



**Cardiff School of Pharmacy and Pharmaceutical Science,
Cardiff University**

**Identification of novel anti-RSV and anti-
enterovirus inhibitors by computer-aided
drug design**

Philosophiae Doctor in Cardiff University

June 2019

Roberto Manganaro

Supervisor: **Prof Andrea Brancale**

DON' T BE AFRAID TO CROSS THE LINE.

STATEMENTS AND DECLARATION

STATEMENT 1 This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD

Signed Roberto Magnum

Date 11/12/2019

STATEMENT 2

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is it being submitted concurrently for any other degree or award (outside of any formal collaboration agreement between the University and a partner organisation)

Signed Roberto Magnum

Date 11/12/2019

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available in the University's Open Access repository (or, where approved, to be available in the University's library and for inter-library loan), and for the title and summary to be made available to outside organisations, subject to the expiry of a University-approved bar on access if applicable.

Signed Roberto Magnum

Date 11/12/2019

DECLARATION

This thesis is the result of my own independent work, except where otherwise stated, and the views expressed are my own. Other sources are acknowledged by explicit references. The thesis has not been edited by a third party beyond what is permitted by Cardiff University's Use of Third Party Editors by Research Degree Students Procedure.

Signed Roberto Magnum

Date 11/12/2019

WORD COUNT 58183

(Excluding summary, acknowledgements, declarations, contents pages, appendices, tables, diagrams and figures, references, bibliography, footnotes and endnotes)

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 642434.

I cannot begin to express my thanks to Prof. Andrea Brancale, who provided me with encouragement and patience throughout the duration of this PhD. He is a fantastic person and a great Supervisor. I'm incredibly grateful to him for the incredible adventure, for his unwavering support and patience during my PhD.

I'm deeply indebted to Dr Marcella Bassetto, whose help cannot be overestimated. She always supported and nurtured me during the PhD. I would also like to extend my deepest gratitude to Dr Salvatore Ferla, for the valuable advice during the last years. Thanks, should also go to Gilda an incredible labmate.

I'd like to recognise the helpful contributions that I received from Moira and many thanks to the rest of Prof Brancale's group, Cinzia, Martin and Cecilia for the pleasant moments spent together.

I also had the great pleasure of working with Birgit, thank you very much for the nice moments and the support during our PhD.

A special thanks to Giulio, a great housemate and a wonderful person, thank you very much for the excellent time spent together, listening to music together and having our healthy dinners.

I cannot leave Cardiff without mentioning Carmine, without him this PhD wouldn't have been the same, thank you for all the evenings and dinners, thank you for your patience, the unrelenting support and your friendship. Thanks to all the beautiful people I met during this adventure.

I'm incredibly grateful to Thierry Langer, an incredible mentor during. I learnt a lot from our talk and thanks for all the helpful advice.

Many thanks to all the ESRs and PIs of ANTIVIRALS ITN, it was a beautiful journey. Special thanks to Claisen Oomen for the incredible organisation of all the ANTIVIRALS meeting, activities and the support.

I want to thank Marie-Louise Jung and all the beautiful people at Prestwick chemical for the incredible secondment I had in Strasbourg and a huge thanks to Prof. Johan Neyts and Dirk Jochmans for allowing me to do the secondment at Rega Institute in Leuven.

I would like to thank my lovely Mom and my lovely Dad, for everything they have done for me. Thanks for your profound belief in me, for your love, your patient, your advice.

I also wish to thank my best friends Valentino, Lorenzo and Alex for all the time we spent on the phone, for all the support.

The completion of my PhD would not have been possible without the support and nurturing of my soulmate, my lovely Marta. You never let me down; you always kept me steady and always extended a lot of assistance. Thanks for your belief in me. Thank you for everything you have done for me.

Abstract

Respiratory Syncytial Virus (RSV), which is the principal etiological cause of LRTIs (Low Respiratory Tract Infections) worldwide, and the virus members of the enterovirus genus, causing a wide range of diseases, ranging from enteric or respiratory infections to hand-foot-and-mouth disease and acute flaccid paralysis, are associated with severe treats for public health. Currently, there are no effective small-molecule antiviral on the market for the treatment or prevention of these viral infections.

In the first part of the project about RSV, the N and F proteins were chosen as targets for the identification of new anti-RSV agents. On the RSV N protein, different computer-aided techniques were used for structure-based virtual screenings, of commercially available drug-like compounds. The two most potent hits were chosen as a starting point for further investigations. A series of analogues were synthesised and evaluated for anti-RSV activity in a virus-cell-based assay. Computer-aided techniques were also used for the rational design of α -helix mimics, able to inhibit the protein-protein interactions between the trimeric inner coiled-coil and the external α -helixes of the F protein. The most promising compound and a series of analogues were synthesised and evaluated for their anti-RSV activities in a virus-cell-based assay.

For the second part of the project on enterovirus, the highly conserved 2C protein among the enterovirus species was investigated. Starting from fluoxetine, which has been identified to inhibit viral replication by targeting 2C protein, and the new resolved EVA 71 2C crystal structure was used to investigate and to elucidate the binding mechanism of fluoxetine, using molecular modelling methodologies. The gained data were used to design, synthesise and evaluate new antiviral agents. Also, different molecular modelling techniques were used to perform a virtual screening in order to identify new potential inhibitors of the ATPase pocket of the 2C protein.

Contents

Part I

Chapter 1: Introduction

1.1	Viruses	3
1.2	Antiviral therapies	7
1.3	Drug discovery and development process	10
1.4	Molecular modelling	12
1.4.1	Quantum mechanics, molecular mechanics and force fields	12
1.4.2	Energy minimisation and conformational search	13
1.5	Computer-Aided Drug Design	14
1.5.1	Structure-based CADD (SB-CADD)	14
1.5.2	Ligand-based CADD (LB-CADD)	16
1.6	Project aims	17
	References	18

Part II

Chapter 2: RSV

2.1	Respiratory Syncytial Virus (RSV)	27
2.1.1	RSV Virion	30
2.1.2	RSV genome organisation	32
2.1.3	RSV Life Cycle	34
2.2	Viral proteins	40
2.3	Project aims	43

Chapter 3: N protein

3.1	N protein	45
3.1.1	Structure-based virtual screening	46
3.1.2	Pharmacophore search	48
3.1.3	Biological evaluation of selected compounds	50
3.1.4	M76 - 1-(2,4-dichlorobenzyl)-1H-pyrazole-3,5-dicarboxylic acid	50

Chapter 4: 2-(2-ethoxyphenyl)quinoline-4-carboxylic acid and 2-(2-(N-(3-chloro-2-methylphenyl)methylsulfonamido)acetamido)benzoic acid

4.1	Synthesis and biological evaluation of 2-(2-ethoxyphenyl) quinoline-4-carboxylic acid	54
4.2	SAR of 2-(2-ethoxyphenyl)quinoline-4-carboxylic acid	57
4.2.1	SAR on the carboxylic group	58
4.2.2	Biological evaluation of the carboxylic acid analogues	62
4.2.3	SAR on the (2-ethoxy)phenyl moiety	63
4.2.4	Biological evaluations of the (2-ethoxy)phenyl moiety analogues	66
4.2.5	SAR on the 6 and 7 positions of the quinolinic ring	68
4.2.6	Biological evaluations of 6 and 7 substituted quinolinic ring analogues	69
4.3	Synthesis of 2-(2-(N-(3-chloro-2-methylphenyl)methylsulfonamido)acetamido)benzoic acid	72
4.3.1	2-(N-(3-chloro-2-methylphenyl)methyl) acetamides analogues	77
4.3.2	2-(2-(N-methylphenyl)acetamido)benzoic acids analogues	79
4.3.3	Biological evaluation	82
4.4	Conclusions	85

Chapter 5: F protein and α -Helix Mimics

5.1	F protein	89
-----	-----------	----

5.2	Rational design of a α -helix mimics	91
5.2.1	Virtual library generation	91
5.2.2	Conformational search and Molecular docking	93
5.3	Synthesis of selected α -helix mimic compound	95
5.3.1	Biological evaluation of selected α -helix mimic	106
5.4	Synthesis of 126 α -helix mimic analogues	107
5.4.1	Biological evaluation of 126 α -helix mimic analogues	110
5.5	Conclusion	111

Chapter 6: Experimental

6.1	General information	114
6.1.1	Molecular modelling	115
6.2	Synthesis of M76	116
6.3	Synthesis of Quinoline structures	119
6.3.1	General procedures 1-3	119
6.3.2	2-(Alkyloxy)acetophenones	120
6.3.3	2-(phenyl)quinoline-4-carboxylic acid analogues	122
6.3.4	Methyl 2-(phenyl)quinoline-4-carboxylate analogues	146
6.4	Synthesis of 2-(2-(N-(3-chloro-2-methylphenyl)methylsulfonamido)acetamido)benzoic acid and derivatives	164
6.4.1	General procedures 4-8	161
6.4.2	Synthesis of 2-(2-(N-(3-chloro-2-methylphenyl)methylsulfonamido)acetamido)benzoic acid	167
6.4.3	2-(N-(3-chloro-2-methylphenyl)methyl)acetamides analogues	171
6.4.4	methyl 2-(2-(N-methylphenyl)acetamido)benzoic acids analogues	176
6.5	Synthesis of the α -helix mimic	183

6.5.1	Synthesis of α -helix mimic building blocks	183
6.5.2	Synthesis of (4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycine	192
6.5.3	Synthesis of (4-(4-(4-amino-3-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycine	197
6.5.4	Synthesis of (4-(4-(4-amino-2-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycine	203
6.5.4	Synthesis of methyl (4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycine derivatives	209
	References	216
	Appendix	231

Part III

Chapter 7: Enterovirus

7.1	Enterovirus	237
7.2	Viral Genome organisation and life-cycle	238
7.3	Therapeutic strategies	241
7.4	Non-structural protein 2C	242
7.5	Non-structural protein 2C inhibitors	245
7.6	Project aims	249

Chapter 8: Fluoxetine

8.1	Fluoxetine	251
8.2	Antiviral evaluation of fluoxetine enantiomers	252
8.2.1	In-vitro binding of fluoxetine enantiomers to 2C protein	253
8.2.2	Antiviral evaluation of fluoxetine enantiomers against other enteroviruses	254
8.3	Identification of fluoxetine binding site	255

8.3.1	Molecular modelling studies	256
8.3.2	Mutations studies on the identified pocket	260
8.4	Fluoxetine fragments	262
8.4.1	Synthesis of fluoxetine fragments	262
8.4.2	Antiviral evaluation of Fluoxetine fragments and 2C binding assay	264
8.5	Fluoxetine analogues	266
8.5.1	Synthesis of fluoxetine analogues	266
8.5.2	Antiviral evaluation of fluoxetine analogues	271
8.6	Conclusions	272

Chapter 9: Guanidine derivatives of fluoxetine

9.1	Design and synthesis of Guanidine derivatives of fluoxetine	275
9.2	Antiviral evaluation of Guanidine derivatives of fluoxetine	279
9.3	Conclusions	291

Chapter 10: N-benzyl-N-phenylfuran-2-carboxamide

10.1	N-benzyl-N-phenylfuran-2-carboxamide	283
10.1.1	Synthesis of N-(4-fluorobenzyl)-N-(4-methoxyphenyl)furan-2-carboxamide	284
10.1.2	Antiviral evaluation of N-(4-fluorobenzyl)-N-(4-methoxyphenyl)furan-2-carboxamide	286
10.2	N-(4-fluorobenzyl)-N-(4-methoxyphenyl)furan-2-carboxamide analogues	287
10.2.1	Synthesis of N-benzyl-2-(methylamino)-N-phenyl acetamides	288
10.2.2	Synthesis of N-(4-fluorobenzyl)-N-(4-methoxyphenyl) furan-2-carboxamide analogues	289
10.2.3	Antiviral evaluation of N-(4-fluorobenzyl)-N-(4-methoxyphenyl)furan-2-carboxamide analogues	290

10.3	Furamide analogues	292
10.3.1	Synthesis of furamide analogues	292
10.3.2	Antiviral evaluation of furamide analogues	294
10.4	Conclusions	296
Chapter 11: 2C ATPase pocket		
11.1	2C protein ATPase activity	298
11.2	Ensemble docking	298
11.2.1	MD simulation and clustering analysis	299
11.2.2	Ensemble Docking	300
11.3	Antiviral evaluation of the selected compounds	301
11.4	Conclusions	302
Chapter 12: Experimental section		
12.1	General information	304
12.2	Molecular dynamics simulations of Fluoxetine enantiomers	305
12.3	Synthesis of Fluoxetine fragments	306
12.4	Synthesis of Fluoxetine analogues	309
12.5	synthesis of Guanidine derivatives of fluoxetine	324
12.6	Synthesis of N-(4-fluorobenzyl)-N-(4-methoxy phenyl)furan-2-carboxamide and related analogues	332
12.6.1	Synthesis of N-benzyl-2-(methylamino)-N-phenyl acetamides	338
12.6.2	Synthesis of furamide analogues	343
12.7	MD simulations of the ATPase pocket	348
	References	349

Appendix	358
Chapter 13: Conclusions	361

List of abbreviations

2D	Two-dimensional
3D	Three-dimensional
6-HB	6-helix-bundle
Å	Angstrom
AAA+	ATPases Associated with various cellular Activities
Acc	H-bond acceptor heavy atom
AFP	Acute Flaccid Paralysis
AM	Acetoxymethyl
AMBER	Assisted Model Building with Energy Refinement
ANP	Acyclic nucleoside phosphonate
Arg	Arginine
ATP	Adenosine triphosphate
AVA	Antiviral activity
BB	Building block
Boc	Tert-butyloxycarbonyl
CADD	Computer-aided drug design
CC ₅₀	Cytotoxic concentration to observe 50% adverse effect
COPD	Chronic obstructive pulmonary disease
CPE	Cytopathic effect
CS-B	Chondroitin dermatan sulfate B
C-TD	C-Terminal Domain
CV	Coxsackievirus
D	Aspartic acid
DB	Database
DCM	Dichloromethane
DIPEA	N,N-Diisopropylethylamine
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
Don	H-bond donor heavy atom
dsDNA	Double strand DNA

E	Glutamic acid
EC ₅₀	Effective concentration to observe 50% activity
EM	Electron microscopy
EMA	European Medicines Agency
F	Phenylalanine
FDA	Food and Drug Administration
GAG	Glycosaminoglycan
GE	Gene end
GFP	Green fluorescent protein
GS	Gene start
H	Histidine
h	Hour
HBV	Hepatitis B virus
HCMV	Human cytomegalovirus
HCV	Hepatitis C virus
HEV	Human enterovirus
HFMD	Hand-foot-and-mouth disease
HIV	Human immunodeficiency virus
HPV	Human papillomavirus
HRA	Heptad repeat A
HRB	Heptad repeat B
HRV	Human rhinovirus
HS	Heparan sulfate
HSV	Herpes simplex virus
HTS	High throughput screening
Hyd	Hydrophobic feature
ICTV	International Committee on Taxonomy of Viruses
IDP	Intrinsically disordered protein
IFN	Interferon
Ile	Isoleucine
IRES	Internal ribosome entry site
IS	Intergenic sequence
ITC	Isothermal titration calorimetry

K	Lysine
kb	Kilo base
LB-CADD	Ligand-based Computer-aided drug design
Le	Leader region
LRTI	Low Respiratory Tract Infection
M	Methionine
MD	Molecular dynamics
MERS	Middle East respiratory syndrome
ML	Metal ligator heavy atom
MMFF	Merck Molecular Force Field
MOE	Molecular Operating Environment
mRNA	Messenger RNA
MS	Mass Spectroscopy
MTS	3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium
MW	Molar weight
NME	New molecular entitie
NMR	Nuclear magnetic resonance
NNRTI	Nonnucleoside reverse transcriptase inhibitor
NPEV	Non-Polio Enterovirus
NRTI	Nucleoside reverse transcriptase inhibitor
NSP	Non-structural protein
nt	Nucleotide
N-TD	N-Terminal Domain
OPLS	Optimized Potentials for Liquid Simulations
ORF	Open Reading Frame
OSBP	Oxysterol-binding protein
PDB	Protien Data Bank
Phe	Phenylalanine
PI	Protease inhibitor
PPI	Protein-Protein interaction
PV	Poliovirus
QMD	Quenched molecular dynamic

R	Arginine
r.t.	Room temperature
RdRp	RNA-dependent-RNA polymerase
Rf	Retention factor
RMSD	Root-mean-square deviation
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
RO	Replication organelle
RSV	Respiratory Syncytial virus
RT	Reverse transcriptase
S	Serine
SAR	Structure-activity relationship
SARS	Severe acute respiratory syndrome
SB-CADD	Structure-based Computer-aided drug design
SI	Selectivity Index
ssDNA	Single strand DNA
SSRI	Selective serotonin reuptake inhibitor
TBTU	O-(Benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium tetrafluoroborate
TEA	Triethylamine
THF	Tetrahydrofuran
TLC	Thin layer chromatography
T _m	Melting temperature
Tr	Trailer region
TSA	Thermal shift assay
UTR	Untranslated region
Val	Valine
VS	Virtual screening
VZV	Varicella zoster virus
Y	Tyrosine

PART I

Chapter 1: Introduction

1.1 Viruses

A virus is an infectious non-living entity. It is small, sophisticated, but not able to replicate without a host cell and evolve by natural selection. Generally, a complete virus particle called virion is composed of the nucleic acid which is the viral genome and the essential infective agent. The genome can be associated with viral proteins called nucleoproteins. The viral genome enclosed in the capsid which, is formed by proteins encoded in the viral genome. The function of the capsid is to protect the viral genome from the intercellular environment. Viruses can also have an envelope which, is a lipid bilayer formed by the host cell membrane.

Since the first observation of an unfilterable pathogen agent infecting tobacco plants by Dmitri Ivanovsky in 1892 and the discovery of the Tobacco Mosaic Virus by Martinus Beijerinck in 1898, 5560 viral species were identified.^{1,2} Following the discovery in 1901 of the first identified human-infecting virus, yellow fever virus, the development in virus discovery technologies lead to the identification of 219 virus species able to infect humans, and the number is rising with a rate of 3-4 new species per year.^{3,4} This increasing number of the discovered human-infecting virus is reflecting the social and economic impact on society. Over the centuries of human history, empires and civilisations have been affected by infectious pathogens. Among these pathogens, viruses were the leading cause of epidemics and pandemics events able to change the course of history, affect the global economy, affecting trades, people movements across the countries and even social behaviours.⁵⁻⁷

Virus classification

The increasing number of viruses led to the need for a classification method. For this reason, in 1966 the International Committee on Taxonomy of Viruses (ICTV) was created to standardising the classification of viruses.⁸ The viruses are classified based on the Linnaean hierarchical system, initially proposed in 1962 by Lwoff, Robert W. Horne, and Paul Tournier.⁹ Currently, the taxonomic system is grouping the viruses based on their shared properties and the viral genome, with the following structure:¹

- One **Phylum** (-viricota)
- Two **Subphylum** (-viricotina)
- Six **Class** (-viricetes)
- 14 **Order** (-virales)
- Five **Suborder** (-virineae)
- 150 **Family** (-viridae)
- 79 **Subfamily** (-virinae)
- 1021 **Genus** (-virus)

- 5560 **Species**

Another type of classification is the Baltimore classification, introduced in 1971 by the Nobel prize winner David Baltimore.¹⁰ All the viruses are divided into groups based on the viral genome and the viral replication method. Based on the viral genome we can identify three significant classes of viruses: DNA viruses, RNA viruses and Retro-transcribing viruses. These viruses are divided into seven groups in the Baltimore classification (see Figure 1).

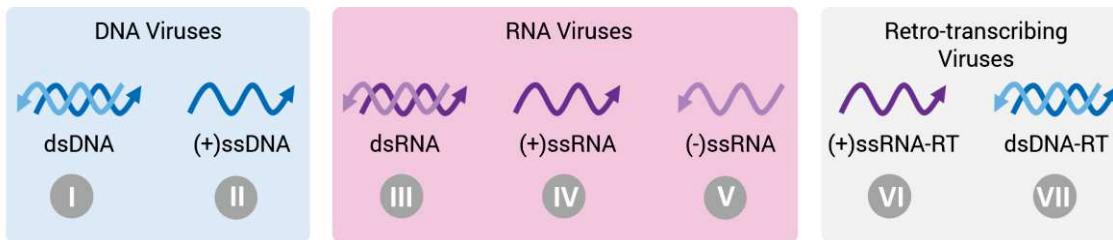


Figure 1. Baltimore classification. Representation of the Baltimore classification. In Baltimore classification viruses are divided into three big classes on the base of their genome: DNA viruses, RNA viruses and retro-transcribing Viruses. Each class is further divided into groups, based on if the genome is a double stranded or single stranded. ds = double stranded. ss = single stranded. (+) = positive. (-) = negative. RT = Reverse Transcriptase.

DNA viruses are classified into two groups. The group I is composed of the double strand DNA (dsDNA) viruses (e.g. Adenoviruses, Herpesviruses, Poxviruses). On the other hand, group II represent the viruses having a + single-stranded DNA genome (+ssDNA); for example, the Parvoviruses are part of this group. The RNA viruses are divided into three groups, called group III, group IV and group V. The group III are the double-stranded RNA (dsRNA) viruses, such Reoviruses. The group IV is formed by positive single-stranded RNA (+ssRNA) viruses (e.g. Picornaviruses, Togaviruses). Finally, the group V is formed by the negative single-stranded RNA (-ssRNA) viruses, the Orthomyxoviruses and Rhabdoviruses are members of this group.

The last two groups, group VI and group VII, in the Baltimore classification are the groups composed by the retro-transcribing viruses. The group VI is composed by the positive single-stranded RNA viruses which, are using the reverse transcriptase to synthesise a DNA intermediate. This group include the retroviruses class. The last group, group VII, is formed by those viruses, such the Hepatitis B virus, which have a double-stranded DNA genome, and which are using an RNA intermediate as a template for the reverse transcriptase for the synthesis of viral DNA.

Viral life cycle

The different species of viruses need a host cell to replicate themselves, to evolve and to be able to infect other hosts. The different viruses shared a typical life cycle process which, can be divided into three stages: viral entry, viral replication, viral shedding.¹¹

The first stage, the viral entry, starts when the virus attaches to the cell. This attachment is mediated by the interaction of specific viral protein on the virion surface and receptors on the host cell membrane. Once the virus is attached to the host cell membrane, the virus must enter the cell to be able to release the viral genome to infect the host cell. This entry stage could occur by membrane fusion, typical of enveloped viruses, and by endocytosis.

All viruses need a further step to release the viral genome inside the cell, called uncoating. Through this process, the proteinaceous shell protecting the viral genome, capsid or core, is disassembled and the viral genome delivery in the host cell.^{12,13}

The viral genome, now inside the host cell, contains the information needed by the virus to produce proteins that are necessary for the assembly of the functional and infective copy of the virus. Viruses, like all biological system, are subjected to the Central Dogma of Molecular Biology proposed by Crick in 1958.^{14,15} The information contained in the DNA Nucleic acid sequence is transfer into protein, through the synthesis of RNA nucleic sequence. The step of RNA production from the DNA is called transcription, and the one in which the RNA produces proteins is called translation (see figure 2).

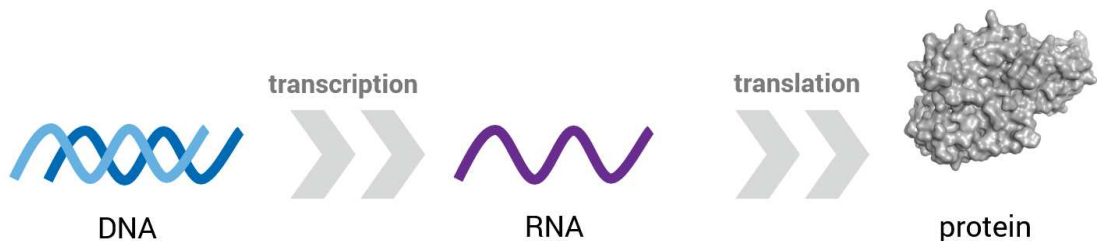


Figure 2. Central Dogma of Molecular Biology. Schematic representation of the central dogma of molecular biology proposed by Crick. The information contained in the DNA sequence is transferred in the proteins through the RNA.

In contrast to the living cells genome, viral genomes are not encoding the complex machinery for the translation process, the ribosomes. For this reason, all viruses need to synthesise mRNA that will be translated into protein by the host cell ribosomes. For convention, this mRNA is called +mRNA.

The different groups of viruses in the Baltimore classification, undergo different transcription steps to produce +mRNA. The difference is based on the type of genome information carried by the infecting virus, as shown in figure 3. The viral +mRNA produced, by the viral genome, will be translated into viral protein components by the host cell ribosome machinery. The viral components can be classified as structural proteins and non-structural proteins. The structural proteins are those that are going to forming the viral particle. The

non-structural are those viral proteins which, are not going to be present inside the virion. The non-structural proteins are mainly proteins needed by the virus for the genome replication.

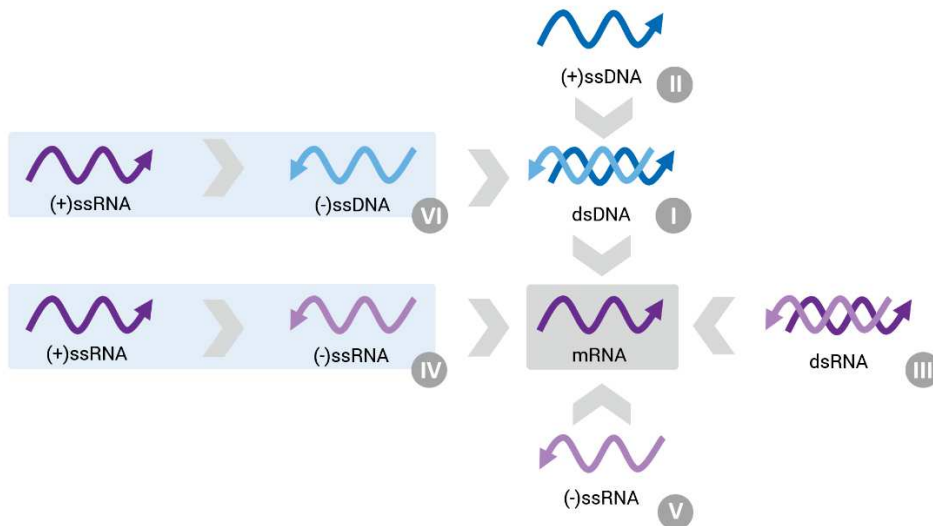


Figure 3. Replication in the different Baltimore classification groups. Schematic representation of the different replication strategies for the different Baltimore virus groups. ds = double stranded. ss = single stranded. (+) = positive. (-) = negative. RT = Reverse Transcriptase.

The general viral replication process diverges for different type of viruses. For example, RNA viruses replicate in the host cell cytoplasm. Instead, DNA viruses and retroviruses need to reach the nucleus to perform DNA synthesis. The different types of virus can use different host cellular organelles during the replication process, like endoplasmic reticulum or the Golgi apparatus.¹⁶⁻¹⁸

Once the viral components are inside the host cell, and the virus has started the genome replication, the new copies of the virus start to be assembled for the final release of the new virions, the viral shedding. The viral shedding can occur via three different processes: 1) budding, 2) exocytosis, 3) via apoptosis. The shedding process depends on the type of virus. The enveloped ones are mainly released from the infected cell via the budding process, in which the virus is forming the viral envelope from the host cell membrane.¹⁹⁻²⁰ The non-enveloped virion is instead released from the infected cell or by exocytosis or by apoptosis of the infected cell.^{11,21}

1.2 Antiviral therapies

Virus-Related infections are a major global threat to human health worldwide. The development of successful antiviral treatment was achieved only for 10 of the 219 human-infecting viruses. Among them, the Human immunodeficiency virus (HIV) has a devastating impact on public health. The virus, which was identified in 1983, infected 76 million people worldwide, and 35 million people died because of the virus-related acquired immunodeficiency syndrome (AIDS).²² In 2017, 1.8 million people contracted the virus, and 1 million people died from AIDS-related issue.²³⁻²⁵ Another virus that has a huge public health impact is the Hepatitis C virus (HCV). The virus is the etiological cause of Hepatitis C, a liver-affecting disease. It is estimated that 71 million people are infected worldwide, and 80% of them have a higher risk to develop cirrhosis or liver cancer.²⁶⁻²⁷ Also, Respiratory virus, causing influenza-like illness, are infecting between 25 and 50 million people every year and during the epidemics of influenza viruses, three to five millions of people are estimated to develop a severe illness, and influenza virus-related death is 290000 to 650000.²⁸⁻³⁰

Ninety-six is the number of antiviral drugs approved so far since the commercialisation of the first antiviral drug in 1963, the Idoxuridine.³¹⁻³³ These drugs have been approved and can be found in the 179 antiviral medications on the market.³⁴

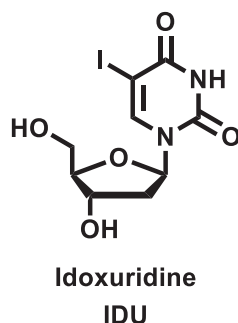


Figure 4. Idoxuridine.

De Clercq and Li, in 2016, reviewed the history of the approved drugs during the last 50 years of the antiviral drug discovery field. As can be seen in Table 1, other than the approved drugs for the treatment of the retrovirus HIV, several other drugs have been approved for the treatment of six DNA viruses, Hepatitis B virus (HBV), varicella zoster virus (VZV), human cytomegalovirus (HCMV), herpes simplex virus (HSV), human papillomavirus (HPV), Variola Virus. On the other hand, among the RNA viruses, treatment is currently available only for hepatitis C virus (HCV), Respiratory Syncytial virus (RSV) and Influenza virus.³⁵

Table 1. Viruses for which antiviral drugs are approved. For each Baltimore virus group are reported the virus families, the human viruses name, the year of discovery, the number of antiviral drugs and their molecular targets.

Group	Family	Human Virus	Year of discovery	Antiviral Drug (n)	Molecular Target(s)
I dsDNA	<i>Hepadnaviridae</i>	HBV	1963	7	DNA polymerase
	<i>Herpesviridae</i>	VZV	1953	5	DNA polymerase, Envelope proteins
		HCMV	1956	4	DNA polymerase
		HSV	1893	9	DNA polymerase, Envelope proteins
	<i>Papillomaviridae</i>	HPV	1965	3	immunomodulators
	<i>Poxviridae</i>	Variola Virus	1931	1	p37 protein
IV (+)ssRNA	<i>Flaviviridae</i>	HCV	1989	16	NS3/4 protease, NS5A, NS5B polymerase
V (-)ssRNA	<i>Orthomyxoviridae</i>	Influenza virus	1933	8	Matrix protein 2, neuraminidase, RNA polymerase, cap-dependent endonuclease
	<i>Paramyxoviridae</i>	RSV	1957	2	RNA polymerase, glycoproteins
VI (+)ssRNA-RT	<i>Retroviridae</i>	HIV	1983	41	Protease, RT, integrase, GP41, CCR5

The ninety-six approved drugs can be divided, on the base of the drug target, into eight classes. The first class is composed of all the approved drugs that are (i) targeting the viral DNA polymerase. These approved drugs can be further divided into five sub-groups based on the type of chemical structure, nucleoside analogues, 5-substituted 2'-deoxyuridine analogues, acyclic guanosine analogues, acyclic nucleoside phosphonate (ANP) analogues and (nonnucleoside) pyrophosphate analogues.

The (ii) hepatitis C virus (HCV) NS5A and NS5B inhibitors compose the second class, and they are targeting the HCV replication. One of the approved NS5B drug, sofosbuvir, is a nucleoside analogue. The class of (iii) reverse transcriptase (RT) inhibitors for the treatment of HIV, can be divided into nucleoside RT inhibitors (NRTIs) and nonnucleoside RT inhibitors (NNRTIs). The (iv) protease inhibitors (PIs) are targeting the viral protease during the virion assembly. (v) Integrase inhibitors are blocking the viral genome integration of HIV. The last three classes are the (vi) entry inhibitors blocking the viral attachment to the cell, (vii) the influenza virus inhibitors which are approved for the treatment of the influenza virus infection, and the (viii) immunostimulators, interferons, oligonucleotides, and antimetabolic inhibitors, this class of drugs are targeting host-cell protein target crucial for the viral replication.

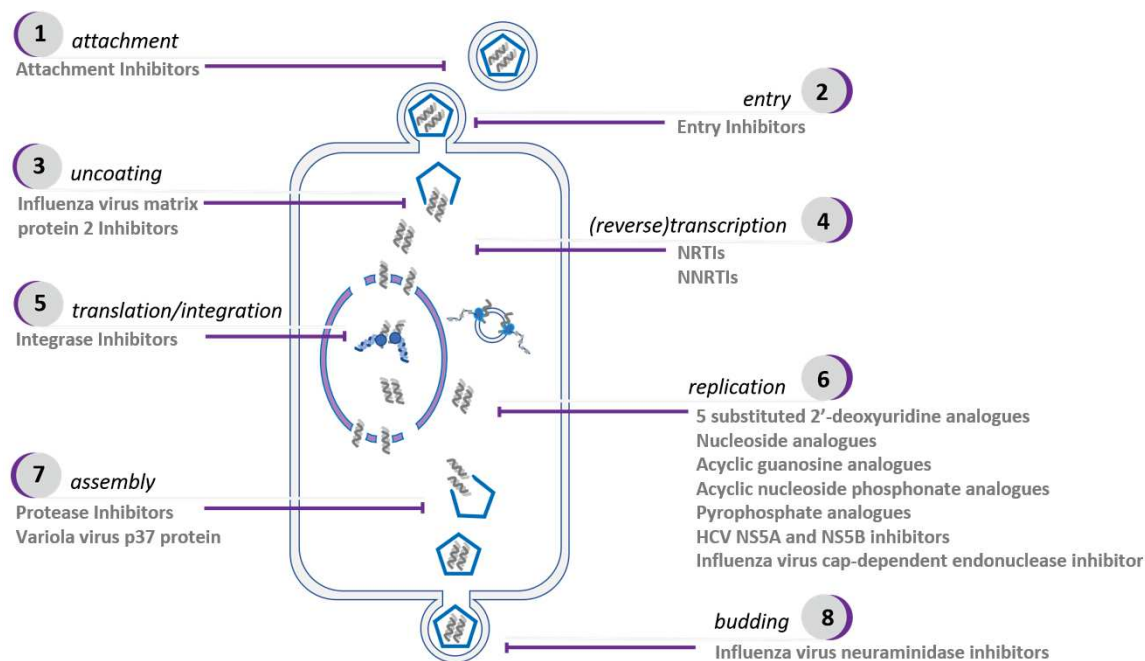


Figure 5. Representation of the viral life-cycle stages targeted by the approval drugs classes

Eleven of the approved drugs are approved for the treatment of different viruses. One example is the Ribavirin drug, which is approved for the treatment of the RNA viruses HCV, RSV, and influenza virus. The drug inhibits the viral replication by targeting the RNA-Dependent-RNA polymerase, which is a typical crucial stage in the viral life cycle for the RNA viruses.

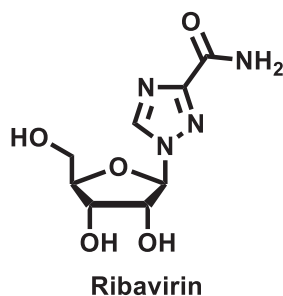


Figure 6. Ribavirin.

Another example is the approved drugs of the acyclic nucleoside phosphonate (ANP) analogues class, were used off-label for the treatment of different DNA virus-related infections.³⁶

Despite the enormous efforts, which have been put in the antiviral drug development and the achievement that has been made in the last decades in this field, viral infections remain a significant threat for human health worldwide. Besides, the threats caused by emerging drug-resistant viruses, antiviral drugs or vaccine against emerging and re-emerging viruses

such Ebola virus, Zika virus, severe acute respiratory syndrome (SARS) and the Middle East respiratory syndrome (MERS) coronaviruses, are not available, pushing the need for the development of new antiviral drugs.³⁷⁻³⁹

1.3 Drug discovery and development process

The drug discovery and development process to bring new molecular entities (NMEs) successfully in the market, is a very long and expensive process, which was estimated to be above \$2.5 billion cost.^{40,41} The process is generally considered to be 10 to 15 years long, but it might take longer due to the increasing complexity required for drug approval from the regulatory agencies, like the Food and Drug Administration (FDA) in the US or the European Medicines Agency (EMA) in Europe.^{42,43} Each therapeutic area has its challenges. In the case of the antiviral drug development, the long time required for the approval of a new drug might be a problem in the cases of life-threatening infections caused by emerging or re-emerging viruses, due to the lack of commercially available treatments. Also, most of this virus is affecting developing countries, reducing the interest of pharmaceutical companies for investment in this therapeutic area.

The development pipeline for bringing the NME into the market generally can be divided into six steps (see figure 7), drug discovery, pre-clinical studies, three phases of clinical trials and the pharmacovigilance.⁴⁴⁻⁴⁶

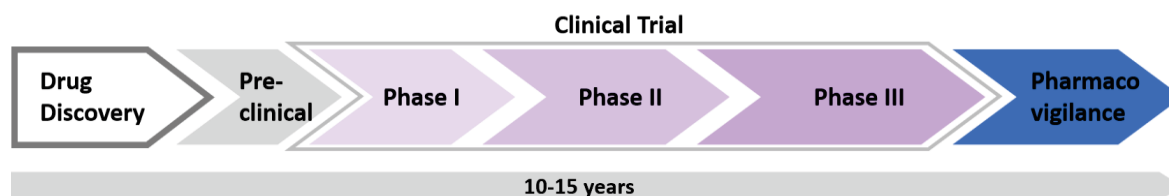


Figure 7. Drug discovery and development process.

In the case of the therapeutic area of antiviral drugs, the discovery step starts with the identification and the validation of a molecular target, which it can be a viral target, or a host cell target essential for the viral replication. Drugs that target a viral protein are related to high specificity and reduced side-effect. The drawbacks of targeting a viral protein are that drug resistant-mutants can emerge after the administration of this drugs and that the drug spectrum of activity is related to how much the protein is conserved among the different serotype inside a family.⁴² After the identification of a druggable target, the development of assays is required to start screening chemical libraries for hit molecule identification. Commonly these screening are performed using High throughput screening (HTS) method, in which thousands of compounds are tested in a short time.⁴⁷ Different antiviral assays are

developed to confirm the activity of the hit molecules and to understand the mode of action. The identified hits are then optimised to obtain lead compounds, with enhanced potency and selectivity in vitro and proven efficacy in animal models. Before reaching the preclinical stage of the development pipeline, the lead compounds are further optimised to enhance the pharmacokinetic properties of the molecules. The candidate compounds, during the preclinical step, are extensively evaluated in vitro and animal models for their safety, before reaching the clinical trials, in which the candidates are tested on humans. The clinical trials are composed of three phases, in which during the phase I, the drug candidates are evaluated on a small group of healthy volunteers to evaluate the safety and the tolerability of the drug candidate. Then, the drug candidate is administered to evaluate the efficacy in a group of patients during phase II of the clinical trials. Finally, the drug candidate is administered to a larger group of patients during phase III of the clinical trials. The drug candidate, once its safety and efficacy are proven and confirmed in human, it can reach the market, but it is continuously under observation for a side effect that was not observed during the clinical trials (pharmacovigilance).⁴⁸

During the last decades, several new techniques and technologies were developed to assist, and more importantly, to reduce time and costs of the process of drug discovery and development. The boom of the computational power, in the last decades, gave to scientists the opportunity to handle a significant amount of data and to simulate complex systems, using tools developed by the computer science field. Nowadays, in the field of drug discovery, several in silico techniques are used both in academia and in pharmaceutical companies, by chemists and biologists to assist them during the research process.

1.4 Molecular modelling

Computational techniques have become a powerful ally of the medicinal chemistry during the drug discovery and development process.⁴⁹ The advances in computer hardware and softwares in the last decades, are helping to handle and process the big amount of data generated during R&D in pharmaceutical field. The software and algorithms are helping to process the data into information and the information into knowledge. Computational methods are widely used worldwide to speed-up, reduce time and costs R&D processes.⁵⁰ The term “Molecular modelling“, which is today largely used in the drug design process, refer to all the computational techniques, such as quantum mechanics, molecular mechanics, energy minimisation, generation of conformations and other computational methods to model, simulate and generate data related to the structure and property of bio-molecular systems.⁵¹

The structural information of small molecules and protein are very important, this because the effect of molecules on specific protein targets are related to specific interactions that the molecules establish, through a series of chemical features with the protein, and they are closely dependent on the 3D structures of the molecules, called also conformations.⁵² Each molecule can assume different conformation on the base of the number of rotatable bonds and each conformation is associated with a specific energy. For this reason, during the modelling process, the prediction of the energies of a possible conformation of the molecules has a central role. Several computational methods were developed to address the calculation of these energies, and among them the quantum mechanics and the molecular mechanics are two widely used methods.⁵³

1.4.1 Quantum mechanics, molecular mechanics and force fields

The quantum mechanics methods are the most precise methods for the calculation of molecular energies. These methods are calculating the physicochemical properties of a molecule at a quantum chemistry level, taking in consideration both the nuclei of the atoms and the electrons and for this reason, the method is extremely time-consuming, and the calculation can be performed on a small number of atoms. The quantum mechanics methods are mainly used for the calculations of molecular properties like the electrostatic potentials, the energy of the molecular orbitals and the calculations of partial atomic charges.⁵⁴

On the other hand, energy minimisation and generation of molecules conformations, are common tasks carried out in molecular modelling using molecular mechanics methods. In Molecular mechanics, the molecules are considered more straightforwardly compared to quantum mechanics. In fact, each atom is considered as one particle with an assigned radius and the bond between the different type atoms are considered as a spring, this

makes it possible to perform the calculations not only on small molecules but also on more complex systems, like proteins, in a relatively short period of time compared to the quantum mechanics calculations.⁵³

The energy related to a specific conformation of a molecule is calculated, in molecular mechanics, using a force field. The term force field refers to the equation and the parametrised value of the different atoms used to calculate the energy. In all the force fields, the calculated energy is the result of the contributions of bonded or covalent terms and non-bonds on non-covalent terms. The bonded or covalent terms are the stretching of the bond between two atoms, the angle and the torsional rotation between atoms. The non-bonded or non-covalent terms are the electrostatic interaction and the van der Waals interactions. Some force fields are also considering other contribution terms and different parametrisation of the atoms and they are optimised for a different set of molecules.⁵¹ AMBER force field is optimised for proteins and nucleic acids, while the MMFF94 is a force field commonly used for small molecules.^{55,56} The licensed Schrodinger suite software has its force field, the OPLS2005 and the last version OPLS3, which are optimised for the medicinal chemistry research both for small molecules and proteins.^{57,58}

1.4.2 Energy minimisation and conformational search

The energy minimisation is an important task performed during molecular modelling, especially when, for example, small molecules, are converted from 2D structures into 3D structures, with the aim of identified conformations of these molecules related with a minimum of potential energy. The energy minimisation became a complicated task when it is applied to molecules with more than one rotatable bond. The term “energy surface” of a molecule refers in molecular modelling to how the potential energy reacts to a change in the conformation of the molecule. The energy surface of a molecule has different points of minimum energy. In those points, any further change of the 3D structure will be related to an increase in the potential energy. The minimisation methods are using an algorithm to find a modification of the conformations related to lower potential energy compared to the initial confirmation. The only drawback is that the process is going “downhill”, and the minimum energy found, could be a local minimum and not a global minimum.⁵¹

One of the applications of the minimisation methods is the conformational search, in which all the minimum-energy conformations are generated for a specific molecule. The conformational search is essential when the molecular modelling techniques are applied to the drug design process. In fact, in the modern molecular modelling software, the conformational search is used for the generation of databases (DB) containing the different conformations of a specific set of molecules, which are used in different computer-aided drug design techniques for the identification of biologically active compounds.⁵⁹

1.5 Computer-Aided Drug Design

Computer-aided drug design (CADD) comprises a series of molecular modelling methodologies applied to rational drug design of biologically active small molecules. Commonly the drug discovery and development process, begin with the identification of hit molecules. The HTS technique is widely used in the pharmaceutical industry for the identification of hit molecules, but this technique is related to high operational costs. Once the hit molecule is confirmed, the chemical structure hit is further derivatised in the process called hit-to-lead, followed by the lead optimisation process. In recent years, CADD methodologies are commonly used for the rational drug design, as useful tools to speed up and reduce costs of the hit-to-lead and lead optimisation processes.^{60,61}

CADD can be split into two main groups, the Structure-based CADD (SB-CADD) and the Ligand-based CADD (LB-CADD). In the Structure-based CADD (SB-CADD), the 3D information of the biological target is used as a starting point to find new or optimise active molecules. In Ligand-based CADD (LB-CADD), only the information related to the biologically active compounds on a specific target is used for the identification of new biologically active compounds.⁶²

1.5.1 Structure-based CADD (SB-CADD)

The structure-based CADD are all the techniques applied to rational drug design where the structures, of the ligand and of target protein, are known. The structure of a protein can be coming from different sources.⁴⁹ The Protein Data Bank (PDB) is the biggest public archive of macromolecules (protein, RNA and DNA) structures, and since the announcement in 1971 the archive contains today 152500 structures.⁶³ These structures are the result of X-ray crystallography, NMR and cryo-electron microscopy (EM).⁶⁴ Another source of protein structure is the homology modelling technique. The unknown 3D structure of a target protein is built starting from a known 3D protein structure (template), knowing only the amino acid sequence of the last one. To obtain a good homology model to be used in SB-CADD, the sequence of the template must have more than 50% of similarity with the target sequence. The homology model protocol consists of the following steps: 1) sequence alignment, 2) Model building, 3) model optimisation 4) model validation. All these steps for the generation of homology model are now incorporated in different softwares.⁶⁵

Another source of protein structure is molecular dynamics (MD) simulations. The MD simulations are used with different scopes, like binding energy analysis, molecular interaction analysis, but also to obtain an ensemble of conformations of the protein that can be used in computer-aided drug design. In fact, the MD is using the Newton's second law to study the movement and the interactions of atoms over time. The integration of the second-order differential equations, showed in figure 8, provide a series of configurations

of the molecular system during the time of the simulations. These configurations represent the trajectories of the simulations, in which the atoms position and their velocities over time are stored.⁶⁶

$$\mathbf{f}_i(t) = m_i \mathbf{a}_i(t) = -\frac{\partial V(\mathbf{x}(t))}{\partial \mathbf{x}_i(t)}$$

Figure 8: Newton's second law. f_i , net force on the atom; m_i , mass of the atom; $a_i(t)$, acceleration of the atom; $x(t)$, vector describing the special position of the interacting atoms; $V(x)$, force field

The SB-CADD can be divided into the *de novo* approach and the virtual screening (VS) approach, and both approaches rely on molecular docking methods. In the first one, molecular docking methods are used to identify fragments able to interact in the binding site, and then the fragment are linked together to give novel ligand.⁶⁷ The virtual screening approach, instead, is used to evaluate the interaction of a high number of compounds in the binding site. This method is very useful to reduce the number of compounds to be screened with more sophisticated docking analyses.⁶⁸ Another method that can be used for virtual screening is the structure-based pharmacophore search. In this method, the analysis of the receptor-ligand interactions is used to define a set of essential 3D features of the ligand (pharmacophore model) for the interaction with the receptor. The resulting pharmacophore model will be used to search a 3D database for similar compounds, before using more sophisticated docking analysis methods.^{69,70}

As was said before, both SB-CADD approaches relay on molecular docking. Molecular docking is a method in which the interactions between a ligand and a binding site of a protein is predicted. In fact, the basic hypothesis for the docking method is that the ability of a ligand to affect a target protein, which has a biological effect, is related to its binding mode with the protein. A compound that has a similar binding mode and shares a similar interaction with the ligand will have similar biological effects.⁷¹

Nowadays, several molecular docking programs are available, and these programs are characterised by different scoring functions, which evaluate the interaction of the ligand with the binding site. The different scoring functions are calculating the free binding energy as a result of different energy contributions, which differ between the different molecular docking programs.^{50,72}

In general, the scoring function is calculating the free binding energy (ΔG_{bind}) of the ligand to the protein by the following form:

$$\Delta G_{\text{bind}} = \Delta G_{\text{solvent}} + \Delta G_{\text{conf}} + \Delta G_{\text{int}} + \Delta G_{\text{rot}} + \Delta G_{\text{tr}} + \Delta G_{\text{vib}}$$

$\Delta G_{\text{solvent}}$ is the solvent contributions to the free binding energy. ΔG_{conf} represents the contribution of the ligand and protein conformational changes. ΔG_{int} represents the contribution of the noncovalent interactions, electrostatic and van der Waals forces, between the protein and the ligand. ΔG_{rot} represents the free energy penalty related to the loss of rotational freedom of the bond. ΔG_{tr} represents the free energy penalty related to the loss of translational and rotational of freedom. ΔG_{vib} is the free energy related to the vibrational modes.

When dealing with a high number of docked compounds, different programs with different scoring functions can be used to evaluate the compounds, and the scoring results of each program can be combined together (*consensus scoring*), this was demonstrated to increase the accuracy of the scoring and ranking.^{73,74}

1.5.2 Ligand-based CADD (LB-CADD)

The ligand-based approaches are the ones that rely only on a set of ligands active on a specific target, this because similar compounds can evoke similar biological effect and interact with the target protein. One of the most common ligand-based approaches is the pharmacophore methods. The core concept of the pharmacophore methods is that the biological activity of a ligand is related with the 3D orientation of a specific set of structural features, associated to pharmacophore features that include positively and negatively ionized areas, hydrogen bond donors, hydrogen bond acceptors and hydrophobic features. One ligand or a set of active ligands can be used to generate a pharmacophore model. This model can be further used to screen a DB, containing a large number of compounds, to find new molecules presenting the same set of features of the model (pharmacophore search).⁷⁵

1.6 Project aims

In this drug discovery project, several techniques were applied to different viral targets to identify small molecules as potential anti-RSV and anti-enterovirus agents.

Among the RSV viral protein, the N protein and the F protein have been chosen as targets for the identification of new viral replication inhibitors.

The N protein, which is part of the ribonucleocapsid complex (RNP), plays a key role in virus replication, acting as a template for transcription and replication, performed by the RdRp (RNA-dependent RNA polymerase). Interactions between the RNP and P protein, which is part of the RdRp, have been reported to be fundamental for viral RNA synthesis.

The F protein is the viral protein responsible the fusion of the viral and cell membranes.

During this project, on both the RSV protein targets, the following activities were performed:

- molecular modelling techniques for the identification of new potential inhibitors;
- synthesis of several compounds for Structure-Activity Relationships (SARs) exploration of the most promising hits;
- Biological evaluation of the synthesised compounds.

For the enterovirus part of the project, the non-structural 2C protein was chosen as a target for the identification of small molecules as anti-enterovirus agents. The 2C protein is highly conserved 2C protein among the enterovirus species and play a crucial role during the viral replication. In this part of the project the following activities were performed:

- molecular modelling techniques used to elucidate the binding mechanism of fluoxetine to the non-structural 2C protein;
- SAR exploration of fluoxetine, through the synthesis of a series of fluoxetine fragments and fluoxetine analogues;
- design and synthesis of novel anti-enterovirus agents;
- biological evaluation of the synthesised compounds;
- molecular modelling techniques for the identification of new potential inhibitors of the ATPase pocket.

References

- 1 Lefkowitz EJ, Dempsey DM, Hendrickson RC, Orton RJ, Siddell SG, Smith DB. Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Res.* **2018**, 46, D708–717. <https://doi.org/10.1093/nar/gkx932>
- 2 Zaitlin M. The Discovery of the Causal Agent of the Tobacco Mosaic Disease. *Discoveries in Plant Biology* **1998**, 105-110. https://doi.org/10.1142/9789812817563_0007
- 3 Woolhouse M, Scott F, Hudson Z, Howey R, Chase-Topping M. Human viruses: discovery and emergence. *Philos .Trans. R. Soc. Lond. B. Biol. Sci.* **2012**, 367, 2864–2871. <https://doi.org/10.1098/rstb.2011.0354>
- 4 Barba M, Hadidi A. Application of Next-Generation Sequencing Technologies to Viroids. *Viroids Satell.* 2017, 401–412. <https://doi.org/10.1016/B978-0-12-801498-1.00038-3>
- 5 Lindahl JF, Grace D. The consequences of human actions on risks for infectious diseases: a review. *Infect. Ecol. Epidemiol.* **2015**, 5, 30048. <https://doi.org/10.3402/iee.v5.30048>
- 6 Turan B, Budhwani H, Fazeli PL, Browning WR, Raper JL, Mugavero MJ, et al. How Does Stigma Affect People Living with HIV? The Mediating Roles of Internalized and Anticipated HIV Stigma in the Effects of Perceived Community Stigma on Health and Psychosocial Outcomes. *AIDS Behav.* **2017**, 21, 283. <https://doi.org/10.1007/s10461-016-1451-5>
- 7 Economic Impact of Zika Virus. *Zika Virus Dis.* **2018**, 137–142. <https://doi.org/10.1016/B978-0-12-812365-2.00012-3>
- 8 Adams MJ, Lefkowitz EJ, King AMQ, Harrach B, Harrison RL, Knowles NJ, et al. 50 years of the International Committee on Taxonomy of Viruses: progress and prospects. *Arch. Virol.* **2017**, 162, 1441–1446. <https://doi.org/10.1007/s00705-016-3215-y>
- 9 Lwoff A, Horne R, Tournier P. A System of Viruses. *Cold Spring Harb. Symp. Quant. Biol.* **1962**, 27, 51–55. <https://doi.org/10.1101/SQB.1962.027.001.008>
- 10 Baltimore D. Expression of animal virus genomes. *Bacteriol Rev.* **1971**, 35, 235–241. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC378387/>

- 11 Dimmock NJ, Easton AJ, Leppard KN. Introduction to modern virology, 7th Edition, *Wiley-Blackwell* **2016**. [https://doi.org/10.1016/S0966-842X\(00\)88934-9](https://doi.org/10.1016/S0966-842X(00)88934-9)
- 12 Kilcher S, Mercer J. DNA virus uncoating. *Virology* **2015**, 578–590. <https://doi.org/10.1016/j.virol.2015.01.024>
- 13 Chang A, Dutch RE. Paramyxovirus Fusion and Entry: Multiple Paths to a Common End. *Viruses* **2012**, 4, 613–636. <https://doi.org/10.3390/v4040613>
- 14 Crick F. Ideas on protein synthesis. *Symp. Soc. Exp. Biol.* **1956**, XII, 139–163.
- 15 Crick F. Central Dogma of Molecular Biology. *Nature* **1970**, 227, 561–563. <https://doi.org/10.1038/227561a0>
- 16 White KA, Enjuanes L, Berkhout B. RNA virus replication, transcription and recombination. *RNA Biol.* **2011**, 8, 182–183. <https://doi.org/10.4161/rna.8.2.15663>
- 17 Kazlauskas D, Krupovic M, Venclovas C. The logic of DNA replication in double-stranded DNA viruses: insights from global analysis of viral genomes. *Nucleic Acids Res.* **2016**, 44, 4551. <https://doi.org/10.1093/nar/gkw322>
- 18 Schmid M, Speiseder T, Dobner T, Gonzalez RA. DNA virus replication compartments. *J. Virol.* **2014**, 88, 1404–1420. <https://doi.org/10.1128/JVI.02046-13>
- 19 Chen BJ, Lamb RA. Mechanisms for enveloped virus budding: Can some viruses do without an ESCRT? *Virology* **2008**, 372, 221–232. <https://doi.org/10.1016/j.virol.2007.11.008>
- 20 Chazal N, Gerlier D. Virus entry, assembly, budding, and membrane rafts. *Microbiol Mol. Biol. Rev.* **2003**, 67, 226–237. <https://doi.org/10.1128/membr.67.2.226-237.2003>
- 21 Bird SW, Kirkegaard K. Escape of non-enveloped virus from intact cells. *Virology* **2015**, 444–449. <https://doi.org/10.1016/j.virol.2015.03.044>
- 22 Fettig J, Riaz N, Wolden SL, Gelblum DY, Eric J. Global Epidemiology of HIV. *Infect Dis. Clin. North Am.* **2014**, 28, 323–337. <https://doi.org/10.1016/j.idc.2014.05.001>
- 23 UNAIDS. Global HIV; AIDS statistics — 2019 fact sheet. <https://www.unaids.org/en/resources/fact-sheet>
- 24 Sharp PM, Hahn BH. Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect. Med.* **2011**, 1, a006841. <https://doi.org/10.1101/cshperspect.a006841>

- 25** HIV Surveillance Report: Diagnoses of HIV Infection in the United States and Dependent Areas. *Centers for Disease Control and Prevention (CDC)* **2017**, vol. 29. <https://www.cdc.gov/hiv/library/reports/hiv-surveillance.html>
- 26** Kish T, Aziz A, Sorio M. Hepatitis C in a New Era: A Review of Current Therapies. *Pharmacy and Therapeutics (P&T)* **2017**, 42, 316–329. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5398625/>
- 27** Bryan-Marrugo OL, Ramos-Jiménez J, Barrera-Saldaña H, Rojas-Martínez A, Vidaltamayo R, Rivas-Estilla AM. History and progress of antiviral drugs: From acyclovir to direct-acting antiviral agents (DAAs) for Hepatitis C. *Med. Univ.* **2015**, 17, 165–174. <https://doi.org/10.1016/j.rmu.2015.05.003>
- 28** WHO. Hepatitis C, <https://www.who.int/news-room/fact-sheets/detail/hepatitis-c>
- 29** WHO. Influenza (Seasonal). [https://www.who.int/news-room/fact-sheets/detail/influenza-\(seasonal\)](https://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal))
- 30** Mousa HAL. Prevention and Treatment of Influenza, Influenza-Like Illness, and Common Cold by Herbal, Complementary, and Natural Therapies. *J Evidence-Based Complement Altern. Med.* **2017**, 22, 166–174. <https://doi.org/10.1177/2156587216641831>
- 31** De Clercq E, Li G. Approved antiviral drugs over the past 50 years. *Clin. Microbiol. Rev.* **2016**, 29, 695–747. <https://doi.org/10.1128/CMR.00102-15>
- 32** Kinch MS, Griesenauer RH. 2017 in review: FDA approvals of new molecular entities. *Drug Discov Today* **2018**, 23, 1469–1473. <https://doi.org/10.1016/j.drudis.2018.05.011>
- 33** Mullard A. 2018 FDA drug approvals. *Nat. Rev. Drug Discov.* **2019**, 18, 85–89. <https://doi.org/10.1038/d41573-019-00014-x>
- 34** Chaudhuri S, Symons JA, Deval J. Innovation and trends in the development and approval of antiviral medicines: 1987–2017 and beyond. *Antiviral Res.* **2018**, 155, 76–88. <https://doi.org/10.1016/j.antiviral.2018.05.005>
- 35** De Clercq E, Li G. Approved Antiviral Drugs over the Past 50 Years. *Clin. Microbiol. Rev.* **2016**, 29, 695-747 <https://doi.org/10.1128/CMR.00102-15>
- 36** De Clercq E. Recent highlights in the development of new antiviral drugs. *Curr. Opin. Microbiol.* **2005**, 8, 552–560. <https://doi.org/10.1016/j.mib.2005.08.010>

- 37** Ryu W-S, Ryu W-S. New Emerging Viruses. *Molecular Virology of Human Pathogenic Viruses* **2017**, 289–302. <https://doi.org/10.1016/B978-0-12-800838-6.00021-7>
- 38** Burrell CJ, Howard CR, Murphy FA, Burrell CJ, Howard CR, Murphy FA. Emerging Virus Diseases. *Fenner and White's Medical Virology* **2017**, 217–225, 5th Edition. <https://doi.org/10.1016/B978-0-12-375156-0.00015-1>
- 39** Rothe C. Emerging Infectious Diseases and the International Traveler. *The Travel and Tropical Medicine Manual* **2017**, 27–35, 5th Edition. <https://doi.org/10.1016/j.ijid.2018.04.3509>
- 40** Di Masi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: New estimates of R&D costs. *J. Health. Econ.* **2016**, 47, 20–33. <https://doi.org/10.1016/j.jhealeco.2016.01.012>
- 41** Prasad V, Mailankody S. Research and development spending to bring a single cancer drug to market and revenues after approval. *JAMA Intern. Med.* **2017**, 177, 1569–1575. <https://doi.org/10.1001/jamainternmed.2017.3601>
- 42** Everts M, Cihlar T, Bostwick JR, Whitley RJ. Accelerating Drug Development: Antiviral Therapies for Emerging Viruses as a Model. *Annu. Rev. Pharmacol. Toxicol.* **2016**, 57, 155–169. <https://doi.org/10.1146/annurev-pharmtox-010716-104533>
- 43** Hering S, Loretz B, Friedli T, Lehr CM, Stieneker F. Can lifecycle management safeguard innovation in the pharmaceutical industry? *Drug Discov. Today* **2018**, 23, 1962–1973. <https://doi.org/10.1016/j.drudis.2018.10.008>
- 44** Hughes JP, Rees SS, Kalindjian SB, Philpott KL. Principles of early drug discovery. *Br. J. Pharmacol.* **2011**, 162, 1239–1249. <https://doi.org/10.1111/j.1476-5381.2010.01127.x>
- 45** Tamimi NAM, Ellis P. Drug development: From concept to marketing! *Nephron. – Clin. Pract.* **2009**, 113, 125–131. <https://doi.org/10.1159/000232592>
- 46** Paul SM, Mytelka DS, Dunwiddie CT, Persinger CC, Munos BH, Lindborg SR, et al. How to improve RD productivity: The pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discov.* **2010**, 9, 203–214. <https://doi.org/10.1038/nrd3078>
- 47** Wildey MJ, Haunso A, Tudor M, Webb M, Connick JH. High-Throughput Screening. *Annu. Rep. Med. Chem.* **2017**, 50, 149–195. <https://doi.org/10.1016/bs.armc.2017.08.004>

- 48** Sinha S, Vohora D. Drug Discovery and Development: An Overview. *Pharmaceutical Medicine and Translational Clinical Research* **2017**,19-32. <https://doi.org/10.1016/B978-0-12-802103-3.00002-X>
- 49** Wang X, Song K, Li L, Chen L. Structure-Based Drug Design Strategies and Challenges. *Curr. Top. Med. Chem.* **2018**, 18, 998–1006. <https://doi.org/10.2174/1568026618666180813152921>
- 50** Pirhadi S, Sunseri J, Ryan Koes D. Open source molecular modeling. *J. Mol. Graph. Model.* **2016**, 69, 127–143. <https://doi.org/10.1016/j.jmgm.2016.07.008>
- 51** Leach AR. Molecular Modelling Principles and Applications. *Harlow PE* **2001**, 2nd Edition. <https://doi.org/10.1021/ci9804241>
- 52** Du X, Li Y, Xia Y-L, Ai S-M, Liang J, Sang P, et al. Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *Int. J. Mol. Sci.* **2016**, 17, E144 <https://doi.org/10.3390/ijms17020144>
- 53** Riniker S. Fixed-Charge Atomistic Force Fields for Molecular Dynamics Simulations in the Condensed Phase: An Overview. *J. Chem. Inf. Model.* **2018**, 58, 565–578. <https://doi.org/10.1021/acs.jcim.8b00042>
- 54** Giese TJ, York DM. Quantum mechanical force fields for condensed phase molecular simulations. *J. Phys. Condens. Matter* **2017**, 29, 383002. <https://doi.org/10.1088/1361-648X/aa7c5c>
- 55** Ponder JW, Case DA. Force fields for protein simulations. *Adv. Protein Chem.* **2003** 66, 27–85. [https://doi.org/10.1016/S0065-3233\(03\)66002-X](https://doi.org/10.1016/S0065-3233(03)66002-X)
- 56** Halgren TA. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, 17, 490–519. [https://doi.org/10.1002/\(SICI\)1096-987X\(199604\)17:5/6<490::AID-JCC1>3.0.CO;2-P](https://doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P)
- 57** Shivakumar D, Harder E, Damm W, Friesner RA, Sherman W. Improving the Prediction of Absolute Solvation Free Energies Using the Next Generation OPLS Force Field. *J. Chem. Theory Comput.* **2012**, 8, 2553–2558. <https://doi.org/10.1021/ct300203w>
- 58** Harder E, Damm W, Maple J, Wu C, Reboul M, Xiang JY, et al. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* **2016**, 12, 281–296. <https://doi.org/10.1021/acs.jctc.5b00864>

- 59** Perola E, Charifson PS. Conformational Analysis of Drug-Like Molecules Bound to Proteins: An Extensive Study of Ligand Reorganization upon Binding. *J. Med. Chem.* **2004**, 47, 2499–2510. <https://doi.org/10.1021/jm030563w>
- 60** Usha T, Shanmugarajan D, Goyal AK, Kumar CS, Middha SK. Recent Updates on Computer-aided Drug Discovery: Time for a Paradigm Shift. *Curr. Top Med. Chem.* **2018**, 17, 3296–3307. <https://doi.org/10.2174/1568026618666180101163651>
- 61** Macalino SJY, Gosu V, Hong S, Choi S. Role of computer-aided drug design in modern drug discovery. *Arch. Pharm. Res.* **2015**, 38, 1686–1701. <https://doi.org/10.1007/s12272-015-0640-5>
- 62** Swift R V, Jusoh SA, Li ES, Amaro RE. Knowledge-Based Methods To Train and Optimize Virtual Screening Ensembles. *J. Chem. Inf. Model.* **2016**, 56, 830–842. <https://doi.org/10.1021/acs.jcim.5b00684>
- 63** Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. *Nucleic Acids Res.* **2000**, 28, 235–242. <https://doi.org/10.1093/nar/28.1.235>
- 64** Szymczyzna BR, Taurog RE, Young MJ, Snyder JC, Johnson JE, Williamson JR. Synergy of NMR, Computation, and X-ray Crystallography for Structural Biology. *Structure* **2009**, 17, 499. <https://doi.org/10.1016/j.str.2009.03.001>
- 65** França TCC. Homology modeling: an important tool for the drug discovery. *J. Biomol. Struct. Dyn.* **2015**, 33, 1780–1793. <https://doi.org/10.1080/07391102.2014.971429>
- 66** De Vivo M, Masetti M, Bottegoni G, Cavalli A. Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* **2016**, 59, 4035–4061. <https://doi.org/10.1021/acs.jmedchem.5b01684>
- 67** Popova M, Isayev O, Tropsha A. Deep Reinforcement Learning for De-Novo Drug Design. *Science Advances* **2018**, 4, eaap7885 <https://doi.org/10.1126/sciadv.aap7885>
- 68** Lionta E, Spyrou G, Vassilatis DK, Cournia Z. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Curr. Top. Med. Chem.* **2014** 14, 1923–1938. <https://doi.org/10.2174/1568026614666140929124445>
- 69** Pirhadi S, Shiri F, Ghasemi JB. Methods and applications of structure based pharmacophores in drug discovery. *Curr. Top. Med. Chem.* **2013**, 13, 1036–1047. <https://doi.org/10.2174/1568026611313090006>

- 70** Ran T, Li W, Peng B, Xie B, Lu T, Lu S, et al. Virtual Screening with a Structure-Based Pharmacophore Model to Identify Small-Molecule Inhibitors of CARM1. *J. Chem. Inf. Model.* **2019**, 59, 522–534. <https://doi.org/10.1021/acs.jcim.8b00610>
- 71** Ferreira LG, Dos Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design strategies. *Molecules* **2015**, 20:13384-13421. <https://doi.org/10.3390/molecules200713384>
- 72** Perricone U, Gulotta MR, Lombino J, Parrino B, Cascioferro S, Diana P, et al. An overview of recent molecular dynamics applications as medicinal chemistry tools for the undruggable site challenge. *Med. Chem. Comm.* **2018**, 9, 920–936. <https://doi.org/10.1039/C8MD00166A>
- 73** Wang R, Wang S. How Does Consensus Scoring Work for Virtual Library Screening? An Idealized Computer Experiment. *J. Chem. Inf. Comput. Sci.* **2001**, 41, 1422-1426. <https://doi.org/10.1021/ci010025x>
- 74** Charifson PS, Corkery JJ, Murcko MA, Walters WP. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J. Med. Chem.* **1999**, 42, 5100–5109. <https://doi.org/10.1021/jm990352k>
- 75** Acharya C, Coop A, Polli JE, Mackerell AD, Jr. Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. *Curr. Comput. Aided Drug Des.* **2011**, 7, 10–22. <https://doi.org/10.2174/157340911793743547>

PART II

Chapter 2: RSV

2.1 Respiratory Syncytial Virus (RSV)

RSV (Respiratory syncytial virus) represents a worldwide health problem. It is the principal etiological cause of LRTIs (Low Respiratory Tract Infections) such as Pneumonia and Bronchiolitis.

RSV is estimated to cause 33 million of LRTIs in children under 5 years of age in 2005, of which 3.4 million were hospitalized (in the first two years of life).¹⁻² The estimated mortality related to RSV was 66.000-199.000 in 2005, of which 99% in developing countries.³ Furthermore, RSV LTRIs are a risk factor in HIV infected patients, immune compromised patients and in elderly people.

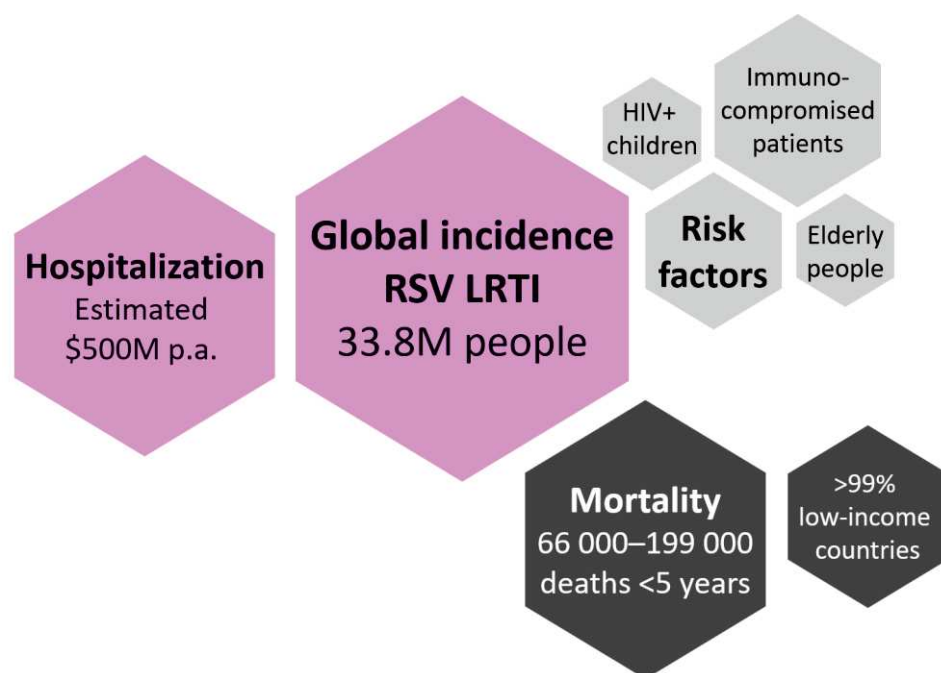


Figure 1. Burden of disease. Representation of information related with RSV low respiratory tract infections (LRTIs), RSV mortality and risk factors.

From a study of hospitalization cost related to bronchiolitis infection in children, the cost in 2002 was estimated to be more than \$500 million per year.⁴ (Fig.1) The high cost of the hospitalization and the lack of an effective cure against the virus make RSV an interesting target for drug discovery. The only approved drug is inhaled Ribavirin, but its use is no longer recommended by the 2014 AAP guidelines for RSV infection due to the ineffectiveness of the treatment.⁵ Another treatment available on the market is a humanized monoclonal antibody, Palivizumab, but due to its high cost it is used only prophylactically in high risk hospitalized children.⁶

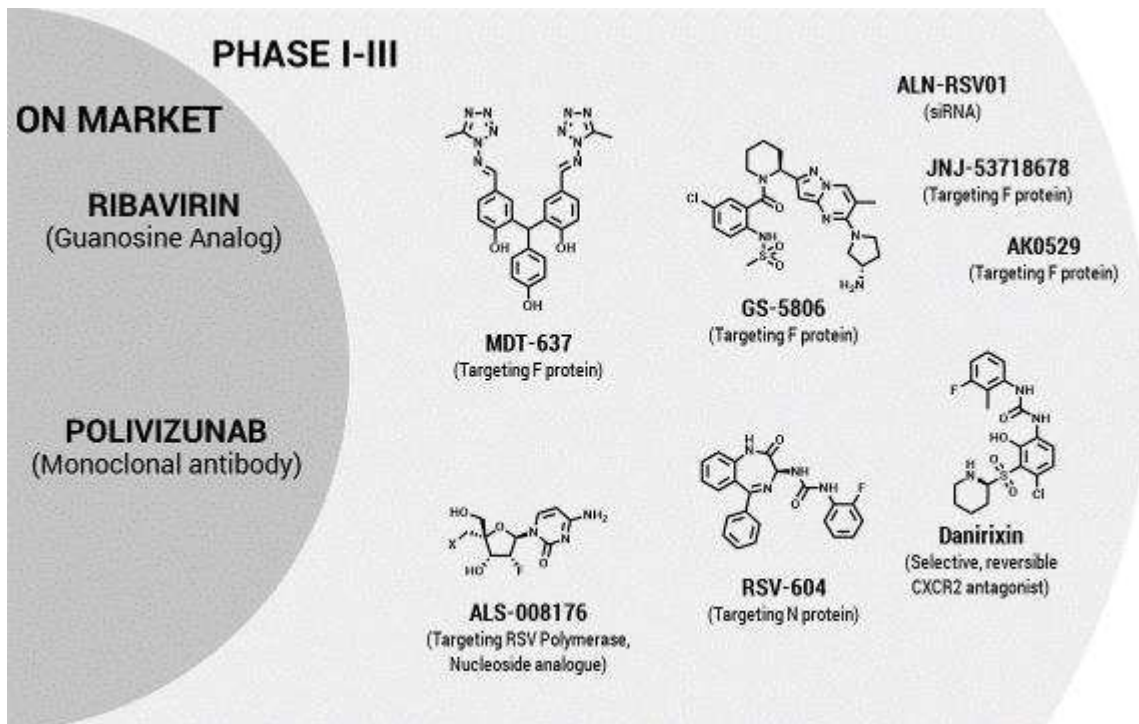


Figure 2. Available Drugs against RSV. representation of RSV drugs available on the market and drugs which are currently in clinical trial (phase I-III)

At the moment 8 antiviral drug candidates (Fig 2) and 4 immunoglobulins are in clinical trials.^{2,7-8}

The current protocol in case of severe bronchiolitis is to administer oxygen and intravenous fluids, due to the fact that in 90% of cases the disease is estimated to resolve after 21 days from the first symptoms.^{5,9}

As reported above, RSV is the most common etiological agent for LRTIs, such Bronchiolitis and Pneumonia (Bronchiolitis is the infection of the bronchioles, Pneumonia is the infection of the lung).

RSV infections in the North-Hemisphere occurred during the winter and the transmission is through person to person contact. The incubation period is 2-8 days and the shedding is approximately 8 days, but it may be 3 weeks in immunocompetent individuals or few months in immunocompromised patient.

RSV infection starts in the nasopharyngeal epithelium and then rapidly spreads to terminal bronchioles thanks to intercellular transmission. A LRTI can lead to formation of edema in the lungs, increased mucus production and epithelial cells necrosis. Necrosis of epithelial cells is followed by regeneration of the tissue that causes swelling and consequently obstruction of the small airways. The symptom of this obstruction is wheezing (a rapid and noisy breath).¹⁰

In 2010 a cohort study, a correlation between severe RSV infections in first years of life and development of long term respiratory diseases (i.e. allergic asthma and airways remodelling) in adulthood was suggested.¹¹ This remodelling was observed to be related not to the virus but to a strong immune response, that could be destructive for infected and neighbour tissues.¹²

RSV belongs to the *Paramyxoviridae* family, *Mononegavirales* order.¹³ RSV was classified as part of the *Pneumovirinae* subfamily.¹⁴ RSV is similar to the other viruses of the *Paramyxovirinae* subfamily, but its genome does not express any HN proteins (Hemagglutinin-Neuraminidase), and the mRNA of accessory proteins is transcribed by ORFs (Open Reading Frames) and is present at 3' terminus, differently from *Paramyxoviridae*.¹⁵ (Fig. 3)

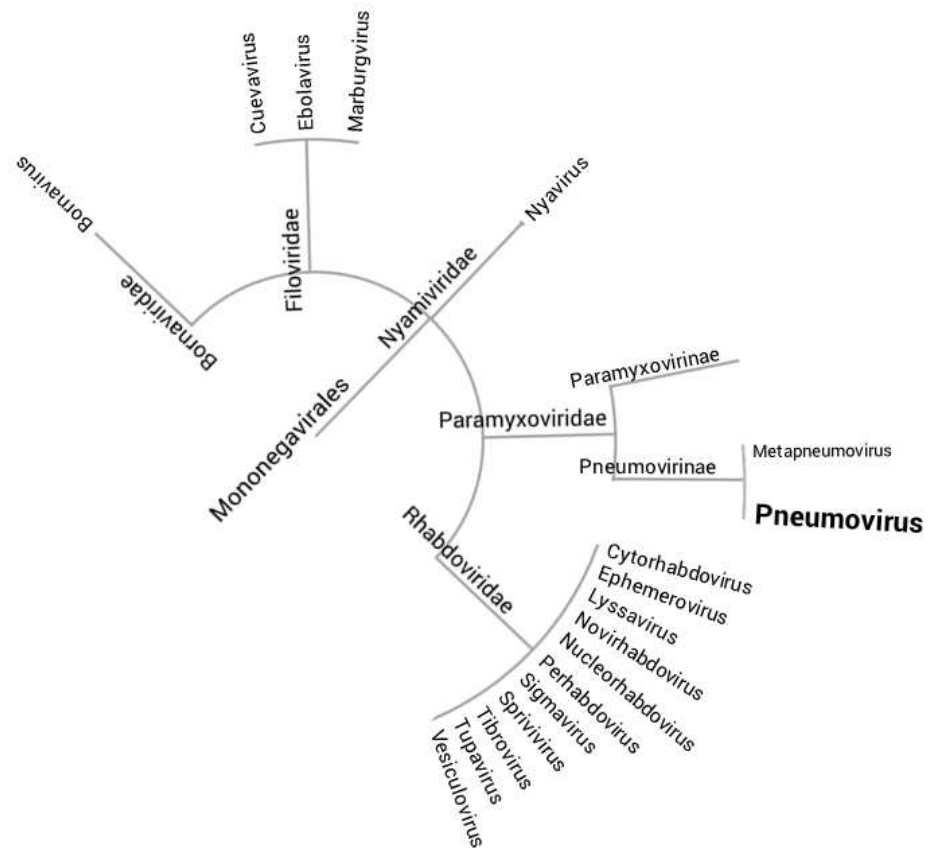


Figure 3. RSV taxonomy tree. Visual representation of the *Pneumovirinae* subfamily, inside the *Mononegavirales* order

2.1.1 RSV Virion

RSV is an enveloped (-) ssRNA virus and takes its name from the fact that it causes the formation of syncytia between infected cell membranes. The virus exists in a single serotype with 2 subgroups, strain A and strain B, which differ in the hypervariable G protein, while the F protein is antigenically similar. Strain A seems to be related to a more severe RSV disease.¹⁶

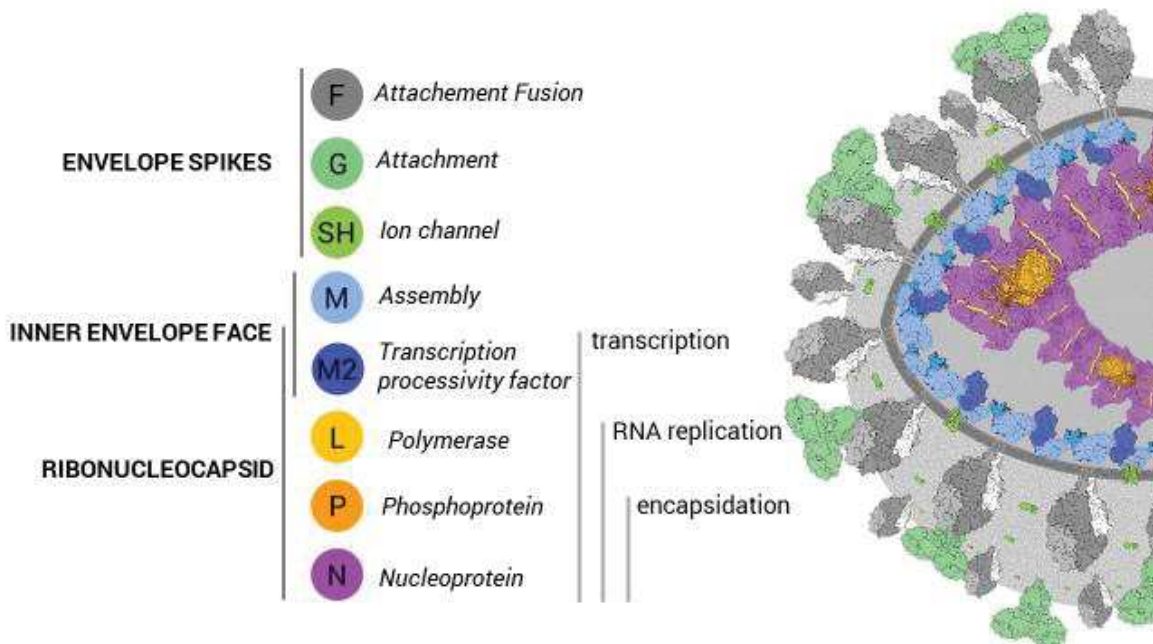


Figure 4. RSV Virion. protein composition of the virion. inside the virion the proteins can be divided into different types: envelope spikes (F, G, SH) proteins, inner envelope face (M, M2) proteins, ribonucleocapsid proteins (N, P, L, M2) proteins. The N and P proteins are essential for the encapsidation of the viral RNA and together with the L and the M2 proteins are taking part in the RNA replication and transcription process.

Only 8 of the 11 proteins encoded by the genome are present in the virion. (Fig. 4) There are 3 membrane proteins, which are considered as *envelope spikes*: F (fusion), G (major surface glycoprotein) and SH (small hydrophobic protein). Under the phospholipidic bilayer there are the M proteins (matrix), which form the *inner envelope face*. The M protein seems to bind the RNP complex (Ribonucleoprotein complex).¹⁷ RNP is a complex of (-) ssRNA and N proteins, that are important for the *encapsidation* of the viral genome; this complex forms flexible left-handed helix.¹⁸ RdRp (RNA-dependent-RNA polymerase) is composed by the L (Large polymerase) protein, and two important co-factors, the P protein (phosphoprotein) and the M2-1 protein: they form the machine necessary for the transcription and replication of the viral RNA. These 8 proteins are necessary for starting the infection.

Two of the remaining proteins are non-structural proteins (NS1 and NS2), important to counteract the IFN innate immune response. The final protein is M2-2, which is synthesized from the viral mRNA of the second ORF of M2 gene, and is a co-factor that leads the switch from transcription to replication in the RSV viral life-cycle. With 11 proteins the RSV is the more complex virus among the *Paramyxoviridae*.

From tomography analysis was observed that virion morphology ranges from filamentous to spherical, with intermediate form and the dimensions are 200 nm to 2 μ m length for filamentous virion. The width was also variable and was observed to be between 70 nm to 190 nm, with an average of 120 nm.¹⁸ (Fig. 5)



Figure 5. Virion morphology. a) filamentous shape; b) spherical shape. The pleomorphicity of the virion was linked to the amount of M proteins on the inner layer. Filamentous virions have up to 85% of the inner membrane surface covered by M proteins and the spherical virion, the M proteins inner layer is covering only the 24% of the surface.

This *pleomorphicity* of virion was linked to the amount of M proteins in the inner layer. For spherical virion, the M proteins inner layer is covering only the 24% of the surface. Whilst the filamentous virions have up to 85% of the inner membrane surface covered by M proteins.¹⁷

The reason for this *pleomorphicity*, is still unknown for RSV. Some studies suggest that this kind of difference in the virion morphology is a strategy of the virus to bluff the immune system, like Archetti bodies produced during influenza A infection.¹⁹ Another hypothesis, is that the host cells can affect the budding process, which lead to difference in morphology in harvested virion, like was observed at the beginning and late in Sendai virus infection.²⁰ Underneath the layer of M proteins there is a second layer composed of M2-1 proteins. The RNP complex runs parallel to the membrane in each virion morphology type and interacts with the M protein layer and M2-1 proteins.¹⁷ M2-1 proteins seem to interact, other than with the RNP complex, with P (Phosphoprotein) protein, and with the N-TD (N Terminal Domain) of the M (Matrix protein).

2.1.2 RSV genome organisation

The RSV genome is 15.2 kb, single stranded, non-segmented, negative sense RNA and contains 10 genes that encode for 11 viral proteins, in order 3'-NS1-NS2-N-P-M-SH-G-F-M2-L-5'. Each virion may contain from 1 to 9 copies of the viral genome, with an average of 1.8 copies *per* virion.

In comparison with other viruses of the *Paramyxovirinae* subfamily, RSV shows a more complex genome, as shown in Figure 6 with two non-structural genes in the 3' terminus of the genome and a gene that encodes for SH (Small Hydrophobic) protein.

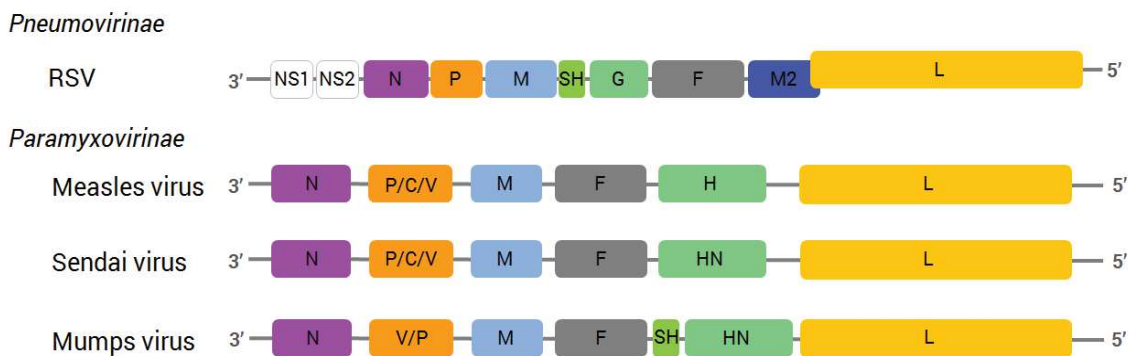


Figure 6. Genome comparison between RSV and other Paramyxovirinae. visual comparison between the RSV, Measles virus, Sendai virus and Mumps virus genomes, RSV has a more complex viral genome compared to other Paramyxoviruses.

The RSV genome is transcribed in 10 different viral mRNAs, each encoding for a single protein. The only exception is the viral RNA of the M2 gene, which contains two ORFs (Open Reading Frames), the first one that is the upstream ORF encodes for the M2-1, meanwhile the other downstream encodes for M2-2 protein. The M2-1 is already present inside the virion, while the M2-2 is synthesized later in the infection.

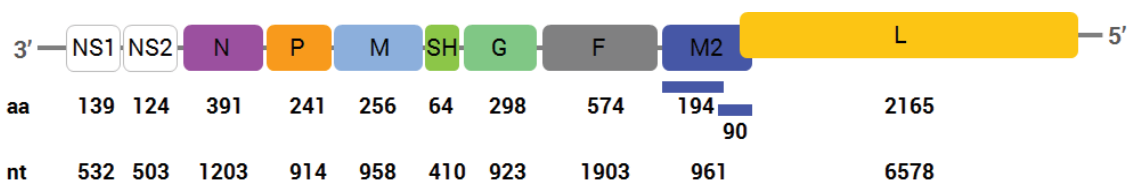


Figure 7. RSV genome. more detailed refiguration of the RSV genome and the relative protein primary sequence mRNA length. aa = amino acids, nt = nucleotides.

A peculiarity of the RSV genome is that there is only one promoter for the transcription and it is localized at 3' terminus which lies in the leader (Le) extragenic region and it is composed

of 44 nt. Another extragenic region is present at the 5' terminus; it is called trailer (Tr) and is formed by 155 nt.²¹ (Fig. 7)

Each gene contains a conserved gene-start (GS), a gene-end (GE) and an intergenic sequence (IS), which directs the polymerase during transcription process.²²⁻²³ (Fig. 8)

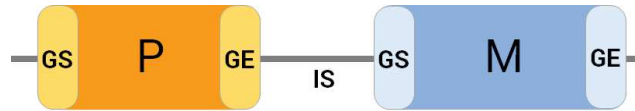


Figure 8. Gene structure. simplified refiguration of the viral RSV gene structure. Gene contains a gene-start (GS) sequence, a gene-end (GE) sequence and an intergenic sequence (IS). These sequences direct the polymerase during transcription process.

The viral polymerase has the tendency to dissociate at gene junctions, and for the reason that is present only promotor at 3' terminus, there is a gradient expression.²⁴⁻²⁵ For these reason the viral genome during evolution has put the protein required in large amount more at the beginning of the genome like for NS1 and NS2, these are important counter agent against innate immune defence of host cells.²⁶

2.1.3 RSV Life Cycle

RSV virus life cycle (Fig 9) occurs in the cytoplasm and can be summarized in three big steps:

- 1) Attachment and entry;
- 2) Transcription, translation and replication;
- 3) Assembly and budding.

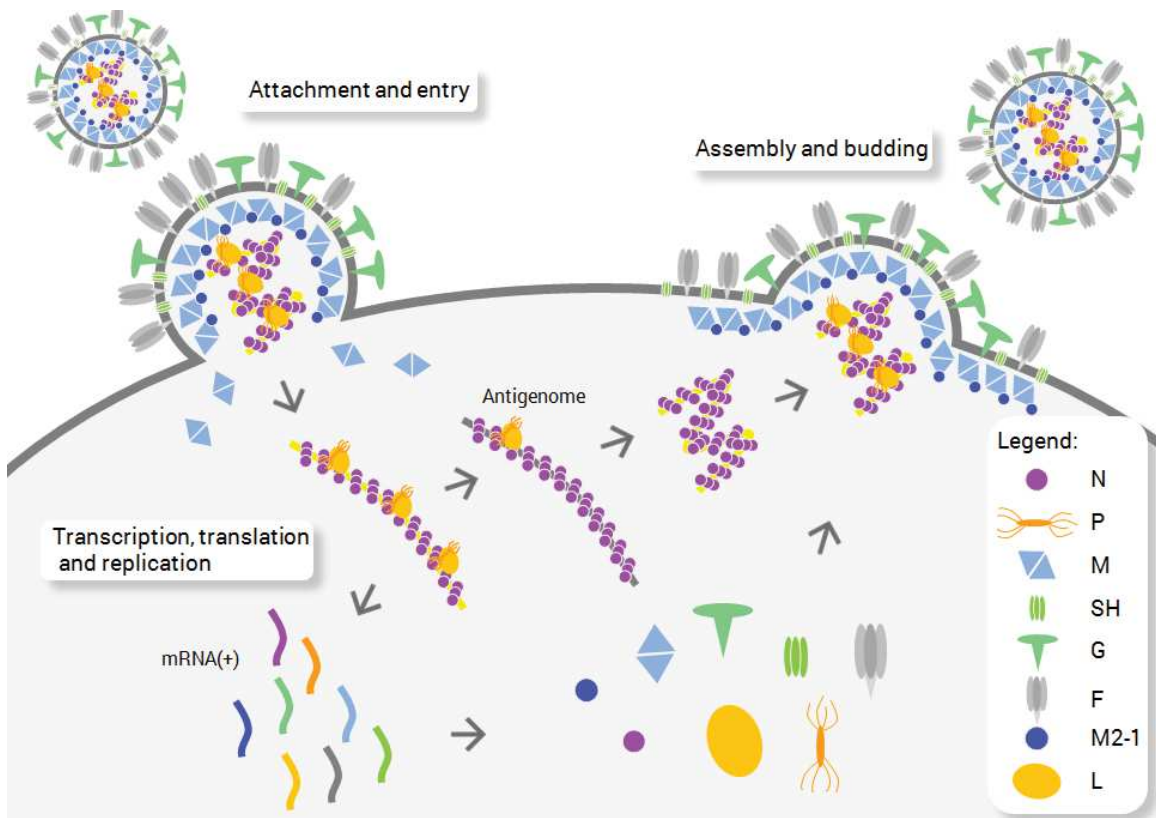


Figure 9. RSV Life cycle. The RSV life cycle can be divided into three steps: 1) Attachment and entry, 2) Transcription, translation and replication, 3) Assembly and budding. During the first step, the F, G and SH proteins are responsible for the virion attachment to the host cell and the fusion between the membranes. The fusion between the viral and the host cell membranes, assure the release of the ribonucleocapsid inside the host cell cytoplasm. During the second step of the viral replication life cycle, the RNA-dependent RNA polymerase (L protein) together with the P protein and the M2-1 and M2-2 co-factor, is responsible for the transcription, translation and replication of the viral elements necessary for the release of new viral copies. Once the viral elements and the newly synthesised viral genome are accumulated in the host cell, the new virions are assembled and released by budding process.

Attachment and entry

In each step of the viral life cycle, different proteins are covering different roles. For the attachment and entry stage, the main role is played by the *envelope spikes*, G, F and SH proteins. RSV infects specifically only ciliated epithelial cells of airways. These cells are polarized and have two different types of cell surface; the first one is called apical; it is

ciliated and exposed to the lumen. The second one is called basolateral and it is exposed to the basement membrane, which is in contact with underlying tissue. RSV entry and budding happen exclusively from the apical surface of epithelial cells.¹²

This specificity of the RSV infection is related to the viral membrane glycoproteins and their ability to bind to glycosaminoglycans (GAGs). This binding was reported to be the first step of the viral infection and was reported that RSV bind especially with HS (heparan sulfate) and CS-B (chondroitin dermatan sulfate). These two types of GAGs were found on the lung cells. GAGs are unbranched polymers of disaccharide units of glucuronic acid or iduronic acid, linked to glucosamine or galactosamine. GAGs can be modified, such as by sulfation. GAGs can be linked to core proteins via an O-glycosidic bond. HS consists of glucuronic acid or iduronic acid linked to *N*-acetylglucosamine, instead CS-B contains iduronic acid linked to *N*-acetylgalactosamine modified by sulfation on C4. The specificity in the binding of RSV to cell surface GAGs seems to be based on GAG structural configuration instead of a simple charge interaction of the virus and sulfate groups.²⁷⁻²⁸

There is a change in the percentage, during life, from CS-B to CS-C in the lungs.²⁹ This change can be related with the fact that most severe RSV LRTIs affect infants.

At the beginning, the G protein was believed to be the principal protein involved in RSV attachment to the cell membrane³⁰, but experiment on RSV mutant lacking of G, SH or both, showed that the virus is still able to infect cells.³⁰⁻³¹ This demonstrates that the F protein, which is responsible for membrane fusion, is able to bind receptors on the cell surface.

The G protein, unique among the *paramyxovirus*, is a heavily glycosylated type II transmembrane protein, which seems to bind carbohydrate receptors on cell surfaces.

The F protein is responsible for the fusion between the virion and the host cell membranes. It is highly conserved in the *Paramyxoviridae* family.³² The F protein is necessary for RSV to enter the cell. The F protein is a homotrimer expressed on the virion envelope. It is expressed as a metastable conformation, which is called pre-fusion conformation. After binding of the virion on the cell surface, the F protein undergoes a series of conformational changes that end with the final stable conformation, called post-fusion.³³⁻³⁶

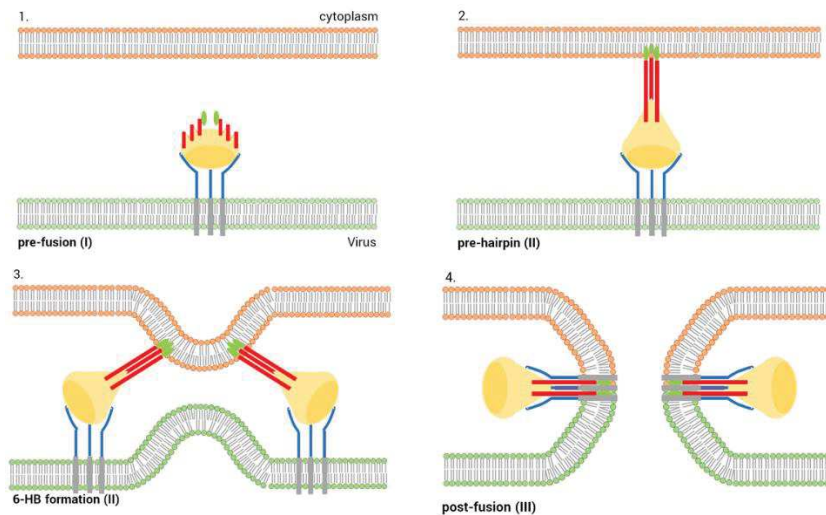


Figure 10. Fusion process. schematic representation of the fusion process performed by the F protein. HRA region is represented in blue is anchored to the viral membrane. HRB region is represented in red and it is responsible for the attachment to the cell membrane.

The F protein is characterized by three regions. HRA (heptad repeat A) at the N-TD, HRB (heptad repeat B) at the C-TD, and head of the protein in between HRA and HRB.

As shown in Figure 10, the F protein is anchored to the virion membrane through a transmembrane domain at the N-TD in proximity to HRA. The HRA (blue) is a region formed by three α -helices of the three F protein monomers. In the pre-fusion conformation, the three α -helices that form HRB are attached to the head of the F protein (I). After the attachment to the HRB region is subjected to a process of refolding, which leads to the formation of a “long trimeric coiled coil” that is able to anchor to the cell membrane (red). This conformation of the F protein is called “pre-hairpin intermediate” (II). This intermediate is highly unstable, and its collapsing is responsible for the membrane’s fusion (III). The final conformation, called post-fusion, is related with the formation of the pore between the viral and cell membranes, is possibly due to the formation of a structure called 6-HB (six helix bundle) also known as hairpin structure (IV). This structure is made by an inner core of HRA surrounded by the three α -helices of HRB. The process of fusion seems to be the results of a cooperation of more than one F proteins.³⁷

The last protein that plays an unclear role in the attachment and entry process is the SH protein. Once the fusion of the virion is completed and its content is inside the host cell, the second stage of RSV life cycle can start.

Transcription, translation and replication

With the transcription, translation and replication stage the virus produces all the elements for the production of new complete virions. Genome encapsidates in the RNP complex (viral

RNA and N proteins) as a flexible left-handed helix is the template for the viral RNA-dependent-RNA polymerase (RdRp) which is composed by L, P, M2-1 proteins.

RSV genome has some peculiarities in comparison with other viruses of the *Paramyxovirinae* subfamily:

- It has 10 genes that encode for 11 proteins, subject to a *Gradient expression*;
- It has a Le (Leader) extragenic region of 44 nt at the 3' terminus and a Tr (Trailer) extragenic region of 155 nt at the 5' terminus;^{21,38}
- Each gene presents a conserved gene-start (GS), gene-end (GE) and an intergenic sequence between genes;³⁹⁻⁴⁰
- It has only one promotor for transcription, which is contained in the Le sequence.

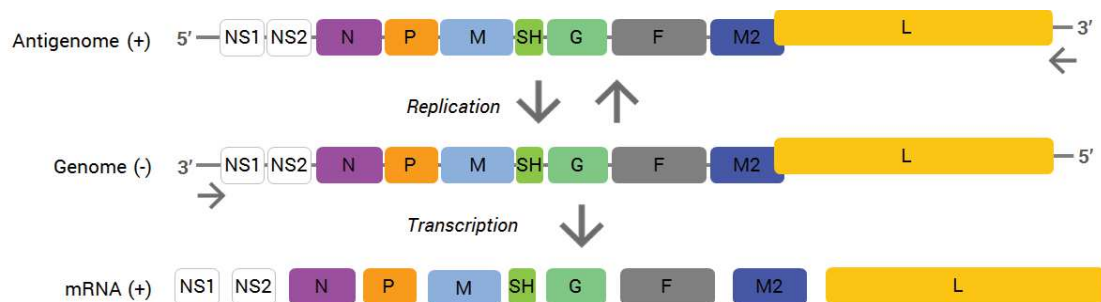


Figure 11. RSV genome transcription and replication. schematic representation of the viral genome transcription and replication performed by the RNA-dependent RNA polymerase. The antigenome is essential for the viral genome replication. The viral genome directly transcribed into mRNAs which are translated into viral proteins.

The viral genome is subject to two different processes by the RdRp: transcription and replication (Fig. 11).

The transcription is the process that the virus uses to produce viral mRNAs. These mRNAs are naked, capped and polyadenylated by the viral polymerase and they are translated by the host cell machinery.

The replication process can be split into two sub-process, the synthesis of an antigenome as template and the replication of the viral genome. Both antigenome and genome are not naked. Viral mRNAs are produced from the promoter on the Le region at 3', with a *start* and *stop* synthesis by the polymerase, which is moving along the genes from the 3' to the 5' terminus. The genome is used by the RdRp as a template for both transcription and replication.

The polymerase has to bind to the RNP, at the promotor region to be functional. Two possible modes have been postulated. In the first one, the 3' terminus and the Nucleocapsid bind to the promotor and generate a structure specific for the polymerase assembly. In the

second mode of action, the polymerase binds to the Nucleocapsid and is able to displace two N proteins from the 3' terminus and interact directly with the promoter.²²

First of all, the virus needs to synthesize all the proteins of the genome. The viral polymerase starts to produce viral sub-genomic RNA starting from the promoter leader region. As was said before the RSV genome presents a *gradient expression* which produces a difference in amount of viral mRNAs produced.

The process of transcription consists in the production of ten mono-cistronic viral mRNAs, which are capped, methylated at the 5' terminus and polyadenylated at the 3' terminus. These mRNAs are then translated from the host cell translation system.

The translation begins from the promoter and then continues through the RNA. Once the polymerase reaches the GE sequence, the newly synthesized mRNA is polyadenylated and released from the polymerase. At this point the polymerase reinitiates the translation of the next gene from the next GS sequence.

To start RSV translation, the viral genome needs three elements:

- A region at the 3' terminus of Le for the recruitment of the polymerase;
- A U rich sequence at the end of the Le;
- The first GS sequence.

The GS (CCCCGUUUUAU) sequence is highly conserved in all RSV genes, except for the GS of L gene.³⁹⁻⁴⁰ The putative function of the GS sequence is to keep the polymerase bonded for enough time to allow the nucleotide to reach the binding site of the polymerase.²²

The GE sequence is a conserved sequence which is followed by a U rich tract, this last one being responsible for the polyadenylation of the nascent mRNA and seems to be related to the process of releasing of the mRNAs from the polymerase.⁴¹ After mRNA releasing, the active site of polymerase is empty and is able to reinitiate the RNA synthesis from the next GS sequence.

The cycle of gene translation is repeated along the entire genome. Each *de novo* synthesized mRNA is translated by the host cell apparatus, which leads to the accumulation of the viral proteins in the host cell.

Once the viral mRNAs are produced, they are translated by the host cell ribosomes in the *Endoplasmic reticulum*. The envelope proteins (G, F and SH), are processed during a passage in the Golgi complex, which allows proteins to be located on the cell membrane. During the infection, there is an increased production of the M2-2 protein. This protein is the co-factor that promotes anti-genome and genome production and inhibits viral mRNA synthesis.

Since RSV genome is negative sense, for its replication the virus needs an anti-genome as template for the genome replication for the new viral progeny.

RSV needs to synthesize a positive-sense genome, called antigenome, to replicate its negative-sense genome. The polymerase is recruited at the promoter in the 3' terminus Le region. Once bonded to the genome, it starts the synthesis along the template, producing a complete positive-sense RNA complement to the RSV genome.

The antigenome has at the 3' terminus the complementary sequence of the Tr, which is 155nt. This sequence contains the promoter for the synthesis of new complete viral genomes and has a higher level of replication than the Le region, producing more replicated genomes than antigenomes.^{22,42} The antigenome is used as template for the synthesis of new copies of the viral genome.

The antigenome and the new copies of the genome are immediately encapsidated. The N protein is able to bind unspecifically to the viral RNA, and protein-protein interactions (PPIs) between N proteins are sufficient for the RNP assembly.⁴³⁻⁴⁴

Assembly and budding

After the virus enters the host cell, replicates its genome and produces all the proteins for the new progeny, new virions are created by budding from the plasma membrane.

During the infection, the formation of Inclusion Bodies (IBs) in the cytoplasm has been observed: these bodies are made by accumulation of viral polymerase and other viral proteins, such as N protein, M protein and non-structural proteins (NS1 and NS2).⁴⁵⁻⁴⁷ These IBs are associated with lipid-raft structures, which are micro-domains in the plasma membrane with high amount of cholesterol and sphingolipids. These lipid rafts seem to be important in the RSV life cycle. All the life cycle seems to turn around this host membrane domain.⁴⁸

RSV virions bind and bud from the apical membrane of polarized epithelial cells of airways tissues; for this reason, RSV infection remain localised in the low respiratory tract. From mutant RSV virions lacking the F and G membrane glycoproteins, the only protein responsible for the specific budding on the apical surface has been identified as the matrix protein (M).⁴⁹⁻⁵⁰ The M protein later in the infection, interact with the cytoplasmic domain of F and G, but this interaction is not responsible for the M protein specificity for the apical membrane. The M protein is believed to bind a specific zone in the plasma membrane which is called lipid raft. This is a microdomain in the lipid bilayer enriched of cholesterol and sphingolipid and seems to be important for the accumulation of viral proteins necessary for viral budding.⁵¹⁻⁵⁴

This leads to the formation of a layer of M proteins that are the anchor point for the encapsidated viral genome (RNP). On the RNP is present L, which is attached to L through the P and M2-1 protein. Once the assembly process is finished and all the required proteins

are located near the membrane, the new virus can bud from the cell surface. This process seems to be carried out by the M protein *inner layer*.⁵⁵

After the assembly of the viral proteins on the inner space of the plasma membrane, the budding is probably triggered by a rearrangement of the M proteins inner layer, such as for the budding of Ebola virus.⁵⁶

2.2 Viral proteins

Envelope spikes

G protein

The G protein is a unique binding protein, different from Hemagglutinins and Hemagglutinin-Neuraminidase among other *Paramyxoviruses*.

