a Tris-glycine SDS running buffer and 200 V constant current for about 1 hour (equipment from BioRad, Hercules, CA, USA). The samples were prepared by diluting 5 μL of eluate with 10 μL of water, followed by the addition of 5 μL of 4X NuPAGE buffer (Invitrogen, Carlsbad, CA, USA) additioned with 0.5 mM TCEP; the mixture was heated at 70°C

for 5 minutes. The run was monitored through the PageRuler™ Prestained Protein Ladder (ThermoFIsher Scientific); the gel was revealed with the Instant*Blue*™ Ultrafast Protein Stain (Sigma-Aldrich) and shaken on an orbital platform shaker. An example of a gel run is shown below in Figure 30.

Figure 30. Photograph of an SDS-PAGE gel with samples deriving from the affinity chromatography, taken with a BioRad ChemiDoc™ Imaging System. 1: PageRuler™ Pre-stained Protein Ladder; 2: supernatant before purification; 3: outflow before elution; 4, 5: initial fractions; 6, 7: central fractions; 8, 9: final fractions.

Based on the results of the electrophoresis analysis, the fractions containing the protein were collected (9-10 mL) and their concentration was checked using a NanoDrop™ 1000 (ThermoFisher Scientific) spectrophotometer (circa 1.9 mg/mL). Then, TEV protease (200 µL of a 10 mg/mL solution) was added to cleave the His-tag from the protein, along with 5 µL of 1 M DTT (final concentration ≈ 100 mM). The resulting solution was transferred in a SnakeSkin™ tube (3.5K MWCO, 22 mm; ThermoFisher Scientific), and dialysed overnight at 4°C in a solution containing 25 mM Hepes·NaOH pH 8.0, 150 Mm NaCl, 0.5 mM DTT.

The dialysed solution was collected and loaded on a BioRad column charged with Ni-NTA resin (0.5-1 mL re-suspended in the dialysis buffer; Qiagen, Hilden, Germany), to remove the His-tag. To prevent the undesired retention of the protein on the resin, 40 mM imidazole was added to the protein solution before the chromatography. Finally, the column was washed with IMAC B, and a sample was kept to be analysed by electrophoresis to make sure that the elution was complete.

The so-obtained protein solution was concentrated to about 1 mL for the next purification step, using 20 mL Vivaspin™ 15R Centrifugal Concentrators equipped with a 10000 Da cut-off filter (Sartorius, Göttingen, Germany), centrifuged at 5000 rpm at 4°C (circa 20 minutes) in an Eppendorf 5804R centrifuge. The resulting solution was additionally centrifuged at high speed (14000 rpm, 15 minutes) in an Eppendorf tube to allow for the removal of the precipitated protein. The supernatant was loaded on a HiLoad 16/600 Superdex 200 exclusion chromatography column (GE Healthcare), previously equilibrated overnight with the desired buffer (25 mM Hepes·NaOH pH 8.0, 150 mM NaCl, 1% glycerol), through a 2 mL loop. The Äkta Explorer 3D purifier was used in this second purification step, too. The collection volume was set at 1 mL and the flow was 0.5 mL/min; the peak corresponding to the protein usually eluted between fractions D12-E2 (96-well plates, serpentine mode), with an average absorbance of 800 mAU at 280 nm. Like before, the most significant fractions were chosen to be analysed through electrophoresis, in the same abovementioned conditions. An example of a gel run is shown below in Figure 31.
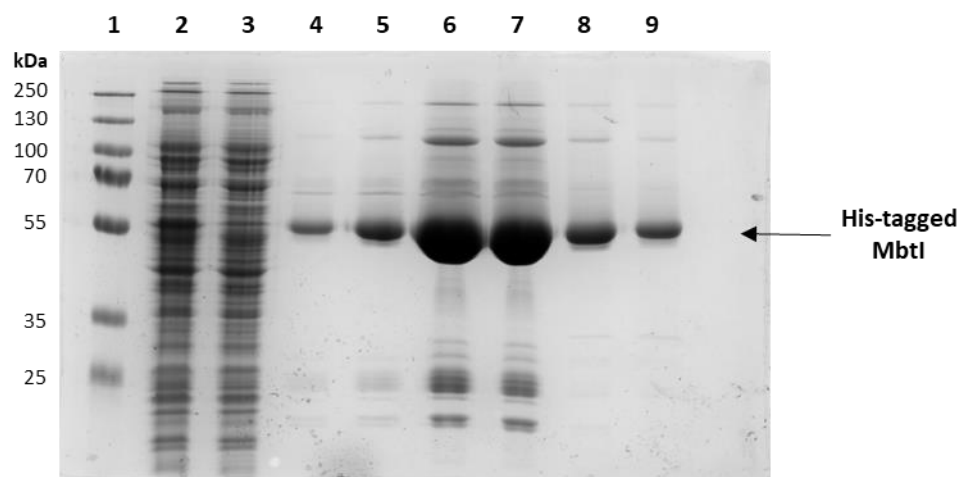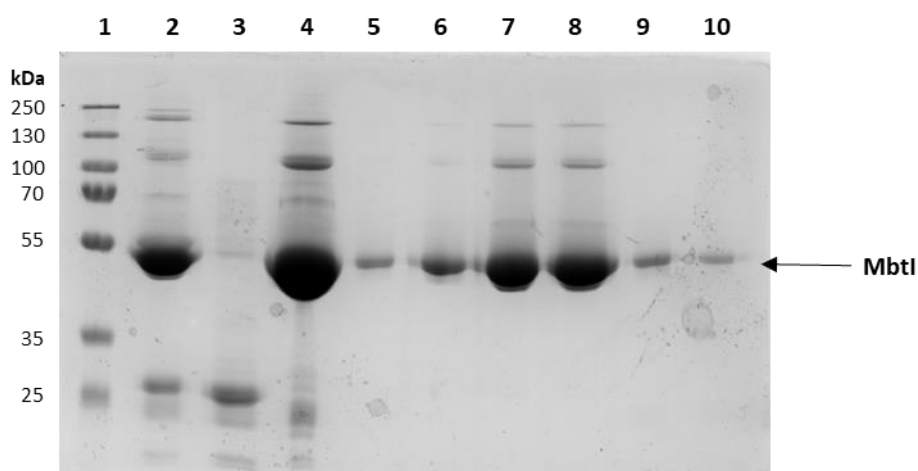

Figure 31. Photograph of an SDS-PAGE gel with samples deriving from the size-exclusion chromatography, taken with a BioRad ChemiDoc™ Imaging System. 1: PageRuler™ Prestained Protein Ladder; 2: protein solution after dialysis; 3: Ni column wash; 4: protein solution before gel filtration; 5, 6: initial fractions; 7, 8: central fractions; 9, 10: final fractions.

The fractions containing the protein were collected (circa 14 mL) and concentrated, to allow the obtainment of the final purified protein solution (usually, about 800 μL, at 22 mg/mL concentration), which was flash-frozen with liquid nitrogen and stored at -80°C.

The buffer solution for the storage of the protein was carefully chosen, considering both literature data and the results of a buffer optimisation analysis, in which the stability of the enzyme was evaluated by instrumental measurement through:

- Dynamic Light Scattering (DLS): the analysis was performed with a DynaPro Plate Reader II (Wyatt Technology, Santa Barbara, CA, USA). This technique allows the determination of the size distribution profile of small particles in suspension. The stability of the protein was tested at 3 mg/mL using 20 different standard buffers; the conditions characterised by a low polydispersity (high homogeneity) and the absence of aggregates after 24 hours were selected as the best options;

- Nano Differential Scanning Fluorimetry (nanoDSF): the analysis was performed with a Prometheus nt48 (NanoTemper, München, Germany). The nanoDSF technology monitors the intrinsic fluorescence signal (no dye) of proteins as a measure of its folding state. A ratiometric measurement of the fluorescent signal is plotted against increasing temperature to determine the Tm. A backscattering measurement is also performed in order to determine conditions where aggregation value appears after the denaturation. The thermal stability of the protein was tested at 3 mg/mL using 20 different standard buffers over a temperature range of 20-95°C (gradient 1°C/min). Denaturation and aggregation appeared at the same time, with a melting temperature on average at 58.2°C, in the best buffer conditions.

In both analyses, a pH value of 8.0 emerged as the best, since lower levels were detrimental for the stability of the protein. The best buffers also contained 50-150 mM NaCl; Tris and Hepes performed analogously, so the latter was chosen, in accordance to the literature conditions.

5.4.2 Crystallisation experiments

Preliminary crystallisation assays. These experiments were performed to define the best conditions to obtain high-quality crystals of MbtI, in its apo form and in complex with selected inhibitors. Starting from the information reported

in the literature (Harrison, Manos-Turvey),[22,28] we attempted an optimisation of the crystallisation conditions by exploring the effect of different variations, including:

- protein concentration: 5-20 mg/mL;

- type and concentration of the precipitating agent: considering the tendency of the protein to crystallise in the presence of polyethylene glycol (PEG), we tried different molecular weights and concentrations, namely 10-20% PEG 2000, 3350, 4000, 6000;

- type, concentration and pH of the buffer solution: 0.15-0.25 M imidazole/malic acid and malic acid/potassium hydroxide pH 5.5-6.5, 0.1-0.2 M acetate buffer pH 4.5-5.5;

- type and concentration of inorganic salts: considering the work of Chi *et al.*, we used 0.2 M potassium thiocyanate instead of a traditional buffer. Moreover, we added different magnesium salts, including $MgSO_4$, $MgCl_2$, $Mg(CH_3COO)_2$, both at low (5-10 mM) and high (0.1 M) concentrations, to the classic PEG-buffer mixture, in an attempt to obtain a crystal structure of the protein with the ionic cofactor embedded in it, which had proven extremely difficult according to the literature;

- time of protein-ligand incubation: both short (30 minutes-2 hours) and long (overnight, 24 hours) incubation periods were tested;

- reservoir and drop volumes: 500 μL or 1 mL for the reservoir, 2-4 μL for the drop, obtained by mixing an equal volume of protein and reservoir solutions;

- temperature: 4°C and 18°C.
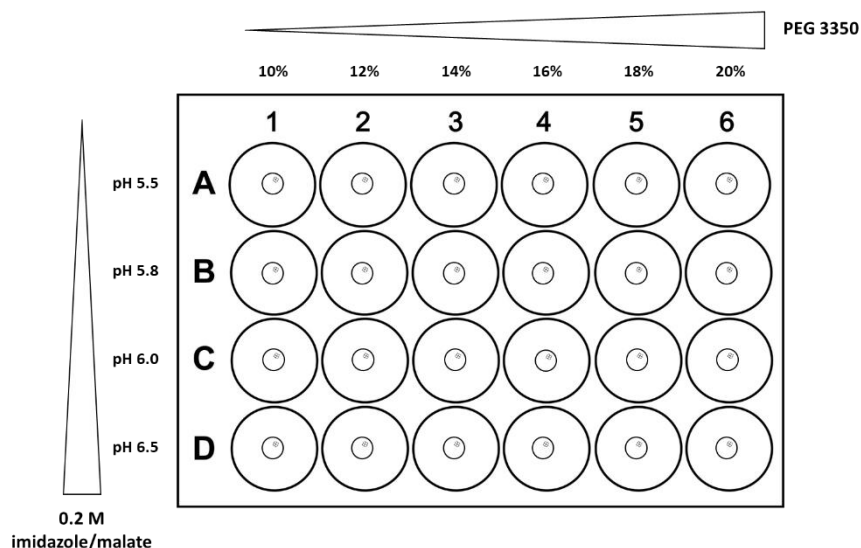
Figure 32. Example of a typical 24-well plate for crystallisation, used for the preliminary experiments on MbtI.

These experiments were performed manually, employing the vapour diffusion technique with the hanging drop or sitting drop methods (Figure 32). All the equipment was purchased from Molecular Dimensions Ltd. (Newmarket, UK), while the chemicals were acquired from Sigma-Aldrich. The protein was produced by Prof. Laurent R. Chiarelli at the Department of Biology and Biotechnology "L. Spallanzani", University of Pavia in Italy. The ligands were selected to meet specific functions, based on their activity on the enzyme:

• positive control: *E,Z*-methyl-AMT was chosen as the reference inhibitor, to validate the crystallisation method;[24]

• natural substrate: some trials were carried out in the presence of chorismic acid (Ba^{2+} salt), although its low stability in solution at 4°C, as well as the enzymatic activity of MbtI, probably made it incompatible with the co-crystallisation process;

• inhibitors: the most active compounds were chosen (**1**, **27**, **54**), and used as sodium salts.

Overall, the best results were obtained at a protein concentration of 20 mg/mL, using 0.2 M imidazole/malate pH 6.0, 10-15% PEG 3350/4000, and storing the crystallisation plates at 4°C. The few crystals formed in these conditions were needle-shaped and diffracted poorly, with resolutions in the range 2.8-3.8 Å; the best sample showed an orthorhombic system with a P2$_1$2$_1$2$_1$ space group (cell

dimensions: *a*: 51.8, *b*: 164.2, *c*: 195.3 Å), which was analogous to that obtained by Harrison, Lott *et al.* (PDB code: 2G5F).[22] The crystals were harvested using Dual Thickness MicroMounts™ (MiTeGen, Ithaca, NY, USA) and stored in Universal V1-Pucks (Unipucks; MiTeGen) under liquid nitrogen. Diffraction data were collected at the SOLEIL Synchrotron in Saint-Aubin, France. Unfortunately, only the apo form of the protein was obtained and no co-crystals were observed, neither with the ligands nor with the magnesium ion.

Optimisation process (Institut Pasteur). An optimised procedure for the production and purification of MbtI was developed, which allowed the obtainment of the enzyme in higher yields. The details of the process are reported in Section 3.4.1.

The protein was subjected to a *de novo* complete screening to identify new crystallisation conditions. Several tests were performed, both on the protein alone, and in complex with the previously described ligands. All experiments were performed in 96-well plates, with the sitting drop vapour diffusion technique. The trials were set up using the Mosquito® crystal Nanolitre Protein Crystallization Robot (TTP Labtech, Melbourne, UK), and the plates were stored in the Rock Imager® 1000 (Formulatrix, Bedford, MA, USA), where pictures of the drops were taken at intervals, following a specified timetable, and then uploaded on a dedicated software for visual checking. Seven screening kits, each comprising 96 different reservoir solutions for a total of 672 conditions, were tested for every protein sample; the following crystallisation screens were used:

- Wizard 1 and 2 (Emerald Biosystems Inc., Bainbridge Island, WA, USA);

- Structure screen 1 and 2 (Molecular Dimensions Ltd.);

- JBS1 to JBS4 (Jena Bioscience, Jena, Germany);

- JBS5 to JBS8 (Jena Bioscience);

- Salt RX (Hampton Research, Aliso Viejo, CA, USA);

- PEGion 1 and 2 (Hampton Research);

- Crystal screen CRYO 1 and 2 (Hampton Research).

The total volume of each drop was 400 nL, obtained by mixing an equal amount of protein and reservoir solutions; the reservoir contained 150 μL. The experiments were performed in duplicate, at 4°C and 18°C.

The protein concentration was always in the range 18-23 mg/mL, while the ligands and the magnesium salts were added at 5 mM, starting from stock aqueous solutions. Different incubation times were tested, but this parameter seemed to have a negligible impact on the co-crystallisation process; during long incubation periods, the protein solution was kept at 4°C, to prevent its degradation. Before each trial, the sample was gradually de-frozen in ice and centrifuged at 14000 rpm for 15 minutes at 4°C to remove traces of precipitated protein.

The growth of the crystals was observed only at 4°C in few conditions with a reasonable reproducibility, between one week to one month after the beginning of the experiment. Most of the samples were plate-shaped, with just few needles, which usually diffracted poorly compared to the plates. The most promising crystals were harvested with CryoLoops™ (Hampton Research) at 4°C and transferred in Universal V1-Pucks (Unipucks; Hampton Research) stored under liquid nitrogen. Different cryo-protectants were tested, namely oil, 25% glycerol and 25% ethylene glycol; overall, glycerol-based cryo-solutions seemed to be tolerated best by the protein, though oil was sometimes a valid option too.

Finally, the best conditions were optimised by designing 24-well and 96-well plates in which the quantities and characteristics (concentration, pH) of the components of the reservoir solution were varied over a small range, using the MatrixMaker robot (Emerald Biosystems Inc.). The drops were made manually (1+1 μL), adopting the hanging drop vapour diffusion method, for the 24-well plates, while the 96-well trials were set up as previously described.

5.4.2.1 Synchrotron X-ray data collection and processing

The collected crystal samples were analysed through X-ray diffraction either at the European Synchrotron Radiation Facility (ESRF) in Grenoble (France) or at the *Source optimisée de lumière d'énergie intermédiaire du LURE*

(SOLEIL) in Saint-Aubin (France). The technical specifications of the two radiation sources are shown below.