The G protein is a 298 aa long type II glycoprotein with a transmembrane sequence at the N-TD (N terminal domain). The G protein ectodomain is post-translationally glycosylated, for the majority with O-oligosaccharide and a small amount of N- oligosaccharide.⁵⁷

In addition, a soluble form of the G protein (GS) is released from infected cells: this seems to be the product of a second highly conserved AUG codon on the G ORF, with the transmembrane domain proteolytically removed.⁵⁸ The function of this soluble form of the G protein is to inhibit the inflammatory cytokines process at the first stage of the infection and as “antigen decoy” to help the virus escaping the immune system.⁵⁹⁻⁶⁰

SH protein

The SH protein is a 65 aa long protein, which forms a transmembrane homopentamer. This protein has a channel function that enables ions and small compounds to pass through the membranes. The SH protein has been found on the plasma membrane and in the Golgi complex.⁶¹⁻⁶² These evidences allows the SH protein to be classified as a viroporin. The SH function is unknown but it is believed to cause the alteration of ionic homeostasis and depolarization of the host cell membranes.⁶³

F protein

The RSV F protein is a type I glycoprotein and is produced as a 574 aa precursor (F0), which is cleaved twice by furin in the process of transportation on the plasma membrane. After the cleavage two proteins are generated, F2 and F1, connected by two disulfide bonds. This cleavage is important to allow the F protein to fuse with the cell membrane. The active F protein, ready for fusion, is a homotrimer of the processed F0.³⁷

Ribonucleocapsid***N protein***

The entire viral genome is coated by N proteins called Ribonucleoprotein (RNP), and it remains coated during transcription and replication, meaning that the RNP is the template for the RdRp. The RNP forms a left handed helix, which is extremely flexible; this flexibility might allow the RNP to be read from the large RdRp, without necessarily removing the N protein from the viral genome.⁴³ This strategy avoids the formation of secondary RNA structures and the possibility to form dsRNA, which could trigger the innate immune response.

RNA-dependent-RNA polymerase***L protein***

The polymerase machine is made by the L protein. In the infected cell, the virus needs to interact with other viral and host proteins, for an efficient and complete RNA synthesis activity.⁶⁴⁻⁶⁵

The L protein is a 2165 aa long multifunctional protein, with different enzymatic domains. This protein is able to identify the promotor sequence, perform RNA synthesis, perform capping and methylation of the 5' terminus of mRNAs and polyadenylate them at 3' terminus. Six conserved regions (I-IV) are postulated to perform the enzymatic activities of the L protein. Some of these regions are connected to specific activities. Regions II and III are believed to form the polymerization domain.⁶⁶⁻⁶⁷

The region V is supposed to be responsible for the guanylyl transferase activity and region IV is believed to have methyltransferase activity.⁶⁸⁻⁶⁹

P protein

The P protein, is a 241 aa long protein, which forms homotetramers in solution and is classified as an intrinsically disordered protein (IDP).⁷⁰ This protein is essential for the correct function of the RdRp activities, and it is responsible for the ability of the polymerase to bind to RNP and move along the encapsidated genome.⁷¹⁻⁷⁶ The P tetramer has a central core of four-helix bundle tetramerization (119–160 aa), a C-TD (161–241 aa) and a N-TD (1-103 aa). These last two are intrinsically disordered regions and are responsible for the binding to other viral proteins.⁷⁷

M2-1 protein

M2-1 is a 194 aa long, basic protein that forms homotetramers in solution.⁷⁸ M2-1 is considered a transcriptase processivity factor that is able to inhibit the termination on

intragenic sequence, helping the polymerase to read the GE and reducing the transcription gradient.⁷⁹⁻⁸¹

Studies on intracellular levels of M2-1 showed that M2-1 is not involved in the switch of the polymerase from the translation process to the replication process.⁸²⁻⁸³ The M2-1 interacts with the P protein and binds specifically to RNA sequences at the Le region, proving that it is loaded on the RNA during the initiation of the transcription process.⁸⁴⁻⁸⁵

M2-2 protein

The M2-2 protein is a 90 aa long protein, which is translated from the second ORF of the M2 gene. This is the co-factor responsible for the switch of the polymerase activity. It has a potent inhibition effect on the RSV RNA translation, possibly explaining the fact that this protein is not present in the virion and is not present in the first stage of the infection.⁸⁶⁻⁸⁸

M protein

The RSV M protein is 250 aa long, and forms homodimers in solution⁸⁹, which represent the active functional proteins for the assembly and the budding process.⁵⁶ The M protein is expressed in high amount in the infected cell and it is able to act as linker between the other proteins necessary for new virions. The Paramyxovirus Matrix protein has also been observed in the nucleus of the infected cells early in the infection, while later the protein is detectable in the cytoplasm.⁹⁰⁻⁹¹ The transient localization of the Matrix protein in the nucleus is believed to be related with the inhibition of nuclear processes or with the prevention of budding of uncomplete virions in early stages of the infection.⁹² Late during the infection a high amount of M protein is required for assembly and budding processes.^{55,93}

2.3 Project aims

Two main targets among RSV proteins have been chosen to identify new inhibitors of the viral replication:

- N protein;
- F protein;

The N protein, which is part of the ribonucleocapsid complex (RNP), plays a key role in virus replication, acting as a template for transcription and replication, performed by the RdRp (RNA-dependent RNA polymerase). Interactions between the RNP and P protein, which is part of the RdRp, have been reported to be fundamental for viral RNA synthesis.

For the N protein, the aim is to design a novel series of anti-RSV compounds using different *in silico* approaches. The attention has been focused on a druggable pocket in the N protein involved in PPIs (protein-protein interaction) with the C-TND of the P protein.⁹⁴, several computer-aided drug design methodologies will be applied to identify and synthesise small molecule as potential anti-RSV agents.

Below, the activities performed during the project on the RSV N protein:

- starting from five available X-ray structures of this protein with the respective ligands molecular modelling techniques were used for the identification of new potential inhibitors;
- Several compounds were identified as potential viral inhibitors through the molecular modelling techniques
- Among the identified compounds, four hits were confirmed in antiviral assays;
- several analogues of the two most potent hits were synthesised for Structure-Activity Relationships (SARs) exploration;
- The synthesised analogues were biologically evaluated.

In the case of the F protein, which is the viral protein responsible the fusion of the viral and cell membranes, the objective was to design α -helix mimics of the HRB. The α -helix mimics compounds are designed to inhibit the assembly of the 6-HB (6-helix-bundle), an important step in the fusion mechanism of the virus for the infection of the cell.

Below, the activities performed during the project on the RSV F protein:

- starting from five available X-ray structure of the F protein in the post-fusion conformation, different cheminformatics tools and molecular modelling techniques were used to design, evaluate and select an α -helix mimic;
- The selected α -helix mimic was synthesised and the optimised synthetic route was used to synthesise a restricted number of analogues.
- All the synthesised compounds and the key intermediates were biologically evaluated.

Chapter 3: N protein

3.1 N protein

RSV N protein is responsible for the encapsidation of the viral genome: it acts as RNP complex, which forms a left-handed helix, which is the template for RNA transcription and replication (Fig. 12). Both these steps need the polymerase to bind to the RNP; this interaction is mediated by the action of the Phosphoprotein (P), binds to N proteins of RNP by the CTD of the P protein. On the other hand, each N protein is organized in four subunits, two NTDs (31 to 252 aa) and CTDs (253 to 360 aa) globular domains and two NTD (31 to 252 aa) and CTD (362 to 391 aa) extensions. The N-P interaction is critical for the viral transcription and replication.

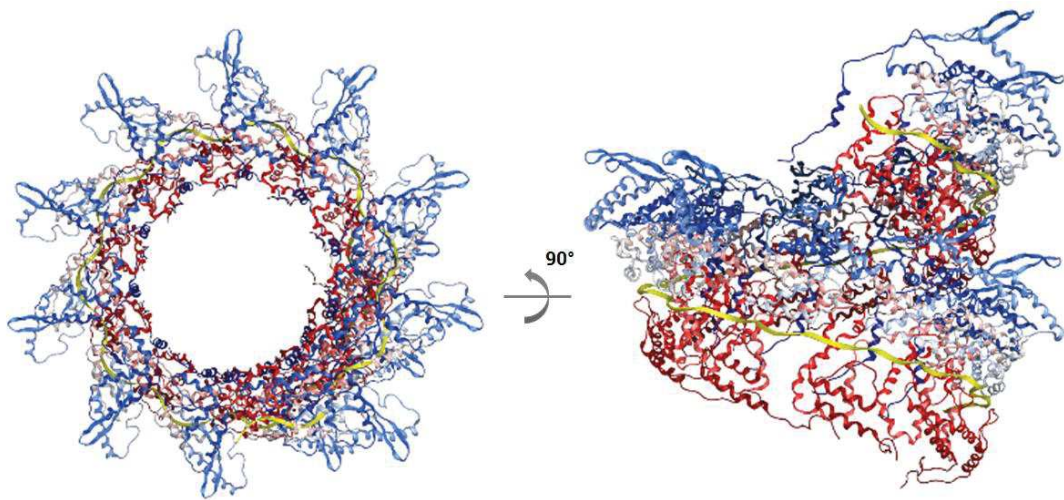


Figure 12. RNP resolved structure. ribbon representation of the ribonucleoprotein (RNP) complex. The N-NTDs are represented as blue ribbons and the N-CTDs are represented as red ribbons. The viral RNA is represented as a yellow ribbon. (Pdb: 4BKK)

In 2012 a pocket was discovered on the N-NTD that is responsible for the interaction with the P-CTD. This pocket is a hydrophobic site surrounded by positively charged residues. From mutagenic analyses, 10 residues were identified as important in the N-P PPI (protein-protein interaction): K46, M50, E128, S131, R132, Y135, K136, R150, H151 and D152. Furthermore, four critical residues of P-CTD were identified as important for the N-P PPI: L238, E239, D240 and F241.⁹⁵

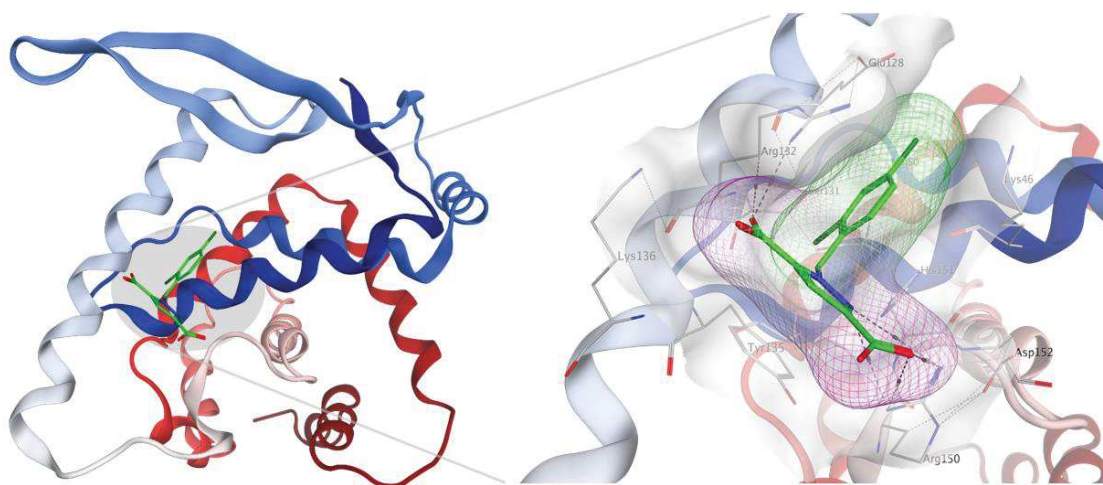


Figure 13. Binding pocket on N-NTD. On the left side, ribbon representation of the N-NTD in complex with the N-P interaction inhibitor M76, with 1-benzyl-1H-pyrazole-3,5-dicarboxylate (BPdC) scaffold (Pdb: 4UCC). On the right side of the figure, a more detailed view of the binding pocket

This information has been confirmed by the resolution of the X-ray crystal structure of N-NTD:P-CTD interaction (Pdb: 4UC8 and 4UC9); only the dipeptide D240-F241 of P-CTD is responsible for the binding with the N-NTD. In particular, residue F241 has a key role in the interaction.⁹⁴ Five compounds with a common 1-benzyl-1H-pyrazole-3,5-dicarboxylate (BPdC) scaffold were found active to inhibit the N-P interaction and to inhibit the RSV viral infection in a cell-based assay (Fig. 13). For three of them a crystal structure in complex with the protein has been resolved (Pdb: 4UCC, 4UCD and 4UCE). All this information was used as a starting point for computer-aided techniques aiming the identification of potential new inhibitors of the N-P interaction.

3.1.1 Structure-based virtual screening

Starting from the crystal structure of N-NTD with the most active compound of the BPdC series (Pdb: 4UCC), a structure-based virtual screening was performed.

Docking of the co-crystallised ligand was performed using Glide SP.⁹⁶⁻⁹⁸ The docked ligand shows the same orientation inside the pocket found in the crystal, with a RMSD (Root-mean-square deviation) of 0.769 Å. This result validates this program for performing a virtual screening.

For the virtual screening (VS), the SPECS library of 342047 commercially available compounds was used.⁹⁹ The VS algorithm (Schrodinger Maestro) generated 906454 poses, from which the top scored 20% of them (181290 poses) were docked using Glide SP. 515038 poses were obtained, which were then rescored using Glide XP, Plants and FlexX.

After combining the results of the rescoring (*consensus scoring*, see Experimental), the poses with the highest ranking were selected and 4481 compounds were then clustered using MACCS Fingerprints and Tanimoto similarity matrix in MOE. In this way, the more dissimilar compound structures were selected, obtaining 2705 compounds, which were analysed through a visual inspection of their predicted binding mode. 24 compounds were finally selected based on their ability to bind deeply inside the hydrophobic pocket and on the basis of the number of interactions (Fig. 14). The selected compounds were bought from the SPECS company and sent for testing in cell-based assays. The compounds structures are showed in Table 1 in the appendix.

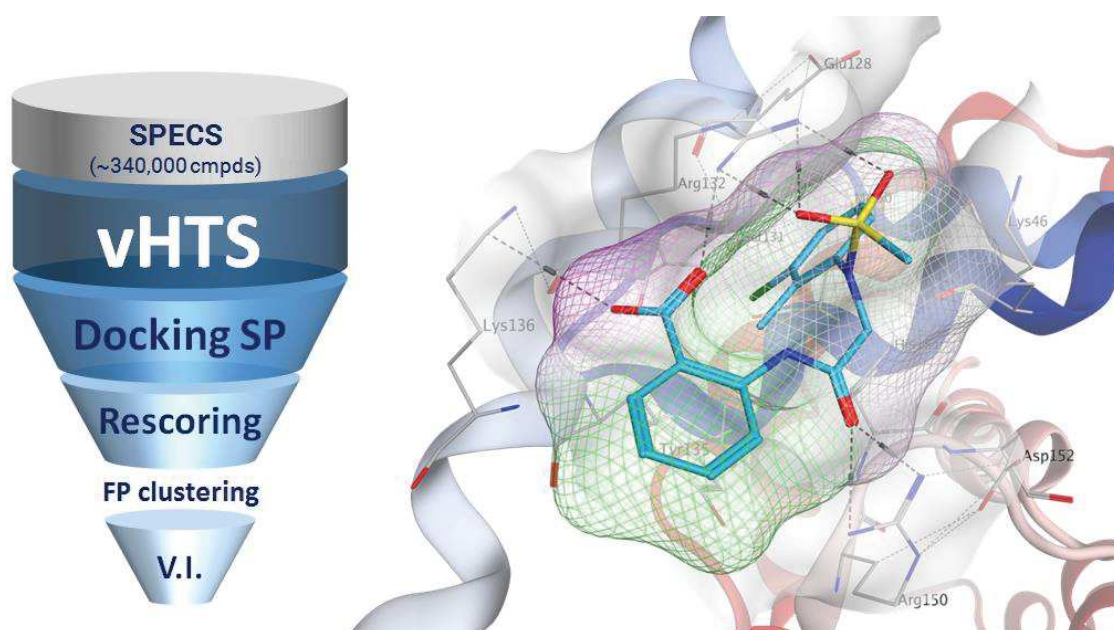


Figure 14. *vHTS workflow and docked pose of a selected compound. Different molecular modelling techniques were employed during the virtual high throughput screening (vHTS) for the identification of new potential N-P interaction inhibitors (left side); On the right side of the figure, the docking pose of one of the selected compounds and its interaction with the binding pocket. (FP clustering = Fingerprint clustering; V.I. = visual inspection)*

3.1.2 Pharmacophore search

A hybrid structure-ligand-based approach was used as well to study the N protein. The information contained in the five available crystal structures was used for the generation of a pharmacophore model (PDB IDs: 4UC8, 4UC9, 4UCC, 4UCD, 4UCE). This model was generated using the PLIFs (Protein-Ligand interaction Fingerprints) tool in MOE.¹⁰⁰ This is a method for summarising the interactions between ligands and proteins using a fingerprint scheme, which is representative of a given database of protein-ligand complexes. From the fingerprints it is possible to generate a pharmacophore model combining all the structural information in the dataset.

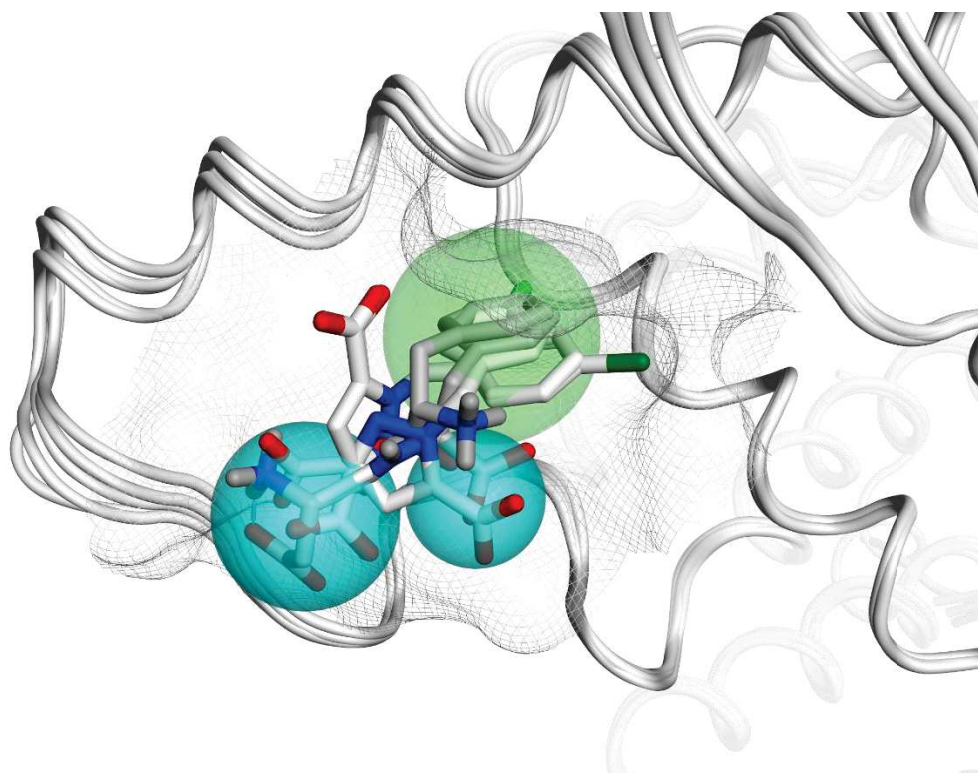


Figure 15. Pharmacophoric filter generated with PLIFs. graphical representation of the pharmacophoric model built using the Protein Ligand interaction fingerprints (PLIFs) technique. The cyan spheres represent H-bond acceptor, heavy atom and anionic heavy atom features and in green a hydrophobic feature.

A pharmacophoric model containing three features was generated (Fig. 15). F1 and F3 (annotated as Anion, H-bond Acceptor and Metal Ligator), are H-bond acceptor (cyan), heavy atom and anionic heavy atom features. The last one, F2 (annotated as Hydrophobic Atom) is a hydrophobic group (green).

This pharmacophore model was used for a pharmacophore search on 1893330 compounds from commercially available libraries (SPECS, ENAMINE and Life Chemicals).¹⁰¹⁻¹⁰² From this search, 65871 compounds were obtained that possess all of the three requested

features. The matching compounds were docked using Glide SP algorithm in a 12 Å radius grid generated, from the 4UCC crystal structure.

The SP docking generated 429753 poses, which were then rescored using Glide XP, Plants and FlexX.

The poses with the best results in the *consensus scoring* were selected and 8918 compounds were obtained. The number of compounds was reduced to 6682 by the elimination of structures with a logD less than 2, and the remaining molecules were clustered using MACCS Fingerprints. From the Tanimoto similarity matrix the more different structures were selected. This process led to 2533 compounds, which were analysed through a visual inspection in the binding pocket. In the end, 20 compounds were selected based on their ability of compounds to bind deeply inside the hydrophobic pocket and based on the number of interactions (Fig. 16). The selected compounds (Table 2 in the appendix) were bought and tested in a cell-based assay.

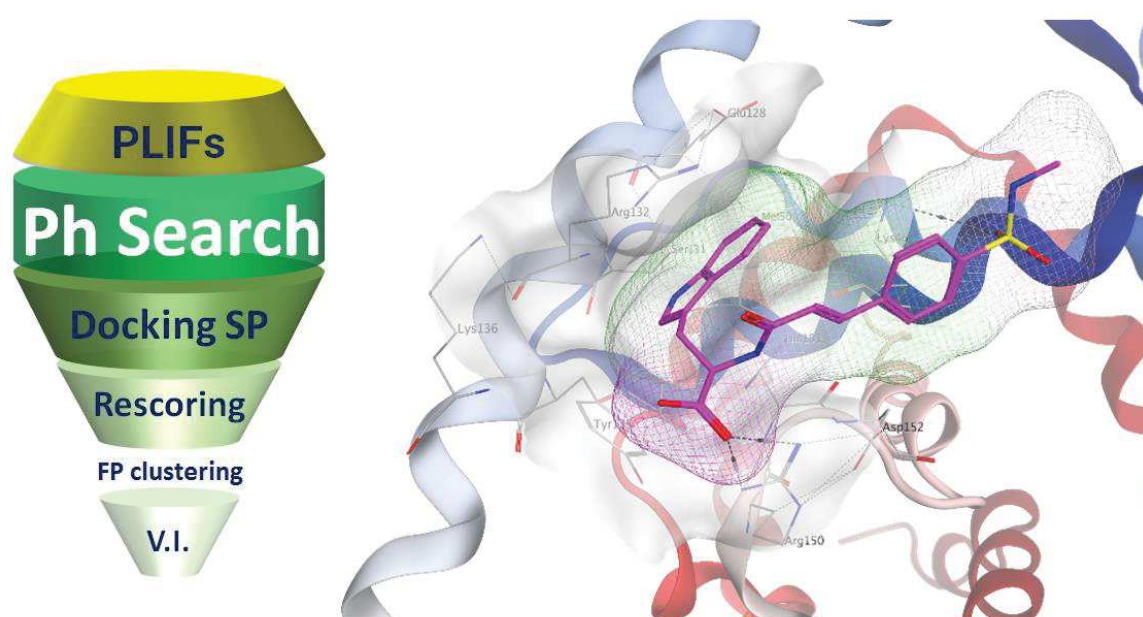


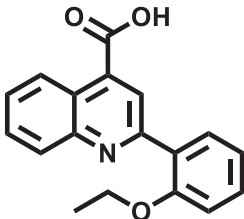
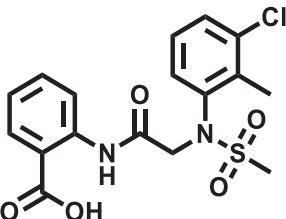
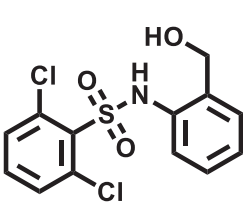
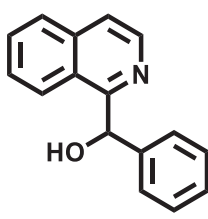
Figure 16. Pharmacophore search workflow and docked pose of a selected compound. schematic representation of the different molecular modelling techniques employed during the Pharmacophore search (Ph search) workflow for the identification of new potential N-P interaction inhibitors (left side); On the right side of the figure, the docking pose of one of the selected compounds and its interaction with the binding pocket. (PLIFs = Protein- Ligand Interaction fingerprints; FP clustering = Fingerprint clustering; V.I. = visual inspection)

3.1.3 Biological evaluation of selected compounds

The compounds selected through the SBDD approaches for their ability to fit and bind inside the N-NTD pocket were tested in the Rega Institute for Medical Research, KU Leuven, Leuven, Belgium, under the supervision of Professor Johan Neyts. The compounds were evaluated in Vero cells co-infected with RSV A2-GFP-1 strain and were evaluated for their ability to inhibit the virus replication. In this assay, the Vero cells are infected with an engineered RSV-A2 virus, in which the green fluorescent protein (GFP) gene was inserted in front of the NS1 gene. This gene is encoding for the GFP, which is a protein used as a reporter of the viral expression. The GFP produced inside the infected cells can give a direct observation of the viral replication, and it is used to evaluate the antiviral activity of the tested molecules during the assay.¹⁰³

From the antiviral evaluation of the selected compounds, four hits were identified in the bioassay to be able to inhibit completely the RSV A2 strain replication (see table 1).

Table 1. Hits identified. Chemical structures of the identified hits. These compounds were found to have antiviral activity on RSV A2-GFP-1 strain. The compounds were tested on cell based antiviral assay (AVA). Data are representative of two independent experiments.

				
	1	2	3	4
EC₅₀ (μM)	64	94	139	318

The most potent compounds, compounds **1** and **2**, were selected as a starting point for further investigations.

3.1.4 M76 - 1-(2,4-dichlorobenzyl)-1H-pyrazole-3,5-dicarboxylic acid

The most potent inhibitor of the 1-benzyl-1H-pyrazole-3,5-dicarboxylate (BPdC) series, 1-(2,4-dichlorobenzyl)-1H-pyrazole-3,5-dicarboxylic acid, M76 (**5**, figure 17), was reported to be able to bind to the N-NTD pocket (Pdb. 4UCC) and to inhibit the RSV replication.⁹⁴ M76 was synthesised to be used as internal reference in the biological evaluation.

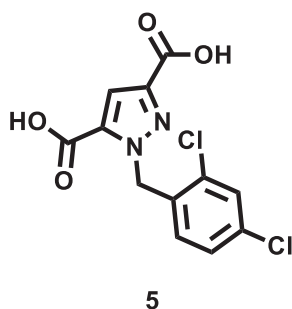
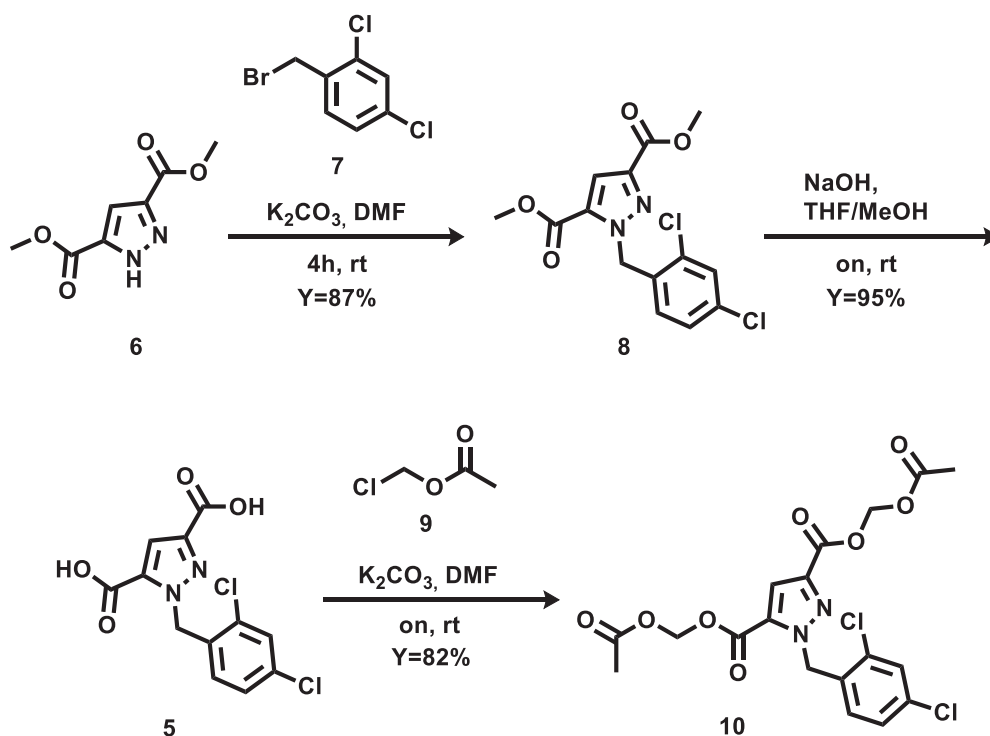


Figure 17. Chemical structure of M76.

M76 was tested in the cell-based assay as acetoxymethyl (AM) esters prodrug to hinder the two carboxylic groups, because the free di-carboxylic acid was reported to be inactive, due to the double negative charge of the compound that could interfere with membrane permeability. For this reason, compound **5** was synthesised as di-AM and as free di-carboxylic acid, along with the di-methyl ester. M76 (**5**) and its derivatives were synthesised through the synthetic pathway reported below (Scheme 1).¹⁰⁴⁻¹⁰⁵



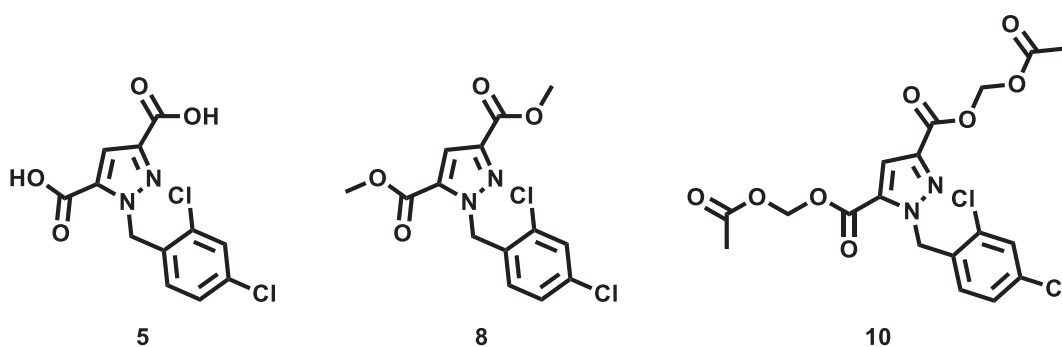
Scheme 1. Synthesis of M76.

The synthesised compounds were biologically evaluated in a cytopathic effect (CPE) antiviral assay in Hep2 cells line, infected with RSV-A2 and RSV-B. The CPE refers to the structural modification of the infected cells by the virus. In the antiviral assay, the CPE is microscopically scored after five days of incubation. The score is used to calculate the EC₅₀ of the tested compounds. The cytotoxicity of the compound was evaluated by quantifying the metabolic activity of uninfected cells treated with the compound. The metabolic activity

was quantified after five days of incubation, using MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium). The reduction of the tetrazolium ring NAD(P)H-dependent cellular oxidoreductase enzymes by the cellular NAD(P)H-dependent cellular oxidoreductase enzymes is giving the formation of a formazan product. The formazan product has an absorbance maximum at 490–500 nm in phosphate-buffered saline, which is used as readout for the cellular metabolic activity.¹⁰⁶⁻¹⁰⁷

The results for these compounds were in line with the data previously reported by Ouizougun-Oubari and are showed in Table 2.

Table 2. Biological evaluation of M76 (5), dimethyl ester derivatives of M76 and the diAM-M76 prodrug. The compounds were biologically evaluated in cell based antiviral assay (AVA) against RSV-A2 and RSV-B strains and for they cytotoxicity. Data are representative of two independent experiments.



Compound		5	8	10
EC ₅₀ (μM)	RSV-A2	>400	>109	>116
	RSV-B	ND	ND	ND
CC ₅₀ (μM)		>400	118	14

The prodrug of M76 showed a higher cytotoxicity in comparison with the reported value, due to the nature of the prodrug which is releasing of formaldehyde and to the different set up of the biological assay, which was two days longer.⁹⁴

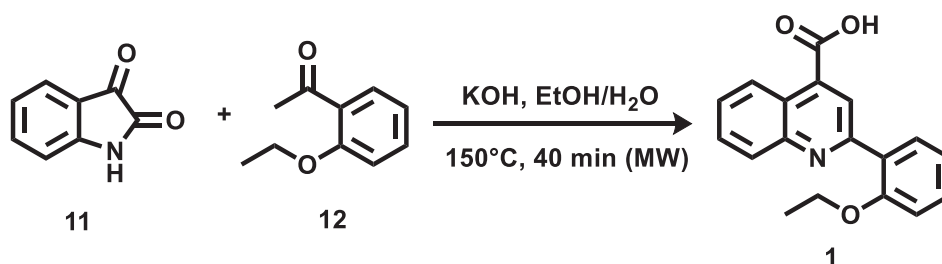
All the compound synthesised in this project were biological evaluated in HEP2 cell line against RSV-A2, RSV-B in a 5 days long CPE antiviral assay and MTS cytotoxicity assay, as discussed previously.

Chapter 4: 2-(2-ethoxyphenyl)quinoline-4-carboxylic acid and 2-(2-(*N*-(3-chloro-2-methylphenyl) methylsulfonamido) acetamido)benzoic acid

4.1 Synthesis and biological evaluation of 2-(2-ethoxyphenyl)quinoline-4-carboxylic acid (1)

To confirm that the antiviral activity of hit **1** is related to the chemical structure provided from the vendor, this analogue was re-synthesised. Several attempts were performed to optimise its preparation and purification. In literature, the Pfitzinger reaction for the synthesis of this scaffold is reported to be performed in ethanol or in a mixture ethanol/water.¹⁰⁸⁻¹⁰⁹ Different attempts were performed and a system of ethanol/water with a 7:1 ratio was chosen as the best system of solvents. The chosen system is ensuring the solubility of the starting materials.

1 was synthesised through an optimised microwave-assisted Pfitzinger reaction (Scheme 2), between isatin **11** and 2-ethoxyacetophenone **12** in basic conditions that provide the compound with a purity >95% and a yield of 50%.



Scheme 2. Synthesis of compound 1.

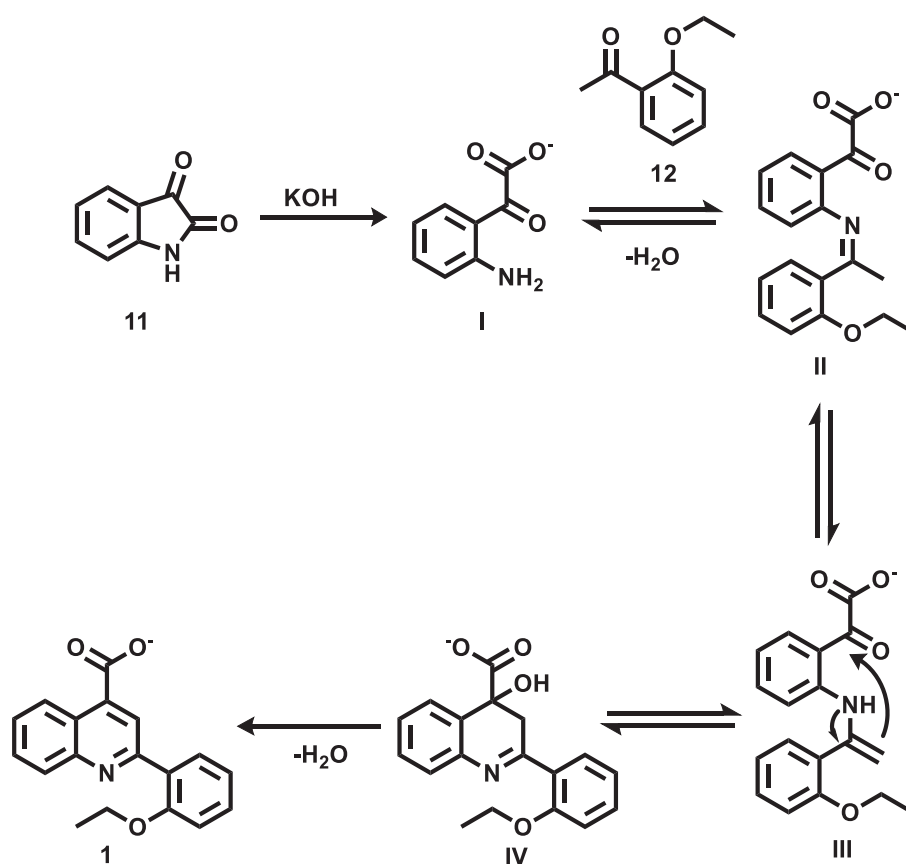
The temperature for the reaction was chosen after performing several attempts. The yields for the different temperatures explored are shown in Table 3.

Table 3. Different attempts for the selection of reaction conditions. (*T* = temperature; *t* = reaction time; *Y* = reaction yield; *MW* = microwave)

T(°C)	t(min)	Y(%)
reflux	36h	23
100 (MW)	60	45
125 (MW)	20	16
150 (MW)	20	32
150 (MW)	40	50

In the Pfitzinger reaction (scheme 3), the basic hydrolysis of the isatin ring leads to the formation of 2-(2-aminophenyl)-2-oxoacetic acid intermediate **I**. At this stage of the reaction, the amine group of **I** is reacting with the carbonyl group of 2-ethoxyacetophenone which is forming the imine intermediate **II**, which is in equilibrium with the enamine intermediate **III**.

Intermediate **III** undergoes a cyclization step (**IV**) and then dehydration gives the final compound **1**.

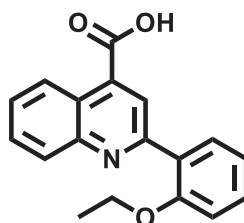


Scheme 3. Pfitzinger reaction.

Several attempts were performed to find better conditions for the synthesis of **1** and for its purification.

The antiviral activity against RSV-A2 of compound **1** was confirmed in HEP2 cells. The compound was also active against RSV-B, showing an EC₅₀ of 15 μM. Information on cytotoxicity (CC₅₀=164 μM) confirms that the activity of the hit is not related to its toxicity. Another critical parameter used to express the efficacy of a drug is the Selectivity Index (SI), which represent the distance between the Antiviral activity (AVA) and the cytotoxicity effect (TOX) of the compound on the cell. The SI is calculated dividing the AVA (EC₅₀) and the TOX (CC₅₀). An ideal compound should have a high SI value, which means that the compound is showing an antiviral effect at a concentration that was much below the concentration in which the compound is showing cytotoxicity.¹¹⁰

Table 4. Biological evaluation of compound 1. Compound 1 was biologically evaluated in cell based antiviral assay (AVA) against RSV-A2 and RSV-B strains and for its cytotoxicity. In brackets the number of independent experiments. SI = selectivity index (CC_{50}/EC_{50})



1

EC₅₀ (μM)	RSV-A2	43 (8)
	RSV-B	15 (2)
CC₅₀ (μM)		164 (4)
SI	RSV-A2	3.8
	RSV-B	11

The compound showed a Selectivity Index (SI) of 3.8 between the antiviral activity and the cytotoxicity against RSV-A2 and an SI of 11 against RSV-B. Starting from these results was decided to investigate further compound 1.

4.2 SAR of 2-(2-ethoxyphenyl)quinoline-4-carboxylic acid (1)

From the docking pose of **1** (Figure 18.a), two major features of the compound are interacting with the N-NTD pocket: the carboxylic group and the 2-ethoxyphenyl moiety. The carboxylic group is interacting with the Arg150 and the 2-ethoxyphenyl moiety is fitting inside the hydrophobic pocket of the N-NTD. The quinolinic ring seems not to make major interactions with the protein but to provide the right orientation for the phenyl moiety.

To understand better the chemical structure relevance for activity, three SAR analyses were designed to investigate the possibility of a further development of the selected hit as anti-RSV agent (Figure 18.b).

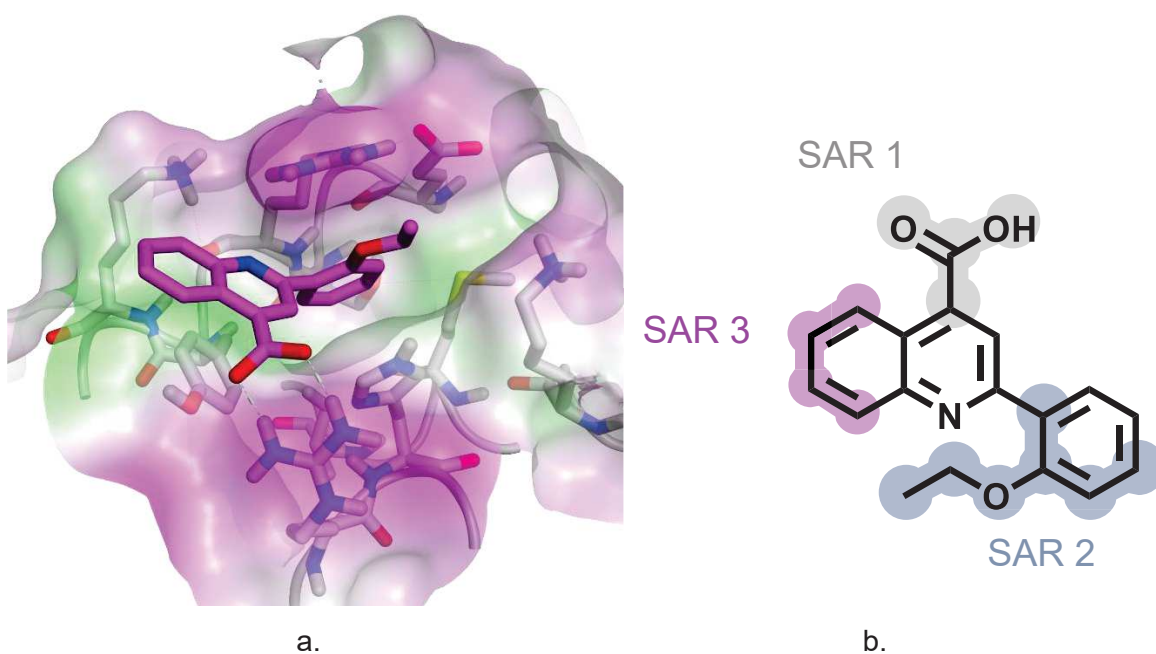


Figure 18. a. Docking pose of compound 1; b. SAR analysis

In SAR1 the interest is to understand the possibility to remove the carboxylic group or to convert it to another chemical group that can lead to an improved membrane permeability. SAR2 was designed to understand the importance of the 2-ethoxy group on the phenyl moiety by removing or replacing it with other similar groups or by changing its position on the phenyl ring. The last SAR analysis is focused on exploring the possibility of increasing the activity of the scaffold by modifying positions 6 and 7 on the quinolone ring.

4.2.1 SAR on the carboxylic group

In this first SAR analysis, several analogues were synthesised with the aim to understand the importance of the carboxylic group for the interaction with the N-NTD pocket (figure 19).

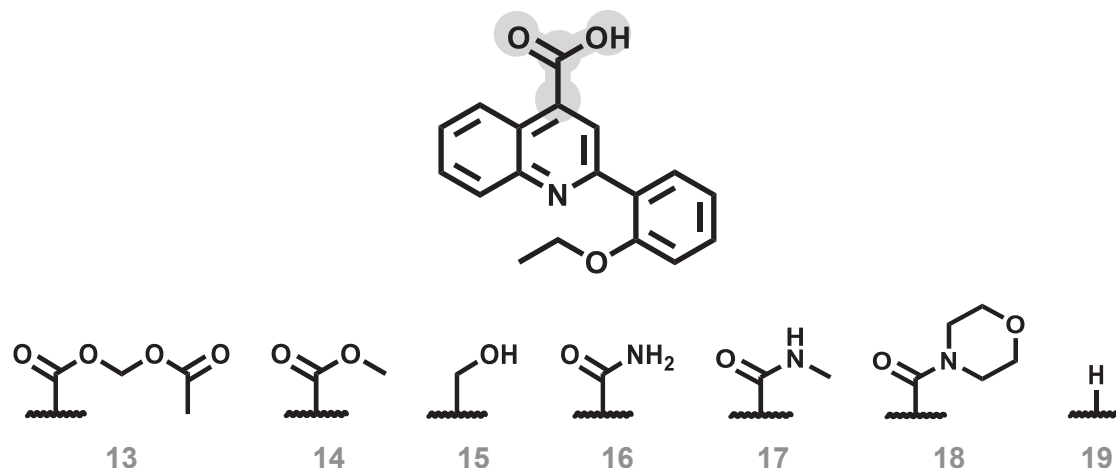
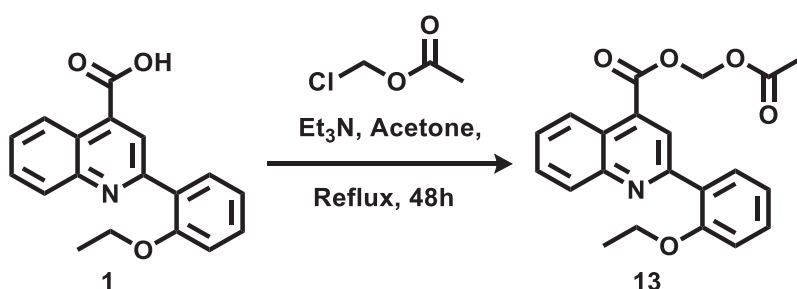


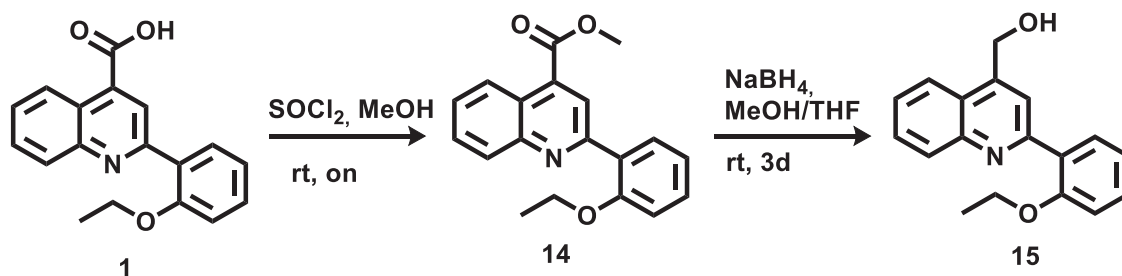
Figure 19. Synthesised analogues of the carboxylic group.

Compound **13** was design based on the approach used for **10**. The acetyloxy methyl ester is a pH-sensitive prodrug that should improve membrane permeability and once the compound enters the cell it should easily restore the free carboxylic group. This compound was synthesised by a nucleophilic displacement of the chloride on chloromethyl acetate in basic conditions (Scheme 4).



Scheme 4. synthesis of compounds 13.

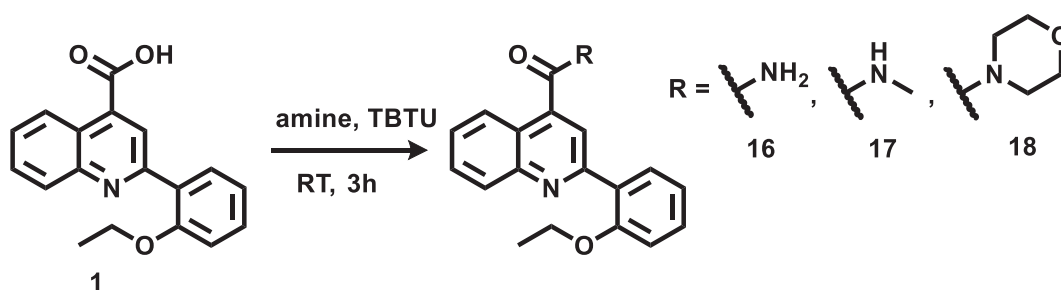
The ester analogue **14**, which was synthesised using an esterification reaction with thionyl chloride in methanol, was also made to see if the ester group could improve the activity of the compound by hindering the carboxylic group (scheme 5).



Scheme 5. Synthesis of compounds **14** and **15**.

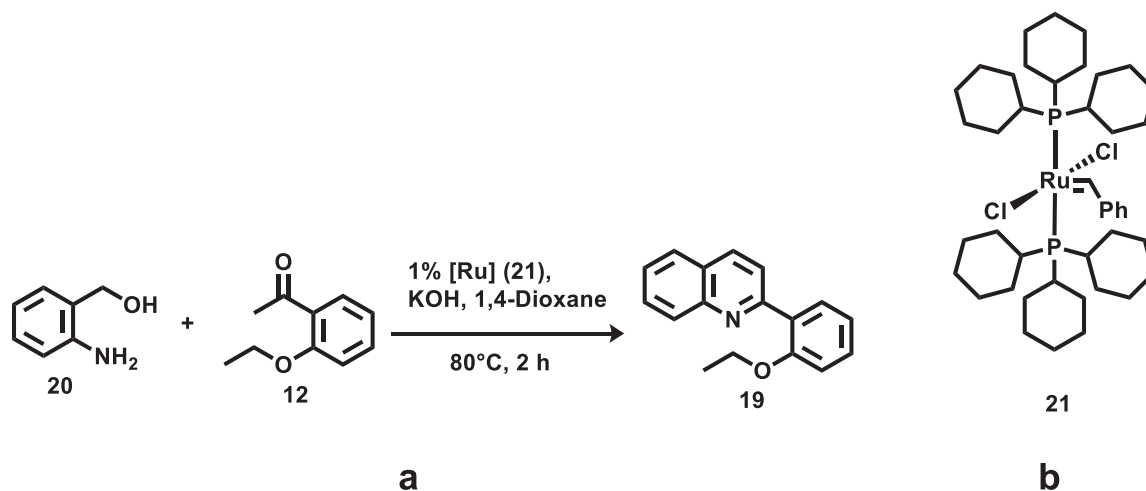
Compound **15** was synthesised by converting the carboxyl ester in the primary alcohol analogue through a sodium borohydride reduction. With this analogue, the importance of the negative charge of the carboxylic group was explored.

Starting from **1**, three amide analogues (**16-18**) were synthesised by a coupling reaction using TBTU and DIPEA (scheme 6).



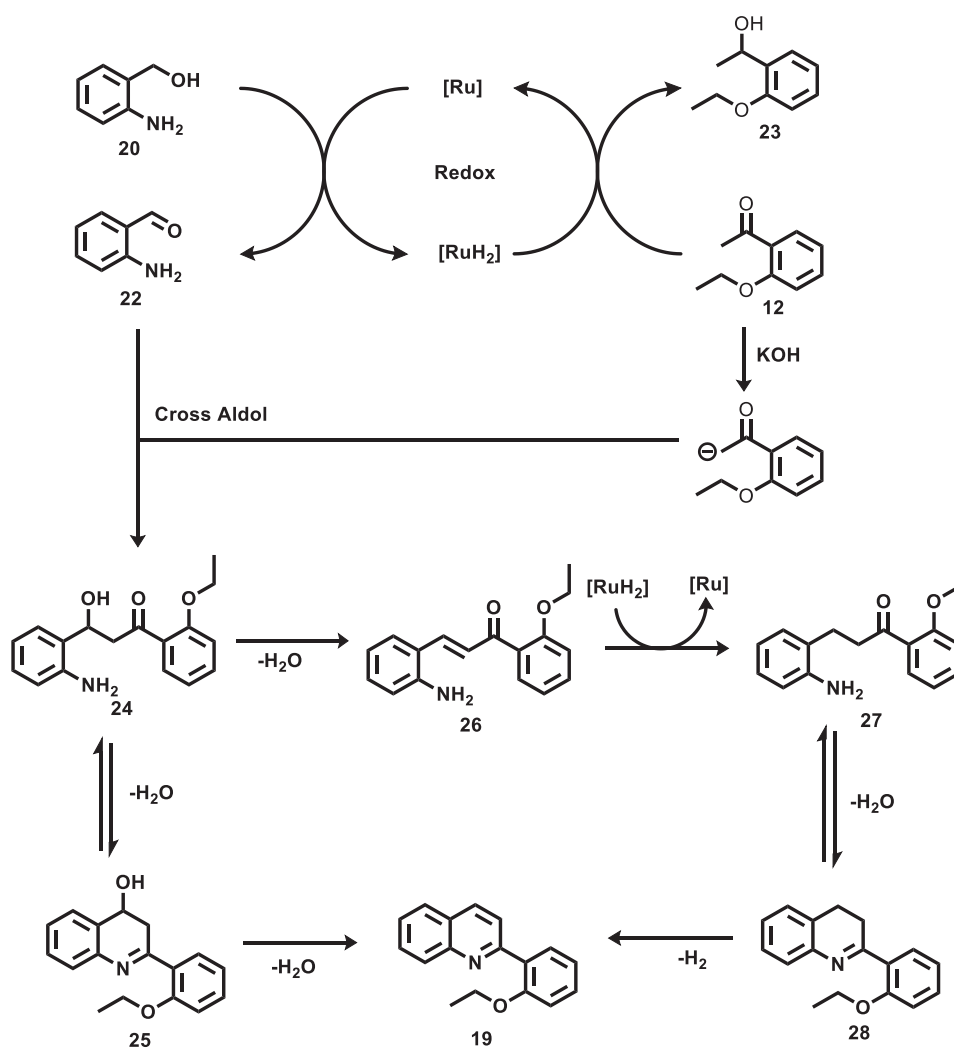
Scheme 6. synthesis of compounds **16**, **17**, **18**. The amines used are respectively ammonia, methylamine and morpholine

The analogue of **1** without the carboxylic group (**19**) was synthesised using a Ruthenium-Catalysed Friedlander protocol with a yield of 32%. The use of Ruthenium in catalytic amount is necessary to generate *in situ* the 2-aminobenzaldehyde **22** by oxidation of the aminobenzylalcohol **20** (scheme 7).¹¹¹



Scheme 7. a. synthesis of compound 17; b. chemical structure of the used Grubbs Catalyst 1st Generation 21

For this reaction, a commercially available 1st generation Grubbs catalyst **21** was used as source of ruthenium (scheme 8).



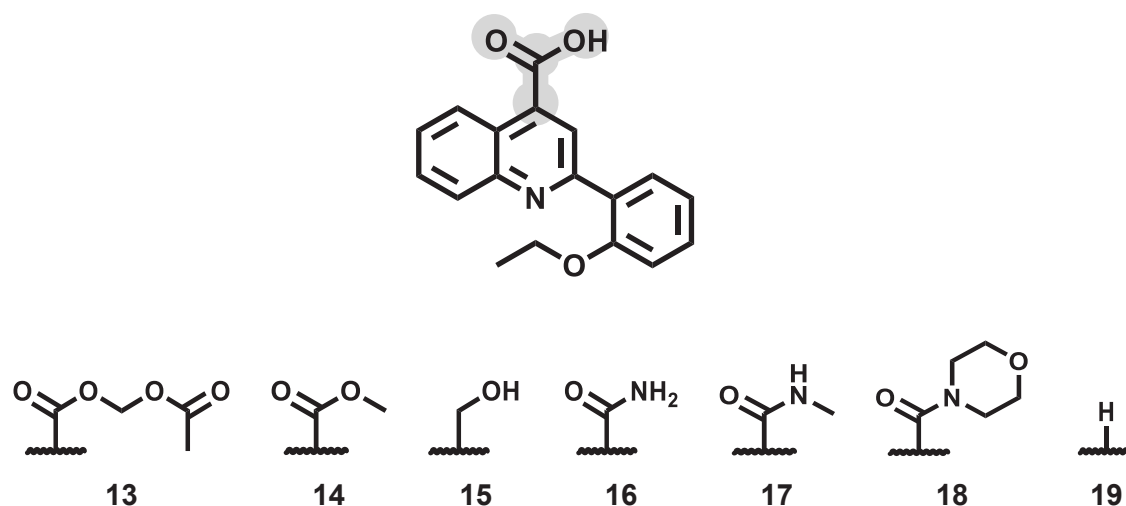
Scheme 8. Ruthenium-Catalysed Friedlander reaction mechanism.

The mechanism of this reaction, as shown in scheme 8, starts with the oxidation of the 2-aminobenzyl alcohol (**20**) into 2-amine benzaldehyde (**22**). The ruthenium catalyst is regenerated through the reduction of the 2-(ethoxy)acetophenone (**12**). The acetophenone in basic conditions is then reacted with the amino benzaldehyde in a cross aldol reaction to form **24**. **24** undergoes two subsequent H₂O eliminations, thus providing the desired quinoline. Another proposed pathway for this reaction is that **24**, by H₂O elimination, is forming the *trans* enone intermediate (**26**), which is then hydrogenated by the [RuH₂] to form intermediate **27**. This could be an alternative way of regeneration of [Ru]. The desired compound is generated through a subsequent H₂O elimination and dehydrogenation.

4.2.2 Biological evaluation of the carboxylic acid analogues

Among the synthesised compounds, **13-19**, as can be inferred from table 5, the pH-sensitive acetyloxy methyl ester prodrug (compound **13**) was associated with cytotoxicity, without showing antiviral activity on RSV-A2 and RSV-B. Compound **15** and **16**, with methyl alcohol and the carboxamide groups respectively instead of the carboxylic acid, were showing retention of antiviral activity against RSV-B.

Table 5. Biological evaluation of the analogues 13-19. Compounds were biologically evaluated in cell based antiviral assay (AVA) against RSV-A2 and RSV-B strains and for its cytotoxicity. Data are representative of two independent experiments. NA = not active. SI = selectivity index (CC_{50}/EC_{50})



Compound	EC ₅₀ RSV-A2 (μ M)	EC ₅₀ RSV-B (μ M)	CC ₅₀ (μ M)	SI RSV-A2	SI RSV-B
1	43	15	164	3.8	11
13	NA	NA	32	-	-
14	NA	3	>300	-	>100
15	44	27	>300	>6.8	>11
16	NA	27	196	-	7.2
17	NA	NA	>300	-	-
18	NA	NA	122	-	-
19	NA	NA	37	-	-

On the other hand, compound **15** and not compound **16** was showing retention of antiviral activity against RSV-A2. Interestingly, the methyl ester derivative of compound **14** was showing the lowest EC₅₀ against RSV-B (3 μ M) but no antiviral effect on RSV-A2. The other three modifications attempted, the *N*-methyl carboxamide (compound **17**), the

morpholine carboxamide (compound **18**) and removed carboxylic group (compound **19**) are related to loss of activity.

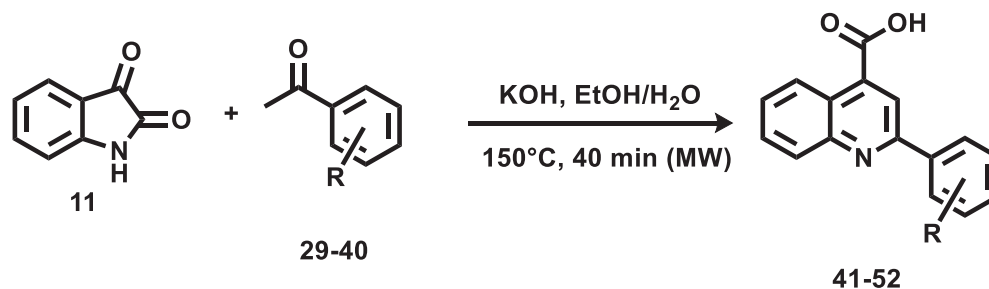
For the further exploration of the chemical structure of compound **1**, was decided to synthesise all the analogues as free carboxylic acid and as methyl esters.

4.2.3 SAR on the (2-ethoxy)phenyl moiety

The second stage of the SAR analysis was to understand the importance of the substitution on the phenyl moiety for the activity of the scaffold. As predicted by the docking study, the phenyl moiety is the one which is interacting in the hydrophobic pocket of N-NTD. Starting from this observation, several analogues were synthesised, using the synthetic approach already optimised for **1**, to investigate the importance of the phenyl moiety substituent. Each analogue was synthesised as free acid and ester derivatives, based on the preliminary results on the SAR on the carboxylic group.

The first analogue synthesised was the unsubstituted compound, to understand the importance of the ethoxy group of **1** for the anti-RSV activity.

The next stage was to investigate the importance of the length of the ethoxy group through the synthesis of methoxy (**42**), isopropoxy (**43**), *n*-propoxy (**44**) and (cyclopropyl)methoxy (**45**) analogues in position 2 of the phenyl moiety.

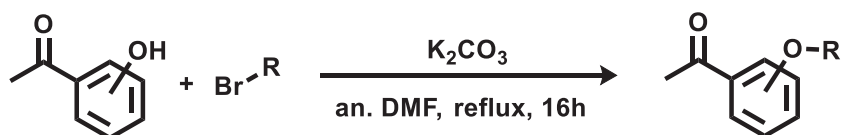


Scheme 9. Synthesis of compounds 41-52.

Table 6. Synthesis of compounds 41-52.

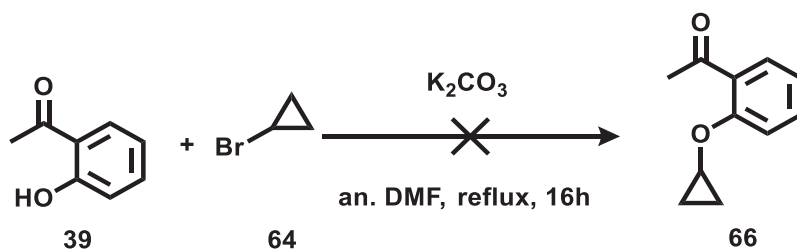
Acetophenone	R	Product	Yield %
29	H	41	66
30	2-OMe	42	58
31	2-O i Pr	43	39
32	2-OPr	44	30
33	2-Cyclopropylmethoxy	45	12
34	3-OEt	46	35
35	4-OEt	47	47
36	3-OMe	48	35
37	4-OMe	49	99
38	2,4-Cl	50	60
39	2-OH	51	20
40	3-OH	52	26

Analogues **41-52** shown in table **6** were synthesised using the same synthetic procedure used for **1**, changing the acetophenone. The acetophenones required were not all commercially available. For this reason, 2-(isopropoxy)acetophenone **31**, 2-(propoxy)acetophenone **32** and 2-((cyclopropyl)methoxy)acetophenone **33** were synthesised by alkylation of 2-hydroxyacetophenone **39**, as shown in scheme 10.¹¹²



Scheme 10. Synthesis of Acetophenones.

With the same procedure, few attempts were performed to synthesise 2-(cyclopropyl)acetophenone **66**. The desired product was not obtained, probably due to the different reactivity of the bromo-cyclopropane. During the reaction, the formation of several unknown by-products was observed (scheme 11).



Scheme 11. Synthesis of 2-(cyclopropyl)acetophenone 66.

The importance of the position of the ethoxy substituent on the phenyl ring was also investigated and the corresponding analogues with a 3-ethoxy and a 4-ethoxy group on the phenyl ring were synthesised. To have a wider overview on the importance of the substituent on the phenyl group, the corresponding analogues with a methoxy group in position 3 and 4 on the phenyl ring were also synthesised.

The most active compound of the series 1-benzyl-1H-pyrazole-3,5-dicarboxylate (BPdC), able to bind into the N-NTD pocket, has 2,4-dichlorobenzyl moiety (Figure 20). Analogue **50** was synthesised to see if it is possible to improve the affinity of the 2-(phenyl)quinolin-4-carboxylic acid scaffold for the hydrophobic pocket.⁹⁴

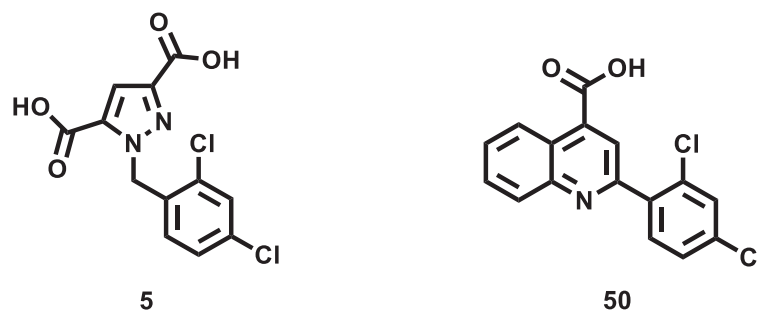
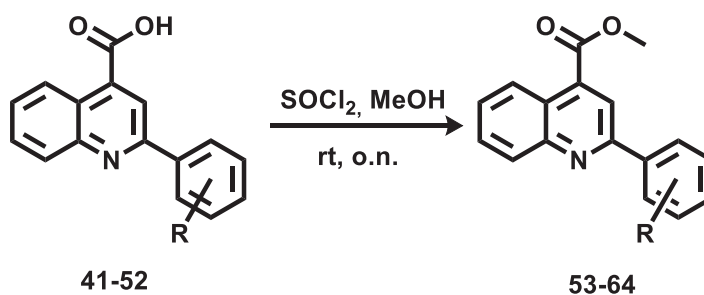


Figure 20. Comparison between 5 and 50.

Hydroxy analogues on position 2 and 3 of the phenyl ring were synthesised to see if introducing a hydrogen donor group could affect the activity of this scaffold.



Scheme 12. Synthesis of compounds 53-64.

Table 7. Synthesis of compounds 53-64.

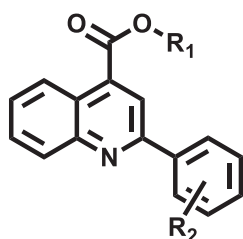
Compound	R	Product	Yield %
41	H	53	72
42	2-OMe	54	59
43	2-OiPr	55	59
44	2-OPr	56	84
45	2-Cyclopropylmethoxy	57	88
46	3-OEt	58	68
47	4-OEt	59	47
48	3-OMe	60	45
49	4-OMeH	61	25
50	2,4-Cl	62	64
51	2-OH	63	30
52	3-OH	64	79

The ester analogues (compounds **53-64**) were synthesised starting from the free acid compounds **41-52** as shown in scheme 12 and table 7.

4.2.4 Biological evaluations of the (2-ethoxy)phenyl moiety analogues

Among the free acid analogues synthesised (see table 8), a loss of activity was observed when removing the substituent on the phenyl ring (compound **41**). Increasing or decreasing the length of the substituent on R₁ was associated with a loss of activity, except for the n-propoxy analogue (compound **44**) which, retain the antiviral activity of compound **1**. Changing the position of the ethoxy group on the phenyl ring (compound **46** and **47**) is correlated with a loss of activity or increasing of cytotoxicity.

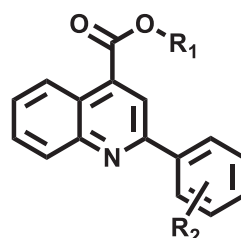
Table 8. Biological evaluation of free acid analogues 41-52. Compounds were biologically evaluated in cell based antiviral assay (AVA) against RSV-A2 and RSV-B strains and for its cytotoxicity. Data are representative of two independent experiments. NA = not active. SI = selectivity index (CC₅₀/EC₅₀)



Compound	R ₁	R ₂	EC ₅₀ RSV- A2 (μM)	EC ₅₀ RSV- B (μM)	CC ₅₀ (μM)	SI RSV- A2	SI RSV- B
1	H	2-OEt	43	15	164	3.8	11
14	Me	2-OEt	NA	3	>300	-	>100
41	H	H	NA	NA	>400	-	-
42	H	2-OMe	NA	NA	>300	-	-
43	H	2-OiPr	151	106	232	1.5	2.2
44	H	2-OnPr	32	35	148	4.6	4.2
45	H	2-Cyclopropyl-methoxy	NA	88	250	-	2.8
46	H	3-OEt	>150	63	206	<1.37	3.2
47	H	4-OEt	70	66	111	1.6	1.7
48	H	3-OMe	NA	NA	>100	-	-
49	H	4-OMe	NA	NA	>100	-	-
50	H	2,4-Cl	>150	204	>300	<2	1.5
51	H	2-OH	>25	68	54	<2.1	0.8
52	H	3-OH	>25	46	41	<1.6	0.9

The same trend can be seen in compound **48** and **49**. The substitution of the 2-ethoxy group with a hydroxy group is giving increased cytotoxicity. Compound **50** having the 2,4-dichlorobenzyl moiety, like compound **5** is not showing retention of the antiviral activity in comparison with compound **1**.

Table 9. Biological evaluation of methyl ester analogues 53-64. Compounds were biologically evaluated in cell based antiviral assay (AVA) against RSV-A2 and RSV-B strains and for its cytotoxicity. Data are representative of two independent experiments. NA = not active. SI = selectivity index (CC_{50}/EC_{50})



Compound	R ₁	R ₂	EC ₅₀ RSV- A2 (μ M)	EC ₅₀ RSV- B (μ M)	CC ₅₀ (μ M)	SI RSV- A2	SI RSV- B
1	H	2-OEt	43	15	164	3.8	11
14	Me	2-OEt					
53	Me	H	NA	66	92	-	1.4
54	Me	2-OMe	NA	NA	15	-	-
55	Me	2-OiPr	9	NA	23	2.5	-
56	Me	2-OnPr	11	16	17	1.5	1.1
57	Me	2-Cyclopropyl-methoxy	10	25	20	2	0.8
58	Me	3-OEt	40	NA	69	1.7	-
59	Me	4-OEt	NA	NA	4	-	-
60	Me	3-OMe	NA	41	137	-	3.3
61	Me	4-OMe	NA	6	18	-	3
62	Me	2,4-Cl	>13	NA	92	<7.1	-
63	Me	2-OH	NA	NA	13	-	-
64	Me	3-OH	NA	NA	26	-	-

For the methyl ester derivatives of **53-64** (see table 9), biological results indicate that all the analogues synthesised are showing an increase in cytotoxicity, with the exception for compound **60** and **62**. Compound **60** seems to retain the antiviral activity against RSV-B only. Instead, compound **62** is retaining the antiviral activity RSV-A2 and is showing an increase of cytotoxicity.

4.2.5 SAR on the 6 and 7 positions of the quinolinic ring

Aiming to reduce the cytotoxicity and enhance the anti-RSV activity of compound **1**, the following five different groups, -Me, Cl, -OMe, -CF₃ and -OCF₃, were selected to be introduced on 6- and 7- positions of the quinolinic ring, keeping the phenyl moiety of the hit compound unchanged. The selected compounds **77-86** were synthesised using the synthetic procedure used for compound **1**, starting from different commercially available 6- or 7-substituted isatines as shown in figure 21 and table 10.

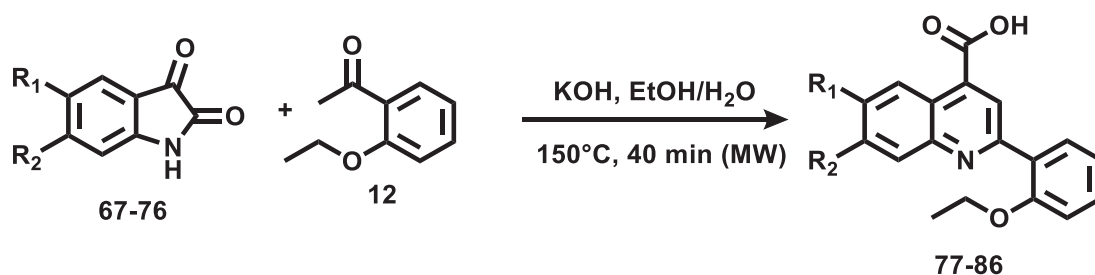


Figure 21. Synthesis of compounds 77-86.

Table 10. Synthesis of compounds 77-86.

Isatin	R ₁	R ₂	Product	Yield %
67	Me	H	77	61
68	Cl	H	78	63
69	OMe	H	79	20
70	CF ₃	H	80	5
71	OCF ₃	H	81	63
72	H	Me	82	34
73	H	Cl	83	20
74	H	OMe	84	17
75	H	CF ₃	85	60
76	H	OCF ₃	86	20

The obtained quinolines **77-86** were, then used for the synthesis of the ester derivatives, compounds **87-96** as shown in figure 22 and in table 11.

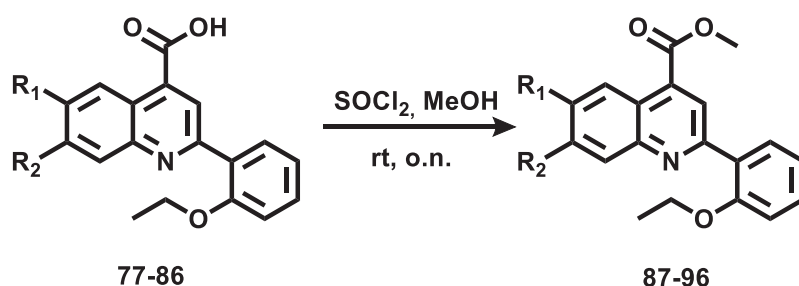


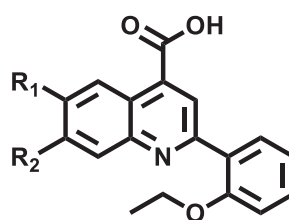
Figure 22. Synthesis of compounds 87-96.

Table 11. Synthesis of compounds 87-96.

Compound	R ₁	R ₂	Product	Yield %
77	Me	H	87	99
78	Cl	H	88	70
79	OMe	H	89	54
80	CF ₃	H	90	17
81	OCF ₃	H	91	56
82	H	Me	92	62
83	H	Cl	93	21
84	H	OMe	94	46
85	H	CF ₃	95	37
86	H	OCF ₃	96	75

4.2.6 Biological evaluations of 6 and 7 substituted quinolinic ring analogues

Table 12. Biological evaluation of free acid analogues 77-86. Compounds were biologically evaluated in cell based antiviral assay (AVA) against RSV-A2 and RSV-B strains and for its cytotoxicity. Data are representative of two independent experiments. NA = not active. SI = selectivity index (CC_{50}/EC_{50})



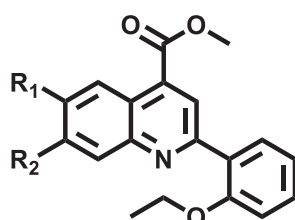
77-86

Compounds	R ₁	R ₂	EC ₅₀ RSV-A2 (μ M)	EC ₅₀ RSV-B (μ M)	CC ₅₀ (μ M)	SI RSV- A2	SI RSV-B
1	H	H	43	15	164	3.8	11
77	Me	H	NA	36	60	-	1.7
78	Cl	H	32	11	42	1.3	3.8
79	OMe	H	NA	NA	>300	-	-
80	CF ₃	H	19	16	9	0.5	0.6
81	OCF ₃	H	25	13	73	2.9	5.6
82	H	Me	NA	NA	>300	-	-
83	H	Cl	169	105	300	1.8	2.8
84	H	OMe	NA	NA	>300	-	-
85	H	CF ₃	36	36	174	4.8	4.8
86	H	OCF ₃	NA	35	229	-	6.5

From the biological data reported in table 12, compounds **77-86** did not show an improvement of antiviral activity on RSV-A2 and RSV-B. For the compounds showing retention of antiviral activity, the cytotoxic effect was enhanced since the SI was below the SI of compound **1**. The only exceptions were compound **85** and **86**, which were showing antiviral activity against RSV-B and cytotoxicity comparable to compound **1**. Between those two compounds only **85** shown anti-RSV-A2 activity, like compound **1**.

Biological evaluation of 6-quinoline substituted methyl ester analogues are reported in table 13. The 6-chloro and the 6-trifluoromethyl substituents, compounds **88** and **90**, do not shown antiviral activity against the two RSV strains. Instead the 6-methyl and the 6-methoxy analogues (compound **87** and **89**) showed an increased antiviral activity but also an enhanced cytotoxicity.

Table 13. Biological evaluation of methyl ester analogues 87-96. Compounds were biologically evaluated in cell based antiviral assay (AVA) against RSV-A2 and RSV-B strains and for its cytotoxicity. Data are representative of two independent experiments. NA = not active. SI = selectivity index (CC_{50}/EC_{50})



87-96

Compounds	R ₁	R ₂	EC ₅₀ RSV-A2 (μ M)	EC ₅₀ RSV-B (μ M)	CC ₅₀ (μ M)	SI RSV- A2	SI RSV- B
1	H	H	43	15	164	3.8	11
87	Me	H	>19	8	25	<1.3	3.1
88	Cl	H	NA	NA	>300	-	-
89	OMe	H	4	4	24	6	6
90	CF ₃	H	NA	NA	>300	-	-
91	OCF ₃	H	65	33	>300	>4.6	9
92	H	Me	11	12	285	25.9	23.7
93	H	Cl	NA	NA	>300	-	-
94	H	OMe	11	23	>300	>27.3	>13
95	H	CF ₃	11	11	>300	>27.3	>27.3
96	H	OCF ₃	3	11	>300	>100	>27.3

The introduction of a trifluoromethoxy group in position 6 (compound **91**) was associated with a retention of antiviral activity, comparable with compound **1**, with no cytotoxicity observed at 300 μM .

Finally, from the biological evaluation of the 7-quinoline substituted methyl ester analogues (compounds **92-96**), except for compound **93**, showed antiviral activity and no cytotoxicity at 300 μM . The most active compound was **96**, with an EC_{50} of 3 μM and 11 μM , respectively against RSV-A2 and RSV-B.

4.3 Synthesis of 2-(2-(*N*-(3-chloro-2-methylphenyl)methylsulfonamido)acetamido)benzoic acid (**2**)

The second most potent hit, identified through the SBDD approach, is the 2-(2-(*N*-(3-chloro-2-methylphenyl)ethylsulfonamido)acetamido)benzoic acid (compound **2**, figure 23 a). From the docking pose, the compound seems to interact with the N-NTD pocket, through an interaction between the carboxylic group and the Lys136. The 3-chloro-2-methylphenyl ring is interacting with the His151 through a π - π interaction, like for the 2-ethoxyphenyl moiety of compound **1**. Also, the carbonyl oxygen of the amide bond and the mesyl group seem to interact, respectively with the Arg150 and the Arg132. (figure 23 b).

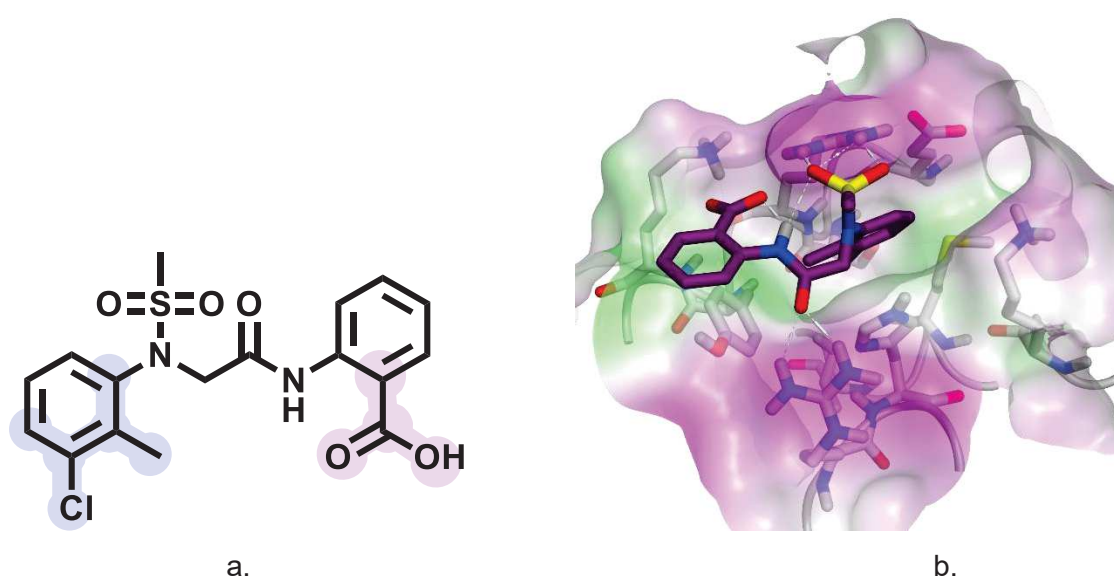


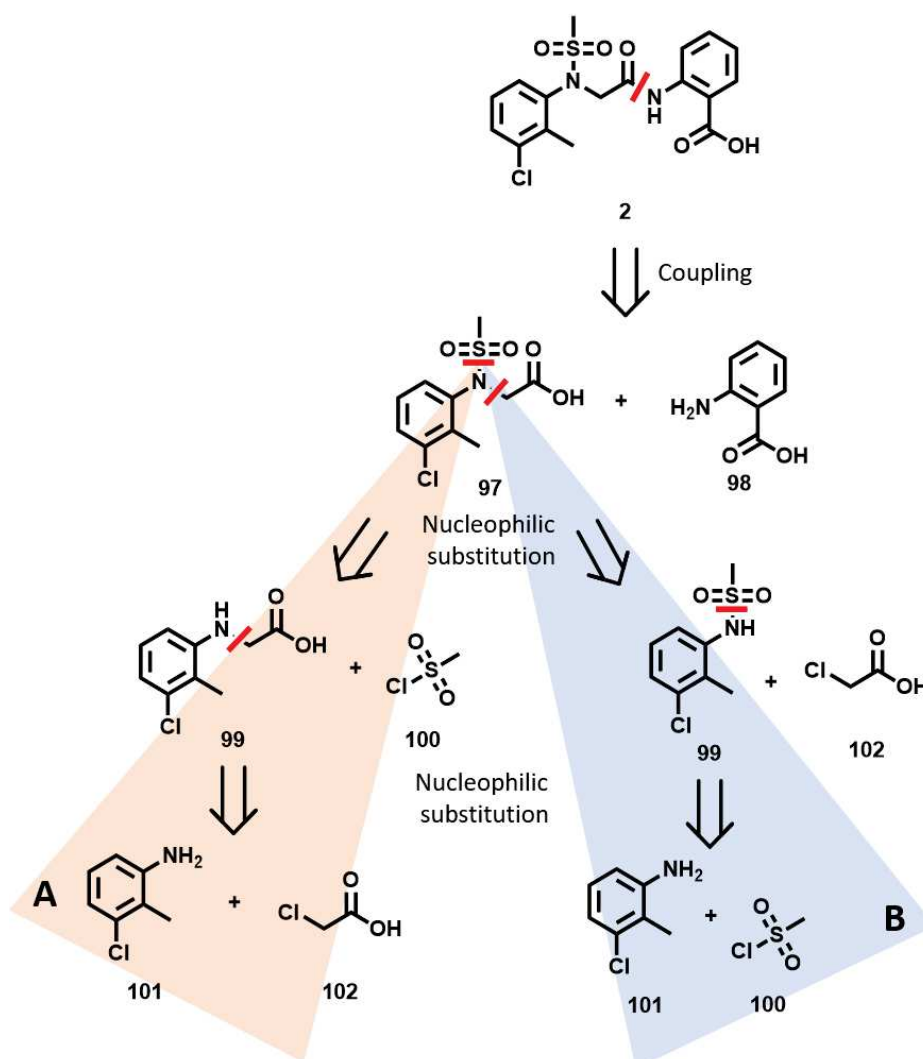
Figure 23. a) Chemical structure; b) Docking pose of compound 2.

In order to confirm the antiviral activity of compound **2**, was decided to synthesise the compound along with a small number of analogues, varying the carboxylic group and its position, and the 3-chloro-2-methylphenyl ring.

From the retrosynthetic analysis, shown in scheme 13, the amide bond of compound **2** can be disconnected to give 2-aminobenzoic acid **98** and the *N*-(3-chloro-2-methylphenyl)-*N*-(methylsulfonyl)glycine **97**. For this last one, either the sulphonamide bond (path A in scheme 13) or the N-CH₂ (path B in scheme 13) can be disconnected, giving two possible synthetic routes for the synthesis of compound **97**. In the path A, the disconnection of **97** is giving origin to the (3-chloro-2-methylphenyl)glycine **99**, which can be subsequently disconnected in 3-chloro-2-methylaniline **101** and 2-chloroacetic acid **102**.

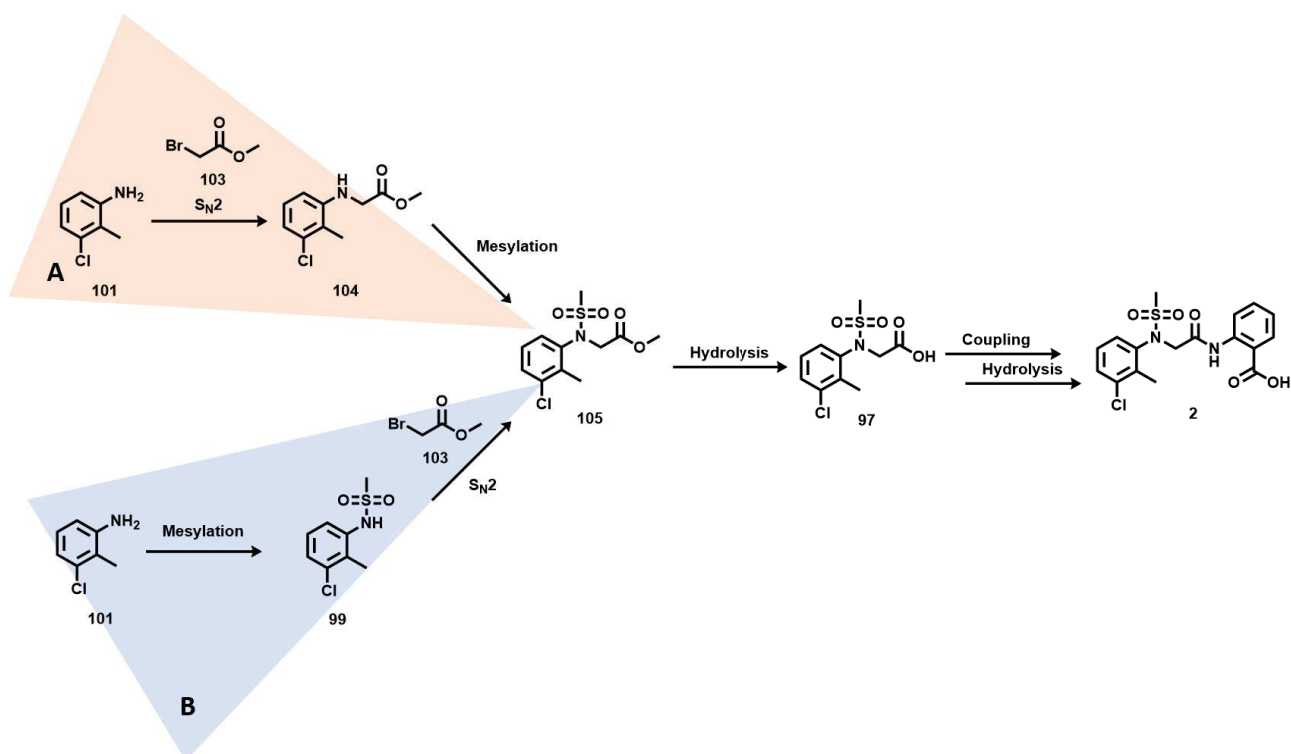
Alternatively, the disconnection of compound **97** in path B of the N-CH₂ bond is giving the 2-chloroacetic acid **102** and *N*-(3-chloro-2-methylphenyl)methanesulfonamide **99**, which

can be subsequently disconnected into 3-chloro-2-methylaniline **101** and mesyl chloride **100**.



Scheme 13. Compound 2 retrosynthetic analysis.

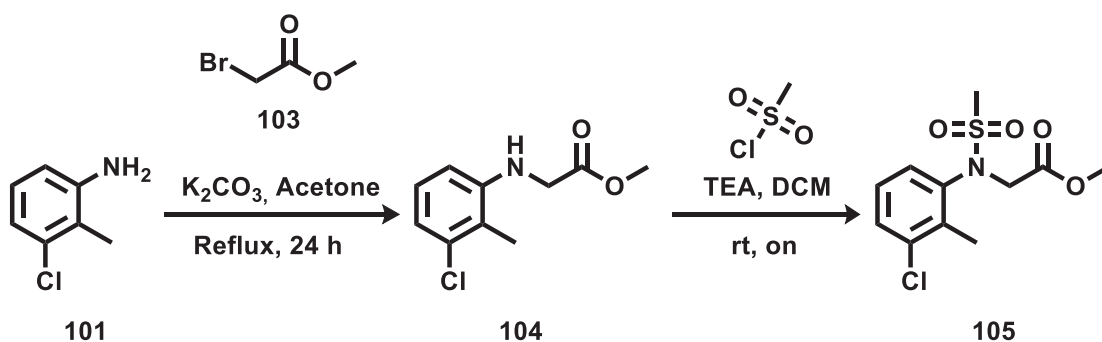
From the retrosynthetic analysis, two synthetic routes were planned to synthesise compound **97**, see scheme 14. Those routes have in common the methyl ester precursor of compound **97**. In synthetic route A, the first step is the nucleophilic substitution (S_N2), followed by the mesylation step. Instead, in the synthetic route B, the mesylation of the 3-chloro-2-methylaniline **101** is the first step, followed by the S_N2 , to give methyl ester precursor **105** of compound **97**.



Scheme 14. Planned synthetic routes, Route A in pink and Route B in light blue.

Synthetic route A

Several attempts to synthesise compound **97** were performed starting from the commercially available 3-chloro-2-methylaniline **101** and the methyl 2-bromoacetate **103**, instead of 2-chloroacetic acid **102**, which is shown in scheme 15. The choice of using methyl 2-bromoacetate **103** was made to avoid any side reaction and because the bromo atom is a better leaving group than the chloro atom in nucleophilic substitution. The best result of the nucleophilic substitution for the formation of the methyl (3-chloro-2-methylphenyl)glycinate **104**, was obtained in acetone in presence of potassium carbonate as a base, the reaction was refluxed for 24 hours. After flash column chromatography the product was isolated in presence of the starting material **101**, in a ratio of 3:2. The obtained mixture was then used for the next mesylation step. After flash column chromatography the desired product **105** was obtained, with an overall yield over two steps of 20%.

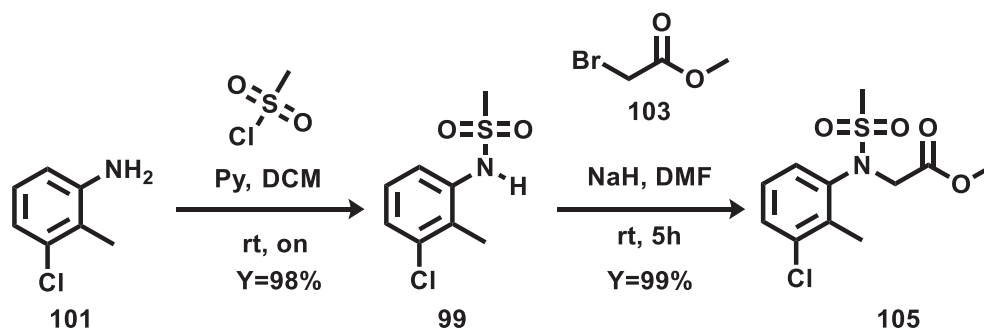


Scheme 15. Synthetic route A.

Synthetic route B

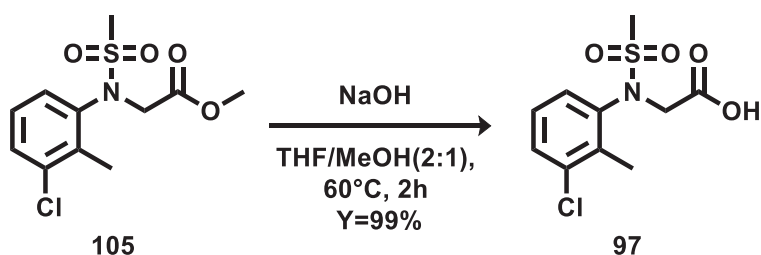
To increase the yield and to reduce the steps of flash chromatography purification, it was decided to try the mesylation of the 3-chloro-2-methylaniline **101** as first step and then the nucleophilic substitution on the methyl 2-bromoacetate **103** (scheme 16).

The mesylation reaction of the aniline **101** was performed in dichloromethane in the presence of pyridine as a base, overnight at room temperature giving the intermediate **99**, pure with a yield of 98%.¹¹³ The desired product was obtained by nucleophilic substitution on methyl 2-bromoacetate by the mesylated aniline in basic condition. This synthetic pathway lead to the synthesis of methyl *N*-(3-chloro-2-methylphenyl)-*N*-(methylsulfonyl)glycinate **105** without requiring any step of flash chromatography purification.



Scheme 16. Synthetic route B.

Compound **97** was finally obtained by basic hydrolysis of the compound **105** as shown in scheme 17.

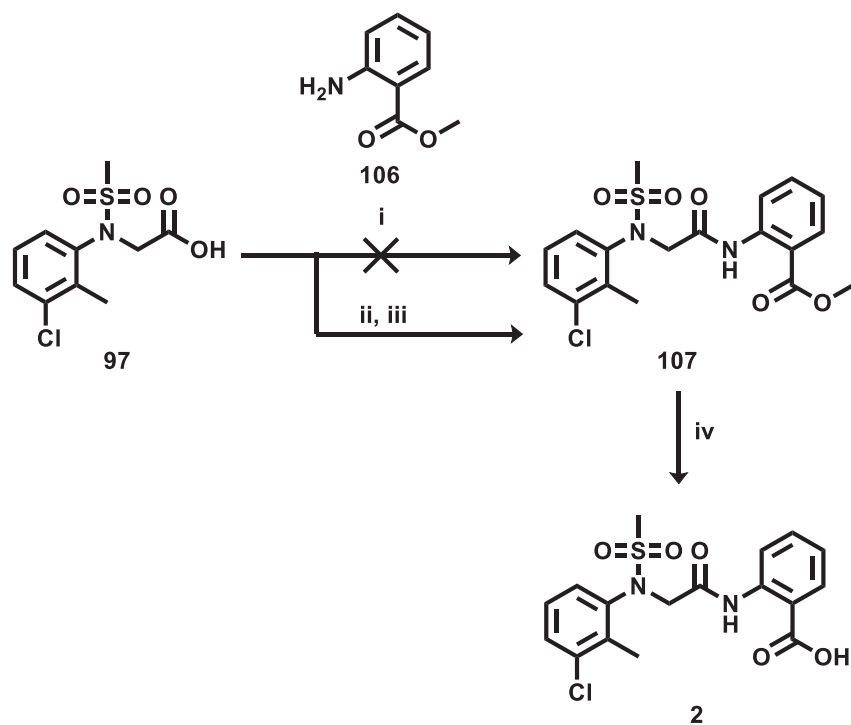


Scheme 17. Hydrolysis of compound 105.

To synthesise the compound **2**, a first attempt of coupling between the free acid **97** and the commercial available methyl 2-aminobenzoate **106** was performed using the TBTU (O-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate) amide coupling reagent. The reaction was attempted several times, increasing the temperature of the reaction, with any success. This can be related to the low reactivity of the anilinic amine,

which is not able to replace the benzotriazole ester formed between the acid **97** and the TBTU.

To overcome this problem, the acid was converted into the acyl chloride derivative using the chlorinating agents, thionyl chloride. The acyl chloride derivative of **97** was then added drop-wise to a DCM solution of methyl 2-aminobenzoate **106**, using triethylamine as base (Scheme 18), giving the methyl ester derivative **107** with a yield of 86% over two steps. Finally compound **2** was obtained by basic hydrolysis of the methyl ester **107**.



Scheme 18. Synthesis of 2-(2-(N-(3-chloro-2-methylphenyl)ethylsulfonamido)acetamido)benzoic acid (**2**). i) TBTU, DIPEA, DMF, on; ii) SOCl₂ neat, reflux, 2h; iii). Aniline, TEA, DCM, from 0°C to rt, on, Y=86% (over 2 steps); iv) NaOH, THF/MeOH (2:1), 60°C, 2h, Y=99%;

4.3.1 2-(*N*-(3-chloro-2-methylphenyl)mesyl) acetamides analogues

Four compounds were designed to investigate the importance of the carboxylic group of the hit compound **2**, as shown in figure 24. In compound **108** and **109** the carboxylic group was respectively removed and substituted with the tetrazole ring, which is the bioisostere of the carboxylic acid group. Compounds **110-111** were designed to investigate if the position of the carboxylic group could affect the antiviral activity of compounds.

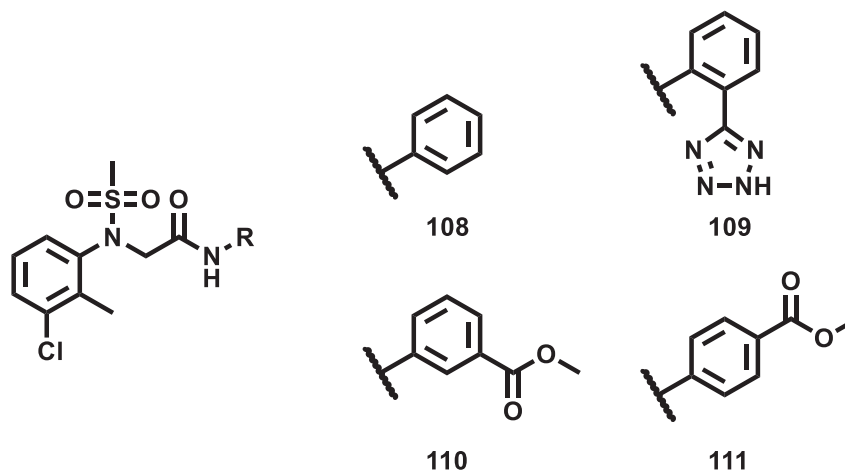
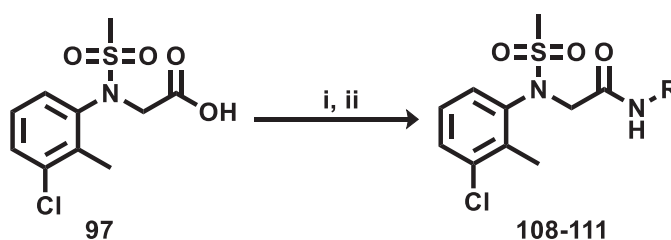


Figure 24. Designed analogues 108-111.

Compounds **108-111** were synthesised, starting from the common intermediate **97**, using the optimised coupling procedure used for the synthesis of compound **2** in which, the acyl chloride derivative of **97** was added drop-wise to a solution of dichloromethane and triethylamine, containing the relative anilines: aniline, 2-(tetrazol-5-yl)aniline, methyl 3-aminobenzoate and methyl 4-aminobenzoate, as shown in scheme 19 and table 14.

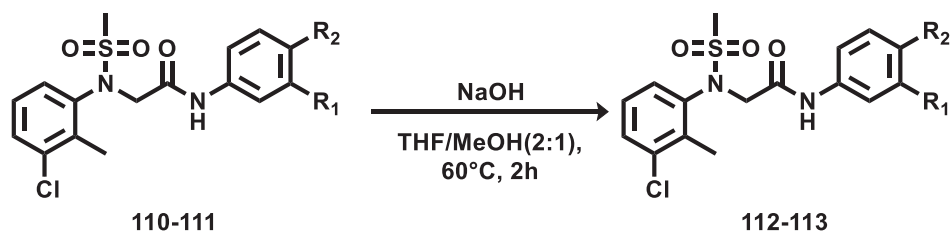


Scheme 19. Synthesis of compounds 108-111. i) SOCl_2 neat, reflux, 2h; ii). R-NH_2 , TEA, DCM, from 0°C to rt, on.

Table. Synthesis of compounds 108-111. The yields are calculated over the two steps of synthesis.

Product	R	Yield %
108	phenyl	72
109	2-(tetrazol-5-yl)phenyl	14
110	methyl 3-benzoate	84
111	methyl 4-benzoate	45

The methyl 3-benzoate **110** and the methyl 4-benzoate **111** were hydrolysed to obtain also the free acid derivatives **112-113**, as shown in scheme 20 and the related yields are reported in table 15.



Scheme 20. Hydrolysis of compounds 112-113.

Table 14. Synthesis of compounds 112-113.

Compound	R ₁	R ₂	Product	R ₁	R ₂	Yield %
110	COOCH ₃	H	112	COOH	H	40
111	H	COOCH ₃	113	H	COOH	28

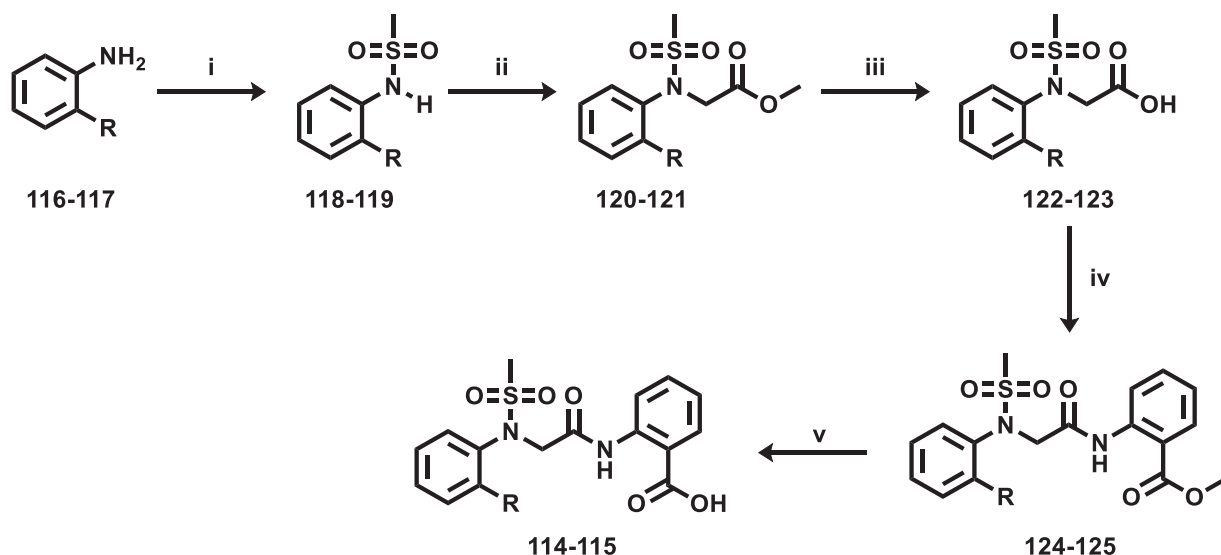
4.3.2 2-(2-(*N*-mesylphenyl)acetamido)benzoic acids analogues

To investigate the importance of the 3-chloro-2-methylphenyl ring of compound **2** for the antiviral activity, two derivatives of compound **2** were designed (see figure 25). In the first one the 3-chloro-2-methylphenyl ring was replaced with a phenyl ring (compound **114**). In the second derivatives of compound **2**, the 3-chloro-2-methylphenyl moiety, which from the docking seems to bind in the hydrophobic pocket of the N protein, was replaced with the 2-etoxyphenyl moiety (compound **115**). This moiety is the same of compound **1**, which from the docking pose seems to bind in the same pocket and to be able to make the same interaction in the pocket.



Figure 25. Compounds **114** and **115**.

The designed compounds **114** and **115** were synthesised using the optimised five steps synthetic route which was used for the synthesis of compound **2**.

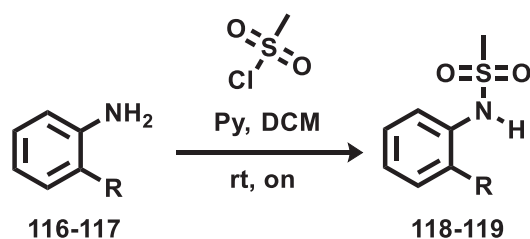


Scheme 21. Synthesis of compounds 114-115. i) mesyl chloride, pyridine, DCM, rt, on; ii) methyl 2-bromoacetate, NaH, DMF, rt, 5h; iii) NaOH, THF/MeOH(2:1), 60°C, 2h; iv) 1)SOCl₂ neat, reflux, 2h; 2). Aniline, TEA, DCM, from 0°C to rt, on; v) NaOH, THF/MeOH(2:1), rt, on.

Synthesis of *N*-mesylphenylamines (118-119)

The two newly synthesised *N*-mesylphenylamines **116** and **117** were synthesised through the mesylating reaction used for the synthesis of 3-chloro-2-methylaniline **101**, with a comparable yield (see table 16).

Table 15. Synthesis of *N*-mesylphenylamines 118-119.

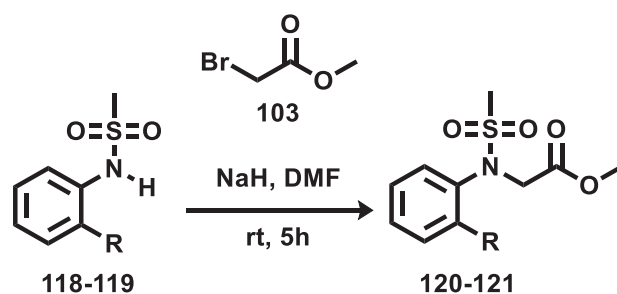


Compound	R ₁	Product	Yield %
116	H	118	99
117	2-ethoxy	119	96

Synthesis of methyl *N*-(mesyl)-*N*-phenylglycinates (120-121)

The next step of synthesis, like for compound **2**, was the S_N2. The sulphonamide, which is deprotonated by the strong base sodium hydride, is displacing the Br atom on the methyl 2-bromoacetate **103** (see table 17).

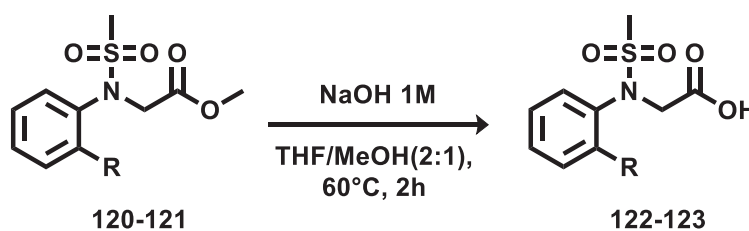
Table 16. Synthesis of methyl *N*-(mesyl)-*N*-phenylglycinates 120-121.



Compound	R	Product	Yield %
118	H	120	96
119	2-ethoxy	121	98

Synthesis of *N*-(mesyl)-*N*-phenylglycines (122-123)

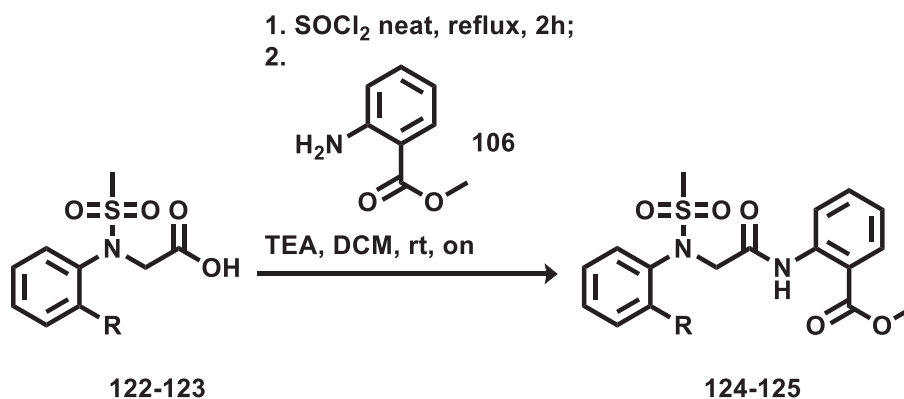
The previously obtained *N*-(mesyl)-*N*-phenylglycinates **120-121**, were then hydrolysed in basic conditions using an aqueous solution of sodium hydroxide (1M) and the reaction performed in solution of tetrahydrofuran and methanol, in a ratio of 2:1 (see table 18), to give the two desired free acids products **122-123** (see table 18).

Table 17. Synthesis of *N*-(mesyl)-*N*-phenylglycine (122-123).