At ESRF, the diffraction data were acquired at the Gemini ID23-1 and ID23-2 beamlines. ID23 is dedicated to macromolecular crystallography experiments; it features a single hutch, fitted with the optical devices needed to deliver the beam to both ID23-1 and ID23-2 end-stations, which operate independently.

o ID23-1 is a tuneable beamline capable of MAD measurements from 5 keV to 20 keV energy (2.07 Å to 0.62 Å). It employs a high de-magnification ratio of 1:4.5 to obtain a small focused beam at the sample position. The X-ray beam profile has a Gaussian shape, with an FWHM (full width at half maximum) dimensions of 30 μm vertically and 40 μm horizontally at the sample position. ID23-1 features a U35 undulator, which allows the tuning of the energy across harmonics.

- Energy range: 5-20 keV

- Beam size: minimum 10x10 μm^2, maximum 45x30 μm^2

- Sample environment: 100 K cryostream, automatic sample changer FlexHCD, mini-kappa goniometer, dehydration device and ReX Rapid Nozzle Exchanger (Arinax, Moirans, France)

- Detector: Pilatus 6M (Dectris, Baden, Switzerland), Rontec Xflash X-ray fluorescence detector (Bruker, Billerica, MA, USA)

- Software: MXCuBE2, ISPyB, ADXV

o ID23-2 is fixed-energy (14.200 keV, 0.873 Å) beamline, delivering a flux of about $2.4x10^{12}$ ph/s in a microfocus beam of 10.0x4.0 μm, which is particularly useful for small crystals. The beam is vertically focussed by beryllium compound refractive lenses and horizontally focussed by a mirror in half-Kirkpatrick-Baez geometry. ID23-2 features a U20.2 undulator, which is a pseudo-monochromatic source, positioned upstream at an angular distance of 1.5 mrad compared to U35.

- Energy: 14.2 keV

- Beam size: 10.0x4.0 μm^2 (microbeam)

- Sample environment: 100 K cryostream, automatic sample changer FlexHCD, dehydration device and ReX Rapid Nozzle Exchanger (Arinax, Moirans, France), micro-diffractometer MD3Up

- Detector: Pilatus 3X2M (Dectris)

- Software: MXCuBE3, ISPyB, ADXV

At SOLEIL, the diffraction data were acquired at PROXIMA 1 (PX-1) and PROXIMA 2A (PX2-A) beamlines.

o PX-1 is dedicated to structural biology experiments; it delivers an intense, nearly parallel and tuneable X-ray beam, allowing for high-resolution acquisitions (< 0.8 Å) and data retrieval from large unit cell dimension crystals. PX-1 has a cryogenically cooled channel cut crystal monochromator and a Kirkpatrick-Baez pair of bi-morph mirrors; the source is a U20 in-vacuum undulator.

- Energy: 5-15 keV

- Beam size: minimum 100x100 μm^2, maximum 250x250 μm^2

- Sample environment: 100 K cryostream, 3-circle kappa goniostat

- Detector: Pilatus 6M (Dectris)

- Software: MXCuBE2, Albula

o PX-2A is a microfocus beamline dedicated to structural biology experiments. As previously mentioned, the small beam allows the recovery of good dataset, even from small and weakly diffracting crystals. PX-2A features a cryogenically cooled channel cut crystal monochromator, a convex pre-focussing mirror and a Kirkpatrick-Baez pair of focussing mirrors; the source of the X-rays are two U24 canted undulators (4.5 mrad).

- Energy: 6-18 keV

- Beam size: 10.0x5.0 μm^2 (microbeam)

- Sample environment: 100 K cryostream, high-capacity sample-changing robot CATS-IRELEC (Irelec, Saint-Martin-d'Hères, France), high-performance micro-diffractometer MD2

- Detector: EIGER X 9M (Dectris)

- Software: MXCuBE2, Albula

Overall, more than 100 crystals were analysed; among them, 36 were deemed of sufficient quality to collect a full set of data (Table 6). The acquisition parameters for the characterisation of the samples (energy, resolution, transmittance, exposure time, oscillation range, number of images) were varied each time, depending on the radiation source and on the crystal itself. Mesh scans were performed to aid the X-ray centring procedure, when necessary. Diffraction plans for data collection were generated automatically at ESRF, and manually through XDSME at SOLEIL.[59] Helical scans were sometimes used, if available and when deemed opportune. The table below shows a list of all such crystals, with concise information regarding each of them.

Data collected at ESRF were automatically processed through ISPyB,[60] a software that features 5 different pipelines, namely:

- Grenoble Automatic Data procEssing (GrenADES);

- EDNA framework Fast Processing System (EDNA Autoprocessing);

- autoPROC;

- XDSAPP;

- XIA2_DIALS.[61]

Each of them independently processes the data, so that the best output can be chosen; they also automatically provide a data collection strategy, which can be used to acquire a complete set of diffraction images. At SOLEIL, these procedures were performed manually, using the XDS-based software XDSME.[59]

After the data indexing and the definition of the space group, the structures were solved through the Molecular Replacement (MR) method, using as models the available MbtI crystals deposited in the Protein Data Bank. The template structures were chosen based on their similarity to the crystal, in terms of cell dimensions and space group. The procedure was performed using Phaser, available in the CCP4 or Phenix suites.[62–64]

The available datasets were scrutinised and then selected to be processed further, based on the resolution and the presence of bound ligands. The graphical visualisation of the *.pdb* and *.mtz* files, and the refitting and rebuilding processes were aided by Coot.[65] The best structures were subjected to multiple cycles of refinement using the software Buster (Global Phasing Ltd.);[44] the molecules were treated as rigid bodies in the first cycles and then normally in the following steps. The options autoncs, TLS and Gelly, with the protgeo_option_chiralrestraint_from_equilib.dat feature, were employed. The restraint dictionaries for the ligands were obtained from Grade (Global Phasing Ltd.) and integrated in the refinement process.[66]

The structure validation was performed with MolProbity and Phenix suite.[64,67] The atomic coordinates will be deposited in the PDB, once the experimental data assembly is completed. The graphical representations were realised with Pymol.[68]

Table 6. Detailed overview of the datasets acquired for the MbtI crystals.

Entry	Conditions		Cell dimensions (Å, °)	Space group	Resolution (Å)	Source
1	*Ligand*	**1**	a: 51.810 b: 164.046 c: 188.047	P2$_1$2$_1$2$_1$	3.48	PX-2A
	Reservoir	15% PEG 4000, 0.2 M imidazole malate pH 6.0, 10 mM MgCl$_2$				
	Cryoprot.	25% glycerol				
2	*Ligand*	**1**	a: 51.852 b: 164.161 c: 195.148	P2$_1$2$_1$2$_1$	2.69	PX-2A
	Reservoir	11% PEG 4000, 0.2 M imidazole malate pH 6.0				
	Cryoprot.	25% glycerol				
3	*Ligand*	**1**	a: 52.037 b: 164.330 c: 195.757	P2$_1$2$_1$2$_1$	2.59	PX-2A
	Reservoir	12% PEG 4000, 0.2 M imidazole malate pH 6.0				
	Cryoprot.	25% glycerol				
4	*Ligand*	**1**	a: 52.056 b: 164.447 c: 195.781	P2$_1$2$_1$2$_1$	3.0	ID23-1
	Reservoir	20% PEG 3350 0.2 M NH$_4^+$ citrate				
	Cryoprot.	25% glycerol				
5	*Ligand*	**1**	a: 52.148 b: 164.167 c: 196.245	P2$_1$2$_1$2$_1$	2.73	ID23-1
	Reservoir	20% PEG 3350 0.2 M NH$_4^+$citrate				
	Cryoprot.	25% ethylene glycol				
6	*Ligand*	Chorismic acid MgCl$_2$	a: 53.112 b: 88.521 c: 100.413 β: 103.124	P2$_1$	1.98	PX-2A
	Reservoir	30% MPD 20% ethanol				
	Cryoprot.	25% glycerol				
7	*Ligand*	Chorismic acid	a: 87.408 b: 115.367 c: 93.939 β: 91.468	P2$_1$	2.96	PX-2A
	Reservoir	20% PEG 3000 0.1 M Na$^+$ citrate pH 5.5				
	Cryoprot.	25% glycerol				
8	*Ligand*	Chorismic acid	a: 88.040 b: 116.900 c: 94.090 β: 91.600	P2$_1$	1.80	PX-2A
	Reservoir	20% PEG 3350 0.2M Na$_2$H$_2$PO$_4$				
	Cryoprot.	oil				

Entry	Conditions		Cell dimensions (Å, °)	Space group	Resolution (Å)	Source
9	*Ligand*	Chorismic acid	a: 88.129 b: 116.749 c: 94.313 β: 91.477	P2$_1$	2.03	PX-2A
	Reservoir	20% PEG 3350 0.2M Na$_2$H$_2$PO$_4$				
	Cryoprot.	25% glycerol				
10	*Ligand*	**54**	a: 88.379 b: 116.318 c: 94.751 β: 91.551	P2$_1$	2.25	PX-2A
	Reservoir	20% PEG 3350 0.2 M Na$_2$SO$_4$				
	Cryoprot.	25% glycerol				
11	*Ligand*	Chorismic acid	a: 88.043 b: 116.899 c: 94.088 β: 91.600	P2$_1$	1.80	PX-2A
	Reservoir	16% PEG 8000 0.04 M KH$_2$PO$_4$ 20% glycerol				
	Cryoprot.	oil				
12	*Ligand*	**27**, MgCl$_2$	a: 90.676 b: 92.578 c: 129.370	P2$_1$2$_1$2$_1$	2.54	PX-2A
	Reservoir	30% MPD 0.02 M CaCl$_2$ 0.1 M acetate buffer pH 4.6				
	Cryoprot.	oil				
13	*Ligand*	**27**	a: 87.787 b: 116.269 c: 93.819 β: 91.731	P2$_1$	1.89	PX-1
	Reservoir	20% PEG 3350 0.2 M NH$_4$H$_2$PO$_4$				
	Cryoprot.	oil				
14	*Ligand*	**27**	a: 87.905 b: 115.511 c: 94.255 β: 91.106	P2$_1$	2.57	PX-1
	Reservoir	20% PEG 3350 0.2 M NH$_4$H$_2$PO$_4$				
	Cryoprot.	25% glycerol				
15	*Ligand*	**27**	a: 87.875 b: 115.419 c: 94.126 β: 91.561	P2$_1$	2.65	PX-1
	Reservoir	20% PEG 3350 0.2 M NH$_4$H$_2$PO$_4$				
	Cryoprot.	25% glycerol				
16	*Ligand*	**54**	a: 87.896 b: 114.884 c: 94.544 β: 91.499	P2$_1$	2.14	PX-2A
	Reservoir	20% PEG 3350 0.2 M Na$_2$SO$_4$				
	Cryoprot.	25% glycerol				
17	*Ligand*	**54**	a: 51.983 b: 164.204 c: 196.757	P2$_1$2$_1$2$_1$	2.21	PX-2A
	Reservoir	20% PEG 3350 0.2 M NH$_4^+$ citrate				
	Cryoprot.	25% glycerol				
18	*Ligand*	**54**	a: 52.109 b: 162.786 c: 195.933	P2$_1$2$_1$2$_1$	2.42	PX-2A
	Reservoir	20% PEG 3350 0.2 M NH$_4^+$ citrate				

Entry	Conditions		Cell dimensions (Å, °)	Space group	Resolution (Å)	Source
	Cryoprot.	oil				
19	*Ligand*	**54**	*a*: 51.913 *b*: 163.412 *c*: 196.078	P2$_1$2$_1$2$_1$	2.74	PX-2A
	Reservoir	20% PEG 3350 0.2 M NH$_4^+$ citrate				
	Cryoprot.	25% glycerol				
20	*Ligand*	**1**, MgCl$_2$	*a*: 51.334 *b*: 98.357 *c*: 168.107	P2$_1$2$_1$2$_1$	3.16	PX-2A
	Reservoir	30% PEG MME 5000 0.2 M (NH$_4$)$_2$SO$_4$ 0.1 M MES pH 6.5				
	Cryoprot.	25% glycerol				
21	*Ligand*	**54**, MgCl$_2$	*a*: 88.080 *b*: 116.485 *c*: 94.449 *β*: 91.632	P2$_1$	2.28	PX-2A
	Reservoir	30% PEG 4000 0.1 M (NH$_4$)$_2$SO$_4$ 0.1 M Na$^+$ citrate pH 5.6				
	Cryoprot.	25% glycerol				
22	*Ligand*	**54**	*a*: 88.304 *b*: 116.339 *c*: 94.805 *β*: 91.600	P2$_1$	2.15	PX-1
	Reservoir	24% PEG 3350 0.150 M Na$_2$SO$_4$				
	Cryoprot.	oil				
23	*Ligand*	Methyl-AMT	*a*: 95.727 *b*: 83.356 *c*: 126.296 *β*: 111.800	P2$_1$	2.77	PX-1
	Reservoir	20% PEG 3350 0.2 M Li$_2$SO$_4$				
	Cryoprot.	oil				
24	*Ligand*	Methyl-AMT	*a*: 86.240 *b*: 107.728 *c*: 94.494 *β*: 92.934	P2$_1$	2.17	PX-2A
	Reservoir	20% PEG 3350 8% Tacsimate				
	Cryoprot.	25% glycerol				
25	*Ligand*	Methyl-AMT	*a*: 86.815 *b*: 109.088 *c*: 95.041 *β*: 93.356	P2$_1$	2.02	PX-2A
	Reservoir	20% PEG 3350 8% Tacsimate				
	Cryoprot.	oil				
26	*Ligand*	Methyl-AMT	*a*: 95.607 *b*: 83.064 *c*: 126.252 *β*: 111.670	P2$_1$	1.65	PX-2A
	Reservoir	16% PEG 3350 0.1 M Na$^+$ citrate pH 5.6 2% Tacsimate				
	Cryoprot.	25% glycerol				
27	*Ligand*	Methyl-AMT	*a*: 95.441 *b*: 83.600 *c*: 127.758 *β*: 111.770	P2$_1$	2.68	PX-1
	Reservoir	20% PEG 3350 0.2 M Li$_2$SO$_4$				
	Cryoprot.	oil				

Entry	Conditions		Cell dimensions (Å, °)	Space group	Resolution (Å)	Source
28	*Ligand*	**54**, MgCl$_2$	*a*: 88.336 *b*: 111.687 *c*: 94.998 *β*: 92.670	P2$_1$	2.09	PX-1
	Reservoir	20% PEG 3350 0.2 M Na$^+$ tartrate				
	Cryoprot.	oil				
29	*Ligand*	**54**, BaCl$_2$	*a*: 87.968 *b*: 115.182 *c*: 94.938 *β*: 91.380	P2$_1$	2.23	PX-1
	Reservoir	20% PEG 3350 0.2 M Na$^+$ tartrate				
	Cryoprot.	25% glycerol				
30	*Ligand*	**54**, CaCl$_2$	*a*: 53.105 *b*: 88.728 *c*: 100.199 *β*: 102.867	P2$_1$	2.12	PX-2A
	Reservoir	10% PEG 4000 0.1 M imidazole pH 8.0 30% MPD				
	Cryoprot.	25% glycerol				
31	*Ligand*	**54**, CaCl$_2$	*a*: 52.911 *b*: 87.719 *c*: 99.960 *β*: 103.323	P2$_1$	2.03	PX-2A
	Reservoir	10% PEG 4000 0.1 M imidazole pH 8.0 30% MPD				
	Cryoprot.	oil				
32	*Ligand*	**54**, CaCl$_2$	*a*: 88.061 *b*: 115.576 *c*: 95.149 *β*: 91.262	P2$_1$	2.89	PX-2A
	Reservoir	20% PEG 3350 0.2 M Na$^+$ tartrate				
	Cryoprot.	oil				
33	*Ligand*	**53**	*a* = *b*: 195.493 *c*: 259.446	I422	2.83	PX-2A
	Reservoir	0.1 M MES pH 6.5 1.75 M MgSO$_4$				
	Cryoprot.	25% glycerol				
34	*Ligand*	**54**	*a* = *b*: 195.213 *c*: 258.483	I422	3.19	PX-2A
	Reservoir	0.1 M MES pH 6.75 1.75 M MgSO$_4$				
	Cryoprot.	25% glycerol				
35	*Ligand*	/	*a* = *b*: 193.970 *c*: 257.030	I422	2.11	PX-2A
	Reservoir	0.1 M MES pH 6.5 1.75 M MgSO$_4$				
	Cryoprot.	oil				
36	*Ligand*	/	*a* = *b*: 196.169 *c*: 259.639	I422	2.63	PX-2A
	Reservoir	1.75 M MgSO$_4$				
	Cryoprot.	25% glycerol				

Table 7. Data collection and refinement statistics for the structures of MbtI with Methyl-AMT, **54** and Mg^{2+}. Statistics for the highest-resolution shell are shown in parentheses.