Compound	R	Product	Yield %
120	H	122	84
121	2-ethoxy	123	99

Synthesis of methyl 2-(2-(*N*-phenylmesyl)acetamido)benzoates (124-125)

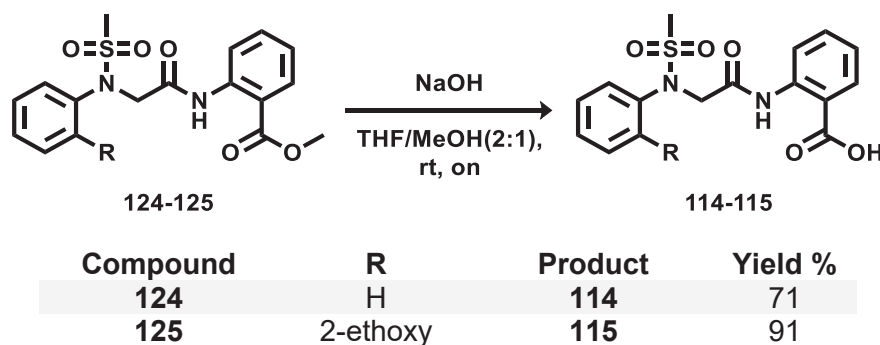
The glycines **122-123** were coupled with the methyl 2-aminobenzoate **106** using the coupling procedure optimised for the synthesis of compound **2**. Both the glycines **122-123** were converted to the respective acyl chloride and then added drop-wise to the solution of methyl 2-aminobenzoate **106** in dichloromethane and triethylamine (see table 19). Two methyl 2-(2-(*N*-phenylmesyl)acetamido)benzoates (**124-125**) were also biologically evaluated along with the free acids **114-115**.

Table 18. Synthesis of methyl 2-(2-(*N*-phenylmesyl)acetamido)benzoates 124-125 (* the yield is referred over two steps).

Compound	R	Product	*Yield %
122	H	124	52
123	2-ethoxy	125	96

Synthesis of 2-(2-(*N*-phenylmesyl)acetamido)benzoic acids (114-115)

Finally, the 2-(2-(*N*-phenylmesyl)acetamido)benzoic acids **114-115** were obtained by hydrolysis in basic condition of compounds **124-125** (see table 20). Two 2-(2-(*N*-phenylmesyl)acetamido)benzoic acids (**114-115**) were also biological evaluated along with the methyl ester precursors **124-125**.

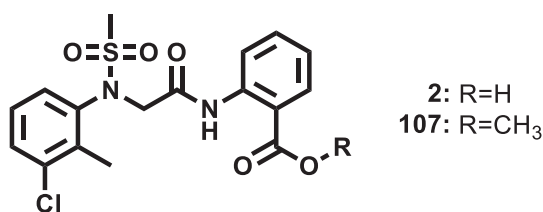
Table 19. Synthesis of 2-(2-(*N*-mesylphenyl)acetamido)benzoic acid 114-115.

4.3.3 Biological evaluation

The newly synthesised 2-(2-(*N*-mesylphenyl)acetamido)benzoic acid compounds were tested for antiviral activity against RSV-A2 and RSV-B CPE antiviral assay and were evaluated for their cytotoxicity in HEp2 cells line.

The re-synthesised compound **2** did not show any antiviral activity in HEp2 cell line against RSV-A2. Instead the compound shown a 58 μM EC_{50} against RSV-B and no cytotoxicity at 300 μM . Besides, the methyl ester analogue compound **107**, shown an EC_{50} against RSV-A2 and RSV-B, respectively of 26 μM and 34 μM with no cytotoxicity at 300 μM (see table 21).

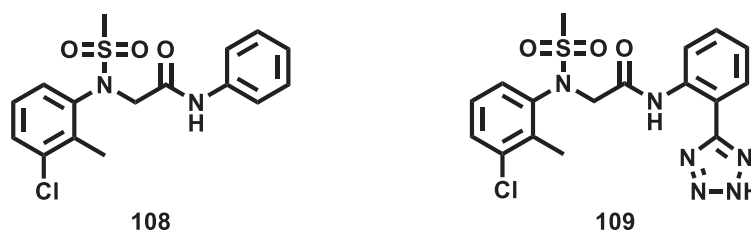
Table 20. Biological results for compounds 2 and 107. Compounds were biologically evaluated in cell based antiviral assay (AVA) against RSV-A2 and RSV-B strains and for its cytotoxicity. Data are representative of two independent experiments. NA = not active. SI = selectivity index ($\text{CC}_{50}/\text{EC}_{50}$)



Compounds	EC_{50} RSV-A2 (μM)	EC_{50} RSV-B (μM)	CC_{50} (μM)	SI RSV-A2	SI RSV-B
2	>300	58	>300	>1	>5.2
107	26	34	>300	>11.5	>8.8

The removal of the carboxylic group was associated with a CC_{50} and a lower SI in comparison with the compound **107**. The replacement of the carboxylic group by the tetrazole ring was associated with a complete loss of antiviral activity (see table 22).

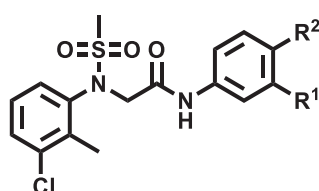
Table 21. Biological results for compounds 108-109. Compounds were biologically evaluated in cell based antiviral assay (AVA) against RSV-A2 and RSV-B strains and for its cytotoxicity. Data are representative of two independent experiments. NA = not active. SI = selectivity index (CC_{50}/EC_{50})



Compounds	EC ₅₀ RSV-A2 (μ M)	EC ₅₀ RSV-B (μ M)	CC ₅₀ (μ M)	SI RSV-A2	SI RSV-B
108	76	35	155	2	4.4
109	NA	NA	>300	-	-

The methyl 3-aminobenzoate **110** and methyl 4-aminobenzoate **111** derivatives of compound **107** lead to an antiviral activity retention but associated to a lower CC_{50} than compound **107**. The free acid derivatives **112** and **113** did not show any antiviral activity (see table 23).

Table 22. Biological results for compounds 110-113. Compounds were biologically evaluated in cell based antiviral assay (AVA) against RSV-A2 and RSV-B strains and for its cytotoxicity. Data are representative of two independent experiments. NA = not active. SI = selectivity index (CC_{50}/EC_{50})



110: R¹=COOCH₃, R²=H;

112: R¹=COOH, R²=H;

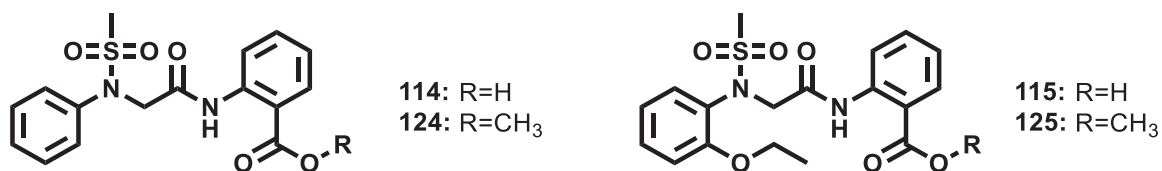
111: R¹=H, R²=COOCH₃;

113: R¹=H, R²=COOH;

Compounds	EC ₅₀ RSV-A2 (μ M)	EC ₅₀ RSV-B (μ M)	CC ₅₀ (μ M)	SI RSV-A2	SI RSV-B
110	36	30	78	2.2	2.6
112	NA	NA	>300	-	-
111	11	11	22	2	2
113	NA	NA	>300	-	-

The two designed modifications on the 3-chloro-2-methylphenyl moiety of compound **2**, did not show any correlation with antiviral activity, neither the methyl ester derivatives **114** and **115**, nor the free acid derivatives **124** and **125** (see table 24).

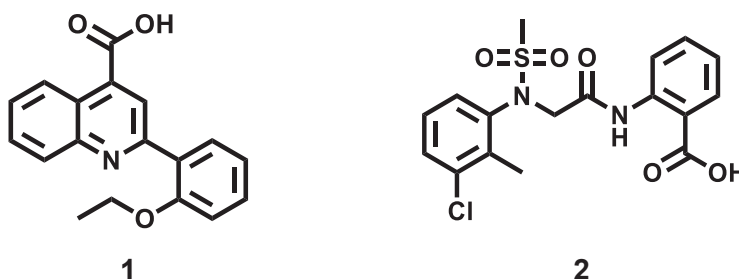
Table 23. Biological results for compounds 114-115, 124-125. Compounds were biologically evaluated in cell based antiviral assay (AVA) against RSV-A2 and RSV-B strains and for its cytotoxicity. Data are representative of two independent experiments. NA = not active. SI = selectivity index (CC_{50}/EC_{50})



Compounds	EC ₅₀ RSV-A2 (μM)	EC ₅₀ RSV-B (μM)	CC ₅₀ (μM)	SI RSV- A2	SI RSV- B
114	NA	NA	>300	-	-
124	NA	103	>300	-	>2.91
115	NA	NA	>300	-	-
125	NA	NA	>300	-	-

4.4 Conclusions

Different structure-based strategies were used to identify new potential compounds able to inhibit RSV viral replication. During the in-silico study, four of the selected compounds were able to inhibit the RSV replication in the cell-based assay fully. The two most active hits, compound **1** and **2**, having a 2-phenylquinoline-4-carboxylic acid and 2-(2-(*N*-phenyl-*N*-mesyl)acetamido)benzoic acid scaffolds respectively, were further investigated.



From the biological results, the antiviral activity of compound **1** against RSV was confirmed. Several derivatives were synthesised to explore the importance of the different features of the 2-phenylquinoline-4-carboxylic acid scaffold with some compound showing improved antiviral activity in comparison to compound **1**.

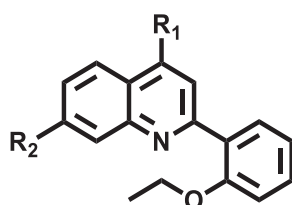
From the analogues synthesised to understand the importance of the carboxylic group, it was observed that the removal of the group was associated with a complete loss of antiviral activity. The replacement of the free acid group of the hit compound **1** with methyl alcohol was associated with retention of antiviral activity. The methyl ester analogue showed an improved activity against only RSV-B. Also, the carboxamide analogue showed retention of antiviral activity against only RSV-B. On the base of those results, was decided to synthesise the other analogue both as free acids and methyl esters.

The analogues synthesised to understand the importance of the phenyl moiety were associated with low CC_{50} and SI values or loss of antiviral activity, either along the free acid series and in the methyl ester series. The only exception was compound **44** having the 2-propoxy group on the phenyl group instead of the 2-ethoxy group. This compound was related to antiviral activity retention against both RSV-A2 and RSV-B. However, the results obtained for its methyl ester derivatives were inconsistent with the results obtained with the free acid compound, in which a low CC_{50} and SI were observed.

The modification introduced on the quinolinic ring in position 6, were associated with low CC_{50} and SI values, both in the free acid and the methyl ester series. The only exception was compound **91**, having the trifluoromethoxy group, showing retention of antiviral activity of compound **1**.

Interestingly the methyl ester series of the 7-substituted quinoline analogues, few compounds showed antiviral activity against both RSV strains in the range of 3-11 μM for RSV-A2 and in a range of 11-23 μM for RSV-B with an SI major of 23.7. These findings, for the 2-phenylquinoline-4-carboxylic acid series, are summarised in table 25.

Table 24. 2-phenylquinoline-4-carboxylic acid derivatives with an improved antiviral activity. Compounds were biologically evaluated in cell based antiviral assay (AVA) against RSV-A2 and RSV-B strains and for its cytotoxicity. Data are representative of two independent experiments. NA = not active. SI = selectivity index (CC_{50}/EC_{50})



Compound	R ₁	R ₂	EC ₅₀ RSV-A2 (μM)	EC ₅₀ RSV-B (μM)	CC ₅₀ (μM)	SI RSV- A2	SI RSV- B
1		H	43	15	164	3.8	11
14		H	NA	3	>300	-	>100
15		H	44	27	>300	>6.8	>11
16		H	NA	27	196	-	7.2
92		Me	11	12	285	25.9	23.7
94		OMe	11	23	>300	>27.3	>13
95		CF ₃	11	11	>300	>27.3	>27.3
96		OCF ₃	3	11	>300	>100	>27.3

Future modifications of 2-phenylquinoline-4-carboxylic acid series could be the replacement of quinoline ring with other heterocyclic aromatic rings, to improve the antiviral activity on RSV for this series of compounds.

The antiviral activity of the second hit, compound **2** was not confirmed. The methyl ester derivative of compound **2** showed a comparable antiviral activity on both RSV stains to compound **1**. All the designed and synthesised 2-(2-(*N*-phenyl-*N*-mesyl)acetamido)benzoic acid analogues did not show any antiviral activity improvements.

Further biological studies are ongoing, raising resistant virus for these compounds, to confirm or not that the N protein binding is the mechanism of action.

Chapter 5: F protein and α -Helix Mimics

5.1 F protein

The F protein is responsible for the viral entry, through the fusion of the viral and cell membranes. The functional protein is a homotrimer of F proteins and has two functional conformations, one called pre-fusion and the second one called post-fusion (Figure 26 a,b). The transformation from the pre- to the post-fusion conformation is driven by a series of conformational rearrangements, as explained in the introduction.

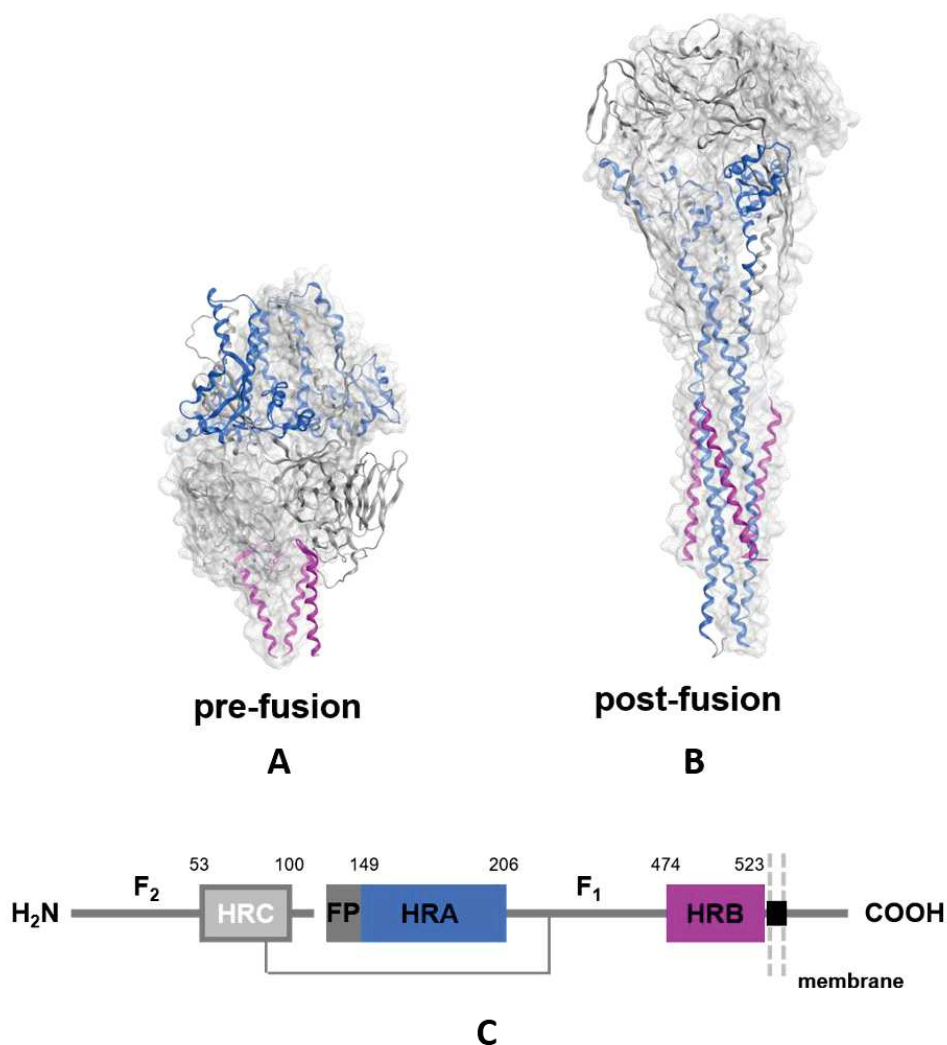


Figure 26. F protein. a) ribbon representation of the pre-fusion F protein conformation; b) ribbon representation of the post-fusion F protein conformation; c) schematic representation of the F protein domains. HRA = Heptad Repeat A domain; HRB = Heptad Repeat B domain, HRC = Heptad Repeat C domain, FP = fusion peptide.

The protein has different domains (Figure 26 c.), but the most important are the HRB and HRA, which are fundamental for the fusion process.³⁷

Since 1980, different series of compounds were published and patented as RSV inhibitors of the F protein (Figure 27), and almost all of them are specifically targeting the 6-HB motif, which is generated during the late stage of the conformational rearrangements.¹¹⁴⁻¹²² The

rearrangements that lead to a stable 6-HB conformation are the most important for viral entry. All these inhibitors cause rapid development of viral strains resistance.¹²³ The formation of the 6-HB seems to be driven by the hydrophobic interaction of the HRB hydrophobic residues.

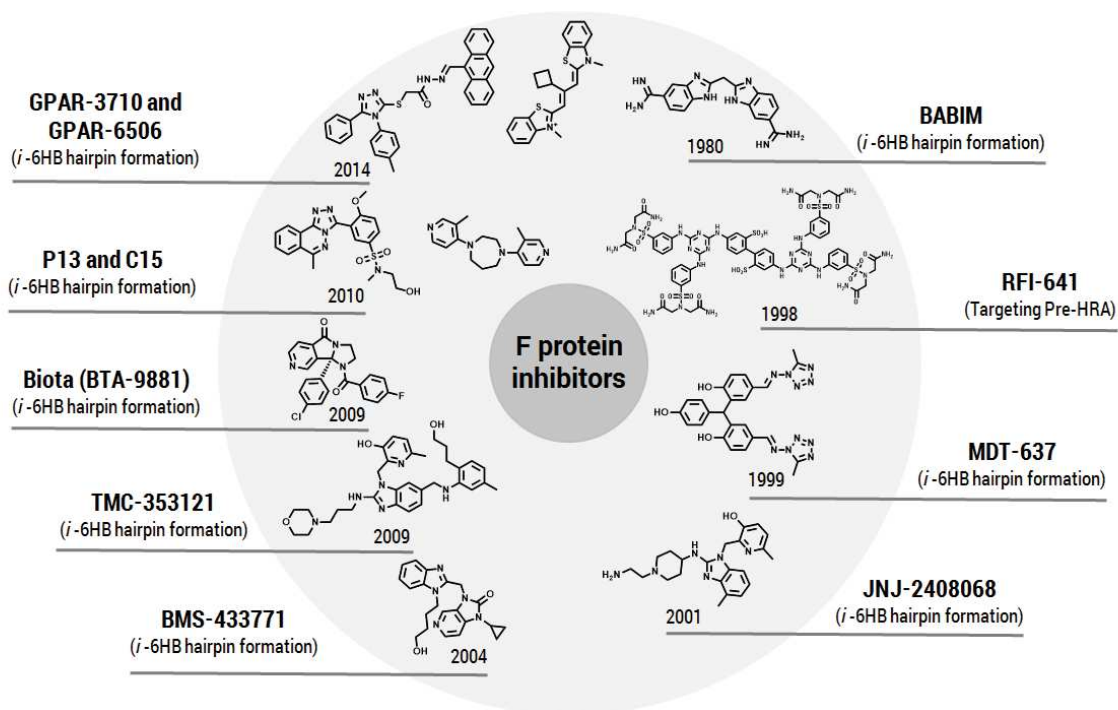


Figure 27. F protein inhibitors.

5.2 Rational design of a α -helix mimics

In 1996, synthetic peptides with the helical structure of the HRB region of the F protein, were reported as potent inhibitors of RSV, in particular, they were able to inhibit the viral fusion in cell-based assays.¹²⁴ The FDASISQVNEKINQSLAFIRKSDELL sequence was reported as the portion of the HRB responsible for the inhibition activity and can be observed in the crystal structure of the post-fusion F protein (Fig. 28; Pdb: 3RRR).

The HIV gp41 protein has a fusion complex similar to the RSV 6-HB motif of the F fusion. A reported strategy for the inhibition of the formation of HIV gp41 6-HB is to inhibit the PPI between the trimeric inner coiled-coil and the external α -helix, using α -helix mimics.¹²⁵ An α -helix mimic is a compound that mimics the topography of the crucial amino acids backbones that are important for the PPI.¹²⁶

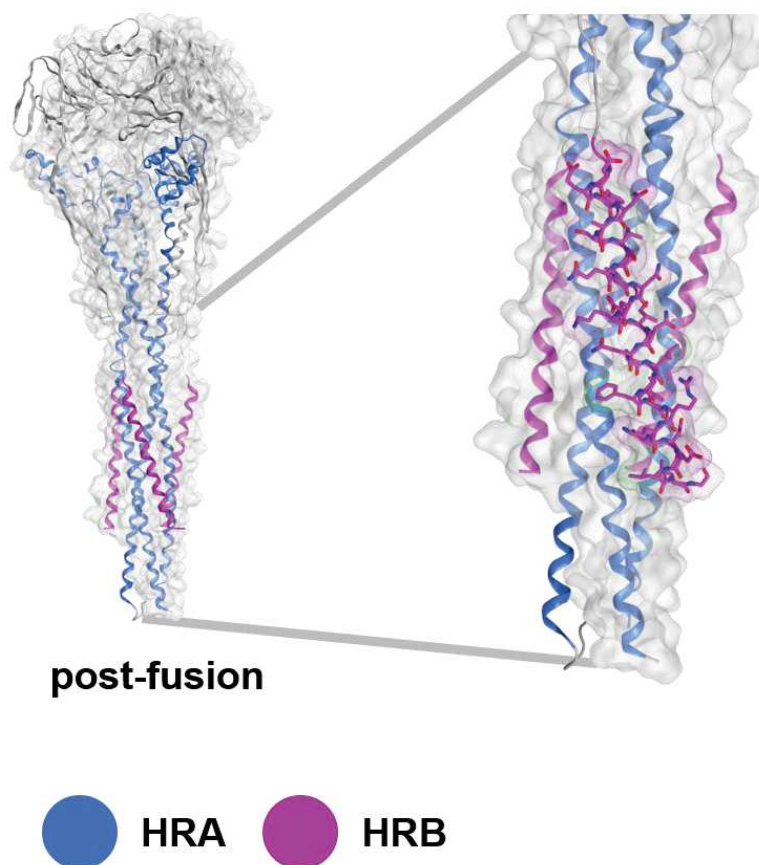


Figure 28. Representation of the 6-HB motif.

5.2.1 Virtual library generation

Starting from this information and based on the model of the fusion process, the FDASISQVN fragment of the HRB helix (Fig. 29) was selected as a template for the rational design of α -helix mimics.

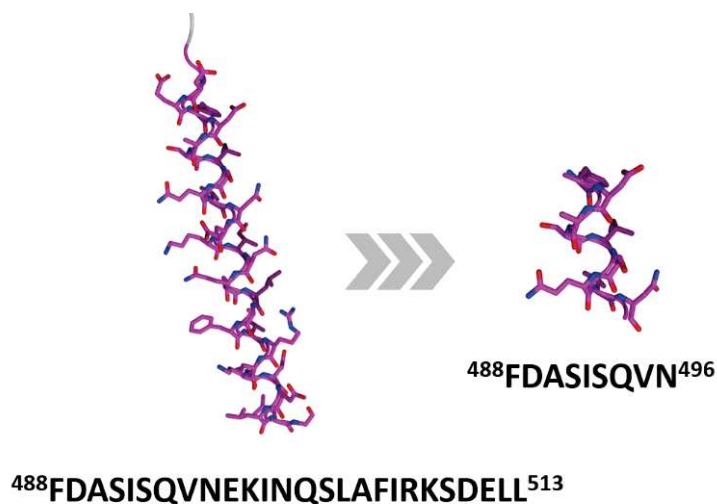


Figure 29. HRB Selected Fragment.

This fragment shows hydrophobic Phe, Ile and Val residues in positions i , $i+4$, $i+7$ and can be speculated that in the hydrophobic driven rearrangement of the F protein, the order of importance of the hydrophobic amino acids is the Phe>Ile>Val.

A series of scaffolds reported from conformational analysis to be able to mimic this kind of positions topologically were selected and explored.¹²⁷⁻¹³³

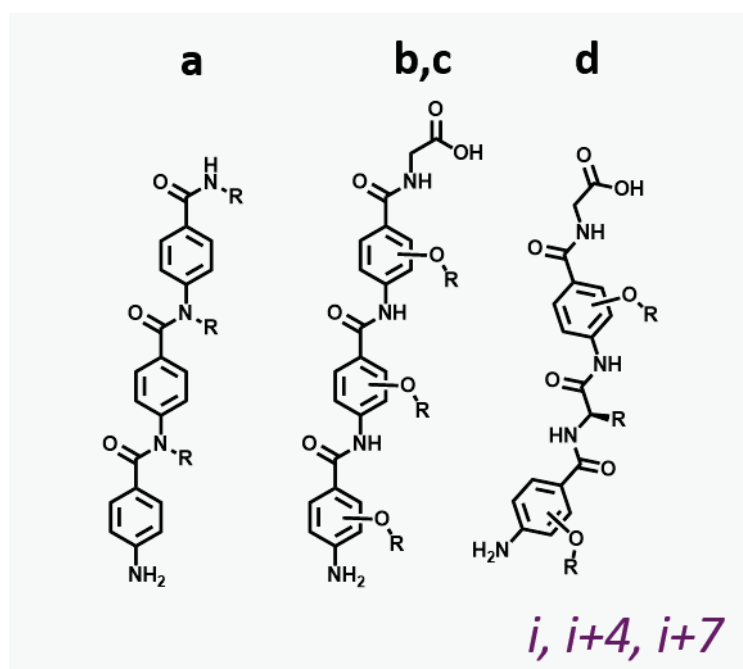


Figure 30. i , $i+4$, $i+7$ mimic scaffold.

Four diverse scaffolds were chosen; their structures are shown in Figure 30. Scaffold **a** is an *N*-alkylated aromatic oligoarylamide, **b** and **c** are 2-*O*-alkylated and 3-*O*-alkylated oligoarylamide, and **d** is a hybrid scaffold, composed of 2,3-*O*-alkylated arylamide and

aminoacids. To mimic the backbones of the hydrophobic residues, benzyl and phenyl moieties were selected for Phe residue. The *sec*-butyl moiety was chosen for Ile residue and an *isopropyl* moiety for Val residue.

Based on the selected scaffolds, a virtual library of compounds was built considering all the possible combinations as shown in Figure 31.

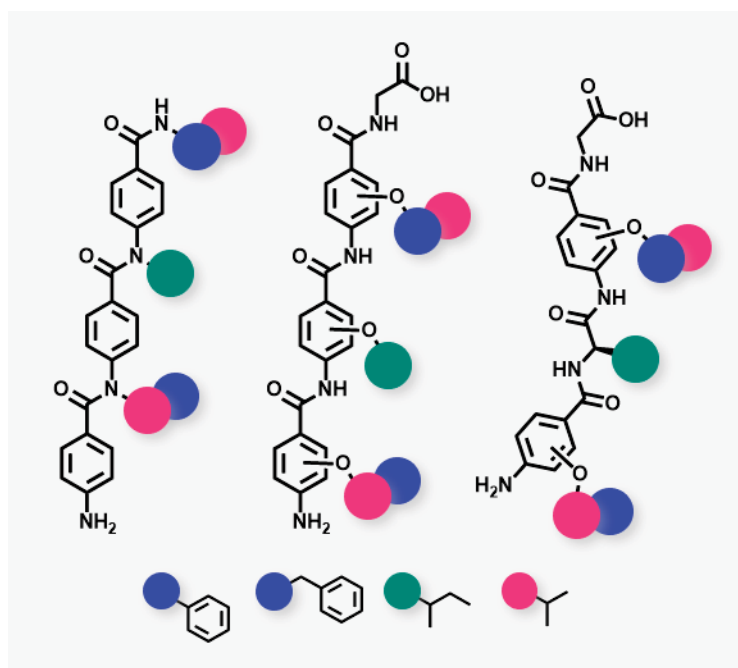


Figure 31. Generated virtual library.

5.2.2 Conformational search and Molecular docking

A conformational search using the QMD (Quenched molecular dynamic) technique was performed on the generated virtual library, using the open-source command-line tool Open3DAlign. The QMD uses short molecular dynamics simulations to find the most stable conformers for each given compound. The program uses the MMFF94 force field for atom assignment and minimisation of the generated structures. The conformers generated with this conformational analysis were manually superimposed to the target sequence and analysed through a visual inspection. Only the O-alkylated scaffolds **a** and **b** and the hybrid scaffold **d** have a topology for the hydrophobic moieties in the correct position to mimic the hydrophobic amino acidic residues of the target sequence (Figure 32).

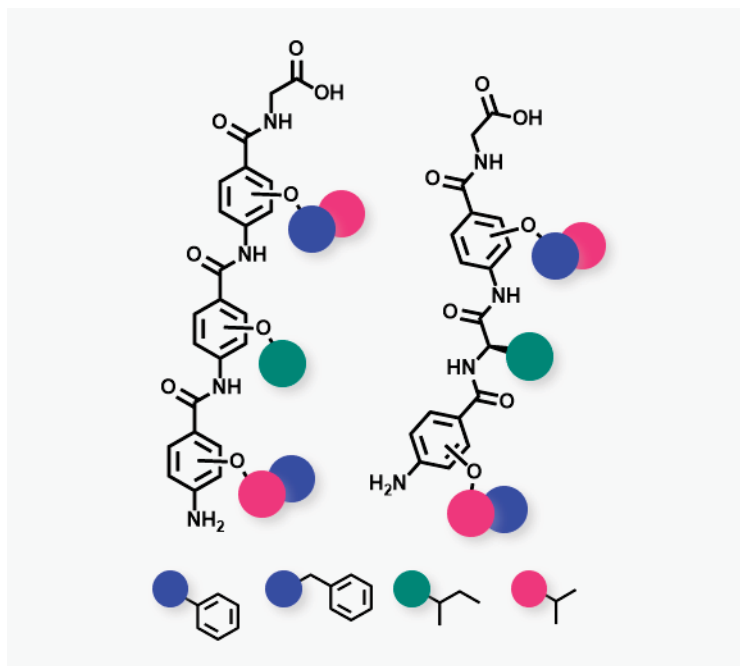


Figure 32. Selected scaffolds.

These compounds were then docked using Glide SP into a grid generated from the 6-HB of 3RRR crystal, after removing HRB peptides from the 6-HB. From the binding modes obtained only **compound 126** (Figure 33) showed the correct orientation in the binding site to mimic the α -helix fragment of HRB. This compound was therefore selected for chemical synthesis and biological evaluation.

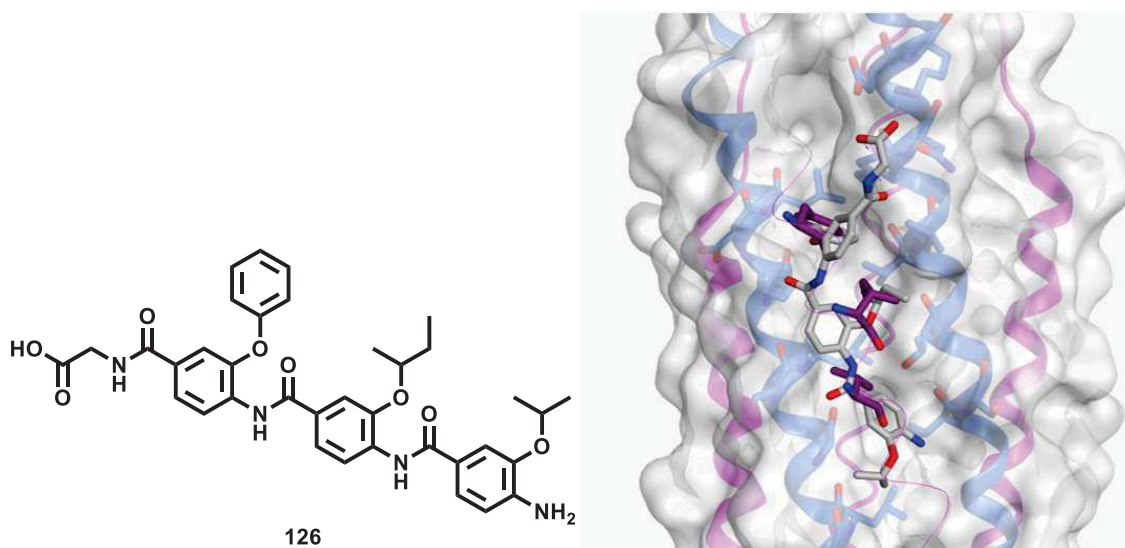


Figure 33. Selected compound 126 and docked pose.

5.3 Synthesis of selected α -helix mimic compound

The selected compound **126** structure, having a 3-O-alkylated oligoarylamide scaffold, can be split into the three 4-amino-3-O-functionalised benzoic acid building blocks (BB) represented in figure 34 by the letter A, B and C, which are the result of the amide bond disconnections showed with red lines in figure 34. Finally, the carboxylic group of the 3-O-alkylated oligoarylamide scaffold is functionalised with a glycine.

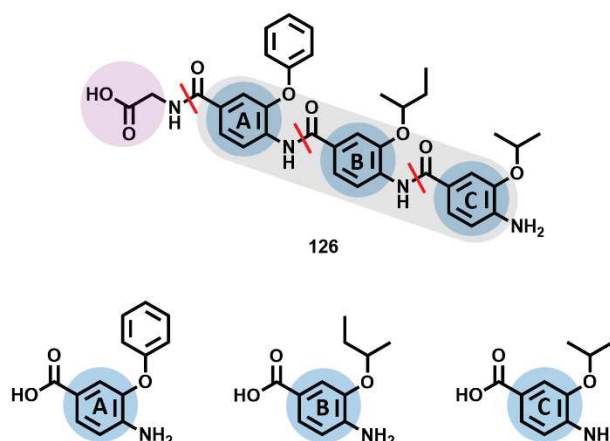


Figure 34. Structural composition of compound 126. 4-amino-3-O-functionalised benzoic acid building blocks A, B and C (in blue); in grey the oligoarylamide scaffold; in purple the glycine; the red lines represent the amide bond disconnections.

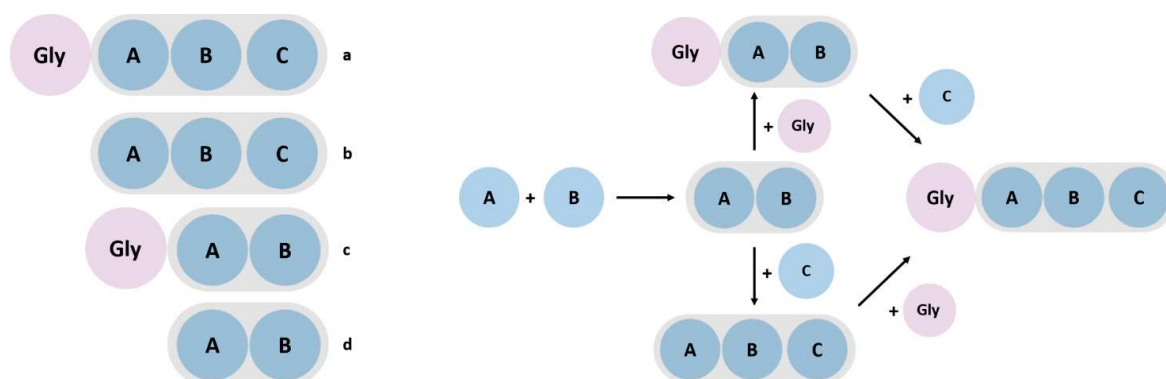
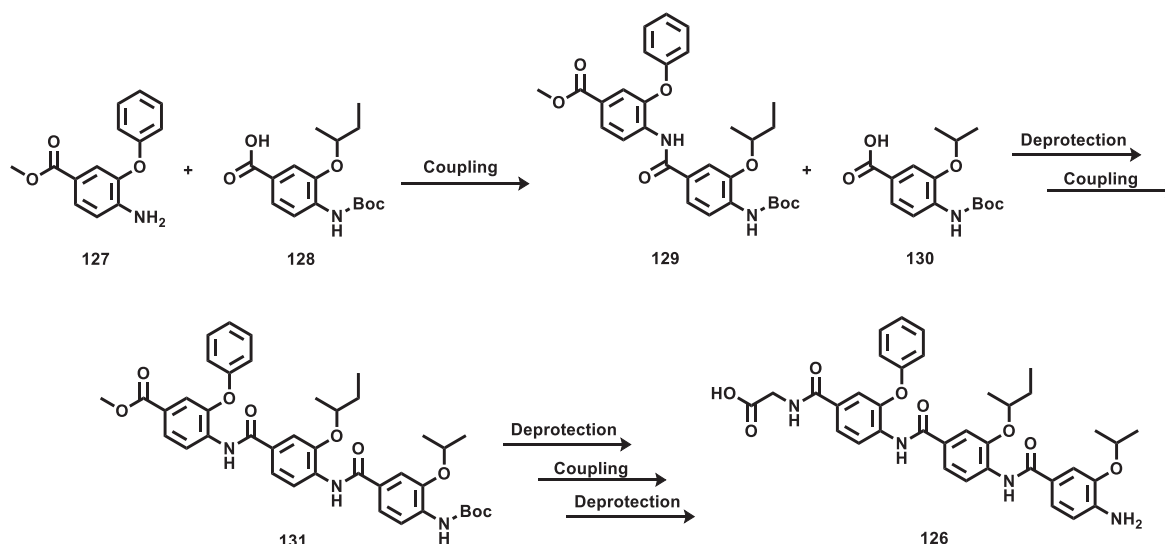


Figure 35. Selected intermediate to be synthesised and evaluated.

To validate the α -helix mimicry approach for the inhibition of the F protein rearrangement was decided to synthesise and biologically evaluate, along with the selected compound **126**, the synthetic intermediates (see figure 35). These intermediates could provide some evidence of the importance of the different hydrophobic backbones introduced in the α -helix mimic compound **126**.

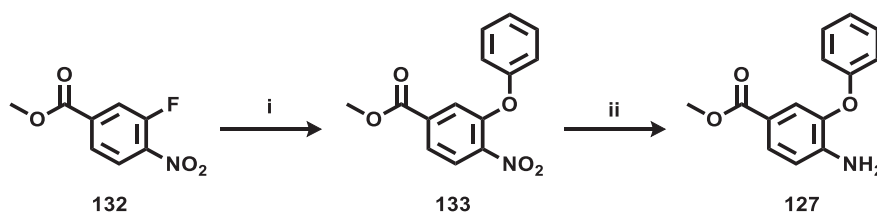
During the planning of the synthesis for the coupling of the three building blocks, it was decided to protect the amino group of the 4-amino-3-O-functionalised benzoic acids B and C.



Scheme 22. Planned synthetic route for compound **126**.

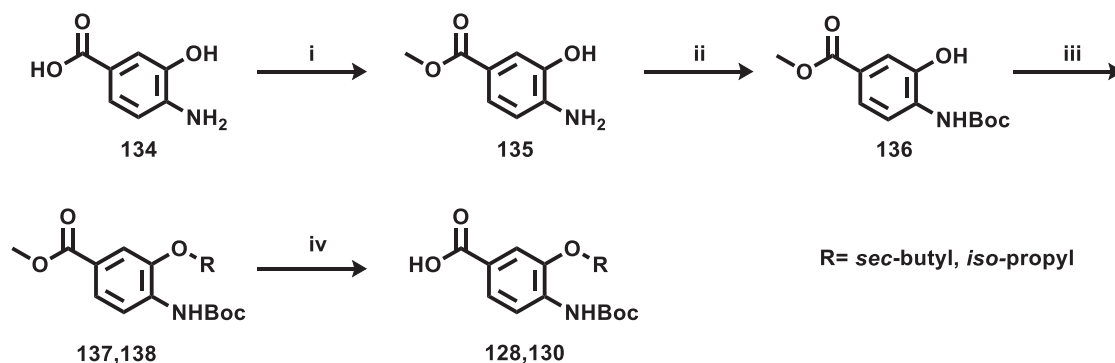
The synthetic route chosen for compound **126**, shown in Scheme 22, consists of three cycles of coupling and deprotection. This synthetic route was designed for the possible future generation of alpha-helix mimics libraries.

Building blocks for the synthesis of the desired compound are not commercially available. Therefore, they were synthesised as shown in Scheme 23 and Scheme 24. **127** was synthesised through the synthesis shown in scheme 24, starting from methyl 3-fluoro-4-nitrobenzoate **132**. First, the fluoro group of **133** was replaced by a phenoxy group with an S_NAr , and then the nitro group was reduced by hydrogenation to give the desired BB **127**.¹³¹



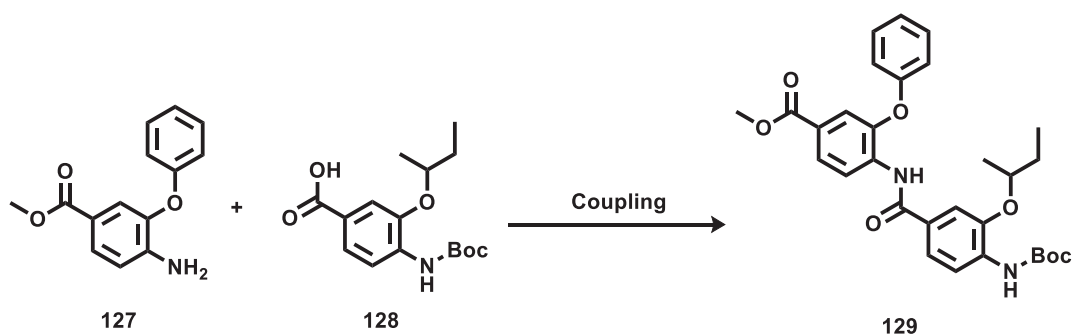
Scheme 23. Synthesis BB **127**; i) Phenol, K_2CO_3 , DMF, rt, o.n., Y=99%; ii) H_2 , Pd/C, MeOH/EtOAc (3:1), rt, o.n., Y=89%.

The other two building blocks **128** and **130** were synthesised both from 4-amino-3-hydroxybenzoic acid **134**. After the protection of the carboxylic acid as methyl ester, the amine group was protected using the *tert*-butyloxycarbonyl protecting group. These two protections were necessary for the next step of *O*-alkylation, which after basic hydrolysis gives the desired building blocks **128** and **130**.¹³⁴⁻¹³⁶



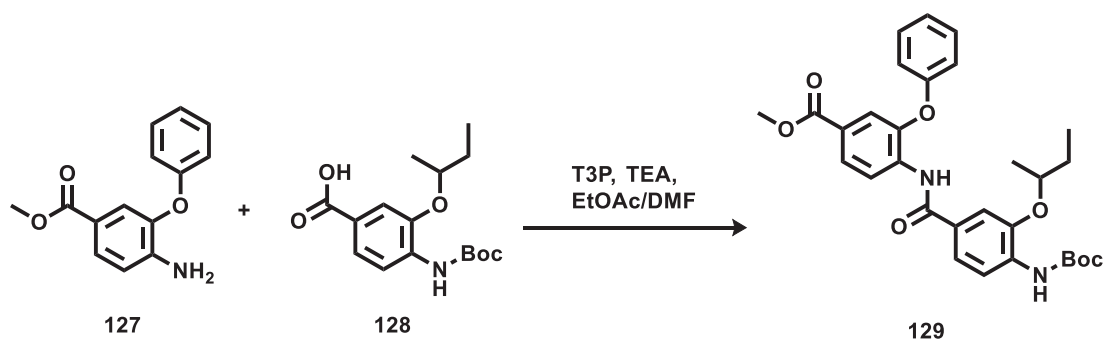
Scheme 24. Synthesis BBs 128, 130. i) SOCl_2 , MeOH, rt, o.n., Y=99%; ii) $(\text{Boc})_2\text{O}$, TEA, DCM, rt, 24h, Y=99%; iii) 2-bromoalkane, NaH, DMF, rt, on, Y=67-70%; iv) NaOH, THF/MeOH (2:1), rt, 3h, Y=88%;

Different reaction conditions were used for the amide coupling of the building blocks **127** and **128** (Scheme 25).

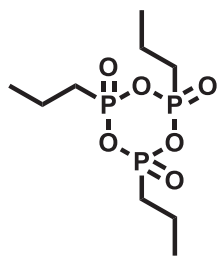


Scheme 25. Coupling between building block 127 and 128.

The first attempt was made using the T3P coupling reagent (Scheme 26). T3P **139** is a propyl phosphonic anhydride which is able to convert the oxygen of the carboxylic group (**128**) in a suitable leaving group, as shown in scheme 27.¹³⁷



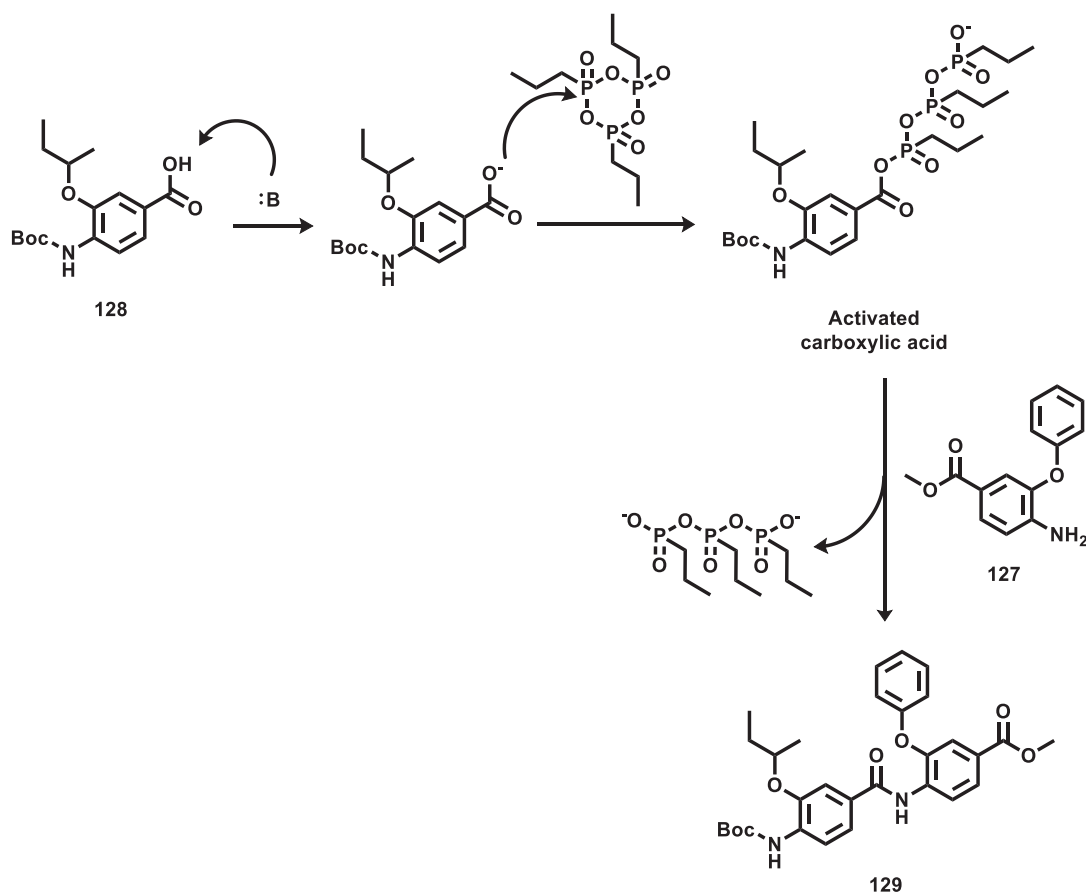
Scheme 26. Coupling attempt with T3P.



139

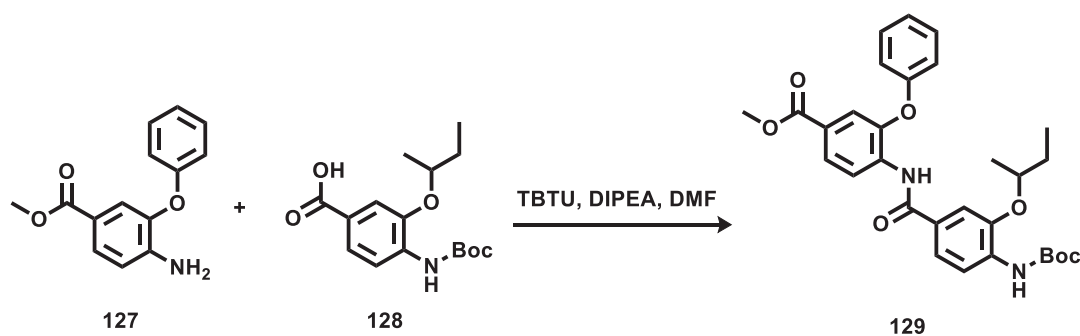
Figure 36. T3P.

The by-product of the leaving group is easily removable from the reaction mixture, making the T3P reagent a promising coupling reagent. Unfortunately, after several attempts, the formation of the desired product was not observed. This result can be related to the low reactivity of the aniline amine **127**, and to the fact that the phenoxy group hinders the amine group.



Scheme 27. T3P reaction mechanism.

Other several attempts were performed using the TBTU (*O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate **140** (Figure 37), amide coupling reagent (Scheme 28).



Scheme 28. Coupling attempt with TBTU.

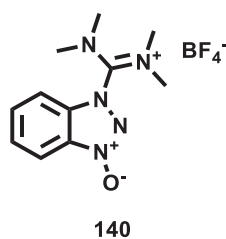
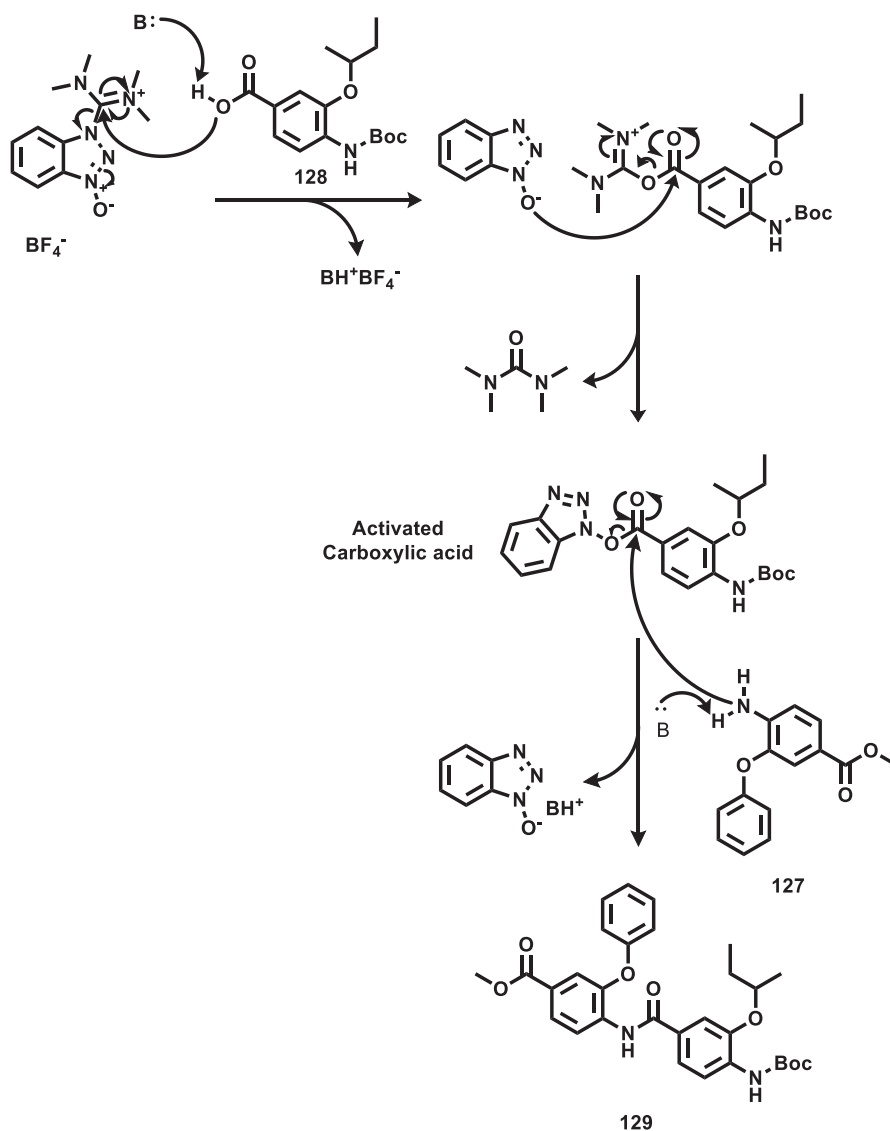


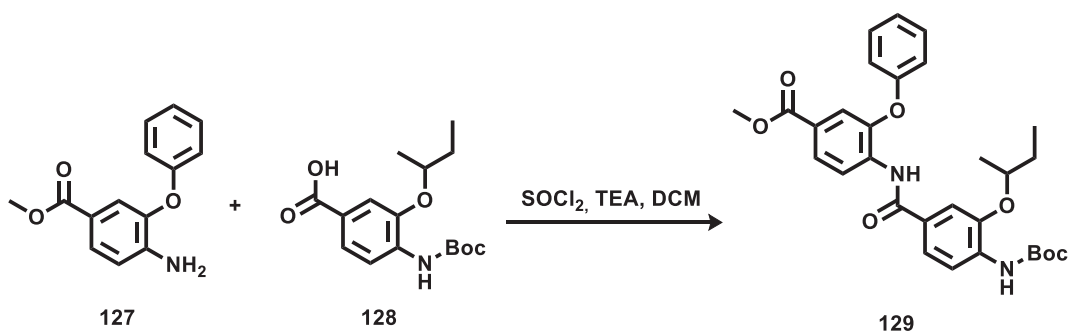
Figure 37. TBTU.

Also, in this case, TBTU is necessary for the activation of the carboxylic acid to transform the oxygen in a leaving group that can be replaced by the amine (Scheme 29). Also, in this case, as for the T3P reactions, the formation of the desired product **129** was not observed. This result, as was supposed for the T3P, can be related to the low reactivity of the aniline amine and for the fact that the amine group is hindered by the phenoxy group.¹³⁸



Scheme 29. TBTU reaction mechanism.

After the results obtained with T3P and TBTU, an attempt using thionyl chloride to generate the acyl chloride in situ was made (Scheme 30).¹³⁹



Scheme 30. Thionyl chloride mediate coupling.

In this case, the desired product **129** was formed, but the reaction gave low yield and several by-products. The major by-product was **128**, with the unprotected amine, due to the acid conditions given by the thionyl chloride in the reaction.

Even if the yield was low, the coupling between the anilinic amine and the acyl chloride seems to work. For this reason, another protecting group or strategy needs to be introduced for the protection of the anilinic amine.

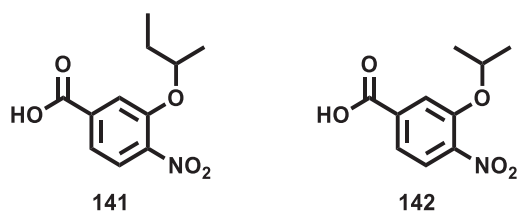
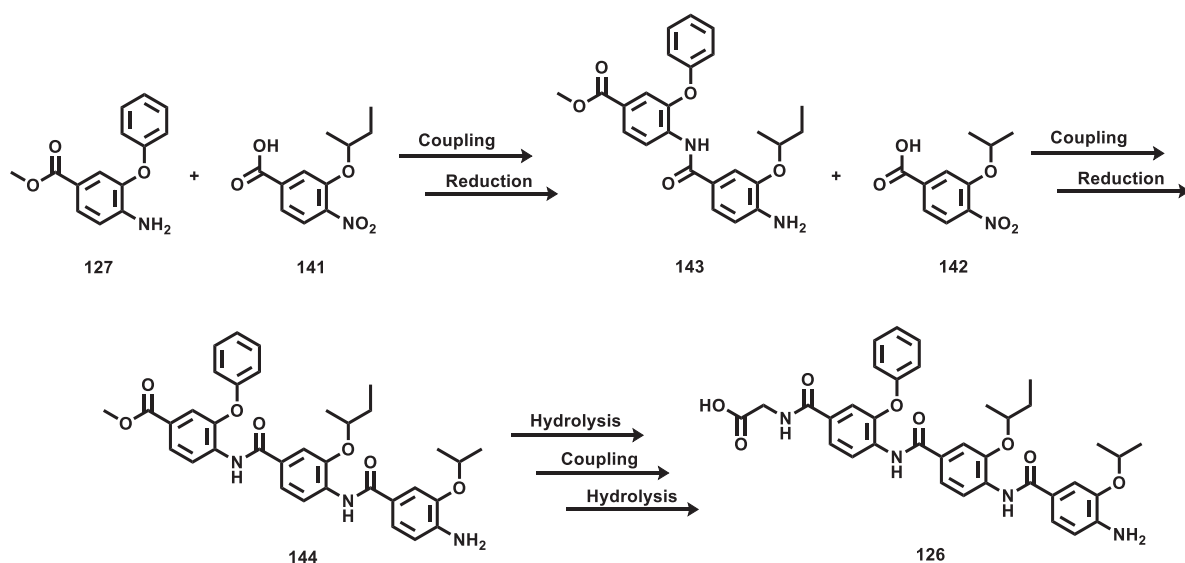


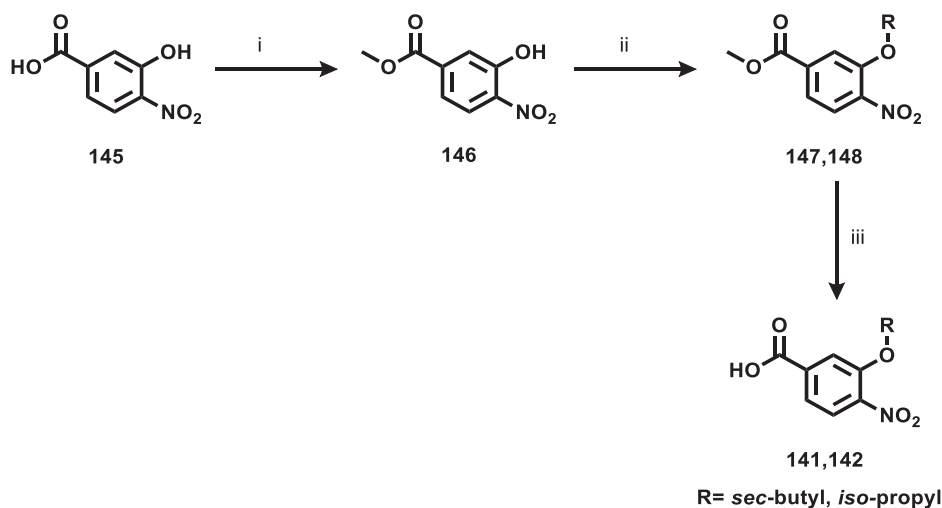
Figure 38. New building blocks B and C.

The previously performed conversion of the BB A **122**, from the nitro precursor **133**, suggested the use of the nitro group as an un-reactive group. This strategy will avoid the formation of any synthetic by-product during the coupling reaction between BBs A and B. For this reason, compounds **141** and **142** were decided to be used instead of the Boc protected BBs, and a new synthetic route was planned, see scheme 31. In this route the oligoarylamide scaffold **144**, formed by the BBs A, B and C are synthesised by two cycles of coupling and reduction of the nitro group to amino group. The final functionalisation of the oligoarylamide scaffold **144** with the glycine is planned to be obtained by basic hydrolysis of the methyl ester, coupling with the methyl glycinate and finally, another step of basic hydrolysis to give compound **126**.



Scheme 31. New planned synthetic route for compound 126.

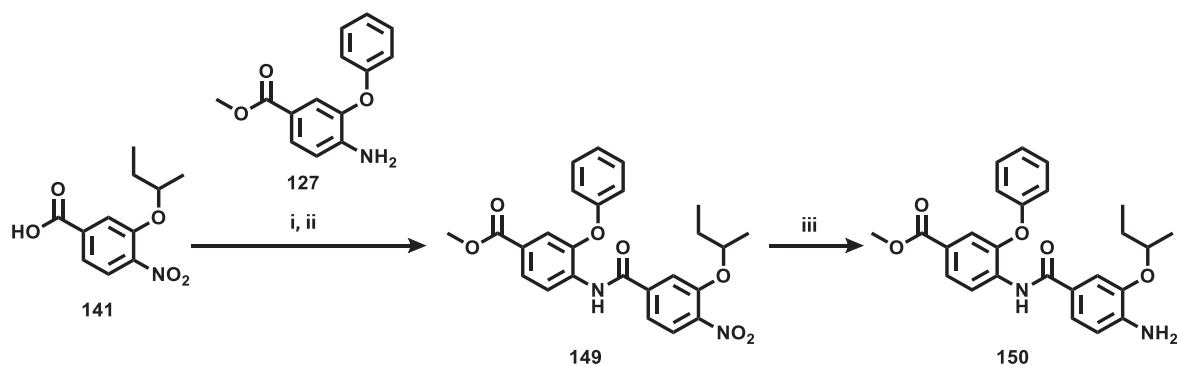
Like the other building blocks (**127**, **128** and **130**), the 3-alkoxy-4-nitrobenzoic acids are not commercially available. As can be seen in scheme 32, compounds **141** and **142** were synthesised in a three steps synthetic pathway.



Scheme 32. Synthesis BBs 141, 142. i) SOCl_2 , MeOH, rt, o.n., Y=98%; ii) bromoalkane, NaH, DMF, 70°C, 6h, Y=84-98%; iii) NaOH, THF/MeOH (2:1), rt, 3h, Y=86-93%.

For the coupling between the building blocks **A** and **B**, was decided to try the coupling procedure developed for the synthesis of compound **2**. The 3-(*sec*-butoxy)-4-nitrobenzoic acid **141** was converted into the acyl chloride and then added drop-wise to a DCM solution containing the BB **A** (**127**) and triethylamine, giving the methyl 4-(3-(*sec*-butoxy)-4-nitrobenzamido)-3-phenoxybenzoate **149** with an overall yield over two steps of 86% (scheme 33).

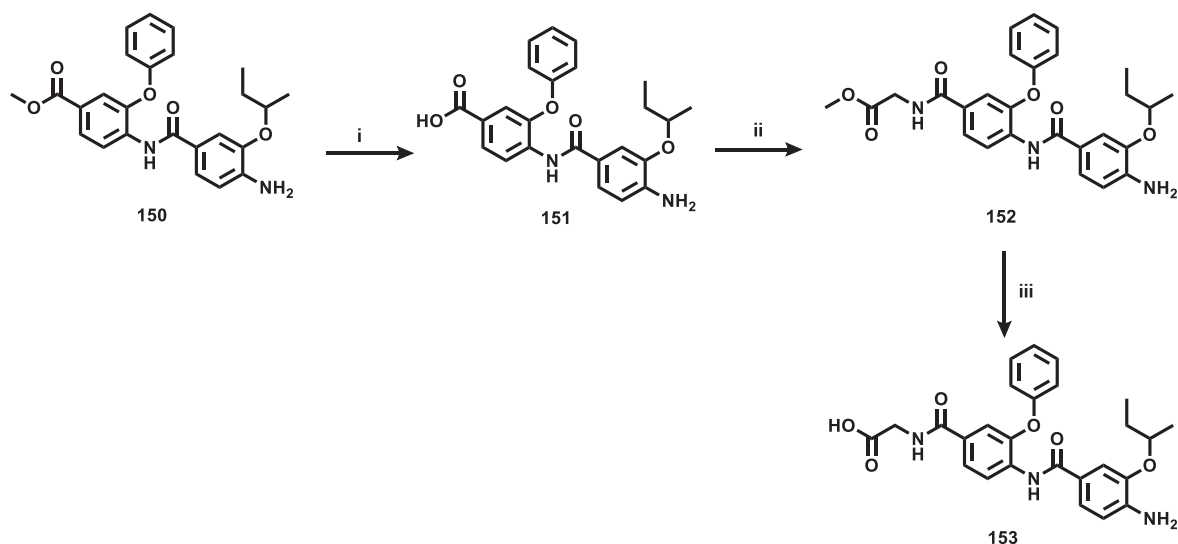
Then the obtained compound **149** was successfully reduced by hydrogenation the relative amine **150**.



Scheme 33. BB A and B coupling and reduction. i) SOCl_2 neat, reflux, 4h; ii) TEA, DCM, rt, o.n., (over 2 steps Y=86%); iii) H_2 , Pd/C, EtOAc, rt, o.n., Y=90%.

The result obtained by the synthesis of the essential intermediate **150**, gave a hint of the feasibility of the applied synthetic strategy for the coupling between the building blocks and the use of the nitro group to mask the anilinic amine. Compound **150** is the crucial synthetic intermediate for the synthesis of the other intermediates.

The methyl 4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoate **150** was hydrolysed in basic condition to give the free acid derivative **151** which, was then functionalised with the glycine. For doing this, the TBTU coupling reagent to activate the carboxylic acid. In this case, the glycine amine is more reactive than the anilinic amine and for this reason, can replace the activated benzotriazole ester. The obtained methyl (4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycinate **152** was then hydrolysed in basic condition to give the free acid derivative **153** (scheme 34).



Scheme 34. Synthesis of compound 153. i) NaOH, THF/MeOH (2:1), rt, 3h, Y=93%; ii) methyl glycinate hydrochloride, TBTU, DIPEA, DMF, rt, o.n., Y=94%; iii) LiOH, THF/MeOH (2:1), rt, on, Y=73%.

The following stage was to synthesise the oligoarylamide **154** with the three hydrophobic features. As shown in scheme 35, this was obtained through the three steps (coupling, reduction of the nitro group and methyl ester hydrolysis) synthesis starting from the intermediate **150**. The reduction step was optimised adding the THF as a co-solvent to increase the solubility of the starting material **155**, which was partially soluble in EtOAc.

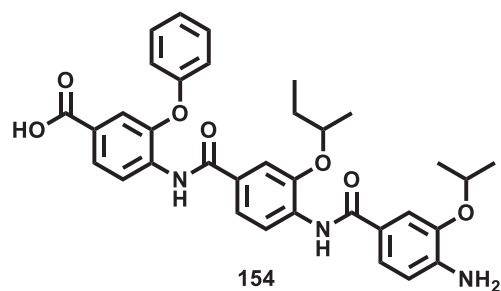
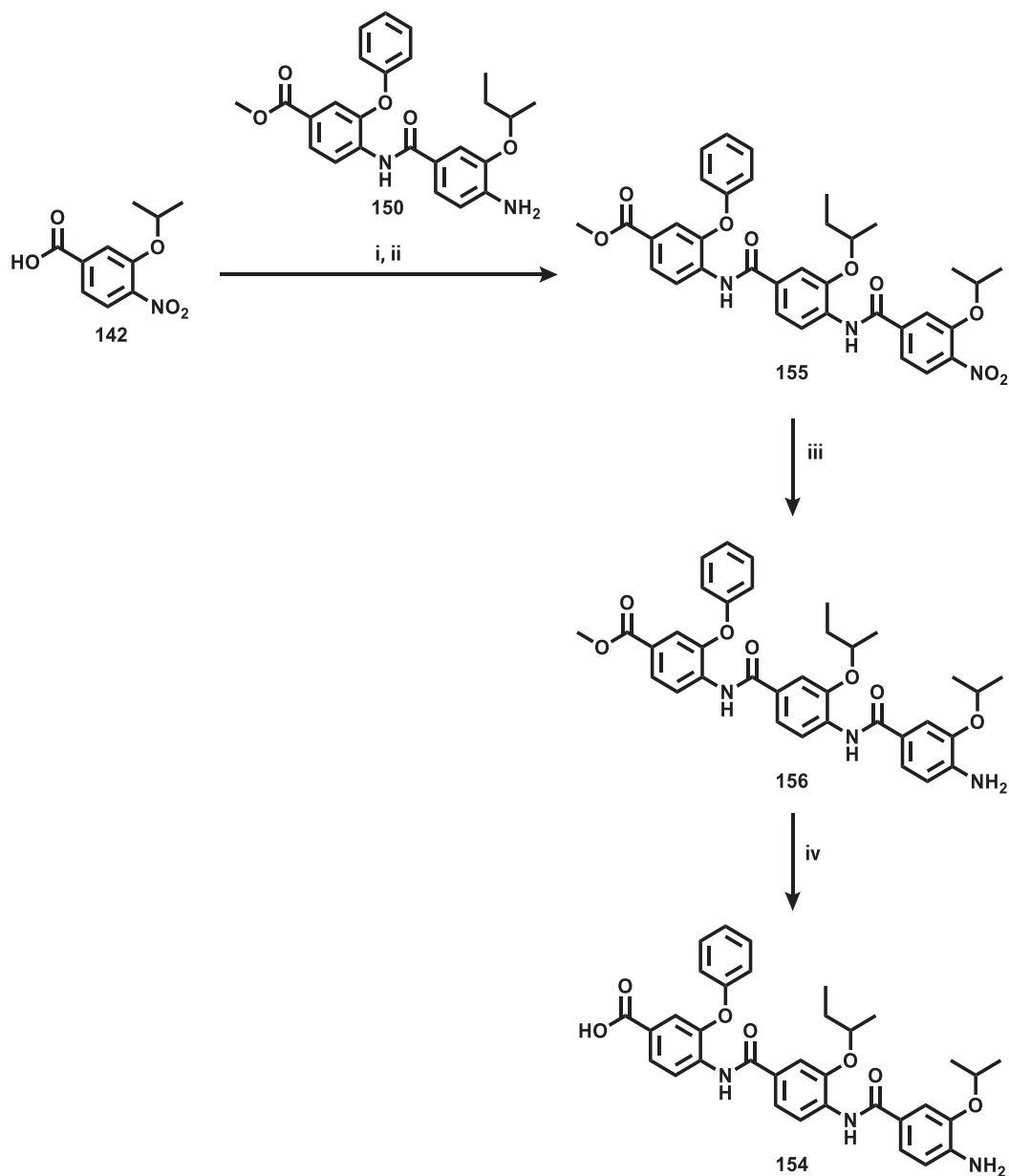
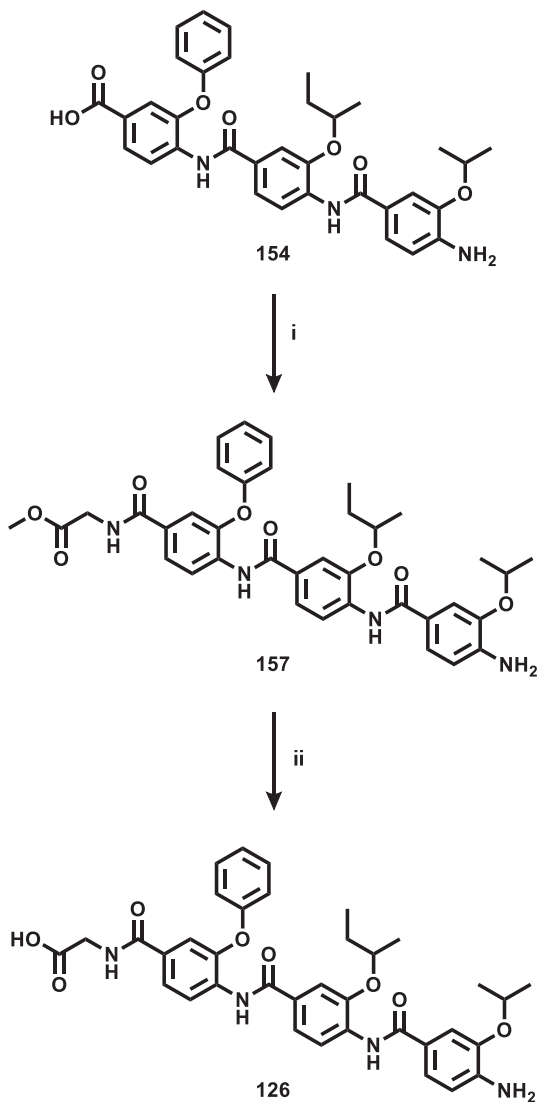


Figure 39. Oligoarylamide 154.



Scheme 35. Synthesis of compound 154. i) SOCl_2 neat, reflux, 4h; ii) TEA, DCM, rt, o.n., (over 2 steps $Y=87\%$); iii) H_2 , Pd/C, EtOAc/THF, rt, o.n., $Y=88\%$; iv) NaOH, THF/MeOH (2:1), rt, 3h, $Y=44\%$.

The α -helix mimic, (4-(4-(4-amino-3-isopropoxybenzamido)-3-(sec-butoxy) benzamido)-3-phenoxybenzoyl)glycine **126**, was obtained by coupling reaction between the intermediate **154**, followed by basic hydrolysis of the methyl ester derivative **157** (Scheme 36).

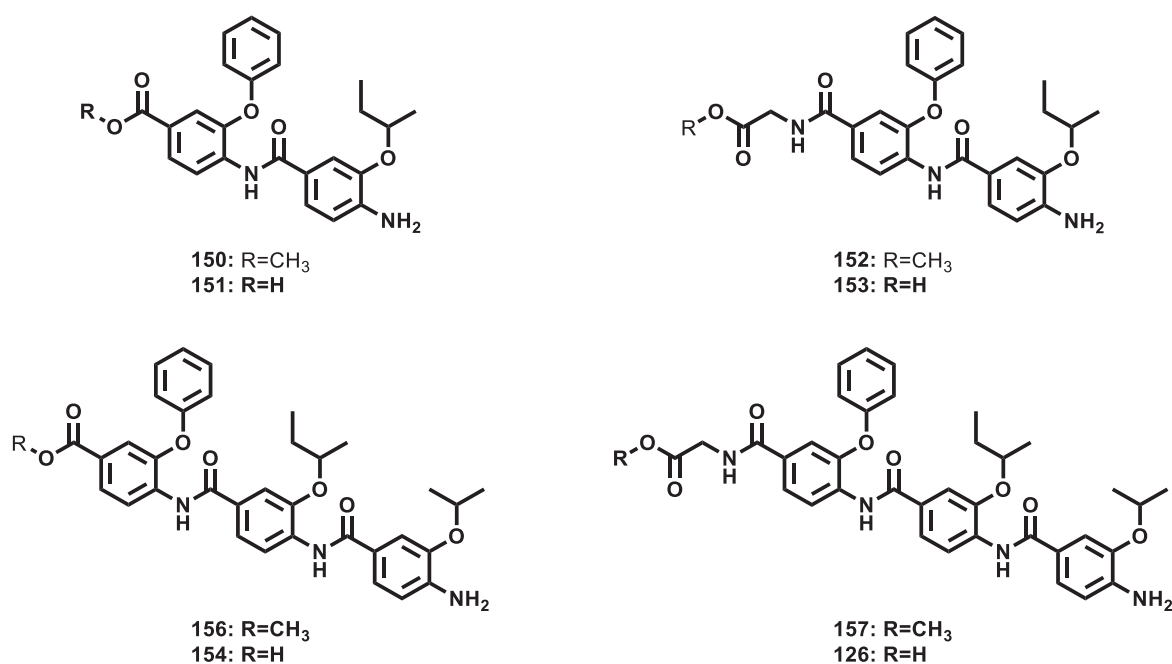


Scheme 36. Synthesis of compound 126. i) methyl glycinate hydrochloride, TBTU, DIPEA, DMF, rt, o.n., Y=99%; ii) LiOH, THF/MeOH (2:1), rt, on, Y=59%

5.3.1 Biological evaluation of selected α -helix mimic

The synthesised α -helix mimic **126** and the selected intermediate were biologically evaluated in CPE antiviral assay in Hep2 cells line, infected with RSV-A2 and RSV-B and the cytotoxicity evaluated by MTS readout (table 26).

Table 25. Antiviral and cytotoxicity data for compound 126 and the selected intermediates. Compounds were biologically evaluated in cell based antiviral assay (AVA) against RSV-A2 and RSV-B strains and for its cytotoxicity. Data are representative of two independent experiments. NA = not active. SI = selectivity index (CC_{50}/EC_{50})



Compounds	EC ₅₀ RSV-A2 (μ M)	EC ₅₀ RSV-B (μ M)	CC ₅₀	SI RSV-A2	SI RSV-B
150	NA	NA	>300	-	-
151	NA	NA	>300	-	-
152	NA	NA	>300	-	-
153	NA	NA	>300	-	-
156	NA	NA	>300	-	-
154	NA	NA	13	-	-
157	NA	NA	>300	-	-
126	12	12	39	3.2	3.2

From the biological evaluation, only the designed α -helix mimic **126** among the synthesised intermediates shown an antiviral activity of 12 μ M against both RSV strains, with a CC_{50} of 39 μ M and an SI of 3.2.

5.4 Synthesis of 126 α -helix mimic analogues

Since the results obtained from the antiviral evaluation of compound **126** and its intermediates in which the addition 4-amino-3-isopropoxybenzamido moiety is associated with the antiviral activity, it was decided to substitute the isopropoxy group in position 3 on the last benzamido moiety (building block **C**) to position 2 (compound **158** in figure 40). This change will give information on the antiviral activity associated with the orientation of the 3-isopropoxy group.

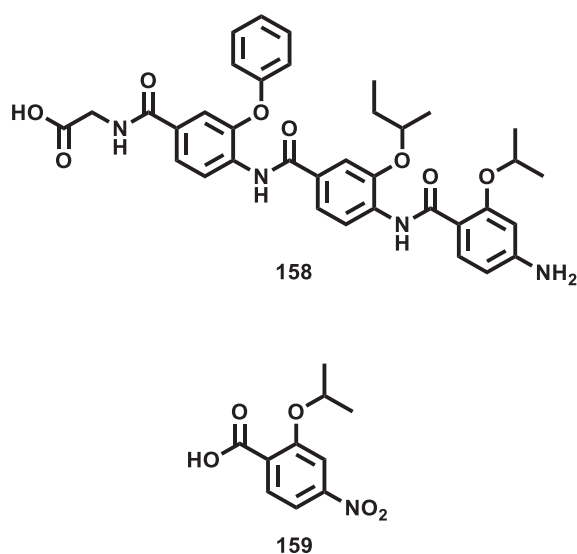
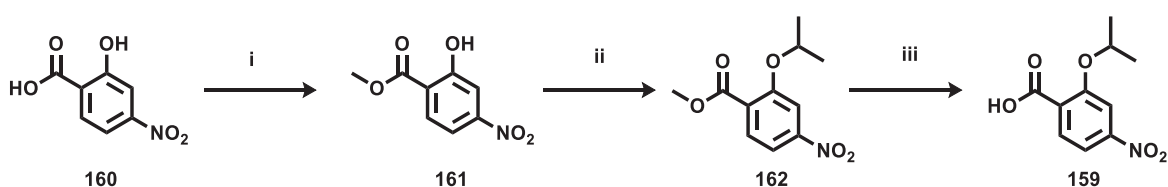


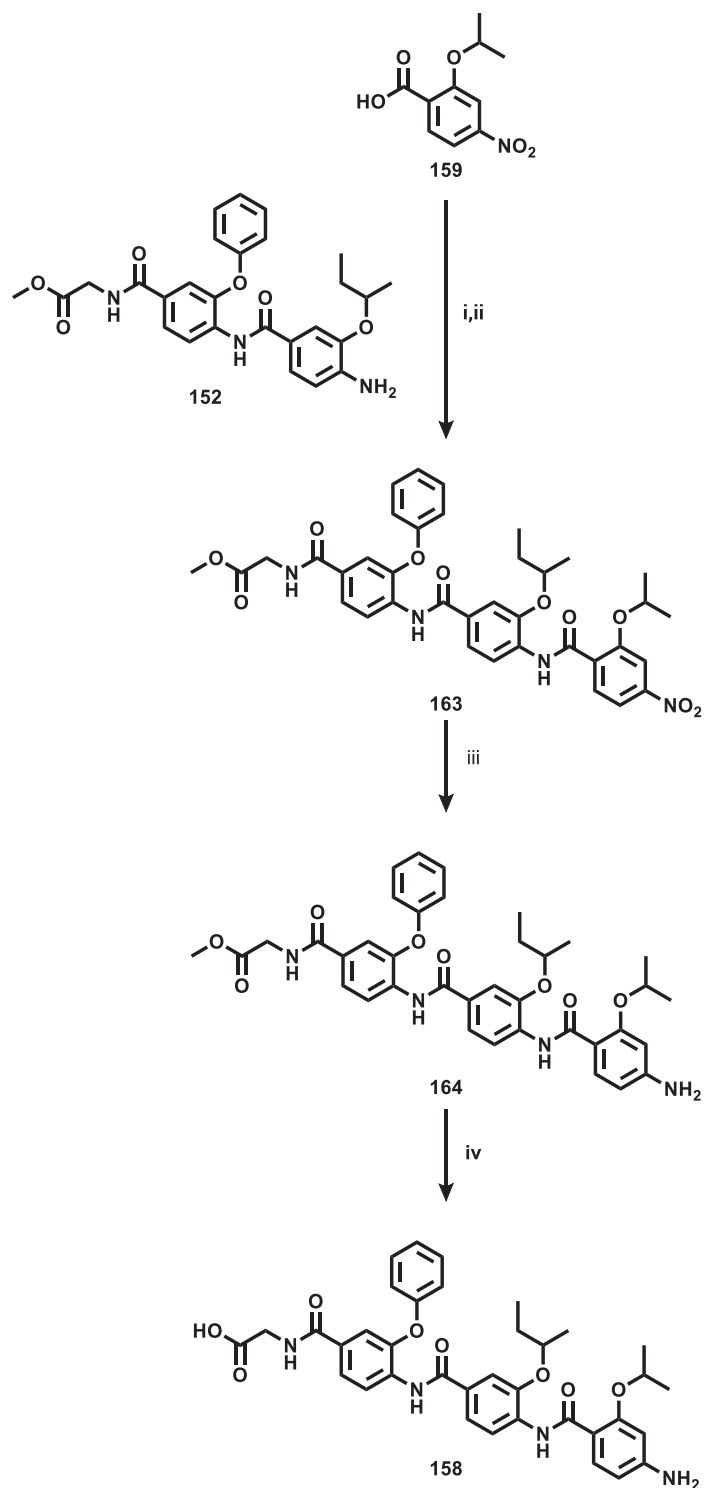
Figure 40. Compound **158** and building block **159**.

For the introduction of the planned modification, it was necessary to synthesise the new building block, the 2-isopropoxy-4-nitrobenzoic acid **159** (see figure 40). The building block was synthesised using the three steps synthetic pathway used for the synthesis of building blocks **141** and **142**, as shown in scheme 37.



Scheme 37. Synthesis BB **159**. i) SOCl_2 , MeOH, rt, o.n., Y=89%; ii) 2-bromopropane, NaH, DMF, 70°C, 6h, Y=90%; iii) NaOH, THF/MeOH (2:1), rt, 3h, Y=92%;

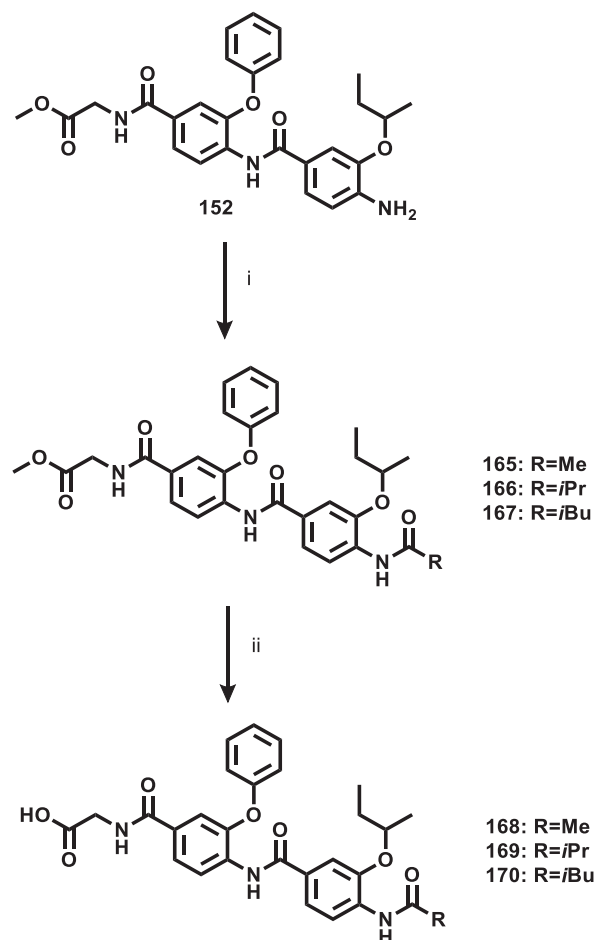
The obtained building block **159**, was then coupled to the intermediate **152**, using the optimised three steps of synthesis as shown in scheme 38 to give the desired product **158**.



Scheme 38. Synthesis of compound 158. i) SOCl_2 neat, reflux, 4h; ii) TEA, DCM, rt, o.n., (over 2 steps $Y=95\%$); iii) H_2 , Pd/C, EtOAc/THF, rt, o.n., $Y=95\%$; iv) LiOH, THF/MeOH (2:1), rt, 3h, $Y=80\%$.

To further investigate the influence of the 4-amino-3-isopropoxybenzamido moiety on the antiviral activity of compound **126**, it was decided to functionalised intermediate **152** with three acyl chlorides, respectively acetyl chloride, isobutyryl chloride and 3-methylbutanoyl chloride.

The acyl chlorides were selected to introduce a methyl, an isopropyl and an isobutyl groups on the compound **152**. The aim was to confirm the importance of the 4-amino-3-isopropoxybenzamido moiety of compound **126**, for the right orientation of the hydrophobic isopropyl group. As can be seen in scheme 39, the derivatives of compound **152** were obtained through the coupling between the selected acyl chloride and the intermediate **152**, followed by methyl ester hydrolysis in basic condition.

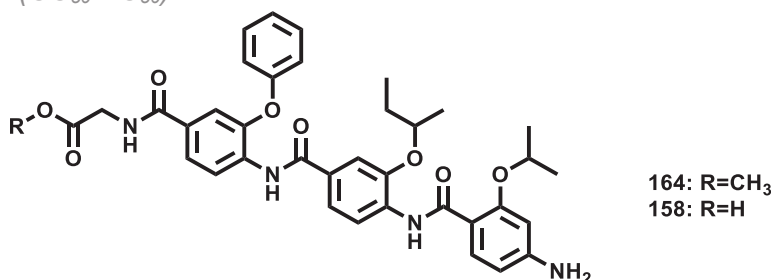


Scheme 39. Synthesis of the derivatives of compound **152**. i) Acyl chloride, TEA, DCM, rt, o.n., Y=60-86%; ii) LiOH, THF/MeOH (2:1), rt, 3h, Y=82-93%.

5.4.1 Biological evaluation of 126 α -helix mimic analogues

From the biological evaluation of the analogues of compound **126**, any of the modification introduced to the original structure give retention or improvement of the antiviral activity (see table 27 and 28). The substitution of the isopropoxy group in position 3 on the last benzamido moiety to position 2 in compound **158** is suggesting that the position of the isopropoxy group is essential for the antiviral activity.

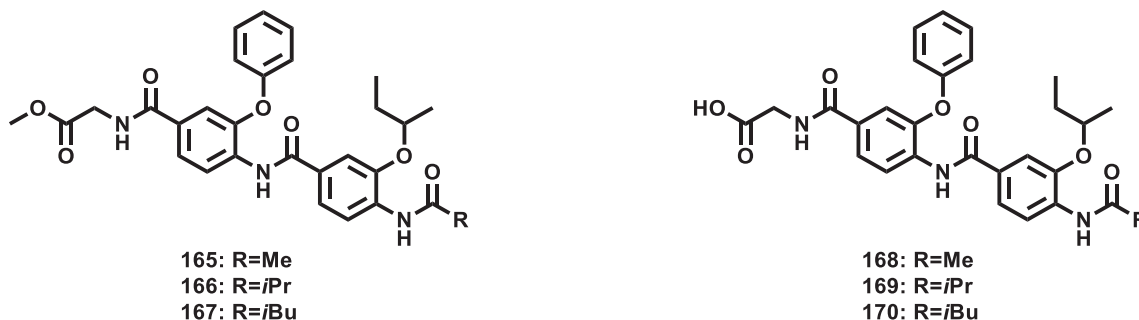
Table 26. Antiviral evaluation of compounds 158 and 164. Compounds were biologically evaluated in cell based antiviral assay (AVA) against RSV-A2 and RSV-B strains and for its cytotoxicity. Data are representative of two independent experiments. NA = not active. SI = selectivity index (CC_{50}/EC_{50})



Compounds	EC ₅₀ RSV-A2	EC ₅₀ RSV-B	CC ₅₀
164	NA	NA	>300
158	NA	NA	65

Also, the three alkylamide derivatives of compound **152** were not associated with any antiviral activity against RSV, suggesting that 4-amino-3-isopropoxybenzamido moiety in compound **126** is important to eventually orientate the isopropoxy group in the right position to bind with the protein.

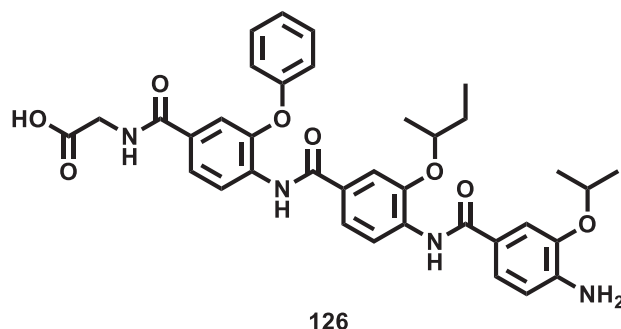
Table 27. Antiviral evaluation of compounds 165 and 170. Compounds were biologically evaluated in cell based antiviral assay (AVA) against RSV-A2 and RSV-B strains and for its cytotoxicity. Data are representative of two independent experiments. NA = not active. SI = selectivity index (CC_{50}/EC_{50})



Compounds	EC ₅₀	EC ₅₀	CC ₅₀
	RSV-A2	RSV-B	
165	NA	NA	>300
166	NA	NA	>300
167	NA	NA	>300
168	NA	NA	>300
169	NA	NA	198
170	NA	NA	>300

5.5 Conclusion

Several chemoinformatic techniques were used to generate and evaluate a focussed virtual library of compounds designed to mimic the hydrophobic backbones in positions *i*, *i*+4, *i*+7 of the ⁴⁸⁸FDASISQVN⁴⁹⁶ fragment. The generated library was used for a structure-based virtual screening on the X-ray structure of the post-fusion F protein. Only compound **126** seemed to be able to have the hydrophobic groups in the proper orientation to mimic the hydrophobic backbones of the Phe488, Ile492 and Val495.



A highly versatile synthetic route was developed for the synthesis of the selected α -helix mimic **126**. The compound and the different intermediates were tested. Only the selected compound showed an antiviral activity, but with a low CC₅₀ and SI.

The modifications on the selected compound suggested that the 4-amino-3-isopropoxybenzamido moiety and the topological orientation of the isopropoxy seem to be crucial for the antiviral activity.

Future modifications on the selected compound could be to improve the chemical-physical properties of the scaffold introducing alkyloxypicolonamido or alkyloxynicotinamido building blocks in the scaffold, to increase solubility. Another modification could be adding another 4-amino-alkyloxybenzamido moiety to the scaffold and eventually optimise the hydrophobic mimicry groups in order to improve the antiviral activity.

Further biological studies are ongoing to validate the hypothesis of the mode of action of the α -helix mimic approach for the inhibition of the F protein as an antiviral strategy.

Chapter 6: Experimental

6.1 General information

All chemicals, reagents and solvents were purchased from Sigma-Aldrich, FluoroChem, Apollo Scientific, TCI UK or purified by standard techniques.

Thin Layer Chromatography

Silica gel plates (Merck Kieselgel 60F254) were used and were developed by the ascending method. After solvent evaporation, compounds were visualised by irradiation with UV light at 254 nm and 366 nm.

Column Chromatography

Column Chromatography was performed using an automated Isolera One System (Biotage), using Biotage pre-packed silica cartridges (SNAP and KP types). Samples were applied as a concentrated solution in the same eluent. Fractions containing the product were identified by TLC, combined and the solvent removed *in vacuo*.

UPLC-MS analysis

UPLC-MS analysis was conducted on a Waters UPLC system with both Diode Array detection and Electrospray (+ve and -ve ion) MS detection. The stationary phase was a Waters Acquity UPLC BEH C18 1.7 μ m 2.1x50mm column. The mobile phase was H₂O containing 0.1% Formic acid (A) and MeCN containing 0.1% Formic acid (B). Column temperature: 40°C. Sample diluent: acetonitrile. Sample concentration 10 μ g/mL. Injection volume 2 μ L. Two methods were used:

Linear gradient standard method A: 90% A (0.1 min), 90%-0% A (2.6 min), 0% A (0.3 min), 90% A (0.1 min); flow rate 0.5 mL/min.

Linear gradient standard method B: 90% A (0.1 min), 90%-0% A (2.1 min), 0% A (0.8 min), 90% A (0.1 min); flow rate 0.5 mL/min.

Linear gradient standard method C: 90% A (0.1 min), 90%-0% A (1.5 min), 0% A (1.4 min), 90% A (0.1 min); flow rate 0.5 mL/min.

NMR Spectroscopy

¹H, ¹³C, NMR spectra were recorded on a Bruker AVANCE 500 spectrometer (500 MHz and 75 MHz respectively) and auto calibrated to the deuterated solvent reference peak. Chemical shifts are given in δ relative to tetramethylsilane (TMS); the coupling constants (J) are given in Hertz. TMS was used as an internal standard (δ = 0 ppm) for ¹H NMR and CDCl₃ served as an internal standard (δ = 77.0 ppm) for ¹³C NMR.

6.1.1 Molecular modelling

A PC 1.80 GHz Intel Xeon (8 cores), running Ubuntu 14.04 LTS was used for molecular modelling studies.

Two molecular modelling softwares were used: Molecular Operating Environment (MOE) 2015.10 and Maestro (Schrödinger version 10).¹⁴⁰

Pharmacophoric filters were created with PLIFs in MOE 2015.10.

Docking experiments were performed using GlideSP module in Maestro, running the default option.

Docking results were rescored using Plants ChemPLP, FlexX and Glide XP scoring function.

The best docked poses were selected using a consensus scoring function: for each scoring function, the function *fquart* was calculated to determine the threshold limit under which relies the best 25% of the scoring results (negative energy value).

Then for each scoring value of each docked pose, the function Sign was calculated. If the docked pose belongs to the best 25% the sign = +1 otherwise sign= -1.

In the end, the three sign values, generated for each scoring values obtained from each docked pose were summed and only the ones with a sign sum equal to +3 were chosen.

The *consensus* scoring function equation for a given docked pose X was calculated as follows:

$$\text{sign sum}(X) = \text{sign}(fquart_A - X_A) + \text{sign}(fquart_B - X_B) + \text{sign}(fquart_C - X_C)$$

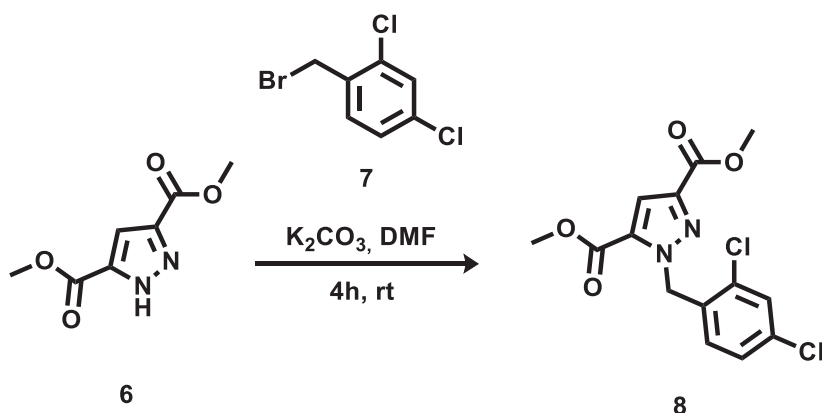
A=GlideXP; B=Plants ChemPLP; C=FlexX

Conformational search was performed with Open3DAlign version 2.3., using the default option.¹⁴¹

6.2 Synthesis of M76

dimethyl 1-(2,4-dichlorobenzyl)-1H-pyrazole-3,5-dicarboxylate (**8**)

(C₁₄H₁₂Cl₂N₂O₄, M.W.: 343.1)



A mixture of dimethyl 1H-pyrazole-3,5-dicarboxylate **6** (0.7 mmol), 1-(bromomethyl)-2,4-dichlorobenzene **7** (0.7 mmol) and K₂CO₃ (1.05 mmol) in anhydrous DMF (1.0 mL) under nitrogen atmosphere was stirred at room temperature for 4 h. The reaction mixture was quenched with 10 ml of water. The aqueous layer was extracted with EtOAc (3 x 10 mL). The organic layers were combined, washed with brine (3 x 30 mL), dried over MgSO₄ and was evaporated at reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP Sil 10g, n-hexane - EtOAc 100:0 v/v increasing to 60:40 v/v in 15 CV) to give pure dimethyl 1-(2,4-dichlorobenzyl)-1H-pyrazole-3,5-dicarboxylate **8** as a white solid.

T.L.C. System: n-hexane – EtOAc 8:2 v/v, R_f: 0.35.

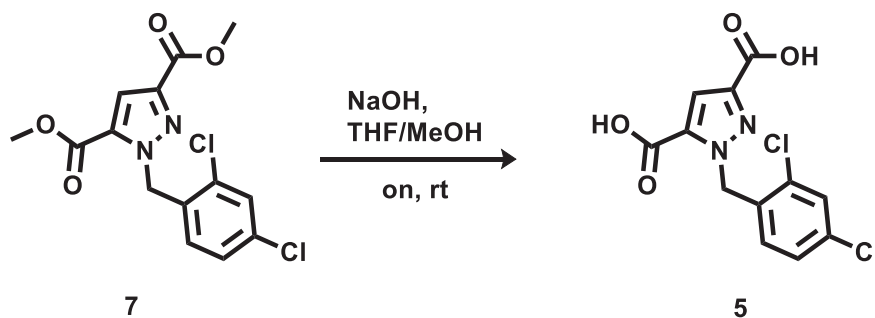
Yield: 211 mg (87%)

Purity: 99%

UPLC-MS method C: Rt: 2.00, MS (ESI)⁺: 343.1 (56%), 345.1(37%), 347.1 (7%) [M+H]⁺, 365.1[M+Na]⁺

¹H NMR (DMSO-d₆), δ: 7.70 (d, ⁴J=2.1 Hz, 1H), 7.42 – 7.36 (m, 2H), 6.76 (d, ³J=8.4 Hz, 1H), 5.88 (s, 2H), 3.83 (s, 3H), 3.82 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 161.56, 159.20 (C=O), 142.55, 134.47, 133.79, 133.65 (C, C-aromatic), 133.14, 130.18, 129.40, 128.28, 113.97(CH, C-aromatic), 53.08(CH₃), 52.96 (CH₂), 52.49 (CH₃).

1-(2,4-dichlorobenzyl)-1H-pyrazole-3,5-dicarboxylic acid (5)**(C₁₂H₈Cl₂N₂O₄, M.W.: 315.1)**

To a solution of dimethyl 1-(2,4-dichlorobenzyl)-1H-pyrazole-3,5-dicarboxylate **8** (0.38 mmol), in THF (2.0 mL), a solution of NaOH (15% m/v, 1.52 mmol) was added and the obtained mixture was stirred at room temperature for 2 days. The reaction mixture diluted with 5 ml of water. The aqueous layer was acidified with a 2N solution of HCl and extracted with EtOAc (3 x 10 mL). The organic layers were combined, washed with brine (3 x 30 mL), dried over MgSO₄. The organic layer was evaporated at reduced pressure to afford the pure 1-(2,4-dichlorobenzyl)-1H-pyrazole-3,5-dicarboxylic acid **5** as a white solid.

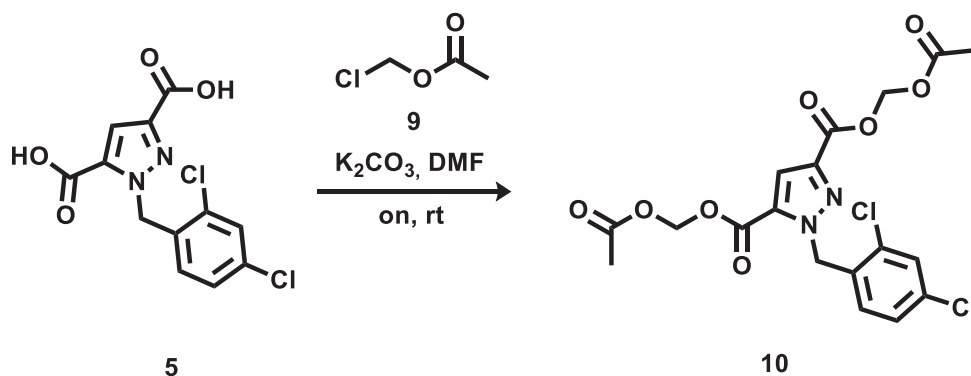
T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f: 0.20.

Yield: 113 mg (95%)

Purity: >95%

¹H NMR (DMSO-d₆), δ: 13.16 (s, 2H), 7.70 (d, ⁴J = 2.1 Hz, 1H), 7.39 (dd, ³J = 8.4 Hz, ⁴J = 2.1 Hz, 1H), 7.26 (s, 1H), 6.71 (d, ³J = 8.4 Hz, 1H), 5.87 (s, 2H).

¹³C NMR (DMSO-d₆), δ: 162.71 (C=O), 160.34 (C=O), 143.55, 135.63, 134.21, 133.50, 133.05 (C, C-aromatic), 130.05, 129.34, 128.26, 113.93 (CH, C-aromatic), 52.70 (CH₂).

bis(acetoxymethyl) 1-(2,4-dichlorobenzyl)-1H-pyrazole-3,5-dicarboxylate (8)**(C₁₈H₁₆Cl₂N₂O₈, M.W.: 459.2)**

A mixture of 1-(2,4-dichlorobenzyl)-1H-pyrazole-3,5-dicarboxylic acid **5** (0.127 mmol), chloromethyl acetate **9** (0.254 mmol) and triethylamine (0.305 mmol) in Acetone (0.5 mL) was stirred at reflux overnight. The reaction mixture was concentrated under reduced pressure and the residue was diluted with 5 ml of water. The aqueous layer was extracted

with EtOAc (3 x 10 mL). The organic layers were combined, washed with brine (3 x 30 mL), dried over MgSO₄. The organic layer was concentrated under reduced pressure and the obtained crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP Sil 10g, n-hexane -EtOAc 100:0 v/v increasing to 60:40 v/v in 15 CV) to give pure bis(acetoxymethyl) 1-(2,4-dichlorobenzyl)-1H-pyrazole-3,5-dicarboxylate (**10**) as a colourless oil.

T.L.C. System: n-hexane – EtOAc 8:2 v/v, Rf: 0.18.

Yield: 48 mg (82%)

Purity: 99%

UPLC-MS method C: Rt: 1.983, MS (ESI)⁺: 481,0[M+Na]⁺

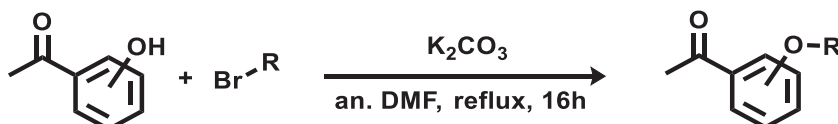
¹H NMR (DMSO-d₆), δ: 7.71 (d, ⁴J = 2.1 Hz, 1H), 7.46 (s, 1H), 7.39 (dd, ³J = 8.4 Hz, ⁴J = 2.1 Hz, 1H), 6.80 (d, ³J = 8.4 Hz, 1H), 5.90 (s, 2H), 5.89(s, 2H), 5.87 (s, 2H), 2.11 (s, 3H), 2.08 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 169.81, 169.68, 159.61, 157.24 (C=O), 141.67, 133.94, 133.77, 133.46, 133.20 (C, C-aromatic), 130.25, 129.45, 128.28, 115.25 (CH, C-aromatic), 80.24, 79.98, 53.29 (CH₂), 20.97, 20.91 (CH₃).

6.3 Synthesis of Quinoline structures

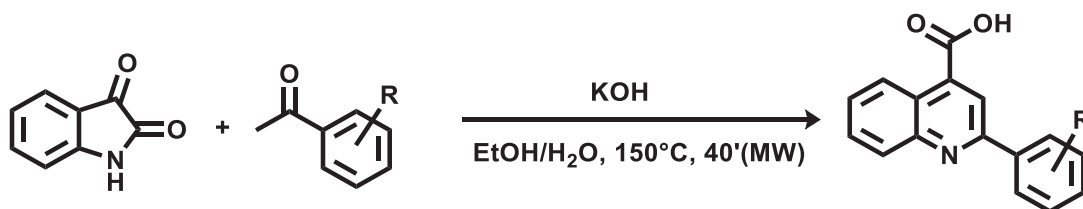
6.3.1 General procedures 1-3

General procedure 1: synthesis of (alkyloxy)acetophenones



A mixture of hydroxyacetophenone (1.47 mmol), bromoalkane (1.62 mmol) and K_2CO_3 (2.20 mmol) in anhydrous DMF (1.4 mL) under nitrogen atmosphere was stirred under reflux for 16 h. The reaction mixture was quenched with 10 ml of NaOH solution (15% p/v). The aqueous layer was extracted with EtOAc (3 x 15 mL). The organic layers were combined, washed with brine (3 x 30 mL) and dried over MgSO_4 . The organic solvent was evaporated at reduced pressure and the crude residue was purified by flash column chromatography to afford pure (alkyloxy)acetophenones.

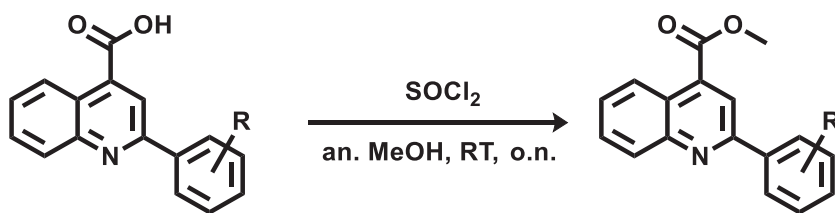
General procedure 2: synthesis of 2-(phenyl)quinoline-4-carboxylic acid analogues



Isatin (2.04 mmol) and KOH (6.12 mmol) were suspended in a solution of EtOH and water (7:1 ratio, 2mL). The reaction mixture was stirring at RT for 10 min and then acetophenone (2.448 mmol) was added. The reaction mixture was irradiated at the MW for 40 min at 150°C .

The reaction mixture was diluted with water (4 mL) and then concentrated under reduced pressure. The residue was diluted with water (10 ml) and washed remove the excess of acetophenone. The aqueous layer was then acidified to pH 5-6 with a solution of HCl (2N). The resulting precipitate was collected by filtration and then purified by re-crystallization from EtOH.