	MbtI-Methyl-AMT	MbtI-54	MbtI-Mg^{2+}
Space group	P2$_1$	P2$_1$	I422
Unit-cell parameters			
a, *b*, *c* (Å)	95.607, 83.064, 126.252	88.336, 111.687, 94.998	193.970, 193.970, 257.030
α, *β*, *γ* (°)	90.000, 111.670, 90.000	90.000, 92.670, 90.000	90.000, 90.000, 90.000
No. molecules in asymmetric unit	4	4	4
Resolution range (Å)	60.84 - 1.65 (1.71 – 1.65)	30.23 – 2.09 (2.17 – 2.09)	30.67 – 2.11 (2.19 – 2.11)
Wavelength (Å)	0.98013	0.97856	0.98012
No. measured reflections	435560 (42891)	216951 (21645)	261679 (25208)
No. unique reflections	217964 (21514)	108508 (10816)	130889 (12603)
Multiplicity	6.9	7.0	23.8
Completeness (%)	98.86 (97.90)	99.87 (99.92)	93.91 (91.33)
Average I/σ(I)	13.85 (1.31)	8.36 (1.47)	18.11 (1.61)
R$_{merge}$	0.0224 (0.4271)	0.0545 (0.3653)	0.0225 (0.4253)
R$_{work}$	0.189 (0.278)	0.210 (0.265)	0.184 (0.264)
R$_{free}$	0.213 (0.294)	0.238 (0.278)	0.208 (0.296)
No. of non-H atoms	14394	13304	14300
Macromolecule	12852	12309	13161
Ligand	100	64	192
Water	1442	931	947
Average *B*-factors	38.32	43.94	54.3
Macromolecule	37.32	43.71	53.60
Ligand	37.95	60.78	91.00
Water	47.30	45.80	57.20
Rms deviations			
Bonds (Å)	0.0129	0.0127	0.0138
Angles (°)	1.55	1.59	1.67
Clashscore	2.25	2.59	2.84
Ramachandran outliers	0.06%	0.00%	0.00%
Ramachandran favoured	99.17%	99.08%	98.11%
Rotamer outliers	1.98%	1.47%	1.44%
C-beta outliers	0.00%	0.00%	0.06%

Bibliography

(1)   Barberis, I.; Bragazzi, N. L.; Galluzzo, L.; Martini, M. The History of Tuberculosis: From the First Historical Records to the Isolation of Koch's Bacillus. *J. Prev. Med. Hyg.* **2017**, *58* (1), E9–E12.

(2)   Organization, W. H. *Global Tuberculosis Report 2019*; World Health Organization: Geneva PP - Geneva.

(3)   Jena, B.; Ludam, R.; Chhotray, P.; Sahu, M. C. Detection of Mycobacterium Tuberculosis with Conventional Microscopic and Culture Methods. *Int. J. Appl. Res.* **2017**, *3* (12), 143–146.

(4)   Sanou, A.; Bañuls, A.-L.; Van Anh, N. T.; Godreuil, S. Mycobacterium Tuberculosis: Ecology and Evolution of a Human Bacterium. *J. Med. Microbiol.* **2015**, *64* (11), 1261–1269. https://doi.org/10.1099/jmm.0.000171.

(5)   Talip, B. A.; Sleator, R. D.; Lowery, C. J.; Dooley, J. S. G.; Snelling, W. J. An Update on Global Tuberculosis (TB). *Infect. Dis. (Auckl).* **2013**, *6*, 39–50. https://doi.org/10.4137/IDRT.S11263.

(6)   Lawn, S. D.; Zumla, A. I. Tuberculosis. *Lancet* **2011**, *378* (9785), 57–72. https://doi.org/https://doi.org/10.1016/S0140-6736(10)62173-3.

(7)   Bailo, R.; Bhatt, A.; Ainsa, J. A. Lipid Transport in Mycobacterium Tuberculosis and Its Implications in Virulence and Drug Development. *Biochem. Pharmacol.* **2015**, *96* (3), 159–167. https://doi.org/10.1016/J.BCP.2015.05.001.

(8)   Riley, L. W. Of Mice, Men, and Elephants: Mycobacterium Tuberculosis Cell Envelope Lipids and Pathogenesis. *J. Clin. Invest.* **2006**, *116* (6), 1475–1478. https://doi.org/10.1172/JCI28734.

(9)   Nicas, M.; Nazaroff, W. W.; Hubbard, A. Toward Understanding the Risk of Secondary Airborne Infection: Emission of Respirable Pathogens. *J. Occup. Environ. Hyg.* **2005**, *2* (3), 143–154.

https://doi.org/10.1080/15459620590918466.

(10) Kumar, V.; Robbins, S. L. *Robbins Basic Pathology*; Saunders/Elsevier: Philadelphia, PA, 2007.

(11) Nahid, P.; Dorman, S. E.; Alipanah, N.; Barry, P. M.; Brozek, J. L.; Cattamanchi, A.; Chaisson, L. H.; Chaisson, R. E.; Daley, C. L.; Grzemska, M.; et al. Official American Thoracic Society/Centers for Disease Control and Prevention/Infectious Diseases Society of America Clinical Practice Guidelines: Treatment of Drug-Susceptible Tuberculosis. *Clin. Infect. Dis.* **2016**, *63* (7), e147–e195. https://doi.org/10.1093/cid/ciw376.

(12) Palomino, J. C.; Martin, A. Drug Resistance Mechanisms in Mycobacterium Tuberculosis. *Antibiot. (Basel, Switzerland)* **2014**, *3* (3), 317–340. https://doi.org/10.3390/antibiotics3030317.

(13) Kaufmann, S. H. E.; Weiner, J.; von Reyn, C. F. Novel Approaches to Tuberculosis Vaccine Development. *Int. J. Infect. Dis.* **2017**, *56*, 263–267. https://doi.org/10.1016/J.IJID.2016.10.018.

(14) Gualano, G.; Capone, S.; Matteelli, A.; Palmieri, F. New Antituberculosis Drugs: From Clinical Trial to Programmatic Use. *Infect. Dis. Rep.* **2016**, *8* (2), 6569. https://doi.org/10.4081/idr.2016.6569.

(15) Chao, A.; Sieminski, P. J.; Owens, C. P.; Goulding, C. W. Iron Acquisition in *Mycobacterium Tuberculosis*. *Chem. Rev.* **2019**, *119* (2), 1193–1220. https://doi.org/10.1021/acs.chemrev.8b00285.

(16) Hameed, S.; Pal, R.; Fatima, Z. Iron Acquisition Mechanisms: Promising Target Against Mycobacterium Tuberculosis. *Open Microbiol. J.* **2015**, *9* (1), 91–97. https://doi.org/10.2174/1874285801509010091.

(17) Malhotra, H.; Patidar, A.; Boradia, V. M.; Kumar, R.; Nimbalkar, R. D.; Kumar, A.; Gani, Z.; Kaur, R.; Garg, P.; Raje, M.; et al. Mycobacterium Tuberculosis Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) Functions as a Receptor for Human Lactoferrin. *Front. Cell. Infect. Microbiol.* **2017**, *7*, 245. https://doi.org/10.3389/fcimb.2017.00245.

(18) Serafini, A.; Pisu, D.; Palù, G.; Rodriguez, G. M.; Manganelli, R. The ESX-3

Secretion System Is Necessary for Iron and Zinc Homeostasis in Mycobacterium Tuberculosis. *PLoS One* **2013**, *8* (10), e78351. https://doi.org/10.1371/journal.pone.0078351.

(19) Tufariello, J. M.; Chapman, J. R.; Kerantzas, C. A.; Wong, K.-W.; Vilchèze, C.; Jones, C. M.; Cole, L. E.; Tinaztepe, E.; Thompson, V.; Fenyö, D.; et al. Separable Roles for Mycobacterium Tuberculosis ESX-3 Effectors in Iron Acquisition and Virulence. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (3), E348-57. https://doi.org/10.1073/pnas.1523321113.

(20) Matsui, T.; Nambu, S.; Ono, Y.; Goulding, C. W.; Tsumoto, K.; Ikeda-Saito, M. Heme Degradation by *Staphylococcus Aureus* IsdG and IsdI Liberates Formaldehyde Rather Than Carbon Monoxide. *Biochemistry* **2013**, *52* (18), 3025–3027. https://doi.org/10.1021/bi400382p.

(21) Meneghetti, F.; Villa, S.; Gelain, A.; Barlocco, D.; Roberto Chiarelli, L.; Rosalia Pasca, M.; Costantino, L. Iron Acquisition Pathways as Targets for Antitubercular Drugs. *Curr. Med. Chem.* **2016**, *23* (35), 4009–4026. https://doi.org/10.2174/0929867323666160607223747.

(22) Harrison, A. J.; Yu, M.; Gårdenborg, T.; Middleditch, M.; Ramsay, R. J.; Baker, E. N.; Lott, J. S. The Structure of MbtI from Mycobacterium Tuberculosis, the First Enzyme in the Biosynthesis of the Siderophore Mycobactin, Reveals It to Be a Salicylate Synthase. *J. Bacteriol.* **2006**, *188* (17), 6081–6091. https://doi.org/10.1128/JB.00338-06.

(23) Kerbarh, O.; Chirgadze, D. Y.; Blundell, T. L.; Abell, C. Crystal Structures of Yersinia Enterocolitica Salicylate Synthase and Its Complex with the Reaction Products Salicylate and Pyruvate. *J. Mol. Biol.* **2006**, *357* (2), 524–534. https://doi.org/10.1016/J.JMB.2005.12.078.

(24) Manos-Turvey, A.; Bulloch, E. M. M.; Rutledge, P. J.; Baker, E. N.; Lott, J. S.; Payne, R. J. Inhibition Studies of Mycobacterium Tuberculosis Salicylate Synthase (MbtI). *ChemMedChem* **2010**, *5* (7), 1067–1079. https://doi.org/10.1002/cmdc.201000137.

(25) Ferrer, S.; Martí, S.; Moliner, V.; Tuñón, I.; Bertrán, J. Understanding the Different Activities of Highly Promiscuous MbtI by Computational

Methods. *Phys. Chem. Chem. Phys.* **2012**, *14* (10), 3482. https://doi.org/10.1039/c2cp23149b.

(26) Ferrer, S.; Martí, S.; Andrés, J.; Moliner, V.; Tuñón, I.; Bertrán, J. Molecular Mechanism of Chorismate Mutase Activity of Promiscuos MbtI. *Theor. Chem. Acc.* **2011**, *128* (4–6), 601–607. https://doi.org/10.1007/s00214-010-0773-z.

(27) Ziebart, K. T.; Toney, M. D. Nucleophile Specificity in Anthranilate Synthase, Aminodeoxychorismate Synthase, Isochorismate Synthase, and Salicylate Synthase. *Biochemistry* **2010**, *49* (13), 2851–2859. https://doi.org/10.1021/bi100021x.

(28) Chi, G.; Manos-Turvey, A.; O'Connor, P. D.; Johnston, J. M.; Evans, G. L.; Baker, E. N.; Payne, R. J.; Lott, J. S.; Bulloch, E. M. M. Implications of Binding Mode and Active Site Flexibility for Inhibitor Potency against the Salicylate Synthase from Mycobacterium Tuberculosis. *Biochemistry* **2012**, *51* (24), 4868–4879. https://doi.org/10.1021/bi3002067.

(29) Kozlowski, M. C.; Tom, N. J.; Seto, C. T.; Sefler, A. M.; Bartlett, P. A. Chorismate-Utilizing Enzymes Isochorismate Synthase, Anthranilate Synthase, and p-Aminobenzoate Synthase: Mechanistic Insight through Inhibitor Design. *J. Am. Chem. Soc.* **1995**, *117* (8), 2128–2140. https://doi.org/10.1021/ja00113a002.

(30) Liu, Z.; Liu, F.; Aldrich, C. C. Stereocontrolled Synthesis of a Potential Transition-State Inhibitor of the Salicylate Synthase MbtI from *Mycobacterium Tuberculosis. J. Org. Chem.* **2015**, *80* (13), 6545–6552. https://doi.org/10.1021/acs.joc.5b00455.

(31) Zhang, X.-K.; Liu, F.; Fiers, W. D.; Sun, W.-M.; Guo, J.; Liu, Z.; Aldrich, C. C. Synthesis of Transition-State Inhibitors of Chorismate Utilizing Enzymes from Bromobenzene *Cis* -1,2-Dihydrodiol. *J. Org. Chem.* **2017**, *82* (7), 3432–3440. https://doi.org/10.1021/acs.joc.6b02801.

(32) Vasan, M.; Neres, J.; Williams, J.; Wilson, D. J.; Teitelbaum, A. M.; Remmel, R. P.; Aldrich, C. C. Inhibitors of the Salicylate Synthase (MbtI) from Mycobacterium Tuberculosis Discovered by High-Throughput Screening.

ChemMedChem **2010**, *5* (12), 2079–2087. https://doi.org/10.1002/cmdc.201000275.

(33) Chiarelli, L. R.; Mori, M.; Barlocco, D.; Beretta, G.; Gelain, A.; Pini, E.; Porcino, M.; Mori, G.; Stelitano, G.; Costantino, L.; et al. Discovery and Development of Novel Salicylate Synthase (MbtI) Furanic Inhibitors as Antitubercular Agents. *Eur. J. Med. Chem.* **2018**, *155*. https://doi.org/10.1016/j.ejmech.2018.06.033.

(34) Gerhard Wolber*, † and; Langer‡, T. LigandScout: 3-D Pharmacophores Derived from Protein-Bound Ligands and Their Use as Virtual Screening Filters. **2004**. https://doi.org/10.1021/CI049885E.

(35) Tuccinardi, T.; Poli, G.; Romboli, V.; Giordano, A.; Martinelli, A. Extensive Consensus Docking Evaluation for Ligand Pose Prediction and Virtual Screening Studies. *J. Chem. Inf. Model.* **2014**, *54* (10), 2980–2986. https://doi.org/10.1021/ci500424n.

(36) Sritharan, M. Iron Homeostasis in Mycobacterium Tuberculosis: Mechanistic Insights into Siderophore-Mediated Iron Uptake. *J. Bacteriol.* **2016**, *198* (18), 2399–2409. https://doi.org/10.1128/JB.00359-16.

(37) Schwyn, B.; Neilands, J. B. Universal Chemical Assay for the Detection and Determination of Siderophores. *Anal. Biochem.* **1987**, *160* (1), 47–56. https://doi.org/10.1016/0003-2697(87)90612-9.

(38) Chiarelli, L. R.; Mori, M.; Beretta, G.; Gelain, A.; Pini, E.; Sammartino, J. C.; Stelitano, G.; Barlocco, D.; Costantino, L.; Lapillo, M.; et al. New Insight into Structure-Activity of Furan-Based Salicylate Synthase (MbtI) Inhibitors as Potential Antitubercular Agents. *J. Enzyme Inhib. Med. Chem.* **2019**, *34* (1), 823–828. https://doi.org/10.1080/14756366.2019.1589462.

(39) Wu, X.; Geng, X.; Zhao, P.; Zhang, J.; Wu, Y.; Wu, A. I $_2$ -Promoted Formal [3+2] Cycloaddition of α-Methylenyl Isocyanides with Methyl Ketones: A Route to 2,5-Disubstituted Oxazoles. *Chem. Commun.* **2017**, *53* (24), 3438–3441. https://doi.org/10.1039/C6CC10275A.

(40) Kadam, K. S.; Jadhav, R. D.; Kandre, S.; Guha, T.; Reddy, M. M. K.; Brahma, M. K.; Deshmukh, N. J.; Dixit, A.; Doshi, L.; Srinivasan, S.; et al. Evaluation of

Thiazole Containing Biaryl Analogs as Diacylglycerol Acyltransferase 1 (DGAT1) Inhibitors. *Eur. J. Med. Chem.* **2013**, *65*, 337–347. https://doi.org/10.1016/J.EJMECH.2013.05.006.

(41) He, G.-X.; Yuan, J.-M.; Zhu, H.-M.; Wei, K.; Wang, L.-Y.; Kong, S.-L.; Mo, D.-L.; Pan, C.-X.; Su, G.-F. Synthesis and Antitumor Evaluation of 2,3-Diarylbenzofuran Derivatives on HeLa Cells. *Bioorg. Med. Chem. Lett.* **2017**, *27* (8), 1660–1664. https://doi.org/10.1016/J.BMCL.2017.03.010.

(42) Curreli, F.; Kwon, Y. Do; Belov, D. S.; Ramesh, R. R.; Kurkin, A. V.; Altieri, A.; Kwong, P. D.; Debnath, A. K. Synthesis, Antiviral Potency, in Vitro ADMET, and X-Ray Structure of Potent CD4 Mimics as Entry Inhibitors That Target the Phe43 Cavity of HIV-1 Gp120. *J. Med. Chem.* **2017**, *60* (7), 3124–3153. https://doi.org/10.1021/acs.jmedchem.7b00179.

(43) Fox, B. G.; Blommel, P. G. Autoinduction of Protein Expression. *Curr. Protoc. protein Sci.* **2009**, *Chapter 5*, Unit-5.23. https://doi.org/10.1002/0471140864.ps0523s56.