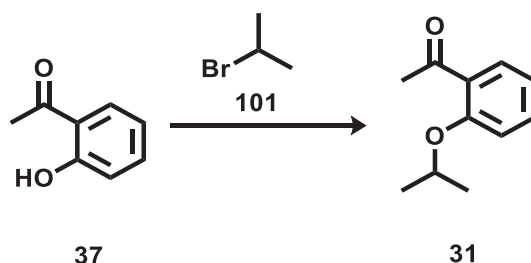
General procedure 3: synthesis of methyl 2-(phenyl)quinoline-4-carboxylate analogues



2-(phenyl)quinoline-4-carboxylic acid (0.17 mmol) was suspended in anhydrous MeOH (2 mL) and SOCl_2 (0.51 mmol) was added dropwise at RT. The reaction was stirred at RT overnight. The reaction was quenched with water (2 mL) and then concentrated under reduced pressure. The residue was then dissolved in a sat. NaHCO_3 solution (8 mL) and extracted with DCM (3 x 10 mL). The organic layers were combined, washed with brine (2 x 20 mL) and dried over MgSO_4 . The organic solvent was evaporated under reduced pressure and the crude residue was purified by flash column chromatography to afford pure methyl 2-(phenyl)quinoline-4-carboxylate.

6.3.2 (Alkyloxy)acetophenones (31-34)

2'-isopropoxyacetophenone (31) ($\text{C}_{11}\text{H}_{14}\text{O}_2$; M.W.: 178.2)



General procedure 1;

Pale yellow solid;

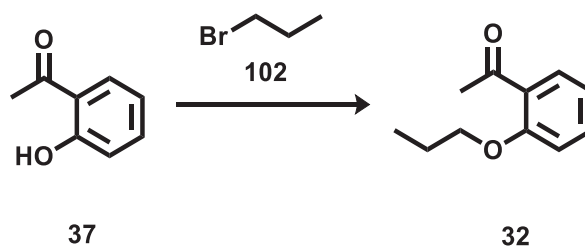
T.L.C. System: n-hexane -EtOAc 7:3 v/v, Rf: 0.67.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 30g (n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV).

Yield: 470 mg (60%)

$^1\text{H-NMR}$ (DMSO-d_6), δ : 7.55 (dd, $^3J = 7.7$ Hz, $^4J = 1.8$ Hz, 1H), 7.50 (td, $^3J = 8.4$, 7.7 Hz, $^4J = 1.8$ Hz, 1H), 7.17 (d, $J = 8.4$ Hz, 1H), 6.98 (td, $^3J = 7.7$ Hz, $^4J = 0.9$ Hz, 1H), 4.79 (hept, $J = 6.0$ Hz, 1H), 2.54 (s, 3H), 1.35 (s, 3H), 1.33 (s, 3H).

2'-propoxyacetophenone (32)
(C₁₁H₁₄O₂; M.W.: 178.2)



General procedure 1;

Pale yellow oil;

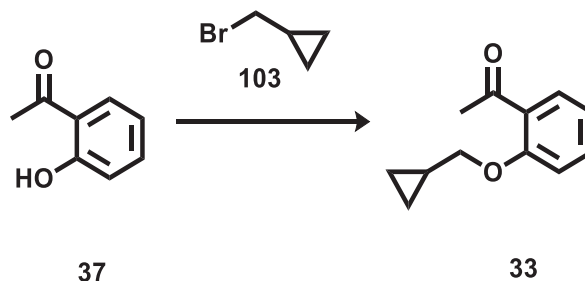
T.L.C. System: n-hexane -EtOAc 7:3 v/v, Rf: 0.67.

Purification: Purification: Flash column chromatography Isolera One system (Biotage),
Cartridge: ZIP KP Sil 10g (n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV).

Yield: 159 mg (60%)

¹H-NMR (DMSO-d₆), δ: 7.59 (dd, ³J = 7.7, ⁴J = 1.9 Hz, 1H), 7.52 (td, ³J = 8.4, ⁴J = 1.9 Hz, 1H), 7.15 (d, J = 8.4 Hz, 1H), 7.01 (td, ³J = 7.7, ⁴J = 0.9 Hz, 1H), 4.11 – 4.03 (m, 2H), 2.56 (s, 3H), 1.87 – 1.76 (m, 2H), 1.03 (t, J = 7.4 Hz, 3H).

2'-(cyclopropylmethoxy)acetophenone (33)
(C₁₂H₁₄O₂; M.W.: 190.2)



General procedure 1;

Pale yellow oil;

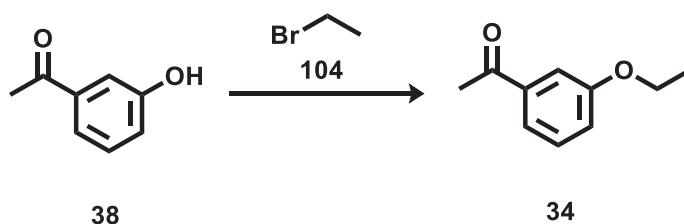
T.L.C. System: n-hexane -EtOAc 7:3 v/v, Rf: 0.67.

Purification: Purification: Flash column chromatography Isolera One system (Biotage),
Cartridge: ZIP KP Sil 10g (n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV).

Yield: 329 mg (78%)

¹H-NMR (DMSO-d₆), δ: 7.58 (d, J = 7.7 Hz, 1H), 7.51 (t, J = 7.7 Hz, 1H), 7.13 (d, J = 8.3 Hz, 1H), 7.01 (t, J = 7.4 Hz, 1H), 3.97 (d, J = 7.0 Hz, 2H), 2.60 (s, 3H), 1.36 – 1.23 (m, 1H), 0.62 – 0.59 (m, 2H), 0.38 (d, J = 4.7 Hz, 2H).

3'-ethoxyacetophenone (34)
(C₁₀H₁₂O₂; M.W.: 164.2)



General procedure 1;

Pale yellow oil;

T.L.C. System: n-hexane -EtOAc 7:3 v/v, R_f: 0.63.

Purification: Purification: Flash column chromatography Isolera One system (Biotage),

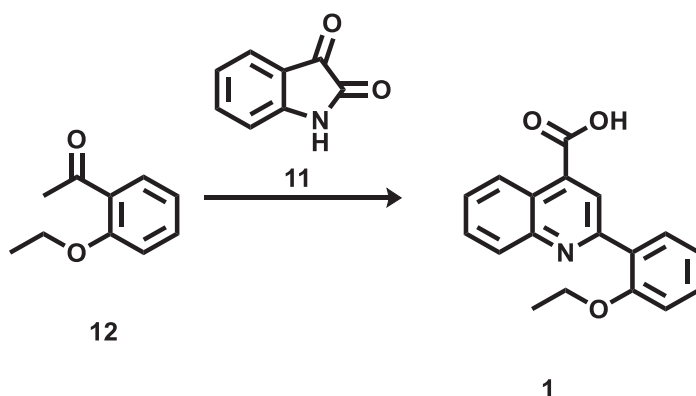
Cartridge: ZIP KP Sil 10g (n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV).

Yield: 134 mg (55%)

¹H NMR (DMSO-d₆), δ: 7.57 – 7.52 (m, 1H), 7.47 – 7.41 (m, 2H), 7.20 (ddd, ³J = 8.2 Hz, ⁴J = 2.7, 0.8 Hz, 1H), 4.09 (q, J = 7.0 Hz, 2H), 2.58 (s, 3H), 1.35 (t, J = 7.0 Hz, 3H).

6.3.3 2-(phenyl)quinoline-4-carboxylic acid analogues (1, 13-19, 41-52, 77-86)

2-(2-ethoxyphenyl)quinoline-4-carboxylic acid (1)
(C₁₈H₁₅NO₃; M.W.: 293.3)



General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f: 0.48.

Purification: re-crystallisation from EtOH

Yield: 300 mg (50%)

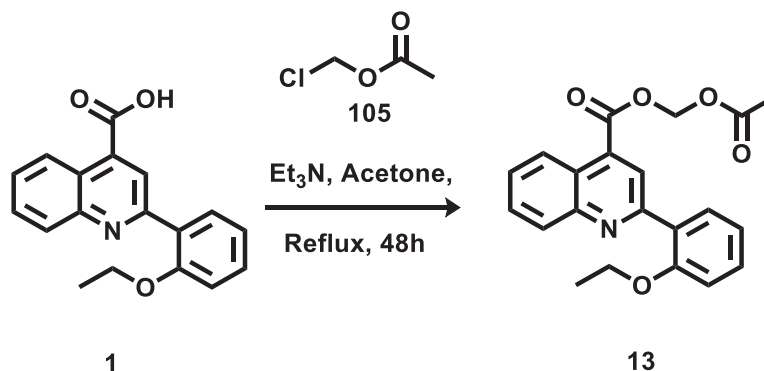
Purity: 99%

UPLC-MS method B: R_t: 1.532, MS (ESI)⁺: 294.1[M+H]⁺

¹H NMR (DMSO-d₆), δ: 13.84 (bs, 1H), 8.75 (dd, ³J = 8.4 Hz, ⁴J = 0.6 Hz, 1H), 8.54 (s, 1H), 8.15 (dd, ³J = 8.4 Hz, ⁴J = 0.6 Hz, 1H), 7.92 (dd, ³J = 7.6 Hz, ⁴J = 1.8 Hz, 1H), 7.84 (td, ³J = 7.7 Hz, ⁴J = 1.4 Hz, 1H), 7.72 (td, ³J = 7.7 Hz, ⁴J = 1.4 Hz, 1H), 7.49 (td, ³J = 7.8 Hz, ⁴J = 1.8 Hz, 1H), 7.21 (d, J = 7.9 Hz, 1H), 7.14 (td, ³J = 7.5 Hz, ⁴J = 0.9 Hz, 1H), 4.18 (q, J = 6.9 Hz, 2H), 1.35 (t, J = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 168.18 (C=O), 156.96, 156.40, 149.00, 135.55 (C, C-aromatic), 131.61, 131.39, 130.21, 130.19, 128.17 (CH, C-aromatic), 128.15 (C, C-aromatic), 125.82, 124.70 (CH, C-aromatic), 123.75 (C, C-aromatic), 121.35, 113.61 (CH, C-aromatic), 64.39 (CH₂), 15.00 (CH₃).

Acetoxymethyl 2-(2-ethoxyphenyl)quinoline-4-carboxylate (13)
(C₂₁H₁₉NO₅; M.W.: 365.4)



A mixture of 2-(2-ethoxyphenyl)quinoline-4-carboxylic acid **1** (0.2 mmol), chloromethyl acetate **105** (0.2 mmol) and triethylamine (0.48 mmol) in acetone (4 mL) was stirred at reflux for 48 hours. The reaction was concentrated under reduced pressure, diluted with water (8 mL). The aqueous solution was extracted with DCM (3 x 10 mL). The combined organic layers were washed with brine (2 x 20 mL), dried over MgSO₄ and evaporated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 5g, n-hexane -EtOAc 100:0 v/v increasing to 70:30 v/v in 12 CV) to give pure acetoxymethyl 2-(2-ethoxyphenyl)quinoline-4-carboxylate (**13**) as colourless oil.

T.L.C. System: n-hexane -EtOAc 8:2 v/v, R_f: 0.41.

Yield: 22 mg (24%)

Purity: 99%

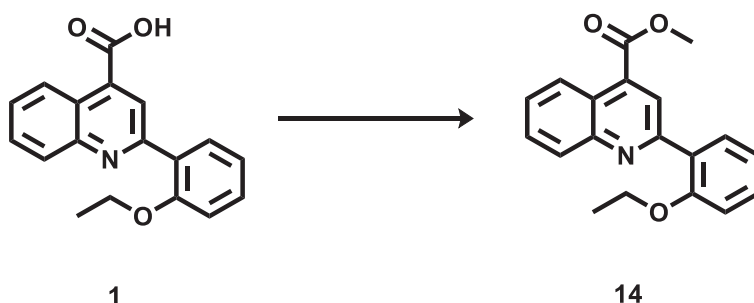
UPLC-MS method C: Rt: 2.143, MS (ESI)⁺: 366.3 [M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.68 (dd, ³J = 8.5 Hz, ⁴J = 0.7 Hz, 1H), 8.61 (s, 1H), 8.18 (dd, ³J = 8.5 Hz, ⁴J = 0.7 Hz, 1H), 7.96 (dd, ³J = 7.6 Hz, ⁴J = 1.7 Hz, 1H), 7.88 (td, ³J = 8.3 Hz, ⁴J = 1.3 Hz, 1H), 7.77 (td, ³J = 8.3 Hz, ⁴J = 1.3 Hz, 1H), 7.50 (td, ³J = 8.4 Hz, ⁴J = 1.8 Hz, 1H),

7.22 (d, $J = 8.1$ Hz, 1H), 7.15 (td, $^3J = 7.4$ Hz, $^4J = 0.8$ Hz, 1H), 6.06 (s, 2H), 4.18 (q, $J = 6.9$ Hz, 2H), 2.15 (s, 3H), 1.38 (t, $J = 6.9$ Hz, 3H).

^{13}C NMR (DMSO- d_6), δ : 169.94, 164.88 (C=O), 156.99, 156.21, 148.96, 132.87 (C, C-aromatic), 131.89, 131.36, 130.60, 130.38, 128.78 (CH, C-aromatic), 127.53 87 (C, C-aromatic), 125.24, 125.17 (CH, C-aromatic), 123.29 (C, C-aromatic), 121.38, 113.49 (CH, C-aromatic), 80.55, 64.40 (CH_2), 20.97, 14.85 (CH_3).

Methyl 2-(2-ethoxyphenyl)quinoline-4-carboxylate (14)
($\text{C}_{19}\text{H}_{17}\text{NO}_3$; M.W.: 307.3)



General procedure 3;

Yellow solid;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.53.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV).

Yield: 30 mg (29%)

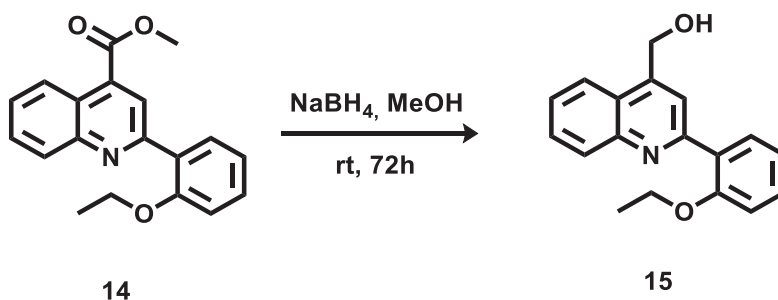
Purity: 99%

UPLC-MS method C: Rt: 2.138, MS (ESI)⁺: 308.1[M+H]⁺

^1H NMR (DMSO- d_6), δ : 8.66 (dd, $^3J = 8.5$ Hz, $^4J = 0.8$ Hz, 1H), 8.57 (s, 1H), 8.17 (dd, $^3J = 8.4$ Hz, $^4J = 0.6$ Hz, 1H), 7.94 (dd, $^3J = 7.6$ Hz, $^4J = 1.8$ Hz, 1H), 7.86 (td, $^3J = 8.4$ Hz, $^4J = 1.4$ Hz, 1H), 7.74 (td, $^3J = 8.3$ Hz, $^4J = 1.3$ Hz, 1H), 7.50 (td, $^3J = 8.4$ Hz, $^3J = 1.8$ Hz, 1H), 7.22 (d, $J = 8.3$ Hz, 1H), 7.14 (td, $^3J = 7.5$ Hz, $^4J = 0.9$ Hz, 1H), 4.19 (q, $J = 6.9$ Hz, 2H), 4.01 (s, 3H), 1.37 (t, $J = 6.9$ Hz, 3H).

^{13}C NMR (DMSO- d_6), δ : 166.81 (C=O), 156.99, 156.29, 148.94, 134.28 (C, C-aromatic), 131.75, 131.34, 130.44, 130.28, 128.47 (CH, C-aromatic), 127.84 (C, C-aromatic), 125.52, 124.81(CH, C-aromatic), 123.38 (C, C-aromatic), 121.37, 113.58 (CH, C-aromatic), 64.42 (CH_2), 53.37, 14.97 (CH_3).

(2-(2-ethoxyphenyl)quinolin-4-yl)methanol (15)
 (C₁₈H₁₇NO₂; M.W.: 279.3)



Methyl 2-(2-ethoxyphenyl)quinoline-4-carboxylate **14** (0.21 mmol) was dissolved in anhydrous MeOH (3 mL). The reaction mixture was cooled to 0°C in an ice-bath. NaBH₄ (0.86 mmol) was added and the reaction was stirred at room temperature for 72 hours. The reaction mixture was concentrated under reduced pressure and the residue was then diluted with water (6 mL). The aqueous layer was extracted with EtOAc (3 x 10 mL). The organic layers were combined, washed with brine (2 x 20 mL) dried over MgSO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 5g, n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV) to give pure acetoxymethyl (2-(2-ethoxyphenyl)quinolin-4-yl)methanol **15** as yellow solid.

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.15.

Yield: 15 mg (25%)

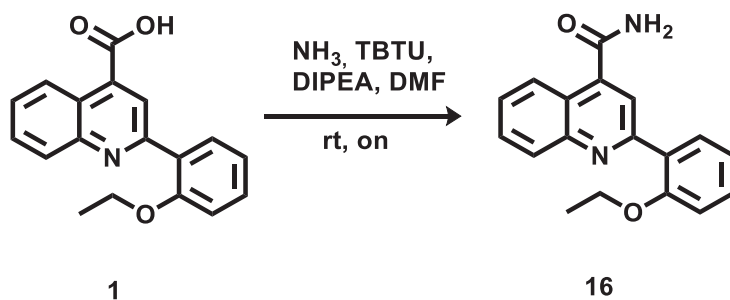
Purity: 97%

UPLC-MS method C: Rt: 1.299, MS (ESI)⁺: 280.0 [M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.13 (s, 1H), 8.10 – 8.01 (m, 2H), 7.82 (d, *J* = 7.4 Hz, 1H), 7.76 (t, *J* = 7.6 Hz, 1H), 7.60 (t, *J* = 7.6 Hz, 1H), 7.45 (t, *J* = 7.6 Hz, 1H), 7.19 (d, *J* = 8.2 Hz, 1H), 7.11 (t, *J* = 7.4 Hz, 1H), 5.57 (t, *J* = 5.4 Hz, 1H), 5.07 (d, *J* = 5.4 Hz, 2H), 4.16 (q, *J* = 6.9 Hz, 2H), 1.35 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 156.83, 156.78, 147.90, 147.07 (C, C-aromatic), 131.53, 130.97, 130.04, 129.51, 126.59 (CH, C-aromatic), 124.82 (C, C-aromatic), 123.72, 121.16, 120.42, 113.55 (CH, C-aromatic), 64.32, 60.31 (CH₂), 15.07 (CH₃).

2-(2-ethoxyphenyl)quinoline-4-carboxamide (16)
 (C₁₈H₁₆N₂O₂; M.W.: 292.3)



A mixture of 2-(2-ethoxyphenyl)quinoline-4-carboxylic acid **1** (0.17 mmol), NH₃ 0.5M in Dioxane (0.34 mmol), TBTU (0.2 mmol), and DIPEA (0.39 mmol) in anhydrous DMF (2 ml) was stirred at room temperature for 3 hours. The reaction mixture was diluted in DCM (10 mL) and washed with a sat. NaHCO₃ solution (8 mL). The aqueous layer was then extracted with DCM (2 x 10 mL). The organic layers were then combined, washed with brine (2 x 20 mL), dried over MgSO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 5g, n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV) to give pure 2-(2-ethoxyphenyl)quinoline-4-carboxamide **16** as white solid.

T.L.C. System: n-hexane -EtOAc 6:4 v/v, R_f: 0.14.

Yield: 96 mg (96%)

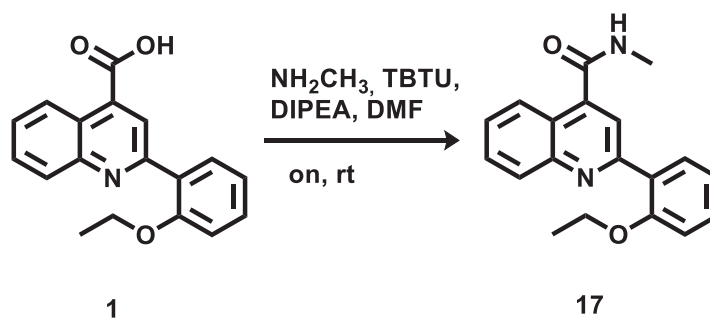
Purity: 95%

UPLC-MS method C: Rt: 1.521, MS (ESI)⁺: 293.1[M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.69-8.61 (m, 1H), 8.19 (s, 1H), 8.11 (d, *J* = 8.4 Hz, 1H), 8.08 (s, 1H), 7.88 (d, *J* = 7.7 Hz, 1H), 7.85 (s, 1H), 7.81 (t, *J* = 7.6 Hz, 1H), 7.66 (t, *J* = 7.6 Hz, 1H), 7.48 (t, *J* = 7.7 Hz, 1H), 7.22 (d, *J* = 8.4 Hz, 1H), 7.14 (t, *J* = 7.6 Hz, 1H), 4.18 (q, *J* = 6.9 Hz, 2H), 1.34 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 169.41 (C=O), 156.89, 156.40, 148.50, 141.98 (C, C-aromatic), 131.52, 131.42, 130.18, 129.89 (CH, C-aromatic), 128.57 (C, C-aromatic), 127.44, 125.79 (CH, C-aromatic), 123.40 (C, C-aromatic), 121.28, 121.25, 113.61 (CH, C-aromatic), 64.37 (CH₂), 15.04 (CH₃).

2-(2-ethoxyphenyl)-N-methylquinoline-4-carboxamide (17)
 (C₁₉H₁₈N₂O₂; M.W.: 306.4)



A mixture of 2-(2-ethoxyphenyl)quinoline-4-carboxylic acid **1** (0.17 mmol), methylamine 2M in THF (0.34 mmol), TBTU (0.2 mmol), and DIPEA (0.39 mmol) in anhydrous DMF (1 ml) was stirred at room temperature overnight. The reaction mixture was diluted in EtOAc (10 mL) and was washed with sat. NH₄Cl solution (3 x 8 mL), sat. NaHCO₃ solution (3 x 8 mL) and brine (3 x 8 mL). The organic layers were then dried over MgSO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 5g, n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 15 CV) to give pure 2-(2-ethoxyphenyl)-N-methylquinoline-4-carboxamide **17** as white solid.

T.L.C. System: n-hexane -EtOAc 6:4 v/v, R_f: 0.18.

Yield: 12 mg (29%)

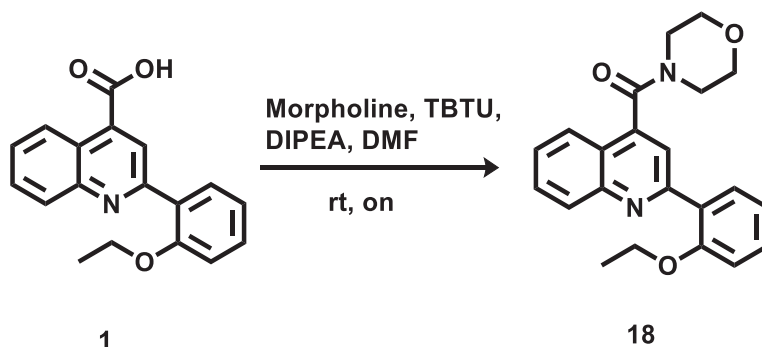
Purity: 99%

UPLC-MS method C: Rt: 1.573, MS (ESI)⁺: 307.2[M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.65 (d, ⁴J = 4.0 Hz, 1H), 8.15 (d, J = 8.3 Hz, 1H), 8.10 (d, J = 8.4 Hz, 1H), 8.02 (s, 1H), 7.86 (d, J = 7.5 Hz, 1H), 7.80 (t, J = 7.6 Hz, 1H), 7.64 (t, J = 7.6 Hz, 1H), 7.47 (t, J = 7.7 Hz, 1H), 7.20 (d, J = 8.3 Hz, 1H), 7.12 (t, J = 7.5 Hz, 1H), 4.17 (q, J = 6.9 Hz, 3H), 2.88 (d, J = 4.5 Hz, 2H), 1.32 (t, J = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 167.80 (C=O), 156.87, 156.47, 148.46, 142.09 (C, C-aromatic), 131.50, 131.40, 130.24, 129.90 (CH, C-aromatic), 128.65 (C, C-aromatic), 127.46, 125.73 (CH, C-aromatic), 123.52 (C, C-aromatic), 121.31, 121.29, 113.67 (CH, C-aromatic), 64.42 (CH₂), 26.59, 15.01 (CH₃).

(2-(2-ethoxyphenyl)quinolin-4-yl)(morpholino)methanone (18)
 (C₂₂H₂₂N₂O₃; M.W.: 362.4)



A mixture of 2-(2-ethoxyphenyl)quinoline-4-carboxylic acid **1** (0.17 mmol), morpholine (0.34 mmol), TBTU (0.2 mmol), and DIPEA (0.39 mmol) in anhydrous DMF (1 ml) was stirred at room temperature overnight. The reaction mixture was diluted in EtOAc (10 mL) and was washed with sat. NH₄Cl solution (3 x 8 mL), sat. NaHCO₃ solution (3 x 8 mL) and brine (3 x 8 mL). The organic layers were then dried over MgSO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 5g, n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 15 CV) to give pure (2-(2-ethoxyphenyl)quinolin-4-yl)(morpholino)methanone (**18**) as white solid.

T.L.C. System: n-hexane -EtOAc 6:4 v/v, Rf: 0.14.

Yield: 44 mg (72%)

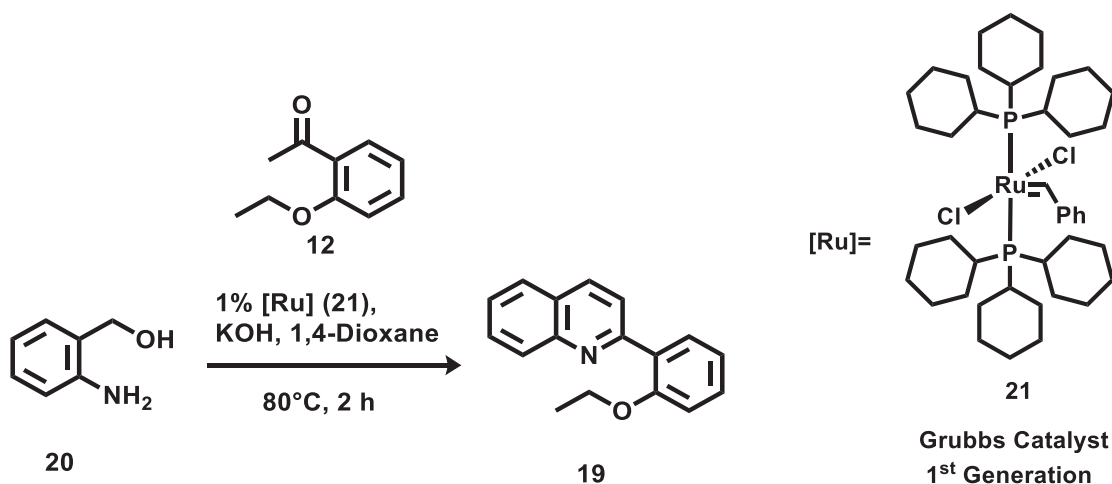
Purity: 99%

UPLC-MS method C: Rt: 1.693, MS (ESI)⁺: 363.3[M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.13 (d, *J* = 8.3 Hz, 1H), 7.90 (s, 1H), 7.88 – 7.80 (m, 3H), 7.68 (t, *J* = 7.3 Hz, 1H), 7.48 (t, *J* = 7.1 Hz, 1H), 7.20 (d, *J* = 8.3 Hz, 1H), 7.13 (t, *J* = 7.3 Hz, 1H), 4.24 – 4.04 (m, 2H), 3.95 – 3.82 (m, 1H), 3.82 – 3.70 (m, 3H), 3.59 – 3.42 (m, 2H), 3.28 – 3.10 (m, 2H), 1.32 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 166.48 (C=O), 156.83, 156.60, 148.26, 141.50 (C, C-aromatic), 131.54, 131.43, 130.59, 130.09 (CH, C-aromatic), 128.69 (C, C-aromatic), 127.90, 125.08 (CH, C-aromatic), 123.00 (C, C-aromatic), 121.24, 120.49, 113.47 (CH, C-aromatic), 66.78, 66.55, 64.38, 47.52, 42.10 (CH₂), 15.08 (CH₃).

2-(2-ethoxyphenyl)quinoline (19)
(C₁₇H₁₅NO; M.W.: 249.3)



A mixture of (2-aminophenyl)methanol **20** (1.46 mmol), 2-ethoxyacetophenone **12** (2.92 mmol), KOH (0.4 mL of a solution 4M in Methanol) and 1st generation Grubbs catalyst **21** (0.0146 mmol, Bis(tricyclohexylphosphine)benzylidene ruthenium(IV) dichloride) in anhydrous dioxane (4.3 mL) was placed in a 7 mL screw-capped vial. The reaction mixture was stirred at 80°C for 2 hours. The reaction mixture was filtered through a short silica gel column (EtOAc) to remove the catalyst and the inorganic salts. The solution obtained from the filtration was concentrated under reduced pressure and the crude was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 30g, n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 15 CV). The solvent was evaporated under reduced pressure. The residue was dissolved in EtOAc (10 mL), washed with water (2 x 8 mL), dried over MgSO₄ and concentrated under reduced pressure. The residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 10g, n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV) to give 2-(2-ethoxyphenyl)quinoline **19** as yellow oil.

T.L.C. System: n-hexane -EtOAc 8:2 v/v, R_f: 0.57.

Yield: 118 mg (32%)

Purity: 99%

UPLC-MS method B: Rt: 1.779, MS (ESI)⁺: 250.2 [M+H]⁺

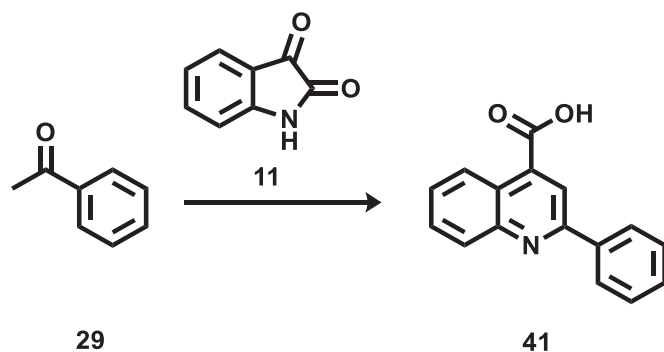
¹H NMR (DMSO-d₆), δ: 8.33 (d, *J* = 8.6 Hz, 1H), 8.05 (d, *J* = 8.4 Hz, 1H), 8.02 – 7.95 (m, 2H), 7.82 (d, *J* = 6.8 Hz, 1H), 7.76 (t, *J* = 7.4 Hz, 1H), 7.60 (t, *J* = 7.3 Hz, 1H), 7.44 (t, *J* = 7.3 Hz, 1H), 7.17 (d, *J* = 8.4 Hz, 1H), 7.10 (t, *J* = 7.4 Hz, 1H), 4.13 (q, *J* = 6.9 Hz, 2H), 1.31 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 156.90, 156.79, 148.18 (C, C-aromatic), 135.58, 131.63, 131.03, 129.91, 129.47 (CH, C-aromatic), 129.30 (C, C-aromatic), 128.12 (CH, C-aromatic), 127.11

(C, C-aromatic), 126.79, 123.72, 121.16, 113.44 (CH, C-aromatic), 64.27 (CH₂), 15.04 (CH₃).

2-phenylquinoline-4-carboxylic acid (41)

(C₁₆H₁₁NO₂; M.W.: 249.3)



General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f: 0.39.

Purification: re-crystallisation from EtOH

Yield: 335 mg (66%)

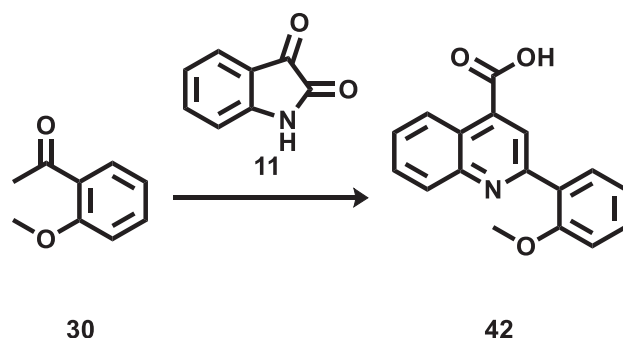
Purity: 99%

UPLC-MS method B: Rt: 1.683, MS (ESI)⁺: 250.1[M+H]⁺

¹H NMR (DMSO-d₆), δ: 14.01 (bs, 1H), 8.67 (d, *J* = 8.5 Hz, 1H), 8.47 (s, 1H), 8.31 (d, *J* = 7.4 Hz, 2H), 8.18 (d, *J* = 8.4 Hz, 1H), 7.87 (t, *J* = 7.6 Hz, 1H), 7.72 (t, *J* = 7.6 Hz, 1H), 7.63-7.51 (m, 3H).

¹³C NMR (DMSO-d₆), δ: 168.14 (C=O), 156.28, 148.87, 138.38, 138.23 (C, C-aromatic), 130.70, 130.47, 130.26, 129.48, 128.25, 127.70, 125.90 (CH, C-aromatic), 123.94 (C, C-aromatic), 119.59 (CH, C-aromatic).

2-(2-methoxyphenyl)quinoline-4-carboxylic acid (42)
(C₁₇H₁₃NO₃; M.W.: 279.3)



General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f: 0.39.

Purification: re-crystallisation from EtOH

Yield: 332 mg (58%)

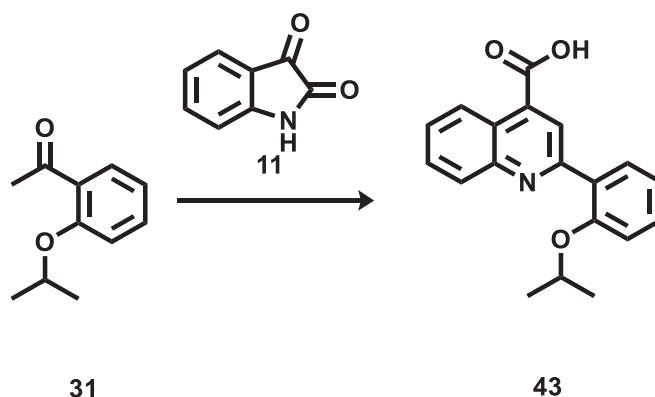
Purity: 99%

UPLC-MS method B: R_t: 1.382, MS (ESI)⁺: 280.0 [M+H]⁺

¹H NMR (DMSO-d₆), δ: 13.91 (bs, 1H), 8.72 (d, *J* = 8.4 Hz, 1H), 8.37 (s, 1H), 8.15 (d, *J* = 8.4 Hz, 1H), 7.87 – 7.82 (m, 2H), 7.72 (td, ³*J* = 8.2 Hz, ⁴*J* = 1.2 Hz, 1H), 7.51 (td, ³*J* = 7.7 Hz, ⁴*J* = 1.8 Hz, 1H), 7.24 (d, *J* = 8.3 Hz, 1H), 7.15 (t, *J* = 7.4 Hz, 1H), 3.89 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 168.20 (C=O), 157.52, 156.50, 148.99, 135.97 (C, C-aromatic), 131.57, 131.47, 130.29, 130.18 (CH, C-aromatic), 128.29 (C, C-aromatic), 128.20, 125.84, 124.34 (CH, C-aromatic), 123.69 (C, C-aromatic), 121.34, 112.55 (CH, C-aromatic), 56.26 (CH₃).

2-(2-isopropoxyphenyl)quinoline-4-carboxylic acid (43)
(C₁₉H₁₇NO₃; M.W.: 307.3)



General procedure 2;

White solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f: 0.5.

Purification: re-crystallisation from EtOH

Yield: 247 mg (39%)

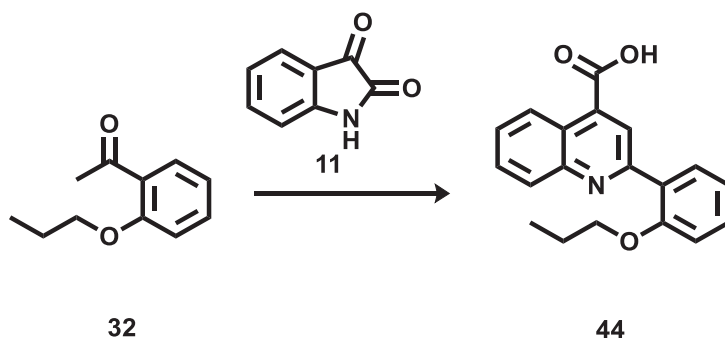
Purity: 99%

UPLC-MS method B: Rt: 1.667, MS (ESI)⁺: 308.1[M+H]⁺

¹H NMR (DMSO-d₆), δ: 13.78 (bs, 1H), 8.75 (d, *J* = 8.5 Hz, 1H), 8.54 (s, 1H), 8.14 (d, *J* = 8.5 Hz, 1H), 7.91 (dd, ³*J* = 7.6 Hz, ⁴*J* = 1.8 Hz, 1H), 7.84 (td, ³*J* = 7.6 Hz, ⁴*J* = 1.3 Hz, 1H), 7.71 (td, ³*J* = 7.6 Hz, ⁴*J* = 1.3 Hz, 1H), 7.48 (td, ³*J* = 7.6 Hz, ⁴*J* = 1.8 Hz, 1H), 7.24 (d, *J* = 8.2 Hz, 1H), 7.13 (t, *J* = 7.5 Hz, 1H), 4.83-4.66 (m, 1H), 1.30 (s, 3H), 1.29 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 168.18 (C=O), 156.55, 155.79, 149.01 (C, C-aromatic), 131.63, 131.50, 130.16, 128.10, 125.88, 125.02 (CH, C-aromatic), 123.97 (C, C-aromatic), 121.33, 115.18, (CH, C-aromatic) 70.79, 22.26 (CH₃).

2-(2-propoxyphenyl)quinoline-4-carboxylic acid (44)
(C₁₉H₁₇NO₃; M.W.: 307.3)



General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f: 0.57.

Purification: re-crystallisation from EtOH

Yield: 52 mg (30%)

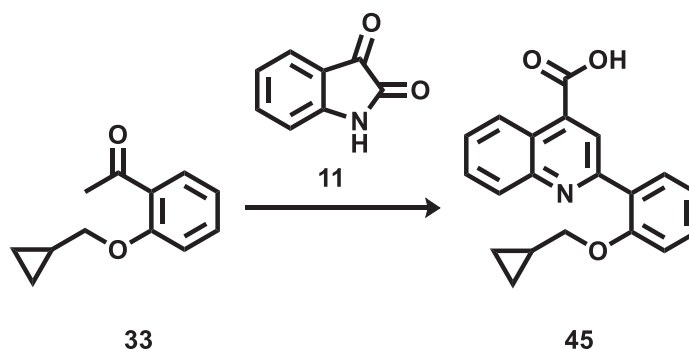
Purity: 99%

UPLC-MS method B: Rt: 1.727, MS (ESI)⁺: 308.1[M+H]⁺

¹H NMR (DMSO-d₆), δ: 13.78 (s, 1H), 8.74 (d, *J* = 8.4 Hz, 1H), 8.52 (s, 1H), 8.15 (d, *J* = 8.4 Hz, 1H), 7.91 (dd, ³*J* = 7.6 Hz, ⁴*J* = 1.7 Hz, 1H), 7.84 (td, ³*J* = 8.3 Hz, ⁴*J* = 1.3 Hz, 1H), 7.72 (td, ³*J* = 8.3 Hz, ⁴*J* = 1.2 Hz, 1H), 7.52 – 7.46 (m, 3H), 7.21 (d, *J* = 8.3 Hz, 1H), 7.14 (t, *J* = 7.5 Hz, 1H), 4.08 (t, *J* = 6.2 Hz, 2H), 1.80 – 1.67 (m, 2H), 0.99 (t, *J* = 7.4 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 168.15 (C=O), 157.08, 156.43, 149.02, 135.67 (C, C-aromatic), 131.60, 131.43, 130.21, 130.19, 128.19 (CH, C-aromatic), 128.14 (C, C-aromatic), 125.84, 124.65 (CH, C-aromatic), 123.75 (C, C-aromatic), 121.26, 113.37 (CH, C-aromatic), 70.15, 22.57 (CH₂), 11.06 (CH₃).

2-(2-(cyclopropylmethoxy)phenyl)quinoline-4-carboxylic acid (45)
 (C₂₀H₁₇NO₃; M.W.: 319.4)



General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f:0.55.

Purification: re-crystallisation from EtOH

Yield: 55 mg (12%)

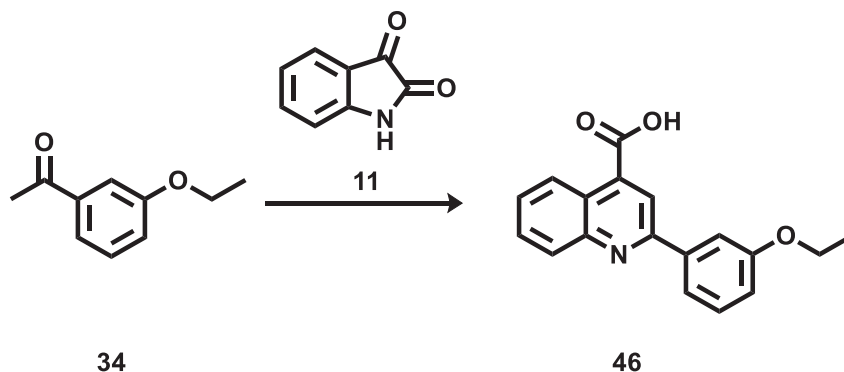
Purity: 98%

UPLC-MS method B: Rt: 1.929, MS (ESI)⁺: 320.19[M+H]⁺

¹H NMR (DMSO-d₆), δ: 13.82 (bs, 1H), 8.74 (d, *J* = 8.4 Hz, 1H), 8.61 (s, 1H), 8.16 (d, *J* = 8.4 Hz, 1H), 7.92 (d, *J* = 7.5 Hz, 1H), 7.84 (t, *J* = 7.5 Hz, 1H), 7.72 (t, *J* = 7.6 Hz, 1H), 7.47 (t, *J* = 7.6 Hz, 1H), 7.20 (d, *J* = 8.3 Hz, 1H), 7.14 (t, *J* = 7.4 Hz, 1H), 4.01 (d, *J* = 6.8 Hz, 2H), 1.31 – 1.12 (m, 1H), 0.58-0.53 (m, 2H), 0.41-0.32 (m, 2H).

¹³C NMR (DMSO-d₆), δ: 168.20 (C=O), 157.01, 156.47, 149.00, 135.71 (C, C-aromatic), 131.55, 131.43, 130.18, 128.33 (CH, C-aromatic), 128.12 (C, C-aromatic), 125.82, 124.77 (CH, C-aromatic), 123.78 (C, C-aromatic), 121.35, 113.85 (CH, C-aromatic), 73.02 (CH₃), 10.50(CH₂), 3.40 (CH₃).

2-(3-ethoxyphenyl)quinoline-4-carboxylic acid (46)
(C₁₈H₁₅NO₃; M.W.: 293.3)



General procedure 2;

White solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, Rf: 0.53.

Purification: re-crystallisation from EtOH

Yield: 58 mg (35%)

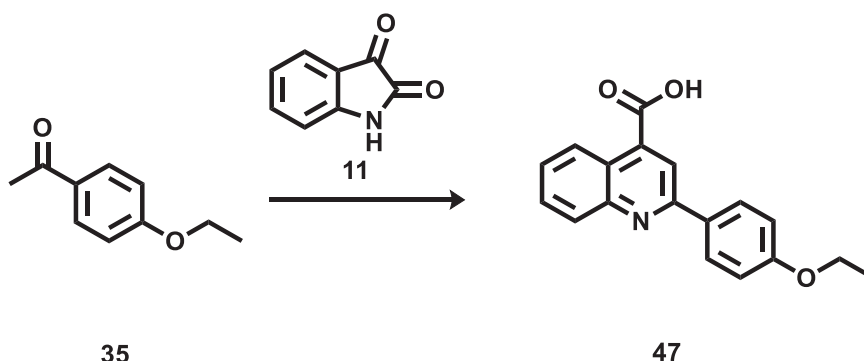
Purity: 99%

UPLC-MS method B: Rt: 1.845, MS (ESI)⁺: 294.1[M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.65 (dd, ³J = 8.4 Hz, ⁴J = 0.7 Hz, 1H), 8.44 (s, 1H), 8.18 (dd, ³J = 8.4 Hz, ⁴J = 0.7 Hz, 1H), 7.91 – 7.80 (m, 3H), 7.71 (td, ³J = 8.3 Hz, ⁴J = 1.2 Hz, 1H), 7.49 (t, J = 7.9 Hz, 1H), 7.14 – 7.07 (m, 1H), 4.17 (q, J = 7.0 Hz, 2H), 1.40 (t, J = 7.0 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 168.14 (C=O), 159.61, 156.09, 148.78, 139.89, 138.42 (C, C-aromatic), 130.68, 130.59, 130.28, 128.25, 125.88 (CH, C-aromatic), 123.99 (C, C-aromatic), 119.99, 119.62, 116.68, 113.35 (CH, C-aromatic), 63.73 (CH₂), 15.17 (CH₃).

2-(4-ethoxyphenyl)quinoline-4-carboxylic acid (47)
(C₁₈H₁₅NO₃; M.W.: 293.3)



General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, Rf: 0.43.

Purification: re-crystallisation from EtOH

Yield: 175 mg (29%)

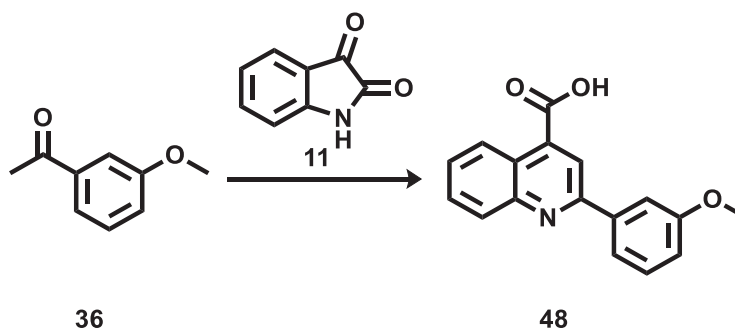
Purity: 99%

UPLC-MS method B: Rt: 1.715, MS (ESI)⁺: 294.0 [M+H]⁺

¹H NMR (DMSO-d₆), δ: 13.95 (bs, 1H), 8.63 (d, *J* = 8.5 Hz, 1H), 8.42 (s, 1H), 8.27 (d, *J* = 8.6 Hz, 2H), 8.13 (d, *J* = 8.4 Hz, 1H), 7.83 (t, *J* = 7.6 Hz, 1H), 7.67 (t, *J* = 7.6 Hz, 1H), 7.11 (d, *J* = 8.6 Hz, 2H), 4.14 (q, *J* = 6.9 Hz, 2H), 1.39 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 168.47 (C=O), 160.64, 155.91, 148.85, 130.73 (C, C-aromatic), 130.47, 129.95, 129.14, 127.55, 126.04 (CH, C-aromatic), 123.68 (C, C-aromatic), 118.86, 115.22 (CH, C-aromatic), 63.73 (CH₂), 15.10 (CH₃).

2-(3-methoxyphenyl)quinoline-4-carboxylic acid (48)
(C₁₇H₁₃NO₃; M.W.: 279.3)



General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f:0.44.

Purification: re-crystallisation from EtOH

Yield: 203 mg (35%)

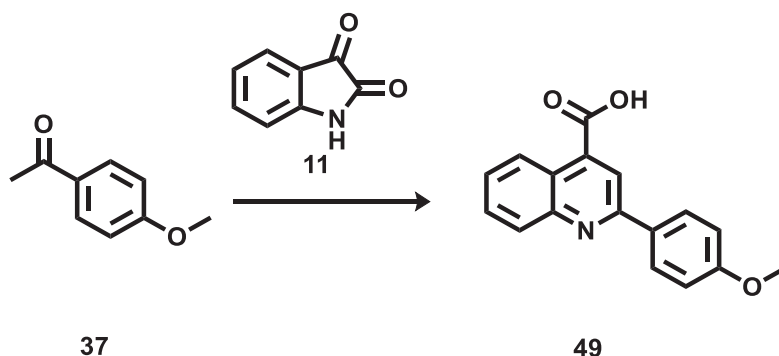
Purity: 99%

UPLC-MS method B: Rt: 1.939, MS (ESI)⁺: 280.04[M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.65 (d, *J* = 8.4 Hz, 1H), 8.45 (s, 1H), 8.18 (d, *J* = 8.4 Hz, 1H), 7.91 – 7.81 (m, 3H), 7.72 (t, *J* = 7.5 Hz, 1H), 7.50 (t, *J* = 8.0 Hz, 1H), 7.12 (d, *J* = 8.0 Hz, 1H), 3.90 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 168.13 (C=O), 160.33, 156.08, 148.77, 139.90, 138.33 (C, C-aromatic), 130.71, 130.60, 130.29, 128.29, 125.86 (CH, C-aromatic), 123.98 (C, C-aromatic), 120.11, 119.66, 116.32, 112.76 (CH, C-aromatic), 55.79 (CH₃).

2-(4-methoxyphenyl)quinoline-4-carboxylic acid (49)
($C_{17}H_{13}NO_3$; M.W.: 279.3)



General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, Rf:0.43.

Purification: re-crystallisation from EtOH

Yield: 564 mg (99%)

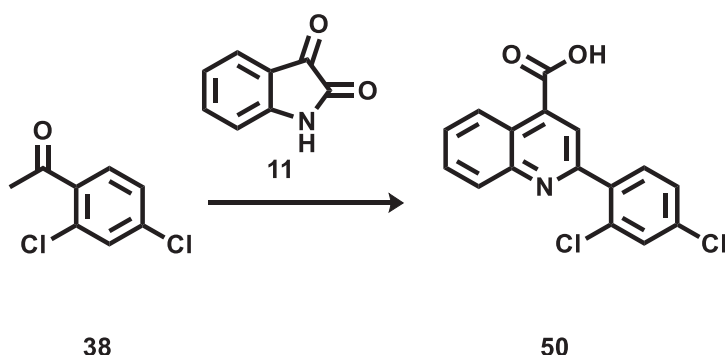
Purity: 99%

UPLC-MS method B: Rt: 1.685, MS (ESI)⁺: 280.0[M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.62 (d, *J* = 8.4 Hz, 1H), 8.41 (s, 1H), 8.28 (d, *J* = 8.4 Hz, 2H), 8.13 (d, *J* = 8.3 Hz, 1H), 7.83 (t, *J* = 7.6 Hz, 1H), 7.67 (t, *J* = 7.6 Hz, 1H), 7.13 (d, *J* = 8.4 Hz, 2H), 3.86 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 168.20 (C=O), 161.40, 155.95, 148.87, 138.11, 130.82 (C, C-aromatic), 130.59, 130.02, 129.18, 127.76, 125.84 (CH, C-aromatic), 123.57 (C, C-aromatic), 119.09, 114.86 (CH, C-aromatic), 55.82 (CH₃).

2-(2,4-dichlorophenyl)quinoline-4-carboxylic acid (50)
($C_{16}H_9Cl_2NO_2$; M.W.: 318.1)



General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, Rf: 0.75.

Purification: re-crystallisation from EtOH

Yield: 194 mg (60%)

Purity: 99%

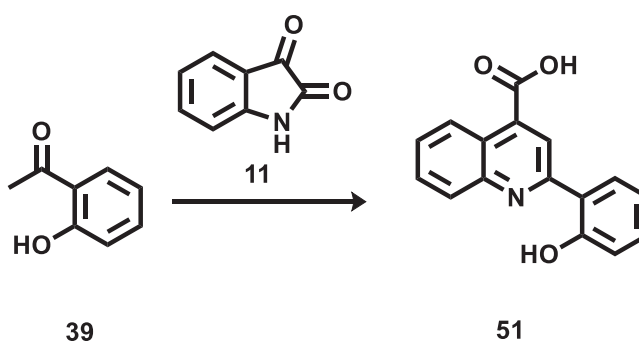
UPLC-MS method B: Rt: 1.96, MS (ESI)⁺: 318.1 (56%), 320.1(37%), 322.1 (7%) [M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.75 (d, *J* = 8.4 Hz, 1H), 8.17 (d, *J* = 8.4 Hz, 2H), 7.90 (dd, ³*J* = 6.9 Hz, ⁴*J* = 1.2 Hz, 1H), 7.84 (d, ⁴*J* = 2.0 Hz, 1H), 7.79 (td, ³*J* = 8.4 Hz, ⁴*J* = 1.2 Hz, 2H), 7.63 (dd, ³*J* = 8.4 Hz, ⁴*J* = 2.0 Hz, 1H).

¹³C NMR (DMSO-d₆), δ: 167.76 (C=O), 155.90, 148.77, 137.75, 137.12, 135.03 (C, C-aromatic), 133.58 (CH, C-aromatic), 132.85 (C, C-aromatic), 130.82, 130.24, 129.93, 129.00, 128.32, 125.99 (CH, C-aromatic), 123.94 (C, C-aromatic), 123.34 (CH, C-aromatic).

2-(2-hydroxyphenyl)quinoline-4-carboxylic acid (51)

(C₁₆H₁₁NO₃; M.W.: 265.3)



General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f:0.36.

Purification: re-crystallisation from EtOH

Yield: 110 mg (20%)

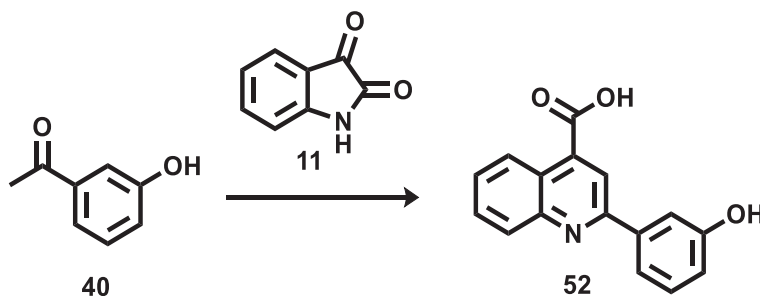
Purity: 99%

UPLC-MS method A: Rt: 2.297, MS (ESI)⁺: 266.0[M+H]⁺

¹H NMR (DMSO-d₆), δ: 14.65 (bs, 2H), 8.59 (d, *J* = 8.0 Hz, 3H), 8.40 (s, 3H), 8.16 (dd, ³*J* = 8.0 Hz, ⁴*J* = 1.2 Hz, 3H), 8.06 (d, *J* = 8.3 Hz, 3H), 7.82 (td, ³*J* = 8.3 Hz, ⁴*J* = 1.2 Hz, 3H), 7.64 (td, ³*J* = 8.2 Hz, ⁴*J* = 1.0 Hz, 3H), 7.39 (td, ³*J* = 8.4 Hz, ⁴*J* = 1.5 Hz, 3H), 7.03 – 6.97 (m, 2H).

¹³C NMR (DMSO-d₆), δ: 168.45 (C=O), 160.41, 157.69, 145.41 (C, C-aromatic), 132.55, 130.97, 128.32, 127.79, 127.42, 127.14 (CH, C-aromatic), 123.98 (C, C-aromatic), 119.50(CH, C-aromatic), 119.47 (C, C-aromatic), 118.39, 116.97 (CH, C-aromatic).

2-(3-hydroxyphenyl)quinoline-4-carboxylic acid (52)
(C₁₆H₁₁NO₃; M.W.: 265.3)



General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f:0.32.

Purification: re-crystallisation from EtOH

Yield: mg (%)

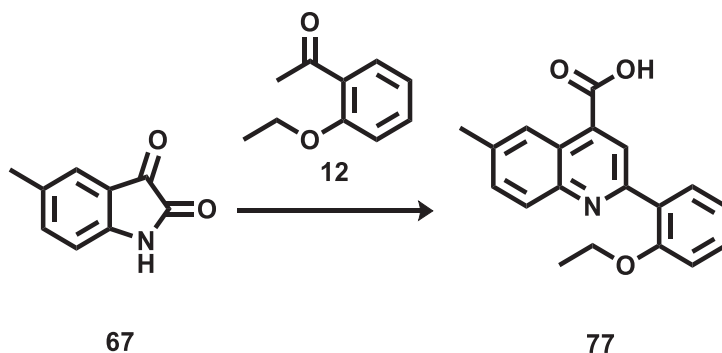
Purity: 99%

UPLC-MS method A: Rt: 2.297, MS (ESI)⁺: 266.0[M+H]⁺

¹H NMR (DMSO-d₆), δ: 9.70 (bs, 1H), 8.67 (d, *J* = 8.4 Hz, 1H), 8.38 (s, 1H), 8.15 (d, *J* = 8.4 Hz, 1H), 7.88 – 7.82 (m, 1H), 7.77 – 7.67 (m, 4H), 7.38 (t, *J* = 7.9 Hz, 1H), 6.94 (dd, ³*J* = 7.8 Hz, ⁴*J* = 2.1 Hz, 1H).

¹³C NMR (DMSO-d₆), δ: 168.10 (C=O), 158.44, 156.27, 148.86, 139.75, 138.11 (C, C-aromatic), 130.65, 130.55, 130.20, 128.16, 125.93 (CH, C-aromatic), 124.01 (C, C-aromatic), 119.63, 118.47, 117.60, 114.21 (CH, C-aromatic).

2-(2-ethoxyphenyl)-6-methylquinoline-4-carboxylic acid (77)
(C₁₉H₁₇NO₃; M.W.: 307.3)



General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f:0.58.

Purification: re-crystallisation from EtOH

Yield: 348 mg (61%)

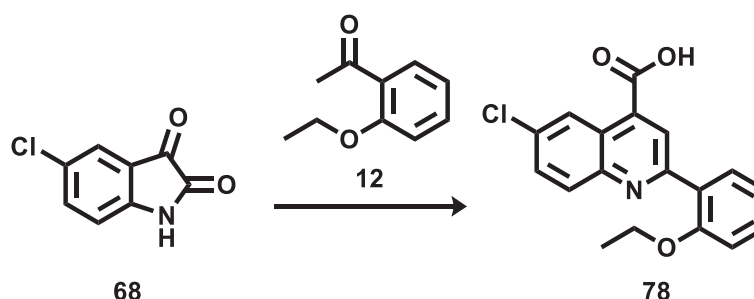
Purity: 99%

UPLC-MS method B: Rt: 1.773, MS (ESI)⁺: 308.18[M+H]⁺

¹H NMR (DMSO-d₆), δ: 13.73 (bs, 1H), 8.50 (s, 1H), 8.49 (s, 1H), 8.03 (d, *J* = 8.6 Hz, 1H), 7.89 (d, *J* = 7.5 Hz, 1H), 7.66 (d, *J* = 8.6 Hz, 1H), 7.45 (t, *J* = 7.7 Hz, 1H), 7.18 (d, *J* = 8.3 Hz, 1H), 7.11 (t, *J* = 7.5 Hz, 1H), 4.15 (q, *J* = 6.9 Hz, 2H), 2.53 (s, 3H), 1.34 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 168.31 (C=O), 156.92, 155.36, 147.68, 137.74, 135.03 (C, C-aromatic), 132.23, 131.38, 131.31, 129.95 (CH, C-aromatic), 128.28 (C, C-aromatic), 124.60, 124.57 (CH, C-aromatic), 123.78 (C, C-aromatic), 121.32, 113.61 (CH, C-aromatic), 64.39 (CH₂), 22.14, 15.00 (CH₃).

6-chloro-2-(2-ethoxyphenyl)quinoline-4-carboxylic acid (78)
(C₁₈H₁₄ClNO₃; M.W.: 327.8)



General procedure 2;

White solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f:0.59.

Purification: re-crystallisation from EtOH

Yield: 344mg (63%)

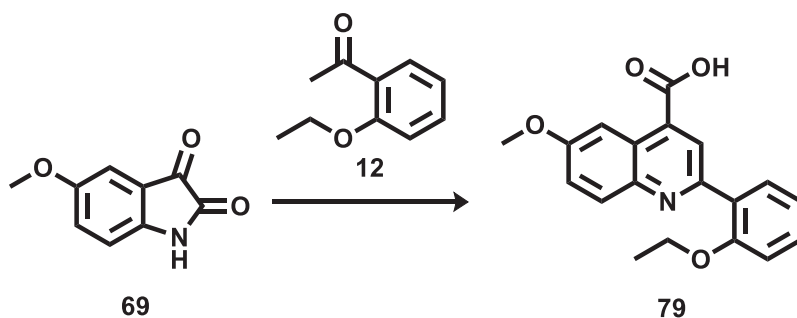
Purity: 95%

UPLC-MS method B: Rt: 2.33, MS (ESI)⁺: 328.15 (75%), 330.15 (25%)[M+H]⁺

¹H NMR (DMSO-d₆), δ: 13.68 (bs, 1H), 8.85 (s, 1H), 8.62 (s, 1H), 8.14 (d, *J* = 8.9 Hz, 1H), 7.90 (d, *J* = 7.6 Hz, 1H), 7.83 (d, *J* = 9.0 Hz, 1H), 7.48 (t, *J* = 7.8 Hz, 1H), 7.20 (d, *J* = 8.3 Hz, 1H), 7.12 (t, *J* = 7.5 Hz, 1H), 4.17 (q, *J* = 6.8 Hz, 2H), 1.34 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 167.71 (C=O), 157.02, 156.93, 147.53, 134.66, 132.76 (C, C-aromatic), 132.24, 131.84, 131.36, 130.64 (CH, C-aromatic), 127.74 (C, C-aromatic), 125.90, 124.84 (CH, C-aromatic), 124.62 (C, C-aromatic), 121.36, 113.63 (CH, C-aromatic), 64.43 (CH₂), 14.98 (CH₃).

2-(2-ethoxyphenyl)-6-methoxyquinoline-4-carboxylic acid (79)
 (C₁₉H₁₇NO₄; M.W.: 323.3)



General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f:0.56.

Purification: re-crystallisation from EtOH

Yield: 108 mg (20%)

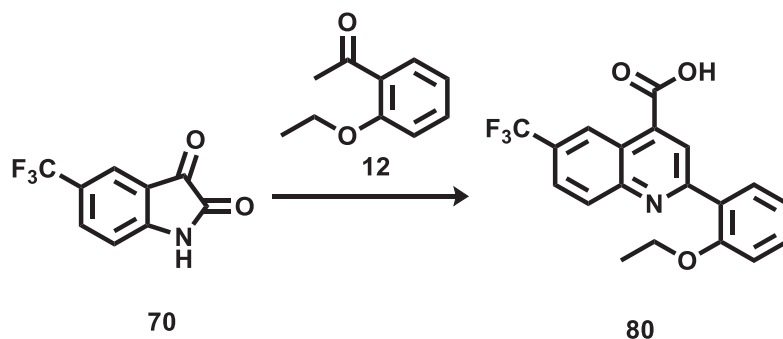
Purity: 97%

UPLC-MS method B: R_t: 1.798, MS (ESI)⁺: 324.2[M+H]⁺

¹H NMR (DMSO-d₆), δ: 13.66 (bs, 1H), 8.56 (s, 1H), 8.21 (d, ⁴J = 2.3 Hz, 1H), 8.06 (d, J = 9.2 Hz, 1H), 7.89 (d, J = 7.6 Hz, 1H), 7.49 (dd, ³J = 9.2 Hz, ⁴J = 2.3 Hz, 1H), 7.45 (t, J = 7.6 Hz, 1H), 7.19 (d, J = 8.2 Hz, 1H), 7.11 (t, J = 7.4 Hz, 1H), 4.16 (q, J = 6.9 Hz, 2H), 3.92 (s, 3H), 1.35 (t, J = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 168.28 (C=O), 158.76, 156.86, 153.55, 145.32, 133.61 (C, C-aromatic), 131.78, 131.20 (CH, C-aromatic), 128.23 (C, C-aromatic), 125.26 (CH, C-aromatic), 125.22 (C, C-aromatic), 122.36, 121.32, 113.62, 104.20 (CH, C-aromatic), 64.38 (CH₂), 55.92, 15.02 (CH₃).

2-(2-ethoxyphenyl)-6-(trifluoromethyl)quinoline-4-carboxylic acid (80)
 (C₁₉H₁₄F₃NO₃; M.W.: 361.3)



General procedure 2;

Orange solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f: 0.58.

Purification: re-crystallisation from EtOH

Yield: 26 mg (5%)

Purity: 97%

UPLC-MS method B: Rt: 2.07, MS (ESI)⁺: 362.2[M+H]⁺

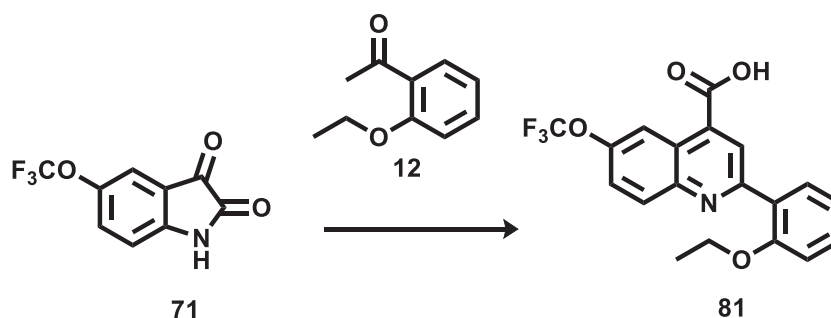
¹H NMR (DMSO-d₆), δ: 14.14 (bs, 1H), 9.25 (s, 1H), 8.73 (s, 1H), 8.35 (d, *J* = 8.7 Hz, 1H), 8.09 (d, *J* = 8.6 Hz, 1H), 7.97 (d, *J* = 7.5 Hz, 1H), 7.53 (t, *J* = 7.5 Hz, 1H), 7.25 (d, *J* = 8.3 Hz, 1H), 7.17 (t, *J* = 7.5 Hz, 1H), 4.21 (q, *J* = 6.9 Hz, 2H), 1.37 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 167.56 (C=O), 159.10, 158.94, 157.18, 150.03 (C, C-aromatic), 132.26, 131.82, 131.50 (CH, C-aromatic), 127.54 (C, C-aromatic), 126.30, 125.48, 124.18 (CH, C-aromatic), 123.10 (C, C-aromatic), 121.45, 113.73 (CH, C-aromatic), 64.50 (CH₂), 14.97 (CH₃).

¹⁹F NMR (DMSO-d₆), δ: -61.36.

2-(2-ethoxyphenyl)-6-(trifluoromethoxy)quinoline-4-carboxylic acid (81)

(C₁₉H₁₄F₃NO₄; M.W.:377.3)



General procedure 2;

white/pale yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f:0.54.

Purification: re-crystallisation from EtOH

Yield: 307 mg (63%)

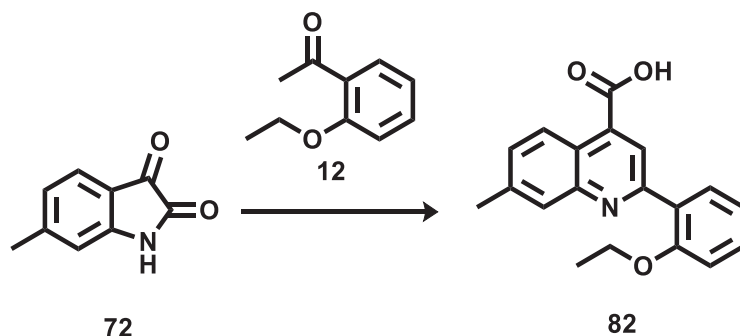
Purity: 99%

UPLC-MS method B: Rt: 2,063, MS (ESI)⁺: 378,3[M+H]⁺

¹H NMR (DMSO-d₆), δ: 14.03 (s, 1H), 8.81 (d, ⁴*J* = 1.4 Hz, 1H), 8.70 (s, 1H), 8.29 (d, *J* = 9.2 Hz, 1H), 7.93 (dd, ³*J* = 7.6 Hz, ⁴*J* = 1.7 Hz, 1H), 7.84 (dd, ³*J* = 9.1 Hz, ⁴*J* = 2.4 Hz, 1H), 7.55 – 7.48 (m, 1H), 7.23 (d, *J* = 8.2 Hz, 1H), 7.15 (t, *J* = 7.5 Hz, 1H), 4.19 (q, *J* = 6.9 Hz, 2H), 1.36 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 167.59 (C=O), 157.34, 157.05, 147.48, 134.89 (C, C-aromatic), 132.91, 131.97, 131.38 (CH, C-aromatic), 127.66 (C, C-aromatic), 126.24 (CH, C-aromatic), 124.23 (C, C-aromatic), 124.01, 121.41, 116.57, 113.68 (CH, C-aromatic), 64.46 (CH₂), 14.97 (CH₃).

¹⁹F NMR (DMSO-d₆), δ: -56.67

2-(2-ethoxyphenyl)-7-methylquinoline-4-carboxylic acid (82)**(C₁₉H₁₇NO₃; M.W.: 307.3)**

General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f: 0.57.

Purification: re-crystallisation from EtOH

Yield: 194mg (34%)

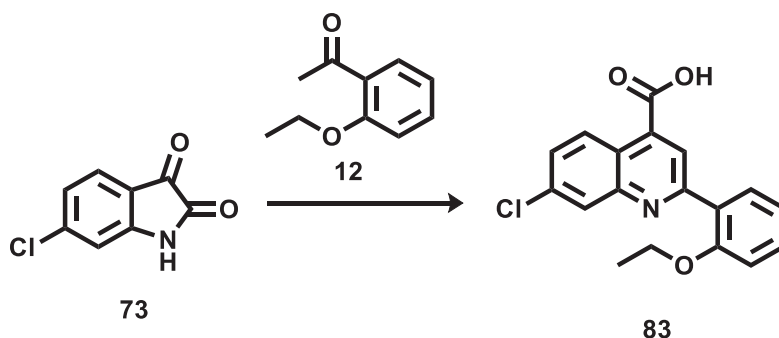
Purity: 99%

UPLC-MS method B: Rt: 1.506, MS (ESI)⁺: 308.1[M+H]⁺

¹H NMR (DMSO-d₆), δ: 13.75 (bs, 1H), 8.63 (d, *J* = 8.7 Hz, 1H), 8.45 (s, 1H), 7.93 (s, 1H), 7.89 (d, *J* = 7.6 Hz, 1H), 7.54 (d, *J* = 8.7 Hz, 1H), 7.47 (t, *J* = 7.8 Hz, 1H), 7.20 (d, *J* = 8.3 Hz, 1H), 7.12 (t, *J* = 7.4 Hz, 1H), 4.16 (q, *J* = 6.9 Hz, 2H), 2.55 (s, 3H), 1.34 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 168.27 (C=O), 156.95, 156.29, 149.30, 140.02, 135.40 (C, C-aromatic), 131.47, 131.35, 130.26, 129.09 (CH, C-aromatic), 128.34 (C, C-aromatic), 125.52, 123.87 (CH, C-aromatic), 121.82 (C, C-aromatic), 121.33, 113.64 (CH, C-aromatic), 64.40 (CH₂), 21.66, 15.01 (CH₃).

7-chloro-2-(2-ethoxyphenyl)quinoline-4-carboxylic acid (83)
($C_{18}H_{14}ClNO_3$; M.W.: 327.8)



General procedure 2;

White solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, Rf:0.60.

Purification: re-crystallisation from EtOH

Yield: 110 mg (20%)

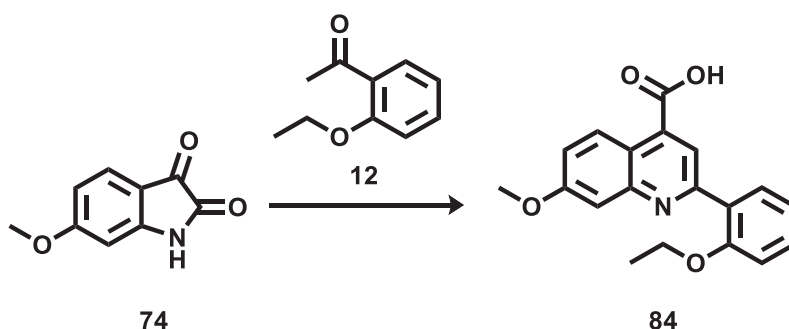
Purity: 99%

UPLC-MS method B: Rt: 2.374, MS (ESI)⁺: 328.1 (75%), 330.1 (25%)[M+H]⁺

¹H NMR (DMSO-d₆), δ: 13.91 (bs, 1H), 8.78 (d, *J* = 9.1 Hz, 1H), 8.56 (s, 1H), 8.17 (s, 1H), 7.89 (d, *J* = 7.7 Hz, 1H), 7.73 (d, *J* = 9.1 Hz, 1H), 7.48 (t, *J* = 7.7 Hz, 1H), 7.20 (d, *J* = 8.3 Hz, 1H), 7.13 (t, *J* = 7.4 Hz, 1H), 4.17 (q, *J* = 6.9 Hz, 2H), 1.34 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 167.81 (C=O), 157.78, 157.05, 149.49, 135.86, 134.77 (C, C-aromatic), 131.93, 131.43, 128.66, 128.58, 128.02 (CH, C-aromatic), 127.72 (C, C-aromatic), 125.12 (CH, C-aromatic), 122.50 (C, C-aromatic), 121.36, 113.65 (CH, C-aromatic), 64.44 (CH₂), 14.98 (CH₃).

2-(2-ethoxyphenyl)-7-methoxyquinoline-4-carboxylic acid (84)
($C_{19}H_{17}NO_4$; M.W.: 323.3)



General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, Rf:0.57.

Purification: re-crystallisation from EtOH

Yield: 92mg (17%)

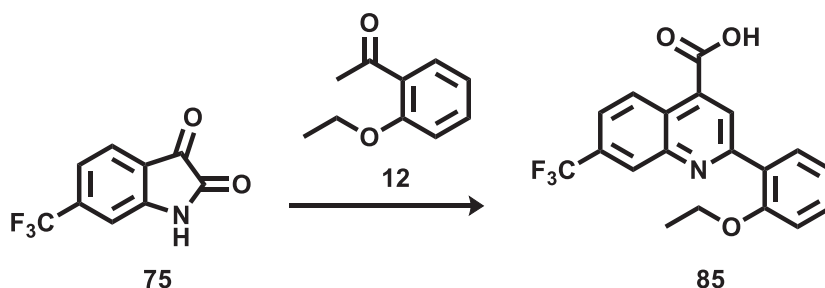
Purity: 98%

UPLC-MS method B: Rt: 1.512, MS (ESI)⁺: 324.2[M+H]⁺

¹H NMR (DMSO-d₆), δ: 13.71 (bs, 1H), 8.65 (d, *J* = 9.3 Hz, 1H), 8.40 (s, 1H), 7.92 (d, *J* = 7.6 Hz, 1H), 7.53 (d, *J* = 2.2 Hz, 1H), 7.48 (t, *J* = 7.7 Hz, 1H), 7.36 (dd, ³*J* = 9.3 Hz, ⁴*J* = 2.5 Hz, 1H), 7.21 (d, *J* = 8.2 Hz, 1H), 7.13 (t, *J* = 7.4 Hz, 1H), 4.17 (q, *J* = 6.9 Hz, 2H), 3.96 (s, 3H), 1.35 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 168.31 (C=O), 160.57, 156.97, 156.58, 150.98, 135.49 (C, C-aromatic), 131.45, 131.36 (CH, C-aromatic), 128.30 (C, C-aromatic), 126.99, 122.36, 121.27, 120.73 (CH, C-aromatic), 118.83 (C, C-aromatic), 113.66, 108.49 (CH, C-aromatic), 64.40 (CH₂), 56.00, 15.01 (CH₃).

2-(2-ethoxyphenyl)-7-(trifluoromethyl)quinoline-4-carboxylic acid (85)
(C₁₉H₁₄F₃NO₃; M.W.: 361.3)



General procedure 2;

Yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f: 0.53.

Purification: re-crystallisation from EtOH

Yield: 302 mg (60%)

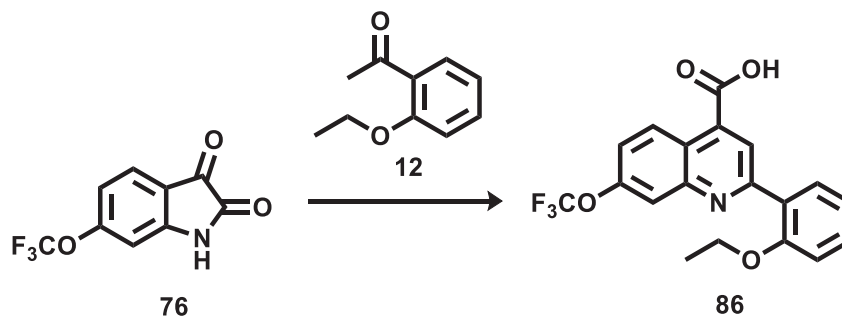
Purity: 97%

UPLC-MS method B: Rt: 2,101, MS (ESI)⁺: 362,2[M+H]⁺

¹H NMR (DMSO-d₆), δ: 14.10 (bs, 1H), 9.00 (d, *J* = 8.9 Hz, 1H), 8.72 (s, 1H), 8.48 (s, 1H), 7.99 (d, *J* = 9.1 Hz, 1H), 7.96 (d, *J* = 7.6 Hz, 1H), 7.52 (t, *J* = 7.6 Hz, 1H), 7.24 (d, *J* = 8.3 Hz, 1H), 7.16 (t, *J* = 7.4 Hz, 1H), 4.20 (q, *J* = 6.9 Hz, 2), 1.37 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 167.59 (C=O), 158.27, 157.11, 147.94, 135.71 (C, C-aromatic), 132.17, 131.46 (CH, C-aromatic), 130.21 (C, C-aromatic), 128.08, 127.50 (CH, C-aromatic), 127.46 (C, C-aromatic), 126.88 (CH, C-aromatic), 125.79, 125.48 (C, C-aromatic), 123.24, 121.43, 113.70 (CH, C-aromatic), 64.49 (CH₂), 14.97 (CH₃).

¹⁹F NMR (DMSO-d₆), δ: -61.36.

2-(2-ethoxyphenyl)-7-(trifluoromethoxy)quinoline-4-carboxylic acid (86)**(C₁₉H₁₄F₃NO₄; M.W.: 377.3)**

General procedure 2;

white/pale yellow solid;

T.L.C. System: EtOAc-MeOH-TEA 8.2:1.6:0.2 v/v, R_f:0.56.

Purification: re-crystallisation from EtOH

Yield: 100 mg (20%)

Purity: 98%

UPLC-MS method B: R_t: 2,098, MS (ESI)⁺: 378,2 [M+H]⁺

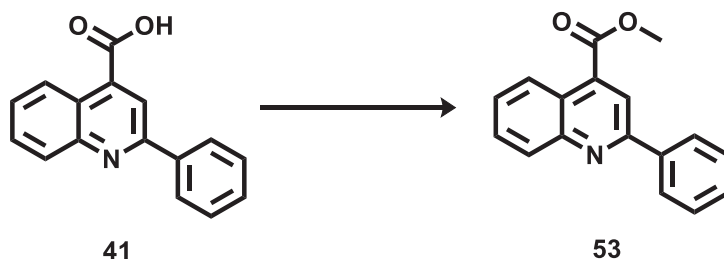
¹H NMR (DMSO-d₆), δ: 14.01 (s, 1H), 8.92 (d, *J* = 9.3 Hz, 1H), 8.63 (s, 1H), 8.06 (s, 1H), 7.94 (dd, ³*J* = 7.6 Hz, ⁴*J* = 1.7 Hz, 1H), 7.74 (dd, ³*J* = 9.1 Hz, ⁴*J* = 2.2 Hz, 1H), 7.56 – 7.47 (m, 1H), 7.23 (d, *J* = 8.2 Hz, 1H), 7.15 (t, *J* = 7.5 Hz, 1H), 4.19 (q, *J* = 6.9 Hz, 2H), 1.36 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 13C NMR (126 MHz, DMSO) δ 167.71 (C=O), 158.16, 157.08, 149.42, 149.21, 135.61 (C, C-aromatic), 132.06, 131.45, 128.79 (CH, C-aromatic), 127.60 (C, C-aromatic), 125.38 (CH, C-aromatic), 122.55 (C, C-aromatic), 121.74, 121.39, 119.97, 113.69 (CH, C-aromatic), 64.48 (CH₂), 14.97 (CH₃).

¹⁹F NMR (DMSO-d₆), δ: 19F NMR (471 MHz, DMSO) δ -56.60.

6.3.4 Methyl 2-(phenyl)quinoline-4-carboxylate analogues (53-64, 87-96)

Methyl 2-phenylquinoline-4-carboxylate (53) (C₁₇H₁₃NO₂; M.W.: 263.3)



General procedure 3;

White solid;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.57.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 70:30 v/v in 12 CV).

Yield: 38 mg (72%)

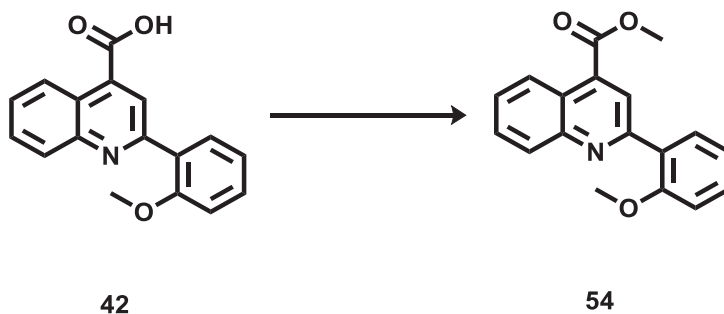
Purity: 99%

UPLC-MS method B: Rt: 2.498, MS (ESI)⁺: 264.09[M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.56 (d, J = 8.4 Hz, 1H), 8.48 (s, 1H), 8.29 (d, J = 7.6 Hz, 2H), 8.18 (d, J = 8.4 Hz, 1H), 7.87 (t, J = 7.6 Hz, 1H), 7.72 (t, J = 7.6 Hz, 1H), 7.62 – 7.51 (m, 3H), 4.03 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 166.78 (C=O), 156.26, 148.83, 138.22, 136.80 (C, C-aromatic), 130.90, 130.56, 130.33, 129.49, 128.52, 127.70, 125.59 (CH, C-aromatic), 123.61 (C, C-aromatic), 119.82 (CH, C-aromatic), 53.43 (CH₃).

Methyl 2-(2-methoxyphenyl)quinoline-4-carboxylate (54) (C₁₈H₁₅NO₃; M.W.: 293.3)



General procedure 3;

Yellow solid;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.47.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV).

Yield: 31 mg (59%)

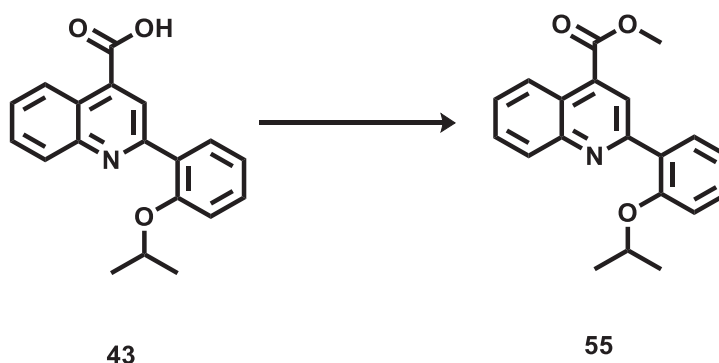
Purity: 99%

UPLC-MS method C: Rt: 2.022, MS (ESI)⁺: 294.1[M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.62 (d, *J* = 8.4 Hz, 1H), 8.37 (s, 1H), 8.17 (d, *J* = 8.4 Hz, 1H), 7.91 – 7.81 (m, 2H), 7.74 (t, *J* = 7.7 Hz, 1H), 7.52 (t, *J* = 7.8 Hz, 1H), 7.25 (d, *J* = 8.3 Hz, 1H), 7.15 (t, *J* = 7.4 Hz, 1H), 4.02 (s, 3H), 3.89 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 166.86 (C=O), 157.55, 156.47, 148.94, 134.80 (C, C-aromatic), 131.66, 131.44, 130.51, 130.27, 128.47 (CH, C-aromatic), 128.17 (C, C-aromatic), 125.56, 124.36 (CH, C-aromatic), 123.33 (C, C-aromatic), 121.38, 112.63 (CH, C-aromatic), 56.31, 53.42(CH₃).

Methyl 2-(2-isopropoxyphenyl)quinoline-4-carboxylate (55)
(C₂₀H₁₉NO₃; M.W.: 321.4)



General procedure 3;

Yellow oil;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, R_f: 0.51.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV).

Yield: 31 mg (59%)

Purity: 99%

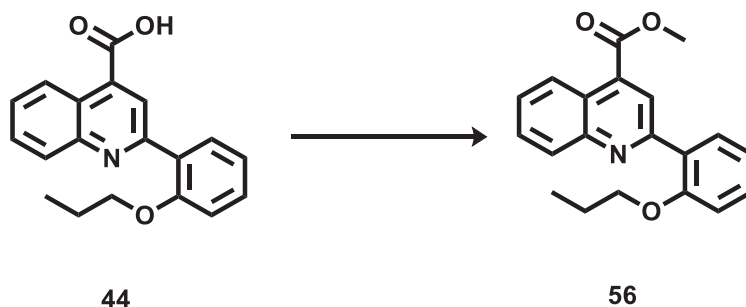
UPLC-MS method C: Rt: 2.212, MS (ESI)⁺: 322.3[M+H]⁺

¹H NMR (CDCl₃-d), δ: 8.81 (d, *J* = 8.5 Hz, 1H), 8.64 (s, 1H), 8.24 (d, *J* = 8.5 Hz, 1H), 7.99 (d, *J* = 7.6 Hz, 1H), 7.76 (t, *J* = 7.6 Hz, 1H), 7.64 (t, *J* = 7.8 Hz, 1H), 7.42 (t, *J* = 7.8 Hz, 1H), 7.14 (t, *J* = 7.5 Hz, 1H), 7.06 (d, *J* = 8.3 Hz, 1H), 4.61 (hept, *J* = 5.9 Hz, 1H), 4.06 (s, 3H), 1.35 (s, 3H), 1.34 (s, 3H).

¹³C NMR CDCl₃-d), δ: 167.10 (C=O), 156.78, 155.75, 149.22, 133.49 (C, C-aromatic), 131.52, 130.67, 130.26 (CH, C-aromatic), 129.79 (C, C-aromatic), 129.36, 127.62, 125.44,

125.35 (CH, C-aromatic), 123.85 (C, C-aromatic), 121.41, 114.74 (CH, C-aromatic), 71.17, 52.54, 22.06 (CH₃).

Methyl 2-(2-propoxyphenyl)quinoline-4-carboxylate (56)
(C₂₀H₁₉NO₃; M.W.: 321.4)



General procedure 3;

Yellow solid;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.54.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV).

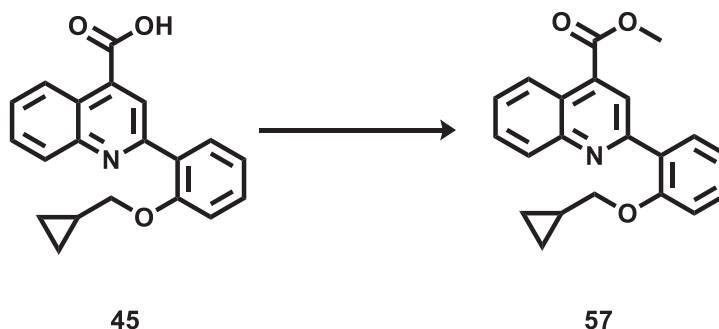
Yield: 23 mg (84%)

Purity: 99%

UPLC-MS method A: Rt: 2.653, MS (ESI)⁺: 322.23[M+H]⁺

¹H NMR (CDCl₃-d), δ: 8.71 (d, *J* = 8.5 Hz, 1H), 8.52 (s, 1H), 8.14 (d, *J* = 8.4 Hz, 1H), 7.88 (d, *J* = 7.6 Hz, 1H), 7.67 (t, *J* = 7.6 Hz, 1H), 7.55 (t, *J* = 7.8 Hz, 1H), 7.33 (t, *J* = 7.8 Hz, 1H), 7.05 (t, *J* = 7.6 Hz, 1H), 6.95 (d, *J* = 8.4 Hz, 1H), 4.00 – 3.90 (m, 5H), 1.78 – 1.66 (m, 2H), 0.95 (t, *J* = 7.4 Hz, 3H).

¹³C NMR (CDCl₃-d), δ: 167.06 (C=O), 156.98, 156.63, 149.24, 133.58 (C, C-aromatic), 131.36, 130.81, 130.25, 129.40 (CH, C-aromatic), 128.70 (C, C-aromatic), 127.64, 125.36, 125.31 (CH, C-aromatic), 123.87 (C, C-aromatic), 121.19, 112.38 (CH, C-aromatic), 70.10, 52.48 (CH₃), 22.62 (CH₂), 10.70 (CH₃).

Methyl 2-(2-(cyclopropylmethoxy)phenyl)quinoline-4-carboxylate (57)**(C₂₁H₁₉NO₃; M.W.: 333.4)**

General procedure 3;

Yellow oil;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.51.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 70:30 v/v in 12 CV).

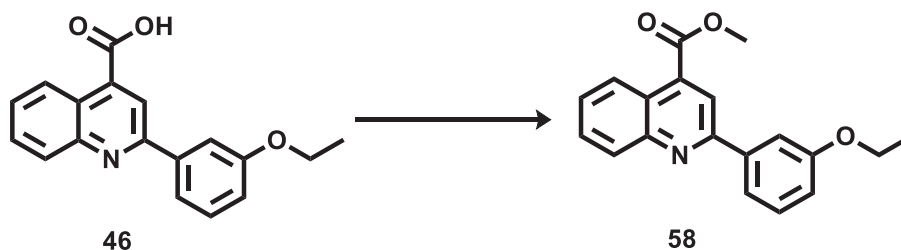
Yield: 26 mg (88%)

Purity: 98%

UPLC-MS method C: Rt: 2.207, MS (ESI)⁺: 334.2 [M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.65 (d, *J* = 8.5 Hz, 1H), 8.63 (s, 1H), 8.16 (d, *J* = 8.4 Hz, 1H), 7.93 (d, *J* = 7.6 Hz, 1H), 7.85 (t, *J* = 7.6 Hz, 1H), 7.73 (t, *J* = 7.7 Hz, 1H), 7.47 (t, *J* = 7.8 Hz, 1H), 7.19 (d, *J* = 8.3 Hz, 1H), 7.13 (t, *J* = 7.4 Hz, 1H), 4.00 (s, 5H), 1.28 – 1.17 (m, 1H), 0.61 – 0.50 (m, 2H), 0.41 – 0.31 (m, 2H).

¹³C NMR (DMSO-d₆), δ: 166.81 (C=O), 157.03, 156.36, 148.96, 134.30 (C, C-aromatic), 131.70, 131.40, 130.41, 130.28, 128.44 (CH, C-aromatic), 127.99 (C, C-aromatic), 125.51, 124.88 (CH, C-aromatic), 123.42 (C, C-aromatic), 121.36, 113.76 (CH, C-aromatic), 72.98 (CH₂), 53.27, 10.48 (CH₃), 3.35 (CH₂).

Methyl 2-(3-ethoxyphenyl)quinoline-4-carboxylate (58)**(C₁₉H₁₇NO₃; M.W.: 307.3)**

General procedure 3;

Yellow solid;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.59

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV).

Yield: 20 mg (68%)

Purity: 99%

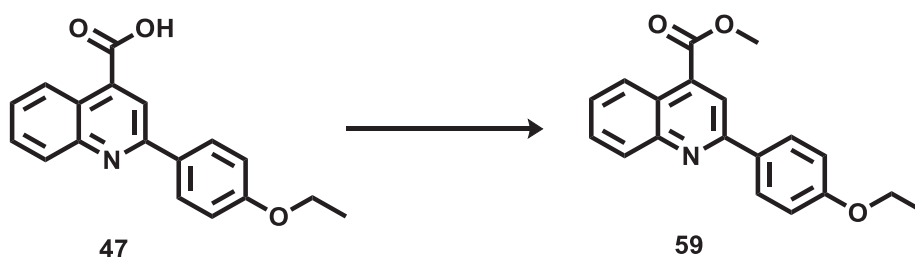
UPLC-MS method A: Rt: 2.603, MS (ESI)⁺: 308.12[M+H]⁺

¹H NMR (CDCl₃-d), δ : 8.67 (d, *J* = 8.5 Hz, 1H), 8.31 (s, 1H), 8.14 (d, *J* = 8.4 Hz, 1H), 7.76 – 7.62 (m, 3H), 7.54 (t, *J* = 7.7 Hz, 1H), 7.36 (t, *J* = 7.9 Hz, 1H), 6.94 (dd, ³*J* = 8.1 Hz, ⁴*J* = 1.7 Hz, 1H), 4.10 (q, *J* = 6.9 Hz, 2H), 3.99 (s, 3H), 1.40 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (CDCl₃-d), δ : 166.85 (C=O), 159.60, 156.53, 149.23, 140.22, 135.54 (C, C-aromatic), 130.37, 129.93, 129.90, 127.83, 125.43 (CH, C-aromatic), 124.09 (C, C-aromatic), 120.46, 119.80, 116.07, 113.46 (CH, C-aromatic), 63.66 (CH₂), 52.73, 14.91 (CH₃).

Methyl 2-(4-ethoxyphenyl)quinoline-4-carboxylate (59)

(C₁₉H₁₇NO₃; M.W.: 307.3)



General procedure 3;

Yellow solid;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.53.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV).

Yield: 24 mg (47%)

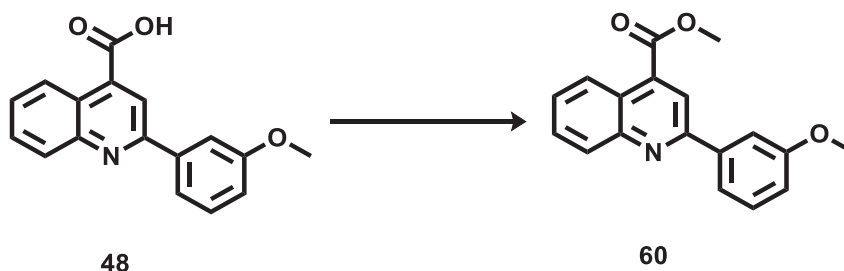
Purity: 99%

UPLC-MS method C: Rt: 2.18, MS (ESI)⁺: 308.2[M+H]⁺

¹H NMR (DMSO-d₆), δ : 8.53 (d, *J* = 8.5 Hz, 1H), 8.44 (s, 1H), 8.26 (d, *J* = 8.9 Hz, 2H), 8.14 (d, *J* = 8.4 Hz, 1H), 7.85 (t, *J* = 8.1 Hz, 1H), 7.69 (t, *J* = 8.1 Hz, 1H), 7.12 (d, *J* = 8.9 Hz, 2H), 4.14 (q, *J* = 7.0 Hz, 2H), 4.04 (s, 3H), 1.38 (t, *J* = 7.0 Hz, 3H).

¹³C NMR (DMSO-d₆), δ : 166.88 (C=O), 160.78, 155.93, 148.83, 136.66 (C, C-aromatic), 130.79 (CH, C-aromatic), 130.48 (C, C-aromatic), 130.09, 129.20, 128.01, 125.54 (CH, C-aromatic), 123.23 (C, C-aromatic), 119.35, 115.29 (CH, C-aromatic), 63.78 (CH₂), 53.39, 15.08 (CH₃).

Methyl 2-(3-methoxyphenyl)quinoline-4-carboxylate (60)
 (C₁₈H₁₅NO₃; M.W.: 293.3)



General procedure 3;

Yellow solid;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.48.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV).

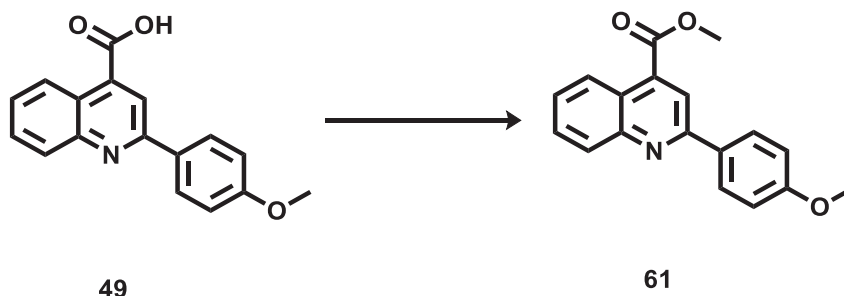
Yield: 48 mg (45%)

Purity: 98%

UPLC-MS: Rt: 2.488, MS (ESI)⁺: 294.1[M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.56 (d, *J* = 8.5 Hz, 1H), 8.49 (s, 1H), 8.20 (d, *J* = 8.5 Hz, 1H), 7.93 – 7.83 (m, 3H), 7.74 (t, *J* = 7.6 Hz, 1H), 7.51 (t, *J* = 7.9 Hz, 1H), 7.13 (d, *J* = 8.1 Hz, 1H), 4.04 (s, 3H), 3.90 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 166.83 (C=O), 160.35, 156.04, 148.72, 139.71, 136.92 (C, C-aromatic), 130.92, 130.62, 130.37, 128.57, 125.57 (CH, C-aromatic), 123.67 (C, C-aromatic), 120.11, 119.92, 116.36, 112.83 (CH, C-aromatic), 55.81, 53.44 (CH₃).

Methyl 2-(4-methoxyphenyl)quinoline-4-carboxylate (61)**(C₁₈H₁₅NO₃; M.W.: 293.3)**

General procedure 3;

White solid;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.43.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP

KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 70:30 v/v in 12 CV).

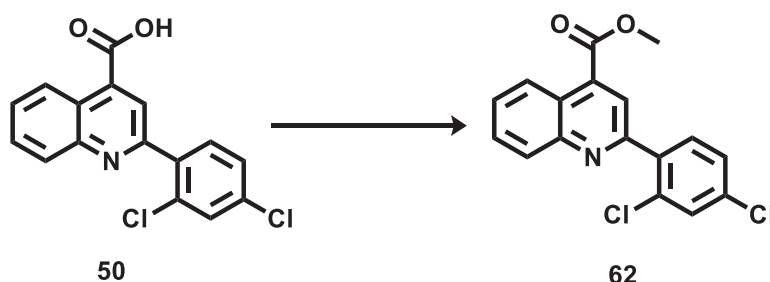
Yield: 24 mg (25%)

Purity: 99%

UPLC-MS method C: Rt: 2.454, MS (ESI)⁺: 294.12[M+H]⁺

¹H NMR (CDCl₃-d), δ: 8.71 (d, *J* = 8.5 Hz, 1H), 8.37 (s, 1H), 8.19 (d, *J* = 8.5 Hz, 3H), 7.75 (t, *J* = 7.7 Hz, 1H), 7.60 (t, *J* = 7.7 Hz, 1H), 7.06 (d, *J* = 8.4 Hz, 2H), 4.08 (s, 3H), 3.90 (s, 3H).

¹³C NMR (CDCl₃-d), δ: 166.97 (C=O), 161.16, 156.28, 149.30, 135.51, 131.40 (C, C-aromatic), 130.13, 129.84, 128.86, 127.40, 125.40 (CH, C-aromatic), 123.68 (C, C-aromatic), 119.93, 114.34 (CH, C-aromatic), 55.44, 52.71 (CH₃).

Methyl 2-(2,4-dichlorophenyl)quinoline-4-carboxylate (62)**(C₁₇H₁₁Cl₂NO₂; M.W.: 332.2)**

General procedure 3;

White solid;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.57.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP

KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV).

Yield: 17 mg (64%)

Purity: 99%

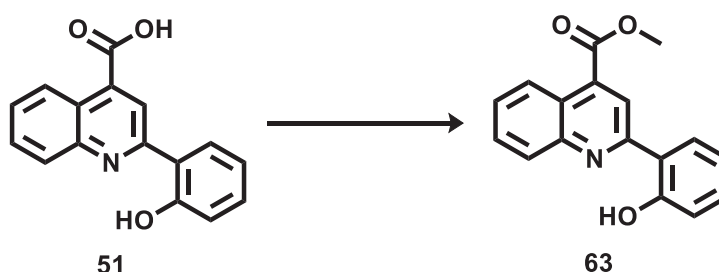
UPLC-MS method A: Rt: 2.683, MS (ESI)⁺: 332.1 (56%), 335.1(37%), 336.1 (7%) [M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.81 (d, *J* = 8.5 Hz, 1H), 8.27 (s, 1H), 8.21 (d, *J* = 8.5 Hz, 1H), 7.80 (t, *J* = 7.6 Hz, 1H), 7.69 (t, *J* = 9.4 Hz, 2H), 7.55 (s, 1H), 7.42 (d, *J* = 8.2 Hz, 1H), 4.06 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 166.52 (C=O), 155.78, 149.10, 137.37, 135.63, 134.74, 133.22 (C, C-aromatic), 132.60, 130.26, 130.11, 129.98, 128.55, 127.67, 125.55 (CH, C-aromatic), 124.08 (C, C-aromatic), 123.83 (CH, C-aromatic), 52.83 (CH₃).

Methyl 2-(2-hydroxyphenyl)quinoline-4-carboxylate (63)

(C₁₇H₁₃NO₃; M.W.: 279.3)



General procedure 3;

Yellow solid;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.48.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 70:30 v/v in 11 CV).

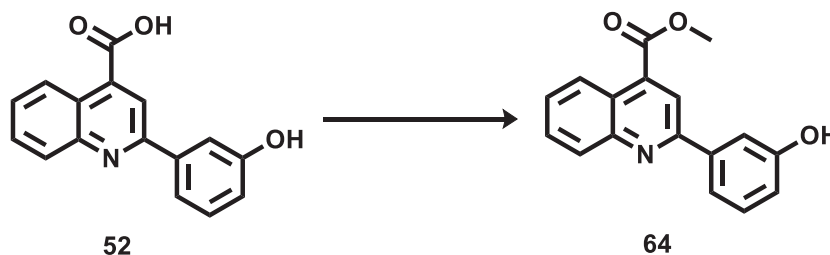
Yield: 11 mg (30%)

Purity: 99%

UPLC-MS method A: Rt: 2.106, MS (ESI)⁺: 280.0[M+H]⁺

¹H NMR (DMSO-d₆), δ: 14.74 (s, 1H), 8.75 (d, *J* = 8.5 Hz, 1H), 8.58 (s, 1H), 8.11 (d, *J* = 8.4 Hz, 1H), 8.02 (d, *J* = 8.2 Hz, 1H), 7.81 (t, *J* = 7.0 Hz, 1H), 7.67 (t, *J* = 7.0 Hz, 1H), 7.42 (t, *J* = 7.0 Hz, 1H), 7.13 (d, *J* = 8.2 Hz, 1H), 7.02 (t, *J* = 7.0 Hz, 1H), 4.13 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 166.35 (C=O), 160.87, 157.46, 145.80, 136.48 (C, C-aromatic), 132.52, 130.77, 128.17, 128.06, 127.07, 125.65 (CH, C-aromatic), 123.46 (C, C-aromatic), 119.16, 119.04, 118.74 (CH, C-aromatic), 118.57 (C, C-aromatic), 53.02 (CH₃).

Methyl 2-(3-hydroxyphenyl)quinoline-4-carboxylate (64)**(C₁₇H₁₃NO₃; M.W.: 279.3)**

General procedure 3;

Orange solid;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.20.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP

KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV).

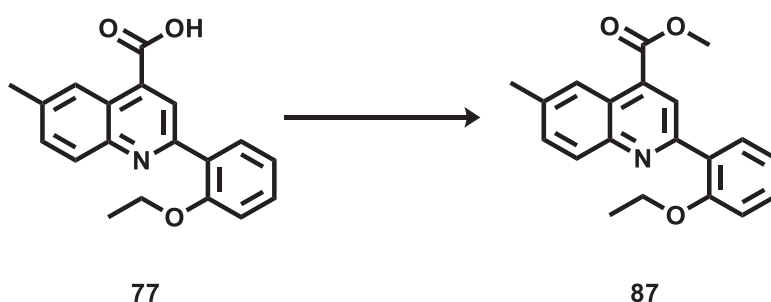
Yield: 40 mg (79%)

Purity: >95%

UPLC-MS method A: Rt: 2.084, MS (ESI)⁺: 280.03[M+H]⁺

¹H NMR (CDCl₃-d), δ: 8.74 (d, *J* = 8.5 Hz, 1H), 8.34 (s, 1H), 8.22 (d, *J* = 8.4 Hz, 1H), 7.75 (d, *J* = 9.2 Hz, 2H), 7.67 – 7.59 (m, 2H), 7.36 (t, *J* = 7.8 Hz, 1H), 6.93 (d, *J* = 7.7 Hz, 1H), 6.61 (s, 1H), 4.07 (s, 3H).

¹³C NMR (CDCl₃-d), δ: 166.81 (C=O), 156.76, 156.63, 148.96, 140.17, 135.88 (C, C-aromatic), 130.24, 130.19, 129.87, 128.00, 125.48 (CH, C-aromatic), 124.14 (C, C-aromatic), 120.82, 119.80, 117.21, 114.65 (CH, C-aromatic), 52.85 (CH₃).

Methyl 2-(2-ethoxyphenyl)-6-methylquinoline-4-carboxylate (87)**(C₂₀H₁₉NO₃; M.W.: 321.4)**

General procedure 3;

White solid;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.51.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP

KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 70:30 v/v in 12 CV).

Yield: 58 mg (99%)

Purity: 99%

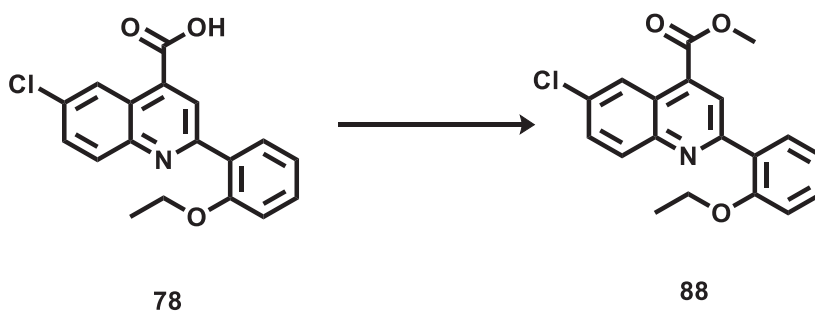
UPLC-MS method B: Rt: 2.606, MS (ESI)⁺: 322.24[M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.51 (s, 1H), 8.41 (s, 1H), 8.05 (d, *J* = 8.5 Hz, 1H), 7.92 (d, *J* = 7.1 Hz, 1H), 7.68 (d, *J* = 8.5 Hz, 1H), 7.47 (t, *J* = 7.4 Hz, 1H), 7.19 (d, *J* = 8.3 Hz, 1H), 7.12 (t, *J* = 7.4 Hz, 1H), 4.16 (q, *J* = 6.9 Hz, 2H), 4.00 (s, 3H), 2.55 (s, 3H), 1.36 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 166.91 (C=O), 156.96, 155.24, 147.63, 138.16, 133.64 (C, C-aromatic), 132.48, 131.53, 131.25, 130.03 (CH, C-aromatic), 127.95 (C, C-aromatic), 124.66, 124.22 (CH, C-aromatic), 123.39 (C, C-aromatic), 121.34, 113.60 (C,H C-aromatic), 64.43, 53.27, 22.12, 14.96 (CH₃).