(44) Bricogne, G.; Blanc, E.; Brandl, M.; Flensburg, C.; Keller, P.; Paciorek, W.; Roversi, P.; Sharff, A.; Smart, O. S.; Vonrhein, C.; et al. BUSTER Version 2.10.3. Global Phasing Ltd.: Cambridge, UK 2017.

(45) Meneely, K. M.; Sundlov, J. A.; Gulick, A. M.; Moran, G. R.; Lamb, A. L. An Open and Shut Case: The Interaction of Magnesium with MST Enzymes. *J. Am. Chem. Soc.* **2016**, *138* (29), 9277–9293. https://doi.org/10.1021/jacs.6b05134.

(46) Claus, S.; Weiler, C.; Schiewe, J.; Friess, W. How Can We Bring High Drug Doses to the Lung? *Eur. J. Pharm. Biopharm.* **2014**, *86* (1), 1–6. https://doi.org/10.1016/J.EJPB.2013.11.005.

(47) Truzzi, E.; Meneghetti, F.; Mori, M.; Costantino, L.; Iannuccelli, V.; Maretti, E.; Domenici, F.; Castellano, C.; Rogers, S.; Capocefalo, A.; et al. Drugs/Lamellae Interface Influences the Inner Structure of Double-Loaded Liposomes for Inhaled Anti-TB Therapy: An in-Depth Small-Angle Neutron Scattering Investigation. *J. Colloid Interface Sci.* **2019**, *541*, 399–406. https://doi.org/10.1016/J.JCIS.2019.01.094.

(48) Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; Cheatham III, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; et al. AMBER, Version 14. San Francisco, CA 2015.

(49) Li, P.; Roberts, B. P.; Chakravorty, D. K.; Merz, K. M. Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for +2 Metal Cations in Explicit Solvent. *J. Chem. Theory Comput.* **2013**, *9* (6), 2733–2748. https://doi.org/10.1021/ct400146w.

(50) Porta, F.; Gelain, A.; Barlocco, D.; Ferri, N.; Marchianò, S.; Cappello, V.; Basile, L.; Guccione, S.; Meneghetti, F.; Villa, S. A Field-Based Disparity Analysis of New 1,2,5-Oxadiazole Derivatives Endowed with Antiproliferative Activity. *Chem. Biol. Drug Des.* **2017**, *90*, 820–839. https://doi.org/10.1111/cbdd.13003.

(51) Skerlj, R. T.; Bourque, E. M. J.; Lansbury, P. T.; Greenlee, W. J.; Good, A. C. Imidazo [1,5-a]Pyrimidinyl Carboxamide Compounds and Their Use in the Treatment of Medical Disorders. WO2017176961 (A1), 2017.

(52) Kemp, M. I.; Woodrow, M. D. Cyanopyrrolidine Dervivatives as Inhibitors for DUBs. WO2017109488 (A1), 2017.

(53) Cremonesi, S.; Luker, T.; Semeraro, T.; Micheli, F. Spirocyclic Derivatives. WO2016042452 (A1), 2016.

(54) Grisostomi, C.; Kast, P.; Pulido, R.; Huynh, J.; Hilvert, D. Efficient in Vivo Synthesis and Rapid Purification of Chorismic Acid Using an Engineered Escherichia Coli Strain. *Bioorg. Chem.* **1997**, *25* (5–6), 297–305. https://doi.org/10.1006/BIOO.1997.1073.

(55) Lowry, O. H.; Rosebrough, N. J.; Farr, A. L.; Randall, R. J. Protein Measurement with the Folin Phenol Reagent. *J. Biol. Chem.* **1951**, *193*, 265–275.

(56) Jasial, S.; Hu, Y.; Bajorath, J. How Frequently Are Pan-Assay Interference Compounds Active? Large-Scale Analysis of Screening Data Reveals Diverse Activity Profiles, Low Global Hit Frequency, and Many Consistently Inactive Compounds. *J. Med. Chem.* **2017**, *60* (9), 3879–3886. https://doi.org/10.1021/acs.jmedchem.7b00154.

(57) Shoichet, B. K. Screening in a Spirit Haunted World. *Drug Discov. Today* **2006**, *11* (13–14), 607–615. https://doi.org/10.1016/J.DRUDIS.2006.05.014.

(58) Dahlin, J. L.; Nissink, J. W. M.; Strasser, J. M.; Francis, S.; Higgins, L.; Zhou, H.; Zhang, Z.; Walters, M. A. PAINS in the Assay: Chemical Mechanisms of Assay Interference and Promiscuous Enzymatic Inhibition Observed during a Sulfhydryl-Scavenging HTS. *J. Med. Chem.* **2015**, *58* (5), 2091–2113. https://doi.org/10.1021/jm5019093.

(59) Legrand, P. XDSME: XDS Made Easier. 2017.

(60) Delageniere, S.; Brenchereau, P.; Launer, L.; Ashton, A. W.; Leal, R.; Veyrier, S.; Gabadinho, J.; Gordon, E. J.; Jones, S. D.; Levik, K. E.; et al. ISPyB: An Information Management System for Synchrotron Macromolecular Crystallography. *Bioinformatics* **2011**, *27* (22), 3186–3192. https://doi.org/10.1093/bioinformatics/btr535.

(61) Monaco, S.; Gordon, E.; Bowler, M. W.; Delagenière, S.; Guijarro, M.; Spruce, D.; Svensson, O.; McSweeney, S. M.; McCarthy, A. A.; Leonard, G.; et al. Automatic Processing of Macromolecular Crystallography X-Ray Diffraction Data at the ESRF. *J. Appl. Crystallogr.* **2013**, *46* (Pt 3), 804–810. https://doi.org/10.1107/S0021889813006195.

(62) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *Phaser* Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40* (4), 658–674. https://doi.org/10.1107/S0021889807021206.

(63) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; et al. Overview of the *CCP* 4 Suite and Current Developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (4), 235–242. https://doi.org/10.1107/S0907444910045749.

(64) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; et al. *PHENIX* : A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (2), 213–221.

https://doi.org/10.1107/S0907444909052925.

(65) Emsley, P.; Cowtan, K. *Coot* : Model-Building Tools for Molecular Graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60* (12), 2126–2132. https://doi.org/10.1107/S0907444904019158.

(66) Smart, O. S.; Holstein, J.; Womack, T. Grade. 2012.

(67) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C.; IUCr. *MolProbity* : All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (1), 12–21. https://doi.org/10.1107/S0907444909042073.

(68) Schrödinger, L. *The PyMOL Molecular Graphics System, Version 1.8*; 2015.

Scientific CV

- **Publications**

1. Cerchia, C.; Nasso, R.; <u>Mori, M.</u>; Villa, S.; Gelain, A.; Capasso, A.; Aliotta, F.; Simonetti, M.; Rullo, R.; Masullo, M.; De Vendittis, E.; Ruocco, M.R.; Lavecchia, A. Discovery of Novel Naphthylphenylketone and Naphthylphenylamine Derivatives as Cell Division Cycle 25B (CDC25B) Phosphatase Inhibitors: Design, Synthesis, Inhibition Mechanism and in Vitro Efficacy against Melanoma Cell Lines. *J Med Chem.*, **2019**, *62*, 7089-7110.

2. Chiarelli, L. R.; <u>Mori, M.</u>; Beretta, G.; Gelain, A.; Pini, E.; Sanmartino, J. C.; Stelitano, G.; Barlocco, D.; Costantino, L.; Lapillo, M.; Poli, G.; Caligiuri, I.; Rizzolio, F.; Bellinzoni, M.; Tuccinardi, T.; Villa, S.; Meneghetti, F. New insight into structure-activity of furan-based salicylate synthase (MbtI) inhibitors as potential antitubercular agents. *J. Enz. Inhib. Med Chem.*, **2019**, *34*, 823-828.

3. Truzzi, E.; Meneghetti, F.; <u>Mori, M.</u>; Costantino, L.; Iannuccelli, V.; Maretti, E.; Domenici, F.; Castellano, C.; Rogers, S.; Capocefalo, A.; Leo, E. Drugs/lamellae interface influences the inner structure of double-loaded liposomes for inhaled anti-TB therapy: An in-depth small-angle neutron scattering investigation. *J Colloid Interface Sci.*, **2019**, *541*, 399-406.

4. Ehrsam, D.; Porta, F.; <u>Mori, M.</u>; Meyer zu Schwabedissen, H.E.; Dalla Via, L.; Garcia-Argaez, A.N.; Basile, L.; Meneghetti, F.; Villa, S.; Gelain, A. Unravelling the Antiproliferative Activity of 1,2,5-oxadiazole Derivatives. *Anticancer Res.*, **2019**, *39*, 3453-3461.