Methyl 6-chloro-2-(2-ethoxyphenyl)quinoline-4-carboxylate (88)

(C₁₉H₁₆ClNO₃; M.W.: 341.8)



General procedure 3;

Yellow solid;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, R_f: 0.51.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 70:30 v/v in 12 CV).

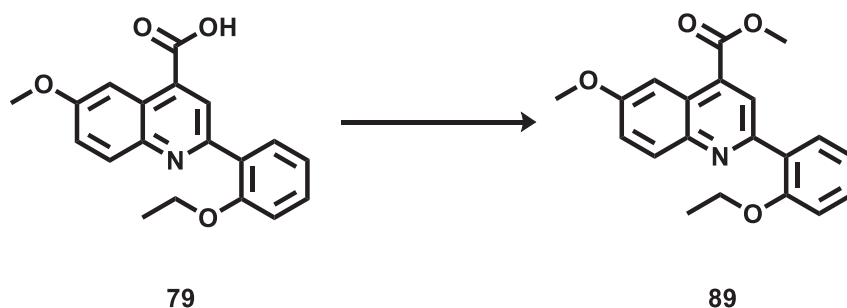
Yield: 39 mg (70%)

Purity: 99%

UPLC-MS method C: Rt: 2.333, MS (ESI)⁺: 342.2 (75%), 344.2 (25%)[M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.73 (s, 1H), 8.65 (s, 1H), 8.16 (d, *J* = 9.0 Hz, 1H), 7.93 (d, *J* = 7.2 Hz, 1H), 7.87 (d, *J* = 9.0 Hz, 1H), 7.50 (t, *J* = 7.6 Hz, 1H), 7.21 (d, *J* = 8.3 Hz, 1H), 7.13 (t, *J* = 7.4 Hz, 1H), 4.18 (q, *J* = 6.9 Hz, 2H), 4.00 (s, 3H), 1.37 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 166.30 (C=O), 157.07, 156.78, 147.46, 133.19, 133.12 (C, C-aromatic), 132.33, 132.02, 131.31, 130.94 (CH, C-aromatic), 127.39 (C, C-aromatic), 125.89, 124.52 (CH, C-aromatic), 124.13 (C, C-aromatic), 121.41, 113.64 (CH, C-aromatic), 64.49 (CH₂), 53.49, 14.94 (CH₃).

Methyl 2-(2-ethoxyphenyl)-6-methoxyquinoline-4-carboxylate (89)**(C₂₀H₁₉NO₄; M.W.: 337.38)**

General procedure 3;

pale yellow solid;

T.L.C. System: n-Hex/EtOAc 9:1 v/v, R_f: 0,37.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 70:30 v/v in 12 CV).

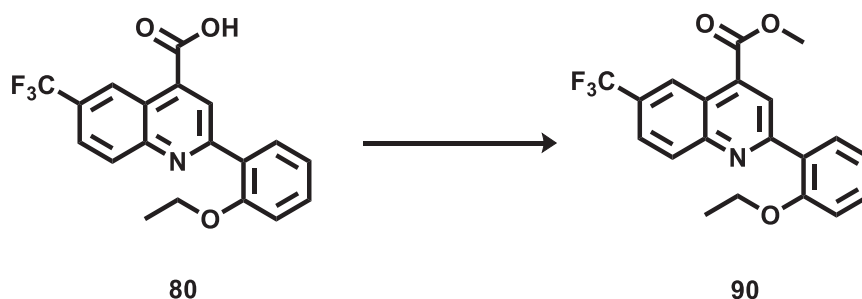
Yield: 28 mg (54%)

Purity: 99%

UPLC-MS method C: R_t:2.14, MS (ESI)⁺: 338.2 [M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.61 (s, 1H), 8.13 (d, ⁴J = 2.6 Hz, 1H), 8.08 (d, J = 9.2 Hz, 1H), 7.93 (dd, ³J = 7.7 Hz, ⁴J = 1.5 Hz, 1H), 7.52 (dd, ³J = 9.2 Hz, ⁴J = 2.6 Hz, 1H), 7.46 (t, J = 7.7 Hz, 1H), 7.20 (d, J = 8.3 Hz, 1H), 7.13 (t, J = 7.5 Hz, 1H), 4.18 (q, J = 6.9 Hz, 2H), 4.01 (s, 3H), 3.94 (s, 3H), 1.38 (t, J = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 166.91 (C=O), 158.99, 156.89, 153.40, 145.28, 132.29 (C, C-aromatic), 131.88, 131.31, 131.13 (CH, C-aromatic), 127.91 (C, C-aromatic), 125.23 (CH, C-aromatic), 124.89 (C, C-aromatic), 122.60, 121.33, 113.59, 103.85 (CH, C-aromatic), 64.41 (CH₂), 55.96, 53.21, 14.97 (CH₃).

Methyl 2-(2-ethoxyphenyl)-6-(trifluoromethyl)quinoline-4-carboxylate (90)**(C₂₀H₁₆F₃NO₃; M.W.: 375.35)**

General procedure 3;

pale yellow solid;

T.L.C. System: n-Hex/EtOAc 9:1 v/v, Rf:0. 0.42.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 70:30 v/v in 12 CV).

Yield: 307 mg (63%)

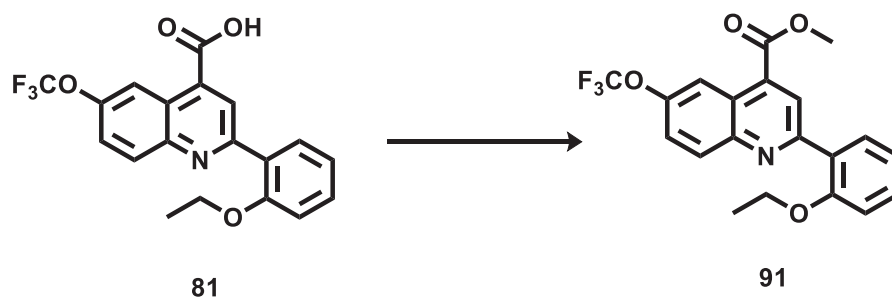
Purity: 99%

UPLC-MS method C: Rt: 2.326, MS (ESI)⁺: 376.3 [M+H]⁺

¹H NMR (CDCl₃-d), δ: 9.13 (s, 1H), 8.67 (s, 1H), 8.25 (d, *J* = 8.8 Hz, 1H), 7.92 (dd, ³*J* = 7.6 Hz, ⁴*J* = 1.7 Hz, 1H), 7.85 (dd, ³*J* = 8.8 Hz, ⁴*J* = 1.9 Hz, 1H), 7.40 – 7.34 (m, 1H), 7.07 (t, *J* = 7.5 Hz, 1H), 6.98 (d, *J* = 8.2 Hz, 1H), 4.09 (q, *J* = 6.9 Hz, 2H), 4.00 (s, 3H), 1.36 (t, *J* = 7.0 Hz, 3H).

¹³C NMR (CDCl₃-d), δ: 166.22(C=O), 158.12, 157.16, 147.89, 134.32 (C, C-aromatic), 132.32, 131.41, 130.54 (CH, C-aromatic), 127.78, 127.62, 127.59 (C, C-aromatic), 127.14, 126.97, 125.38 (CH, C-aromatic), 123.50 (C, C-aromatic), 121.45, 113.68 (CH, C-aromatic), 64.53 (CH₂), 53.56, 14.93 (CH₃).

¹⁹F NMR (CDCl₃-d), δ: -61.36.

Methyl 2-(2-ethoxyphenyl)-6-(trifluoromethoxy)quinoline-4-carboxylate (91)**(C₂₀H₁₆F₃NO₄; M.W.: 391.35)**

General procedure 3;

white solid;

T.L.C. System: n-Hex/EtOAc 9:1 v/v, R_f: 0. 0.47.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge:

ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 70:30 v/v in 12 CV).

Yield: 58 mg (56%)

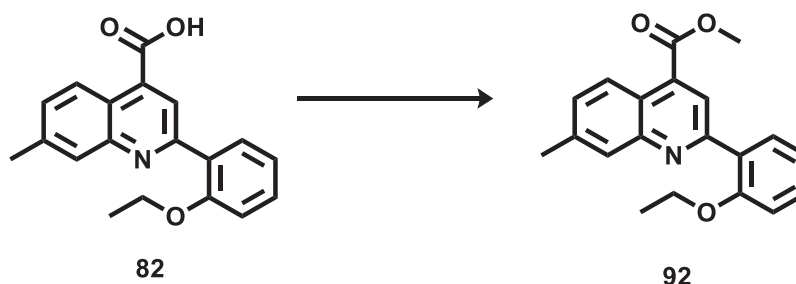
Purity: 99%

UPLC-MS method B: Rt: 2.347, MS (ESI)⁺: 392.2 [M+H]⁺

¹H NMR (DMSO-*d*₆), δ : 8.72 (s, 1H), 8.71 (s, 1H), 8.30 (d, $J = 9.2$ Hz, 1H), 7.95 (dd, $^3J = 7.6$ Hz, $^4J = 1.7$ Hz, 1H), 7.86 (dd, $^3J = 9.1$ Hz, $^4J = 2.4$ Hz, 1H), 7.55 – 7.48 (m, 1H), 7.23 (d, $J = 8.3$ Hz, 1H), 7.15 (t, $J = 7.5$ Hz, 1H), 4.20 (q, $J = 6.9$ Hz, 2H), 4.02 (s, 3H), 1.38 (t, $J = 6.9$ Hz, 4H).

¹³C NMR (DMSO-*d*₆), δ : 166.23 (C=O), 157.18, 157.10, 147.43, 133.67 (C, C-aromatic), 132.99, 132.10, 131.32 (CH, C-aromatic), 127.35 (C, C-aromatic), 126.18, 124.21 (CH, C-aromatic), 123.83 (C, C-aromatic), 121.42, 116.29, 113.66 (CH, C-aromatic), 64.51 (CH₂), 53.51, 14.93 (CH₃).

¹⁹F NMR (DMSO-*d*₆), δ : -56.60.

Methyl 2-(2-ethoxyphenyl)-7-methylquinoline-4-carboxylate (92)**(C₂₀H₁₉NO₃; M.W.: 321.38)**

General procedure 3;

pale yellow solid;

T.L.C. System: n-Hex/EtOAc 9:1 v/v, R_f: 0.034.

Purification: Flash column chromatography Isolera One system (Biotage),

Cartridge:

ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 70:30 v/v in 12 CV).

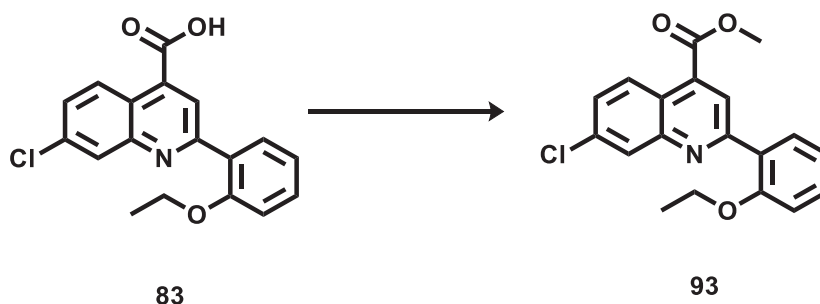
Yield: 58 mg (62%)

Purity: 99%

UPLC-MS method C: R_t: 2.185, MS (ESI)⁺: 322.3 [M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.55 (d, *J* = 8.7 Hz, 1H), 8.50 (s, 1H), 7.96 (s, 1H), 7.92 (dd, ³*J* = 7.6 Hz, ⁴*J* = 1.8 Hz, 1H), 7.58 (dd, ³*J* = 8.7 Hz, ⁴*J* = 1.7 Hz, 1H), 7.49 (ddd, ³*J* = 8.4, 7.4 Hz, ⁴*J* = 1.8 Hz, 1H), 7.21 (d, *J* = 8.3 Hz, 1H), 7.14 (td, ³*J* = 7.5 Hz, ⁴*J* = 0.9 Hz, 1H), 4.18 (q, *J* = 6.9 Hz, 2H), 4.00 (s, 3H), 2.56 (s, 3H), 1.37 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 166.90 (C=O), 156.99, 156.18, 149.25, 140.32, 134.04 (C, C-aromatic), 131.61, 131.29, 130.57, 129.17 (CH, C-aromatic), 128.03 (C, C-aromatic), 125.19, 123.96 (CH, C-aromatic), 121.44 (C, C-aromatic), 121.35, 113.63 (CH, C-aromatic), 64.44 (CH₂), 53.28, 21.65, 14.97 (CH₃).

Methyl 7-chloro-2-(2-ethoxyphenyl)quinoline-4-carboxylate (93)**(C₁₉H₁₆ClNO₃; M.W.: 341.8)**

General procedure 3;

White solid;

T.L.C. System: n-hexane -EtOAc 8:2 v/v, R_f: 0.51.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP

KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 70:30 v/v in 12 CV).

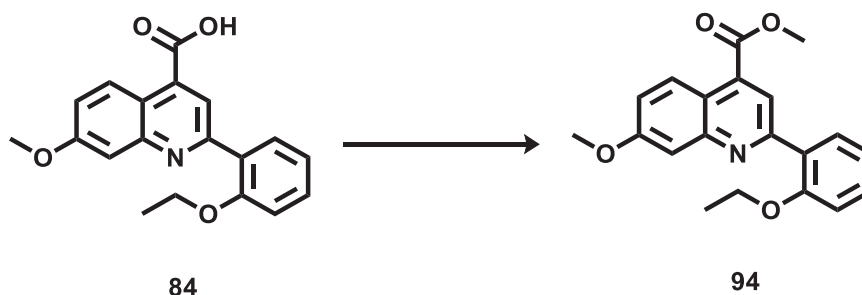
Yield: 11 mg (21%)

Purity: 99%

UPLC-MS method C: Rt: 2.353, MS (ESI)⁺: 342.2 (75%), 344.2 (25%)[M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.68 (d, *J* = 9.1 Hz, 1H), 8.60 (s, 1H), 8.19 (d, ⁴*J* = 1.4 Hz, 1H), 7.92 (d, *J* = 7.5 Hz, 1H), 7.76 (dd, ³*J* = 9.1 Hz, ⁴*J* = 1.8 Hz, 1H), 7.50 (t, *J* = 7.6 Hz, 1H), 7.21 (d, *J* = 8.3 Hz, 1H), 7.13 (t, *J* = 7.4 Hz, 1H), 4.18 (q, *J* = 6.9 Hz, 2H), 4.00 (s, 3H), 1.36 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 166.41 (C=O), 157.64, 157.10, 149.42, 135.03, 134.32 (C, C-aromatic), 132.10, 131.37, 128.89, 128.77, 127.67 (CH, C-aromatic), 127.38 (C, C-aromatic), 125.23 (CH, C-aromatic), 122.08 (C, C-aromatic), 121.39, 113.64 (CH, C-aromatic), 64.49 (CH₂), 53.47, 14.94 (CH₃).

Methyl 2-(2-ethoxyphenyl)-7-methoxyquinoline-4-carboxylate (94)**(C₂₀H₁₉NO₄; M.W.: 337.38)**

General procedure 3;

pale yellow solid;

T.L.C. System: n-Hex/EtOAc 9:1 v/v, Rf:0. 0.37.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge:

ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 70:30 v/v in 12 CV).

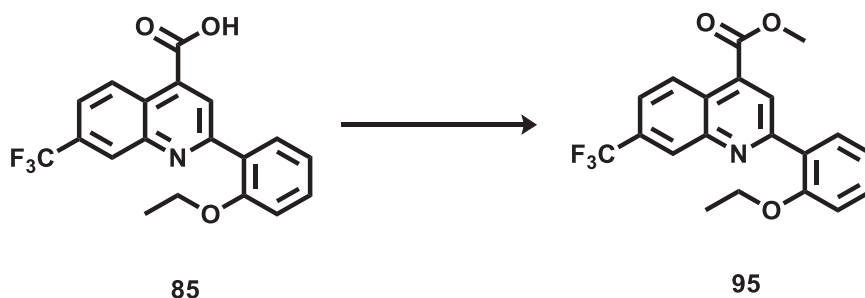
Yield: 24 mg (46%)

Purity: 99%

UPLC-MS method C: Rt: 2.08, MS (ESI)⁺: 338.3 [M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.56 (d, *J* = 9.3 Hz, 1H), 8.43 (s, 1H), 7.95 (dd, ³*J* = 7.6 Hz, ⁴*J* = 1.7 Hz, 1H), 7.55 (d, *J* = 2.6 Hz, 1H), 7.48 (ddd, ³*J* = 9.0, 7.4 Hz, ⁴*J* = 1.8 Hz, 1H), 7.38 (dd, ³*J* = 9.3 Hz, ⁴*J* = 2.7 Hz, 1H), 7.21 (d, *J* = 8.0 Hz, 1H), 7.14 (td, ³*J* = 7.6 Hz, ⁴*J* = 0.8 Hz, 1H), 4.18 (q, *J* = 6.9 Hz, 2H), 4.00 (s, 3H), 3.97 (s, 3H), 1.37 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 166.95 (C=O), 160.72, 157.01, 156.49, 150.96, 134.13 (C, C-aromatic), 131.60, 131.31 (CH, C-aromatic), 127.99 (C, C-aromatic), 126.68, 122.40, 121.30, 121.05 (CH, C-aromatic), 118.45 (C, C-aromatic), 113.64, 108.58 (CH, C-aromatic), 64.44 (CH₂), 56.04, 53.29, 14.97 (CH₃).

Methyl 2-(2-ethoxyphenyl)-7-(trifluoromethyl)quinoline-4-carboxylate (95)**(C₂₀H₁₆F₃NO₃; M.W.: 375.35)**

General procedure 3;

yellow solid;

T.L.C. System: n-Hex/EtOAc 9:1 v/v, R_f: 0.042.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge:

ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 70:30 v/v in 12 CV).

Yield: 39 mg (37%)

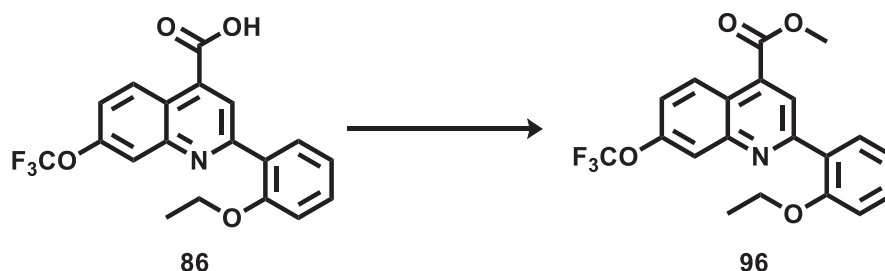
Purity: 99%

UPLC-MS method C: R_t: 2.355, MS (ESI)⁺: 376.3 [M+H]⁺

¹H NMR (DMSO-*d*₆), δ: 8.90 (d, *J* = 8.9 Hz, 1H), 8.75 (s, 1H), 8.49 (s, 1H), 8.04 – 7.96 (m, 2H), 7.56 – 7.50 (m, 1H), 7.24 (d, *J* = 8.3 Hz, 1H), 7.16 (t, *J* = 7.5 Hz, 1H), 4.21 (q, *J* = 6.9 Hz, 2H), 4.03 (s, 3H), 1.39 (t, *J* = 6.9 Hz, 3H).

¹³C NMR (DMSO-*d*₆), δ: 166.22 (C=O), 158.12, 157.16, 147.89, 134.32 (C, C-aromatic), 132.32, 131.41 (CH, C-aromatic), 130.54 (C, C-aromatic), 127.78, 127.60 (CH, C-aromatic), 127.14 (C, C-aromatic), 126.97 (CH, C-aromatic), 125.38 (C, C-aromatic), 123.50, 121.45, 113.68 (CH, C-aromatic), 64.53 (CH₂), 53.56, 14.93 (CH₃).

¹⁹F NMR (DMSO-*d*₆), δ -61.36.

Methyl 2-(2-ethoxyphenyl)-7-(trifluoromethoxy)quinoline-4-carboxylate (96)**(C₂₀H₁₆F₃NO₄; M.W.: 391.35)**

General procedure 3;

off-white solid;

T.L.C. System: n-Hex/EtOAc 9:1 v/v, R_f: 0.47.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge:

ZIP KP Sil 5g (n-hexane -EtOAc 100:0 v/v increasing to 70:30 v/v in 12 CV).

Yield: 39 mg (75%)

Purity: 99%

UPLC-MS method C: R_t: 2.362, MS (ESI)⁺: 392.3 [M+H]⁺

¹H NMR (DMSO-d₆), δ: 8.82 (d, *J* = 9.3 Hz, 1H), 8.65 (s, 1H), 8.07 (d, ⁴*J* = 1.1 Hz, 1H), 7.96 (dd, ³*J* = 7.7 Hz, ⁴*J* = 1.7 Hz, 1H), 7.75 (dd, ³*J* = 9.3 Hz, ⁴*J* = 1.9 Hz, 1H), 7.52 (ddd, ³*J* = 8.4, 7.4 Hz, ⁴*J* = 1.8 Hz, 1H), 7.23 (d, *J* = 8.0 Hz, 1H), 7.15 (td, ³*J* = 7.6 Hz, ⁴*J* = 0.9 Hz, 1H), 4.20 (q, *J* = 6.9 Hz, 2H), 4.02 (s, 3H), 1.38 (t, *J* = 6.9 Hz, 3H).

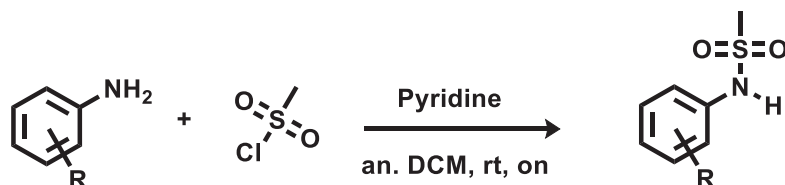
¹³C NMR (DMSO-d₆), δ: 166.36 (C=O), 158.02, 157.12, 149.36, 134.35 (C, C-aromatic), 132.19, 131.40, 128.51 (CH, C-aromatic), 127.30 (C, C-aromatic), 125.44 (CH, C-aromatic), 122.15 (C, C-aromatic), 121.98, 121.40, 120.02, 113.67 (CH, C-aromatic), 64.51 (CH₂), 53.51, 14.93 (CH₃).

¹⁹F NMR (DMSO-d₆), δ: -56.60.

6.4 Synthesis of 2-(2-(*N*-(3-chloro-2-methylphenyl)methylsulfonamido) acetamido)benzoic acid and derivatives

6.4.1 General procedures 4-8

General procedure 4: synthesis of *N*-mesylphenylamines



Aniline (4.254 mmol) and pyridine (4.680 mmol) were dissolved in anhydrous DCM (10 ml) under a nitrogen atmosphere. Methane sulfonyl chloride (4.254 mmol) was added dropwise at 0°C. The reaction mixture was stirred at rt overnight. After the complete conversion of the starting material aniline, the reaction mixture was quenched with a aqueous solution of NaOH (3M). The aqueous layer was extracted with DCM (3x20 ml) to remove the pyridine. Then the aqueous layer was acidified with HCl conc. and the resulting precipitate was collected by filtration. The recovered precipitate was dried under vacuum and used without further purification.

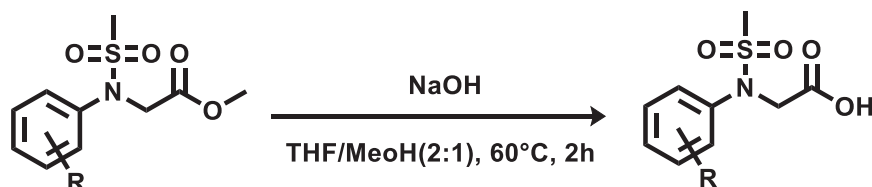
General procedure 5: synthesis of methyl *N*-mesyl-*N*-phenylglycinates



To a solution of *N*-phenylmethanesulfonamide (0.789 mmol) in anhydrous DMF (2 ml), NaH (0.789 mmol) was added under nitrogen atmosphere. The reaction mixture was stirred for 10 min before adding drop-wise a solution of methyl 2-bromoacetate (0.658 mmol) in anhydrous DMF (1 ml). The resulting reaction mixture was stirred at rt for 5 hours. The reaction was quenched with water (8 ml) and then the aqueous layer was extracted with EtOAc (3x10 ml). The organic layers were combined, washed with brine (3 x 30 mL) and dried over MgSO₄. The organic solvent was evaporated at reduced pressure and the crude

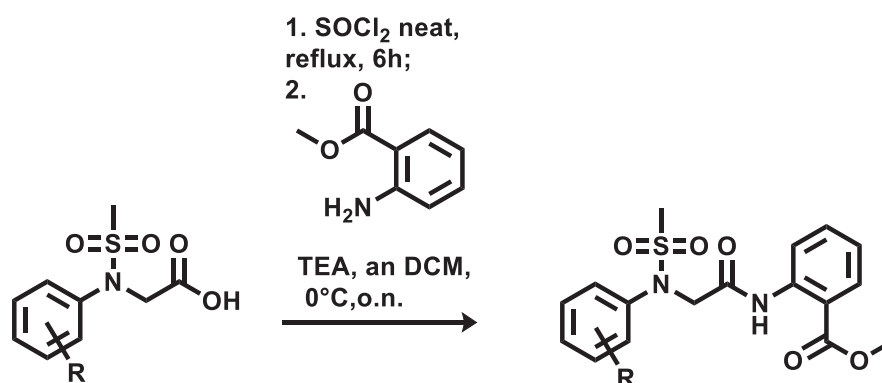
residue was purified by flash column chromatography to afford pure methyl *N*-(methylsulfonyl)-*N*-phenylglycinate.

General procedure 6: synthesis of *N*-mesyl-*N*-phenylglycines



To a solution of *N*-(methylsulfonyl)-*N*-phenylglycinate (3.78 mmol) in tetrahydrofuran and methanol (0.1 M, ratio 2:1) was added a solution 1N of NaOH (6 mL). The reaction mixture was stirred at 60°C for 2 hours. The reaction mixture was concentrated under reduced pressure and the residue was dissolved in water (20 mL). The aqueous layer was acidified with HCl 2N and then extracted with EtOAc (3x20mL). The organic layers were combined, washed with Brine and then dried over MgSO₄ and concentrated under reduced pressure to give *N*-(methylsulfonyl)-*N*-phenylglycine that were used for the following step without further purification.

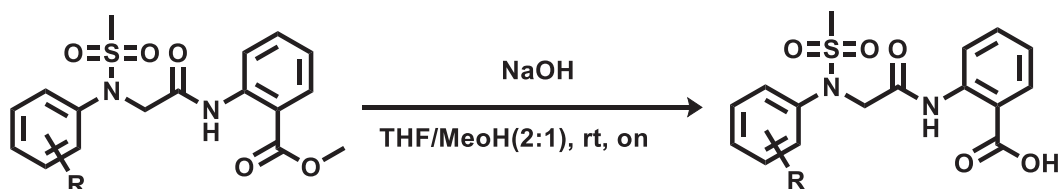
General procedure 7: synthesis of methyl 2-(2-(*N*-phenylmesyl)acetamido)benzoates



N-(methylsulfonyl)-*N*-phenylglycine (0.5 mmol) in SOCl₂ (0.5 ml) was refluxed for 6 hours. SOCl₂ was eliminated from the reaction mixture under reduced pressure. The resulting residue was dissolved in anhydrous DCM (0.5 ml) and added drop-wise to a solution, at 0°C, of methyl 2-aminobenzoate (0.6 mmol), triethylamine (1.0 mmol) in anhydrous DCM (3 ml) under nitrogen atmosphere. The reaction was allowed to warm-up to rt and was stirred overnight. The reaction mixture was quenched with water and the organic layer was extracted with DCM (3x10ml). The organic layers were combined, washed with brine (3 x

30 mL) and dried over MgSO_4 . The organic solvent was evaporated at reduced pressure and the crude residue was purified by flash column chromatography to afford pure methyl 2-(2-(*N*-phenylmethylsulfonamido)acetamido)benzoate.

General procedure 8: synthesis of 2-(2-(*N*-phenylmesyl)acetamido)benzoic acids

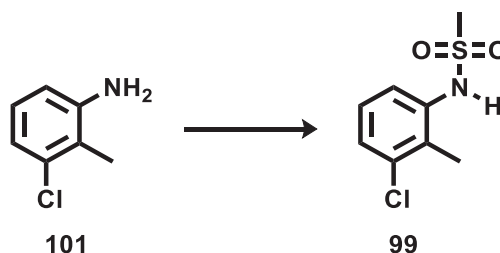


To a solution of methyl 2-(2-(*N*-phenylmethylsulfonamido)acetamido)benzoate (0.244 mmol) in tetrahydrofuran and methanol (0.05 M, ratio 2:1) was added a solution 1N of NaOH (0.39 mL). The reaction mixture was stirred overnight at room temperature. The reaction mixture was concentrated under reduced pressure and the residue was dissolved in water (20 mL). The aqueous layer was acidified with HCl 2N and then extracted with DCM (3x10mL). The organic layers were combined, washed with Brine, then dried over MgSO_4 and concentrated under reduced pressure. The crude residue was purified by flash column chromatography to afford pure 2-(2-(*N*-phenylmethylsulfonamido)acetamido)benzoic acid.

6.4.2 Synthesis of 2-(2-(*N*-(3-chloro-2-methylphenyl)methylsulfonyl)acetamido)benzoic acid (2)

N-(3-chloro-2-methylphenyl)methanesulfonamide (99)

(C₈H₁₀ClNO₂S; M.W.: 219.68)



General procedure 4;

White powder;

T.L.C. System: n-Hex/EtOAc 7:3 v/v, Rf: 0.37.

Yield: 941 mg (98%)

Purity: 99%

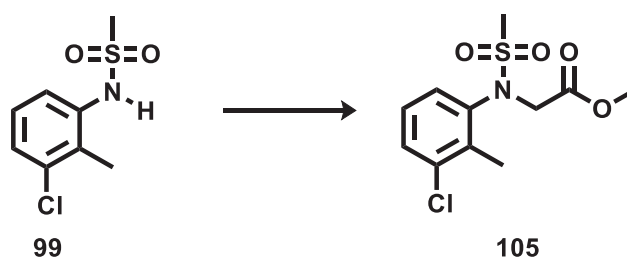
UPLC-MS method C: Rt: 1.657, MS (ESI): 218.1 [M-H]⁻

¹H NMR (CDCl₃-d), δ: 7.41 (d, *J* = 8.0 Hz, 1H), 7.31 (d, *J* = 8.0 Hz, 1H), 7.19 (t, *J* = 8.1 Hz, 1H), 6.43 (s, 1H), 3.06 (s, 3H), 2.42 (s, 3H).

¹³C NMR (CDCl₃-d), δ: 135.84, 135.66, 129.78 (C, C-aromatic), 127.43, 122.14 (CH, C-aromatic), 40.09, 15.03 (CH₃).

Methyl *N*-(3-chloro-2-methylphenyl)-*N*-(methylsulfonyl)glycinate (105)

(C₁₁H₁₄ClNO₄S; M.W.: 291.75)



General procedure 5;

pale yellow solid;

T.L.C. System: n-Hex/EtOAc 7:3 v/v, Rf: 0.38.

Yield: 1.68g (99%)

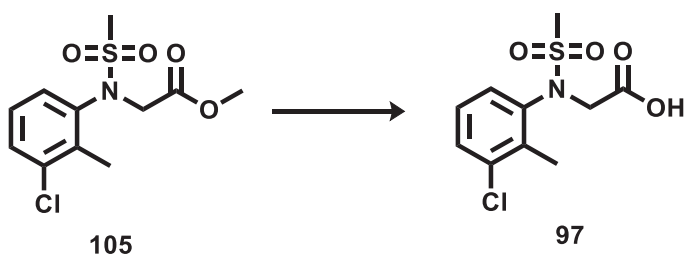
Purity: 99%

UPLC-MS method C: Rt: 1.825, MS (ESI)⁺: 292.0 (75%), 294.0 (25%)[M+H]⁺, 311.2 [M+Na]⁺
¹H NMR (DMSO-d₆), δ: 7.54 (d, *J* = 8.0 Hz, 1H), 7.50 (d, 1H), 7.31 (t, *J* = 8.0 Hz, 1H), 4.47 (dd, *J* = 88.8, 18.1 Hz, 2H), 3.65 (s, 3H), 3.15 (s, 3H), 2.40 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 169.94 (C=O), 140.81, 137.89, 134.93 (C, C-aromatic), 129.86, 128.70, 128.02 (CH, C-aromatic), 52.69 (CH₂), 52.53, 39.30, 16.46 (CH₃).

***N*-(3-chloro-2-methylphenyl)-*N*-(methylsulfonyl)glycine (97)**

(C₁₀H₁₂ClNO₄S; M.W.: 277.72)



General procedure 6;

White powder;

T.L.C. System: DCM/MeOH 95:5 v/v, R_f: 0.23.

Yield: 846 mg (81%)

Purity: 99%

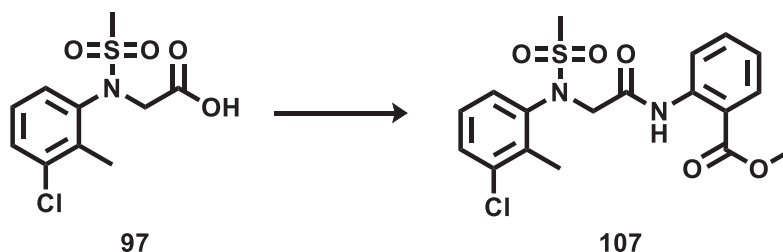
UPLC-MS method C: Rt: 1.617, MS (ESI)⁻: 276.1 [M-H]⁻

¹H NMR (DMSO-d₆), δ: 12.83 (bs, 1H), 7.56 (d, *J* = 8.0 Hz, 1H), 7.49 (d, *J* = 8.0 Hz, 1H), 7.30 (t, *J* = 8.0 Hz, 1H), 4.34 (dd, *J* = 123.8, 18.0 Hz, 2H), 3.15 (s, 3H), 2.39 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 170.84, 140.98 (C=O), 137.92, 134.88 (C, C-aromatic), 129.78, 128.73, 127.95 (CH, C-aromatic), 52.82 (CH₂), 39.52, 16.46 (CH₃).

Methyl 2-(2-(*N*-(3-chloro-2-methylphenyl)methylsulfonamido)acetamido)benzoate
(107)

(C₁₈H₁₉ClN₂O₅S; M.W.: 410.87)



General procedure 7;

White powder;

T.L.C. System: n-Hex/EtOAc 7:3 v/v, Rf: 0.25.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: SNAP KP-Sil 25g (n-hexane -EtOAc 100:0 v/v increasing to 60:40 v/v in 20 CV).

Yield: 346 mg (84%)

Purity: 99%

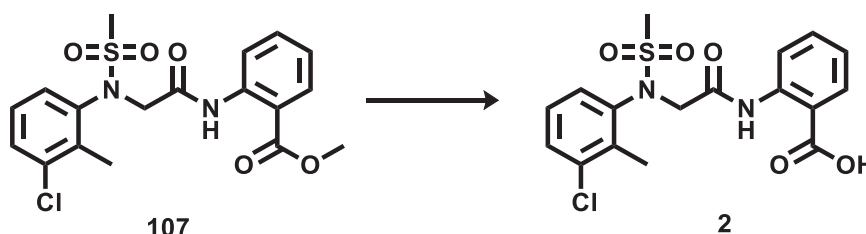
UPLC-MS method C: Rt: 1.988, MS (ESI)⁺: 411.2 (75%), 413.2 (25%)[M+H]⁺, 433.2 [M+Na]⁺

¹H NMR (DMSO-d₆), δ: 10.62 (s, 1H), 8.14 (d, *J* = 7.7 Hz, 1H), 7.90 (dd, ³*J* = 7.9 Hz, ⁴*J* = 1.5 Hz, 1H), 7.66 – 7.56 (m, 2H), 7.50 (d, *J* = 7.4 Hz, 1H), 7.31 (t, *J* = 8.0 Hz, 1H), 7.28 – 7.20 (m, 1H), 4.55 (dd, *J* = 115.5, 17.1 Hz, 2H), 3.83 (s, 3H), 3.23 (s, 3H), 2.41 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 167.73 (C=O), 167.44 (C=O), 140.96, 139.03, 137.70, 134.95 (C, C-aromatic), 134.32, 130.98, 129.86, 128.91, 128.04, 124.30, 122.13 (C, C-aromatic), 119.15 (CH, C-aromatic), 55.26 (CH₂), 52.88, 39.71, 16.47 (CH₃).

2-(2-(*N*-(3-chloro-2-methylphenyl)methylsulfonamido)acetamido)benzoic acid (2)

(C₁₇H₁₇ClN₂O₅S; M.W.: 396.84)



General procedure 8;

White powder;

T.L.C. System: DCM/MeOH 97:3 v/v, Rf: 0.35.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge:

SNAP KP-Sil 25g (DCM-MeOH 100:0 v/v increasing to 90:10 v/v in 16 CV).

Yield: 39 mg (40%)

Purity: 99%

UPLC-MS method C: Rt: 1.8, MS (ESI)⁺: 397.2 (75%), 399.2 (25%) [M+H]⁺, 419.2 [M+Na]⁺

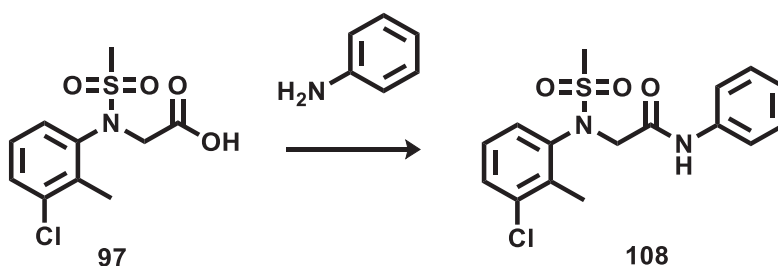
¹H NMR (DMSO-d₆), δ: 11.86 (bs, 1H), 8.42 (d, *J* = 7.9 Hz, 1H), 7.99 (dd, ³*J* = 7.9 Hz, ⁴*J* = 1.5 Hz, 1H), 7.68 (d, *J* = 7.7 Hz, 1H), 7.57 – 7.51 (m, 1H), 7.49 (d, *J* = 7.7 Hz, 1H), 7.30 (t, *J* = 8.0 Hz, 1H), 7.15 (t, *J* = 7.2 Hz, 1H), 4.53 (dd, *J* = 106.0, 17.0 Hz, 2H), 3.23 (s, 3H), 2.42 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 169.82 (C=O), 167.21 (C=O), 141.05, 141.01, 140.51, 137.62, 134.99 (C, C-aromatic), 133.85, 133.82, 131.59, 129.84, 128.98, 128.01, 123.40, 120.41 (CH, C-aromatic), 55.74 (CH₂), 39.51, 16.56 (CH₃).

6.4.3 2-(*N*-(3-chloro-2-methylphenyl)mesyl) acetamides analogues (108-113)

2-(*N*-(3-chloro-2-methylphenyl)methylsulfonamido)-*N*-phenylacetamide (108)

(C₁₆H₁₇ClN₂O₃S; M.W.: 352.83)



General procedure 7;

off-white solid;

T.L.C. System: n-Hex/EtOAc 7:3 v/v, Rf: 0.21.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: SNAP KP-Sil 25g (n-hexane -EtOAc 100:0 v/v increasing to 50:50 v/v in 21 CV).

Yield: 126 mg (71%)

Purity: 99%

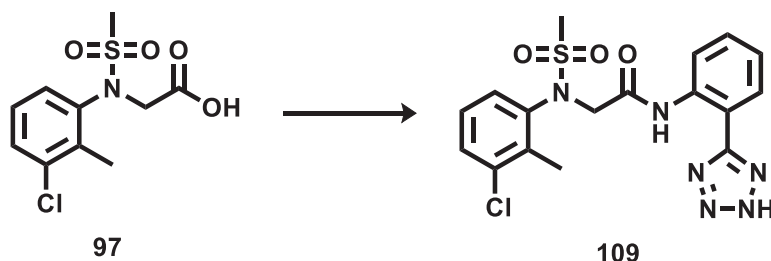
UPLC-MS method C: Rt: 1.836, MS (ESI)⁺: 353.2 (75%), 355.2 (25%) [M+H]⁺, 375.2 [M+Na]⁺

¹H NMR (DMSO-d₆), δ: 10.03 (s, 1H), 7.59 (d, *J* = 8.0 Hz, 1H), 7.54 (d, *J* = 7.7 Hz, 2H), 7.50 (d, *J* = 7.4 Hz, 1H), 7.33 – 7.28 (m, 3H), 7.06 (t, *J* = 7.4 Hz, 1H), 4.44 (dd, *J* = 208.1, 16.9 Hz, 2H), 3.26 (s, 3H), 2.40 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 167.07 (C=O), 140.95, 138.90, 137.69, 134.89 (C, C-aromatic), 129.88, 129.29, 129.07, 128.05, 124.10, 119.70 (CH, C-aromatic), 54.53 (CH₂), 40.57, 16.29 (CH₃).

***N*-(2-(1H-tetrazol-5-yl)phenyl)-2-(*N*-(3-chloro-2-methylphenyl)methylsulfonamido)acetamide (109)**

(C₁₇H₁₇ClN₆O₃S; M.W.: 420.87)



General procedure 7;

pale yellow solid;

T.L.C. System: DCM/MeOH 95:5 v/v, Rf: 0.25.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: SNAP KP-Sil 25g (DCM-MeOH 100:0 v/v increasing to 99:1 v/v in 24 CV).

Yield: 29 mg (14%)

Purity: 99%

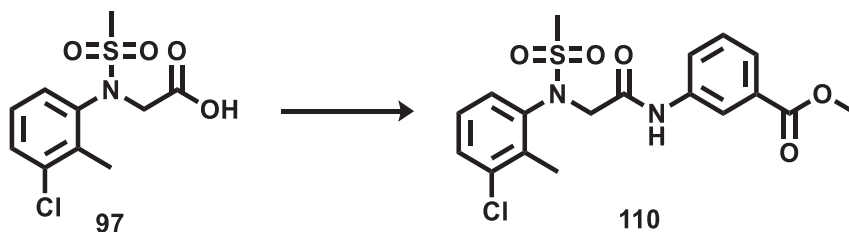
UPLC-MS method C: Rt: 1.735, MS (ESI)⁺: 421.2 (75%), 423.2 (25%) [M+H]⁺, 443.2 [M+Na]⁺

¹H NMR (DMSO-d₆), δ: 12.48 (s, 1H), 8.47 (d, *J* = 7.8 Hz, 1H), 8.18 (dd, ³*J* = 7.8 Hz, ⁴*J* = 1.3 Hz, 1H), 7.73 (d, *J* = 7.5 Hz, 1H), 7.47 (d, *J* = 7.5 Hz, 1H), 7.32 – 7.25 (m, 2H), 7.17 – 7.12 (m, 1H), 4.59 (dd, *J* = 144.8, 16.9 Hz, 2H), 3.29 (s, 3H), 2.45 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 166.86 (C=O), 141.05, 137.88, 136.17, 134.92 (C, C-aromatic), 129.80, 129.09, 128.60, 127.98, 127.90, 123.76, 120.13 (CH, C-aromatic), 118.95 (C, C-aromatic), 55.62 (CH₂), 40.02, 16.58 (CH₃).

Methyl 3-(2-(*N*-(3-chloro-2-methylphenyl)methylsulfonamido)acetamido)benzoate (110)

(C₁₈H₁₉ClN₂O₅S; M.W.: 410.87)



General procedure 7;

White powder;

T.L.C. System: n-Hex/EtOAc 7:3 v/v, Rf: 0.28.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: SNAP KP-Sil 25g (n-hexane -EtOAc 100:0 v/v increasing to 60:40 v/v in 28 CV).

Yield: 110 mg (37%)

Purity: 99%

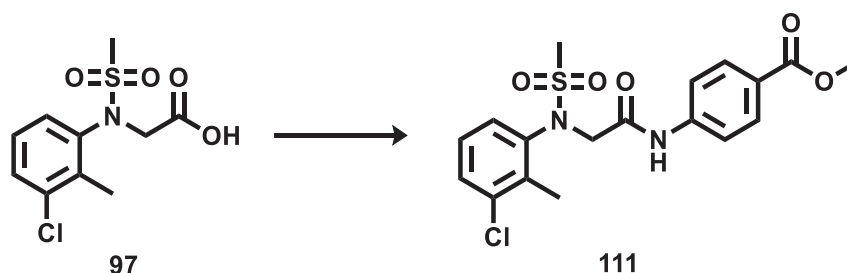
UPLC-MS method C: Rt: 1.842, MS (ESI)⁺: 411.3 (75%), 413.3 (25%) [M+H]⁺, 433.2 [M+Na]⁺

¹H NMR (DMSO-d₆), δ: 10.29 (s, 1H), 8.26 (s, 1H), 7.75 (d, ³J = 8.1 Hz, ⁴J = 2.1 Hz, 1H), 7.66 (d, J = 8.0 Hz, 1H), 7.60 (d, J = 7.9 Hz, 1H), 7.51 – 7.44 (m, 2H), 7.31 (t, J = 8.0 Hz, 1H), 4.46 (dd, J = 203.3, 17.0 Hz, 2H), 3.86 (s, 3H), 3.26 (s, 3H), 2.41 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 167.46 (C=O), 166.46 (C=O), 140.92, 139.31, 137.64, 134.89, 130.67 (C, C-aromatic), 129.90, 129.85, 129.15, 128.07, 124.70, 124.13, 120.04 (CH, C-aromatic), 54.51 (CH₂), 52.71, 40.55, 16.29 (CH₃).

Methyl 4-(2-(*N*-(3-chloro-2-methylphenyl)methylsulfonamido)acetamido)benzoate (111)

(C₁₈H₁₉ClN₂O₅S; M.W.: 410.87)



General procedure 7;

off-white solid;

T.L.C. System: n-Hex/EtOAc 7:3 v/v, Rf: 0.27.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: SNAP KP-Sil 25g (n-hexane -EtOAc 100:0 v/v increasing to 50:50 v/v in 25 CV).

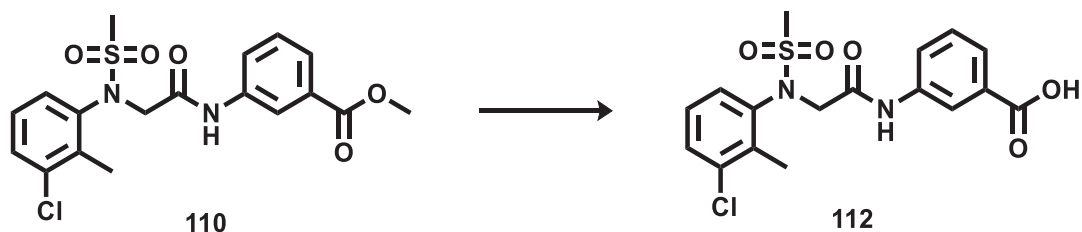
Yield: 186 mg (45%)

Purity: 99%

UPLC-MS method C: Rt: 1.853, MS (ESI)⁺: 411.3 (75%), 413.3 (25%) [M+H]⁺, 433.3 [M+Na]⁺

¹H NMR (DMSO-d₆), δ: 10.39 (s, 1H), 7.92 (d, J = 8.8 Hz, 2H), 7.69 (d, J = 8.8 Hz, 2H), 7.60 (d, J = 7.9 Hz, 1H), 7.50 (d, J = 7.4 Hz, 1H), 7.30 (t, J = 8.0 Hz, 1H), 4.49 (dd, J = 194.8, 17.2 Hz, 2H), 3.83 (s, 3H), 3.25 (s, 3H), 2.41 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 167.77 (C=O), 166.21 (C=O), 143.28, 140.93, 137.67, 134.90 (C, C-aromatic), 130.84, 129.90, 129.13, 128.06 (CH, C-aromatic), 124.79 (C, C-aromatic), 119.11 (CH, C-aromatic), 54.57 (CH₂), 52.38, 40.50, 16.30 (CH₃).

3-(2-(N-(3-chloro-2-methylphenyl)methylsulfonamido)acetamido)benzoic acid (112)**(C₁₇H₁₇ClN₂O₅S; M.W.: 396.84)**

General procedure 8;

White powder;

T.L.C. System: DCM/MeOH 97:3 v/v, R_f: 0.27.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge:

SNAP KP-Sil 25g (DCM-MeOH 100:0 v/v increasing to 90:10 v/v in 16 CV).

Yield: 23 mg (28%)

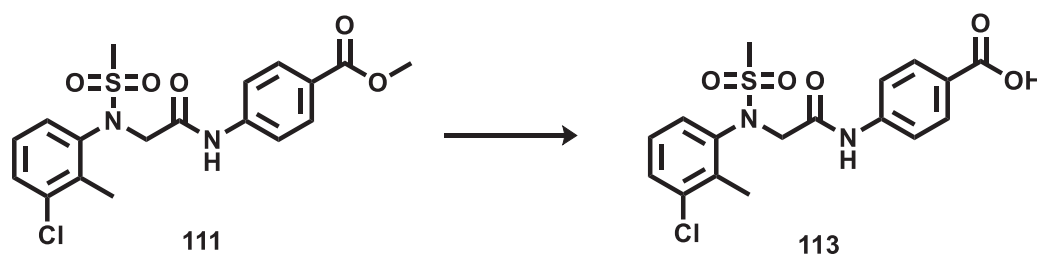
Purity: 99%

UPLC-MS method C: Rt: 1.709, MS (ESI)⁺: 397.2 (75%), 399.2 (25%) [M+H]⁺, 419.2[M+Na]⁺

¹H NMR (DMSO-d₆), δ: 13.00 (bs, 1H), 10.24 (s, 1H), 8.22 (s, 1H), 7.72 (d, *J* = 8.1 Hz, 1H), 7.64 (d, *J* = 7.7 Hz, 1H), 7.60 (d, *J* = 8.0 Hz, 1H), 7.49 (d, *J* = 8.1 Hz, 1H), 7.43 (t, *J* = 7.9 Hz, 1H), 7.30 (t, *J* = 8.1 Hz, 1H), 4.45 (dd, *J* = 206.3, 17.1 Hz, 2H), 3.25 (s, 3H), 2.40 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 167.52 (C=O), 167.39 (C=O), 140.93, 139.15, 137.65, 134.88, 131.86 (C, C-aromatic), 129.89, 129.60, 129.14, 128.07, 124.88, 123.74, 120.33 (CH, C-aromatic), 54.50 (CH₂), 40.57, 16.29 (CH₃).

4-(2-(*N*-(3-chloro-2-methylphenyl)methylsulfonamido)acetamido)benzoic acid (113)
(C₁₇H₁₇ClN₂O₅S; M.W.: 396.84)



General procedure 8;

White powder;

T.L.C. System: DCM/MeOH 97:3 v/v, R_f: 0.25.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: SNAP KP-Sil 25g (DCM-MeOH 100:0 v/v increasing to 90:10 v/v in 16 CV).

Yield: 68 mg (70%)

Purity: 99%

UPLC-MS method C: Rt: 1.717, MS (ESI)⁺: 397.3 (75%), 399.3 (25%) [M+H]⁺

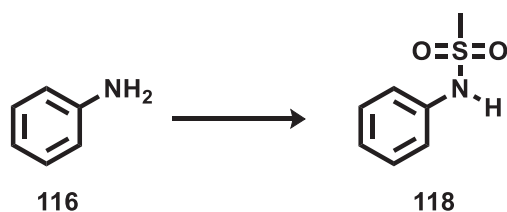
¹H NMR (DMSO-*d*₆), δ: 12.75 (bs, 1H), 10.35 (s, 1H), 7.90 (d, *J* = 8.8 Hz, 2H), 7.66 (d, *J* = 8.8 Hz, 2H), 7.59 (d, *J* = 8.0 Hz, 1H), 7.49 (d, *J* = 8.1 Hz, 1H), 7.30 (t, *J* = 8.2 Hz, 1H), 4.48 (dd, *J* = 197.1, 17.2 Hz, 2H), 3.25 (s, 3H), 2.40 (s, *J* = 8.2 Hz, 3H).

¹³C NMR (DMSO-*d*₆), δ: 167.68 (C=O), 167.30 (C=O), 142.91, 140.93, 137.67, 134.89 (C, C-aromatic), 130.94, 129.90, 129.12, 128.06 (CH, C-aromatic), 125.98 (C, C-aromatic), 118.98 (CH, C-aromatic), 54.56 (CH₂), 40.51, 16.30 (CH₃).

6.4.4 Methyl 2-(2-(*N*-mesylphenyl)acetamido)benzoic acids analogues (114-115, 124-125)

N-phenylmethanesulfonamide (118)

(C₇H₉NO₂S; M.W.:171.21)



General procedure 4;

White powder;

T.L.C. System: n-Hex/EtOAc 7:3 v/v, Rf:0. 0.33.

Yield: 1.45g (99%)

Purity: 99%

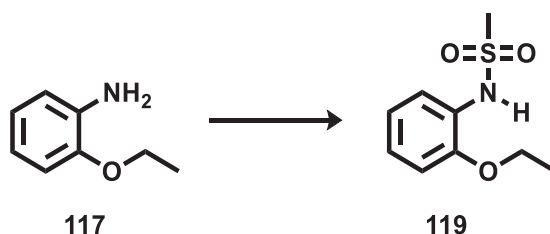
UPLC-MS method C: Rt: 1.338, MS (ESI)⁺: 171.9 [M+H]⁺

¹H NMR (DMSO-d₆), δ: 9.72 (s, 1H), 7.36 – 7.31 (m, 2H), 7.25 – 7.18 (m, 2H), 7.11 (t, *J* = 7.4 Hz, 1H), 2.98 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 138.86 (C, C-aromatic), 129.76, 124.30, 120.24 (CH, C-aromatic), 39.65 (CH₃).

N-(2-ethoxyphenyl)methanesulfonamide (119)

(C₉H₁₃NO₃S; M.W.: 215.27)



General procedure 4;

White powder;

T.L.C. System: n-Hex/EtOAc 7:3 v/v, Rf:0. 0.35.

Yield: 1.75g (96%)

Purity: 99%

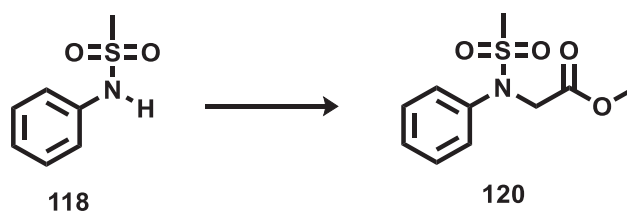
UPLC-MS method C: Rt: 1.608, MS (ESI)⁺: 216.1 [M+H]⁺, 235.1 [M+Na]⁺

¹H NMR (DMSO-d₆), δ: 8.81 (s, 1H), 7.25 (dd, ³J = 7.8 Hz, ⁴J = 1.6 Hz, 1H), 7.18 (ddd, ³J = 8.2, 7.5 Hz, ⁴J = 1.7 Hz, 1H), 7.05 (dd, ³J = 8.3 Hz, ⁴J = 1.2 Hz, 1H), 6.91 (td, ³J = 7.7 Hz, ⁴J = 1.3 Hz, 1H), 4.08 (q, J = 7.0 Hz, 2H), 2.95 (s, 3H), 1.37 (t, J = 7.0 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 152.28 (C, C-aromatic), 127.20, 126.51 (CH, C-aromatic), 126.28 (C, C-aromatic), 120.84, 113.01 (CH, C-aromatic), 64.25 (CH₂), 40.63, 14.99 (CH₃).

Methyl *N*-(methylsulfonyl)-*N*-phenylglycinate (120)

(C₁₀H₁₃NO₄S; M.W.: 243.28)



General procedure 5;

pale yellow powder;

T.L.C. System: n-Hex/EtOAc 7:3 v/v, Rf: 0.38.

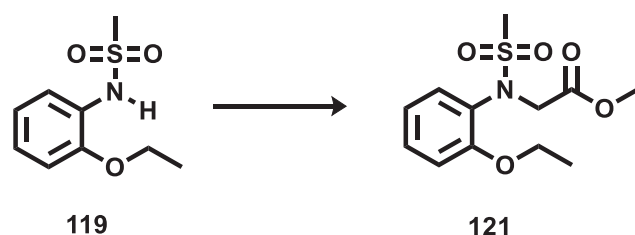
Yield: 1.55g (96%)

Purity: 99%

UPLC-MS method C: Rt: 1.564, MS (ESI)⁺: 244.1 [M+H]⁺, 266.0 [M+Na]⁺

¹H NMR (DMSO-d₆), δ: 7.48 – 7.40 (m, 4H), 7.34 (m, 1H), 4.53 (s, 2H), 3.65 (s, 3H), 3.10 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 170.18 (C=O), 140.67 (C, C-aromatic), 129.73, 128.02, 127.95 (CH, C-aromatic), 52.53 (CH₃), 52.52 (CH₂), 39.28 (CH₃).

Methyl *N*-(2-ethoxyphenyl)-*N*-(methylsulfonyl)glycinate (121)**(C₁₂H₁₇NO₅S; M.W.: 287.33)**

General procedure 5;

purple solid;

T.L.C. System: n-Hex/EtOAc 7:3 v/v, R_f: 0.35.

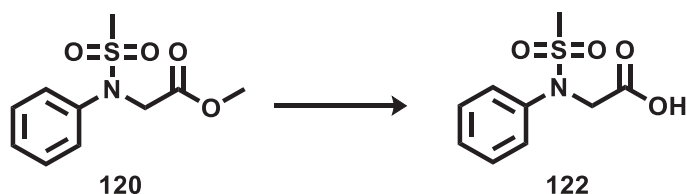
Yield: 1.55 g (98%)

Purity: 99%

UPLC-MS method C: Rt: 1.717, MS (ESI)⁺: 288.0 [M+H]⁺, 310.1 [M+Na]⁺

¹H NMR (DMSO-d₆), δ: 7.41 (dd, ³J = 7.8 Hz, ⁴J = 1.7 Hz, 1H), 7.33 (ddd, ³J = 8.3, 7.5 Hz, ⁴J = 1.8 Hz, 1H), 7.10 (dd, ³J = 8.3 Hz, ⁴J = 1.2 Hz, 1H), 6.95 (td, ³J = 7.6 Hz, ⁴J = 1.3 Hz, 1H), 4.35 (s, 2H), 4.11 (q, J = 7.0 Hz, 2H), 3.62 (s, 3H), 3.04 (s, 3H), 1.37 (t, J = 7.0 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 170.15 (C=O), 155.35 (C, C-aromatic), 133.56, 130.38 (CH, C-aromatic), 127.63 (C, C-aromatic), 120.64, 113.48 (CH, C-aromatic), 64.13 (CH₃), 52.32(CH₃), 51.19 (CH₂), 40.07, 14.99 (CH₃).

N*-(methylsulfonyl)-*N*-phenylglycine (122)*(C₉H₁₁NO₄S; M.W.: 229.25)**

General procedure 6;

off-White solid;

T.L.C. System: DCM/MeOH 95:5 v/v, R_f: 0.21

Yield: 1.23g (84%)

Purity: 99%

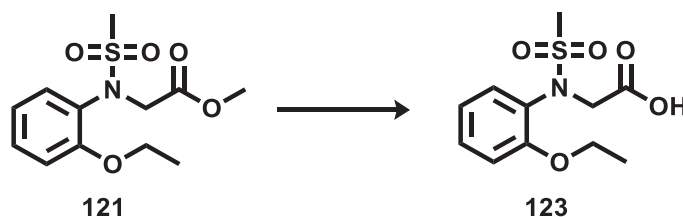
UPLC-MS method C: Rt: 1.287, MS (ESI)⁻: 228.2 [M-H]⁻

¹H NMR (DMSO-d₆), δ: 12.92 (bs, 1H), 7.49 – 7.39 (m, 4H), 7.35 – 7.30 (m, 1H), 4.41 (s, 2H), 3.09 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 171.11 (C=O), 140.84 (C, C-aromatic), 129.68, 127.83, 127.80 (CH, C-aromatic), 52.64 (CH₂), 39.41 (CH₃).

***N*-(2-ethoxyphenyl)-*N*-(methylsulfonyl)glycine (123)**

(C₁₁H₁₅NO₅S; M.W.: 273.30)



General procedure 6;

off-White solid;

T.L.C. System: DCM/MeOH 95:5 v/v, R_f: 0.21.

Yield: 1.46 g (99%)

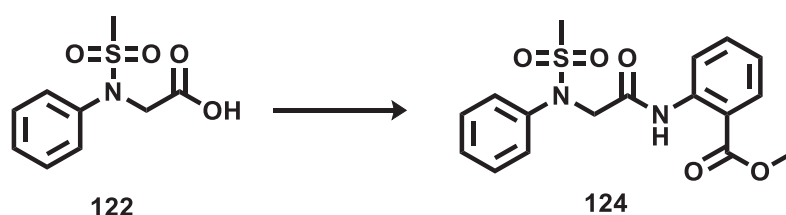
Purity: 99%

UPLC-MS method C: Rt: 1.484, MS (ESI)⁺: 274.0 [M+H]⁺, 296.0 [M+Na]⁺

MS (ESI)⁻: 272.1 [M-H]⁻

¹H NMR (DMSO-d₆), δ: 12.71 (s, 1H), 7.42 (dd, ³J = 7.7 Hz, ⁴J = 1.7 Hz, 1H), 7.33 (ddd, ³J = 8.3, 7.6 Hz, ⁴J = 1.7 Hz, 1H), 7.09 (dd, ³J = 8.3 Hz, ⁴J = 1.1 Hz, 1H), 6.94 (td, ³J = 7.6 Hz, ⁴J = 1.3 Hz, 1H), 4.24 (s, 2H), 4.11 (q, J = 7.0 Hz, 2H), 3.03 (s, 3H), 1.37 (t, J = 7.0 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 171.05 (C=O), 155.37 (C, C-aromatic), 133.70, 130.25 (CH, C-aromatic), 127.76 (C, C-aromatic), 120.59, 113.45 (CH, C-aromatic), 64.11, 51.22 (CH₂), 40.22, 15.00 (CH₃).

Methyl 2-(2-(*N*-phenylmethylsulfonamido)acetamido)benzoate (124)**(C₁₇H₁₈N₂O₅S; M.W.: 362.4)**

General procedure 7;

pale yellow powder;

T.L.C. System: n-Hex/EtOAc 7:3 v/v, Rf: 0.24.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge:

SNAP KP-Sil 25g (n-hexane -EtOAc 100:0 v/v increasing to 60:40 v/v in 28 CV).

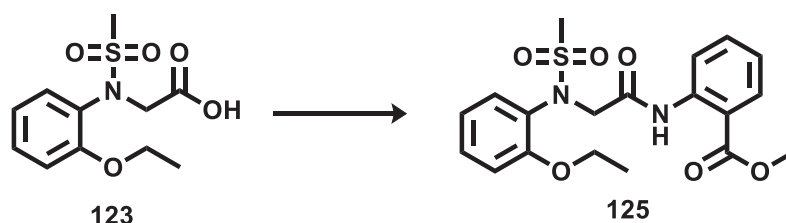
Yield: 249 mg (52%)

Purity: 99%

UPLC-MS method C: Rt: 1.815, MS (ESI)⁺: 363.3 [M+H]⁺, 385.2 [M+Na]⁺

¹H NMR (DMSO-d₆), δ: 11.31 (s, 1H), 8.42 (d, *J* = 8.4 Hz, 1H), 7.98 (dd, ³*J* = 8.0 Hz, ⁴*J* = 1.5 Hz, 1H), 7.65 – 7.58 (m, 3H), 7.46 (t, *J* = 7.8 Hz, 2H), 7.36 (t, *J* = 7.4 Hz, 1H), 7.21 (t, *J* = 7.6 Hz, 1H), 4.59 (s, 2H), 3.93 (s, 3H), 3.13 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 167.94 (C=O), 167.71 (C=O), 140.79, 139.79 (C, C-aromatic), 134.75, 131.20, 129.81, 128.37, 128.19, 123.94, 120.88 (CH, C-aromatic), 117.25 (C, C-aromatic), 55.47 (CH₂), 53.09, 37.48 (CH₃).

Methyl 2-(2-(*N*-(2-ethoxyphenyl)methylsulfonamido)acetamido)benzoate (125)**(C₁₉H₂₂N₂O₆S; M.W.: 406.45)**

General procedure 7;

yellow powder;

T.L.C. System: n-Hex/EtOAc 7:3 v/v, Rf: 0.25.

Yield: 1.75 g (96%)

Purity: 99%

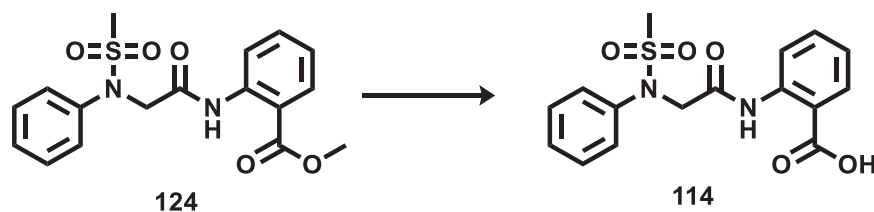
UPLC-MS method C: Rt: 1.886, MS (ESI)⁺: 407.3 [M+H]⁺, 429.3 [M+Na]⁺

¹H NMR (DMSO-d₆), δ: 11.28 (s, 1H), 8.45 (d, *J* = 8.4 Hz, 1H), 7.99 (d, *J* = 8.0 Hz, 1H), 7.67 (d, *J* = 7.8 Hz, 1H), 7.62 (t, *J* = 7.9 Hz, 1H), 7.35 (t, *J* = 7.9 Hz, 1H), 7.22 (t, *J* = 7.6 Hz, 1H), 7.14 (d, *J* = 8.3 Hz, 1H), 6.97 (t, *J* = 7.6 Hz, 1H), 4.43 (s, 2H), 4.16 (q, *J* = 6.9 Hz, 2H), 3.93 (s, 3H), 3.11 (s, 3H), 1.41 (t, *J* = 7.0 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 168.22 (C=O), 167.88 (C=O), 155.14 (C, C-aromatic), 139.86 (C, C-aromatic), 134.75, 133.39, 131.25, 130.57 (CH, C-aromatic), 127.89 (C, C-aromatic), 123.88, 120.91, 120.85 (CH, C-aromatic), 117.19 (C, C-aromatic), 113.73 (CH, C-aromatic), 64.23, 54.70 (CH₂), 53.13, 39.01, 14.98 (CH₃).

2-(2-(*N*-phenylmethylsulfonamido)acetamido)benzoic acid (114)

(C₁₆H₁₆N₂O₅S; M.W.: 348.37)



General procedure 8;

pale yellow powder;

T.L.C. System: DCM/MeOH 97:3 v/v, R_f: 0.25.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: SNAP KP-Sil 25g (DCM-MeOH 100:0 v/v increasing to 90:10 v/v in 16 CV).

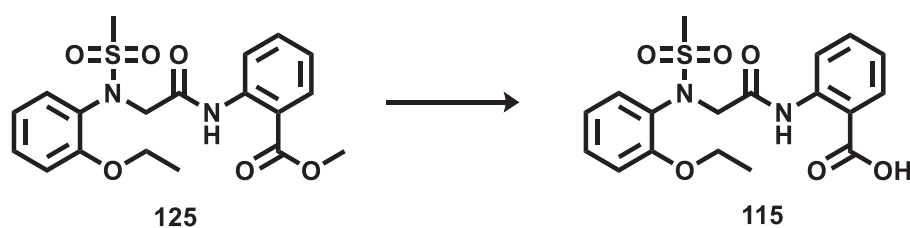
Yield: 71 mg (71%)

Purity: 99%

UPLC-MS method C: R_t: 1.663, MS (ESI)⁺: 349.2 [M+H]⁺, 371.2 [M+Na]⁺

¹H NMR (DMSO-d₆), δ: 12.05 (s, 1H), 8.59 (d, *J* = 8.5 Hz, 1H), 8.02 (dd, ³*J* = 8.0 Hz, ⁴*J* = 1.6 Hz, 1H), 7.69 – 7.64 (m, 2H), 7.59 (ddd, ³*J* = 8.7, 7.4 Hz, ⁴*J* = 1.7 Hz, 1H), 7.45 (t, *J* = 7.8 Hz, 2H), 7.38 – 7.33 (m, 1H), 7.17 (td, ³*J* = 8.0, ⁴*J* = 1.2 Hz, 1H), 4.56 (s, 2H), 3.10 (s, *J* = 4.2 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 170.07 (C=O), 167.78 (C=O), 140.88, 140.86 (C, C-aromatic), 134.75, 131.68, 129.78, 128.44, 128.34, 123.47, 119.86 (CH, C-aromatic), 116.54 (C, C-aromatic), 55.77 (CH₂), 36.90 (CH₃).

2-(2-(*N*-(2-ethoxyphenyl)methylsulfonamido)acetamido)benzoic acid (115 from 125)**(C₁₈H₂₀N₂O₆S; M.W.: 392.43)**

General procedure 8;

White powder;

T.L.C. System: DCM/MeOH 97:3 v/v, Rf: 0.28.

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge:

SNAP KP-Sil 25g (DCM-MeOH 100:0 v/v increasing to 90:10 v/v in 16 CV).

Yield: 76 mg (91%)

Purity: 99%

UPLC-MS method C: Rt: 1.758, MS (ESI)⁺: 393.3 [M+H]⁺, 415.3 [M+Na]⁺

¹H NMR (DMSO-d₆), δ: 13.90 (bs, 1H), 12.08 (s, 1H), 8.63 (d, *J* = 8.4 Hz, 1H), 8.04 (dd, ³*J* = 7.9 Hz, ⁴*J* = 1.6 Hz, 1H), 7.78 (dd, ³*J* = 7.9 Hz, ⁴*J* = 1.6 Hz, 1H), 7.60 (t, *J* = 7.9 Hz, 1H), 7.35 (t, *J* = 7.9 Hz, 1H), 7.18 (t, *J* = 7.6 Hz, 1H), 7.14 (d, *J* = 8.3 Hz, 1H), 6.94 (t, *J* = 7.6 Hz, 1H), 4.41 (s, 2H), 4.15 (q, *J* = 7.0 Hz, 2H), 3.08 (s, 3H), 1.41 (t, *J* = 7.0 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 170.08 (C=O), 168.42 (C=O), 155.04, 140.96 (C, C-aromatic), 134.74, 133.50, 131.67, 130.59 (CH, C-aromatic), 127.91 (C, C-aromatic), 123.43, 120.73, 119.92 (CH, C-aromatic), 116.52 (C, C-aromatic), 113.72 (CH, C-aromatic), 64.21, 54.97 (CH₂), 38.68, 14.98 (CH₃).

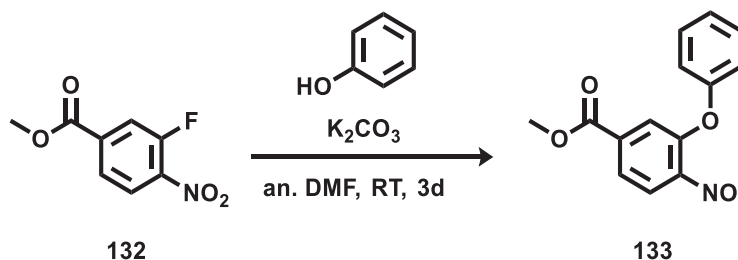
6.5 Synthesis of the α -helix mimic 126

6.5.1 Synthesis of α -helix mimic building blocks

Methyl 4-amino-3-phenoxybenzoate (127)

Methyl 4-nitro-3-phenoxybenzoate (133)

(C₁₄H₁₁NO₅; M.W.: 273.24)



A mixture of methyl 3-fluoro-4-nitrobenzoate **132** (1.25 mmol), phenol (2.5 mmol), Potassium Carbonate (2.5 mmol) in anhydrous DMF (3 ml) was stirred at room temperature for three days. The reaction mixture was diluted with water (7 mL) and was extracted with EtOAc (3 x 10 mL). The organic layers were combined and then dried over MgSO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 25g, isocratic method with DCM in 11 CV) to give pure methyl 4-nitro-3-phenoxybenzoate **133** as a yellow oil.

T.L.C. System: DCM, R_f: 0.35.

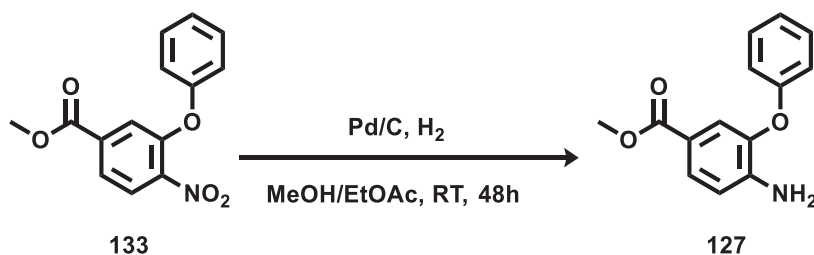
Yield: 1.23g (87%)

Purity: 99%

UPLC-MS method C: Rt: 2.003, MS (ESI)⁺: 274.2[M+1]⁺

¹H NMR (DMSO-d₆), δ : 8.20 (d, J = 8.5 Hz, 1H), 7.86 (dd, 3J = 8.4 Hz, 4J = 1.7 Hz, 1H), 7.53 – 7.46 (m, 3H), 7.30 (t, J = 7.4 Hz, 1H), 7.19 (dd, 3J = 8.6 Hz, 4J = 0.9 Hz, 2H), 3.84 (s, 3H).

¹³C NMR (DMSO-d₆), δ 164.66 (C=O), 155.34, 149.97, 144.00, 135.17 (C, C-aromatic), 131.03, 126.63, 125.76, 124.69, 120.49, 119.80 (CH, C-aromatic), 53.40 (CH₃).

Methyl 4-amino-3-phenoxybenzoate (127)**(C₁₄H₁₃NO₃; M.W.: 243.26)**

A solution of methyl 4-nitro-3-phenoxybenzoate **133** (0.9 mmol) in MeOH/EtOAc (ratio 3:1, 12 ml) was stirred under H₂ atmosphere in presence of Pd/C (15%) at RT for 48 hours. The reaction mixture was filtered on Celite. The filtered solution was concentrated under reduced pressure and the obtained residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 25g, n-hexane -EtOAc 100:0 v/v increasing to 40:60 v/v in 11 CV) to give pure Methyl 4-amino-3-phenoxybenzoate **127** as a colourless oil.

T.L.C. System: n-hexane -EtOAc 7:3 v/v, R_f: 0.35.

Yield: 290 mg (99%)

Purity: 99%

UPLC-MS method C: Rt: 1.845, MS (ESI)⁺: 244.1[M+1]⁺

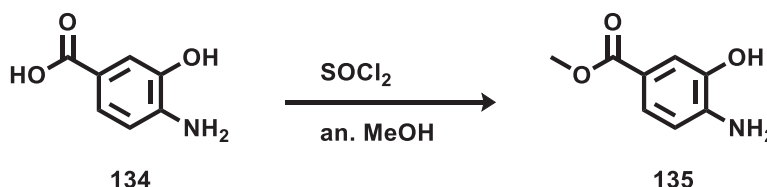
¹H NMR (DMSO-d₆), δ: 7.55 (dd, ³J = 8.4 Hz, ⁴J = 1.9 Hz, 1H), 7.40 – 7.35 (m, 2H), 7.26 (d, ⁴J = 1.9 Hz, 1H), 7.13 – 7.09 (m, 1H), 6.99 – 6.95 (m, 2H), 6.83 (d, J = 8.4 Hz, 1H), 5.91 (s, 2H), 3.71 (s, 3H).

¹³C NMR (DMSO-d₆), δ 166.32 (C=O), 157.43, 145.98, 141.65 (C, C-aromatic), 130.37, 127.30, 123.44, 120.70, 118.00 (CH, C-aromatic), 116.75 (C, C-aromatic), 114.74 (CH, C-aromatic), 51.86 (CH₃).

Boc protected 3-(alkyloxy)-4-(amino)benzoic acid (128,130)

Methyl 4-amino-3-hydroxybenzoate (135)

(C₈H₉NO₃; M.W.: 167.16)



To a solution of 4-amino-3-hydroxybenzoic acid **134** (6.5 mmol) in anhydrous methanol (20 mL), thionyl chloride (19.5 mmol) was added drop-wise at 0°C. The reaction mixture was stirred at room temperature for 24 hours. The mixture was concentrated under reduced pressure and the obtained residue was then diluted with a sat. NaHCO₃ solution (60 mL) and extracted with EtOAc (3x80 mL). The organic layers were combined, washed with Brine and then dried over MgSO₄ and concentrated under reduced pressure to give methyl 4-amino-3-hydroxybenzoate **135**, that were used for the following step without further purification.

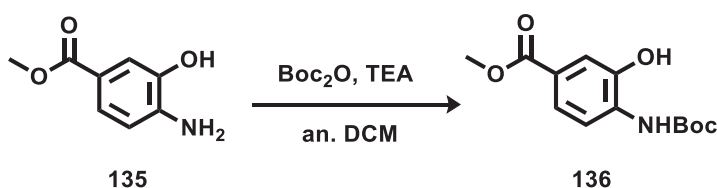
T.L.C. System: n-hexane -EtOAc 7:3 v/v, R_f: 0.7.

Yield: 1.13 g (99%)

¹H NMR (DMSO-d₆), δ: 9.41 (s, 1H), 7.34 – 7.17 (m, 2H), 6.68 – 6.50 (m, 1H), 5.40 (s, 2H), 3.73 (s, 3H).

Boc protected methyl 4-(amino)-3-hydroxybenzoate (136)

(C₁₃H₁₇NO₅; M.W.: 267.28)



To a solution of methyl 4-amino-3-hydroxybenzoate **135** (1.79 mmol) and triethylamine (1.97 mmol) in anhydrous DCM (6 mL), a solution of di-tert-butyl decarbonate (1.97 mmol) in anhydrous DCM (1.2 mL) was added drop-wise. The reaction mixture was stirred at room temperature for 24 hours. To the reaction mixture water (7 mL) was added and extracted with DCM (3x8 mL). The organic layers were combined, washed with Brine and then dried over MgSO₄ and concentrated under reduced pressure. The obtained residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 25g,

n-hexane -EtOAc 95:5 v/v increasing to 40:60 v/v in 11 CV) to give pure Boc protected methyl 4-(amino)-3-hydroxybenzoate **136** as a white solid.

T.L.C. System: n-hexane -EtOAc 7:3 v/v, Rf: 0.45.

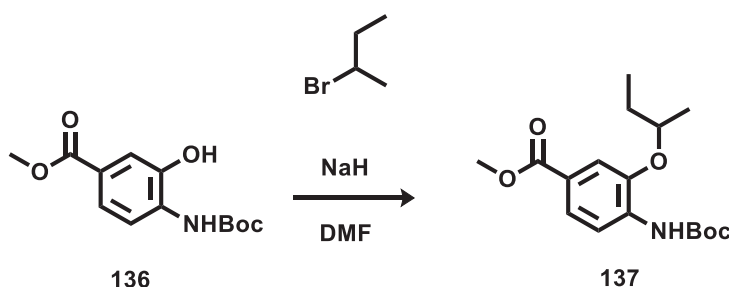
Yield: 486 mg (99%)

Purity: >95%

¹H NMR (CDCl₃-d), δ : 7.81 (d, ⁴J = 1.8 Hz, 1H), 7.77 (dd, ³J = 8.4 Hz, ⁴J = 1.8 Hz, 1H), 6.80 – 6.75 (m, 1H), 4.19 (s, 2H), 3.88 (s, 3H), 1.59 (s, 9H).

Boc protected methyl 3-(*sec*-butoxy)-4-(amino)benzoate (**137**)

(C₁₇H₂₅NO₅; M.W.: 323.39)

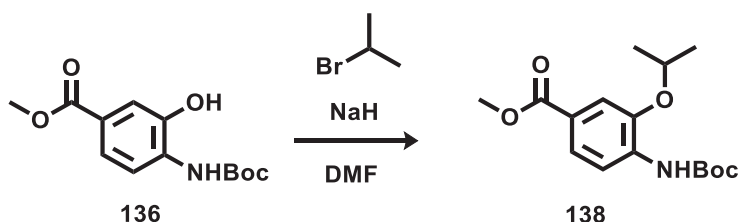


A solution of Boc protected methyl 4-(amino)-3-hydroxybenzoate **136** (0.75 mmol) and sodium hydride (0.82 mmol) in anhydrous DMF (10 mL) was stirred at room temperature for 15 minutes, then 2-bromobutane (1.8 mmol) was added drop-wise and the reaction mixture was stirred at room temperature for 19 hours. The reaction mixture was diluted with water (15 mL) and then extracted with EtOAc (3x20 mL). The organic layers were combined, washed with Brine and then dried over MgSO₄ and concentrated under reduced pressure. The obtained residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 25g, n-hexane -EtOAc 95:5 v/v increasing to 60:40 v/v in 11 CV) to give pure Boc protected methyl 3-(*sec*-butoxy)-4-(amino)benzoate **137** as a white solid.

T.L.C. System: n-hexane -EtOAc 7:3 v/v, Rf: 0.56.

Yield: 163 mg (67%)

¹H NMR (DMSO-d₆), δ : 8.00 (d, *J* = 8.5 Hz, 1H), 7.95 (s, 1H), 7.54 (dd, ³*J* = 8.5 Hz, ⁴*J* = 1.7 Hz, 1H), 7.48 (d, ⁴*J* = 1.6 Hz, 1H), 4.50 (h, *J* = 6.0 Hz, 1H), 3.84 (d, *J* = 16.2 Hz, 3H), 1.83 – 1.69 (m, 1H), 1.69 – 1.57 (m, 1H), 1.49 (s, 9H), 1.26 (d, *J* = 6.0 Hz, 3H), 0.93 (t, *J* = 7.4 Hz, 3H).

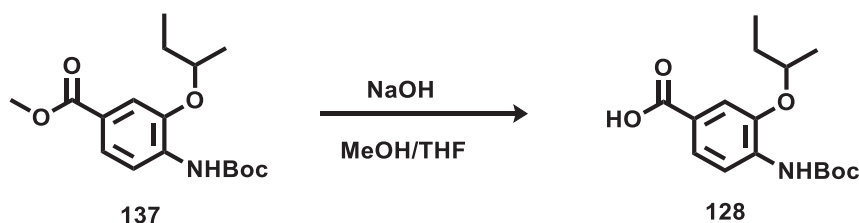
Boc protected methyl 3-(*iso*-propoxy)-4-(amino)benzoate (138)**(C₁₆H₂₃NO₅; M.W.: 309.36)**

A solution of Boc protected methyl 4-(amino)-3-hydroxybenzoate **136** (0.75 mmol) and sodium hydride (0.82 mmol) in anhydrous DMF (10 mL) was stirred at room temperature for 15 minutes, then 2-bromopropane (1.8 mmol) was added drop-wise and the reaction mixture was stirred at room temperature for 19 hours. The reaction mixture was diluted with water (15 mL) and then extracted with EtOAc (3x20 mL). The organic layers were combined, washed with Brine and then dried over MgSO₄ and concentrated under reduced pressure. The obtained residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 25g, n-hexane -EtOAc 95:5 v/v increasing to 60:40 v/v in 11 CV) to give pure Boc protected methyl 3-(*iso*-propoxy)-4-(amino)benzoate **138** as a white solid.

T.L.C. System: n-hexane -EtOAc 7:3 v/v, R_f: 0.56.

Yield: 162 mg (70%)

¹H NMR (CDCl₃-d), δ: 8.20 (d, *J* = 8.4 Hz, 1H), 7.66 (dd, ³*J* = 8.5 Hz, ⁴*J* = 1.6 Hz, 1H), 7.54 (d, ⁴*J* = 1.6 Hz, 1H), 4.72 (hept, *J* = 6.0 Hz, 1H), 3.91 (s, 3H), 1.57 (s, 10H), 1.42 (s, 3H), 1.41 (s, 3H).

Boc protected 3-(*sec*-butoxy)-4-(amino)benzoic acid (128)**(C₁₆H₂₃NO₅; M.W.: 309.36)**

To a solution of Boc protected methyl 3-(*sec*-butoxy)-4-(amino)benzoate **137** (0.18 mmol) in tetrahydrofuran and methanol (ratio 2:1) was added a solution 1N of NaOH (0.3 mL). The reaction mixture was stirred at 60°C for 4 hours. The reaction mixture was concentrated under reduced pressure and the residue was dissolved in water (5 mL). The aqueous layer was acidified with HCl 2N and then extracted with EtOAc (3x8mL). The organic layers were combined, washed with Brine and then dried over MgSO₄ and concentrated under reduced

pressure to give Boc protected 3-(sec-butoxy)-4-(amino)benzoic acid **128**, that were used for the following step without further purification.

T.L.C. System: n-hexane -EtOAc 7:3 v/v, Rf: 0.23.

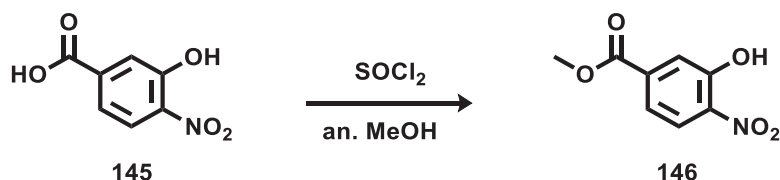
Yield: 39 mg (88%)

¹H NMR (DMSO-d₆), δ : 12.81 (bs, 1H), 7.95 (d, J = 8.4 Hz, 1H), 7.92 (s, 1H), 7.52 (dd, 3J = 8.4 Hz, 4J = 1.7 Hz, 1H), 7.47 (d, 4J = 1.6 Hz, 1H), 4.48 (h, J = 6.1 Hz, 1H), 1.82 – 1.69 (m, 1H), 1.69 – 1.58 (m, 1H), 1.49 (s, 9H), 1.27 (d, J = 6.0 Hz, 3H), 0.94 (t, J = 7.4 Hz, 3H).

3-(alkoxy)-4-nitrobenzoic acids (141,142)

methyl 3-hydroxy-4-nitrobenzoate (146)

(C₈H₇NO₅; M.W.: 197.15)



To a solution of methyl 3-hydroxy-4-nitrobenzoate **145** (3.46 mmol) in anhydrous methanol (10 mL), thionyl chloride (16.39 mmol) was added drop-wise at 0°C. The reaction mixture was stirred at room temperature for 24 hours. The mixture was concentrated under reduced pressure and the obtained residue was then diluted with a sat. NaHCO₃ solution (60 mL) and extracted with EtOAc (3x80 mL). The organic layers were combined, washed with Brine, dried over MgSO₄ and concentrated under reduced pressure to give methyl 3-hydroxy-4-nitrobenzoate **146** as a yellow solid, which was used without further purification.

T.L.C. System: n-Hex/EtOAc 8:2 0,44

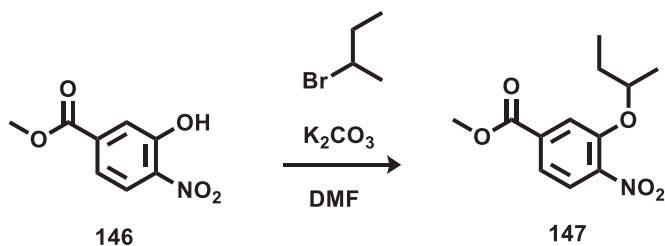
Yield: 1.07 g (97%)

Purity: 99%

UPLC-MS method C: Rt: 1.691, MS (ESI)⁺: -

¹H NMR (DMSO-d₆), δ : 11.46 (s, 1H), 7.94 (d, J = 8.5 Hz, 1H), 7.67 (d, J = 1.7 Hz, 1H), 7.47 (dd, 3J = 8.5 Hz, 4J = 1.7 Hz, 1H), 3.87 (s, 3H).

¹³C NMR (DMSO-d₆), δ : 165.20 (C=O), 151.78, 140.59, 134.97 (C, C-aromatic), 125.96, 119.87 (CH, C-aromatic), 53.19 (CH₃).

methyl 3-(sec-butoxy)-4-nitrobenzoate (147)**(C₁₂H₁₅NO₅; M.W.: 253.25)**

A solution of methyl 3-hydroxy-4-nitrobenzoate **146** (5.07 mmol) and potassium carbonate (8.12 mmol) in anhydrous DMF (0.7 M) was stirred at room temperature for 15 minutes, then 2-bromobutane (**107**) (6.09 mmol) was added drop-wise and the reaction mixture was stirred overnight at 70°C. The reaction mixture was diluted with EtOAc (20 mL) and then extracted with HCl 1N solution (4x20 mL). The organic layer was then washed with Brine, dried over MgSO₄ and concentrated under reduced pressure to give methyl 3-(sec-butoxy)-4-nitrobenzoate **147** as a yellow/orange solid, which was used without further purification. T.L.C. System: n-hexane -EtOAc 8:2 v/v, R_f: 0.46.

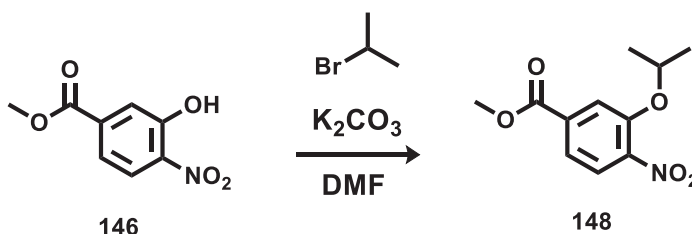
Yield: 1.25 g (97%)

Purity: 99%

UPLC-MS method C: Rt: 2.024, MS (ESI)⁺: 195[C₁₀H₁₂NO₃ + 1]⁺, 224[C₁₀H₁₀NO₅ + 1]⁺

¹H NMR (DMSO-d₆), δ: 7.95 (d, *J* = 8.3 Hz, 1H), 7.77 (d, *J* = 1.6 Hz, 1H), 7.63 (dd, ³*J* = 8.3 Hz, ⁴*J* 1.6 Hz, 1H), 4.73 (h, *J* = 6.0 Hz, 1H), 3.91 (s, 3H), 1.71 – 1.60 (m, 2H), 1.27 (d, *J* = 6.1 Hz, 3H), 0.92 (t, *J* = 7.4 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 165.28 (C=O), 150.33, 143.90, 134.53 (C, C-aromatic), 125.47, 121.54, 116.69 (CH, C-aromatic), 77.52 (CH), 53.31 (CH₃), 28.81 (CH₂), 19.09, 9.56 (CH₃).

methyl 3-isopropoxy-4-nitrobenzoate (148)**(C₁₁H₁₃NO₅; M.W.: 239.23)**

A solution of methyl 3-hydroxy-4-nitrobenzoate **146** (5.07 mmol) and potassium carbonate (8.12 mmol) in anhydrous DMF (0.7 M) was stirred at room temperature for 15 minutes, then 2-bromopropane (**101**) (6.09 mmol) was added drop-wise and the reaction mixture was stirred overnight at 70°C. The reaction mixture was diluted with EtOAc (20 mL) and then

washed with HCl 1N solution (4x20 mL). The organic layer was then washed with Brine, dried over MgSO₄ and concentrated under reduced pressure to give methyl 3-isopropoxy-4-nitrobenzoate **148** as a yellow solid, which was used without further purification.

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.46.

Yield: 1,02 g (84%)

Purity: 99%

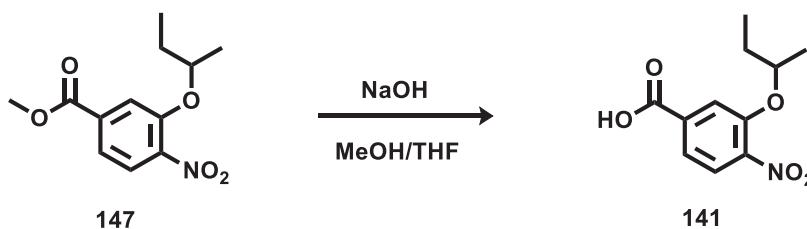
UPLC-MS method C: Rt: 1.946, MS (ESI)⁺: 181.1[C₉H₁₀NO₃ +1]⁺

¹H NMR (DMSO-d₆), δ: 7.94 (d, *J* = 8.3 Hz, 1H), 7.78 (d, ⁴*J* = 1.6 Hz, 1H), 7.63 (dd, ³*J* = 8.3 Hz, ⁴*J* = 1.6 Hz, 1H), 4.91 (hept, *J* = 6.0 Hz, 1H), 3.90 (s, 3H), 1.30 (d, *J* = 6.0 Hz, 6H).

¹³C NMR (DMSO-d₆), δ: 165.26 (C=O), 150.08, 143.98, 134.52 (C, C-aromatic), 125.44, 121.64, 116.96 (CH, C-aromatic), 73.06 (CH), 53.28, 21.92 (CH₃).

3-(sec-butoxy)-4-nitrobenzoic acid (**141**)

(C₁₁H₁₃NO₅; M.W.: 239.23)



To a solution of methyl 3-(sec-butoxy)-4-nitrobenzoate **147** (4.41 mmol) in tetrahydrofuran and methanol (0.12 M, ratio 2:1) was added a solution 1N of NaOH (8.82 mL). The reaction mixture was stirred at room temperature for 2 hours. The reaction mixture was concentrated under reduced pressure and the residue was dissolved in water (10 mL). The aqueous layer was acidified with HCl 2N and then extracted with DCM (3x15mL). The organic layers were combined, washed with Brine, dried over MgSO₄ and concentrated under reduced pressure to give 3-(sec-butoxy)-4-nitrobenzoic acid **141**, as a yellow powder which was used without further purification.

T.L.C. System: DCM/MeOH 95:5 v/v, Rf: 0.3.

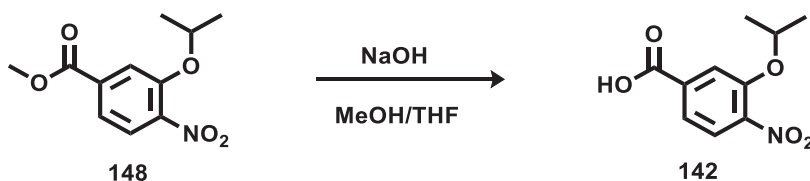
Yield: 1.12 g (97%)

Purity: 99%

UPLC-MS method C: Rt: 1.807, MS (ESI)⁺: 262.8[M+Na]⁺

¹H NMR (DMSO-d₆), δ: 13.60 (bs, 1H), 7.92 (d, *J* = 8.3 Hz, 1H), 7.75 (d, ⁴*J* = 1.5 Hz, 1H), 7.61 (dd, ³*J* = 8.3 Hz, ⁴*J* = 1.5 Hz, 1H), 4.77 – 4.65 (m, 1H), 1.73 – 1.58 (m, 2H), 1.27 (d, *J* = 6.1 Hz, 3H), 0.92 (t, *J* = 7.4 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 166.29 (C=O), 150.31, 143.65, 135.87 (C, C-aromatic), 125.33, 121.59, 116.74 (CH, C-aromatic), 77.42 (CH), 28.83 (CH₂), 19.11, 9.58 (CH₃).

3-isopropoxy-4-nitrobenzoic acid (142)**(C₁₀H₁₁NO₅; M.W.: 225.2)**

To a solution of methyl 3-isopropoxy-4-nitrobenzoate **148** (4.3 mmol) in tetrahydrofuran and methanol (0.12 M, ratio 2:1) was added a solution 1N of NaOH (8.6 mL). The reaction mixture was stirred at room temperature for 2 hours. The reaction mixture was concentrated under reduced pressure and the residue was dissolved in water (10 mL). The aqueous layer was acidified with HCl 2N and then extracted with DCM (3x15mL). The organic layers were combined, washed with Brine, dried over MgSO₄ and concentrated under reduced pressure to give 3-isopropoxy-4-nitrobenzoic acid **142**, as a yellow solid which was used without further purification.

T.L.C. System: DCM/MeOH 95:5 v/v, Rf: 0.3.

Yield: 840 mg (86%)

Purity: 99%

UPLC-MS method C: Rt: 1.723, MS (ESI)⁺: 224.1[M-1]-

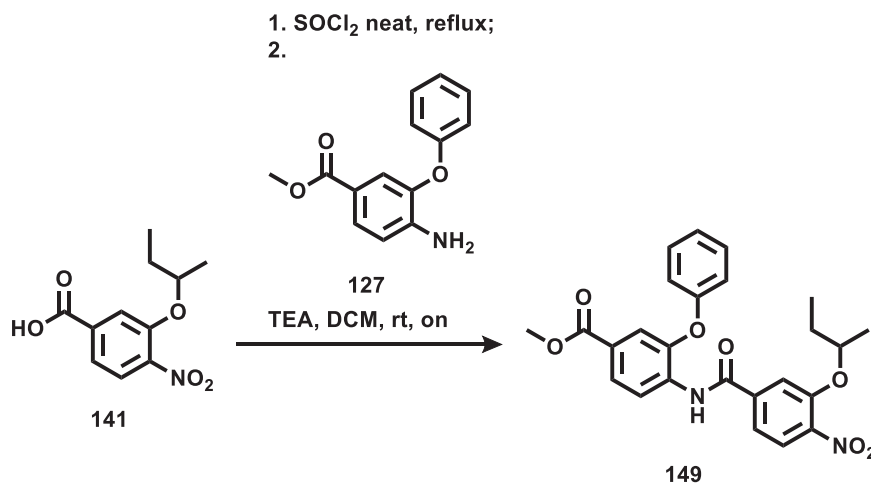
¹H NMR (DMSO-d₆), δ: 13.60 (bs, 1H), 7.91 (d, *J* = 8.3 Hz, 1H), 7.76 (d, ⁴*J* = 1.6 Hz, 1H), 7.62 (dd, ³*J* = 8.3 Hz, ⁴*J* = 1.6 Hz, 1H), 4.90 (hept, *J* = 6.0 Hz, 1H), 1.30 (d, *J* = 6.0 Hz, 6H).

¹³C NMR (DMSO-d₆), δ: 166.29 (C=O), 150.07, 143.73, 135.88 (C, C-aromatic), 125.31, 121.70, 117.01 (CH, C-aromatic), 72.92 (CH), 21.95 (CH₃).

6.5.2 Synthesis of (4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycine (**150**)

methyl 4-(3-(sec-butoxy)-4-nitrobenzamido)-3-phenoxybenzoate (**149**)

(C₂₅H₂₄N₂O₇; M.W.: 464.47)



3-(sec-butoxy)-4-nitrobenzoic acid **141** (1.78 mmol) in SOCl₂ (1.5 ml) was refluxed for 2 hours. SOCl₂ was eliminated from the reaction mixture under reduced pressure. The resulting residue was dissolved in anhydrous DCM (3 ml) and added drop-wise to a solution, at 0°C, of methyl 4-amino-3-phenoxybenzoate **127** (2.14 mmol), triethylamine (3.57 mmol) in anhydrous DCM (10 ml) under nitrogen atmosphere. The reaction was allowed to warm-up to room temperature and was stirred for 3h. The reaction mixture was diluted with DCM and extracted with 1N HCl solution (3x15ml). The organic layer was washed with brine (3 x 30 mL) and dried over MgSO₄. The organic solvent was evaporated at reduced pressure and the crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 5g, n-hexane -EtOAc 100:0 v/v increasing to 80:20 v/v in 16 CV) to give methyl 4-(3-(sec-butoxy)-4-nitrobenzamido)-3-phenoxybenzoate **149** as white powder.

T.L.C. System: n-Hex/EtOAc 8:2 v/v, Rf: 0.4

Yield: 828 mg

Purity: 80%

UPLC-MS method C: Rt: 2.22, MS (ESI)⁺: 465.3[M+1]⁺/244.1[M+1]⁺, 463.2[M-1]⁻

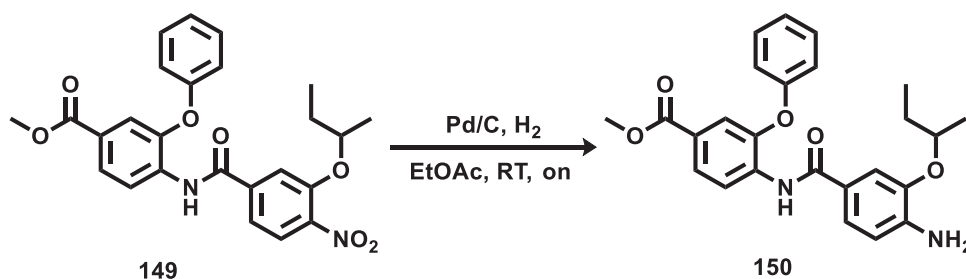
¹H NMR (DMSO-d₆), δ:

Major species: 10.33 (s, 1H), 8.04 (d, *J* = 8.4 Hz, 1H), 7.92 (d, *J* = 8.3 Hz, 1H), 7.83 (dd, ³*J* = 8.4 Hz, ⁴*J* = 1.9 Hz, 1H), 7.62 (d, ⁴*J* = 1.3 Hz, 1H), 7.48 (td, ³*J* = 4.0 Hz, ⁴*J* = 1.6 Hz, 2H), 7.44 – 7.38 (m, 2H), 7.20 – 7.15 (m, 1H), 7.06 (dd, ³*J* = 8.6 Hz, ⁴*J* = 0.9 Hz, 2H), 4.65 (h, *J* = 6.0 Hz, 1H), 3.82 (s, 3H), 1.72 – 1.59 (m, 2H), 1.26 (d, *J* = 6.1 Hz, 3H), 0.91 (t, *J* = 7.4 Hz, 3H).

Minor species (methyl 4-amino-3-phenoxybenzoate): 7.55 (dd, $^3J = 8.4$ Hz, $^4J = 1.9$ Hz, 1H), 7.42 – 7.33 (m, 2H), 7.26 (d, $^4J = 1.9$ Hz, 1H), 7.16 – 7.05 (m, 1H), 7.01 – 6.92 (m, 2H), 6.83 (d, $J = 8.4$ Hz, 1H), 5.91 (s, 2H), 3.71 (s, 3H).

Methyl 4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoate (150)

($C_{25}H_{26}N_2O_5$; M.W.: 434.49)



A solution of methyl methyl 4-(3-(sec-butoxy)-4-nitrobenzamido)-3-phenoxybenzoate **149** (1.44 mmol) in anhydrous EtOAc (0.5 M) was stirred under H_2 atmosphere in presence of Pd/C (15%) at room temperature for overnight. The reaction mixture was filtered on Celite. The filtered solution was concentrated under reduced pressure and the obtained residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 80g, n-hexane-EtOAc 100:0 v/v increasing to 80:20 v/v in 16 CV) to give pure methyl 4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoate **150** as a white solid.

T.L.C. System: n-Hex/EtOAc 8:2 v/v, Rf: 0,25

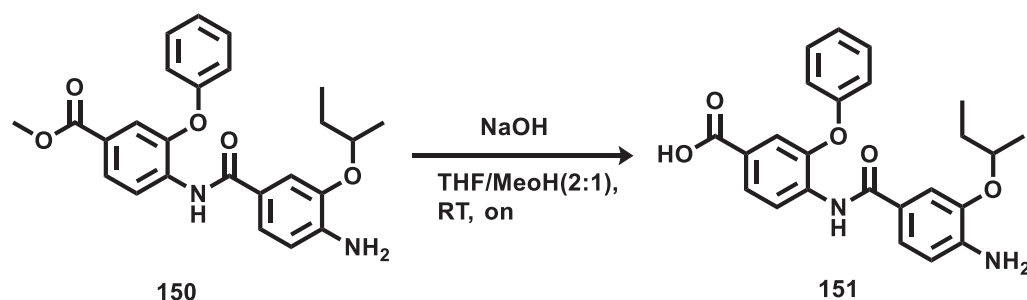
Yield: 538 mg (86% over two steps)

Purity: 99%

UPLC-MS method C: Rt: 2.086, MS (ESI)⁺: 435.3[M+1]⁺

1H NMR (DMSO- d_6), δ : 9.44 (s, 1H), 8.18 (d, $J = 8.5$ Hz, 1H), 7.79 (dd, $^3J = 8.5$ Hz, $^4J = 1.9$ Hz, 1H), 7.45 – 7.38 (m, 3H), 7.32 – 7.26 (m, 2H), 7.20 – 7.15 (m, 1H), 7.11 – 7.07 (m, 2H), 6.65 (d, $J = 8.2$ Hz, 1H), 5.39 (s, 2H), 4.31 (h, $J = 6.0$ Hz, 1H), 3.80 (s, 3H), 1.74 – 1.64 (m, 1H), 1.64 – 1.54 (m, 1H), 1.22 (d, $J = 6.1$ Hz, 3H), 0.93 (t, $J = 7.4$ Hz, 3H).

^{13}C NMR (DMSO- d_6), δ : 165.84 (C=O), 165.57 (C=O), 156.47, 148.12, 143.79, 143.75, 135.31 (C, C-aromatic), 130.60 (CH, C-aromatic), 125.96 (C, C-aromatic), 125.35, 124.46, 124.14, 122.26 (CH, C-aromatic), 121.02 (C, C-aromatic), 119.18, 119.02, 113.24, 112.88 (CH, C-aromatic), 75.67 (CH), 52.62 (CH₃), 29.03 (CH₂), 19.56, 10.06 (CH₃).

4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoic acid (151)**(C₂₄H₂₄N₂O₅; M.W.: 420.46)**

To a solution of methyl 4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoate **150** (1.0 mmol) in tetrahydrofuran and methanol (0.03 M, ratio 2:1) was added a solution 1N of NaOH (1.1 mL). The reaction mixture was stirred overnight at room temperature. The reaction mixture was concentrated under reduced pressure and the residue was dissolved in water (10 mL). The aqueous layer was acidified with HCl 2N and then extracted with DCM (3x15mL). The organic layers were combined, washed with Brine, dried over MgSO₄ and concentrated under reduced pressure to give 4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoic acid **151**, as a pale-yellow solid which was used without further purification. T.L.C. System: DCM/MeOH 97:3 v/v, Rf:0.3

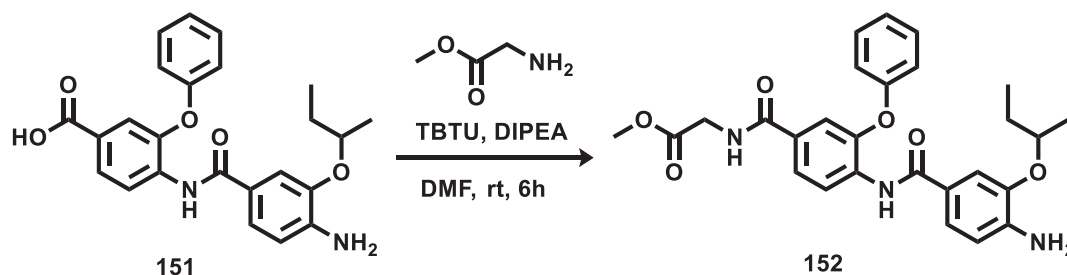
Yield: 459 mg (99%)

Purity: 99%

UPLC-MS method C: Rt: 1.848, MS (ESI)⁺: 421.3[M+1]⁺

¹H NMR (DMSO-d₆), δ : 12.92 (bs, 1H), 9.41 (s, 1H), 8.13 (d, $J = 8.4$ Hz, 1H), 7.76 (dd, $^3J = 8.4$ Hz, $^4J = 1.9$ Hz, 1H), 7.45 – 7.38 (m, 3H), 7.32 – 7.26 (m, 2H), 7.20 – 7.15 (m, 1H), 7.12 – 7.05 (m, 2H), 6.64 (d, $J = 8.1$ Hz, 1H), 5.37 (s, 2H), 4.32 (h, $J = 6.0$ Hz, 1H), 1.73 – 1.64 (m, 1H), 1.64 – 1.54 (m, 1H), 1.22 (d, $J = 6.1$ Hz, 3H), 0.93 (t, $J = 7.4$ Hz, 3H).

¹³C NMR (DMSO-d₆), δ : 166.93 (C=O), 165.54 (C=O), 156.51, 148.20, 143.79, 143.69, 134.81 (C, C-aromatic), 130.57 (CH, C-aromatic), 127.32 (C, C-aromatic), 125.38, 124.41, 124.12, 122.22 (CH, C-aromatic), 121.10 (C, C-aromatic), 119.31, 119.08, 113.24, 112.89 (CH, C-aromatic), 75.67 (CH), 29.04 (CH₂), 19.57, 10.07 (CH₃).

Methyl (4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycinate (152)**(C₂₇H₂₉N₃O₆; M.W.: 491.54)**

A mixture of 4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoic acid **151** (0.92 mmol), methyl glycinate (1.1 mmol), TBTU (1.1 mmol), and DIPEA (3.67 mmol) in anhydrous DMF (0.2 M) was stirred at room temperature for 6 hours. The reaction mixture was diluted in EtOAc (20 mL) and was washed with HCl 1N solution (3 x 20 mL), sat. NaHCO₃ solution (3 x 20 mL) and brine (3 x 20 mL). The organic layers were then dried over MgSO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP Sil 25g, n-hexane - DCM 100:0 v/v increasing to 0:100 v/v in 5 CV and DCM – MeOH 100:0 v/v increasing to 97:3 v/v in 13 CV) to give pure (methyl (4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycinate **152** as yellow/orange solid.

T.L.C. System: DCM/MeOH 97:3 v/v, Rf: 0.37

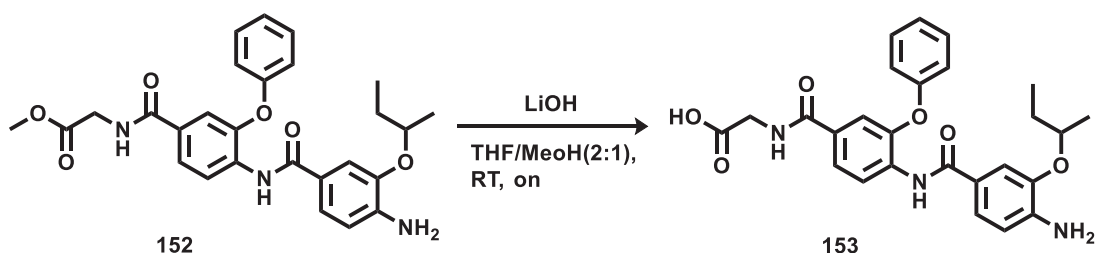
Yield: 256 mg (57%)

Purity: 96%

UPLC-MS method C: Rt: 1.816, MS (ESI)⁺: 492.2[M+1]⁺

¹H NMR (DMSO-d₆), δ: 9.39 (s, 1H), 8.95 (t, *J* = 5.8 Hz, 1H), 8.07 (d, *J* = 8.4 Hz, 1H), 7.73 (dd, ³*J* = 8.5 Hz, ⁴*J* = 2.0 Hz, 1H), 7.50 (d, ⁴*J* = 2.0 Hz, 1H), 7.42 – 7.35 (m, 2H), 7.30 – 7.24 (m, 2H), 7.16 – 7.11 (m, 1H), 7.07 – 7.01 (m, 2H), 6.63 (d, *J* = 8.1 Hz, 1H), 5.35 (s, 2H), 4.35 – 4.26 (m, 1H), 3.98 (d, *J* = 5.8 Hz, 2H), 3.65 (s, 3H), 1.75 – 1.64 (m, 1H), 1.64 – 1.53 (m, 1H), 1.22 (d, *J* = 6.1 Hz, 3H), 0.93 (t, *J* = 7.4 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 170.83 (C=O), 165.93 (C=O), 165.54 (C=O), 156.90, 148.09, 143.76, 143.57, 133.92 (C, C-aromatic), 130.48 (CH, C-aromatic), 130.43 (C, C-aromatic), 124.54, 124.02, 123.37, 122.19 (CH, C-aromatic), 121.16 (C, C-aromatic), 118.62, 118.48, 113.19, 112.89 (CH, C-aromatic), 75.64 (CH), 52.21 (CH₃), 41.69, 29.04 (CH₂), 19.57, 10.07 (CH₃).

(4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycine (153)**(C₂₆H₂₇N₃O₆; M.W.: 477.52)**

(methyl (4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycinate **152** (0.039 mmol) was solubilised in tetrahydrofuran and methanol (0.025 M, ratio 2:1) and a solution 1N of LiOH (0.04 mL) was added. The reaction mixture was stirred overnight at room temperature. The reaction mixture was concentrated under reduced pressure and the residue was dissolved in water (10 mL). The aqueous layer was acidified with HCl 2N and then extracted with DCM (3x15mL). The organic layers were combined, washed with Brine, dried over MgSO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP Sil 5g, n-hexane - DCM 100:0 v/v increasing to 0:100 v/v in 5 CV and DCM – MeOH 100:0 v/v increasing to 9:1 v/v in 11 CV) to give pure (4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycine **153** as yellow powder.

T.L.C. System: DCM/MeOH 9:1 v/v. Rf 0.23

Yield: 13.6 mg (73%)

Purity: 97%

UPLC-MS method C: Rt: 1.723, MS (ESI)⁺: 478.3[M+1]⁺

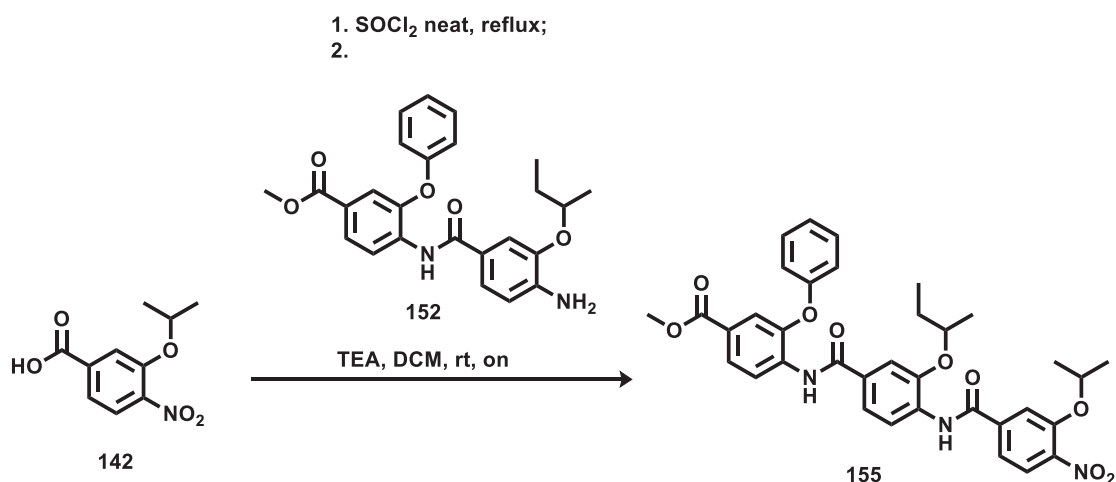
¹H NMR (DMSO-d₆), δ: 12.58 (bs, 1H), 9.38 (s, 1H), 8.83 (t, *J* = 5.6 Hz, 1H), 8.06 (d, *J* = 8.4 Hz, 1H), 7.73 (d, *J* = 8.4 Hz, 1H), 7.50 (s, *J* = 1.1 Hz, 1H), 7.38 (t, *J* = 7.8 Hz, 2H), 7.31 – 7.22 (m, 2H), 7.13 (t, *J* = 7.3 Hz, 1H), 7.04 (d, *J* = 7.9 Hz, 2H), 6.63 (d, *J* = 8.0 Hz, 1H), 5.34 (s, 2H), 4.30 (m, 1H), 3.88 (d, *J* = 5.6 Hz, 2H), 1.75 – 1.64 (m, 1H), 1.64 – 1.53 (m, 1H), 1.22 (d, *J* = 6.0 Hz, 3H), 0.93 (t, *J* = 7.4 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 171.76 (C=O), 165.78 (C=O), 165.54 (C=O), 156.94, 148.06, 143.76, 143.55, 133.82, 130.72 (C, C-aromatic), 130.47, 124.54, 123.98, 123.36, 122.18 (CH, C-aromatic), 121.18 (C, C-aromatic), 118.66, 118.43, 113.18, 112.89 (CH, C-aromatic), 75.64 (CH), 41.69, 29.04 (CH₂), 19.57, 10.07 (CH₃).

6.5.3 Synthesis of (4-(4-(4-amino-3-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycine (126)

Methyl 4-(3-(sec-butoxy)-4-(3-isopropoxy-4-nitrobenzamido)benzamido)-3-phenoxybenzoate (155)

(C₃₅H₃₅N₃O₉; M.W.: 641.67)



3-isopropoxy-4-nitrobenzoic acid **142** (0.228 mmol) in SOCl₂ (0.5 ml) was refluxed for 2 hours. SOCl₂ was eliminated from the reaction mixture under reduced pressure. The resulting residue was dissolved in anhydrous DCM (3 ml) and added drop-wise to a solution, at 0°C, of methyl 4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoate **152** (0.346 mmol), triethylamine (0.576 mmol) in anhydrous DCM (3 ml) under nitrogen atmosphere. The reaction was allowed to warm-up to room temperature and was stirred overnight. The reaction mixture was diluted with DCM and extracted with 1N HCl solution (3x10ml). The organic layer was washed with brine (3 x 10 mL), dried over MgSO₄ and evaporated at reduced pressure. The obtained crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 5g, DCM -MeOH 100:0 v/v increasing to 95:5 v/v in 6 CV) to methyl 4-(3-(sec-butoxy)-4-(3-isopropoxy-4-nitrobenzamido)benzamido)-3-phenoxybenzoate **155** as yellow powder.

T.L.C. System: DCM/MeOH 95:5 v/v, Rf: 0.32

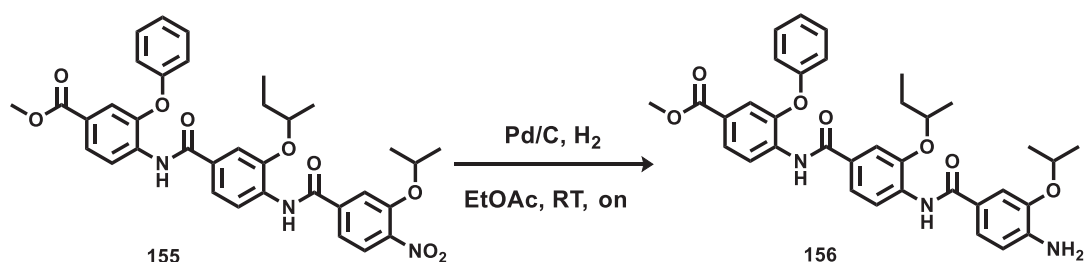
Yield: 195 mg (87%)

¹H NMR (DMSO-d₆), δ: 10.02 (s, 1H), 9.69 (s, 1H), 8.07 (d, *J* = 8.4 Hz, 1H), 7.98 (d, *J* = 8.4 Hz, 2H), 7.82 (dd, ³*J* = 8.4 Hz, ⁴*J* = 1.9 Hz, 1H), 7.79 (d, *J* = 1.3 Hz, 1H), 7.58 (dd, ³*J* = 8.3 Hz, ⁴*J* = 1.6 Hz, 1H), 7.53 – 7.46 (m, 3H), 7.44 – 7.39 (m, 2H), 7.18 (t, *J* = 7.4 Hz, 1H), 7.08 (dd, ³*J* = 8.7 Hz, ⁴*J* = 1.0 Hz, 2H), 4.99 – 4.90 (m, 1H), 4.52 – 4.44 (m, 1H), 3.82 (s, 3H), 1.77 – 1.67 (m, 1H), 1.67 – 1.59 (m, 1H), 1.34 (d, *J* = 5.3 Hz, 6H), 1.26 (d, *J* = 6.0 Hz, 3H), 0.93 (t, *J* = 7.4 Hz, 3H).

^{13}C NMR (DMSO- d_6), δ : 176.62 (C=O), 165.81 (C=O), 164.21 (C=O), 150.22, 149.68, 149.20, 142.80, 139.66, 134.54, 131.64, 131.40 (C, C-aromatic), 130.60 (CH, C-aromatic), 127.56, 127.23 (C, C-aromatic), 125.75, 125.46, 125.24, 124.46, 123.71, 120.61, 120.00, 119.48, 118.99, 115.59, 113.52 (CH, C-aromatic), 76.53, 72.99 (CH), 52.74 (CH₃), 28.97 (CH₂), 22.03, 19.27, 9.86 (CH₃).

Methyl 4-(4-(4-amino-3-*iso*-propoxybenzamido)-3-(*sec*-butoxy)benzamido)-3-phenoxybenzoate (156)

(C₃₅H₃₇N₃O₇; M.W.: 611.69)



A solution of methyl 4-(3-(*sec*-butoxy)-4-(3-isopropoxy-4-nitrobenzamido)benzamido)-3-phenoxybenzoate **155** (0.268 mmol) in anhydrous EtOAc (0.025 M) was stirred under H₂ atmosphere in presence of Pd/C (15%) at room temperature overnight. The reaction mixture was filtered on Celite. The filtered solution was concentrated under reduced pressure and the obtained residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 10g, DCM - MeOH 100:0 v/v increasing to 97:3 v/v in 8 CV) to give pure methyl 4-(4-(4-amino-3-isopropoxybenzamido)-3-(*sec*-butoxy)benzamido)-3-phenoxybenzoate **156** as a white solid.

T.L.C. System: DCM/MeOH 97:3 v/v, R_f: 0.22

Yield: 145 mg (88%)

Purity: 99%

UPLC-MS method C: Rt: 2.228, MS (ESI)⁺: 612.3[M+1]⁺

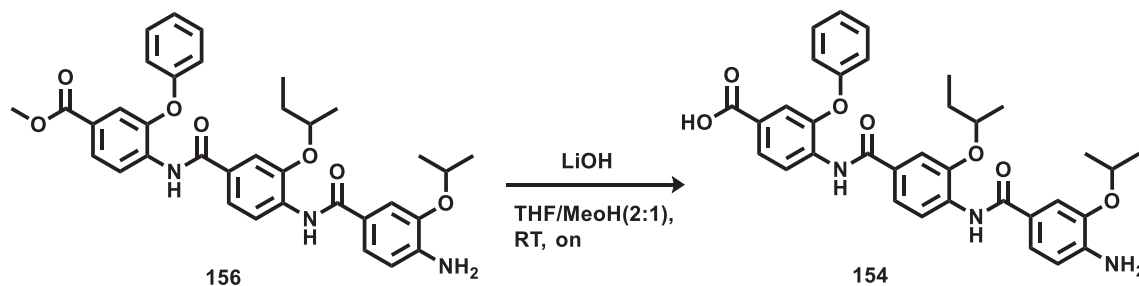
^1H NMR (DMSO- d_6), δ : 9.93 (s, 1H), 8.92 (s, 1H), 8.24 (d, $J = 8.5$ Hz, 1H), 8.09 (d, $J = 8.4$ Hz, 1H), 7.81 (dd, $^3J = 8.4$ Hz, $^4J = 1.9$ Hz, 1H), 7.53 – 7.45 (m, 3H), 7.45 – 7.38 (m, 2H), 7.36 – 7.29 (m, 2H), 7.22 – 7.14 (m, 1H), 7.11 – 7.06 (m, 2H), 6.72 (d, $J = 8.2$ Hz, 1H), 5.45 (s, 2H), 4.65 – 4.56 (m, 1H), 4.56 – 4.47 (m, 1H), 3.82 (s, 3H), 1.80 – 1.73 (m, 1H), 1.73 – 1.64 (m, 1H), 1.32 (dd, $^3J = 6.0$ Hz, $^4J = 2.0$ Hz, 6H), 1.29 (d, $J = 6.1$ Hz, 3H), 0.96 (t, $J = 7.4$ Hz, 3H).

^{13}C NMR (DMSO- d_6), δ : 176.62 (C=O), 165.79 (C=O), 165.41 (C=O), 164.89, 156.48, 149.01, 147.40, 143.86, 143.69, 134.70, 132.76 (C, C-aromatic), 130.58 (CH, C-aromatic), 129.38, 127.01 (C, C-aromatic), 125.48, 125.25, 124.43, 121.78 (CH, C-aromatic), 121.23

(C, C-aromatic), 121.01, 120.33, 119.45, 119.00, 113.25, 113.02 (CH, C-aromatic), 76.55, 70.84 (CH), 52.71 (CH₃), 28.99 (CH), 22.42, 22.38, 19.32, 9.84 (CH₃).

4-(4-(4-amino-3-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxybenzoic acid (154)

(C₃₄H₃₅N₃O₇; M.W.: 597.67)



methyl 4-(4-(4-amino-3-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxy benzoate **156** (0.196 mmol) was solubilised in tetrahydrofuran and methanol (0.07 M, ratio 2:1) and a solution 1N of NaOH (0.11 mL) was added. The reaction mixture was stirred overnight at room temperature. The reaction mixture was concentrated under reduced pressure and the residue was dissolved in water (10 mL). The aqueous layer was acidified with HCl 2N and then extracted with DCM (3x15mL). The organic layers were combined, washed with Brine, dried over MgSO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP Sil 10g, DCM – MeOH 100:0 v/v increasing to 9:1 v/v in 11 CV) to give pure 4-(4-(4-amino-3-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxy benzoic acid **154** as white powder.

T.L.C. System: DCM/MeOH 95:5 v/v, Rf: 0.27

Yield: 52 mg (44%)

Purity: 97%

UPLC-MS method C: Rt: 2.003, MS (ESI)⁺: 598.3 [M+1]⁺

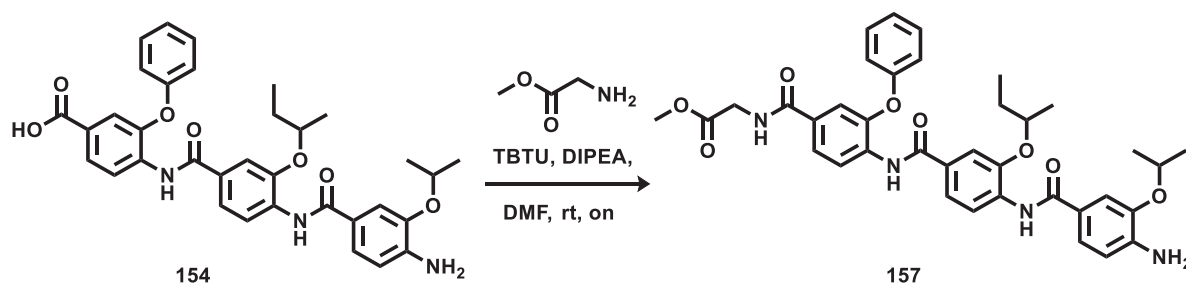
¹H NMR (DMSO-d₆), δ: 13.01 (bs, 1H), 9.90 (s, 1H), 8.92 (s, *J* = 6.6 Hz, 1H), 8.24 (d, *J* = 8.8 Hz, 1H), 8.04 (d, *J* = 8.4 Hz, 1H), 7.78 (dd, ³*J* = 8.4 Hz, ⁴*J* = 1.9 Hz, 1H), 7.51 – 7.47 (m, 2H), 7.45 (d, *J* = 1.9 Hz, 1H), 7.44 – 7.38 (m, 2H), 7.34 – 7.29 (m, 2H), 7.19 – 7.14 (m, 1H), 7.10 – 7.06 (m, 2H), 6.72 (d, *J* = 8.1 Hz, 1H), 5.45 (s, 2H), 4.64 – 4.55 (m, 1H), 4.55 – 4.47 (m, 1H), 1.82 – 1.72 (m, 1H), 1.72 – 1.62 (m, 1H), 1.32 (dd, ³*J* = 6.0 Hz, ⁴*J* = 2.0 Hz, 6H), 1.30 (d, *J* = 6.1 Hz, 3H), 0.96 (t, *J* = 7.4 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 176.62 (C=O), 166.87 (C=O), 165.37 (C=O), 164.89, 156.51, 149.08, 147.41, 143.85, 143.69, 134.20, 132.71 (C, C-aromatic), 130.55 (CH, C-aromatic), 129.44 (C, C-aromatic), 125.45, 125.29, 124.38, 121.77 (CH, C-aromatic), 121.24 (C, C-

aromatic), 120.98, 120.34, 119.57, 119.06, 113.25, 113.02, 112.99 (CH, C-aromatic), 76.54, 70.84 (CH), 29.00 (CH₂), 22.42, 22.39, 19.33, 9.84 (CH₃).

Methyl (4-(4-(4-amino-3-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycinate (157)

(C₃₇H₄₀N₄O₈; M.W.: 668.75)



A mixture of 4-(4-(4-amino-3-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxybenzoic acid **154** (0.067 mmol), methyl glycinate (0.08 mmol), TBTU (0.08 mmol), and DIPEA (2.68 mmol) in anhydrous DMF (0.05 M) was stirred at room temperature for 6 hours. The reaction mixture was diluted in EtOAc (8 mL) and was washed with HCl 1N solution (3 x 10 mL), sat. NaHCO₃ solution (3 x 10 mL) and brine (3 x 10 mL). The organic layers were then dried over MgSO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 5g, n-hexane - DCM 100:0 v/v increasing to 0:100 v/v in 5 CV and DCM - MeOH 100:0 v/v increasing to 97:3 v/v in 14 CV) to give pure methyl (4-(4-(4-amino-3-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycinate **157** as white powder.

T.L.C. System: DCM/MeOH 97:3 v/v, Rf: 0.38

Yield: 49.5 mg (99%)

Purity: 99%

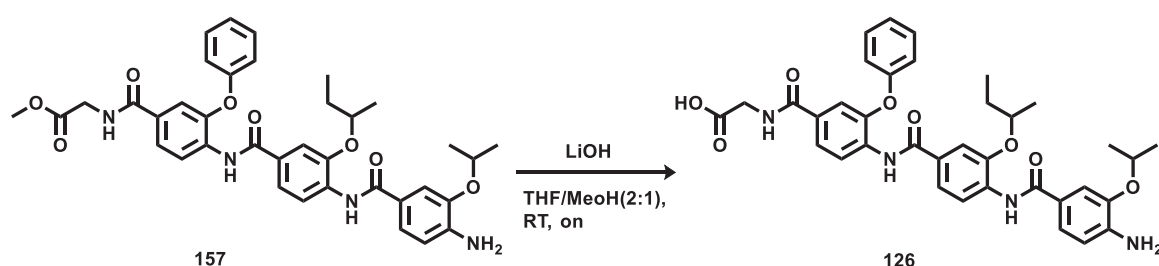
UPLC-MS method C: Rt: 1.992, MS (ESI)⁺: 669.3[M+1]⁺

¹H NMR (CDCl₃-d), δ: 8.74 – 8.71 (m, 2H), 8.67 (s, 1H), 8.61 (d, *J* = 8.4 Hz, 1H), 7.58 (dd, ³*J* = 8.6 Hz, ⁴*J* = 1.9 Hz, 1H), 7.53 (d, ⁴*J* = 1.7 Hz, 1H), 7.45 (dd, ³*J* = 13.2 Hz, ⁴*J* = 1.8 Hz, 1H), 7.42 – 7.37 (m, 1H), 7.30 – 7.23 (m, 1H), 7.19 (t, *J* = 7.4 Hz, 1H), 7.10 – 7.06 (m, 2H), 6.73 (d, *J* = 8.1 Hz, 1H), 6.57 (t, *J* = 5.0 Hz, 1H), 4.72 – 4.63 (m, 1H), 4.58 – 4.49 (m, 1H), 4.21 (s, 2H), 4.20 (s, 2H), 3.79 (s, 3H), 1.83 (s, 1H), 1.74 (s, 1H), 1.39 (d, *J* = 6.0 Hz, 6H), 1.37 (d, *J* = 6.1 Hz, 3H), 1.02 (t, *J* = 7.5 Hz, 3H).

^{13}C NMR ($\text{CDCl}_3\text{-d}$), δ : 176.62 (C=O), 170.47 (C=O), 166.27 (C=O), 165.12 (C=O), 165.03, 155.79, 146.49, 145.76, 144.86, 141.44, 133.28, 132.99 (C, C-aromatic), 130.32 (CH, C-aromatic), 128.99, 128.68 (C, C-aromatic), 124.59 (CH, C-aromatic), 124.04 (C, C-aromatic), 122.56, 119.89, 119.14, 118.71, 118.57, 117.02, 113.62, 112.33, 111.80 (CH, C-aromatic), 76.62, 70.87 (CH), 52.51 (CH_3), 41.76, 29.22 (CH_2), 22.22, 22.21, 19.32, 9.70 (CH_3).

(4-(4-(4-amino-3-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycine (126)

($\text{C}_{36}\text{H}_{38}\text{N}_4\text{O}_8$; M.W.: 654.72)



methyl (4-(4-(4-amino-3-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycinate **157** (0.06 mmol) was solubilised in tetrahydrofuran and methanol (0.04 M, ratio 2:1) and a solution 1N of LiOH (0.06 mL) was added. The reaction mixture was stirred overnight at room temperature. The reaction mixture was concentrated under reduced pressure and the residue was dissolved in water (10 mL). The aqueous layer was acidified with HCl 2N and then extracted with DCM (3x15mL). The organic layers were combined, washed with Brine, dried over MgSO_4 and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 5g, DCM – MeOH 100:0 v/v increasing to 9:1 v/v in 15 CV) to give pure (4-(4-(4-amino-3-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl) glycine **126** as yellow powder.

T.L.C. System: DCM/MeOH 9:1 v/v, Rf: 0.23

Yield: 23 mg (59%)

Purity: 96%

UPLC-MS method C: Rt: 1.881, MS (ESI) $^+$: 665.3[M+1] $^+$

^1H NMR (DMSO-d_6), δ : 12.59 (bs, 1H), 9.87 (s, 1H), 8.91 (s, 1H), 8.87 (t, $J = 5.9$ Hz, 1H), 8.23 (d, $J = 8.9$ Hz, 1H), 7.98 (d, $J = 8.4$ Hz, 1H), 7.76 (dd, $^3J = 8.4$ Hz, $^4J = 1.9$ Hz, 1H), 7.54 (d, $^4J = 1.9$ Hz, 1H), 7.49 – 7.45 (m, 2H), 7.41 – 7.36 (m, 2H), 7.34 – 7.29 (m, 2H), 7.13 (t, $J = 7.4$ Hz, 1H), 7.04 (dd, $^3J = 8.6$ Hz, $^4J = 0.9$ Hz, 2H), 6.72 (d, $J = 8.1$ Hz, 1H), 5.44 (s, 2H), 4.59 (hept, $J = 6.0$ Hz, 1H), 4.51 (h, $J = 6.0$ Hz, 1H), 3.90 (d, $J = 5.8$ Hz, 2H),

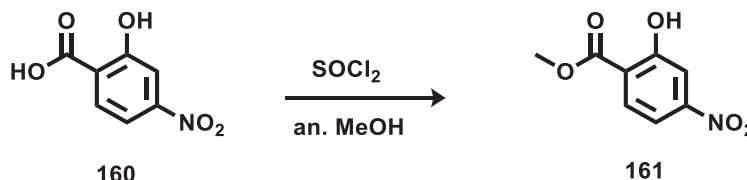
1.80 – 1.72 (m, 1H), 1.72 – 1.64 (m, 1H), 1.32 (dd, $^3J = 6.0$ Hz, $^4J = 1.9$ Hz, 6H), 1.29 (d, $J = 6.1$ Hz, 3H), 0.96 (t, $J = 7.4$ Hz, 3H).

^{13}C NMR (DMSO- d_6), δ : 171.73 (C=O), 165.74 (C=O), 165.33 (C=O), 164.88 (C=O), 156.94, 148.92, 147.38, 143.84, 143.69, 133.20, 132.62, 131.71 (C, C-aromatic), 130.45 (CH, C-aromatic), 129.48 (C, C-aromatic), 125.80, 123.95, 123.33 (CH, C-aromatic), 121.76 (C, C-aromatic), 121.26, 120.94, 120.33, 118.90, 118.41, 113.25, 113.01, 112.93 (CH, C-aromatic), 76.52, 70.84 (CH), 41.70, 29.00 (CH₂), 22.42, 22.38, 19.33, 9.85 (CH₃).

6.5.4 Synthesis of (4-(4-(4-amino-2-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycine (158)

Methyl 2-hydroxy-4-nitrobenzoate (161)

(C₈H₇NO₅; M.W.: 197.15)



To a solution of 2-hydroxy-4-nitrobenzoic acid **160** (5.46 mmol) in anhydrous methanol (0.5 M), thionyl chloride (16.39 mmol) was added drop-wise at 0°C. The reaction mixture was stirred at room temperature overnight. The mixture was concentrated under reduced pressure and the obtained residue was then diluted with a sat. NaHCO₃ solution (60 mL) and extracted with EtOAc (3x60 mL). The organic layers were combined, washed with Brine and then dried over MgSO₄ and concentrated under reduced pressure to give methyl 2-hydroxy-4-nitrobenzoate **161** as a yellow powder, that was used without further purification. T.L.C. System: n-hexane - n-Hex/EtOAc 8:2 v/v, R_f: 0.5.

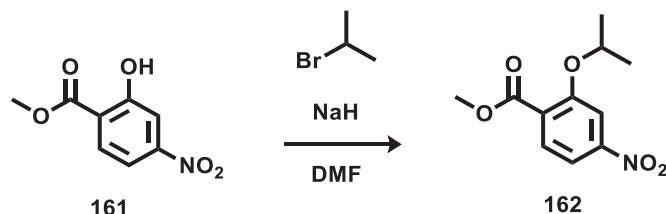
Yield: 961 mg (89%)

Purity: 97%

UPLC-MS method C: Rt: 1.803, MS (ESI)⁺: 196.3[M-1]-

¹H NMR (DMSO-d₆), δ: 10.97 (s, 1H), 7.92 (d, *J* = 8.6 Hz, 1H), 7.74 (d, ⁴*J* = 2.2 Hz, 1H), 7.71 (dd, ³*J* = 8.6 Hz, ⁴*J* = 2.2 Hz, 1H), 3.89 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 167.03 (C=O), 159.21, 151.14 (C, C-aromatic), 132.34 (CH, C-aromatic), 121.71 (C, C-aromatic), 113.94, 112.38 (CH, C-aromatic), 53.18 (CH₃).

Methyl 2-hydroxy-4-nitrobenzoate (162)**(C₁₁H₁₃NO₅; M.W.: 239.23)**

A solution of methyl 2-hydroxy-4-nitrobenzoate **161** (2.33 mmol) and potassium carbonate (2.79 mmol) in anhydrous DMF (0.7 M) was stirred at room temperature for 15 minutes, then 2-bromopropane (3.73 mmol) was added drop-wise and the reaction mixture was stirred overnight at 70°C. The reaction mixture was diluted with EtOAc (20 mL) and then extracted with HCl 1N solution (4x20 mL). The organic layer was then washed with Brine, dried over MgSO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP Sil 50g, n-Hexane – EtOAc 100:0 v/v increasing to 80:20 v/v in 10 CV) to give pure methyl 2-hydroxy-4-nitrobenzoate **162** as a yellow solid.

T.L.C. System: n-hexane -EtOAc 8:2 v/v, Rf: 0.46.

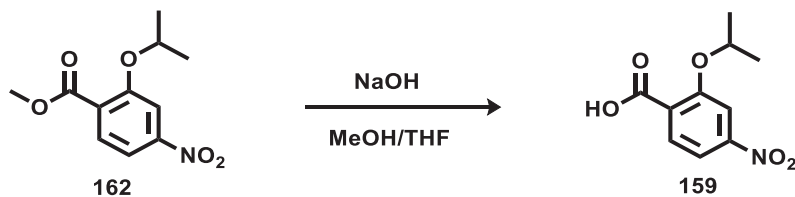
Yield: 504 mg (90%)

Purity: 96%

UPLC-MS method C: Rt: 1.946, MS (ESI)⁺: 181.1[C₉H₁₀NO₃ +1]⁺

¹H NMR (DMSO-d₆), δ: 7.88 (d, ⁴J = 1.7 Hz, 1H), 7.85 – 7.80 (m, 2H), 4.86 (hept, J = 6.0 Hz, 1H), 3.84 (s, 3H), 1.31 (d, J = 6.0 Hz, 6H).

¹³C NMR (DMSO-d₆), δ: 165.79 (C=O), 156.85, 150.62 (C, C-aromatic), 131.67 (CH, C-aromatic), 128.32 (C, C-aromatic), 115.48, 110.09 (CH, C-aromatic), 72.54 (CH), 52.88, 21.97 (CH₃).

2-isopropoxy-4-nitrobenzoic acid (159)**(C₁₀H₁₁NO₅; M.W.: 225.2)**

To a solution of methyl 2-hydroxy-4-nitrobenzoate **162** (1.89 mmol) in tetrahydrofuran and methanol (0.12 M, ratio 2:1) was added a solution 1N of NaOH (3.8 mL). The reaction mixture was stirred at room temperature overnight. The reaction mixture was concentrated under reduced pressure and the residue was dissolved in water (10 mL). The aqueous layer was acidified with HCl 2N and then extracted with DCM (3x15mL). The organic layers were combined, washed with Brine, dried over MgSO₄ and concentrated under reduced pressure to give 2-isopropoxy-4-nitrobenzoic acid **159**, as an off-white powder which was used without further purification.

T.L.C. System: DCM/MeOH 95:5 v/v, R_f: 0.3.

Yield: 394 mg (92%)

Purity: 97%

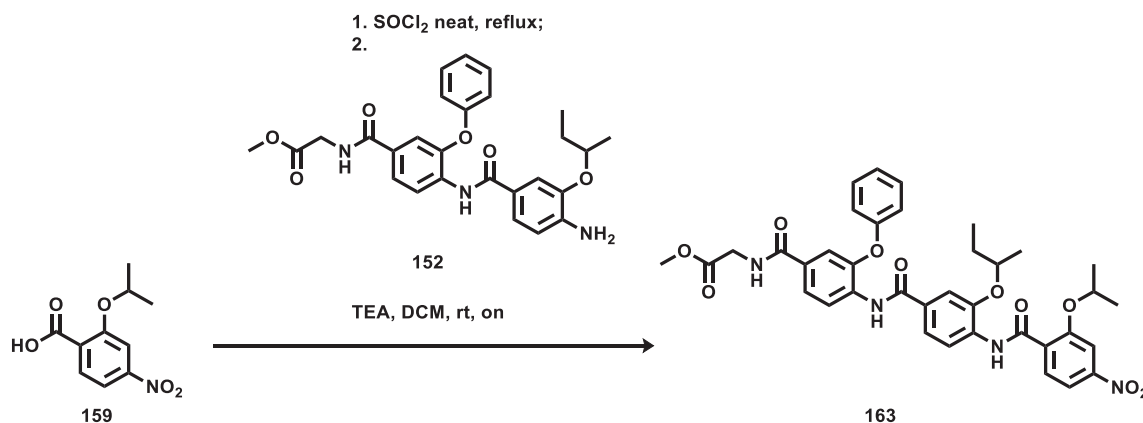
UPLC-MS method C: Rt: 1.684, MS (ESI)⁻: 224.1[M-1]⁻

¹H NMR (DMSO-d₆), δ: 13.31 (bs, 1H), 7.90 (d, *J* = 2.0 Hz, 1H), 7.86 (dd, ³*J* = 8.4 Hz, ⁴*J* = 2.1 Hz, 1H), 7.81 (d, *J* = 8.4 Hz, 1H), 4.89 (hept, *J* = 6.0 Hz, 1H), 1.36 (d, *J* = 6.0 Hz, 6H).

¹³C NMR (DMSO-d₆), δ: 167.07 (C=O), 156.40, 150.14 (C, C-aromatic), 131.15 (CH, C-aromatic), 130.19 (C, C-aromatic), 115.45, 109.84 (CH, C-aromatic), 72.29 (CH), 22.00 (CH₃).

Methyl (4-(3-(sec-butoxy)-4-(2-isopropoxy-4-nitrobenzamido)benzamido)-3-phenoxybenzoyl)glycinate (163)

(C₃₇H₃₈N₄O₁₀; M.W.: 698.73)



2-isopropoxy-4-nitrobenzoic acid **159** (0.135 mmol) in SOCl₂ (0.5 ml) was refluxed for 2 hours. SOCl₂ was eliminated from the reaction mixture under reduced pressure. The resulting residue was dissolved in anhydrous DCM (3 ml) and added drop-wise to a solution, at 0°C, of methyl (4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycinate **152** (0.203 mmol), triethylamine (0.406 mmol) in anhydrous DCM (3 ml) under nitrogen atmosphere. The reaction was allowed to warm-up to room temperature and was stirred overnight. The reaction mixture was diluted with DCM and extracted with 1N HCl solution (3x10ml). The organic layer was washed with brine (3 x 10 mL), dried over MgSO₄ and evaporated at reduced pressure. The obtained crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP Sil 10g, DCM -MeOH 100:0 v/v increasing to 99:1 v/v in 10 CV) to give pure methyl (4-(3-(sec-butoxy)-4-(2-isopropoxy-4-nitrobenzamido)benzamido)-3-phenoxybenzoyl)glycinate **163** as yellow powder.

T.L.C. System: DCM/MeOH 99:1 v/v, Rf: 0.22

Yield: 134.8 mg (95%)

Purity: 99%

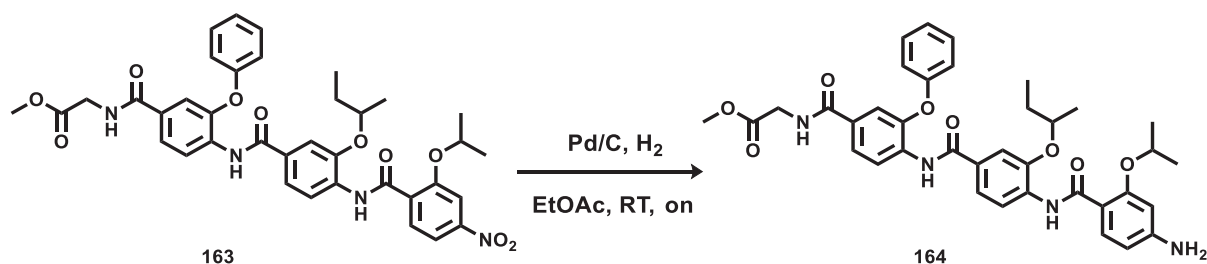
UPLC-MS method C: Rt: 2.213, MS (ESI)⁺: 699.3[M+1]⁺, 721.3[M+Na]⁺

¹H NMR (DMSO-d₆), δ: 9.98 (s, 1H), 9.94 (s, 1H), 9.00 (t, *J* = 5.8 Hz, 1H), 8.46 (d, *J* = 8.7 Hz, 1H), 8.18 (d, *J* = 8.6 Hz, 1H), 8.01 (d, ⁴*J* = 1.8 Hz, 1H), 7.98 (d, *J* = 8.4 Hz, 1H), 7.95 (dd, ³*J* = 8.6 Hz, ⁴*J* = 2.1 Hz, 1H), 7.77 (dd, ³*J* = 8.4 Hz, ⁴*J* = 1.9 Hz, 1H), 7.54 (d, ⁴*J* = 1.9 Hz, 1H), 7.53 – 7.48 (m, 2H), 7.41 – 7.36 (m, 2H), 7.13 (t, *J* = 7.4 Hz, 1H), 7.04 (d, *J* = 7.7 Hz, 2H), 5.10 – 5.01 (m, 1H), 4.56 – 4.48 (m, 1H), 3.99 (d, *J* = 5.8 Hz, 2H), 3.65 (s, 3H), 1.84 – 1.72 (m, 1H), 1.72 – 1.59 (m, 1H), 1.45 (d, *J* = 6.0 Hz, 3H), 1.43 (d, *J* = 6.0 Hz, 3H), 1.31 (d, *J* = 6.1 Hz, 3H), 0.92 (t, *J* = 7.4 Hz, 3H).

^{13}C NMR (DMSO- d_6), δ : 170.81 (C=O), 165.88 (C=O), 165.23 (C=O), 162.25 (C=O), 156.91, 156.04, 150.79, 149.00, 147.36 (C, C-aromatic), 133.26 (CH, C-aromatic), 133.24, 131.55, 131.53 (C, C-aromatic), 130.45 (CH, C-aromatic), 129.16 (C, C-aromatic), 125.91, 123.98, 123.37, 121.02, 120.61, 118.91, 118.42, 116.28, 113.23, 110.59 (CH, C-aromatic), 77.62, 74.42 (CH), 52.22(CH₃), 41.71, 29.22 (CH₂), 22.10, 22.09, 19.84, 10.26 (CH₃).

Methyl (4-(4-(4-amino-2-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycinate (164)

(C₃₇H₄₀N₄O₈; M.W.: 668.75)



A solution of methyl (4-(3-(sec-butoxy)-4-(2-isopropoxy-4-nitrobenzamido)benzamido)-3-phenoxybenzoyl)glycinate **163** (0.16 mmol) in anhydrous ethyl acetate and tetrahydrofuran (0.004 M, ratio 3:1) was stirred under H₂ atmosphere in presence of Pd/C (15%) at room temperature overnight. The reaction mixture was filtered on Celite. The filtered solution was concentrated under reduced pressure and the obtained residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP Sil 10g, DCM - MeOH 100:0 v/v increasing to 97:3 v/v in 8 CV) to give pure methyl (4-(4-(4-amino-2-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycinate **164** as a yellow solid.

T.L.C. System: DCM/MeOH 97:3 v/v, R_f: 0.36

Yield: 102 mg (95%)

Purity: 99%

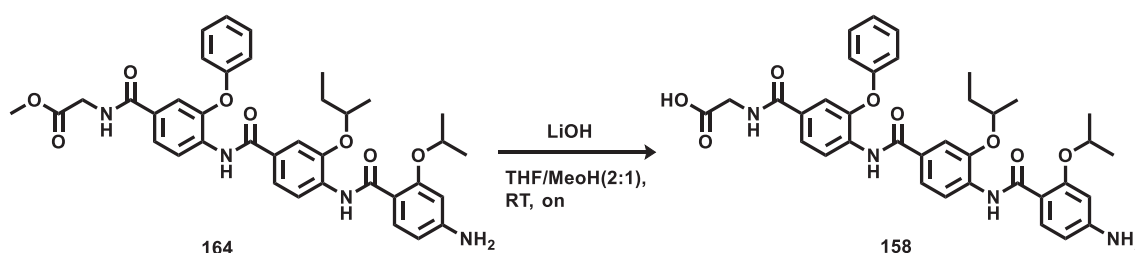
UPLC-MS method C: Rt: 1.958, MS (ESI)⁺: 669.4[M+1]⁺

^1H NMR (DMSO- d_6), δ : 10.00 (s, 1H), 9.85 (s, J = 6.2 Hz, 1H), 8.99 (t, J = 5.8 Hz, 1H), 8.46 (d, J = 8.4 Hz, 1H), 8.00 (d, J = 8.4 Hz, 1H), 7.78 – 7.72 (m, 2H), 7.53 (d, 4J = 1.9 Hz, 1H), 7.49 – 7.42 (m, 2H), 7.42 – 7.35 (m, 2H), 7.13 (t, J = 7.4 Hz, 1H), 7.04 (d, J = 8.5 Hz, 2H), 6.34 (d, J = 1.8 Hz, 1H), 6.28 (dd, 3J = 8.6 Hz, 4J = 1.9 Hz, 1H), 5.91 (s, 2H), 4.67 (hept, J = 6.1 Hz, 1H), 4.46 (h, J = 6.2 Hz, 1H), 3.99 (d, J = 5.8 Hz, 2H), 3.65 (s, 3H), 1.85 – 1.71 (m, 1H), 1.71 – 1.60 (m, 1H), 1.42 (d, J = 6.1 Hz, 3H), 1.40 (d, J = 6.1 Hz, 3H), 1.31 (d, J = 6.1 Hz, 3H), 0.93 (t, J = 7.4 Hz, 3H).

¹³C NMR (DMSO-d₆), δ 170.82 (C=O), 165.90 (C=O), 165.36 (C=O), 163.88 (C=O), 158.08, 156.91, 154.71, 148.86, 146.95 (C, C-aromatic), 133.65 (CH, C-aromatic), 133.38, 133.13, 131.36 (C, C-aromatic), 130.45 (CH, C-aromatic), 128.78 (C, C-aromatic), 125.71, 123.98, 123.37, 121.07, 120.14, 118.87, 118.42, 113.33 (CH, C-aromatic), 109.82 (C, C-aromatic), 107.62, 98.94 (CH, C-aromatic), 77.86, 72.79 (CH), 52.22 (CH₃), 41.71, 29.36 (CH₂), 22.49, 22.47, 20.01, 10.29 (CH₃).

(4-(4-(4-amino-2-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycine (158)

(C₃₆H₃₈N₄O₈; M.W.: 654.72)



methyl (4-(4-(4-amino-2-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxy benzoyl)glycinate **164** (0.06 mmol) was solubilised in tetrahydrofuran and methanol (0.04 M, ratio 2:1) and a solution 1N of LiOH (0.06 mL) was added. The reaction mixture was stirred overnight at room temperature. The reaction mixture was concentrated under reduced pressure and the residue was dissolved in water (10 mL). The aqueous layer was acidified with HCl 2N and then extracted with DCM (3x15mL). The organic layers were combined, washed with Brine, dried over Na₂SO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP Sil 10g, DCM – MeOH 100:0 v/v increasing to 97:3 v/v in 14 CV) to give pure (4-(4-(4-amino-2-isopropoxybenzamido)-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycine **158** as pale yellow powder.

T.L.C. System: DCM/MeOH 9:1 v/v, R_f: 0.2

Yield: 32 mg (80%)

Purity: 99%

UPLC-MS method C: Rt: 1.851, MS (ESI)⁺: 655.3[M+1]⁺

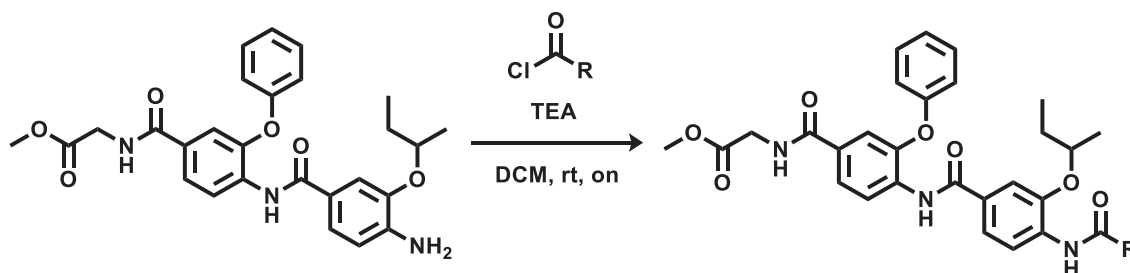
¹H NMR (DMSO-d₆), δ: 12.56 (s, 1H), 10.00 (s, 1H), 9.84 (s, 1H), 8.91 – 8.82 (m, 1H), 8.46 (d, J = 8.4 Hz, 1H), 7.99 (d, J = 8.4 Hz, 1H), 7.78 – 7.69 (m, 2H), 7.54 (d, ⁴J = 1.5 Hz, 1H), 7.50 – 7.42 (m, 2H), 7.38 (t, J = 8.0 Hz, 2H), 7.13 (t, J = 7.4 Hz, 1H), 7.03 (d, J = 7.8 Hz, 2H), 6.33 (d, ⁴J = 1.7 Hz, 1H), 6.28 (dd, ³J = 8.6 Hz, ⁴J = 1.8 Hz, 1H), 5.90 (s, 2H), 4.67 (hept, J = 6.0 Hz, 1H), 4.46 (h, J = 6.1 Hz, 1H), 3.89 (d, J = 4.0 Hz, 2H), 1.83 – 1.71 (m,

1H), 1.71 – 1.59 (m, 1H), 1.42 (d, $J = 6.1$ Hz, 3H), 1.40 (d, $J = 6.1$ Hz, 3H), 1.31 (d, $J = 6.1$ Hz, 3H), 0.93 (t, $J = 7.4$ Hz, 3H).

^{13}C NMR (DMSO- d_6), δ : 171.73 (C=O), 165.74 (C=O), 165.35 (C=O), 163.88 (C=O), 158.08, 156.96, 154.71, 148.83, 146.95 (C, C-aromatic), 133.65 (CH, C-aromatic), 133.28, 133.12, 131.65 (C, C-aromatic), 130.44 (CH, C-aromatic), 128.79 (C, C-aromatic), 125.71, 123.94, 123.35, 121.06, 120.14, 118.91, 118.38, 113.32 (CH, C-aromatic), 109.82 (C, C-aromatic), 107.62, 98.94 (CH, C-aromatic), 77.85, 72.79 (CH), 41.74, 29.36 (CH₂), 22.49, 22.47, 20.01, 10.29 (CH₃).

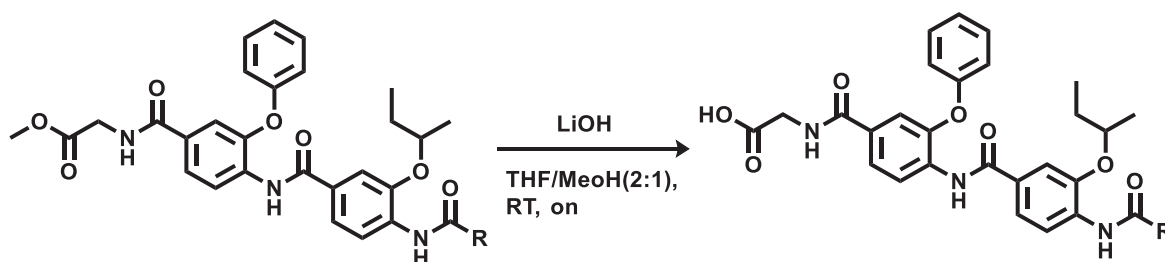
6.5.5 Synthesis of methyl (4-(4-amino-3-(*sec*-butoxy)benzamido)-3-phenoxybenzoyl)glycine derivatives(168-170)

General procedure 9: synthesis of methyl (4-(4-amino-3-(*sec*-butoxy)benzamido)-3-phenoxybenzoyl)glycinate derivatives



The Acyl chloride was added drop-wise to a solution, at 0°C, of 4-(4-amino-3-(*sec*-butoxy)benzamido)-3-phenoxybenzoic acid (0.108 mmol), triethylamine (0.216 mmol) in anhydrous DCM (0.1 M) under nitrogen atmosphere. The reaction was allowed to warm-up to rt and was stirred overnight. The reaction mixture was concentrated at reduced pressure and the residue was purified by flash column chromatography to afford pure methyl (4-(4-amino-3-(*sec*-butoxy)benzamido)-3-phenoxybenzoyl)glycinate derivative.

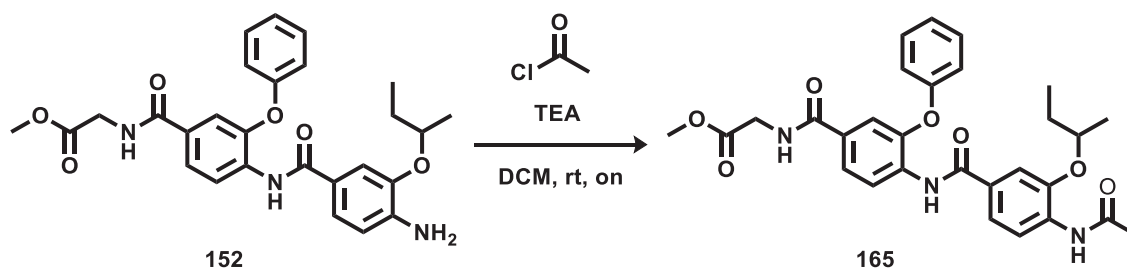
General procedure 10: synthesis of (4-(4-amino-3-(*sec*-butoxy)benzamido)-3-phenoxybenzoyl)glycine derivatives



To a solution of methyl (4-(4-amino-3-(*sec*-butoxy)benzamido)-3-phenoxybenzoyl)glycinate derivative (0.037 mmol) in tetrahydrofuran and methanol (0.05 M, ratio 2:1) was added a solution 1N of LiOH (0.04 mL). The reaction mixture was stirred overnight at room

temperature. The reaction mixture was concentrated under reduced pressure and the residue was dissolved in water (10 mL). The aqueous layer was acidified with HCl 2N and then extracted with DCM (3x10mL). The organic layers were combined, washed with Brine, then dried over MgSO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography to afford pure (4-(4-amino-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycine derivative.

Methyl (4-(4-acetamido-3-(sec-butoxy)benzamido)-3-phenoxybenzoyl)glycinate (165)
(C₂₉H₃₁N₃O₇; M.W.: 533.58)



Procedure: 9

white powder;

T.L.C. System: DCM/MeOH 98:2 v/v, Rf: 0.28

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (DCM -MeOH 100:0 v/v increasing to 96:4 v/v in 11 CV).

Yield: 34 mg (60%)

Purity: 99%

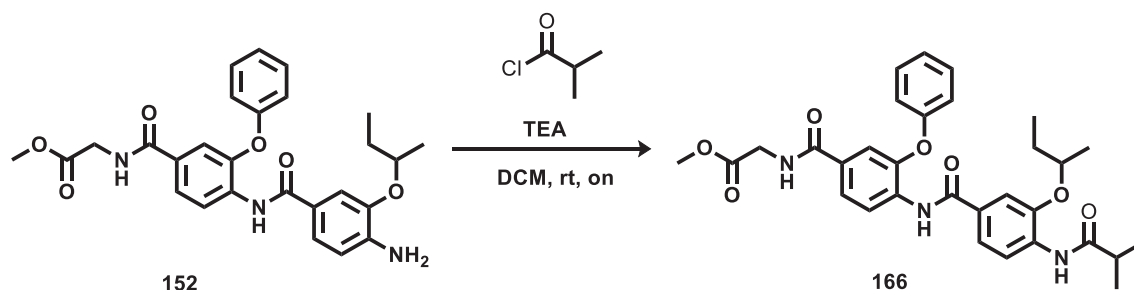
UPLC-MS method C: Rt: 1.871, MS (ESI)⁺: 534.2[M+1]⁺, 556.2[M+Na]⁺

¹H NMR (DMSO-d₆), δ: 9.85 (s, 1H), 9.03 – 8.95 (m, 2H), 8.11 (d, *J* = 8.1 Hz, 1H), 7.97 (d, *J* = 8.4 Hz, 1H), 7.75 (dd, ³*J* = 8.4 Hz, ⁴*J* = 1.9 Hz, 1H), 7.52 (d, ⁴*J* = 1.9 Hz, 1H), 7.43 – 7.34 (m, 4H), 7.13 (t, *J* = 7.4 Hz, 1H), 7.03 (d, *J* = 8.5 Hz, 2H), 4.47 – 4.38 (m, 1H), 3.98 (d, *J* = 5.8 Hz, 2H), 3.65 (s, 3H), 2.14 (s, 3H), 1.82 – 1.70 (m, 1H), 1.70 – 1.56 (m, 1H), 1.26 (d, *J* = 6.1 Hz, 3H), 0.94 (t, *J* = 7.4 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 176.62 (C=O), 170.81 (C=O), 169.29 (C=O), 165.88 (C=O), 165.34, 156.88, 148.94, 133.29, 132.40, 131.42, 130.65 (C, C-aromatic), 130.44, 125.78, 123.98, 123.34, 120.90, 120.64, 118.84, 118.44, 113.17 (CH, C-aromatic), 76.74 (CH), 52.22(CH₃), 41.70 (CH₂), 40.59 (CH₃), 28.88 (CH₂), 19.43, 10.18 (CH₃).

Methyl (4-(3-(*sec*-butoxy)-4-isobutyramidobenzamido)-3-phenoxybenzoyl)glycinate (166)

(C₃₁H₃₅N₃O₇; M.W.: 561.63)



Procedure: 9

yellow solid;

T.L.C. System: DCM/MeOH 98:2 v/v, Rf: 0.28

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (DCM -MeOH 100:0 v/v increasing to 96:4 v/v in 11 CV).

Yield: 68 mg (86%)

Purity: 99%

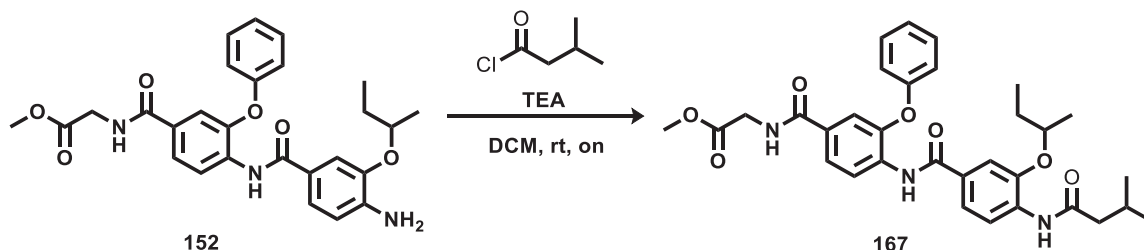
UPLC-MS method C: Rt: 2.026, MS (ESI)⁺: 562.3[M+1]⁺, 584.3[M+Na]⁺

¹H NMR (DMSO-d₆), δ: 9.86 (s, 1H), 8.99 (t, *J* = 5.8 Hz, 1H), 8.85 (s, 1H), 8.07 (d, *J* = 8.3 Hz, 1H), 7.97 (d, *J* = 8.4 Hz, 1H), 7.75 (dd, ³*J* = 8.4 Hz, ⁴*J* = 1.9 Hz, 1H), 7.53 (d, ⁴*J* = 1.9 Hz, 1H), 7.44 – 7.33 (m, 4H), 7.16 – 7.10 (m, 1H), 7.03 (dd, ³*J* = 8.6 Hz, ⁴*J* = 0.9 Hz, 2H), 4.42 (h, *J* = 6.0 Hz, 1H), 3.99 (d, *J* = 5.8 Hz, 2H), 3.65 (s, 3H), 2.80 (hept, *J* = 6.8 Hz, 1H), 1.80 – 1.69 (m, 1H), 1.69 – 1.59 (m, 1H), 1.26 (d, *J* = 6.1 Hz, 3H), 1.11 (d, *J* = 6.8 Hz, 6H), 0.94 (t, *J* = 7.4 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 175.85 (C=O), 170.81 (C=O), 165.89 (C=O), 165.32 (C=O), 156.88, 148.96, 147.78, 133.27, 132.32, 131.45 (C, C-aromatic), 130.45 (CH, C-aromatic), 129.82 (C, C-aromatic), 125.84, 123.98, 123.34, 121.27, 120.69, 118.87, 118.43, 113.24 (CH, C-aromatic), 76.77 (CH), 52.22 (CH), 41.71(CH₂), 35.18 (CH₃), 28.95 (CH₂), 20.01, 19.86, 19.44, 10.03 (CH₃).

Methyl (4-(3-(*sec*-butoxy)-4-(3-methylbutanamido)benzamido)-3-phenoxybenzoyl)glycinate (167)

(C₃₂H₃₇N₃O₇; M.W.: 575.66)



Procedure: 9

white powder;

T.L.C. System: DCM/MeOH 98:2 v/v, Rf: 0.28

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (DCM -MeOH 100:0 v/v increasing to 96:4 v/v in 11 CV).

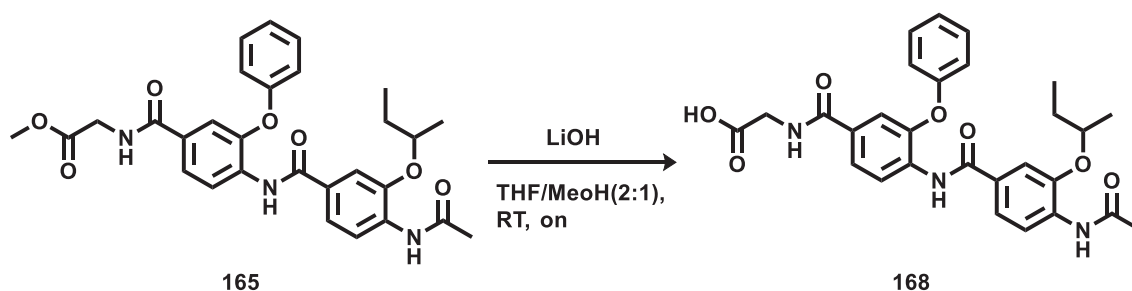
Yield: 65 mg (81%)

Purity: 99%

UPLC-MS method C: Rt: 2.077, MS (ESI)⁺: 576.4[M+1]⁺, 598.3[M+Na]⁺

¹H NMR (DMSO-d₆), δ: 9.90 (s, 1H), 9.04 (t, *J* = 5.8 Hz, 1H), 8.94 (s, 1H), 8.11 (d, *J* = 8.3 Hz, 1H), 8.03 (d, *J* = 8.4 Hz, 1H), 7.81 (dd, ³*J* = 8.4 Hz, ⁴*J* = 1.5 Hz, 1H), 7.58 (d, *J* = 1.5 Hz, 1H), 7.50 – 7.39 (m, 4H), 7.18 (t, *J* = 7.3 Hz, 1H), 7.08 (d, *J* = 7.9 Hz, 2H), 4.47 (h, *J* = 5.9 Hz, 1H), 4.04 (d, *J* = 5.7 Hz, 2H), 3.70 (s, 3H), 2.37 (d, *J* = 6.4 Hz, 2H), 2.17 – 2.04 (m, 1H), 1.85 – 1.74 (m, 1H), 1.74 – 1.63 (m, 1H), 1.31 (d, *J* = 6.0 Hz, 3H), 0.99 (s, 9H).

¹³C NMR (DMSO-d₆), δ: 171.46 (C=O), 170.81 (C=O), 165.89 (C=O), 165.35 (C=O), 156.89, 148.91, 147.76, 133.30, 132.22, 131.43 (C, C-aromatic), 130.44 (CH, C-aromatic), 129.87 (C, C-aromatic), 125.78, 123.97, 123.36, 121.46, 120.61, 118.89, 118.41, 113.21 (CH, C-aromatic), 76.68 (CH), 52.22 (CH), 45.83, 41.71, 28.93 (CH₂), 26.16, 22.68, 22.67, 19.44, 10.09 (CH₃).

(4-(4-acetamido-3-(*sec*-butoxy)benzamido)-3-phenoxybenzoyl)glycine (168)**(C₂₈H₂₉N₃O₇; M.W.: 519.55)**

Procedure 10

pale yellow powder;

T.L.C. System: DCM/MeOH 95:5 v/v, Rf: 0.33

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (DCM -MeOH 100:0 v/v increasing to 95:5 v/v in 16 CV).

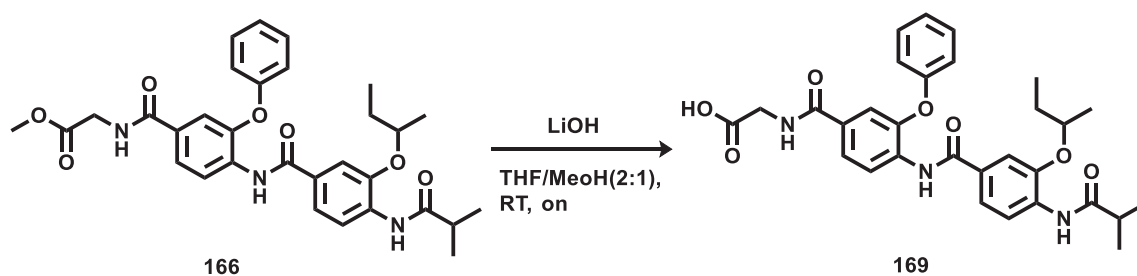
Yield: 17 mg (93%)

Purity: 98%

UPLC-MS method C: Rt: 1.748, MS (ESI)⁺: 520.3[M+1]⁺

¹H NMR (DMSO-d₆), δ: 12.55 (bs, 1H), 9.84 (s, 1H), 8.98 (s, 1H), 8.87 – 8.80 (m, 1H), 8.11 (d, *J* = 8.1 Hz, 1H), 7.96 (d, *J* = 8.4 Hz, 1H), 7.75 (dd, ³*J* = 8.4 Hz, ⁴*J* = 1.9 Hz, 1H), 7.53 (d, ⁴*J* = 1.9 Hz, 1H), 7.43 – 7.34 (m, 4H), 7.15 – 7.09 (m, 1H), 7.05 – 7.00 (m, 2H), 4.42 (dq, *J* = 12.1, 6.1 Hz, 1H), 3.88 (d, *J* = 5.7 Hz, 2H), 2.14 (s, 3H), 1.83 – 1.70 (m, 1H), 1.70 – 1.58 (m, 1H), 1.26 (d, *J* = 6.1 Hz, 3H), 0.94 (t, *J* = 7.4 Hz, 3H).

¹³C NMR (DMSO-d₆), δ: 171.70 (C=O), 169.30 (C=O), 165.70 (C=O), 165.33 (C=O), 156.93, 148.92, 133.18, 132.38, 131.74 (C, C-aromatic), 130.44, 125.79, 123.94, 123.31, 120.96, 120.64, 118.88, 118.40, 113.16 (CH, C-aromatic), 76.74 (CH), 41.76 (CH₂), 40.58 (CH₃), 28.88 (CH₂), 19.43, 10.18 (CH₃).

(4-(3-(*sec*-butoxy)-4-isobutyramidobenzamido)-3-phenoxybenzoyl)glycine (169)**(C₃₀H₃₃N₃O₇; M.W.: 547.61)**

Procedure 10

white powder;

T.L.C. System: DCM/MeOH 95:5 v/v, R_f: 0.33

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (DCM -MeOH 100:0 v/v increasing to 95:5 v/v in 16 CV).

Yield: 39 mg (82%)

Purity: 99%

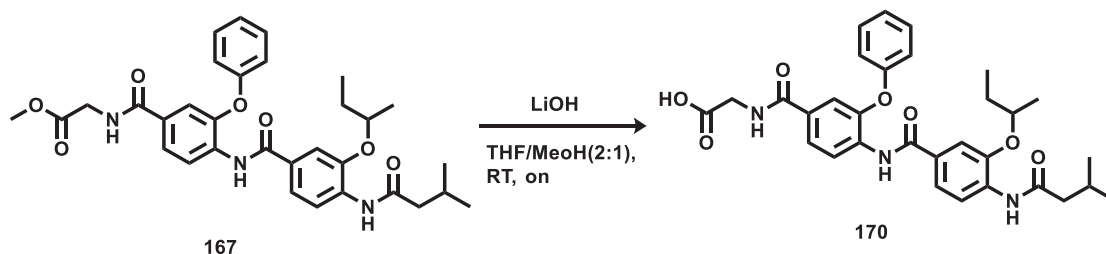
UPLC-MS method C: R_t: 1.94, MS (ESI)⁺: 548.3[M+1]⁺

¹H NMR (DMSO-*d*₆), δ: 12.59 (bs, 1H), 9.85 (s, 1H), 8.92 – 8.74 (m, 2H), 8.06 (d, *J* = 8.2 Hz, 1H), 7.96 (d, *J* = 8.4 Hz, 1H), 7.75 (dd, ³*J* = 8.4 Hz, ⁴*J* = 1.8 Hz, 1H), 7.53 (d, ⁴*J* = 1.8 Hz, 1H), 7.44 – 7.34 (m, 4H), 7.12 (t, *J* = 7.4 Hz, 1H), 7.02 (d, *J* = 7.8 Hz, 2H), 4.46 – 4.37 (m, 1H), 3.89 (d, *J* = 5.8 Hz, 2H), 2.80 (hept, *J* = 6.7 Hz, 1H), 1.80 – 1.70 (m, 1H), 1.69 – 1.57 (m, 1H), 1.26 (d, *J* = 6.1 Hz, 3H), 1.11 (d, *J* = 6.8 Hz, 6H), 0.94 (t, *J* = 7.4 Hz, 3H).

¹³C NMR (DMSO-*d*₆), δ: 175.85 (C=O), 171.73 (C=O), 165.72 (C=O), 165.31 (C=O), 156.93, 148.92, 147.78, 133.17, 132.31, 131.74 (C, C-aromatic), 130.43 (CH, C-aromatic), 129.83 (C, C-aromatic), 125.82, 123.94, 123.33, 121.26, 120.68, 118.91, 118.39, 113.23 (CH, C-aromatic), 76.77 (CH), 41.73(CH₂), 35.18 (CH), 28.95 (CH₂), 20.01, 19.86, 19.44, 10.03 (CH₃).

**(4-(3-(*sec*-butoxy)-4-(3-methylbutanamido)benzamido)-3-phenoxybenzoyl)glycine
(170)**

(C₃₁H₃₅N₃O₇; M.W.: 561.63)



Procedure 10

white powder;

T.L.C. System: DCM/MeOH 95:5 v/v, Rf: 0.33

Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: ZIP KP Sil 5g (DCM -MeOH 100:0 v/v increasing to 95:5 v/v in 16 CV).

Yield: 39 mg (87%)

Purity: 99%

UPLC-MS method C: Rt: 1.979, MS (ESI)⁺: 562.3[M+1]⁺

¹H NMR (DMSO-*d*₆), δ: 12.59 (bs, 1H), 9.84 (s, 1H), 8.88 (s, 1H), 8.84 (t, *J* = 5.8 Hz, 1H), 8.06 (d, *J* = 8.2 Hz, 1H), 7.96 (d, *J* = 8.4 Hz, 1H), 7.75 (dd, ³*J* = 8.4 Hz, ⁴*J* = 1.9 Hz, 1H), 7.53 (d, ⁴*J* = 1.9 Hz, 1H), 7.44 – 7.32 (m, 4H), 7.12 (t, *J* = 7.4 Hz, 1H), 7.02 (d, *J* = 8.6 Hz, 2H), 4.42 (h, *J* = 6.0 Hz, 1H), 3.88 (d, *J* = 5.8 Hz, 2H), 2.32 (d, *J* = 5.8 Hz, 2H), 2.12 – 1.99 (m, 1H), 1.80 – 1.69 (m, 1H), 1.69 – 1.56 (m, 1H), 1.25 (d, *J* = 6.1 Hz, 3H), 0.95 (s, 9H).

¹³C NMR (DMSO-*d*₆), δ: 171.71 (C=O), 171.46 (C=O), 165.69 (C=O), 165.34 (C=O), 156.94, 148.88, 133.18, 132.20, 131.75 (C, C-aromatic), 130.43 (CH, C-aromatic), 129.88 (C, C-aromatic), 125.79, 123.93, 123.33, 120.60, 118.92, 118.36, 113.20 (CH, C-aromatic), 76.68 (CH), 45.83, 41.79, 28.93 (CH₂), 26.16 (CH), 22.69, 22.67, 19.44, 10.09 (CH₃).

References

- 1 Feldman A.S., He Y., Moore M.L., Hershenson M.B., Hartert T.V. Toward primary prevention of asthma: Reviewing the evidence for early-life respiratory viral infections as modifiable risk factors to prevent childhood asthma. *Am. J. Respir. Crit. Care Med.* **2015**, 91, 34–44. <https://doi.org/10.1164/rccm.201405-0901PP>
- 2 Mazur N.I., Martínón-Torres F., Baraldi E., Fauroux B., Greenough A.; Heikkinen, T. et al. Lower respiratory tract infection caused by respiratory syncytial virus: current management and new therapeutics. *Lancet Respir. Med.* **2015**, 2600, 1–13. [https://doi.org/10.1016/S2213-2600\(15\)00255-6](https://doi.org/10.1016/S2213-2600(15)00255-6)
- 3 Draft C., Mortality G.B.D., Collaborators D., Burden T.G. et al. Global, regional, and national age–sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **2014**, 385, 117–71. [https://doi.org/10.1016/S0140-6736\(14\)61682-2](https://doi.org/10.1016/S0140-6736(14)61682-2)
- 4 Pelletier A.J., Mansbach J.M., Camargo C. Direct medical costs of bronchiolitis hospitalizations in the United States. *Pediatrics* **2006**, 118, 2418–2423. <https://doi.org/10.1542/peds.2006-1193>
- 5 Ralston S.L., Lieberthal A.S., Meissner H.C. et al. Clinical Practice Guideline : The Diagnosis , Management , and Prevention of Bronchiolitis. *Am. Acad. Pediatr.* **2014**,134, 1474-1502. <https://doi.org/10.1542/peds.2014-2742>
- 6 Meng J., Stobart C.C., Hotard A.L., Moore M.L. An Overview of Respiratory Syncytial Virus. *PLoS Pathog.* **2014**, 10, e1004-10016. <https://doi.org/10.1371/journal.ppat.1004016>
- 7 Heylen E., Neyts, J., Jochmans D. Drug candidates and model systems in respiratory syncytial virus antiviral drug discovery. *Biochem. Pharmacol.* **2017**,127, 1–12. <https://doi.org/10.1016/j.bcp.2016.09.014>
- 8 Simões E.A.F., Bont L., Manzoni P., Fauroux B., Paes B., Figueras-Aloy J. et al. Past, Present and Future Approaches to the Prevention and Treatment of Respiratory Syncytial Virus Infection in Children. *Infect. Dis. Ther.* 2018;7(1):87–120. <https://doi.org/10.1007/s40121-018-0188-z>

- 9** Borchers A.T., Chang C., Gershwin M.E., Gershwin L.J. Respiratory syncytial virus—a comprehensive review. *Clin. Rev. Allergy Immunol.* **2013**, 45, 331–379. <https://doi.org/10.1007/s12016-013-8368-9>
- 10** Piedimonte G., Perez M.K. Respiratory syncytial virus infection and bronchiolitis. *Pediatr. Rev.* **2014**, 35, 519–530. <https://doi.org/10.1542/pir.35-12-519>
- 11** Sigurs N., Aljassim F., Kjellman B., Robinson P.D., Sigurbergsson F., Bjarnason R. et al. Asthma and allergy patterns over 18 years after severe RSV bronchiolitis in the first year of life. *Thorax* **2010**, 65, 1045–1052. <https://doi.org/10.1136/thx.2009.121582>
- 12** Zhang L., Peeples M.E., Boucher R.C., Collins P.L., Pickles R.J. Respiratory Syncytial Virus Infection of Human Airway Epithelial Cells Is Polarized, Specific to Ciliated Cells and without Obvious Cytopathology *J. Virol.* **2002**, 76, 5654–5666. <https://doi.org/10.1128/JVI.76.11.5654-5666.2002>
- 13** Kingsbury D.W., Bratt M.A., Choppin P.W., Hanson R.P., Hosaka Y., Ter Meulen V. et al. Paramyxoviridae. *Intervirology* **1978**, 10, 137–152. <https://doi.org/10.1159/000148979>
- 14** Gotoh B., Komatsu T., Takeuchi K., Yokoo J. Paramyxovirus accessory proteins as interferon antagonists. *Microbiol. Immunol.* **2001**, 45, 787–800. <https://doi.org/10.1111/j.1348-0421.2001.tb01315.x>
- 15** Curran J. A role for the Sendai virus P protein trimer in RNA synthesis. *J. Virol.* **1998**, 72, 4274–4280. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC109657/>
- 16** Gilca R., De Serres .G, Tremblay M., Vachon M.L., Leblanc E., Bergeron M.G., et al. Distribution and Clinical Impact of Human Respiratory Syncytial Virus Genotypes in Hospitalized Children over 2 Winter Seasons. *J. Infect. Dis.* **2006**, 193, 54–58. <https://doi.org/10.1086/498526>
- 17** Kiss G, Holl JM, Williams GM, Alonas E, Vanover D, Lifland AW, et al. Structural analysis of respiratory syncytial virus reveals the position of M2-1 between the matrix protein and the ribonucleoprotein complex. *J. Virol.* **2014**, 88, 7602–17. <https://doi.org/10.1128/JVI.00256-14>
- 18** Liljeroos L, Krzyzaniak MA, Helenius A, Butcher SJ. Architecture of respiratory syncytial virus revealed by electron cryotomography. *Proc. Natl. Acad. Sci. U S A.* **2013**;110 (27):11133–11138. <https://doi.org/10.1073/pnas.1309070110>

- 19** Vijayakrishnan S, Loney C, Jackson D, Suphamungmee W, Rixon FJ, Bhella D. Cryotomography of Budding Influenza A Virus Reveals Filaments with Diverse Morphologies that Mostly Do Not Bear a Genome at Their Distal End. *PLoS Pathog.* **2013**;9(6), 1-11. <https://doi.org/10.1371/journal.ppat.1003413>
- 20** Haywood AM. Membrane uncoating of intact enveloped viruses. *J. Virol.* **2010**;84(21):10946–55. <https://doi.org/10.1128/JVI.00229-10>
- 21** Mink MA, Stec DS, Collins PL. Nucleotide sequences of the 3' leader and 5' trailer regions of human respiratory syncytial virus genomic RNA. *Virology* **1991**;185(2):615–24. [https://doi.org/10.1016/0042-6822\(91\)90532-G](https://doi.org/10.1016/0042-6822(91)90532-G)
- 22** Cowton VM, McGivern DR, Fearn R. Unravelling the complexities of respiratory syncytial virus RNA synthesis. *J. Gen. Virol.* **2006**;87(7):1805–21. <https://doi.org/10.1099/vir.0.81786-0>
- 23** Fearn R, Collins PL. Model for polymerase access to the overlapped L gene of respiratory syncytial virus. *J. Virol.* **1999**;73(1):388–97. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC103844/>
- 24** Kuo L, Fearn R, Collins PL. Analysis of the gene start and gene end signals of human respiratory syncytial virus: quasi-templated initiation at position 1 of the encoded mRNA. *J. Virol.* **1997**;71(7):4944–53. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC191725/>
- 25** Fearn R, Collins PL, Peeples ME. Functional analysis of the genomic and antigenomic promoters of human respiratory syncytial virus. *J. Virol.* **2000**;74(13):6006–14. <https://doi.org/10.1128/JVI.74.13.6006-6014.2000>
- 26** Atreya PL, Peeples ME, Collins PL. The NS1 protein of human respiratory syncytial virus is a potent inhibitor of minigenome transcription and RNA replication. *J. Virol.* **1998**;72(2):1452–61. <https://www.ncbi.nlm.nih.gov/pubmed/9445048>
- 27** Hallak LK, Collins PL, Knudson W, Peeples ME. Iduronic acid-containing glycosaminoglycans on target cells are required for efficient respiratory syncytial virus infection. *Virology*. **2000**;271(2):264–75. <https://doi.org/10.1006/viro.2000.0293>
- 28** Hallak LK, Spillmann D, Collins PL, Peeples ME. Glycosaminoglycan sulfation requirements for respiratory syncytial virus infection. *J. Virol.* **2000**;74(22):10508–13. <https://doi.org/10.1128/JVI.74.22.10508-10513.2000>

- 29** Schmid K, Grundboeck-Jusco J, Kimura A, Tschopp FA, Zollinger R, Binette JP, et al. The distribution of the glycosaminoglycans in the anatomic components of the lung and the change in concentration of these macromolecules during development and aging. *Biochim. Biophys. Acta.* **1982**; 716:178–87. [https://doi.org/10.1016/0304-4165\(82\)90266-5](https://doi.org/10.1016/0304-4165(82)90266-5)
- 30** Levine S, Kaliaber-Franco R, Paradiso PR. Demonstration that glycoprotein G is the attachment protein of respiratory syncytial virus. *J. Gen. Virol.* **1987**;68(9):2521–4. <https://doi.org/10.1099/0022-1317-68-9-2521>
- 31** Techaarpornkul S, Barretto N, Peeples ME. Functional Analysis of Recombinant Respiratory Syncytial Virus Deletion Mutants Lacking the Small Hydrophobic and / or Attachment Glycoprotein Gene. *J. Virol.* **2001**;75(15):6825–34. <https://doi.org/10.1128/JVI.75.15.6825-6834.2001>
- 32** Harris J, Werling D. Binding and entry of respiratory syncytial virus into host cells and initiation of the innate immune response. *Cell Microbiol.* **2003**;5(10):671–80. <https://doi.org/10.1046/j.1462-5822.2003.00313.x>
- 33** Swanson KA, Settembre EC, Shaw CA, Dey AK, Rappuoli R, Mandl CW, et al. Structural basis for immunization with postfusion respiratory syncytial virus fusion F glycoprotein (RSV F) to elicit high neutralizing antibody titers. *Proc. Natl. Acad. Sci.* **2011**; 108(23):9619–24. <https://doi.org/10.1073/pnas.1106536108>
- 34** McLellan JS, Yang Y, Graham BS, Kwong PD. Structure of respiratory syncytial virus fusion glycoprotein in the postfusion conformation reveals preservation of neutralizing epitopes. *J. Virol.* **2011**; 85(15):7788–96. <https://doi.org/10.1128/JVI.00555-11>
- 35** McLellan JS, Chen M, Leung S, Graepel KW, Du X, Yang Y, Zhou T, Baxa U, Yasuda E, Beaumont T, Kumar A, Modjarrad K, Zheng Z, Zhao M, Xia N, Kwong PD, Graham BS. Structure of RSV Fusion Glycoprotein Trimer Bound to a Prefusion-Specific Neutralizing Antibody. *Science* **2013**; 340: 1113–1117. <https://doi.org/10.1126/science.1234914>
- 36** Yin H-S, Paterson RG, Wen X, Lamb RA, Jardetzky TS. Structure of the uncleaved ectodomain of the paramyxovirus (hPIV3) fusion protein. *Proc. Natl. Acad. Sci. U S A.* **2005**;102(26):9288–93. <https://doi.org/10.1073/pnas.0503989102>
- 37** Melero JA, Mas V. The Pneumovirinae fusion (F) protein: A common target for vaccines and antivirals. *Virus Res.* **2015**;1–8. <https://doi.org/10.1016/j.virusres.2015.02.024>

- 38** Collins PL, Mink M a, Stec DS. Rescue of synthetic analogs of respiratory syncytial virus genomic RNA and effect of truncations and mutations on the expression of a foreign reporter gene. *Proc. Natl. Acad. Sci. U S A.* **1991**;88(21):9663–7. <https://doi.org/10.1073/pnas.88.21.9663>
- 39** Collins PL, Dickens LE, Buckler-White A, Olmsted R, Spriggs MK, Camargo E, et al. Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order. *Proc. Natl. Acad. Sci. U S A.* **1986**; 83(13):4594–8. <https://doi.org/10.1073/pnas.83.13.4594>
- 40** Collins PL, Olmsted RA, Spriggs MK, Johnson PR, Buckler-White AJ. Gene overlap and site-specific attenuation of transcription of the viral polymerase L gene of human respiratory syncytial virus. *Proc. Natl. Acad. Sci. U S A.* **1987**;84(15):5134–8. <https://doi.org/10.1073/pnas.84.15.5134>
- 41** Whelan SPJ, Barr JN, Wertz GW. Transcription and replication of nonsegmented negative-strand RNA viruses. *Curr. Top Microbiol. Immunol.* **2004**; 283:61–119. https://doi.org/10.1007/978-3-662-06099-5_3
- 42** Fearn R, Collins PL, Peeples ME. Functional analysis of the genomic and antigenomic promoters of human respiratory syncytial virus. *J. Virol.* 2000;74(13):6006–14. <https://doi.org/10.1128/jvi.74.13.6006-6014.2000>
- 43** Bhella D, Ralph A, Murphy LB, Yeo RP. Significant differences in nucleocapsid morphology within the Paramyxoviridae. *J. Gen. Virol.* **2002**;83: 1831–9. <https://doi.org/10.1099/0022-1317-83-8-1831>
- 44** Murphy LB, Loney C, Murray J, Bhella D, Ashton P, Yeo RP. Investigations into the amino-terminal domain of the respiratory syncytial virus nucleocapsid protein reveal elements important for nucleocapsid formation and interaction with the phosphoprotein. *Virology.* **2003**;307(1):143–53. [https://doi.org/10.1016/s0042-6822\(02\)00063-6](https://doi.org/10.1016/s0042-6822(02)00063-6)
- 45** Lifland AW, Jung J, Alonas E, Zurla C, Crowe JE, Santangelo PJ. Human respiratory syncytial virus nucleoprotein and inclusion bodies antagonize the innate immune response mediated by MDA5 and MAVS. *J. Virol.* **2012**;86(15):8245–58. <https://doi.org/10.1128/JVI.00215-12>
- 46** Baviskar PS, Hotard AL, Moore ML, Oomens AGP. The Respiratory Syncytial Virus Fusion Protein Targets to the Perimeter of Inclusion Bodies and Facilitates Filament

Formation by a Cytoplasmic Tail -Dependent Mechanism. *J. Virol.* **2013**; 87:10730–41. <https://doi.org/10.1128/JVI.03086-12>

47 García J, García-Barreno B, Vivo A, Melero JA. Cytoplasmic Inclusions of Respiratory Syncytial Virus-Infected Cells: Formation of Inclusion Bodies in Transfected Cells That Coexpress the Nucleoprotein, the Phosphoprotein, and the 22K Protein Title. *Virology.* **1993**;195(1):243–7. <https://doi.org/10.1006/viro.1993.1366>

48 Chang TH, Segovia J, Sabbah A, Mgbemena V, Bose S. Cholesterol-rich lipid rafts are required for release of infectious human respiratory syncytial virus particles. *Virology* **2012**;422(2):205–13. <https://doi.org/10.1016/j.virol.2011.10.029>

49 Batonick M, Oomens AGP, Wertz GW. Human respiratory syncytial virus glycoproteins are not required for apical targeting and release from polarized epithelial cells. *J. Virol.* **2008**;82(17):8664–72. <https://doi.org/10.1128/JVI.00827-08>

50 Brock SC, Heck JM, McGraw PA, Crowe JE Jr. The Transmembrane Domain of the Respiratory Syncytial Virus F Protein Is an Orientation-Independent Apical Plasma Membrane Sorting Sequence *J. Virol.* **2005**;79(19):12528–35. <https://doi.org/10.1128/JVI.79.19.12528-12535.2005>

51 Marty A, Meanger J, Mills J, Shields B, Ghildyal R. Association of matrix protein of respiratory syncytial virus with the host cell membrane of infected cells. *Arch. Virol.* **2004**;149(1):199–210. <https://doi.org/10.1007/s00705-003-0183-9>

52 Henderson G, Murray J, Yeo RP. Sorting of the Respiratory Syncytial Virus Matrix Protein into Detergent-Resistant Structures Is Dependent on Cell-Surface Expression of the Glycoproteins. *Virology.* **2002**;300(2):244–54. <https://doi.org/10.1006/viro.2002.1540>

53 Yeo DS, Chan R, Brown G, Ying L, Sutejo R, Aitken J, Tan BH, Wenk MR, Sugrue RJ. Evidence that selective changes in the lipid composition of raft-membranes occur during respiratory syncytial virus infection. *Virology.* **2009**; 386(1):168–82. <https://doi.org/10.1016/j.virol.2008.12.017>

54 Brown G, Jeffree CE, McDonald T, Rixon HW, Aitken JD, Sugrue RJ. Analysis of the interaction between respiratory syncytial virus and lipid-rafts in Hep2 cells during infection. *Virology.* **2004**;327(2):175–85. <https://doi.org/10.1016/j.virol.2004.06.038>

55 Harrison MS, Sakaguchi T, Schmitt AP. Paramyxovirus assembly and budding: Building particles that transmit infections. *Int. J. Biochem. Cell. Biol.* **2010**;42(9):1416–29. <https://doi.org/10.1016/j.biocel.2010.04.005>

- 56** Bornholdt ZA, Noda T, Abelson DM, Halfmann P, Wood MR, Kawaoka Y, Saphire EO. Structural rearrangement of ebola virus vp40 begets multiple functions in the virus life cycle. *Cell* **2013**;154(4):763–74. <https://doi.org/10.1016/j.cell.2013.07.015>
- 57** Wertz GW, Collins PL, Huang Y, Gruber C, Levine S, Ball LA. Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein. *Proc. Natl. Acad. Sci. U S A.* **1985**; 82:4075–9. <https://doi.org/10.1073/pnas.82.12.4075>
- 58** Roberts SR, Lichtenstein D, Ball LA, Wertz GW. The membrane-associated and secreted forms of the respiratory syncytial virus attachment glycoprotein G are synthesized from alternative initiation codons. *J. Virol.* **1994**;68(7):4538–4546. <https://www.ncbi.nlm.nih.gov/pubmed/8207828>
- 59** Polack FP, Irusta PM, Hoffman SJ, Schiatti MP, Melendi GA, Delgado MF, et al. The cysteine-rich region of respiratory syncytial virus attachment protein inhibits innate immunity elicited by the virus and endotoxin. *Proc. Natl. Acad. Sci. U S A.* **2005**; 102(25):8996–9001. <https://doi.org/10.1073/pnas.0409478102>
- 60** Bukreyev A, Yang L, Fricke J, Cheng L, Ward JM, Murphy BR, et al. The secreted form of respiratory syncytial virus G glycoprotein helps the virus evade antibody-mediated restriction of replication by acting as an antigen decoy and through effects on Fc receptor-bearing leukocytes. *J. Virol.* **2008**;82(24):12191–204. <https://doi.org/10.1128/JVI.01604-08>
- 61** Kochva U, Leonov H, Arkin IT. Modeling the structure of the respiratory syncytial virus small hydrophobic protein by silent-mutation analysis of global searching molecular dynamics. *Protein Sci.* **2003**;12(12):2668–74. <https://doi.org/10.1110/ps.03151103>
- 62** Gan SW, Tan E, Lin X, Yu D, Wang J, Tan GM, Vararattanavech A, Yeo CY, Soon CH, Soong TW, Pervushin K, Torres J. The small hydrophobic protein of the human respiratory syncytial virus forms pentameric ion channels. *J. Biol. Chem.* **2012**; 287(29):24671–89. <https://doi.org/10.1074/jbc.M111.332791>
- 63** Nieva JL, Madan V, Carrasco L. Viroporins: structure and biological functions. *Nat Rev. Microbiol.* **2012**;10(8):563–74. <https://doi.org/10.1038/nrmicro2820>
- 64** Burke E, Dupuy L, Wall C, Barik S. Role of cellular actin in the gene expression and morphogenesis of human respiratory syncytial virus. *Virology.* **1998**;252(1):137–48 <https://doi.org/10.1006/viro.1998.9471>

- 65** Burke E, Mahoney NM, Almo SC, Barik S. Profilin is required for optimal actin-dependent transcription of respiratory syncytial virus genome RNA. *J. Virol.* **2000**;74(2):669–75. <https://doi.org/10.1128/jvi.74.2.669-675.2000>
- 66** Poch O, Sauvaget I, Delarue M, Tordo N. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J.* **1989**; 8(12):3867–74. <https://doi.org/10.1002/j.1460-2075.1989.tb08565.x>
- 67** Poch O, Blumberg BM, Bougueleret L, Tordo N. Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. *J. Gen. Virol.* **1990**;71(5):1153–62. <https://doi.org/10.1099/0022-1317-71-5-1153>
- 68** Ferron F, Longhi S, Henrissat B, Canard B. Viral RNA- polymerases – a predicted 2'-O-ribose methyltransferase domain shared by all Mononegavirales. *Trends Biochem. Sci.* **2002**; 27(5):222–4. [https://doi.org/10.1016/s0968-0004\(02\)02091-1](https://doi.org/10.1016/s0968-0004(02)02091-1)
- 69** Liuzzi M, Mason SW, Cartier M, Mccollum RS, Dansereau N, Lapeyre N, et al. Inhibitors of Respiratory Syncytial Virus Replication Target Cotranscriptional mRNA Guanylylation by Viral RNA-Dependent RNA Polymerase. *J. Virol.* **2005**; 79(20):13105–15. <https://doi.org/10.1128/JVI.79.20.13105-13115.2005>
- 70** Noval MG, Esperante SA, Molina IG, Chemes LB, Prat-Gay G. Intrinsic disorder to order transitions in the scaffold phosphoprotein P from the respiratory syncytial virus RNA-polymerase complex. *Biochemistry* **2016**; 55(10):1441-54. <https://doi.org/10.1021/acs.biochem.5b01332>
- 71** Mazumder B, Barik S. Requirement of Casein Kinase II-Mediated Phosphorylation for the Transcriptional Activity of Human Respiratory Syncytial Viral Phosphoprotein P: Transdominant Negative Phenotype of Phosphorylation-Defective P Mutant. *Virology.* **1994**; 205(1):104–11. <https://doi.org/10.1006/viro.1994.1624>
- 72** Barik S, McLean T, Dupuy LC. Phosphorylation of Ser232 directly regulates the transcriptional activity of the P protein of human respiratory syncytial virus: phosphorylation of Ser237 may play an accessory role. *Virology.* **1995**; 213(2):405–12. <https://doi.org/10.1006/viro.1995.0013>
- 73** Grosfeld H, Hill MG, Collins PL. RNA replication by respiratory syncytial virus (RSV) is directed by the N, P, and L proteins; transcription also occurs under these conditions but

requires RSV superinfection for efficient synthesis of full-length mRNA. *J. Virol.* **1995**; 69(9):5677–86. <https://www.ncbi.nlm.nih.gov/pubmed/7637014>

74 Yu Q, Hardy RW, Wertz GW. Functional cDNA clones of the human respiratory syncytial (RS) virus N, P, and L proteins support replication of RS virus genomic RNA analogs and define minimal trans-acting requirements for RNA replication. *J. Virol.* **1995**; 69(4):2412–9. <https://www.ncbi.nlm.nih.gov/pubmed/7884888>

75 Dupuy LC, Dobson S, Bitko V, Barik S. Casein kinase 2-mediated phosphorylation of respiratory syncytial virus phosphoprotein P is essential for the transcription elongation activity of the viral polymerase; phosphorylation by casein kinase 1 occurs mainly at Ser(215) and is without effect. *J. Virol.* **1999**; 73(10):8384–92. <https://www.ncbi.nlm.nih.gov/pubmed/10482589>

76 Mellon MG, Emerson SU. Rebinding of transcriptase components (L and NS proteins) to the nucleocapsid template of vesicular stomatitis virus. *J. Virol.* **1978**; 27(3):560–567. <https://www.ncbi.nlm.nih.gov/pubmed/212581>

77 Kingston RL, Hamel DJ, Gay LS, Dahlquist FW, Matthews BW. Structural basis for the attachment of a paramyxoviral polymerase to its template. *Proc. Natl. Acad. Sci. U S A* **2004**; 101(22):8301–6. <https://doi.org/10.1073/pnas.0402690101>

78 Tanner SJ, Ariza A, Richard C-A, Kyle HF, Dods RL, Blondot M-L, et al. Crystal structure of the essential transcription antiterminator M2-1 protein of human respiratory syncytial virus and implications of its phosphorylation. *Proc. Natl. Acad. Sci.* **2014**; 111(4):1580–5. <https://doi.org/10.1073/pnas.1317262111>

79 Hardy RW, Wertz GW. The product of the respiratory syncytial virus M2 gene ORF1 enhances readthrough of intergenic junctions during viral transcription. *J. Virol.* **1998**; 72(1):520–6. <https://www.ncbi.nlm.nih.gov/pubmed/9420254>

80 Hardy RW, Harmon SB, Wertz GW. Diverse gene junctions of respiratory syncytial virus modulate the efficiency of transcription termination and respond differently to M2-mediated antitermination. *J. Virol.* **1999**; 73(1):170–176. <https://ncbi.nlm.nih.gov/pubmed/9847319>

81 Sutherland KA, Collins PL, Peeples ME. Synergistic Effects of Gene-End Signal Mutations and the M2-1 Protein on Transcription Termination by Respiratory Syncytial Virus. *Virology* **2001**; 288(2):295–307. <https://doi.org/10.1006/viro.2001.1105>

- 82** Collins PL, Hill MG, Cristina J, Grosfeld H. Transcription elongation factor of respiratory syncytial virus, a nonsegmented negative-strand RNA virus. *Proc. Natl. Acad. Sci. U S A.* **1996**; 93(1):81–5. <https://doi.org/10.1073/pnas.93.1.81>
- 83** Fearn R, Collins PL. Role of the M2-1 transcription antitermination protein of respiratory syncytial virus in sequential transcription. *J. Virol.* **1999**; 73(7):5852–64. <https://www.ncbi.nlm.nih.gov/pubmed/10364337>
- 84** Mason SW, Aberg E, Lawetz C, Whitehead P, Liuzzi M, Delong R. Interaction between Human Respiratory Syncytial Virus (RSV) M2-1 and P Proteins Is Required for Reconstitution of M2-1-Dependent RSV Minigenome Activity. *J. Virol.* **2003**;77(19):10670–6. <https://doi.org/10.1128/JVI.77.19.10670-10676.2003>
- 85** Cuesta I, Geng X, Asenjo A, Villanueva N. Structural phosphoprotein M2-1 of the human respiratory syncytial virus is an RNA binding protein. *J. Virol.* **2000**;74(21):9858–67. <https://doi.org/10.1128/jvi.74.21.9858-9867.2000>
- 86** Bermingham A, Collins PL. The M2-2 protein of human respiratory syncytial virus is a regulatory factor involved in the balance between RNA replication and transcription. *Proc. Natl. Acad. Sci. U S A.* **1999**; 96:11259–64. <https://doi.org/10.1073/pnas.96.20.11259>
- 87** Cheng X, Park H, Zhou H, Jin H. Overexpression of the M2-2 Protein of Respiratory Syncytial Virus Inhibits Viral Replication. *J. Virol.* **2005**; 79(22):13943–52. <https://doi.org/10.1128/JVI.79.22.13943-13952.2005>
- 88** Ahmadian G, Randhawa JS, Easton J. Expression of the ORF-2 protein of the human respiratory syncytial virus M2 gene is initiated by a ribosomal termination-dependent reinitiation mechanism. *EMBO J.* **2000**;19(11):2681–9. <https://doi.org/10.1093/emboj/19.11.2681>
- 89** Förster A, Maertens GN, Farrell PJ, Bajorek M. Dimerization of Matrix Protein Is Required for Budding of Respiratory Syncytial Virus. *J. Virol.* **2015**; 89(8):4624–35. <https://doi.org/10.1128/JVI.03500-14>
- 90** Ghildyal R, Baulch-Brown C, Mills J, Meanger J. The matrix protein of Human respiratory syncytial virus localises to the nucleus of infected cells and inhibits transcription. *Arch. Virol.* **2003**; 148:1419–29. <https://doi.org/10.1007/s00705-003-0112-y>
- 91** Peeples ME, Wang C, Gupta KC, Coleman N. Nuclear entry and nucleolar localization of the Newcastle disease virus (NDV) matrix protein occur early in infection and

do not require other NDV proteins. *J. Virol.* **1992**; 66(5): 3263–9. <https://www.ncbi.nlm.nih.gov/pubmed/1560547>

92 Ghildyal R, Ho A, Jans DA. Central role of the respiratory syncytial virus matrix protein in infection. *FEMS Microbiol. Rev.* **2006**; 30(5):692–705. <https://doi.org/10.1111/j.1574-6976.2006.00025.x>

93 Ghildyal R, Ho A, Dias M, Soegiyono L, Bardin PG, Tran KC, et al. The Respiratory Syncytial Virus Matrix Protein Possesses a Crm1-Mediated Nuclear Export Mechanism. *J. Virol.* **2009**; 83(11):5353–62. <https://doi.org/10.1128/JVI.02374-08>

94 Ouizougoun-oubari M, Pereira N, Tarus B, Galloux M, Lassoued S, Fix J, et al. A Druggable Pocket at the Nucleocapsid/Phosphoprotein Interaction Site of Human Respiratory Syncytial Virus. *J. Virol.* **2015**; 89(21): 11129-43 <https://doi.org/10.1128/JVI.01612-15>

95 Galloux M, Tarus B, Blazevic I, Fix J, Duquerroy S, Eléouët J-F. Characterization of a viral phosphoprotein binding site on the surface of the respiratory syncytial nucleoprotein. *J. Virol.* **2012**;86(16):8375–87. <https://doi.org/10.1128/JVI.00058-12>

96 Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J. Med. Chem.* **2006**; 49: 6177–96. <https://doi.org/10.1021/jm051256o>

97 Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**; 47:1750–9. <https://doi.org/10.1021/jm030644s>

98 Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**; 47:1739–49. <https://doi.org/10.1021/jm030643o>

99 SPECS. <http://www.specs.net/>

100 Molecular Operating Environment (MOE), 2015.10; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2016. <https://www.chemcomp.com/Products.htm>

101 Enamine Ltd. <http://www.enamine.net/>

- 102** Life chemicals. <http://www.lifechemicals.com/>
- 103** Turoverov K, Uversky V, Stepanenko O, Verkhusha V, Kuznetsova I. Fluorescent Proteins as Biomarkers and Biosensors: Throwing Color Lights on Molecular and Cellular Processes. *Curr. Protein Pept. Sci.* **2008**; 9(4): 338–69. <https://doi.org/10.2174/138920308785132668>
- 104** Gendron R, Francisco S, Ca US, Hughes AD. US 2012/0308587 A1.
- 105** Nudelman A, Gnizi E, Katz Y, Azulai R, Cohen-Ohana M. Prodrugs of butyric acid Novel derivatives possessing increased aqueous solubility and potential for treating cancer and blood diseases. *Eur. J. Med. Chem.* **2001**;36(1): 63-74. [https://doi.org/10.1016/S0223-5234\(00\)01199-5](https://doi.org/10.1016/S0223-5234(00)01199-5)
- 106** Mirabelli C, Jaspers M, Boon M, Jorissen M, Koukni M, Bardiot D, et al. Differential antiviral activities of respiratory syncytial virus (RSV) inhibitors in human airway epithelium. *J. Antimicrob Chemother.* **2018**; 73(7): 1823–9. <https://doi.org/10.1093/jac/dky089>
- 107** Cory AH, Owen TC, Barltrop JA, Cory JG. Use of an Aqueous Soluble Tetrazolium/Formazan Assay for Cell Growth Assays in Culture. *Cancer Commun.* **1991**; 3(7): 207–12. <https://doi.org/10.3727/095535491820873191>
- 108** El Ashry ESH, Ramadan ES, Abdel Hamid H, Hagar M. Microwave-Assisted Synthesis of Quinoline Derivatives from Isatin. *Synth. Commun.* **2005**; 35(17): 2243–50. <https://doi.org/10.1002/chin.200603153>
- 109** Li J, Chen J, Gui C, Zhang L, Qin Y, Xu Q, Zhang J, Liu H, Shen X, Jiang H. Discovering novel chemical inhibitors of human cyclophilin A. Virtual screening, synthesis and bioassay. *Bioorg. Med. Chem.* **2006**; 14(7): 2209-24. <https://doi.org/10.1016/j.bmc.2005.11.006>
- 110** Pritchett JC, Naesens L, Montoya J. Treating HHV-6 Infections. In: Human Herpesviruses HHV-6A, HHV-6B & HHV-7. *Diagnosis and Clinical Management* **2014**, 311-331 <https://doi.org/10.1016/C2011-0-08453-3>
- 111** Mierde HV, Van Der Voort P, De Vos D, Verpoort F. A Ruthenium-Catalyzed Approach to the Friedländer Quinoline Synthesis. *Eur. J. of Org. Chem.* **2008**; 9: 1625-1631. <https://doi.org/10.1002/ejoc.200701001>

- 112** Tourteau A, Andrzejak V, Body-malapel M, Lemaire L, Lemoine A, Mansouri R, et al. 3-Carboxamido-5-aryl-isoxazoles as new CB2 agonists for the treatment of colitis. *Bioorg. Med. Chem.* **2013**; 21(17):5383–94. <https://doi.org/10.1016/j.bmc.2013.06.010>
- 113** Venning ARO, Bohan PT, Alexanian EJ. Palladium-Catalyzed, Ring-Forming Aromatic C–H Alkylations with Unactivated Alkyl Halides. *J. Am. Chem. Soc.* **2015**; 137(11): 3731–4. <https://doi.org/10.1021/jacs.5b01365>
- 114** Dubovi EJ, Geratz JD, Tidwell RR. Inhibition of respiratory syncytial virus by bis(5-amidino-2-benzimidazolyl)methane. *Virology.* **1980**; 103:502–4. [https://doi.org/10.1016/0042-6822\(80\)90207-x](https://doi.org/10.1016/0042-6822(80)90207-x)
- 115** Razinkov V, Gazumyan A, Nikitenko A, Ellestad G, Krishnamurthy G. RFI-641 inhibits entry of respiratory syncytial virus via interactions with fusion protein. *Chem. Biol.* **2001**; 8(7): 645–59. [https://doi.org/10.1016/s1074-5521\(01\)00042-4](https://doi.org/10.1016/s1074-5521(01)00042-4)
- 116** Huntley CC, Weiss WJ, Gazumyan A, Buklan A, Feld B, Hu W, et al. RFI-641, a potent respiratory syncytial virus inhibitor. *Antimicrob. Agents Chemother.* **2002**; 46(3): 841–847. <https://doi.org/10.1128/aac.46.3.841-847.2002>
- 117** Morton CJ, Cameron R, Lawrence LJ, Lin B, Lowe M, Luttick A, et al. Structural characterization of respiratory syncytial virus fusion inhibitor escape mutants: homology model of the F protein and a syncytium formation assay. *Virology.* **2003**; 311(2): 275–88. [https://doi.org/10.1016/s0042-6822\(03\)00115-6](https://doi.org/10.1016/s0042-6822(03)00115-6)
- 118** Cianci C, Yu K, Combrink K, Sin N, Pearce B, Wang A, et al. Orally Active Fusion Inhibitor of Respiratory Syncytial Virus Orally Active Fusion Inhibitor of Respiratory Syncytial Virus. *Antimicrob. Agents Chemother.* **2004**; 48(2): 413–422. <https://doi.org/10.1128/aac.48.2.413-422.2004>
- 119** Cianci C, Langley DR, Dischino DD, Sun Y, Yu KL, Stanley A, Roach J, Li Z, Dalterio R, Colonno R, Meanwell NA, Krystal M. Targeting a binding pocket within the trimer-of-hairpins: small-molecule inhibition of viral fusion. *Proc. Natl. Acad. Sci. U S A.* **2004**; 101(42): 15046–15051. <https://doi.org/10.1073/pnas.0406696101>
- 120** Douglas JL, Panis ML, Ho E, Lin KY, Krawczyk SH, Grant DM, Cai R, Swaminathan S, Chen X, Cihlar T. Small molecules VP-14637 and JNJ-2408068 inhibit respiratory syncytial virus fusion by similar mechanisms. *Antimicrob. Agents Chemother.* **2005**; 49(6): 2460–2466. <https://doi.org/10.1128/AAC.49.6.2460-2466.2005>

- 121** Roymans D, De Bondt HL, Arnoult E, Geluykens P, Gevers T, Van Ginderen M, et al. Binding of a potent small-molecule inhibitor of six-helix bundle formation requires interactions with both heptad-repeats of the RSV fusion protein. *Proc. Natl. Acad. Sci.* **2010**; 107(1): 308–13. <https://doi.org/10.1073/pnas.0910108106> x
- 122** Lundin A, Bergström T, Bendrioua L, Kann N, Adamiak B, Trybala E. Two novel fusion inhibitors of human respiratory syncytial virus. *Antiviral Res.* **2010**; 88(3): 317–24. <https://doi.org/10.1016/j.antiviral.2010.10.004>
- 123** Yan D, Lee S, Thakkar VD, Luo M, Moore ML, Plemper RK. Cross-resistance mechanism of respiratory syncytial virus against structurally diverse entry inhibitors. *Proc. Natl. Acad. Sci.* **2014**; 111(33): e3441–3449. <https://doi.org/10.1073/pnas.1405198111>
- 124** Lambert DM, Barney S, Lambert AL, Guthrie K, Medinas R, Davis DE, Bucy T, Erickson J, Merutka G, Petteway SR Jr. Peptides from conserved regions of paramyxovirus fusion (F) proteins are potent inhibitors of viral fusion. *Proc. Natl. Acad. Sci. U S A.* **1996**; 93: 2186–91. <https://doi.org/10.1073/pnas.93.5.2186>
- 125** Whitby LR, Boyle KE, Cai L, Yu X, Gochin M, Boger DL. Discovery of HIV fusion inhibitors targeting gp41 using a comprehensive α -helix mimetic library. *Bioorganic Med. Chem. Lett.* **2012**; 22(8): 2861–5. <https://doi.org/10.1016/j.bmcl.2012.02.062>
- 126** Azzarito V, Long K, Murphy NS, Wilson AJ. Inhibition of α -helix-mediated protein-protein interactions using designed molecules. *Nat. Chem.* **2013**; 5(3): 161–73. <https://doi.org/10.1038/nchem.1568>
- 127** Cummings CG, Hamilton AD. Disrupting protein-protein interactions with non-peptidic, small molecule α -helix mimetics. *Curr. Opin. Chem. Biol.* **2010**; 14(3): 341–6. <https://doi.org/10.1016/j.cbpa.2010.04.001>
- 128** Azzarito V, Miles JA, Fisher J, Edwards TA, Warriner SL, Wilson AJ. Stereocontrolled protein surface recognition using chiral oligoamide proteomimetic foldamers. *Chem. Sci.* **2015**; 6(4): 2434–43. <https://doi.org/10.1039/c4sc03559c>
- 129** Xin D, Ko E, Perez LM, Ioerger TR, Burgess K. Evaluating minimalist mimics by exploring key orientations on secondary structures (EKOS). *Org. Biomol. Chem.* **2013**; 11(44): 7789–801. <https://doi.org/10.1039/c3ob41848k>
- 130** Lao BB, Drew K, Guarracino DA, Brewer TF, Heindel DW, Bonneau R, Arora PS. Rational design of topographical helix mimics as potent inhibitors of protein-protein interactions. *J. Am. Chem. Soc.* **2014**; 136(22): 7877–88. <https://doi.org/10.1021/ja502310r>

- 131** Murphy NS, Prabhakaran P, Azzarito V, Plante JP, Hardie MJ, Kilner CA, Warriner SL, Wilson AJ. Solid-phase methodology for synthesis of O-alkylated aromatic oligoamide inhibitors of α -helix-mediated protein-protein interactions. *Chemistry* **2013**; 19(18): 5546–50. <https://doi.org/10.1002/chem.201204098>
- 132** Jayatunga MK, Thompson S, Hamilton AD. α -Helix mimetics: Outwards and upwards. *Bioorganic Med. Chem. Lett.* **2014**; 24(3): 717–24. <https://doi.org/10.1016/j.bmcl.2013.12.003>
- 133** Wilson AJ. Helix mimetics: Recent developments. *Prog. Biophys. Mol. Biol.* **2015**; 119(1): 33–40. <https://doi.org/10.1016/j.pbiomolbio.2015.05.001>
- 134** Soulère L, Aldrich C, Daumke O, Gail R, Kissau L, Wittinghofer A, Waldmann H. Synthesis of GTP-Derived Ras Ligands. *Chembiochem.* **2004**; 5(10): 1448–53. <https://doi.org/10.1002/cbic.200400133>
- 135** Ahn J, Han S. Facile synthesis of benzamides to mimic an α -helix. *Tetrahedron Letters* **2007**; 48(20): 3543–3547 <https://doi.org/10.1016/j.tetlet.2007.03.108>
- 136** Ahn J-M. US 2010/0178324 A1.
- 137** Waghmare A, Hindupur R, Pati H. Propylphosphonic anhydride (T3P®): An expedient reagent for organic synthesis. *Review Journal of Chemistry* **2014**, 4(2): 53–131 <https://doi.org/10.1134/S2079978014020034>
- 138** Leggio A, Belsito EL, Luca G De, Gioia ML Di, Leotta V, Romio E. One-pot synthesis of amides from carboxylic acids activated using thionyl chloride. *RSC Adv.* **2016**; 6: 34468–34475. <https://doi.org/10.1039/C5RA24527C>
- 139** Eccles KS, Elcoate CJ, Lawrence SE, Maguire AR. Convenient and robust one-pot synthesis of symmetrical and unsymmetrical benzyl thioethers from benzyl halides using thiourea. *Arkivoc.* **2010**; IX: 216–228. <https://doi.org/10.3998/ark.5550190.0011.921>
- 140** Schrödinger Release 2014-4: Maestro, version 10.0, Schrödinger, LLC, New York, NY, 2014.
- 141** Tosco P, Balle T, Shiri F. Open3DALIGN: An open-source software aimed at unsupervised ligand alignment. *J. Comput. Aided. Mol. Des.* **2011**; 25(8): 777–783. <https://doi.org/10.1007/s10822-011-9462-9>

Appendix

All compounds were purchased from SPECS, Enamine Ltd and Life Chemicals Inc.