5. <u>Mori, M.</u>; Sammartino, J.C.; Costantino, L.; Gelain, A.; Meneghetti, F.; Villa, S.; Chiarelli, L.R. An Overview on the Potential Antimycobacterial Agents Targeting Serine/Threonine Protein Kinases from Mycobacterium tuberculosis. *Curr. Top. Med. Chem.*, **2019**, *19*, 646-661.

6. Gelain, A.; <u>Mori, M.</u>; Meneghetti, F.; Porta, F.; Basile, L.; Marverti, G.; Asai, A.; Hyeraci, M.; García-Argáez, A.N.; Dalla Via, L.; Guccione, S.; Villa, S. Exploring the Biological Activity of a Library of 1,2,5-Oxadiazole Derivatives Endowed With Antiproliferative Activity. *Anticancer Res.*, **2019**, *39*, 135-144.

7. Gelain, A.; <u>Mori, M.</u>; Meneghetti, F.; Villa, S. Signal Transducer and Activator of Transcription Protein 3 (STAT3): an Update on its Direct Inhibitors as Promising Anticancer Agents. *Curr. Med. Chem.*, **2019**, *26*, 5165-5206.

8. Pini, E.; Poli, G.; Tuccinardi, T.; Chiarelli, L.R.; <u>Mori, M.</u>; Gelain, A.; Costantino, L.; Villa, S.; Meneghetti, F.; Barlocco, D. New Chromane-Based Derivatives as Inhibitors of Mycobacterium tuberculosis Salicylate Synthase (MbtI): Preliminary Biological Evaluation and Molecular Modeling Studies. *Molecules*, **2018**, *23*, E1506.

9. Rossetti, A.; Landoni, S.; Meneghetti, F.; Castellano, C.; <u>Mori, M.</u>; Colombo Dugoni, G.; Sacchetti, A. Application of chiral bi- and tetra-dentate bispidine-derived ligands in the copper(II)-catalyzed asymmetric Henry reaction. *New J. Chem.*, **2018**, *42*, 12072-12081.

10. Chiarelli, L.R.; <u>Mori, M.</u>; Barlocco, D.; Beretta, G.; Gelain, A.; Pini, E.; Porcino, M.; Mori, G.; Stelitano, G.; Costantino, L.; Lapillo, M.; Bonanni, D.; Poli, G.; Tuccinardi, T.; Villa, S.; Meneghetti, F. Discovery and Development of Novel Salicylate Synthase (MbtI) Furanic Inhibitors as Antitubercular Agents. *Eur. J. Med. Chem.*, **2018**, *155*, 754-763.

11. Ferraboschi, P.; Sala, M.; Stradi, R.; Ragonesi, L.; Gagliardi, C.; Lanzarotti, P.; Ragg, E.M.; <u>Mori, M.</u>; Meneghetti, F. Full spectroscopic characterization of two crystal pseudopolymorphic forms of the antiandrogen cortexolone 17α-propionate for topic application. *Steroids*, **2017**, *128*, 95-104.

---

12. Legnani, L.; Colombo, D.; Venuti, A.; Pastori, C.; Lopalco, L.; Toma, L.; <u>Mori, M.</u>; Grazioso, G.; Villa, S. Diazabicyclo analogues of maraviroc: synthesis, modeling, NMR studies and antiviral activity. *MedChemComm*, **2016**, *8*, 422-433.

- **Oral presentations**

1. <u>Mori, M.</u>; Villa, S.; Chiarelli, L. R.; Stelitano, G.; Gelain, A.; Tuccinardi, T.; Bellinzoni, M.; Meneghetti, F. Structural insights for the development of novel antituberculars targeting the iron uptake pathway. XXVI National Meeting in Medicinal Chemistry and XII Young Medicinal Chemists' Symposium, Milano, 2019.

2. <u>Mori, M.</u>; Villa, S.; Chiarelli, L.R.; Gelain, A.; Costantino, L.; Tuccinardi, T.; Lapillo, M.; Meneghetti, F. Design and synthesis of novel furan-based MbtI inhibitors as potential antitubercular agents. Proceedings of the Merck Young Chemists Symposium, Milano Marittima, 2017.

- **Posters**

1. <u>Mori, M.</u>; Chiarelli, L. R.; Stelitano, G.; Gelain, A.; Tuccinardi, T.; Bellinzoni, M.; Meneghetti, F.; Villa, S. Exploring the role of the heterocyclic moiety of 5-(4-nitrophenyl)-2-furoic acid, a promising inhibitor of MbtI from *Mycobacterium tuberculosis*. XXVI National Meeting in Medicinal Chemistry and XII Young Medicinal Chemists' Symposium, Milano, 2019.

2. <u>Mori, M.</u>; Villa, S.; Chiarelli, L. R.; Stelitano, G.; Gelain, A.; Tuccinardi, T.; Bellinzoni, M.; Meneghetti, F. Multi-target antitubercular drugs: identification of novel dual MbtI and MptpB inhibitors. XXVI National Meeting in Medicinal Chemistry and XII Young Medicinal Chemists' Symposium, Milano, 2019.

3. Villa, S.; <u>Mori, M.</u>; Meneghetti, F.; Pedretti, A.; Regazzoni, L.; Colombo, D.; Asai, A., Gelain, A. 5,6-Dimethyl-1,3-dihydrobenzo[*c*][1,2,5]thiadiazole-2,2-dioxide as new STAT3 inhibitor. XXVI National Meeting in Medicinal Chemistry and XII Young Medicinal Chemists' Symposium, Milano, 2019.

4. Gelain, A.; <u>Mori, M.</u>; Meneghetti, F.; Asai, A.; Villa, S. Unravelling the interaction between an antiproliferative 1,2,5-oxadiazole derivative and STAT3. XXXIX Convegno Nazionale della Divisione di Chimica Organica della Società Chimica Italiana, Torino, 2019.

5. Gelain, A.; <u>Mori, M.</u>; Villa, S.; Chiarelli, L.R.; Meneghetti, F. Design, synthesis and biological evaluation of novel antitubercular agents targeting the Salicylate Synthase (MbtI). Convegno Nazionale della divisione di Chimica Organica della Società Chimica Italiana, Milano, 2018.

6. Villa, S.; Chiarelli, L.R.; <u>Mori, M.</u>; Barlocco, D.; Gelain, A.; Pini, E.; Mori, G.; Costantino, L.; Tuccinardi, T.; Meneghetti, F. Design, synthesis and biological evaluation of novel antitubercular agents targeting the Salicylate Synthase (MbtI). National Meeting on Medicinal Chemistry, Palermo, 2018.

7. <u>Mori, M.</u>; Villa, S.; Gelain, A.; Beretta, G.; Chiarelli, L.R.; Costantino, L.; Leo, E.; Iannuccelli, V.; Tuccinardi, T.; Castellano, C.; Domenici, F.; Meneghetti, F. New synergic treatments for inhaled anti-TB therapy: a preliminary SANS study. Recent Development in Pharmaceutical Analysis, Rimini, 2017.

8. Meneghetti, F.; Castellano, C.; <u>Mori, M.</u>; Leo, E.; Iannuccelli, V.; Costantino, L.; Domenici F. SANS study of lipid-based nanocarriers as inhaled drug delivery systems for anti-tuberculosis chemotherapy. 2nd Symposium on Medicinal Chemistry for Global Health, 2017.

9. Villa, S.; Chiarelli, L.R.; <u>Mori, M.</u>; Gelain, A.; Costantino, L.; Bonanni, D.; Tuccinardi, T.; Meneghetti, F. Design and synthesis of novel antituberculars targeting the salicylate synthase (MbtI). 2nd Symposium on Medicinal Chemistry for Global Health, 2017.

10. Gelain, A.; Villa, S.; <u>Mori, M.</u>; Chiarelli, L.R.; Costantino, L.; Tuccinardi, T.; Meneghetti, F. Synthesis of Novel Furan-based Antituberculars Potentially Targeting the Methionine Aminopeptidases (MtMetAP1a). 6th World Congress on Medicinal Chemistry and Drug Design, 2017.

11. Villa, S.; Gelain, A.; Barlocco, D.; Pini, E.; Beretta, G.; <u>Mori, M.</u>; Costantino, L.; Chiarelli, L.R.; Pasca, M.R.; Bonanni, D.; Tuccinardi, T.; Meneghetti, F. New inhibitors of salicylate synthase (MbtI) from *M. tuberculosis*: a promising strategy to antitubercular drugs. XXIV National Meeting in Medicinal Chemistry, 2016.