Table 1

Compound number	Chemical structure	Compound number	Chemical structure
171		177	
172		178	
173		179	
174		180	
175		181	
176		182	

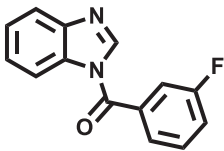
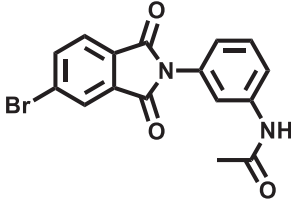
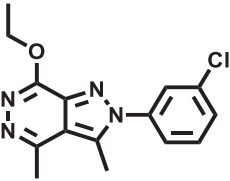
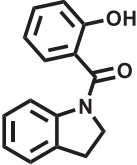
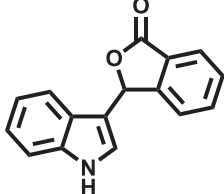
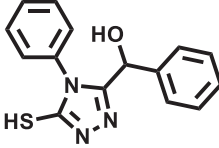
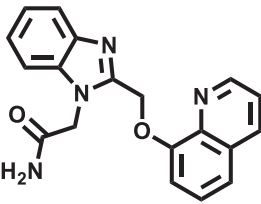
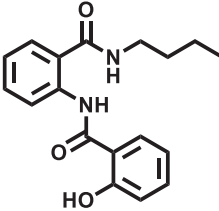
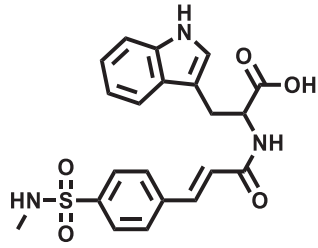
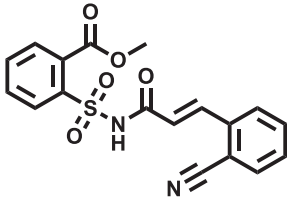
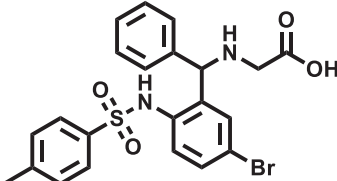
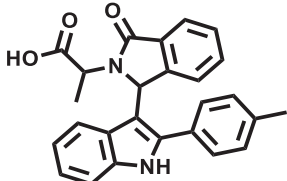
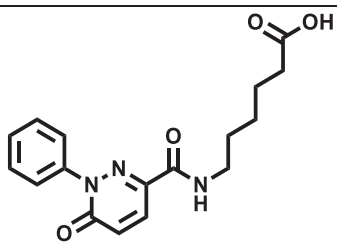
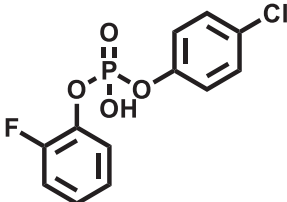
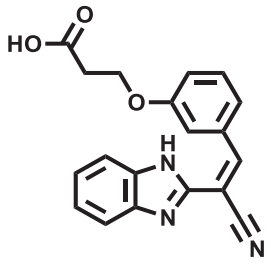
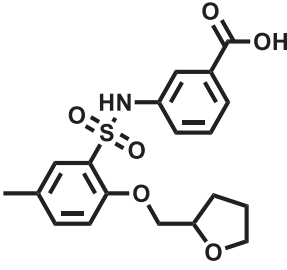
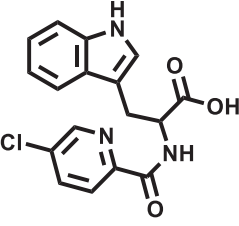
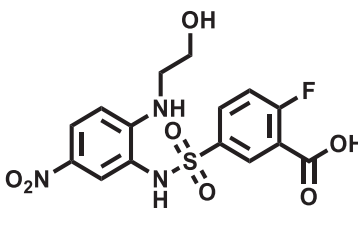
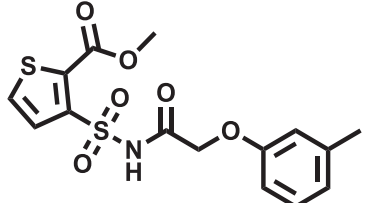
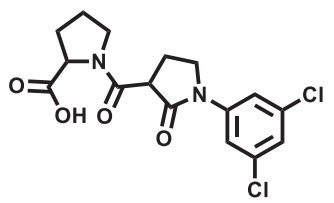
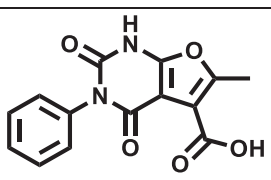
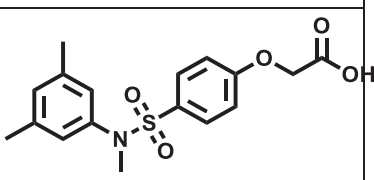
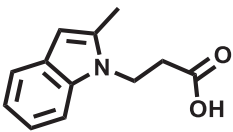
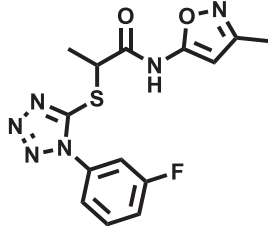
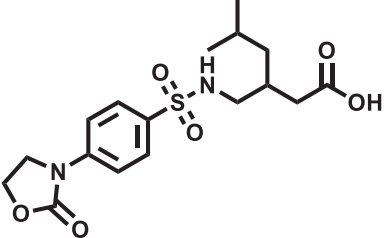
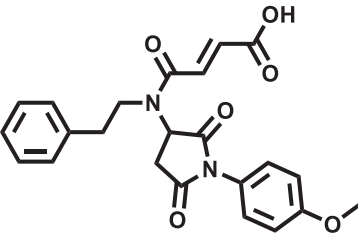
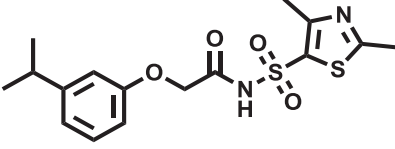
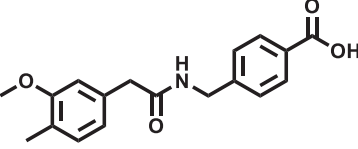
183		187	
184		188	
185		189	
186		190	

Table 2

Compound number	Chemical structure	Compound number	Chemical structure
191		198	
192		199	
193		200	
194		201	
195		202	
196		203	
197		204	

205		208	
206		209	
207		210	

PART III

Chapter 7: Enterovirus

7.1 Enterovirus

Enteroviruses are part of the Picornaviridae family. They are positive single-stranded RNA, non-enveloped virus member of the Group IV in the Baltimore classification, in which the viral genome is inside an icosahedral capsid. Poliovirus, discovered in 1908, is the most studied enterovirus, and it is the etiological cause of Poliomyelitis, also known as Polio, an infectious disease of the central nervous system. The Poliovirus is affecting mainly children, and during the infection it spread and damage the motor neurons, leading in the worst cases in total paralysis. Since the 1960s, in which the Sabin's vaccine against the three Poliovirus serotype was licensed, an extensive vaccine campaign to eradicate Polio was performed. Today, poliovirus is declared eradicated in the developed countries but remains a battle in Nigeria, Syria, Afghanistan and Pakistan. Furthermore, at the beginning of 2018, an outbreak of Polio was reported in Papua New Guinea, a county defined Polio-free since 2000.¹

The three serotypes Poliovirus, with the other Non-Polio Enteroviruses (NPEVs), are part of the 260 serotypes among the seven human pathogens species in the enterovirus genus. One hundred of those serotypes can be divided into Enterovirus A-D and the other 160 into Rhinovirus A-C. Those NPEVs have emerged in the last decades as severe treats for public health.

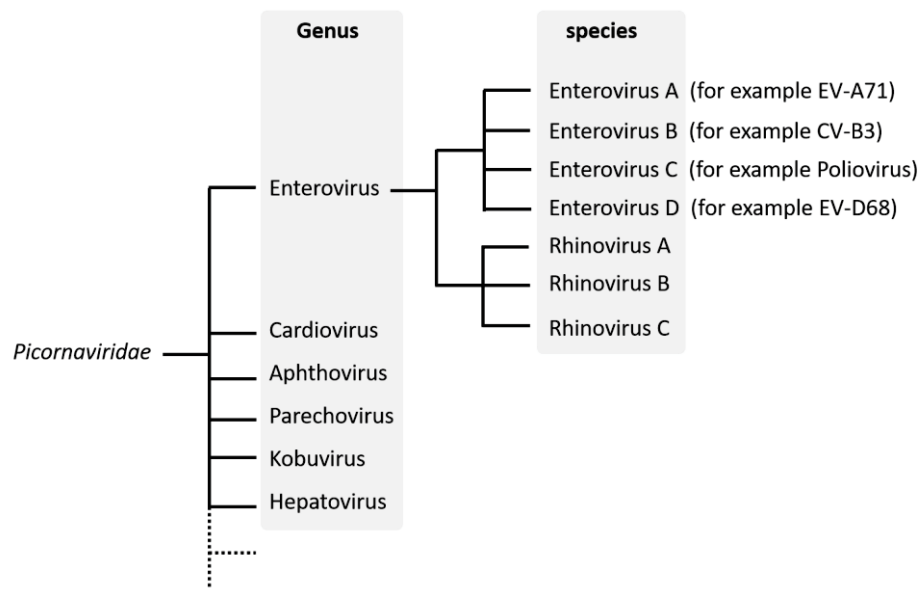


Figure 1. Enterovirus genus and human pathogen species. graphic representation of the Non-Polio Enteroviruses (NPEVs) species, in which the seven human pathogens species (Enterovirus A-D and Rhinovirus A-C) include 260 different viral serotypes.

Few examples are EVA-71, the primary cause of Hand-foot-and-mouth disease (HFMD). EVD-68, causes severe respiratory disease and Acute Flaccid Paralysis (AFP) and cranial nerve dysfunction in children. Coxsackieviruses are associated with aseptic meningitis but

also related to the aetiology of type 1 diabetes. Rhinoviruses are the cause of the common cold, but they can also be related to asthma and development of chronic obstructive pulmonary disease (COPD).²

Enteroviruses transmission occurs through the faecal-oral or respiratory routes and during the infection they can spread and infect other different tissues and organs and, as said before, they can infect the central nervous system. Except for the Anti-polio vaccines and two recently approved vaccine against EVA-71 in China, there is no antiviral therapy available on the market.³⁻⁴

7.2 Viral Genome organisation and life-cycle

The viral genome of enterovirus is ~7.5 kb long positive single-stranded RNA. The genome can be divided into three parts: 5'-untranslated region (5'UTR), the coding region and a 3'-untranslated region (3'UTR) (figure 2).

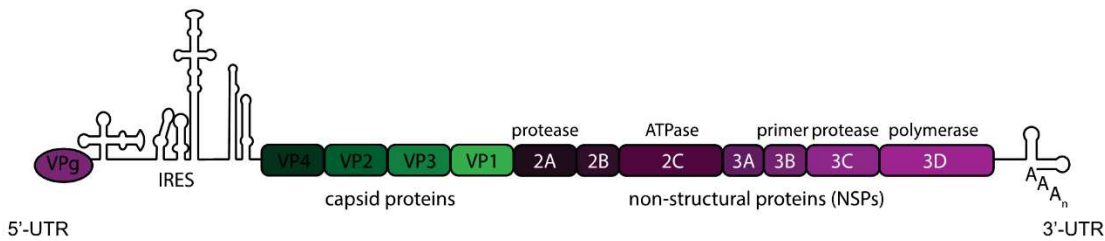


Figure 2. Enterovirus genome organisation. Schematic representation of the enterovirus genome which encodes 11 viral proteins. At the 5'UTR contains six stem-loop intramolecular base pairing structures, the internal ribosome entry site (IRES) and it is covalently bounded to the VPg protein, a protein essential for the 3Dpol. At the 3'UTR the viral genome has three stem-loop structures which are involved in the viral replication (Designed by Denise Seitner).

The 5'UTR contains six stem-loop intramolecular base pairing structures which are involved in the viral RNA synthesis and include internal ribosome entry site (IRES) which is essential for the viral protein translation.⁵ The 5'UTR is covalently bound to the VPg protein, which is the primer for the viral RNA- dependent RNA polymerase 3Dpol. The coding region has only one open reading frame (ORF) and is encoding for a ~2,100 amino acids polyprotein. The viral polyprotein is successively cleaved by the viral proteinase 2Apro, 3Cpro and 3CDpro into four capsid proteins and six non-structural proteins (NSPs) and several functional cleavage intermediates. The last part of the viral genome is 3'UTR. This region is highly conserved and has three stem-loop structures which are essential for the viral replication.

As said before the polyprotein is cleaved in ten proteins that can be four structural and six non-structural proteins. The four structural proteins, VP1-4, are the structural subunit of the

icosahedral capsid of the virion. VP1-3 are on the surface of the capsid, VP4 instead is internalised in the virion. Those proteins are the functional unit for the assembly of the viral capsid of the new virions. The enterovirus life-cycle, like for the other viruses, can be divided into three main stages: 1) Entry process; 2) Translation and genome replication; 3) Virion assembly and release (figure 3).⁴

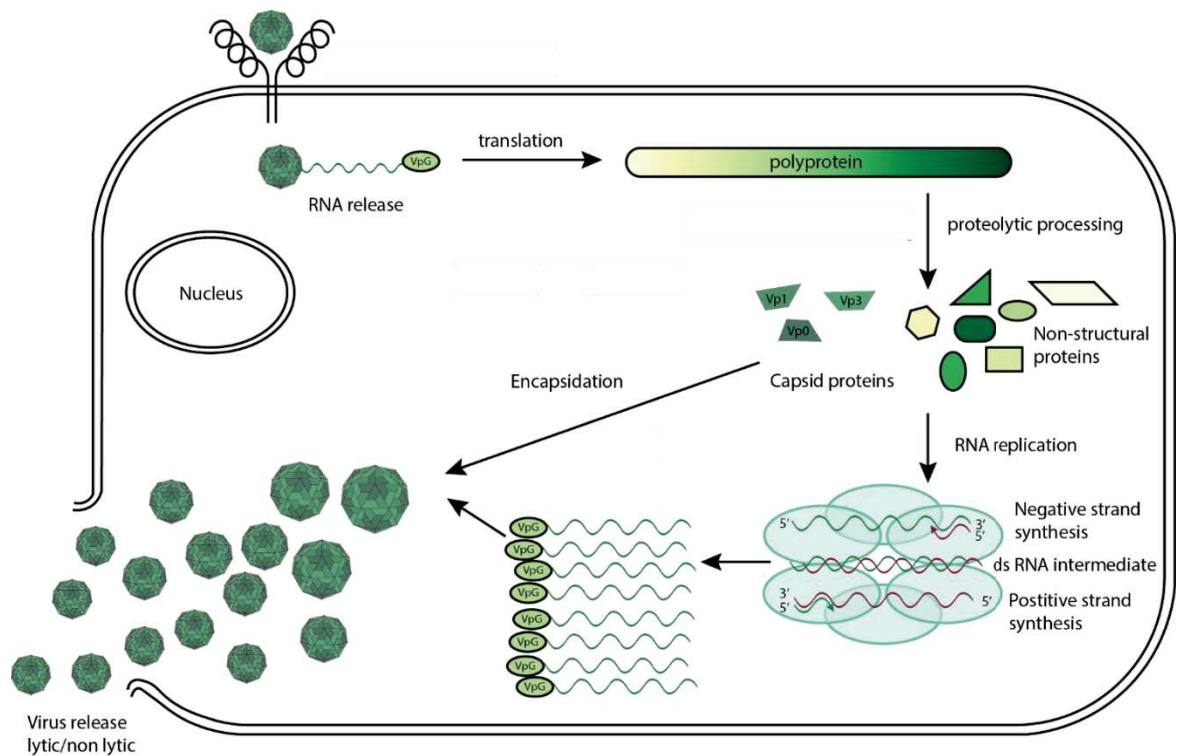


Figure 3. Viral cycle representation. The viral life-cycle of enterovirus can be divided into three steps: 1) entry process; 2) translation and genome replication; 3) virion assembly and release. The viruses enter the cell through two different mechanisms, both mediated by the viral capsid proteins: by binding to the cell membrane or by endocytosis. Once the virus enters or is bonded to the cell, the viral genome is released by the viral capsid protein rearrangement. The viral proteins, produced by the auto-proteolytic process of the polyprotein, which is obtained by the translation of the viral genome, are involved in the formation of the replication organelles (ROs). The ROs are essential for the viral genome replication. Once the viral genome is replicated and the structural proteins are produced, the new virions are auto-assembled and released by lytic or non-lytic mechanisms. (Designed by Denise Seitner)

Entry process

The virus to begin the infectious life-cycle must release the viral genome inside the host cell. For doing this, the virus binds to specific receptors expressed on the cell surface. Several different NPEV receptors have been recognised, explaining the various spectrum of diseases caused by enteroviruses. Those receptors can be classified as uncoating or attachment receptors, in which the first type is triggering the rearrangement of the capsid's

proteins to release the viral genome. The second type is promoting the endocytosis of the viral particle. Then in the endosome, the low pH is stimulating the capsid's proteins uncoating process. Both pathways end with the release of the viral genome into the cytosol of the host cell.

Translation and genome replication

Now that the viral genome is in the cytosol, it is translated by the host cell ribosomes into the polyprotein. As soon as the viral proteins are produced by the auto-proteolytic process of the polyprotein, the production of new virion can start. The newly produced non-structural proteins 3A and 2BC intermediate, together with several host-cell proteins, are driving the formation of the replication organelles (ROs), in which the viral genome is replicated by RNA-dependent RNA polymerase 3Dpol. The viral polymerase is working with 2C, 3CDpro and 3AB for the synthesis of a negative RNA strand as a double-strand RNA genome and act as a template for the viral genome replication.

Virion assembly and release

Now that the viral genome and the structural protein have been produced, the virion assembly stage can begin. The VP0 (precursor of VP2 and VP4), VP1 and VP3 structural proteins, formed through the proteolytical activity of the non-structural protein 3CDpro, are self-oligomerising around the newly replicated viral-RNA. Then the viral-RNA is cleaving the VP0 into VP2 and VP4 to give the mature virion. The lysis of the cell releases the new virions which are now able to infect other cells and organs.

7.3 Therapeutic strategies

There is an unmet need for antiviral drugs for NPEV and Polio infections, due to the lack of antiviral therapy on the market. Several compounds were evaluated in clinical trials to treat rhinovirus infections.

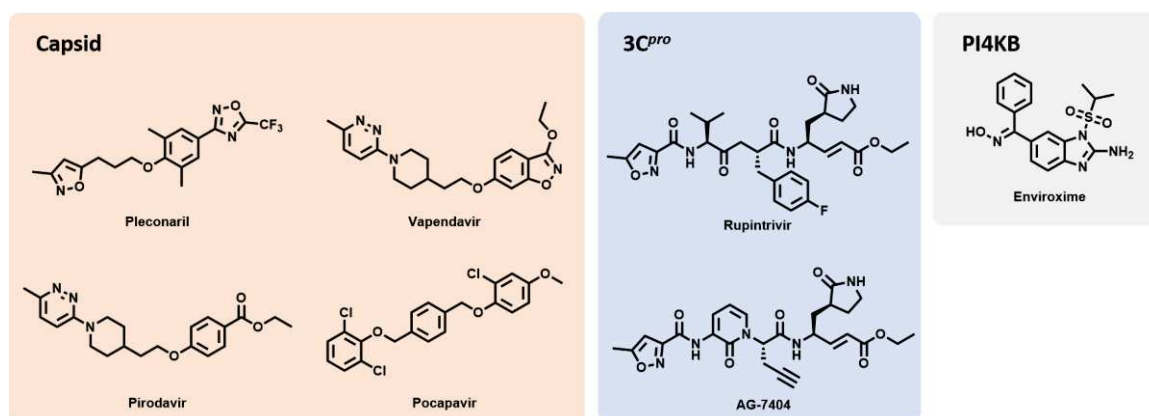


Figure 4. Enteroviruses Drug candidates. In the figure are reported the chemical structures of the drug candidates in clinical trials for the treatment of rhinovirus infections. On the left side of the figure are shown four capsid inhibitors: Pleconaril, Vapendavir, Pirodavir and Pocapavir. In the middle two inhibitors of the viral protease 3C^{pro}: Rupintrivir and AG-7404. On the right side, the chemical structure of Enviroxime a PI4KB inhibitor.

Mainly these compounds were capsid inhibitors, able to inhibit the release process of the viral genome by blocking the uncoating process of the virus. Other compounds were protease inhibitors targeting the viral protease 3C^{pro}. Another compound which was clinically evaluated was Enviroxime. This compound inhibits the phosphatidylinositol 4-kinase- β (PI4KB), a host factor which is crucial in the formation of the ROs for the viral genome replication. Those compounds never reached the market, because they did not show efficacy or low toxicity.^{4,6}

During the last decades, several compounds were reported to be able to inhibit the enteroviruses replication by targeting the previously described target but also other targets, either viral or host factors. Among the viral targets, several compounds were identified through drug repurposing screenings, to inhibit the 3D^{pol} RNA-dependent-RNA polymerase and others the non-structural protein 2C, which is taking part in several steps of the viral replication life-cycle. The inhibitions of another host factor target, the oxysterol-binding protein (OSBP) which is involved together with PI4KB in the formation of the Ros, was reported to block the viral replication. The inhibition of three host factors was reported to be useful as antiviral strategies, interfering with the viral uncoating or the capsid assembly. The inhibition of the *N*-myristoyltransferase 1 and the depletion of Glutathione was demonstrated to interfere with the uncoating process. Heat shock protein 90 (HSP90),

involved with the assembly process of the virions, was demonstrated to interfere with the capsid assembly.⁷

7.4 Non-structural protein 2C

Among the human enterovirus species, the non-structural 2C protein, encoded in the P2 region, is highly conserved, and it is composed of 329 amino acids. This protein was extensively studied during the years, different interacting and functional regions on the 2C protein have been identified (figure 5).

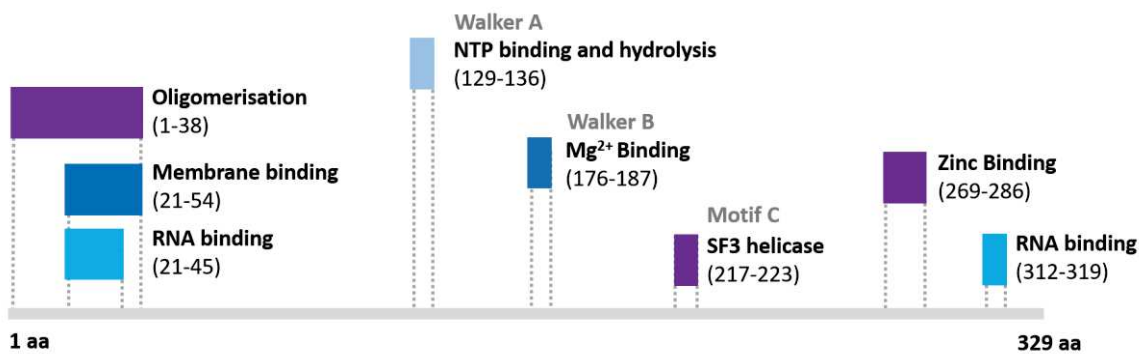


Figure 5. Representation of the interacting and functional motifs on the 2C protein. The non-structural 2C protein has RNA binding domains between the residues 21-54 and 312-319 and a zinc binding domain between 269-286 residues. The viral 2C protein has also highly conserved motifs, Walker A, Walker B and the motif C, which are part of the ATPase domain. Between the residues 1-54 of the N-TD, the protein has two regions which are essential for the interaction with the membranes and for the oligomerisation.

A region crucial for the interaction with the membranes and the oligomerisation was identified between 21-54 residues of the N-terminal domain.⁸⁻¹⁰ This region, together with another region, close to the C-terminal domain (312-319 residues), were identified as RNA binding domains.¹¹⁻¹² The zinc binding motif, another region identified in the CTD and located between 269-286 residues, seem to be involved either in interaction with other protein or with RNA.¹³ The 2C protein also has two highly conserved motifs which are part of the ATPase domain. The first motif, also called Walker A, located between 129 and 136, is involved in the binding with the ATP and hydrolysis of the ATP into ADP, related with energy release. The second motif, called Walker B, is involved in the coordination of the Mg²⁺.¹⁴⁻¹⁶ The ATPase domain of 2C was demonstrated to be crucial for the RNA replication process and proliferation of enteroviruses.^{14,16-17}

The protein has a highly conserved helicase motif, called motif C, located between the 217-223 residues.¹⁸ For these three conserved motifs, Walker A, Walker B and motif C, the 2C protein can be classified as a member of the superfamily3 helicases.¹⁹ The 2C protein, for

which was observed hexameric oligomerisation, was associated with RNA helicases and ATP-independent RNA chaperones function.²⁰⁻²³ The 2C protein, by directly interacting with other viral and host proteins, is also essential for the rearrangement of intracellular membranes, the formation of ROs and the viral genome encapsidation.²⁴⁻²⁸

Crystal structure of the non-structural 2C protein

In 2017, Guan et al. released a truncated X-ray crystal structure of the EV71 2C protein, consisting of an ATPase domain, a zinc finger, and a long C-terminal α -helix.

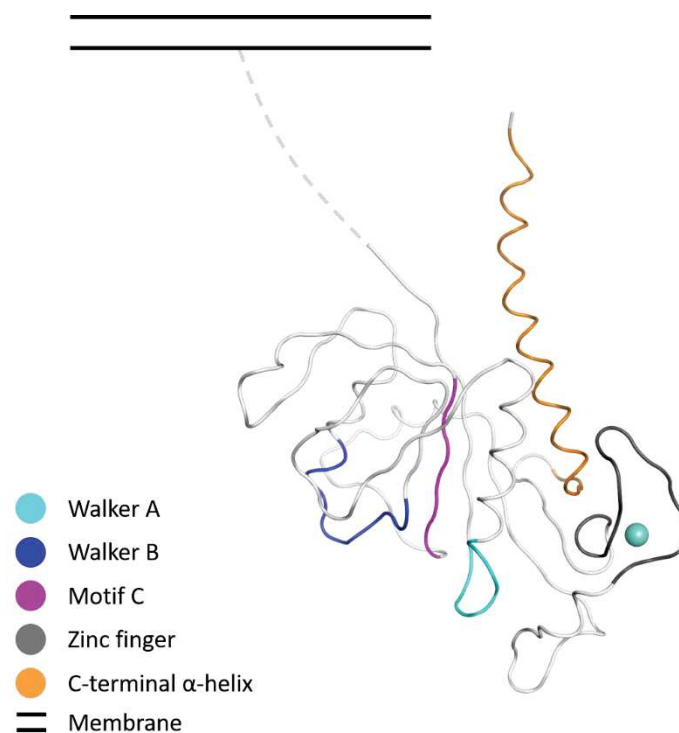


Figure 6. Representation EV-A71 2C protein. Ribbon representation of the crystal structure of the truncated non-structural 2C protein (116-329 residues, PDB ID:5GRB-chain A). Walker A, Walker B and the motif C are respectively represented in cyan, blue and purple. Zinc finger motif is represented in grey and the C-terminal α -helix in orange.

The ATPase domain is composed of five parallel β strands, forming a β -sheet, surrounded by three α -helices, forming the α/β Rossman fold, typical of enzymes binding to nucleotide cofactors. The Walker A, Walker B motifs and motif-C are located in the ATPase domain, as shown in figure 6. The Zinc finger domain is formed by the 267 to 298 residues, and the zinc atom is coordinated by three cysteine residues, C270, C281 and C286 respectively, and the carbonyl of S282 backbone. The last domain is the long C-terminal α -helix, formed by the residues 299-329. This α -helix seems to be involved in the stabilisation of the hexameric configuration of the 2C proteins and the residues 312-319 is known to interact with the RNA.

Even if the crystal of EV-A71 2C protein were obtained by removing the N-terminal domain (1-116 residues), which is responsible for the membrane binding and oligomerisation, the polymerisation of the 2C proteins was observed when co-crystallised with adenosine 5'-O-(3-thiotriphosphate (ATP- γ -S)). The ATP- γ -S occupies the pocket formed by the neighboring chain A and chain F. The triphosphate group of the ATP- γ -S is interacting with the G132, G134, K135, S136 (Walker A motif) and with D176, D177 (Walker B motif) of chain A.

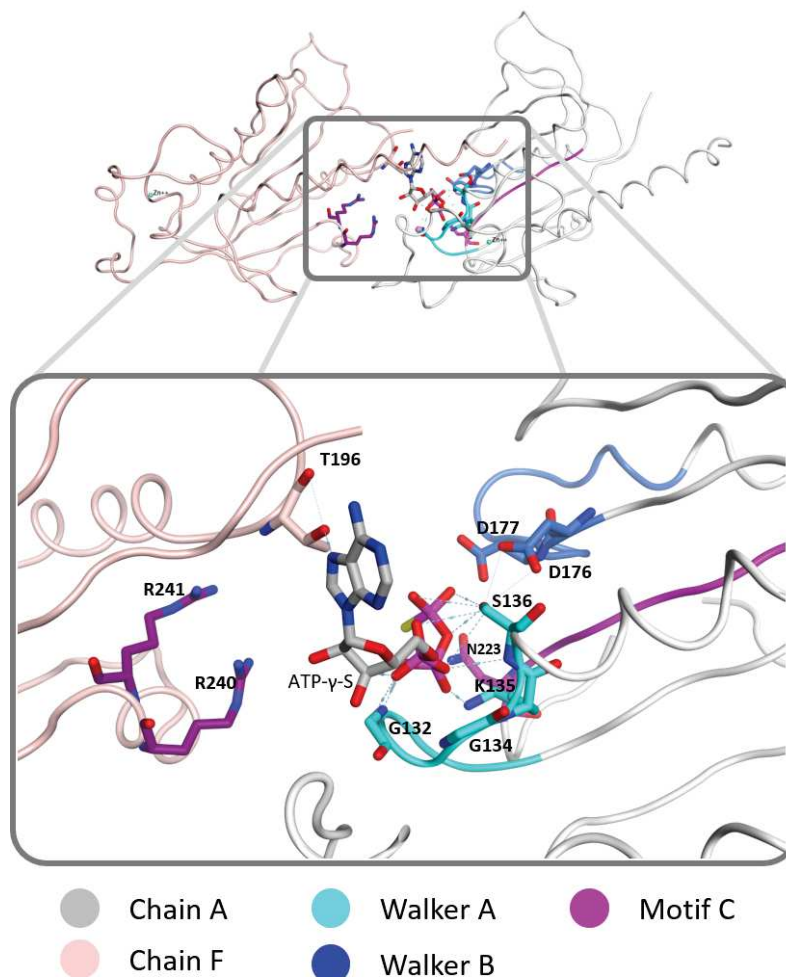


Figure 7. ATP- γ -S binding to the ATPase pocket formed. In the top part of the figure, the structure of two neighbouring 2C proteins and the ATP- γ are shown. In the bottom part of the figure, the ATPase binding pocket and the amino acids involved in the interaction with the ATP- γ are shown. Walker A, Walker B and the motif C are respectively represented in cyan, blue and purple

The N223 of motif C is located behind the Walker B motif. The chain F is interacting with the ATP- γ -S through a hydrogen bond between the T196 of the ATP- γ -S base. The chain F is also providing two arginines, close to the phosphate group, the R240 and the R241. These two arginines, through mutation studies, were demonstrated to be essential for the ATPase activity, suggesting that they might function as arginine finger during the hydrolysis.²⁹

The released structure of the EV-A71 2C protein represents the starting point for the understanding of the mode of action of the known 2C inhibitors and the development of new inhibitors using structure-based drug design approaches.

7.5 Non-structural protein 2C inhibitors

Over the past decades, several compounds have been identified to inhibit the viral replication of different enterovirus species. From the isolation and sequencing analysis of resistant-enterovirus mutant to these compounds, mutations on the 2C protein were observed, suggesting the 2C protein as a target for these inhibitors.

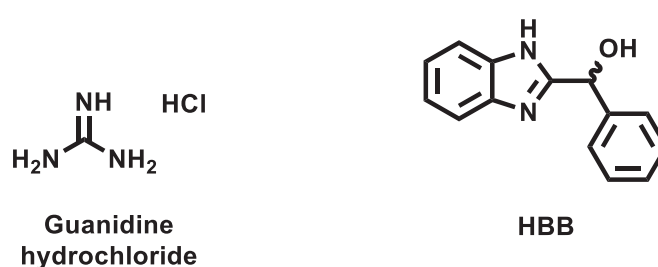


Figure 8. Chemical structure of guanidine and 2-(α -Hydroxybenzyl)-Benzimidazole (HBB).

The first two inhibitors which were identified to target the 2C protein, were guanidine and the 2-(α -Hydroxybenzyl)-Benzimidazole (HBB). Both of the compounds were found to be active against poliovirus-1 (PV-1), echovirus-9 (E-9), coxsackievirus A9 (CVA9) and CVB3.³⁰⁻³⁵ Guanidine was able to block the initiation step of the negative-sense RNA but not the replication of the positive one and at higher concentration, guanidine was inhibiting the ATP hydrolysis.³⁶⁻³⁸ The compound was also reported to inhibit the association of the protein precursor, 2BC, with the membrane structures.³⁹

Also, for those two compounds, dependent-mutant were identified. These mutants required the presence of guanidine or HBB, to be able to replicate correctly.⁴⁰

Other benzimidazole were found to target the 2C protein, the {1-(2,6-difluorophenyl)-6-trifluoromethyl-1H,3H-thiazolo[3,4-a]benzimidazole (TBZE-029) and the 1-(4-fluorophenyl)-2-(4-imino-1,4-dihydropyridin-1-yl)methylbenzimidazole (MRL-1237) (figure 9).^{18,41} These two compounds were able to inhibit the viral replication of CVB3 and showed cross-resistance to the human enterovirus-B species (HEV-B) resistant to guanidine and HBB. Most of the mutations were found to be located within the three ATPase domain and especially on a loop, 224-229 residues.^{15,30,32,42-43}

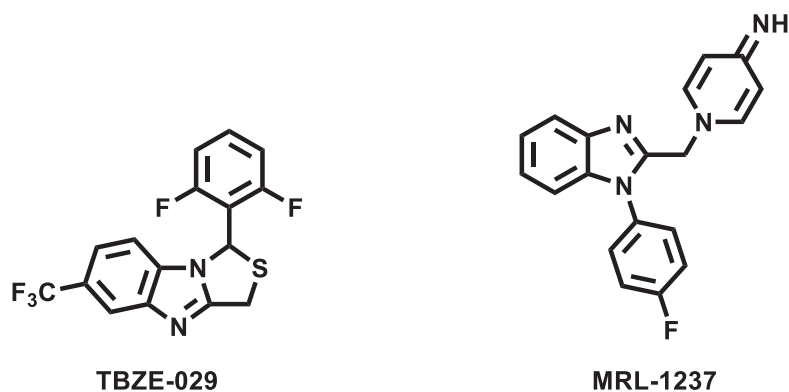
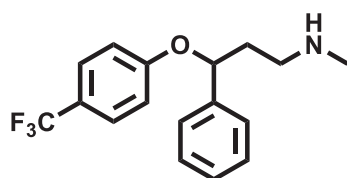


Figure 9. Chemical structure of TBZE-029 and MRL-1237.

In 2012, during a screening of a library of FDA-approved drugs, fluoxetine ((*RS*)-*N*-methyl-3-phenyl-3-(4-(trifluoromethyl)phenoxy)propan-1-amine), was identified as a potent inhibitor of CVB3.⁴⁴



Fluoxetine

Figure 10 Chemical structure of fluoxetine

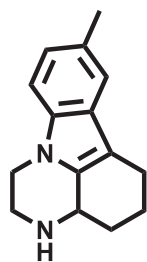
In the same study, norfluoxetine, the active metabolite of fluoxetine, was achieving a comparable antiviral activity. Fluoxetine was then confirmed to inhibit HEV-B and D species replication but not the HEV-A and HEV-C, and during the study, the same mutation that conferred resistance to TBZE-029 were conferring cross-resistance to fluoxetine.

Interestingly, one of the CVB3 mutant, the A224V-I227V-A229V, was cross-resistance to guanidine, HBB and MRL-1237 and the mutations are summarised in table 1.⁴⁵

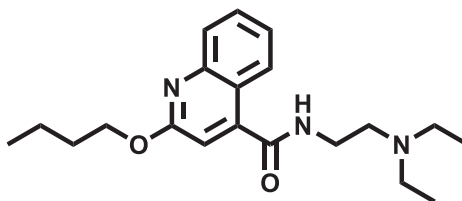
Table 1. Mutation and cross-resistant mutant for Guanidine. TBZE-029, MRL-1237 and fluoxetine against EV-A71 and CVB3. +++: high resistance; ++: moderate resistance; +: low resistance; /: no resistance; Dep: dependent for viral replication; No Dep: not dependent for viral replication; -: not evaluated; NA: not active.

		Mutation	Fluoxetine	MRL-1237	Gua HCl	HBB	TBZE-029	Ref
Gua HCl	EV-A71	I127V		-	++	-	-	ref 46
		T173M		-	-	-	-	
		M187I	NA	-	-	-	-	
		M193L		-	+++	-	-	
		S224C		-	-	-	-	
		M193L + I127V		-	+++	-	-	
TBZE-029	CVB3 Nancy	A224V	/	/	++	/	/	ref 18
		I227V	+++	++	/	/	/	
		A229V	No Dep	Dep	Dep	Dep	Dep	
		A224V + I227V	+++	+++	++	++	+++	
		A224V + A229V	No Dep	Dep	Dep	Dep	Dep	
		I227V + A229V	/	+++	++	++	++	
		A224V + I227V + A229V	+++	+++	+++	+++	+++	
MRL-1237	CVB3 Nancy	A224V	-	++	/	-	-	ref 18
		I227V	-	+++	+++	-	-	

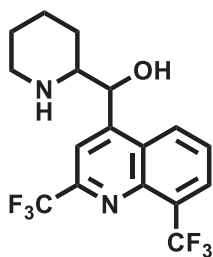
Pirlindole, dibucaine, mefloquine and zuclopenthixol were also found, during another screening of FDA-approved compounds, to be able to inhibit the viral replication of CVB3 and EV-D68. Pirlindole and dibucaine also showed a weak antiviral activity against EV-A71. Cross-resistance of the A224V-I227V-A229V 2C CVB3 mutant to Pirlindole, dibucaine and zuclopenthixol was observed, suggesting that also these compounds are targeting the 2C protein.⁴⁷⁻⁴⁸



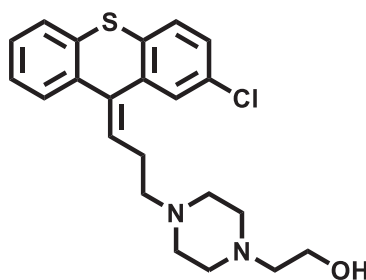
Pirlindole



Dibucaine



Mefloquine



Zuclopenthixol

Figure 11. Chemical structure of pirlindole, dibucaine, mefloquine and zuclopenthixol

7.6 Project aims

The 2C protein is highly conserved 2C protein among the enterovirus species and is essential during the viral replication. Several compounds have been identified to inhibit viral replication by targeting 2C protein. Among them, fluoxetine, a selective serotonin reuptake inhibitor, was showing a strong antiviral effect against HEV-B and D species whereas HEV-A and C species are not inhibited. The information available on the mutation-sensible to fluoxetine, together with the new resolved EVA 71 2C crystal structure were used to investigate the region of the protein close to the fluoxetine resistant-mutations.

Below, the activities performed during the project on the non-structural 2C protein:

- Different molecular modelling techniques were used to investigate and to elucidate the binding mechanism of fluoxetine;
- Different fragments and analogues of fluoxetine were synthesised and biologically evaluated;
- Mutation studies were performed to further explore the potential binding mechanism of fluoxetine;
- The gained information on fluoxetine were used to design pan-enterovirus inhibitors;
- The designed compounds were synthesised and biologically evaluated.

Furthermore, different molecular modelling techniques will be used to perform a virtual screening in order to identify new potential inhibitors of the ATPase pocket. This project, on the 2C protein, was carried out in close collaboration with Lisa Bauer, PhD student in the group of Professor Frank J.M. van Kuppeveld at Utrecht University, and Birgit Zonsics, PhD student in the group of Professor Andrea Brancale.

Chapter 8: Fluoxetine

8.1 Fluoxetine

Fluoxetine was the blockbuster drug of Eli Lilly. Introduced on the market in 1987 under the market name Prozac, it was the first selective serotonin reuptake inhibitor (SSRI) approved by the FDA. It is used for the treatment of major depression and anxiety disorders, but it is also indicated for different disorders: bulimia, obsessive-compulsive disorder, bipolar and panic disorders.⁴⁹

In 2012, during a drug-repurposing screening, fluoxetine was identified as an inhibitor of the viral replication of Coxsackievirus B-3 (CVB-3), member of the Human enterovirus B (HEV-B) genus, with an EC_{50} of 2.3 μ M and a CC_{50} of 25 μ M, showing a selectivity index of approximately 10 in HeLa-RW cells.⁴⁴ Fluoxetine, in further studies, showed an antiviral activity also against HEV-D, but not against HEV-A, HEV-C and rhinovirus species.⁵⁰ Afterwards, mutations on the CVB-3 non-structural protein 2C were reported to confer resistance to fluoxetine. These mutations were close to the Walker C motif and were on the ²²⁴AGSINA²²⁹ loop (A224V, I227V and A229V) which, is highly conserved among the HEV-B and HEV-D species but not the HEV-A and HEV-C.⁵¹ Thus, the antiviral activity shown by fluoxetine may be a result of direct interaction with the non-structural 2C protein. The fact also confirmed the specific interaction of fluoxetine with 2C protein, that other SSRIs like citalopram, paroxetine, and sertraline, did not show any antiviral effect on CVB-3.⁴⁵

In 2106, fluoxetine was successfully used to treat a child with X-linked agammaglobulinemia (XLA), affected by life-threatening chronic enterovirus encephalitis, caused by CVB-2.⁵²

However, the association between the use of SSRIs and an enhanced risk of bleeding, together with other non-neurological and pharmacological off-target effects of fluoxetine are limiting the use as an anti-enterovirus agent, especially in newborns.⁵³⁻⁵⁵

That evidence is making fluoxetine the perfect candidate for the elucidation of the molecular mechanism of the 2C inhibitors, sensible to the mutation on ²²⁴AGSINA²²⁹ loop and able to inhibit viral replication. The elucidation of the molecular mechanism of inhibition of 2C protein by fluoxetine will provide essential information for the development of potential enterovirus pan-inhibitors.

The FDA-approved SSRI Fluoxetine is a racemic mixture of (*RS*)-*N*-methyl-3-phenyl-3-(4-(trifluoromethyl) phenoxy)propan-1-amine, in which both enantiomers were demonstrated to have the same SSRI activity, in several in vitro and in vivo preclinical models.⁵⁶ The two fluoxetine enantiomers were purchased and biologically evaluated to understand if the antiviral activity of fluoxetine is independent of the stereochemical configuration.

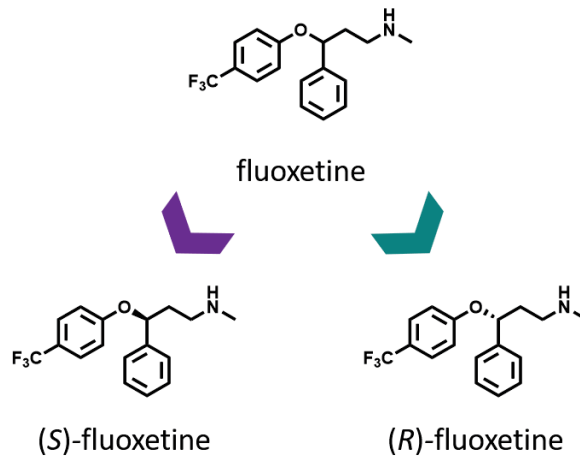


Figure 12. Fluoxetine enantiomers.

8.2 Antiviral evaluation of fluoxetine enantiomers

The antiviral evaluation of the compounds was performed in Virology Division, Department of Infectious Diseases and Immunology at Utrecht University, under the supervision of Professor Frank J.M. van Kuppeveld. The fluoxetine racemate and both the enantiomers were tested in multicycle CPE-reduction assay in Hela R19 cell-line infected with CVB-3 virus. In this assay, the cytopathic effect induced by the virus after three days and the related cell viability is used to calculate the EC_{50} of the tested compound. The cell viability is evaluated using MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium), as a readout for the cellular metabolic activity. The uninfected cell, treated with the compounds, are used for the evaluation of the CC_{50} , using the MTS as a readout of the assay. The racemic mixture inhibited CVB3 with an EC_{50} of $3.2 \pm 0.95 \mu\text{M}$ and a CC_{50} of $29.32 \pm 0.35 \mu\text{M}$, with a selectivity index of 9. Those results were in line with the previously reported EC_{50} and CC_{50} on CVB-3 (figure 13).⁵⁰

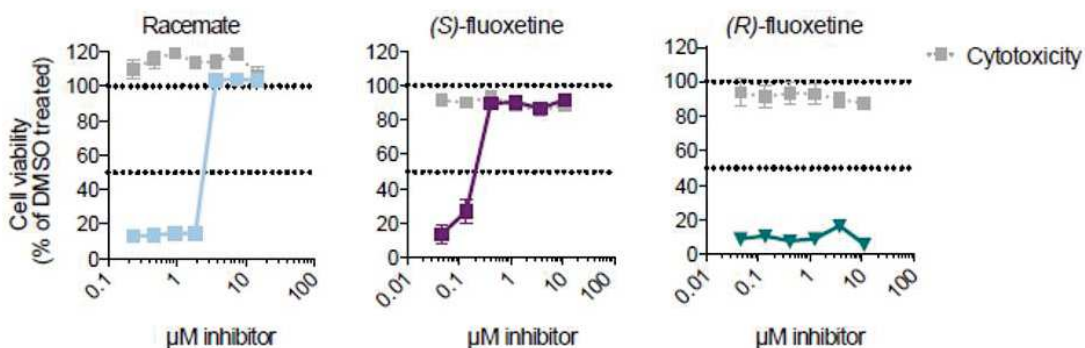


Figure 13. Antiviral evaluation of fluoxetine enantiomers. Both fluoxetine enantiomers and the racemate were tested in a multicycle CPE-reduction assay in Hela R19 cell-line infected with CVB-3 virus. MTS assay was used for the readout. The same assay was used on uninfected cells for the evaluation of CC_{50} . Data are shown from one experiment representative of at least two independent experiments.

Interestingly the S-enantiomer inhibited the viral replication with an EC_{50} of $0.4 \pm 0.15 \mu\text{M}$ and a CC_{50} of $28.63 \pm 1.40 \mu\text{M}$. On the other hand, the R-enantiomer of fluoxetine did not show any antiviral activity.

To confirm the results obtained for the compounds by the multicycle CPE-reduction assay, the compounds were evaluated in a single-cycle assay. In this assay, the HeLa R19 cells are infected with RLuc-CVB3. The RLuc-CVB3 is a genetically engineered Coxsackievirus B-3, in which a gene encoding for *Renilla luciferase* is inserted upstream of the capsid coding region. Seven hours post-infection, the intracellular activity of the luciferase proteins, produced during the viral replication, it is used as a sensitive readout for the calculation of the antiviral activity of the tested compounds.

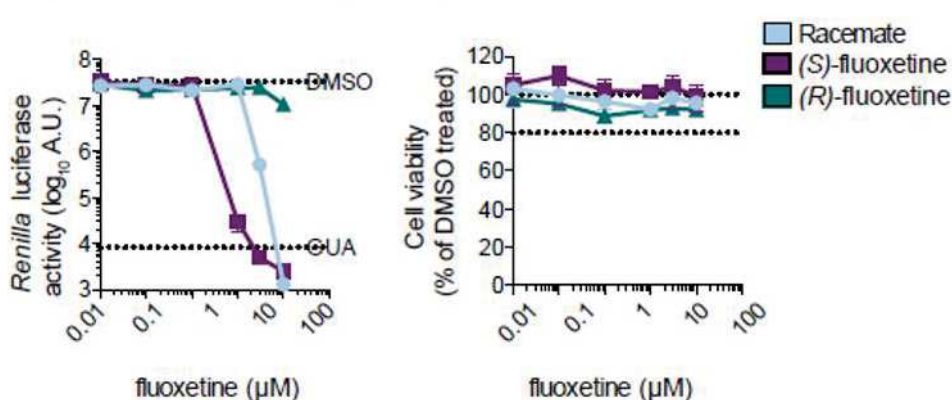


Figure 14. Antiviral evaluation of fluoxetine enantiomers by single-cycle assay. Both fluoxetine enantiomers and the racemate were tested in a multicycle CPE-reduction assay in HeLa R19 cell-line infected with RLuc-CVB3. Data are shown from one experiment representative of at least two independent experiments.

As shown in figure 14, the single-cycle assay confirmed that only the S-enantiomer inhibit the viral replication, showing a potency ~5-fold higher ($EC_{50}=0.42 \pm 0.17 \mu\text{M}$) than the racemic mixture ($EC_{50}=2.02 \pm 0.94 \mu\text{M}$).

8.2.1 *In-vitro* binding of fluoxetine enantiomers to 2C protein

The *in-vitro* binding assays of the compounds were performed in the CNRS at Aix-Marseille Université, under the supervision of Professor Bruno Coutard. To test the potential selectivity of the (S)-fluoxetine as an antiviral agent, both enantiomers were evaluated in two different binding assays, thermal shift assay (TSA) and isothermal titration calorimetry (ITC). These assays were performed on a monomeric 2C protein, generated by removing the amphipathic helix at the N-terminal domain (CV-B3 2C Del36).⁴⁷

In the thermal shift assay (TSA), the binding of a ligand to the protein is affecting the melting temperature (T_m) of the protein which is directly related with a variation of the thermal stability of the protein. The racemic mixture showed a dose-dependent increase of the ΔT_m

of 2C protein from 10 μM to 250 μM was observed; then at higher concentrations, the ΔT_m decreased (figure 15).

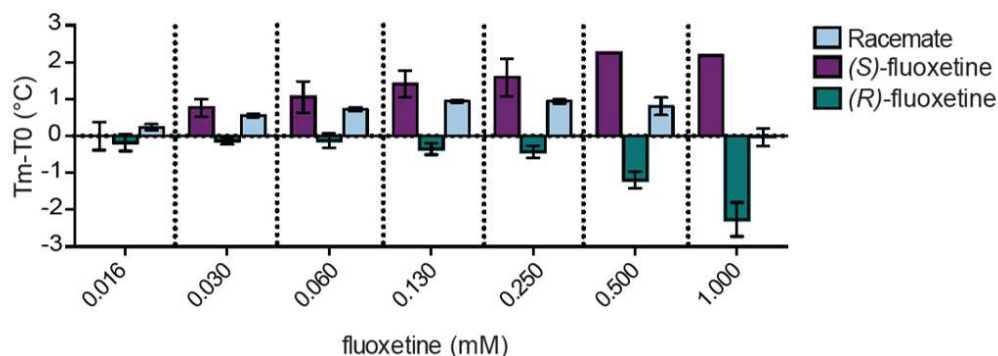


Figure 15. TSA evaluation of the fluoxetine enantiomers. The (S)-fluoxetine shown to thermally stabilize the 2C protein and the (R)-fluoxetine showed a dose-dependent destabilizing effect on the protein. Data are shown from one experiment representative of at least three independent experiments

The (S)-fluoxetine shown to thermally stabilize the 2C protein with no destabilization at high concentrations of compound, suggesting a direct binding of the enantiomer to the 2C protein. On the other hand, the (R)-fluoxetine has a dose-dependent destabilizing effect on the protein which could explain the effect observed for the racemic mixture at a concentration >250 μM .

Then, the binding of the fluoxetine enantiomers was assessed using the isothermal titration calorimetry (ITC) technique. This technique is extensively used to evaluate the protein-ligand interactions, but also for other macromolecule interactions. The ITC is a quantitative technique that can be used to determine the thermodynamics of binding interactions (variation of enthalpy, ΔH) and the kinetics of enzyme-catalyzed reactions (dissociation equilibrium constant K_d).⁵⁷

A K_d of $\sim 9.5 \mu\text{M}$ was obtained for the (S)-fluoxetine, confirming the binding of the S-enantiomer to 2C protein. Instead, when the 2C protein was titrated with (R)-fluoxetine, the aggregation of the protein was observed, and it was not possible to obtain a precise K_d value.

8.2.2 Antiviral evaluation of fluoxetine enantiomers against other enteroviruses

Fluoxetine, as a racemic mixture, was previously reported to inhibit CVB3 (strain Nancy) and EV-D68 (strain Fermon), which represent the HEV-B and HEV-D species.^{44,50}

After finding that the (S)-fluoxetine, among the fluoxetine enantiomers, is the only responsible for the antiviral activity of the racemic mixture, was decided to test the compounds on a panel of representative enteroviruses species (EV-A71, CVB3, EV-D68, PV-1) and rhinoviruses (HRV-A2, HRV-B14).

Table 2. Antiviral evaluation of the fluoxetine enantiomers against other enterovirus species. Data represents mean values \pm SD calculated from three independent experiments and are expressed in μ M. NA = not active. SI = selectivity index (CC_{50}/EC_{50}).

	EC ₅₀ EV-A71 (BrCr)	EC ₅₀ CVB3 (Nancy)	EC ₅₀ PV-1 (Sabin1)	EC ₅₀ EV-D68 (Fermon)	EC ₅₀ HRV- A2	EC ₅₀ HRV- B14	CC ₅₀
(RS)- fluoxetine	NA	2.02 \pm 0.52	NA	1.85 \pm 0.10	NA	NA	29.32 \pm 0.35
(S)- fluoxetine	NA	0.42 \pm 0.17	NA	0.67 \pm 0.22	7.95 \pm 0.39	6.34 \pm 1.02	28.63 \pm 1.40
(R)- fluoxetine	NA	NA	NA	NA	NA	NA	23.63 \pm 1.40
SI_{Racemic}	-	14.51	-	21.72	-	-	
SI_{(S)-fluoxetine}	-	71.56	-	42.73	3.60	4.52	

As expected, fluoxetine showed antiviral activity only against CVB3 and EV-D68 species, and neither the fluoxetine racemic mixture nor the enantiomers showed antiviral activity against EV-A71 and Poliovirus-1 (PV-1) at 30 μ M (see table 2). The previously observed enhanced antiviral effect of (S)-fluoxetine against CVB3, was also observed against EV-D68, with an EC₅₀ of 0.67 \pm 0.22 μ M.

The compounds were also tested against rhinovirus A2 (HRV-A2) and rhinovirus B14 (HRV-B14). Only the (S)-fluoxetine was able to inhibit the viral replication of the two rhinoviruses, with an EC₅₀ 7.95 \pm 0.39 on HRV-A2 and 6.34 \pm 1.02 on HRV-B14. Taking in consideration the results obtained testing the (S)-fluoxetine against the two species of rhinovirus, the antiviral activity of (S)-fluoxetine against the other enteroviruses tested, EV-A71 and PV-1, at a concentration higher than the CC₅₀ cannot be excluded.

8.3 Identification of fluoxetine binding site

The mode of action and the interactions between the fluoxetine and the 2C protein are unknown. The stereospecific antiviral activity of the (S)-fluoxetine and the reported mutation on the ²²⁴AGSINA²²⁹ loop of 2C protein, which confers resistance to the compound are suggesting a specific interaction between the (S)-fluoxetine and the 2C protein. Several in-silico modelling approaches were used to elucidate the fluoxetine binding site and

eventually the mode of action. Unfortunately, the crystal structure of CVB3-2C protein, which is sensitive to the (S)-fluoxetine, is not available. In 2017, a truncated (116-329 amino acids) crystal structures of the EV-A71 2C protein was released.²⁹ The EV-A71 2C protein has a sequence identity and similarity with the CVB3 2C protein of 62% and 80%, respectively. The EV-A71 2C protein crystal structure, which is non-sensible to the (S)-fluoxetine, was used as a starting point for molecular modelling studies, to understand the interaction between the protein and the compound.

8.3.1 Molecular modelling studies

The molecular modelling studies were done in collaboration with Birgit Zonsics, PhD student in Professor Brancale's group, which provided the homology models and performed the docking of the compounds. Chain A of the published EV-A71 2C structure (PDB: 5GRB) was used as a template to generate the homology model of CVB3 2C protein. The generation of the homology model was necessary because the EV-A71 2C protein is not sensible to (S)-fluoxetine, and the crystal structure of EV-A71 2C protein is an ideal starting point for the generation of the homology model.

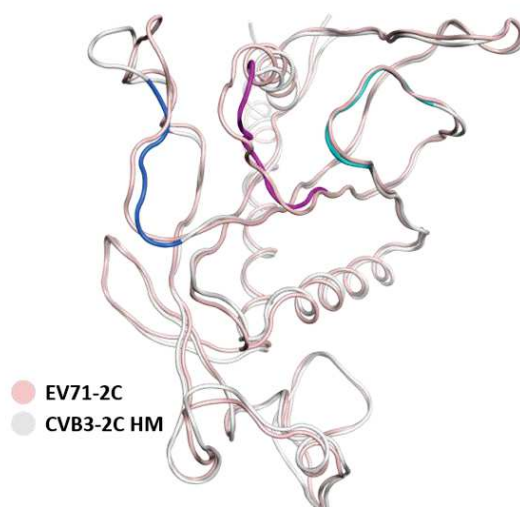


Figure 16. Homology model of the 2C protein of CVB3. The model was built on the crystal structure of EV-A71. Ribbon of ²²⁴AGSINA²²⁹ loop are in blue, the 175-183 loop are in violet. In pink the ribbon of non-structure 2C protein of EV71 and in grey the homology model of CVB3.

As can be seen in figure 17, around the ²²⁴AGSINA²²⁹ loop, two potential binding sites were identified (site A and site B in figure 17) on the homology model of CVB3 2C protein. Site A is formed by the ²²⁴AGSINA²²⁹ loop and hydrophilic residues (D245, R295 and R296). The site is also composed by a lipophilic pocket with the potentiality to bind the

trifluoromethylphenyl moiety of fluoxetine. Site B, on the other hand, is located between the ²²⁴AGSINA²²⁹ loop and the flexible 175-183 loop part of Walker B motif, which is the motif responsible for binding to Mg²⁺ to assure proper function of the ATPase activity of the 2C protein.

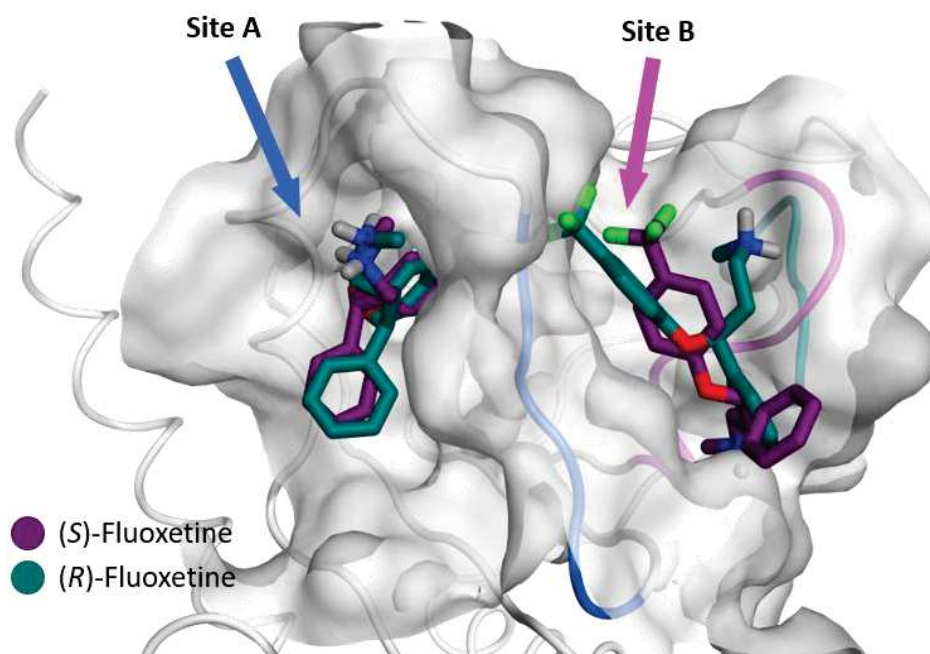


Figure 17. *S- and R-enantiomers of fluoxetine docked into sites A and B of the homology model. Ribbon of 224AGSINA229 loop are in blue, the 175-183 loop are in violet. (S)-fluoxetine in purple and the (R)-fluoxetine in bottle green.*

Due to the flexible nature of the loops, the binding of the compound to these two sites located around the ²²⁴AGSINA²²⁹ loop could lead to variation in the shape and size of the binding pocket. Molecular Dynamics (MD) studies were decided to be used to assess the change of the loops during the interaction with the pockets.

Both enantiomers of fluoxetine were docked in site A and site B, as can be seen in figure 12, and then used to perform the molecular dynamic (MD) simulations. Both enantiomers docked in site A has the trifluoromethylphenyl moiety interacting in the hydrophobic pocket. The hydrophobic pocket was defined by the residues L178, C179, V187 and F190. In site B, instead, the enantiomers were interacting superficially and were exposed to the solvent. The best docking pose, showed in figure 17, were used as a starting point for the MD simulations.

MD simulations are a useful tool to address various issues, such as loop flexibility and the potential shape changes that can occur during the interaction between the protein and a ligand. Three independent molecular dynamics simulations of 100 ns were performed, for

each enantiomer for each site (see figure 18), to investigate the behaviour of the fluoxetine enantiomers, binding to the two potential binding sites.

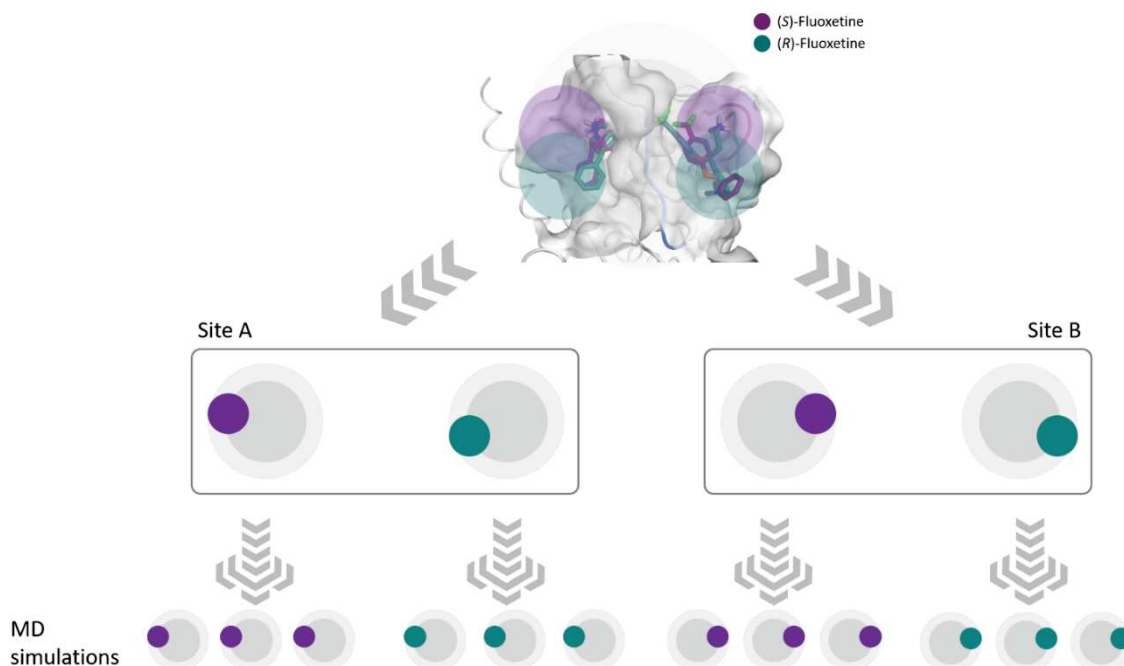


Figure 18. Scheme of the MD simulations performed on site A and site B. For each site and both enantiomers a MD simulation was performed. MD simulation were performed in triplicates.

To estimate the binding affinity, to potentially discriminate and identify the binding pocket of (S)-fluoxetine, molecular mechanics energies combined with generalised Born and surface area continuum solvation (MM/GBSA) method was applied to each MD simulation.

$$G = E_{\text{binding}} + E_{\text{el}} + E_{\text{vdW}} + G_{\text{pol}} + G_{\text{np}} - TS$$

The free-energy (G) is obtained considering conformational energy terms (the MM part, represented by E_{binding} , E_{el} and E_{vdW} , considering the potential energy of a molecular structure), and a solvation free-energy term, which is the sum of a polar component and a non-polar contribution ($G_{\text{pol}} + G_{\text{np}}$).

This method can be located between the docking method and the free-energy perturbation (FEP). The first one is extremely efficient, but not accurate. It is a perfect tool to discriminate between binding and non-binding ligands. The other method, the FEP, is very accurate but it is a very time-consuming method.

The $\Delta G_{\text{binding}}$, calculated through the MM/GBSA method, is obtained after splitting the trajectory file of the MD simulation into snapshots. for each snapshot is calculated the free energy of the binding of the ligand: ⁵⁸

$$\Delta G_{\text{binding}} = E_{\text{complex}} - (E_{\text{ligand}} + E_{\text{receptor}})$$

From the results obtained analysing the MD simulations of (*S*)-fluoxetine, a stronger binding on the site A was suggested, as can be seen in table 3, with a lower $\Delta G_{\text{binding}}$ in comparison with site B. During the MD simulation of the (*S*)-fluoxetine in site B, dissociation of the ligand from the protein was observed.

Table 3. Binding energies of the protein ligand complexes during MD simulations. The indicated values are calculated $\Delta G_{\text{binding}}$ (kJ/mol) * indicate dissociation of the ligand from the protein.

	Site	MD1	MD2	MD3
(S)-fluoxetine	A	-41.63	-42.28	-41.54
	B	-34.79	-29.88	-30.83*
(R)-fluoxetine	A	-29.71*	-42.70	-19.85*
	B	-29.59*	-42.95	-27.38

The (*R*)-fluoxetine analysis of the $\Delta G_{\text{binding}}$ is suggesting that the binding to the protein is not stable, indeed in both site A and B, dissociation of the R-enantiomer was observed. This result is in line with the lack of antiviral activity of the R-enantiomer of fluoxetine against the 2C protein of CVB3.

In each MD simulations of the (*S*)-fluoxetine in site A, the compound is oriented with the trifluoromethylphenyl moiety inside a hydrophobic pocket, defined by the residues L126, L178, C179, V187, F190, I227, A229, L238, F242. A hydrogen bond repeatedly formed between the amino acid D245 and the positively charged amino group of (*S*)-fluoxetine was also observed, as can be seen in figure 19.

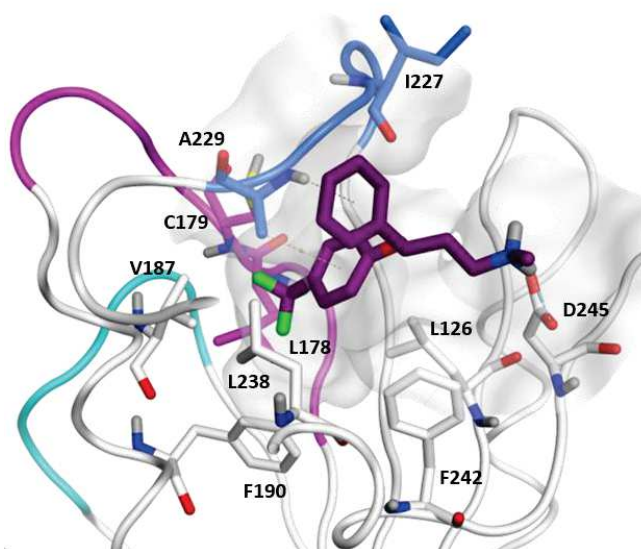


Figure 19. View of (*S*)-fluoxetine in site A as identified in the MD simulation. Ribbon and carbon atoms of ²²⁴AGSINA²²⁹ loop are in blue, the 175-183 loop are in violet.

The MD simulations revealed a stereospecific binding selectivity for the (S)-fluoxetine to 2C protein and the identified pocket, composed by a hydrophobic part and a potential binding point for the positively charged amino group. The speculated binding pocket could also be supported by the fact that the triple mutation A224V-I227V-A229V (AVIVAV) on the ²²⁴AGSINA²²⁹ loop, in which the residues A229 is interacting with the ligand, was previously shown to provide resistance to the racemic fluoxetine mixture.⁵⁰

8.3.2 Mutations studies on the identified pocket

For the validation of the identified binding pocket, several CVB3 mutant viruses were made, and those viruses were used for antiviral evaluation of the (S)-fluoxetine. BF738735, known to be a broad-spectrum anti-enterovirus which is acting as an inhibitor of the cellular protein PI4KIII β , was used as a reference.⁵⁹

The first step was to understand the single contribution of the three reported mutations on the A224V-I227V-A229V (AVIVAV) on the ²²⁴AGSINA²²⁹ loop. For this reason, the mutant of CVB3 virus, containing the single mutations (A224V, I227V and A229V) were made. As can be deduced from figure 20, the only amino acid giving resistance to the (S)-fluoxetine is the CVB3 I227V mutant.

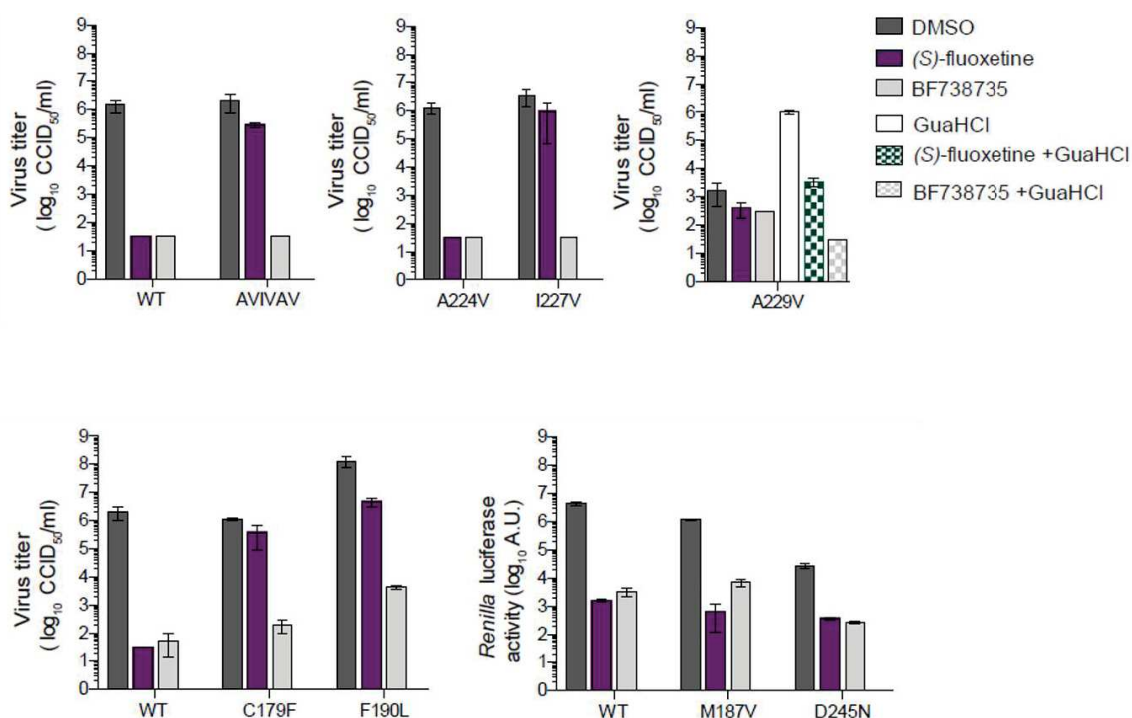


Figure 20 Antiviral evaluation of (S)-fluoxetine against the CVB3 mutant viruses. Data are shown from one experiment representative of at least two independent experiments.

Previously, the dependency of the CVB3 A229V mutant to several compounds to were reported. These compounds were, Guanidine, HBB, TBZE-029 and MRL-1237, did not

inhibit the mutant virus but instead, they were required for efficient viral replication.¹⁸ The test of the (S)-fluoxetine against the CVB3 A229V mutant, conferred neither dependency nor resistance. However, when the mutant virus was treated with both (S)-fluoxetine and Guanidine a reduction of the viral replication, in comparison to the Guanidine treated virus, was observed.

Two other mutant viruses of CVB3 were made to investigate the critical residues inside the hydrophobic pocket. These mutants were C179F, V187M and F190L and they raise resistance to the (S)-fluoxetine, except for the V187M mutation.

Finally, the importance of D245 amino acid was investigated. As was said before, this residue should be able to interact with the amino group on the (S)-fluoxetine. The substitution of the glutamic acid to an asparagine did not confer resistance.

The result obtained is suggesting that (S)-fluoxetine binding might be different from the predicted one because the D245N and A229V are not raising resistance. On the other hand, the C179F and F190L mutations are confirming the interaction of the trifluoromethylphenyl moiety within the hydrophobic pocket. Another possibility is that the (S)-fluoxetine could reach the C179 and F190 through the 158-164 loop. However, in the homology model and the EV-A71 crystal structure, a potential pocket between the 175-183 loop (violet ribbon in figure 21) and the 158-164 loop (cyan ribbon in figure 21) was not present. A possible pocket, generating between these two loops, make the explanation of the mutation on the ²²⁴AGSINA²²⁹ loop (blue ribbon in figure 21) more complicated.

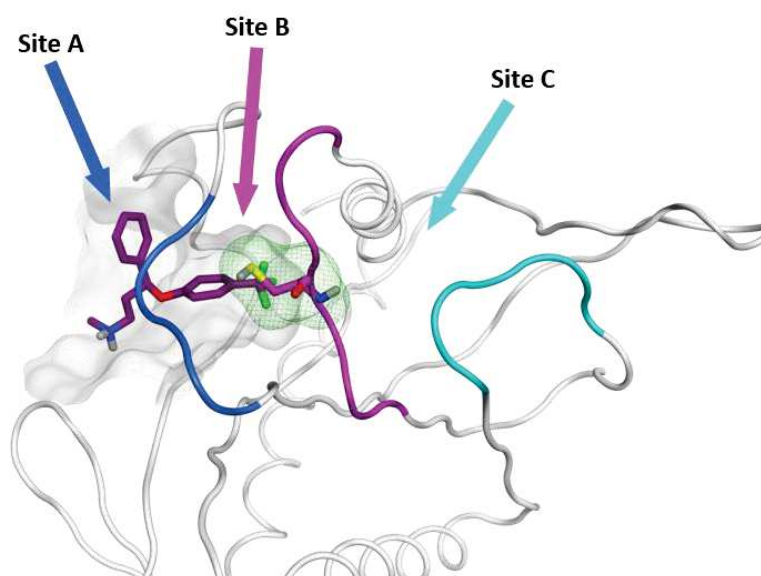


Figure 21 The three possible entrances of (S)-fluoxetine. C179 is represented in green line surface. Ribbon and carbon atoms of ²²⁴AGSINA²²⁹ loop are in blue, the 175-183 loop are in violet and the 158-164 loop are in cyan.

8.4 Fluoxetine fragments

Maira Lorenzo Lopez, an Erasmus student in Professor Brancale's group, has contribute towards the synthesis of the compounds of this part of the project. The contribution of the different chemical features on the fluoxetine were explored by testing six fragments. Those fragments, which were not all commercially available like *N*-methyl-3-(4-(trifluoromethyl)phenoxy)propan-1-amine **211** and 1-(benzyloxy)-4-(trifluoromethyl)benzene **214** or quite expensive like the *N*-methyl-3-phenylpropan-1-amine **212**, were synthesised.

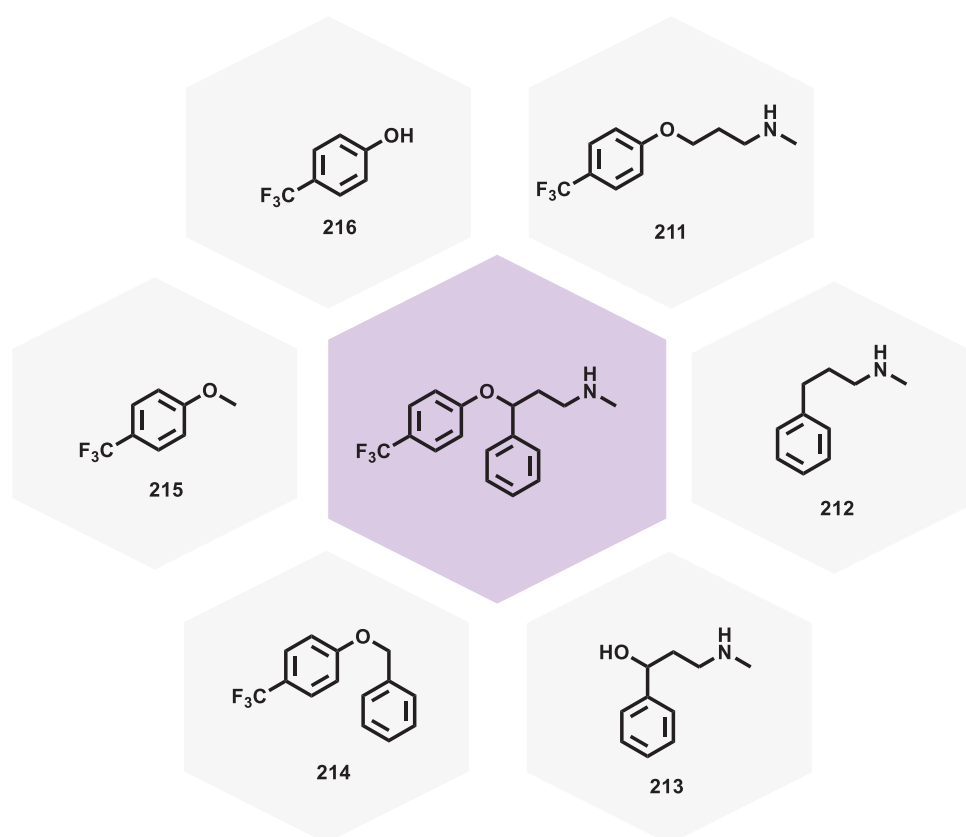
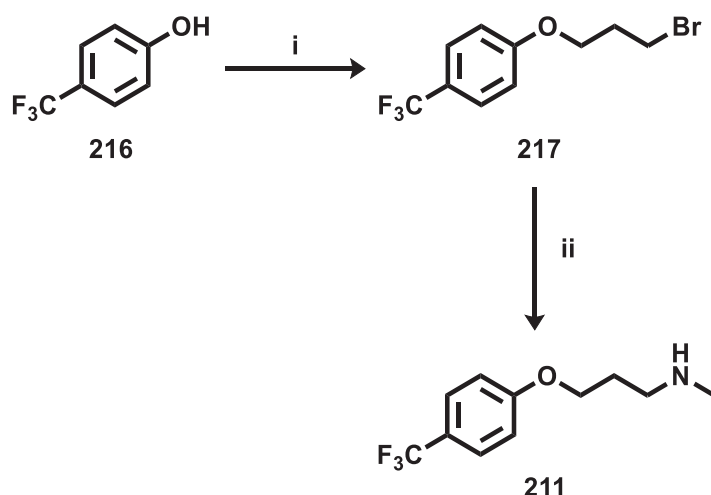


Figure 22. Chemical structure of the fluoxetine selected fragments.

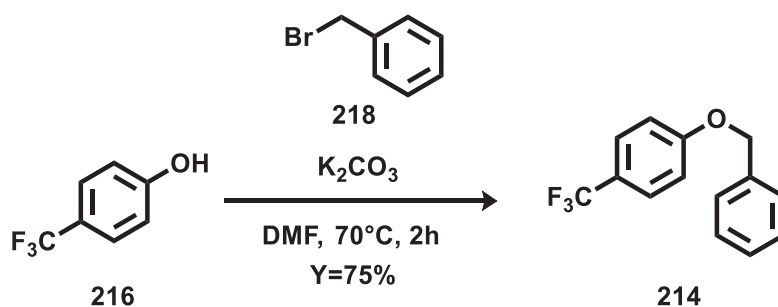
8.4.1 Synthesis of fluoxetine fragments

The *N*-methyl-3-(4-(trifluoromethyl)phenoxy)propan-1-amine **211** fragment was synthesised through a two-step synthesis (see scheme 1). In the first step, an S_N2 between the 4-(trifluoromethyl)phenol **216** and the dibromopropane occur in basic conditions. Successively, the methylamine is used to replace the bromo on the 1-(3-bromopropoxy)-4-(trifluoromethyl)benzene **217** intermediate.



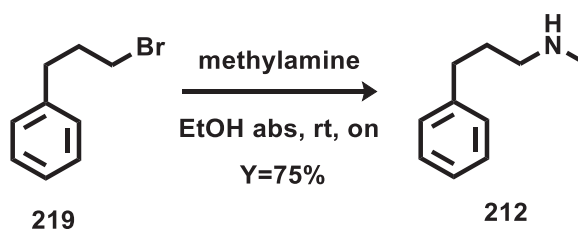
Scheme 1 Synthesis *N*-methyl-3-(4-(trifluoromethyl)phenoxy)propan-1-amine **211**. i) 1,3-dibromopropane, K₂CO₃, DMF, 70°C, 2h, Y=47%; ii) Methylamine, EtOH absolute, rt, on, Y=69%.

The 1-(benzyloxy)-4-(trifluoromethyl) benzene **214** fragment was obtained by S_N1, between the 4-(trifluoromethyl)phenol **216** and the benzylbromide **218** (scheme 2).



Scheme 2 Synthesis of 1-(benzyloxy)-4-(trifluoromethyl) benzene **214** fragment

The last not commercially available fragment synthesised was *N*-methyl-3-phenylpropan-1-amine **212**. This compound was obtained by the methylamine nucleophilic replacement on the bromo atom of the (3-bromopropyl)benzene **219**.



Scheme 3 Synthesis of *N*-methyl-3-phenylpropan-1-amine **212**

8.4.2 Antiviral evaluation of Fluoxetine fragments and 2C binding assay

The six fragments were evaluated in a multicycle-CPE antiviral assay against CVB3. As can be seen in figure 23, the only fragment that showed a weak antiviral activity against CVB3 was the *N*-methyl-3-(4-(trifluoromethyl)phenoxy)propan-1-amine **211**. However, the antiviral activity was close to the cytotoxicity effect on the HeLa R19.

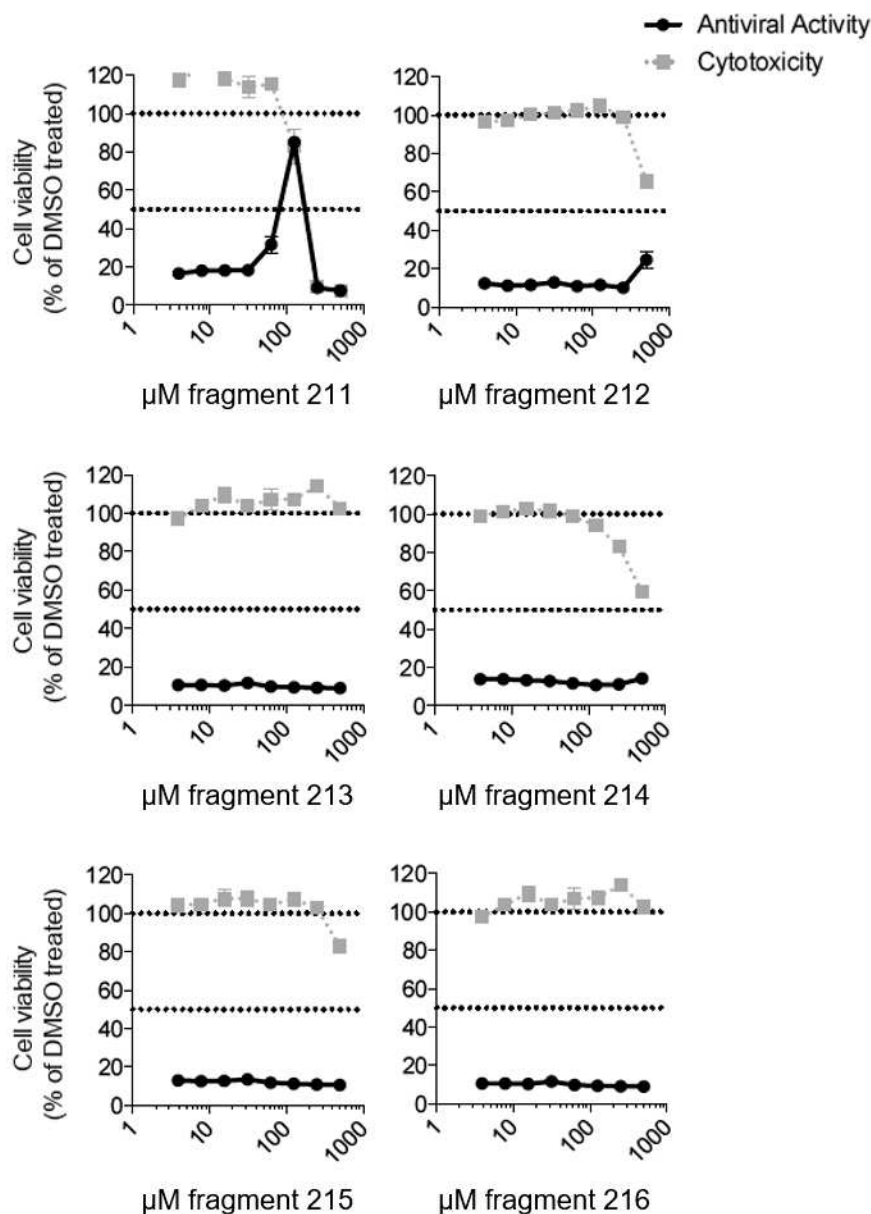


Figure 23. Antiviral evaluation of 211-216 fragments. The fluoxetine fragments were tested in a multicycle CPE-reduction assay in HeLa R19 cell-line infected with CVB-3 virus. MTS assay was used for the readout. The same assay was used on un-infected cells for the evaluation of CC_{50} . Data are shown from one experiment representative of at least two independent experiments.

Fragment **211**, which contain the methylamino group and the trifluoromethylphenyl moiety, was tested against EV-A71 to evaluate if the antiviral effect is related to the cytotoxicity and no-antiviral effect was observed. Like for fluoxetine, fragment **211** did not show antiviral effect against the EV-A71. The obtained result suggested that the antiviral activity is specific to the fragment.

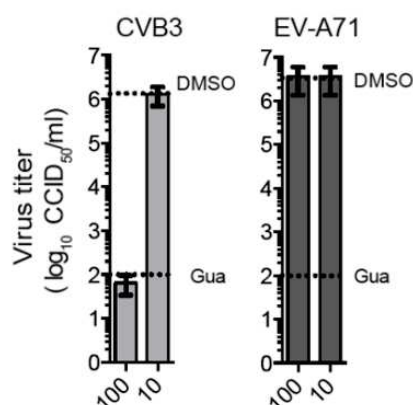


Figure 24. Antiviral evaluation of 211 fragment against CVB3 and EV-A71. Data are shown from one experiment representative of at least two independent experiments

Finally, the fragments were evaluated in TSA assay. No fragment showed a binding to the 2C protein. Also, the **211** did not stabilise the protein at 400 μ M. This result could be implicated to the eventuality that the concentration of the fragment might be much higher than the concentration in which the fragment was tested.

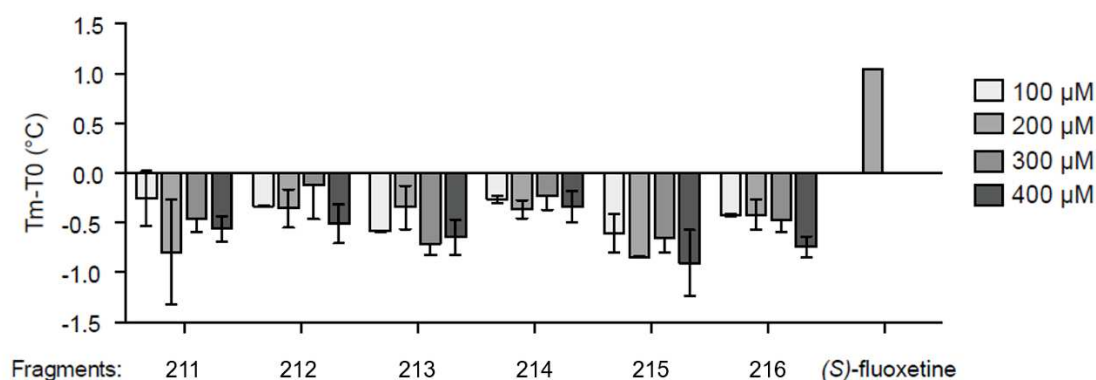


Figure 25. TSA evaluation of the fragments. Data are shown from one experiment representative of at least three independent experiments

8.5 Fluoxetine analogues

As was said before, the repurpose of fluoxetine could be difficult due neurological effect as SSRI and to the pharmacological off-target effect. Four analogues of fluoxetine were synthesised to try to separate the neurological effect from the antiviral activity. The analogues have different substituted phenyl moieties. Those compounds were chosen because they had reduced SSRI activity.⁴⁹

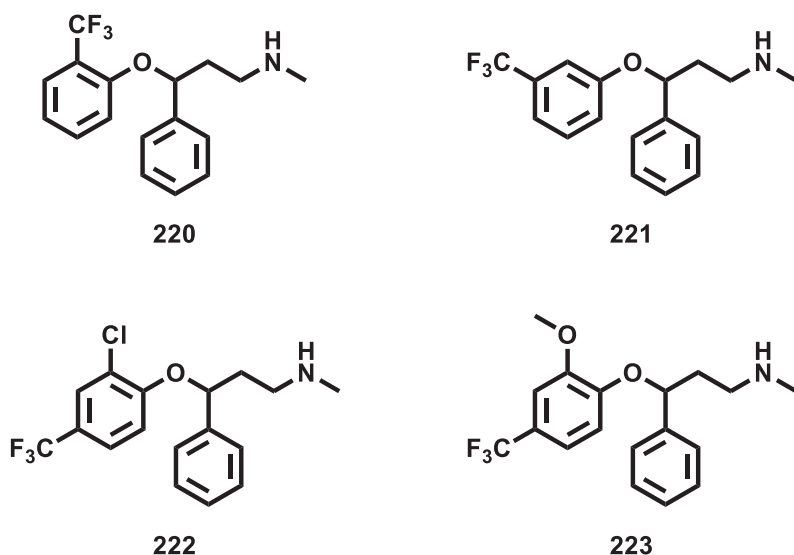
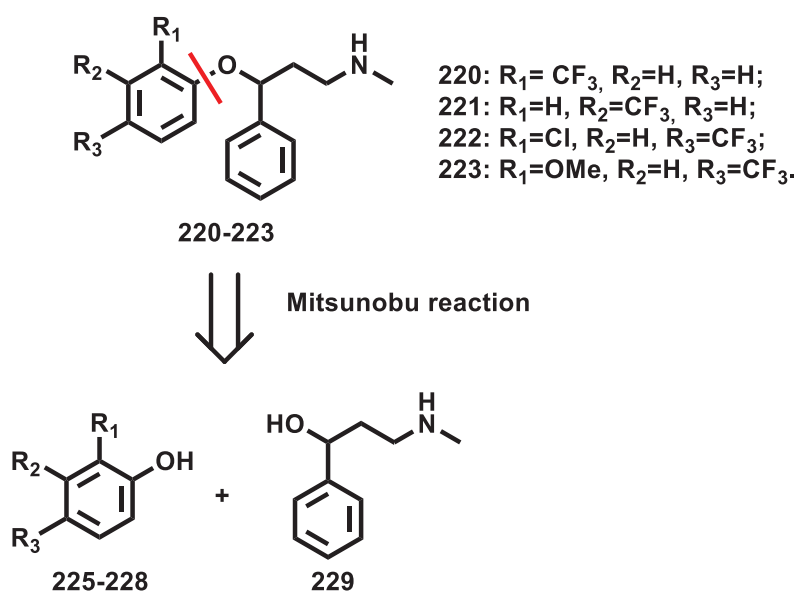


Figure 26. Selected fluoxetine analogues.

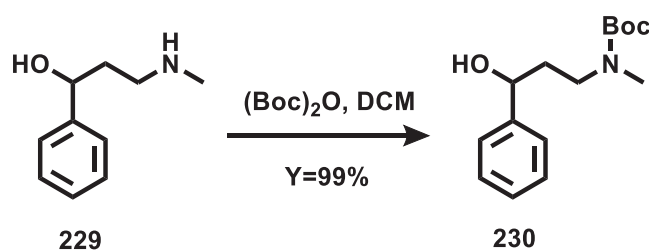
8.5.1 Synthesis of fluoxetine analogues

From the retrosynthetic analysis (scheme 4), the compounds **220-223** could be synthesised through the Mitsunobu reaction between the phenols **225-228** and the 3-(methylamino)-1-



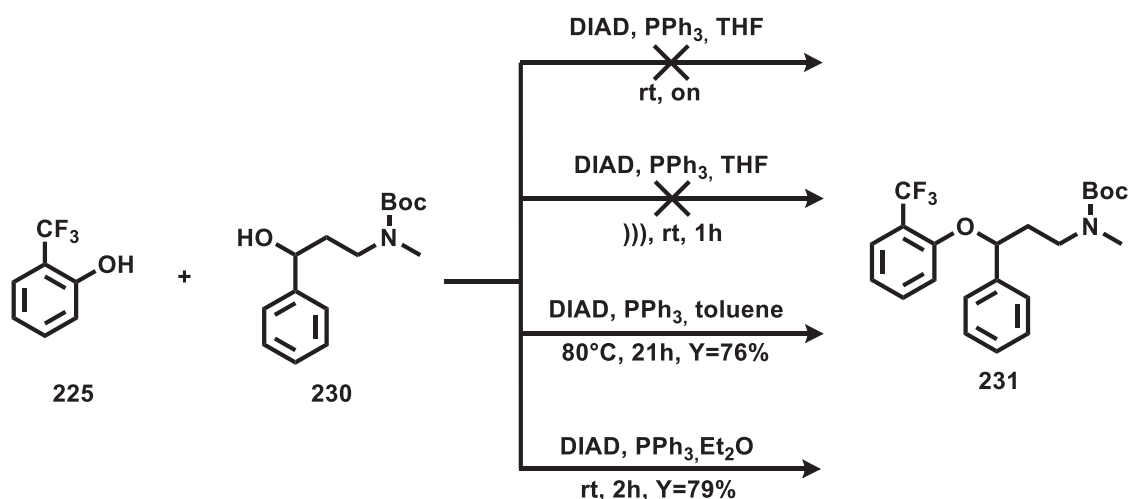
Scheme 4. Retrosynthetic analysis for compounds 220-223.

phenylpropan-1-ol **229**. To avoid any potential by-products during the Mitsunobu reaction, the amino group of the starting material **229**, was protected with a tert-butoxycarbonyl protecting group (scheme 5).



Scheme 5. Synthesis of the amino protected compound 230.

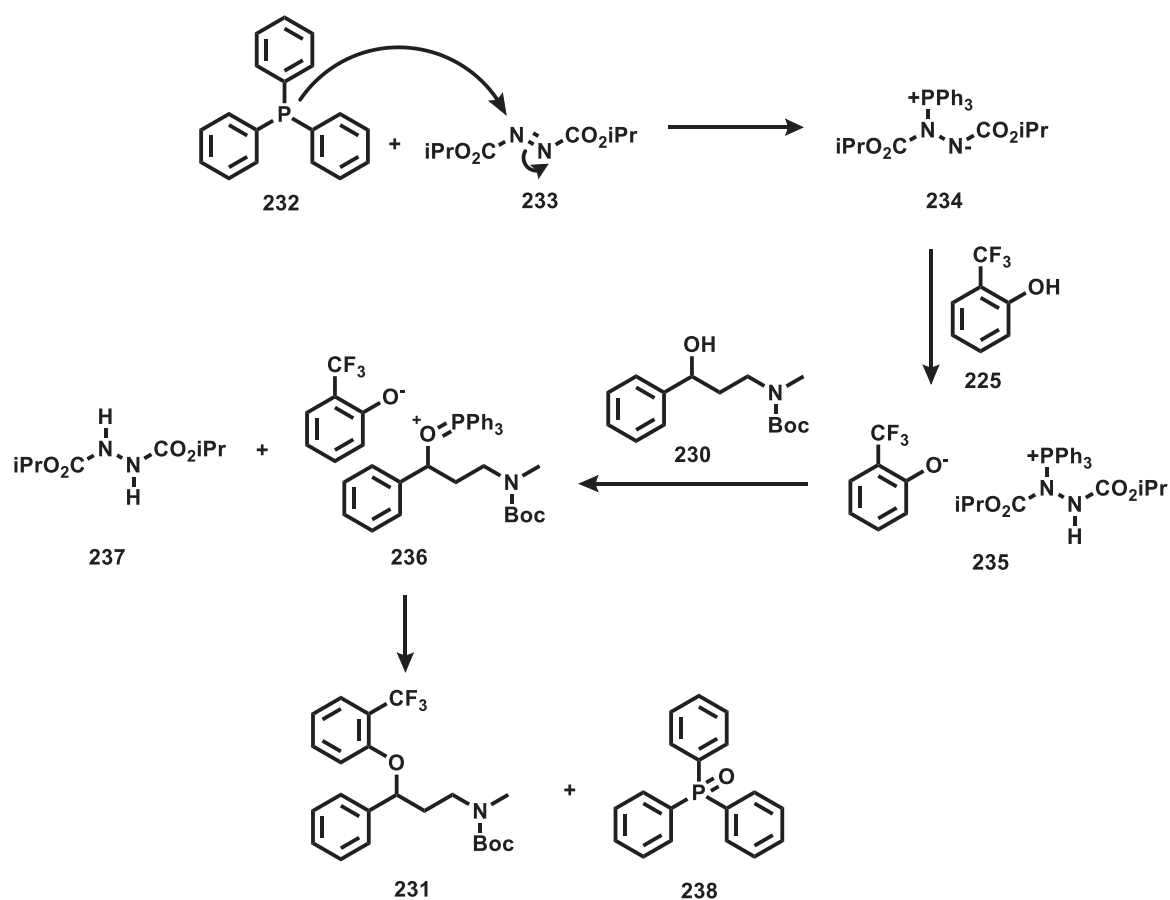
The obtained compound **230**, was then used for the Mitsunobu reaction for the formation of the alkyl-aryl ether bond formation with the 2-(trifluoromethyl)phenol **225**. The Mitsunobu reaction is a very versatile tool used for dehydration coupling between an alcohol and an acid or a nucleophile. In a first attempt, the reaction was performed, dissolving the phenol **225**, the alcohol **230** and triphenylphosphine (PPh_3) in tetrahydrofuran (THF) (scheme 6). The resulting solution was then cooled to 0°C before the addition of Diisopropyl azodicarboxylate (DIAD), and the reaction left at room temperature. Formation of the desired product was not observed, and different attempts changing the molar ratio of the starting materials were unsuccessful.



Scheme 6. Synthesis of the Boc protected N-methyl-3-phenyl-3-(2-(trifluoromethyl)phenoxy)propan-1-amine 231.

Another attempt was performed using a Mitsunobu reaction under sonication condition and increasing the concentration to 3.0 M. This procedure is reported to be extremely useful for sterically hindered substrates and the sonication process together with the increased concentration is reported to speed-up the Mitsunobu reaction. Unfortunately, also this

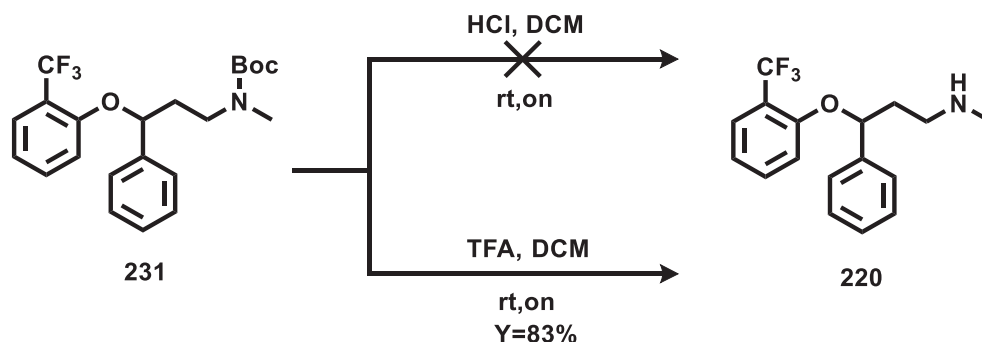
attempt was unsuccessful.⁶⁰ Finally, the formation of the desired procedure was obtained with a good yield using a reported procedure for similar compounds. In this procedure, toluene was used instead of THF and the reaction was heated at 80°C for 21h. After a work-up and then purification of the crude, the desired product was obtained with a yield of 79%.⁶¹ To try to avoid the use of toluene and the needs to do a work-up before the purification, another procedure reported in the literature for a similar compound was used. In this procedure, diethyl ether was used as a solvent, and the DIAD and the PPh₃ were added before the phenol and the alcohol, aiming to perform the betaine salt. This reaction occurs at room temperature and with a comparable yield after purification.⁶²



Scheme 7. Reaction Mechanism of the Mitsunobu reaction.

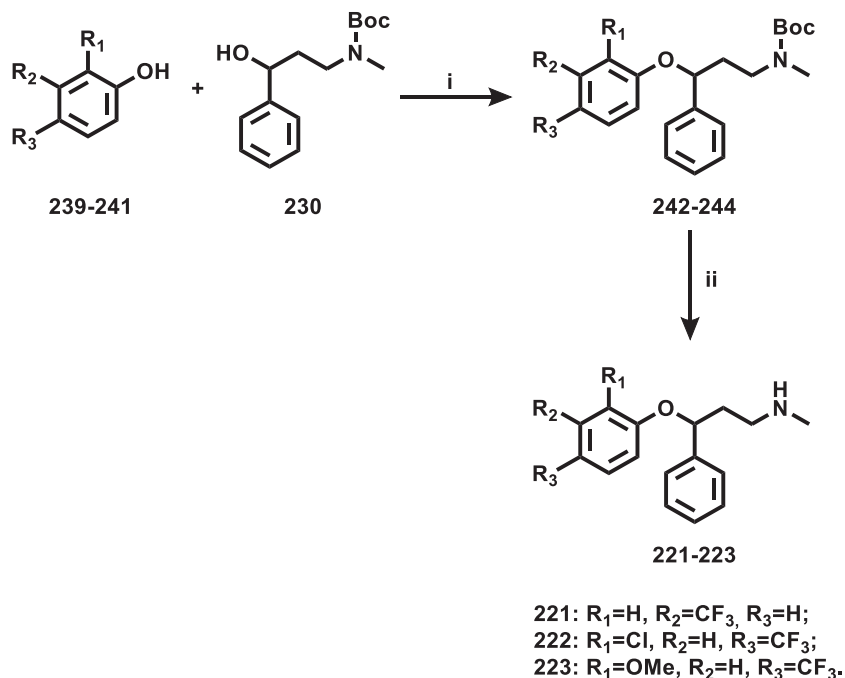
As shown in scheme 7, the reaction starts with the formation of quaternary phosphonium salt, also called betaine salt **234**. The betaine salt is then deprotonating the 2-(trifluoromethyl)phenol, to give the ion pair **235**, which subsequently react with the alcohol **230** to form the alkoxy-phosphonium salt **236** and the diethyl hydrazinedicarboxylate **237**. In the end, the Boc protected *N*-methyl-3-phenyl-3-(2-(trifluoromethyl)phenoxy)propan-1-amine **231** and triphenylphosphine oxide **238** are obtained by an S_N2 displacement of the alkoxy-phosphonium salt **236** by the phenol **225**.⁶³

The last step of synthesis was the deprotection of the Boc protected compound **231** (see scheme 8). This step was carried out using a 2 M solution of hydrogen chloride (HCl) in diethyl ether. By monitoring the reaction by TLC, a partial conversion of the starting material **231** was observed. Increasing the equivalents of the HCl did not provide further conversion of the **231**. Indeed, the formation of the phenol **225** was observed which could be due to the ether cleavage side-reaction when using strong acid like HCl.



Scheme 8. Deprotection of Boc protected *N*-methyl-3-phenyl-3-(2-(trifluoromethyl)phenoxy)propan-1-amine **231**.

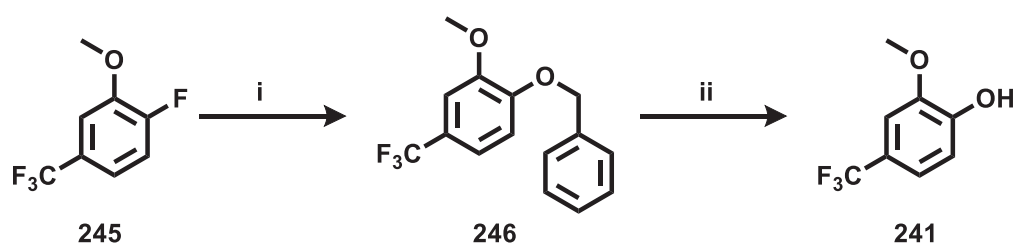
This side-reaction was avoided, changing the HCl to the trifluoroacetic acid, a weaker acid, which gives after liquid chromatography the desired *N*-methyl-3-phenyl-3-(2-(trifluoromethyl)phenoxy)propan-1-amine **220** with a yield of 83%.



Scheme 9. Synthesis of the fluoxetine analogues **221-223**. i) DIAD, PPh₃, Et₂O, rt, 2h, Y=60-64%; ii) TFA, DCM, rt, on, Y=62-64%.

The other fluoxetine analogues **221-223** were synthesised using the two steps synthetic route optimised for the synthesis of compound **220**, as showed in scheme 9. The desired analogues were obtained with comparable yields to the ones obtained during the synthesis of **220**.

For the synthesis of the 3-(2-methoxy-4-(trifluoromethyl)phenoxy)-*N*-methyl-3-phenylpropan-1-amine **223**, the synthesis of the 2-methoxy-4-(trifluoromethyl)phenol was required. Starting from 1-fluoro-2-methoxy-4-(trifluoromethyl)benzene **245**, 1-(benzyloxy)-2-methoxy-4-(trifluoromethyl)benzene **246** was obtained through a S_NAr . The benzyloxy benzene **246** was then subjected to a hydrogenation step which leads to the desired product **241** (scheme 10).



Scheme 10. Synthesis of 2-methoxy-4-(trifluoromethyl)phenol **241**. i) benzyl alcohol, NaH, DMF, rt, 5h, Y=85%; ii) H₂, Pd/C, EtOAc, rt, on, Y=99%

To investigate the importance and the contribution of the fluoxetine amino group for the antiviral activity was decided to synthesise the acetyl analogue **247**.

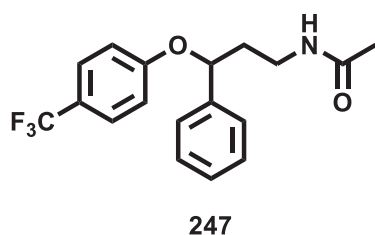
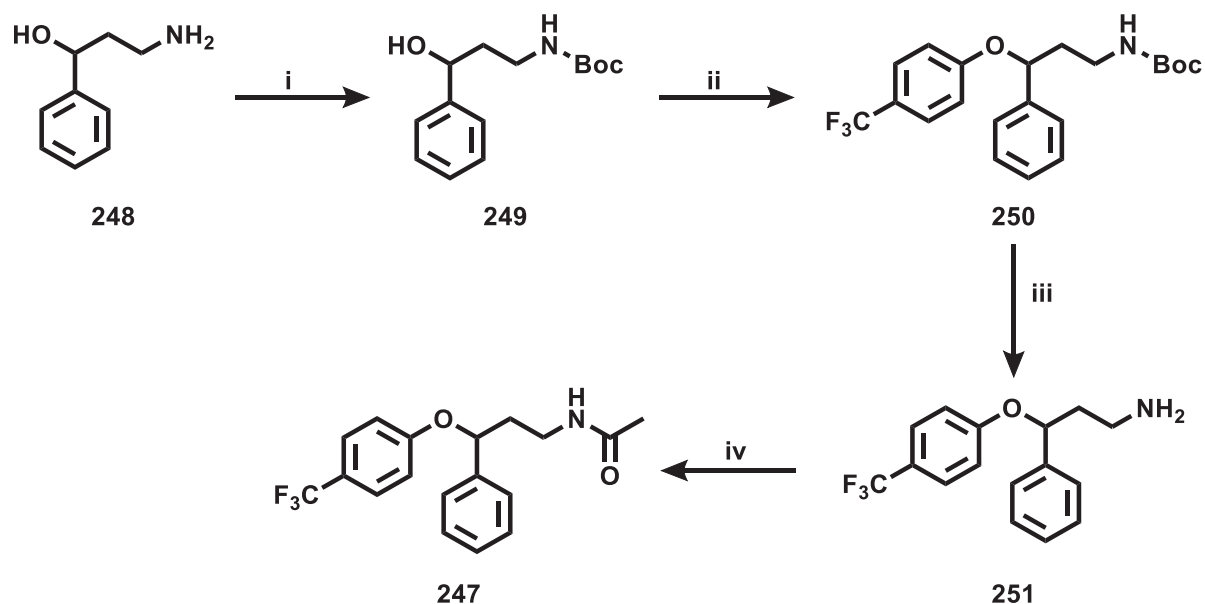


Figure 27. Acetyl analogue of fluoxetine.

This compound was synthesised by acetylation of the 3-phenyl-3-(4-(trifluoromethyl)phenoxy)propan-1-amine **251** (norfluoxetine), as showed in scheme 11. The norfluoxetine was synthesised using the three-steps synthetic route used to synthesise the other fluoxetine analogues, starting from the 3-amino-1-phenylpropan-1-ol **248** instead of 3-(methylamino)-1-phenylpropan-1-ol **229**.



Scheme 11. Synthesis of the fluoxetine analogues 247. i) (Boc)₂O, DCM, rt, on, Y=87%; ii) 4-(trifluoromethyl)phenol, DIAD, PPh₃, Et₂O, rt, on, Y=69%; iii) TFA, DCM, rt, on, Y=64% iv) acetyl chloride, TEA, DCM, rt, 2h, Y=98%.

8.5.2 Antiviral evaluation of fluoxetine analogues

The fluoxetine analogues **220-223** and **247**, were evaluated in the multicycle-CPE antiviral assay on CBV3. As reported in table 4, no antiviral activity was observed for any of the fluoxetine analogues synthesised. Except for compounds **220** and **247**, the other compounds were associated with increased cytotoxic effect.

Table 4. Antiviral evaluation and cytotoxicity assessment of compounds 220-223 and 247. Compounds were tested in a multicycle CPE-reduction assay in Hela R19 cell-line infected with CVB-3 virus. MTS assay was used for the readout. The same assay was used on un-infected cells for the evaluation of CC₅₀. Data represents mean values ± SD calculated from three independent experiments and are expressed in μM. NA = not active.

Compound	EC ₅₀ CVB3 (μM)	CC ₅₀ (μM)
220	NA	28.86±0.86
221	NA	12.56±1.79
222	NA	3.14±0.07
223	NA	18.41±1.26
247	NA	>30

8.6 Conclusions

In the first part of this project, the antiviral activity of the fluoxetine enantiomers was explored. The (*S*)-fluoxetine is the only enantiomer of the racemic mixture which is associated with an antiviral activity, which was found to be ~5-fold higher (EC_{50} 0.42 ± 0.17 μ M) than the racemic mixture (EC_{50} 2.02 ± 0.94 μ M). This result was also confirmed by TSA and ITC assays, confirming that only the (*S*)-fluoxetine is binding to the 2C non-structural protein of CVB3. Also, the *S*-enantiomer showed enhanced antiviral activity, compared to the racemate, against CVB3 and EV-D68. Interestingly, the active enantiomer was found to be also active against the rhinoviruses HRV-A2 and HRV-B14.

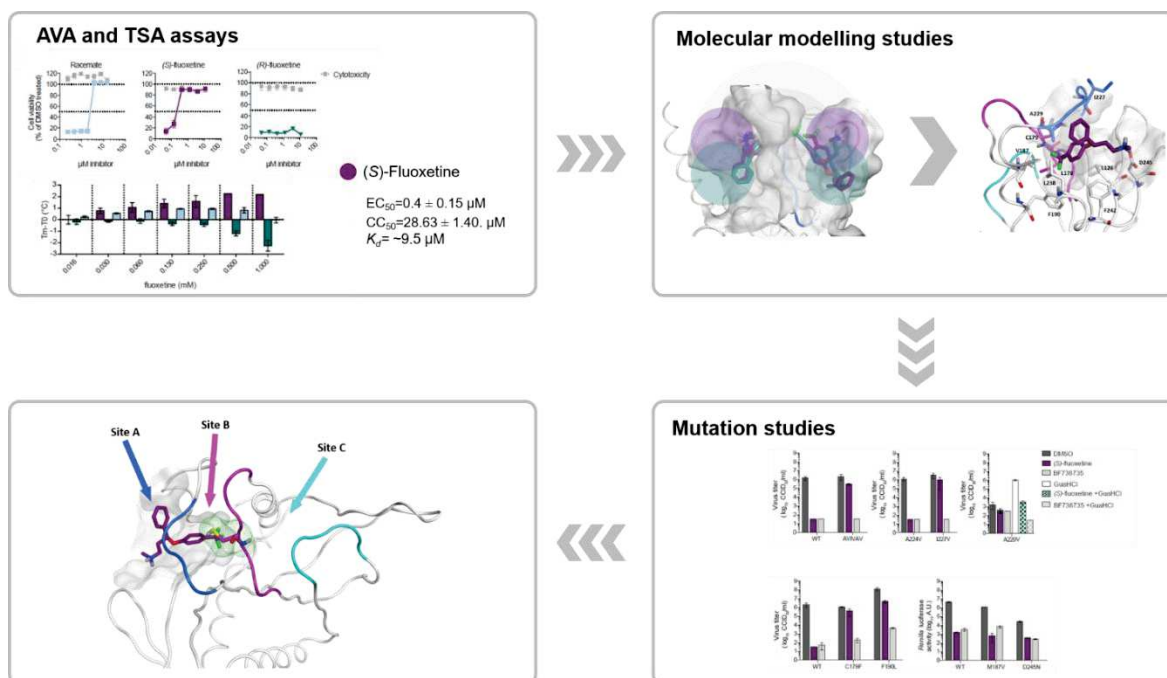


Figure 28. Summary picture.

Starting from the information acquired on the stereospecificity of the antiviral activity of the (*S*)-fluoxetine, molecular modelling techniques were used to elucidate the binding mode of fluoxetine on 2C protein. The available crystal structure of the EV-A71 2C protein was used to generate the homology model of the CVB3 2C protein. The model was then used to dock the enantiomers in two potential pockets, presents near the ²²⁴AGSINA²²⁹ loop, and was then subjected to MD simulations. The results obtained suggested that only site A was able to make a stable interaction with the *S*-enantiomer. Stable interactions were observed for the *R*-enantiomer neither in site A nor in site B.

Unfortunately, from mutations study, the interaction of the (*S*)-fluoxetine with the hydrophobic pocket identified are crucial for the antiviral activity. On the other hand, the mutant viruses having the D245N and A229V mutation, respectively, did not show any

resistance. These results are suggesting that the binding of the (S)-fluoxetine could be slightly different or another entrance for the hydrophobic, potentially between 175-183 and 158-164 loops cannot be excluded.

Further, to investigate the importance of the different chemical features of the fluoxetine, six fragments were tested in the antiviral assay. The results suggested that the 4-trifluoromethylphenyl moiety and the amino group are crucial for the antiviral activity.

Trying to separate the SSRI from the antiviral activity, four analogues of fluoxetine were synthesised and biologically evaluated against the CVB3. No antiviral activity was detected, suggesting that the hydrophobic pocket in which part of the (S)-fluoxetine is interacting, is selective for the 4-trifluoromethylphenyl moiety, and any of the changes introduced with the analogues are allowed.

Finally, the importance of the amino group for the interaction with the protein 2C was confirmed by the acetylated analogues of fluoxetine. The compound in which, the amino group of fluoxetine is involved in an amide bond, is losing the antiviral activity.

Instead, the resolution of a crystal structure, co-crystallised with the (S)-fluoxetine could confirm or not the binding mode and the predicted pocket. The crystal structure could help to understand the difference among the 2C protein of the different enterovirus species, and the explanation of the selectivity of fluoxetine for HEV-B and HEV-D species will help for the rational design of new enterovirus pan-inhibitors.

Chapter 9: Guanidine derivatives of fluoxetine

9.1 Design and synthesis of Guanidine derivatives of fluoxetine

The guanidine hydrochloride is one of the most studied inhibitor of enterovirus replication since 1964. The compound is interfering with an early stage of the viral replication. The mechanism of action proposed is that the guanidine is inhibiting only the initiation of the negative-strand RNA synthesis which prevents the positive-strand RNA synthesis.³⁶

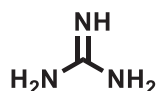


Figure 29. Guanidine.

Mutations on the 2C protein characterise nearly all Guanidine-raised mutants.^{30,33} Those mutations are providing resistance also on HBB, MRL-1237 and TBZE-029, known to be also 2C inhibitors. Some of these mutations are making the mutants dependent on the presence of guanidine for viral replication.^{15,32,43} The A224V-I227V-A229V mutations on the ²²⁴AGSINA²²⁹ loop, the same which makes the CVB3 mutant resistant to fluoxetine were also giving resistance to guanidine, and in particular the A229V mutation alone or in combination with A224V, were making the mutant dependent on guanidine for the viral replication.¹⁸

The previously gained information about the important of the amino group of the fluoxetine analogue for the antiviral activity, together with knowledge that the common mutations on the ²²⁴AGSINA²²⁹ loop are making CVB3 mutants to guanidine, led to the design of the 1-(3-phenyl-3-(4-(trifluoromethyl)phenoxy)propyl)guanidine **252** compound.

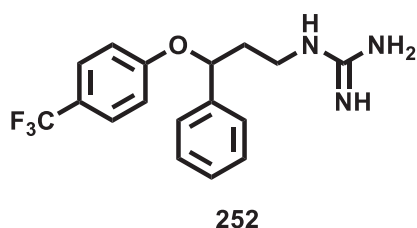
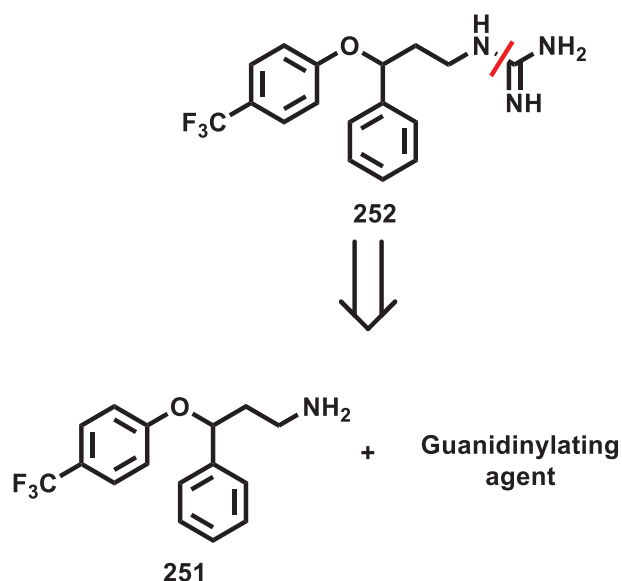


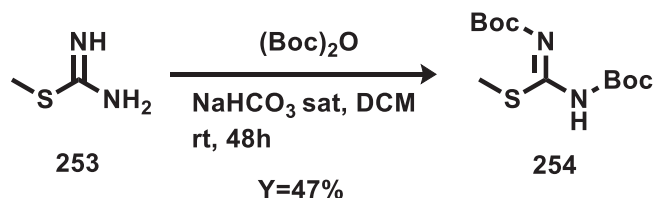
Figure 30. Guanidine analogue of fluoxetine.

From a retrosynthetic analysis, compound **252** could be synthesised starting from the previously synthesised norfluoxetine **251** and a guanidinylation agent (schem12). The *N,N'*-bis(Boc)-*S*-methylisothiourea guanidinylation agent **254** was decided to be used. It was not commercially available, and its synthesis was required.



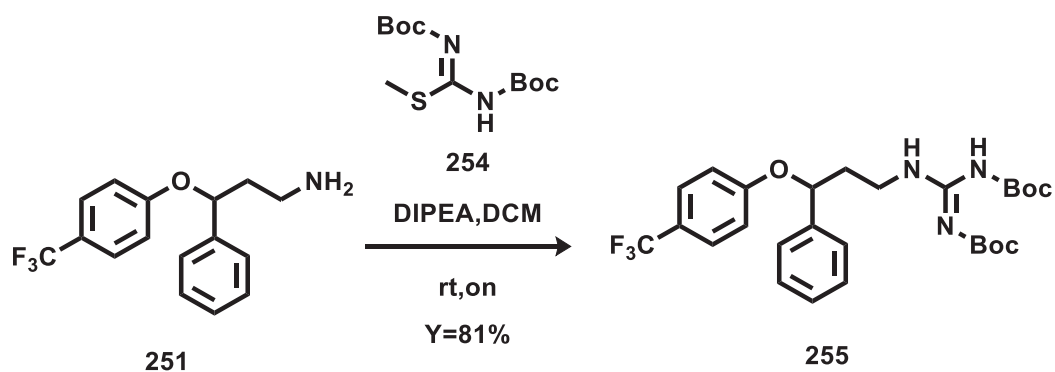
Scheme 12. Retrosynthetic analysis for 1-(3-phenyl-3-(4-(trifluoromethyl)phenoxy)propyl)guanidine **252**.

The *N,N'*-bis(Boc)-*S*-methylisothiurea **254** was synthesised by Boc protecting the *S*-methyl isothiurea (**253**), following the reported procedure, shown in scheme 13.⁶⁴



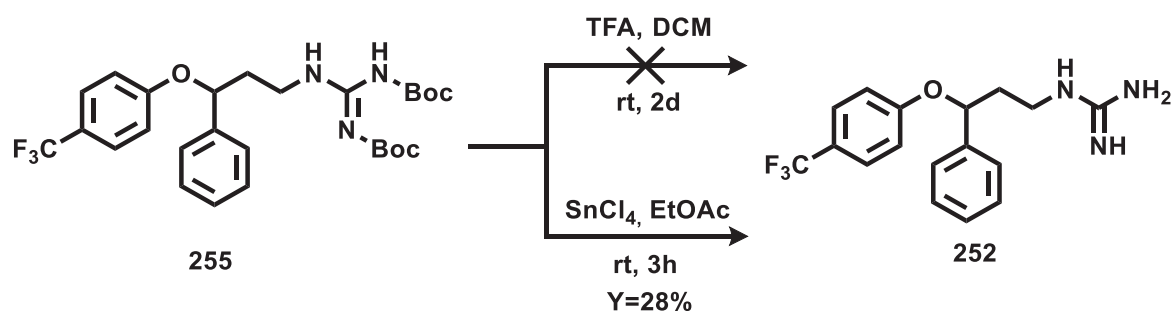
Scheme 13. Synthesis of *N,N'*-bis(Boc)-*S*-methylisothiurea **254**.

The guanidinyllating agent **254** was then used for the synthesis *N,N'*-Bis(tert-butoxycarbonyl)-(1-(3-phenyl-3-(4-(trifluoromethyl)phenoxy)propyl))guanidine **255**. This step of synthesis consisted of a nucleophilic displacement of the thiomethyl group of **254** by the amino group of the norfluoxetine (**251**) in the presence of DIPEA as a base (scheme 14).⁶⁵ The last step of the synthesis was the di-Boc deprotection of the intermediate **255**. A first attempt was performed using the same procedure used for the Boc deprotection of norfluoxetine, increasing the amount of acid (scheme 15). Even by increasing the amount of TFA, only the mono Boc deprotected **255** was recovered.



Scheme 14. Synthesis of *N',N''*-Bis(tert-butoxycarbonyl)-(1-(3-phenyl-3-(4-(trifluoromethyl)phenoxy)propyl))guanidine **255**.

As an alternative to TFA, stannic chloride (SnCl_4) was used, a reagent which was reported to remove both the Boc protecting groups under mild and neutral conditions.⁶⁶⁻⁶⁷ The reaction was performed in ethyl acetate (EtOAc) at room temperature, and the desired 1-(3-phenyl-3-(4-(trifluoromethyl)phenoxy)propyl)guanidine (**252**) was obtained as a free-base (scheme 15).



Scheme 15. Boc deprotection of intermediate **255**.

To modify the distance between the guanidinium group added to the fluoxetine was decided to synthesise the 1-(2-phenyl-2-(4-(trifluoromethyl)phenoxy)ethyl)guanidine **256**, with a methylene bridge less than the previously synthesised guanidine compound **252**.

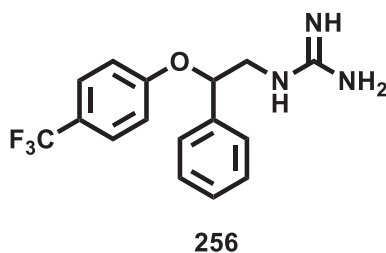
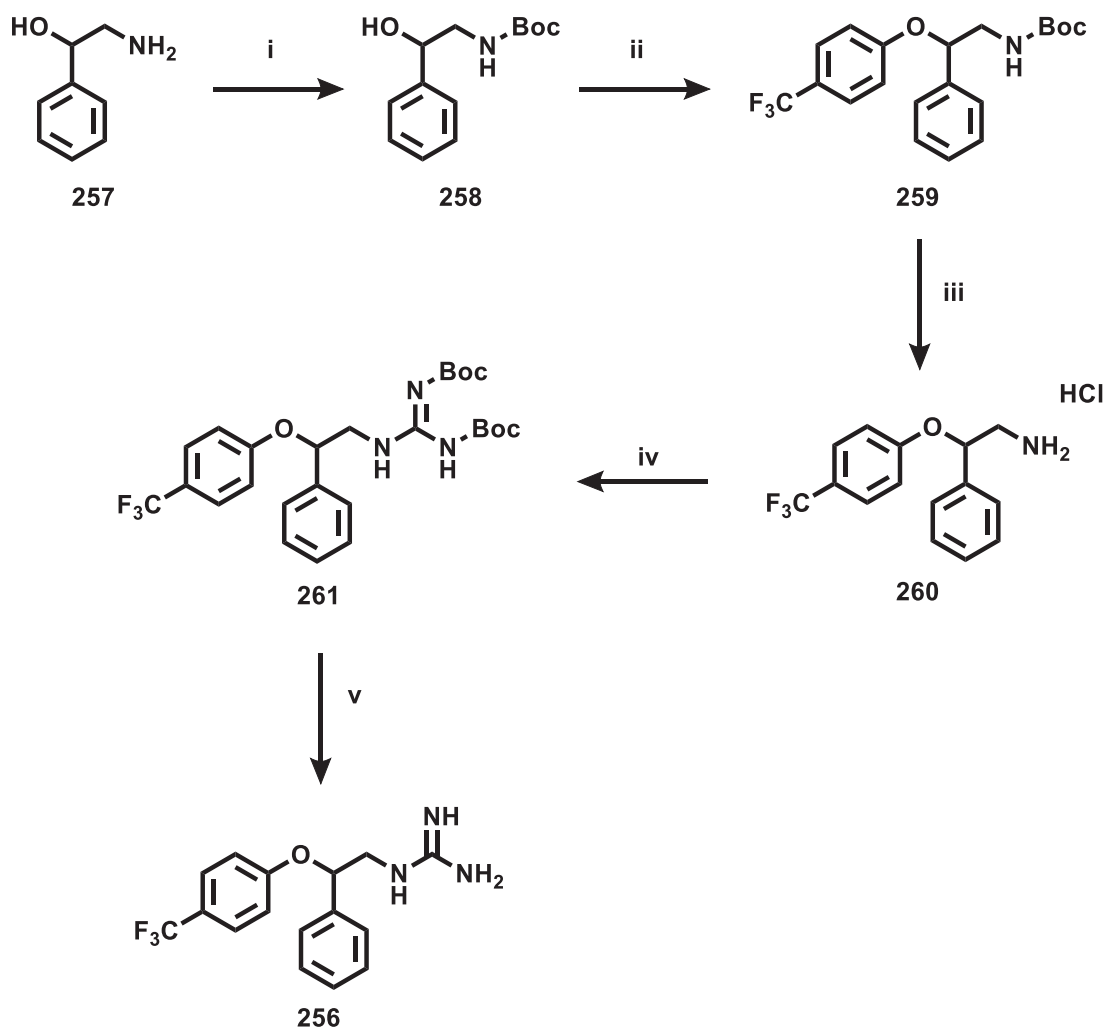


Figure 31. Guanidine analogue of fluoxetine with a methylene bridge less.

The planned compound was synthesised through the five-step synthetic route, showed in scheme 16, starting from the 2-amino-1-phenylethan-1-ol **257**. The synthetic intermediate **260** was obtained by Boc deprotection of the precursor **259**, using the HCl. In this case, the desired product precipitate during the reaction, giving the desired product after filtration.



Scheme 16. Synthesis of 1-(2-phenyl-2-(4-(trifluoromethyl)phenoxy)ethyl)guanidine **256.** i) (Boc)₂O, TEA, DCM, rt, on, Y=93%; ii) 4-(trifluoromethyl)phenol, DIAD, PPh₃, Et₂O, rt, on; iii) HCl, DCM, rt, 24h, Y=37%(over 2 steps); iv) 205, DIPEA,DCM, rt, on, Y=46%; v) SnCl₄, EtOAc, rt, on, Y=95%

9.2 Antiviral evaluation of Guanidine derivatives of fluoxetine

As can be deduced from table 5, the introduction from the guanidine group is related to a ~7-fold increase of the antiviral activity in comparison with the fluoxetine racemic mixture. The compound **252** is equipotent to the (S)-fluoxetine, with an EC₅₀ of 0.41 ± 0.27 μM and a CC₅₀ >50 μM. The analogue of norfluoxetine with a methylene bridge less **260** showed to be active in the same range of activity of the fluoxetine, and the introduction of the guanidine group **256** is related with a ~3-fold increased antiviral activity.

Table 5. Antiviral evaluation of the guanidine analogues. Compounds were tested in a multicycle CPE-reduction assay in Hela R19 cell-line infected with CVB-3 virus. MTS assay was used for the readout. The same assay was used on un-infected cells for the evaluation of CC₅₀. Data represents mean values ± SD calculated from three independent experiments and are expressed in μM. SI = selectivity index (CC₅₀/EC₅₀).

Compound	EC ₅₀ CVB3 (μM)	CC ₅₀ (μM)	SI
(RS)-fluoxetine	3.2 ± 0.95	29.32±0.35	9.1
(S)-fluoxetine	0.4 ± 0.15	28.63±1.40	71.6
252	0.41±0.27	>50	>121.9
260	4.20±0.93	32.26±0.10	7.7
256	1.22±0.15	>50	>40.9

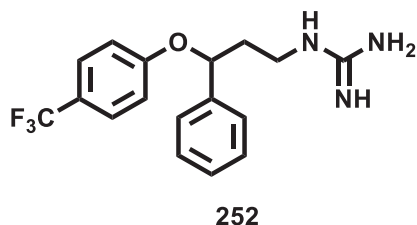
The guanidine analogues of fluoxetine **252** and **256** were then tested on a panel of enterovirus. The tested compounds had the same trend the other enterovirus as (S)-fluoxetine. The compound was equipotent to (S)-fluoxetine against EV-D68, and they did not show any antiviral activity against the tested HEV-A and HEV-C species. Finally, compound **252** was the only one showing antiviral activity against HRV-B14.

Table 6. Antiviral evaluation of the guanidine analogues on other enteroviruses. Compounds were tested in a multicycle CPE-reduction assay in Hela R19 cell-line infected with different enteroviruses. MTS assay was used for the readout. The same assay was used on un-infected cells for the evaluation of CC_{50} . Data represents mean values \pm SD calculated from three independent experiments and are expressed in μ M. NA = not active.

Compound	EC ₅₀ EV-D68 (Fermon)	EC ₅₀ EV-A71 (BrCr)	EC ₅₀ PV-1 (Sabin1)	EC ₅₀ HRV-A2	EC ₅₀ HRV-B14
(S)-fluoxetine	0.67 \pm 0.22	NA	NA	7.95 \pm 0.39	6.34 \pm 1.02
252	0.34 \pm 0.05	NA	NA	NA	2.11 \pm 0.34
256	0.55 \pm 0.04	NA	NA	NA	NA

9.3 Conclusions

Starting from the information available on the antiviral activity of guanidine, and the stereospecific antiviral activity of the (*S*)-fluoxetine, a novel guanidine analogue of fluoxetine was designed.



The designed compound **252**, synthesised as a racemic mixture, was found to be equipotent against CVB3 in comparison with the (*S*)-fluoxetine and the same trend of antiviral activity against other enteroviruses.

The influence of the methylene bridge length between the amino group and the rest of the chemical feature of norfluoxetine and the guanidine analogue **252** was also explored. The synthesised analogues of norfluoxetine and compound **252** were associated with retention of antiviral activity equipotent to the racemic mixture of fluoxetine. Also, in compound **256**, the guanidine analogue with a methylene bridge less than **252**, showed the same trend of fluoxetine against the other enterovirus species. The analogue **256** was active against EV-D68, but not against EV-A71 and PV-1.

Further, the biological evaluation will be performed to confirm that these analogues of fluoxetine are affected by the same mutations that are conferring to CVB3 resistance to fluoxetine.

Chapter 10: *N*-benzyl-*N*-phenylfuran-2-carboxamide

10.1 *N*-benzyl-*N*-phenylfuran-2-carboxamide

In 2016, Zuo et al. reported and patented several chemical structures, which were identified by a screening campaign on 85585 compounds, to be able to inhibit the viral replication of CVB3.^{48,68} Selection of resistant-CVB3 mutants, followed by sequence analysis, for those compounds, revealed that the 2C protein was containing most of the mutations observed, especially mutations at the 179 and 227 positions on the protein sequence. Furthermore, the C179F CVB3 mutant showed resistance against all the identified compounds, confirming the 2C protein as the target for the identified compounds. Interestingly the I227V and the C179F mutations were the same ones found to confer resistance to CB3 against the (*S*)-fluoxetine.

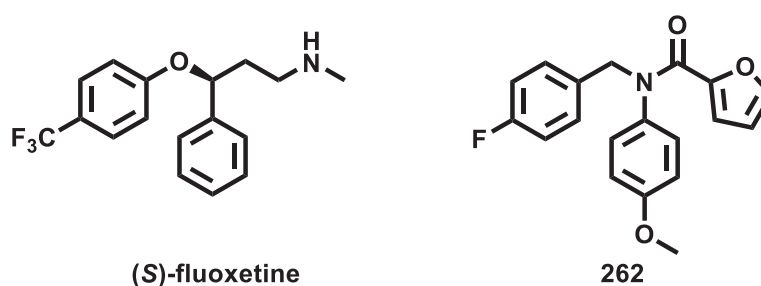


Figure 32. (*S*)-fluoxetine and compound 262.

One of the compounds, the *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)furan-2-carboxamide **262**, reported inhibiting a panel of HEV-B, but not HEV-A and HEV-C. The compound was reported to have an EC₅₀ of 2.9 μM and a CC₅₀ of 45.5 μM. This compound was raising mutations for which the (*S*)-fluoxetine was also sensible, suggesting that, probably, the 4-fluorobenzyl moiety might interact in the same pocket in which the 4-trifluoromethyl moiety of fluoxetine is interacting. Besides, compound **262** has a second aromatic ring and an H-bond acceptor heteroatom in the same orientation as (*S*)-fluoxetine.

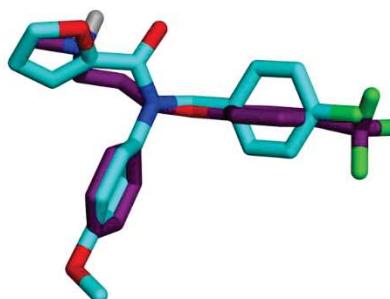


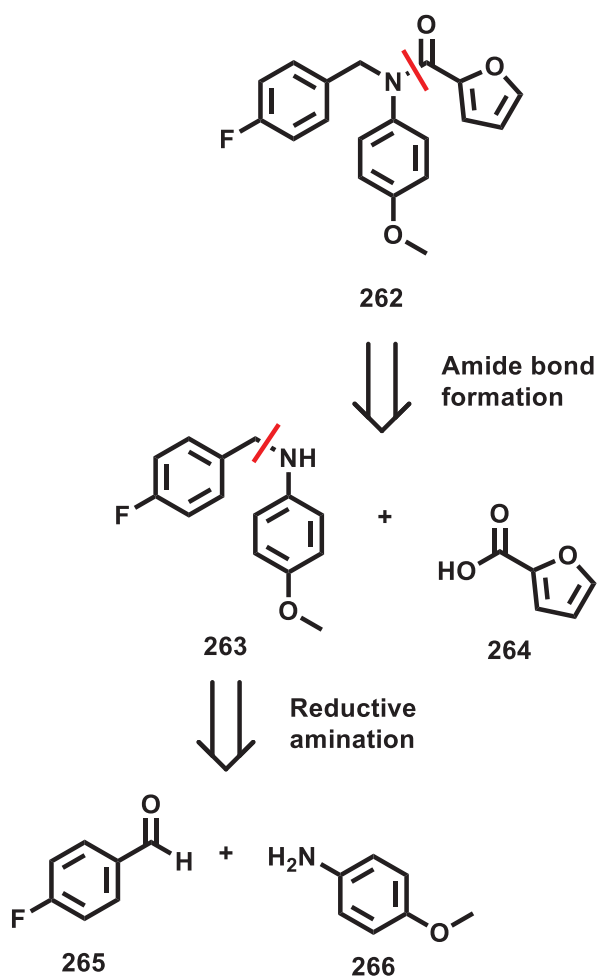
Figure 33. Superimposition of (*S*)-fluoxetine (purple) and compound 262 (cyan).

The similarity of the chemical features and their orientation between the (*S*)-fluoxetine and compound **262** (figure 32 and 33) and the lack of chiral center, made the *N*-(4-fluorobenzyl)-

N-(4-methoxyphenyl)furan-2-carboxamide **262** the perfect candidate for the further development of an enterovirus pan-inhibitors.

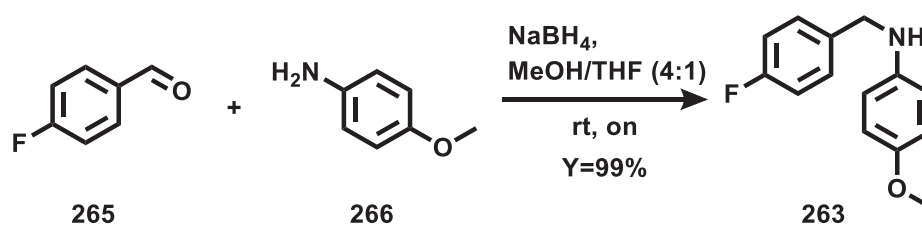
10.1.1 Synthesis of *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)furan-2-carboxamide

The synthesis of compound **262** was required because the compound is not commercially available. From the retrosynthetic analysis (scheme 17), compound 262 can be synthesised by a two steps synthesis, consisting from a reductive amination step between the 4-fluorobenzaldehyde **265** and the 4-methoxyaniline **266**, followed by a coupling reaction between amine, obtained by the reductive amination, and the furan-2-carboxylic acid **264**.



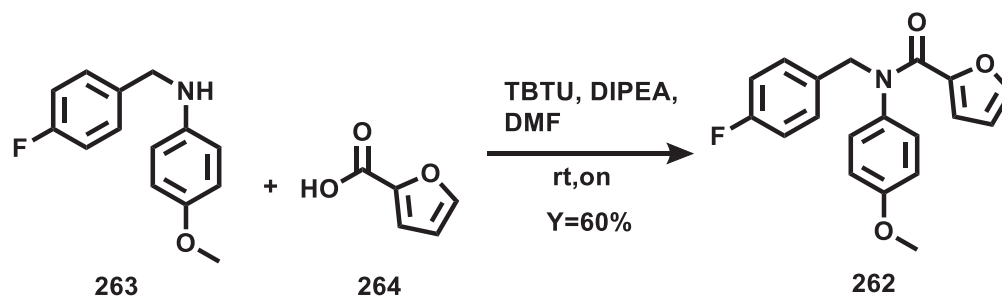
Scheme 17. Retrosynthetic analysis for compound 262.

The reductive amination for the synthesis of the *N*-(4-fluorobenzyl)-4-methoxyaniline **263**, was performed by solubilising the benzaldehyde **265** and the aniline **266** in a solution of methanol and tetrahydrofuran in a ratio of 4:1. The use of THF as a co-solvent was required to increase the solubility of the starting material, which is crucial for the formation of the imine intermediate. The imine formation was monitored by TLC, followed by the addition of sodium borohydride as reductive agent., to give the desired product with a yield of 99%.



Scheme 18. Synthesis of the N-(4-fluorobenzyl)-4-methoxyaniline 263.

Then, the *N*-(4-fluorobenzyl)-4-methoxyaniline **263** was used for the coupling reaction, with the furan-2-carboxylic acid **264**, using TBTU as coupling reagent. The reaction gives the desired *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)furan-2-carboxamide **262**, with a yield of 60%, as shown in scheme 19.



Scheme 19. Synthesis of the N-(4-fluorobenzyl)-N-(4-methoxyphenyl) furan-2-carboxamide 262.

10.1.2 Antiviral evaluation of *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)furan-2-carboxamide

The compound was evaluated in the multicycle-CPE antiviral assay, showing an antiviral activity on CVB3 comparable to the one reported in the paper.⁴⁸ The EC₅₀ of 1.12±0.28 μM found for compound **262**, is ~3-fold more potent compared to the fluoxetine and ~3-fold less potent than the (S)-fluoxetine and guanidine analogue **252** (table 7).

Table 7. Antiviral evaluation of compound 262. Compounds were tested in a multicycle CPE-reduction assay in Hela R19 cell-line infected with CVB-3 virus. MTS assay was used for the readout. The same assay was used on un-infected cells for the evaluation of CC₅₀. Data represents mean values ± SD calculated from three independent experiments and are expressed in μM. SI = selectivity index (CC₅₀/EC₅₀).

Compound	EC ₅₀ CVB3 (μM)	CC ₅₀ (μM)	SI
(RS)-fluoxetine	3.2 ±0.95	29.32 ±0.35	9.1
(S)-fluoxetine	0.4 ±0.15	28.63 ±1.40	71.6
252	0.41 ±0.27	>50	>121.9
262	1.12 ±0.28	>50	>44.6

Compound **262** was also tested against other representative enterovirus species. The compound showed the same trend observed for the (S)-fluoxetine and the guanidine analogue **252**, and with the data previously reported.⁴⁸ Compound **262** was able to inhibit the viral replication of EV-D68, HRV-A2 and HRV-B14 with EC₅₀ comparable to the (S)-fluoxetine and antiviral activity against EV-A71 and PV-1 was not observed.

Table 8. Antiviral evaluation of compound 262 against other enteroviruses. Compounds were tested in a multicycle CPE-reduction assay in Hela R19 cell-line infected with different enteroviruses. MTS assay was used for the readout. The same assay was used on un-infected cells for the evaluation of CC₅₀. Data represents mean values ± SD calculated from three independent experiments and are expressed in μM. NA = not active.

Compound	EC ₅₀ EV-D68 (Fermon)	EC ₅₀ EV-A71 (BrCr)	EC ₅₀ PV-1 (Sabin1)	EC ₅₀ HRV-A2	EC ₅₀ HRV-B14	CC ₅₀
(RS)-fluoxetine	1.85 ±0.10	NA	NA	NA	NA	29.32 ±0.35
(S)-fluoxetine	0.67 ±0.22	NA	NA	7.95 ±0.39	6.34 ±1.02	28.63 ±1.40
252	0.34 ±0.05	NA	NA	NA	2.11 ±0.34	>50
262	0.31 ±0.06	NA	NA	15.15 ±1.13	10.85 ±1.86	>50

10.2 *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)furan-2-carboxamide analogues

To explore the shared chemical features between compound **262** and the (*S*)-flouxetine (see figure 34), three analogues were designed. These *N*-benzyl-2-(methylamino)-*N*-phenylacetamide compounds were designed to introduce the methylamino group instead of the furan ring of compound **262**. The first of this series of the compound was **267**, in which the methylamino group replaced the furan ring. In compound **268** and **269**, in addition to the methylamine group introduced in **267**, the trifluoromethyl substituent was introduced on the benzyl ring and on the phenyl ring respectively, instead of the fluoro and methoxy substituent, to understand how the trifluoromethyl group could affect the antiviral activity of the compound.

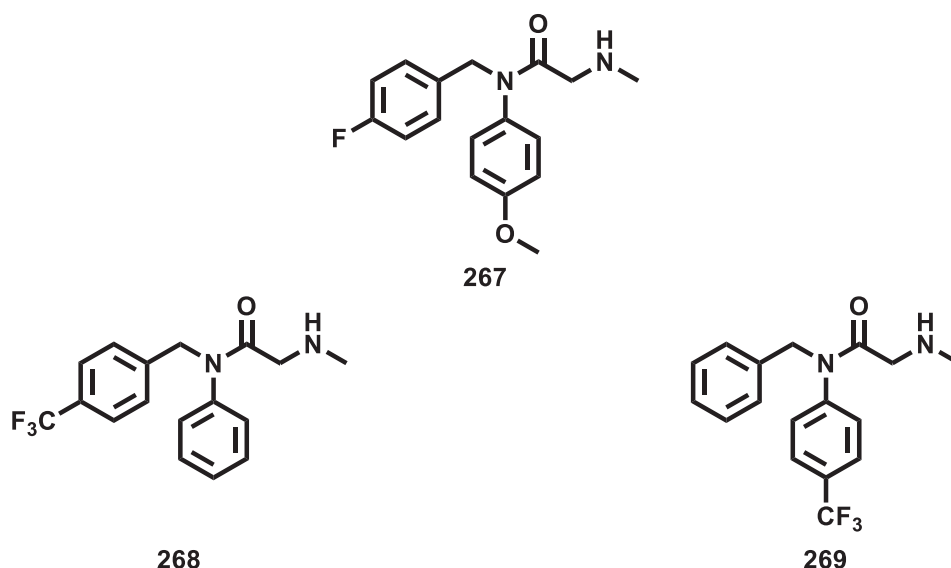


Figure 34. Designed analogues of **262** having the methyl amine group instead of the furan ring.

In order to investigate the possibility to introduce the trifluoromethyl feature of the (*S*)-flouxetine in compound **262**, two other analogues, compound **268** and **269** were designed. In compound **270**, the *N*-(4-fluorobenzyl)-4-methoxyaniline moiety was replaced by the *N*-(4-(trifluoromethyl)benzyl)aniline moiety. In this compound, the trifluoromethyl group should be in the same position as the trifluoromethyl group of the (*S*)-flouxetine. On the other compound **271**, instead, the *N*-(4-fluorobenzyl)-4-methoxyaniline moiety was replaced by the *N*-benzyl-4-(trifluoromethyl)aniline moiety. In this case, the trifluoromethyl group should not affect the antiviral activity, because it is located on the aniline ring (see figure 35).

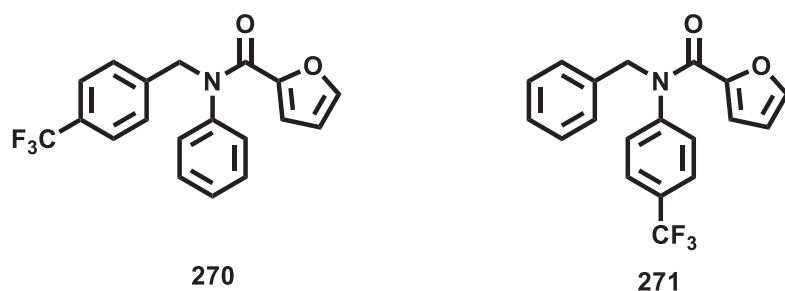
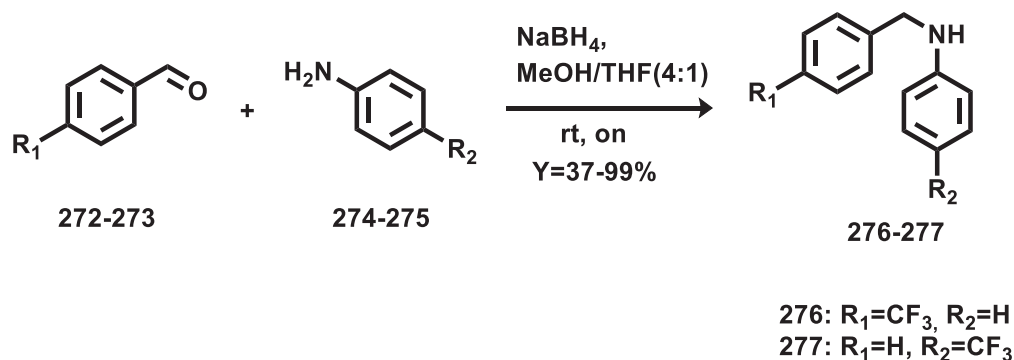


Figure 35. Designed analogues of 262 with the trifluoromethyl group.

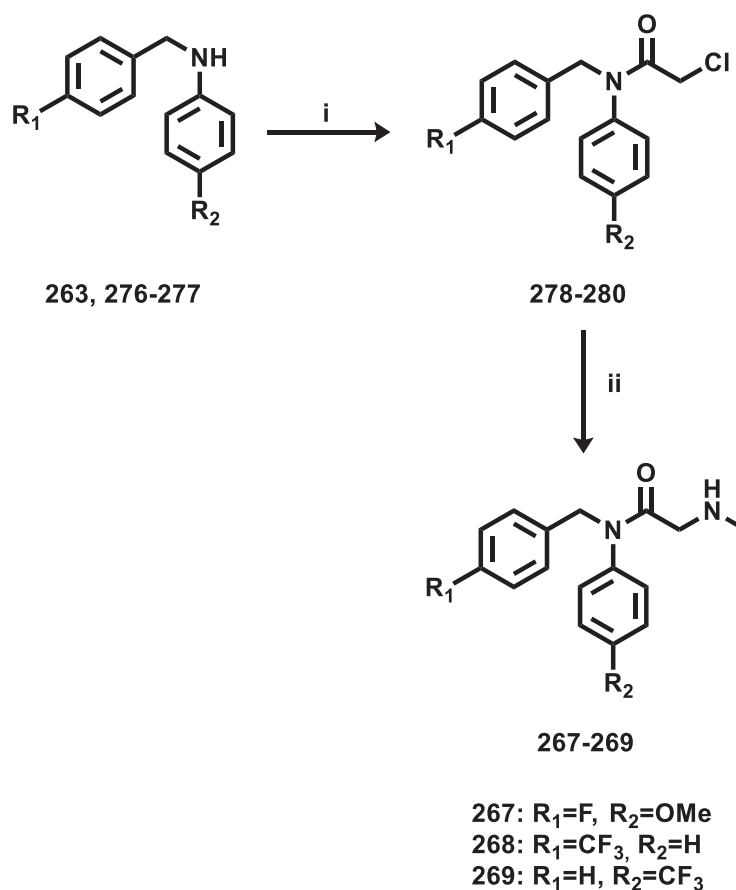
10.2.1 Synthesis of *N*-benzyl-2-(methylamino)-*N*-phenyl acetamides

To obtain the *N*-benzyl-2-(methylamino)-*N*-phenyl acetamides **267-269**, the *N*-(4-(trifluoromethyl)benzyl)aniline **276** and *N*-benzyl-4-(trifluoromethyl)aniline **277** were synthesised by reductive amination, using the same procedure for the synthesis of *N*-(4-fluorobenzyl)-4-methoxyaniline **263** (scheme 20).



Scheme 20. Synthesis of *N*-(4-(trifluoromethyl)benzyl)aniline **276** and *N*-benzyl-4-(trifluoromethyl)aniline **277**.

The *N*-benzyl-2-(methylamino)-*N*-phenyl acetamides **267-269** were then synthesised by a two-step synthesis (scheme 21). In the first step, the *N*-benzylanilines **263** and **276-277** were used to prepare the *N*-benzyl-2-chloro-*N*-phenylacetamides **278-280**, by amide bond formation between the *N*-benzylanilines and 2-bromoacetyl chloride in basic conditions. Finally, the chloro atom of the *N*-benzyl-2-chloro-*N*-phenylacetamides **278-280**, was displaced by the methylamine to give the desired *N*-benzyl-2-(methylamino)-*N*-phenyl acetamides **267-269**.



Scheme 21. Synthesis of the *N*-benzyl-2-(methylamino)-*N*-phenylacetamides 267-269.
i) 2-bromoacetyl chloride, TEA, DCM, rt, 1h, Y=64-87%; *ii)* methylamine, EtOH, rt, on, Y=64-66%.

10.2.2 Synthesis of *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl) furan-2-carboxamide analogues

The two others analogue of compound **262**, in which the *N*-(4-fluorobenzyl)-4-methoxyaniline moiety was replaced by the *N*-(4-(trifluoromethyl)benzyl)aniline moiety (**270**) and *N*-benzyl-4-(trifluoromethyl)aniline moiety (**271**) (figure 36), were synthesised starting from the previously synthesised *N*-benzylanilines **276-277**.

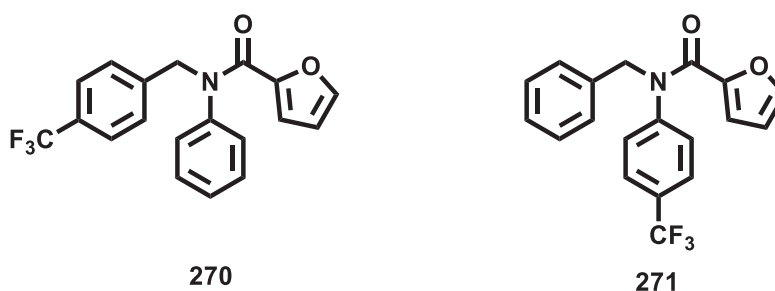
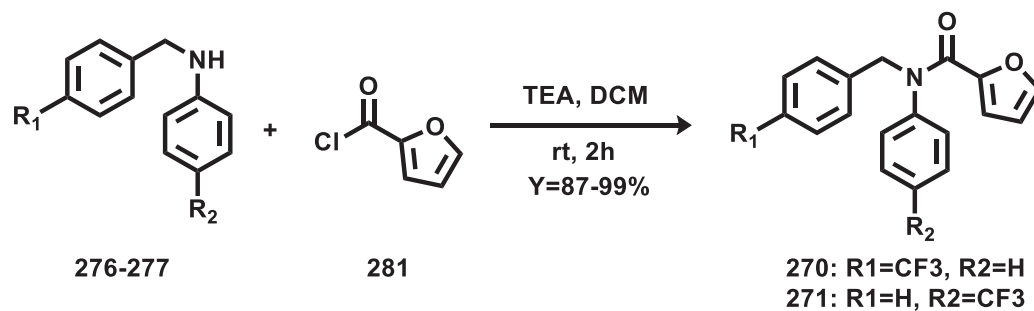


Figure 36. Designed analogues of 262 with the trifluoromethyl group.

The *N*-benzylanilines **276-277** were coupled with the furan-2-carbonyl chloride **215**, in basic condition to give the desired *N*-phenyl-*N*-(4-(trifluoromethyl)benzyl)furan-2-carboxamide **270** and *N*-benzyl-*N*-(4-(trifluoromethyl)phenyl)furan-2-carboxamide **271** (scheme 22). The furan-2-carbonyl chloride was used as activated carboxylic acid to avoid the use of a coupling agent.



Scheme 22. Synthesis of *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl) furan-2-carboxamide analogues **270-271**.

10.2.3 Antiviral evaluation of *N*-(4-fluorobenzyl)-*N*-(4-methoxy phenyl)furan-2-carboxamide analogues

The newly synthesised *N*-benzyl-2-(methylamino)-*N*-phenyl acetamides **267-269** and the *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl) furan-2-carboxamide analogues **270-271** were biological evaluated on the multicycle-CPE antiviral assay. The substitution of the furanic ring with the methylamine group in compound **267**, as can be seen from table 9, was not related to antiviral activity. Also, the other designed compound **268-269** were not active on CVB3. The results obtained on *N*-benzyl-2-(methylamino)-*N*-phenyl acetamides **267-269** are suggesting that probably the furan ring of compound **262** is interacting with 2C in a different way than the (*S*)-fluoxetine.

Table 9. Antiviral evaluation of compounds **267-271**. Compounds were tested in a multicycle CPE-reduction assay in Hela R19 cell-line infected with CVB-3 virus. MTS assay was used for the readout. The same assay was used on un-infected cells for the evaluation of CC₅₀. Data represents mean values \pm SD calculated from three independent experiments and are expressed in μ M. SI = selectivity index (CC₅₀/EC₅₀).

	EC ₅₀ CVB3 (μ M)	CC ₅₀ (μ M)	SI
262	1.12 \pm 0.28	>50	>44.6
267	NA	>50	-
268	NA	>50	-
269	NA	>50	-
270	0.84 \pm 0.03	>50	>50
271	2.42 \pm 0.42	>50	>20

Compound **270**, in which the *N*-(4-fluorobenzyl)-4-methoxyaniline moiety of **262** was replaced by the *N*-(4-(trifluoromethyl)benzyl)aniline moiety, showed retention of antiviral activity compared to **262**. The *N*-benzyl-4-(trifluoromethyl)aniline moiety of compound **271** was also related with retention of antiviral activity but with a ~2-fold lower potency compared with **262**.

Table 10. Antiviral evaluation of compounds 267-271 against different enterovirus species. Compounds were tested in a multicycle CPE-reduction assay in Hela R19 cell-line infected with different enteroviruses. MTS assay was used for the readout. The same assay was used on un-infected cells for the evaluation of CC_{50} . Data represents mean values \pm SD calculated from three independent experiments and are expressed in μ M. NA = not active.

Compound	EC_{50} EV-D68 (Fermon)	EC_{50} EV-A71 (BrCr)	EC_{50} PV-1 (Sabin1)	EC_{50} HRV-A2	EC_{50} HRV-B14	CC_{50}
262	0.31 \pm 0.06	NA	NA	15.15 \pm 1.13	10.85 \pm 1.86	>50
270	1.395 \pm 0.2	0.39 \pm 0.005	21 \pm 0.85	NA	4.76 \pm 2.8	>50
271	3.16 \pm 0.12	NA	NA	NA	NA	>50

Compound **270** and **271**, in which retention of antiviral activity was observed, were tested against the other representative enterovirus species. Surprisingly, the introduction of the trifluoromethyl group on the benzylic ring in compound **270**, provide a pan-inhibitory antiviral activity to the compound. This pan-inhibitory activity observed for compound **270** was not observed in compound **271**, in which the trifluoromethyl group is located on the aniline ring.

10.3 Furamide analogues

After the discovery that the scaffold of the compound **262** is suitable for pan-inhibitory antiviral activity, the role of the furamide moiety was further investigated. Different analogues were designed in order to identify a potential candidate group for the replacement of the furamide moiety (figure 37).

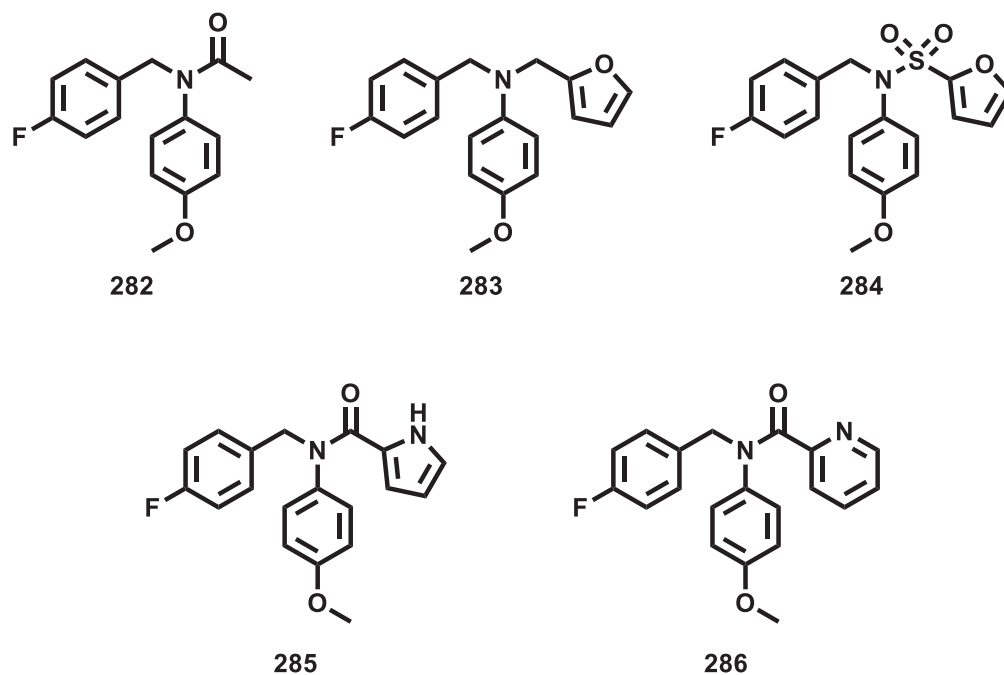
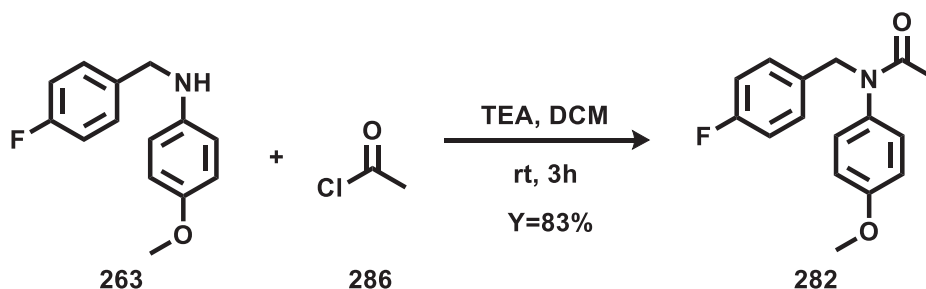


Figure 37. Furamide analogues.

In order to understand the importance of the furan ring for the antiviral activity, it was replaced by a methyl group in compound **282**. Furthermore, in compound **283** and **284**, the role of the amide bond was investigated, by replacing the carbonyl group with a methylene bridge and a sulfonyl group respectively. Lastly, the furan ring was replaced by a pyrrole and pyridine in compound **285** and **286**.

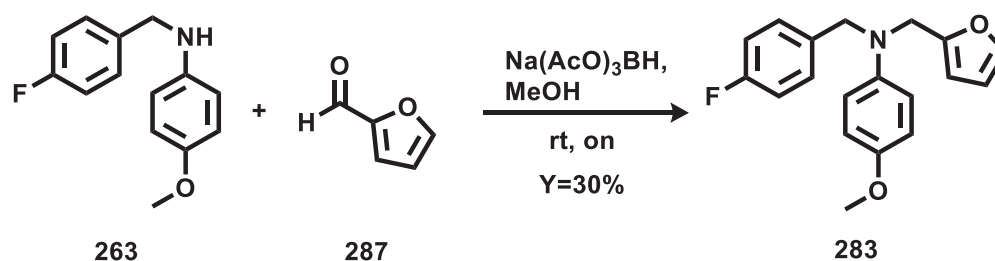
10.3.1 Synthesis of furamide analogues

The *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)acetamide (**282**) was synthesised by acetylation of the *N*-(4-fluorobenzyl)-4-methoxyaniline **263** in the presence of triethylamine as a base (scheme 23).



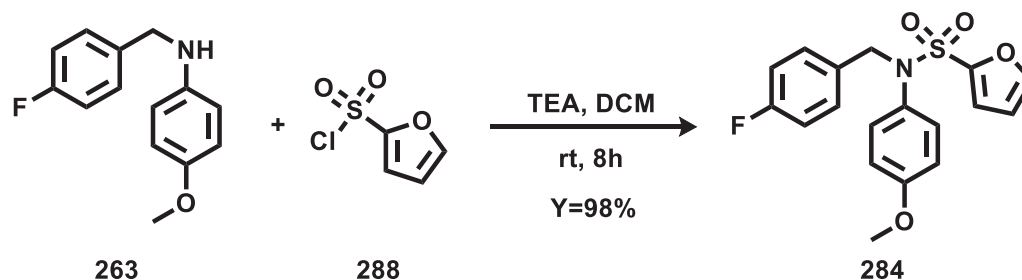
Scheme 23. Synthesis of *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)acetamide **282**.

The reductive amination between the *N*-(4-fluorobenzyl)-4-methoxyaniline **263** and the furan-2-carbaldehyde **287**, using the sodium triacetoxyborohydride as a reductive agent, give the desired *N*-(4-fluorobenzyl)-*N*-(furan-2-ylmethyl)-4-methoxyaniline **283**.



Scheme 24. Synthesis of *N*-(4-fluorobenzyl)-*N*-(furan-2-ylmethyl)-4-methoxyaniline **283**.

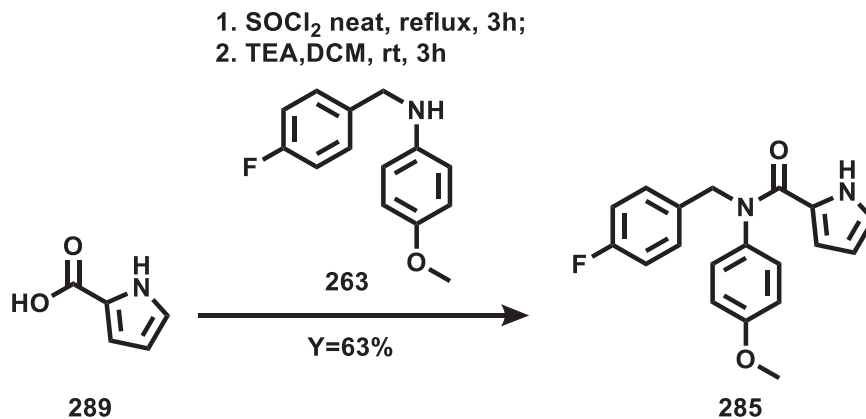
The furan-2-sulfonamide analogue **284** was synthesised by S_N2 displacement between the *N*-benzylaniline **263** and the furan-2-sulfonyl chloride **288**, in the presence of triethylamine as a base (scheme 25).



Scheme 25. Synthesis of *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)furan-2-sulfonamide **284**.

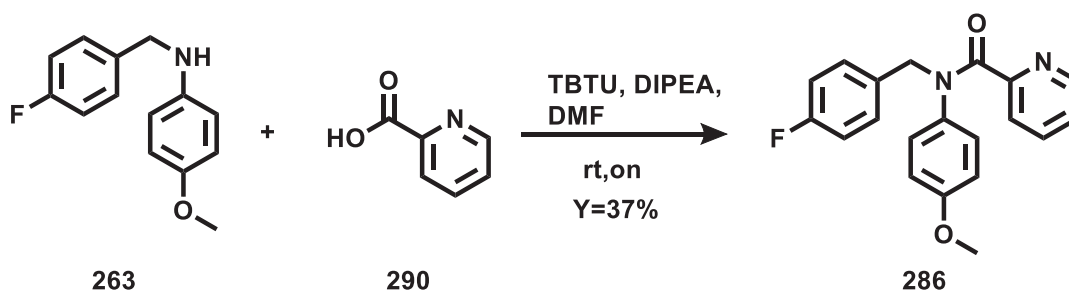
For the synthesis of the pyrometer-2-carboxamide analogue **285**, the first attempt of coupling between the 1H-pyrrole-2-carboxylate acid **289** and the *N*-(4-fluorobenzyl)-4-methoxyaniline **263** was performed using TBTU as a coupling reagent, but the formation of the desired product was not observed. The desired *N*-(4-fluorobenzyl)-*N*-(4-

methoxyphenyl)-1H-pyrrole-2-carboxamide **285** was obtained by converting the carboxylic acid of **289** in the respective acrylic chloride, which was then coupled with the *N*-benzylaniline **263** (scheme 26).



Scheme 26. Synthesis of *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)-1H-pyrrole-2-carboxamide **285**.

The last furamide analogue, the *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)picolinamide **286**, was obtained by Aimee bond formation between the *N*-benzylaniline **263** and the picolinic acid, using TBTU as coupling reagent.



Scheme 27. Synthesis of *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)picolinamide **286**.

10.3.2 Antiviral evaluation of furamide analogues

From the antiviral evaluation of the furamide analogues, reported in table 11, the substitution of the furan ring with a methyl group in compound **282** and the substitution of the carbonyl group with a methylene bridge in compound **283** were related with loss of antiviral activity. Also, the replacement of the furanamide with the furansulfonamide in compound **284** loss of antiviral activity was observed. From this information, can be deduced that the furan ring is important for the interaction with the protein, like also the carbonyl group part of the amine bond.

The pyrrole analogue **285** was associated with antiviral activity retention with an EC₅₀ comparable to **262**. A loss of 5-fold in antiviral activity was observed when a pyridine ring replaced the furan ring.

Table 11. Antiviral evaluation of compounds 282-286. Compounds were tested in a multicycle CPE-reduction assay in Hela R19 cell-line infected with CVB-3 virus. MTS assay was used for the readout. The same assay was used on un-infected cells for the evaluation of CC₅₀. Data represents mean values ± SD calculated from three independent experiments and are expressed in μM. SI = selectivity index (CC₅₀/EC₅₀).

Compound	EC ₅₀ CVB3 (μM)	CC ₅₀ (μM)	SI
262	1.12 ±0.28	>50	>44.6
282	NA	>50	-
283	NA	>50	-
284	NA	>50	-
285	1.04 ±0.16	>50	>16
286	6.18 ±0.12	>50	>7.8

The two compounds, **285** and **286**, active on CVB3, were also tested on the representative panel of enteroviruses. The compounds were only able to inhibit the viral replication of EV-D68. They were also not showing antiviral activity on the HRV species.

Table 12. Antiviral evaluation of compounds 285-286 against different enterovirus species. Compounds were tested in a multicycle CPE-reduction assay in Hela R19 cell-line infected with different enteroviruses. MTS assay was used for the readout. The same assay was used on un-infected cells for the evaluation of CC₅₀. Data represents mean values ± SD calculated from three independent experiments and are expressed in μM. NA = not active.

Compound	EC ₅₀ EV-D68 (Fermon)	EC ₅₀ EV-A71 (BrCr)	EC ₅₀ PV-1 (Sabin1)	EC ₅₀ HRV-A2	EC ₅₀ HRV-B14	CC ₅₀
262	0.31 ±0.06	NA	NA	15.15 ±1.13	10.85 ±1.86	>50
285	1.39 ±0.42	NA	NA	NA	NA	>50
286	11.82 ±0.56	NA	NA	NA	NA	>50

10.4 Conclusions

The *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)furan-2-carboxamide **262** was reported to inhibit the viral replication of CVB3, by targeting the non-structural protein 2C. Like fluoxetine, compound **262** was able to inhibit HEV-B species but not HEV-A and HEV-C. Both compounds were affected by the same mutations, C179F and the mutations on the AGSINA loop, suggesting that both compounds might potentially share the same mechanism of action. Furthermore, the chemical structure similarity observed between compound **262** and (*S*)-fluoxetine, was suggesting the possibility to avoid the chiral centre on fluoxetine for further development, by using the scaffold of compound **262**. Several analogues of compound **262** were designed to introduce the relevant features of fluoxetine for the antiviral activity into compound **262** and to explore the importance of the furamide group for the antiviral activity.

Interestingly, the replacement of the fluoro group on the benzylic ring of compound **262** by the trifluoromethyl group, was associated with pan-inhibitor antiviral activity, suggesting that the scaffold of **262** is suitable for the development of a pan-enterovirus inhibitor of the viral replication. Besides, the antiviral evaluation of the tested analogues of **262** is suggesting that the amide bond is essential for retention of antiviral activity. The role of the heterocyclic furan ring was also explored, by replacing the furan by a pyrrole or a pyridine ring. In both cases, the antiviral activity was retained, but the pyrrole ring was associated with an EC₅₀ comparable to **262**.

These results obtained on the *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)furan-2-carboxamide **262** analogues, represent a promising starting point for further development of a novel class of enterovirus pan-inhibitor.

Chapter 11: 2C ATPase pocket

11.1 2C protein ATPase activity

The 2C protein is classified as AAA+ (ATPases Associated with various cellular Activities) SF3 helicase superfamily, in which six monomeric 2C proteins assemble into a hexameric structure to form a ring. It was observed by electron microscopy in different enteroviruses for the 2C protein.^{19,20,23,29} The 2C hexamer structure is responsible for the ATP-dependent RNA helicase activity, which it unwinds the viral RNA in 3'-to-5' direction.²¹ The hexameric helicases are characterised to have six ATPase sites, forming between the hexamer units, and this is also true for the 2C protein hexamers. As for other helicases, the ATPase activity of the 2C protein is coupled with the translocase activity of the hexamer.⁶⁹⁻⁷⁰

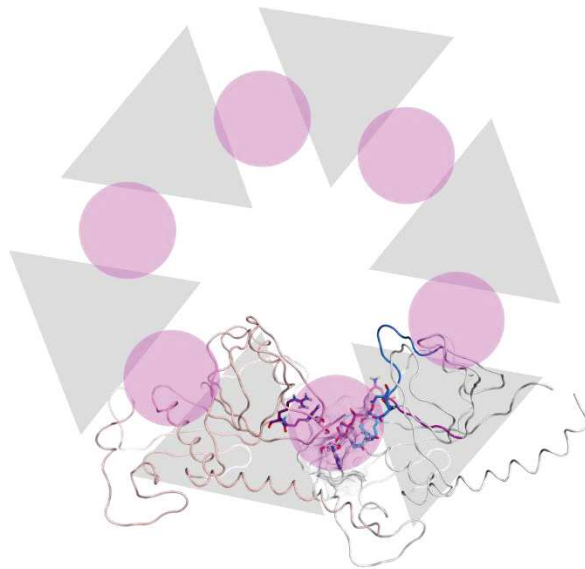


Figure 38. Representation of 2C protein hexameric ring. 2C unit represented as a grey triangle and the ATPase site as a purple circle.

The ATPase pockets, as can be seen in figure 38, are located between two 2C units, and the hydrolysis of the ATP occur through a movement of the two units. This movement brings the arginine finger close to the triphosphate moiety of the ATP during the hydrolysis.

For the identification of new inhibitors of the ATPase activity of the 2C protein, molecular dynamics simulations were used to generate several conformations of the ATPase pocket. These conformations were then used to perform an ensemble docking approach.

11.2 Ensemble docking

In general, during a molecular docking approach, several generated-conformations of a ligand are docked into a rigid receptor model and evaluated for their binding affinity. In a situation in which the pocket in the receptor could have different conformations, MD-generated conformation ensemble of the protein can be used to identify new inhibitor

candidate, addressing the flexibility of the binding site.⁷¹ The MD simulations can generate a high number of protein conformations and performing an ensemble docking using all the generated conformations is related with a very high molecular docking cost in term of time and computing cost, especially when docking a large set of ligands. This issue can be addressed by selecting the more representative and significant conformations using the RMSD clustering approach, in which each representative structure conformation of a cluster is representative for each of the cluster member.⁷²⁻⁷³

11.2.1 MD simulation and clustering analysis

Chain A and chain F of the published EV-A71 2C structure (PDB: 5GRB), in which the ATP- γ -S is interacting with the Walker A and B motifs, and the arginine finger is close to the triphosphate moiety, were used to prepare the model system for the MD simulations. The system was used to perform three independent 500ns MD simulations, recording the frames every 20 ps during the simulations, obtaining three trajectories of 25000 frames each. To reduce the computational cost only the last 200 ns of each MD simulation were used for the clustering analysis. The density-based spatial clustering of applications with noise (DBSCAN) algorithm was used to cluster the atom-positional RMSD of all atoms in the binding site close to the ATP- γ -S as the similarity criterion in each of the obtained frame.⁷⁴⁻⁷⁵ This algorithm is grouping the frames in the high-density region into clusters, and the frames in the low-density region are considered as noise. Eight clusters were obtained considering all the frames with a neighboring distance of 1Å ($\epsilon=1\text{\AA}$) with 30 as a minimum number of neighboring frames (figure 39).

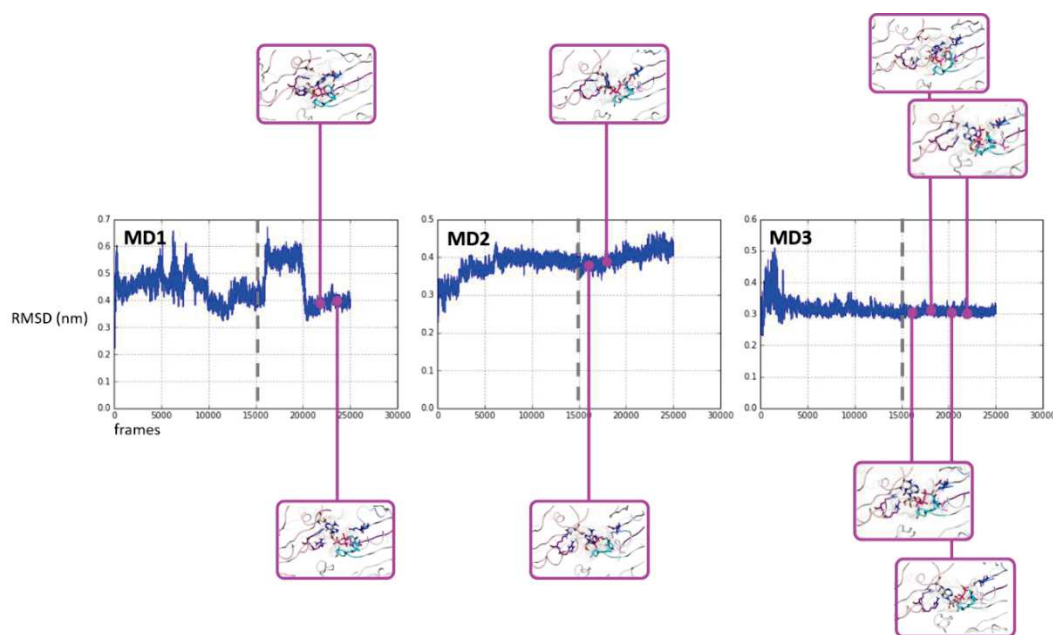


Figure 39. Centroid frames obtained from the clusterisation analysis of the three MD trajectories. The three graphs represent the atom-positional RMSD of all atoms in the binding site close to the ATP- γ -S. In pink the centroid frames of the most representative clusters.

For each of the obtained cluster, the centroid frames were used as ensemble conformations of the ATPase pocket for the ensemble docking.

11.2.2 Ensemble Docking

The centroid frames, which they were obtained from the cluster analysis, they were used for the ensemble docking. The Glide Virtual Screening Workflow was used to virtual screening the SPECS library of 342047 commercially available compounds on the eight centroid frames. The best 25% docked compounds with the VS algorithm were then docked using Glide SP and rescored using Glide XP. The obtained docking results from each of the centroid frame were merged, and only the ligands obtained as a result in more than one centroid frame were taking in consideration, giving 2776 compounds (figure 40).

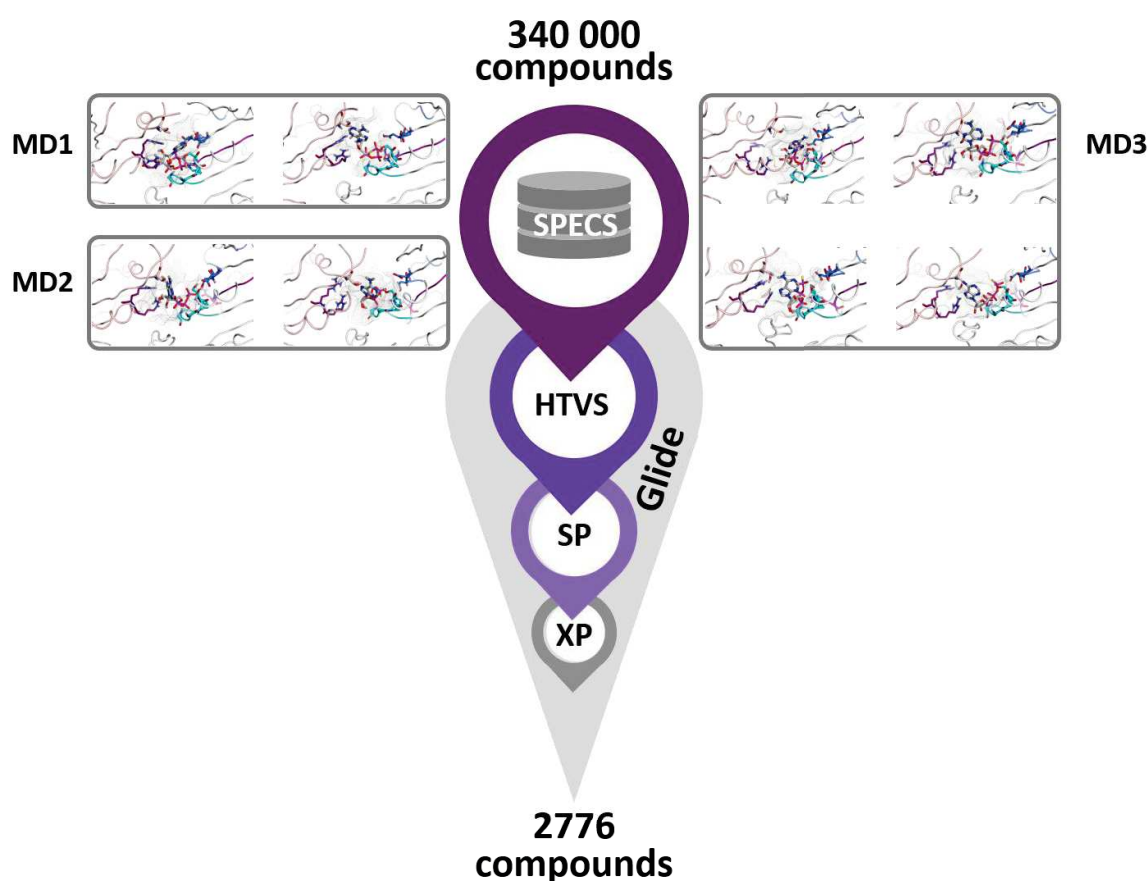


Figure 40. Ensemble docking workflow representation. Glide Virtual Screening Workflow was used to virtual screening the SPECS library of commercially available compounds on the eight centroid frames. MD = Molecular Dynamics simulations. Glide HTVS = Glide Virtual Screening algorithm, Glide SP = Glide Standard Precision algorithm. Glide XP = Glide Extra Precision algorithm.

The compounds were then clustered using MACCS Fingerprints and Tanimoto similarity matrix to select the more different compound structures. From the compounds obtained

from the fingerprints clusterisation, only compounds with less than one carboxylic group were selected, a log P lower than three and with a molar weight (MW) higher than 250 Dalton, from which 50 compounds were selected to be visually inspected. In the end, 18 compounds were selected from the visual inspection on the base of their ability to bind to the residues of the Walker A motif with which the triphosphate moiety is interacting. The compounds were bought and biological evaluated (see the appendix for the chemical structure of the selected compounds).

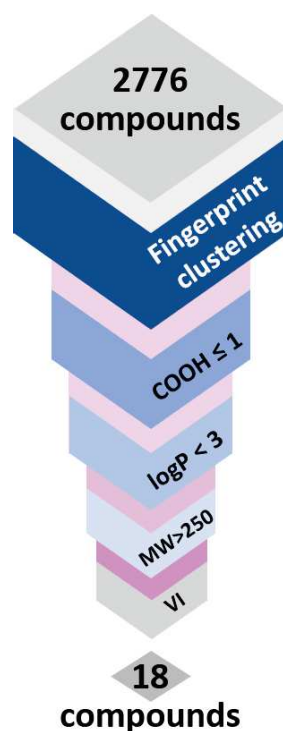


Figure 41. Selection workflow of the ensemble docking results. Compounds were clustered using MACCS Fingerprints and Tanimoto similarity matrix. Then, only compounds with less than one carboxylic group were selected, a log P lower than three and with a molar weight (MW) higher than 250 Dalton. Finally, 18 compounds were selected after visual inspection.

11.3 Antiviral evaluation of the selected compounds

The selected compounds were evaluated for their antiviral activity on CVB3 and EV-A71 in the multicycle-CPE antiviral assay, and only compound **291** showed an EC₅₀ of 15.3 μM against CVB3 and an EC₅₀ of 35 μM against EV-A71, with a CC₅₀ higher than 70 μM.

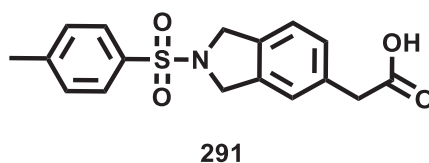


Figure 42. Chemical structure of compound 291.

11.4 Conclusions

In this part of the project, MD simulations were performed to address the flexibility related to the ATPase pocket. The MD trajectories were clustered, and the centroid frames were used to perform an ensemble docking, and eighteen compounds were selected. The selected compounds were bought and biologically evaluated on CVB3 and EV-A71. One of the selected compounds showed antiviral activity against both viruses. Further biological investigations are ongoing in order to determine if the compound is targeting the 2C protein. If the 2C protein is confirmed to be the target, this compound represents a starting point for future structural modifications of the compound.

Chapter 12: Experimental section

12.1 General information

All chemicals, reagents and solvents were purchased from Sigma-Aldrich, FluoroChem, Apollo Scientific, TCI UK or purified by standard techniques.

Thin Layer Chromatography

Silica gel plates (Merck Kieselgel 60F254) were used and were developed by the ascending method. After solvent evaporation, compounds were visualised by irradiation with UV light at 254 nm and 366 nm.

Column Chromatography

Column Chromatography was performed using an automated Isolera One System (Biotage), using Biotage pre-packed silica cartridges (SNAP and KP types). Samples were applied as a concentrated solution in the same eluent. Fractions containing the product were identified by TLC, combined and the solvent removed *in vacuo*.

UPLC-MS analysis

UPLC-MS analysis was conducted on a Waters UPLC system with both Diode Array detection and Electrospray (+'ve and -'ve ion) MS detection. The stationary phase was a Waters Acquity UPLC BEH C18 1.7 μ m 2.1x50mm column. The mobile phase was H₂O containing 0.1% Formic acid (A) and MeCN containing 0.1% Formic acid (B). Column temperature: 40°C. Sample diluent: acetonitrile. Sample concentration 10 μ g/mL. Injection volume 2 μ L. Two methods were used:

Linear gradient standard method A: 90% A (0.1 min), 90%-0% A (2.6 min), 0% A (0.3 min), 90% A (0.1 min); flow rate 0.5 mL/min.

Linear gradient standard method B: 90% A (0.1 min), 90%-0% A (2.1 min), 0% A (0.8 min), 90% A (0.1 min); flow rate 0.5 mL/min.

Linear gradient standard method C: 90% A (0.1 min), 90%-0% A (1.5 min), 0% A (1.4 min), 90% A (0.1 min); flow rate 0.5 mL/min.

NMR Spectroscopy

¹H, ¹³C, ¹⁹F NMR spectra were recorded on a Bruker AVANCE 500 spectrometer (500 MHz ¹H, 150 MHz ¹³C and 471 MHz ¹⁹F respectively) and auto calibrated to the deuterated solvent reference peak. Chemical shifts are given in δ relative to tetramethylsilane (TMS); the coupling constants (J) are given in Hertz. TMS was used as an internal standard (δ = 0 ppm) for ¹H NMR and CDCl₃ served as an internal standard (δ = 77.0 ppm) for ¹³C NMR.

Molecular modelling

A PC equipped with a 1.80 GHz Intel Xeon CPU (8 cores), 32GB of RAM and a GPU Zotac GeForce GTX 1080-Ti Mini with 11GB GDDR5X of dedicated RAM, running Ubuntu 14.04 LTS was used for molecular modelling studies. The molecular dynamics simulations were performed using Desmond package (Schrödinger Release 2018-1: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2018. Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY, 2018).⁷⁶

The ensemble docking was performed using Glide package (Schrödinger Release 2016-4: Glide, Schrödinger, LLC, New York, NY, 2016.)⁷⁷⁻⁷⁹

VMD software (<http://www.ks.uiuc.edu/Research/vmd/>) was used to convert the Desmond output format of the trajectories into .dcd trajectory format.⁸⁰ The Jupyter notebook was used as interactive command shell Python coding (Python v3.6.8).⁸¹ The MD analysis and the clustering analysis were performed using the SciKit-learn Python library and SciPy Python library. PyPlot Python library was used to generate MD simulation analysis and clustering analysis graph. Pictures of molecular modelling were prepared using MOE.⁸²

12.2 Molecular dynamics simulations of Fluoxetine enantiomers

The docking pose of (R)-fluoxetine and (S)-fluoxetine into 2C-site A and 2C-site B of CVB3 were used to prepare the model systems for the MD simulations. OPLS3 was used as a force field. Each of the protein-ligand complex systems was prepared, placing the protein-ligand complex in a cubic box (buffer 10Å) and filling the space using water molecules. TIP3P model was used as a solvent model. The negative charges on the protein were neutralized, adding Na⁺ atoms to the system. Magnesium chloride (10 mM) was added to the box to simulate physiological conditions.

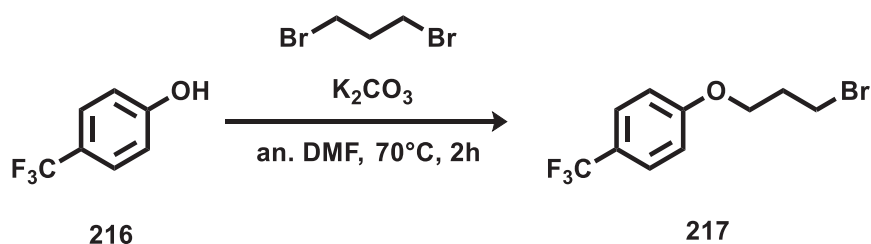
The protein-ligand complex systems were then equilibrated for 112ps at 10 K in an NVT ensemble and then simulated for 48 ps at a constant pressure of 1 atm using the NPT ensemble. The equilibrated systems were then MD simulated for 100 ns at constant temperature (300 K) and pressure recording snapshots every 160 ps. The RMSD values of C α atoms during the simulations were used as an indication of structural stability and simulation integrity (figure 43 in the appendix).

The estimated $\Delta G_{\text{binding}}$ was calculated using the Desmond command-line script `thermal_mmgbsa.py`. After splitting the trajectory file of the MD simulation into snapshots, the script is calculating the average computed binding energy of the ligand.

12.3 Synthesis of Fluoxetine fragments

Synthesis of 1-(3-bromopropoxy)-4-(trifluoromethyl)benzene (**217**)

(C₁₀H₁₀BrF₃O, M.W.: 283.09)



To a solution of 1,3-dibromopropane (6.17 mmol) and potassium carbonate (4.63 mmol) in DMF (3 mL), 4-(trifluoromethyl)phenol **216** (3.08 mmol) in DMF (1.2 mL) were added dropwise. The reaction mixture was stirred at room temperature for two hours and then heated to 70°C for another two hours. The mixture was filtrated, diluted with ethyl acetate (10 ml) and washed with water (3x10 ml). The organic layer was dried over sodium sulphate and evaporated under reduced pressure. The residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP Sil 50g, n-hexane - EtOAc 100:0 v/v increasing to 70:30 v/v in 20 CV) to give pure 1-(3-bromopropoxy)-4-(trifluoromethyl)benzene (**217**) as a yellow oil.

T.L.C. System: n-Hexane/EtOAc 80:20 v/v, R_f: 0.7

Yield: 207 mg (47%)

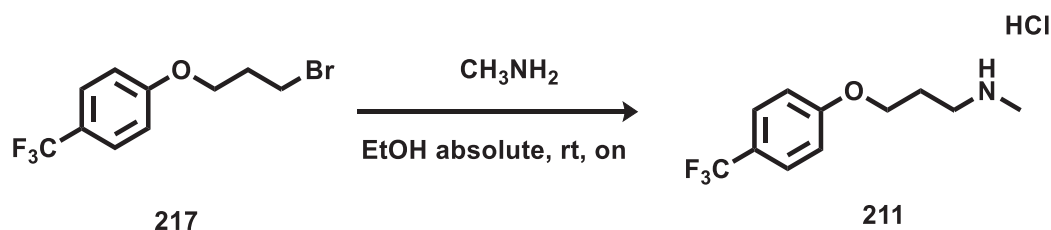
Purity: >95%

UPLC-MS method C: Rt: 2.175, MS (ESI)⁺: ND

¹H NMR (DMSO) δ: 7.65 (d, *J* = 8.7 Hz, 2H), 7.14 (d, *J* = 8.7 Hz, 2H), 4.17 (t, *J* = 6.0 Hz, 2H), 3.68 (t, *J* = 6.6 Hz, 2H), 2.28 (p, *J* = 6.3 Hz, 2H).

¹³C NMR (DMSO) δ: 161.67 (C, C-aromatic), 127.44 (q, *J* = 3.7 Hz, CH, C-aromatic), 125.02 (q, *J* = 271.0 Hz, CF₃), 121.73 (q, *J* = 32.1 Hz, C, C-aromatic), 115.44 (CH, C-aromatic), 32.08, 31.50 (CH₂).

¹⁹F NMR (DMSO) δ: -59.83 (s, 3F).

Synthesis of *N*-methyl-3-(4-(trifluoromethyl) phenoxy) propan-1-amine (211)**(C₁₁H₁₅ClF₃NO, M.W.: 269.69)**

A solution of 1-(3-bromopropoxy)-4-(trifluoromethyl) benzene **171** (0.80mmol) in EtOH absolute (0.8mL) were added dropwise at 0°C to a round bottom flask containing methylamine in absolute ethanol (2mL). The reaction mixture was stirred at room temperature overnight. The mixture was filtrated. The obtained residue was dissolved in DCM (10 ml). The organic layer was washed with sat. aq. NaHCO₃ and brine, dried over Na₂SO₄ and concentrated in vacuo. The residue was treated with HCl in diethyl ether. the resulting solid was then filtered and washed with diethyl ether to give 131 mg of *N*-methyl-3-(4-(trifluoromethyl) phenoxy) propan-1-amine hydrochloride (**164**) salt as a white powder. T.L.C. System: DCM/MeOH 95:5 v/v, Rf: 0.16

Yield: 131 mg (69%)

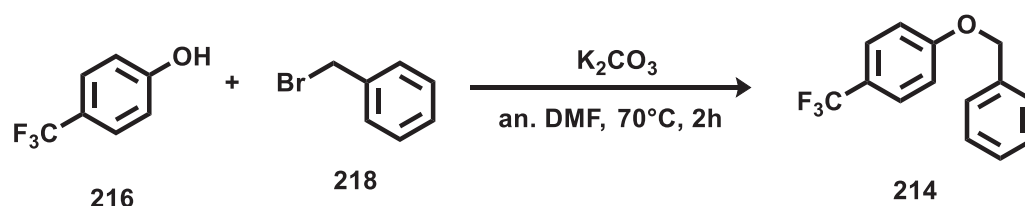
Purity: >95%

UPLC-MS method C: Rt: 1.521, MS (ESI)⁺: 234.1 [M+Na]⁺.

¹H NMR (DMSO) δ: 8.65 (s, 2H), 7.68 (d, *J* = 8.6 Hz, 2H), 7.13 (d, *J* = 8.6 Hz, 2H), 4.16 (t, *J* = 6.1 Hz, 2H), 3.06 (d, *J* = 7.4 Hz, 2H), 2.58 (s, 3H), 2.13 – 2.04 (m, 2H).

¹³C NMR (DMSO) δ: 161.58 (C, C-aromatic), 127.43 (q, *J* = 3.7 Hz, CH, C-aromatic), 125.02 (q, *J* = 271.1 Hz, CF₃), 121.76 (q, *J* = 32.2 Hz, C, C-aromatic), 115.48 (CH, C-aromatic), 65.58, 46.14(CH₂), 33.07 (CH₃), 25.71 (CH₂).

¹⁹F NMR (DMSO) δ: -59.78 (s, 3F).

Synthesis of 1-(benzyloxy)-4-(trifluoromethyl)benzene(214)**(C₁₄H₁₁F₃O, M.W.: 252.24)**

To a solution of 4-(trifluoromethyl) phenol **216** (1.23 mmol), in DMF (3mL), (bromomethyl)benzene **218** (1.55 mmol) and potassium carbonate (4,93 mmol) were

added. The obtained mixture was stirred at 105°C for four hours. After the reaction was complete, the mixture was filtrated, diluted with ethyl acetate (10 ml) and washed with water (3x10 ml). The organic layer was dried over sodium sulphate and evaporated under reduced pressure. The residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP Sil 10g, n-hexane -DCM 100:0 v/v increasing to 0:100 v/v in 12 CV) to give pure 1-(benzyloxy)-4-(trifluoromethyl)benzene **214** as a white powder.

T.L.C. System: n-Hexane/EtOAc 80:20 v/v, Rf: 0.77

Yield: 212 mg (75%)

Purity: >95%

UPLC-MS method C: Rt: 2.177, MS (ESI)⁺: ND

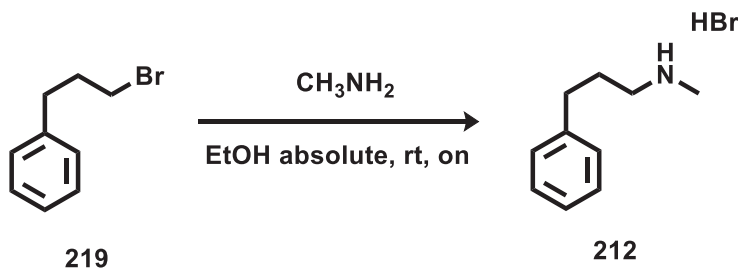
¹H NMR (DMSO) δ: 7.66 (d, *J* = 8.5 Hz, 2H), 7.49 – 7.44 (m, 2H), 7.44 – 7.38 (m, 2H), 7.38 – 7.32 (m, 1H), 7.20 (d, *J* = 8.5 Hz, 2H), 5.20 (s, 2H).

¹³C NMR (DMSO) δ: 161.64, 136.88 (C, C-aromatic), 128.98, 128.50, 128.25 (CH, C-aromatic), 127.42 (q, *J* = 3.7 Hz, CH, C-aromatic), 125.02 (q, *J* = 271.1 Hz, CF₃), 121.74 (q, *J* = 32.1 Hz, (C, C-aromatic)), 115.77 (CH, C-aromatic), 70.00 (CH₂).

¹⁹F NMR (DMSO) δ: -59.81 (s, 3F).

Synthesis of *N*-methyl-3-phenylpropan-1-amine (**212**)

(C₁₀H₁₆BrN, M.W.: 230.15)



To a round bottom flask containing methylamine in absolute ethanol (2.18 mL), a solution of (3-bromopropyl) benzene **219** (1.25 mmol) in EtOH absolute (0.87mL) was added dropwise at 0°C. The reaction mixture was stirred at room temperature overnight. The solid residue was filtered and washed with EtOH absolute to give *N*-methyl-3-phenylpropan-1-amine **212** as white powder.

T.L.C. System: DCM/MeOH 95:5 v/v, Rf: 0.42

Yield: 195 mg (75%)

Purity: >95%

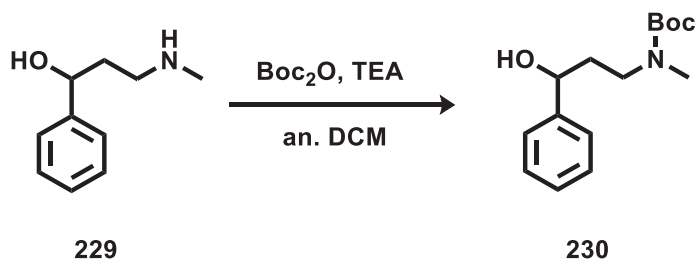
¹H NMR (DMSO) δ: 8.48 (s, 2H), 7.35 – 7.27 (m, 2H), 7.26 – 7.17 (m, 3H), 2.91 – 2.85 (m, 2H), 2.65 (t, *J* = 7.7 Hz, 2H), 2.55 (s, 3H), 1.94 – 1.85 (m, 2H).

¹³C NMR (DMSO) δ: 141.13 (C, C-aromatic), 128.92, 128.73, 126.57 (CH, C-aromatic), 48.30 (CH₂), 32.88 (CH₃), 32.32, 27.57 (CH₂).

12.4 Synthesis of Fluoxetine analogues

Synthesis of Boc protected 3-(methylamino)-1-phenylpropan-1-ol (**230**)

($C_{15}H_{23}NO_3$; M.W.: 265.35)



To a solution of 3-(methylamino)-1-phenylpropan-1-ol **229** (12.1 mmol) in DCM (48mL), di-tert-butyl dicarbonate (13.3 mmol) was added, the reaction mixture was stirred at room temperature overnight. The reaction mixture was then washed with NH_4Cl (3x30ml), brine (3x30ml), dried over Na_2SO_4 and concentrated under reduce pressure to give tert-butyl (3-hydroxy-3-phenylpropyl) (methyl)carbamate **230** as a colorless oil which, was used without further purification.

T.L.C. System: n-Hex/EtOAc 80:20 v/v, Rf: 0.28

Yield: 3.16 g (99%)

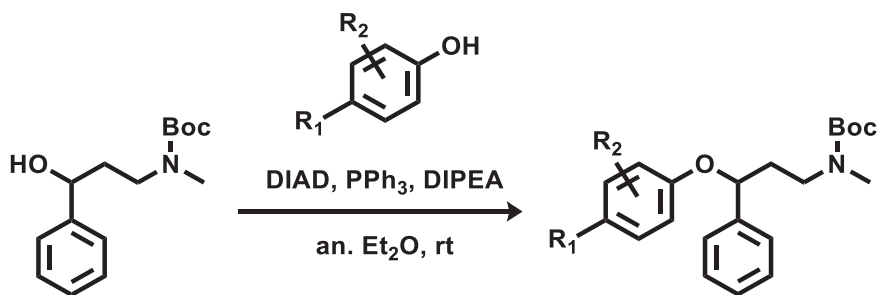
Purity: >95%

UPLC-MS method C: Rt: 1.883, MS (ESI)⁺: 288.1 $[M+Na]^+$, 192.1 $[C_{11}H_{14}NO_2 \cdot]^+$

¹H NMR (DMSO-d₆), δ : 7.36 – 7.29 (m, 4H), 7.30 – 7.20 (m, 1H), 5.22 (d, $J = 4.4$ Hz, 1H), 4.52 (dd, $J = 11.0, 6.2$ Hz, 1H), 3.18 (br s, 2H), 2.76 (s, 3H), 1.78 (br s, 2H), 1.35 (br s, 9H).

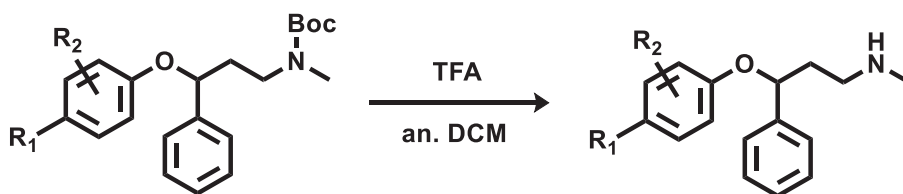
¹³C NMR (DMSO-d₆), δ : 155.22 (C=O), 146.33 (C, C-aromatic), 128.48, 127.17, 126.13 (CH, C-aromatic), 78.66 (C), 70.67 (CH), 46.03, 37.91 (CH₂), 34.46, 28.53 (CH₃).

General procedure 11: Synthesis of Boc protected N-methyl-3-phenoxy-3-phenylpropan-1-amines

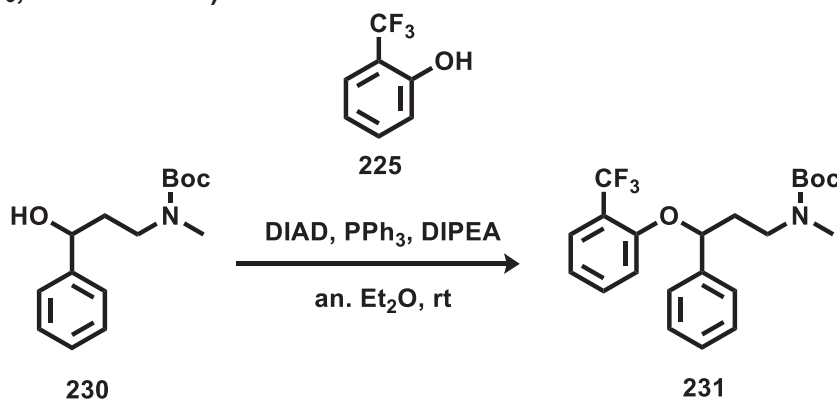


To a solution of triphenylphosphine (1.70 mmol) in diethyl ether (3.39 mL) cooled at 0°C, diisopropyl azodicarboxylate (1.14 mmol) in diethyl ether (0.57 mL) was added. After twenty minutes 3-(trifluoromethyl)phenol (1.36 mmol) in diethyl ether (2.26 mL) was added before adding the solution of tert-butyl (3-hydroxy-3-phenylpropyl)(methyl)carbamate (1.13 mmol) in diethyl ether (2.83 mL). The reaction was stirred at room temperature for two hours. The organic solvent was evaporated under reduced pressure and the crude residue was purified by flash column chromatography to afford tert-butyl (3-(phenoxy)-3-phenylpropyl)(methyl)carbamate.

General procedure 12: Synthesis of N-methyl-3-phenoxy-3-phenylpropan-1-amines



To a solution of tert-butyl (3-phenyl-3-(4-(trifluoromethyl) phenoxy)propyl)carbamate (0.84 mmol) in DCM (4.2 mL) at 0°C TFA (4.20 mmol) was added. The reaction was stirred at room temperature for two hours. The reaction mixture was diluted with DCM (8 mL) and was washed with sat. NaHCO₃ solution (3 x 10 mL) and brine (3 x 10 mL). The organic layers were then dried over Na₂SO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography to afford 3-phenyl-3-(4-(trifluoromethyl) phenoxy) propan-1-amine.

tert-butyl methyl(3-phenyl-3-(2-(trifluoromethyl)phenoxy)propyl)carbamate (231)**(C₂₂H₂₆F₃NO₃; M.W.: 409.45)**

General procedure 11

Colorless oil

T.L.C. System: n-Hexane/EtOAc 8:2 v/v, R_f: 0.5

Purification: Purification: Flash column chromatography Isolera One system (Biotage),
 Cartridge: SNAP KP-Sil 50g (n-hexane-DCM 100:0 v/v increasing to 0:100 v/v in 5 CV and
 DCM/MeOH 100:0 v/v increasing to 97:3 v/v in 10 CV).

Yield: 303 mg (79%)

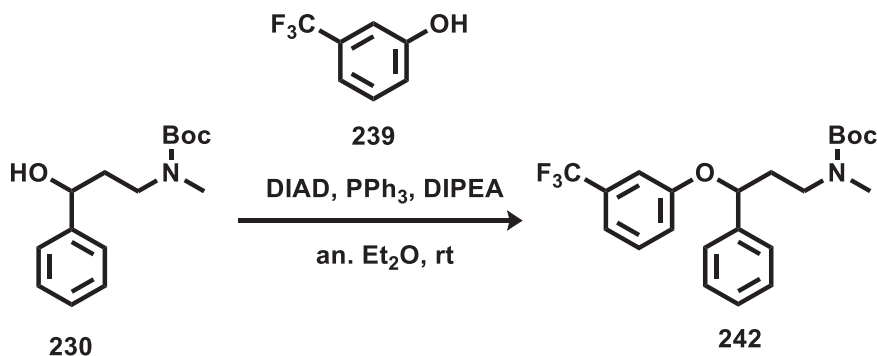
Purity: >95%

UPLC-MS method C: Rt: 2.332, MS (ESI)⁺: 192.1[C₁₁H₁₄NO₂.]⁺

¹H NMR (DMSO-d₆), δ: 7.59 (dd, ³J = 7.7 Hz, ⁴J = 1.2 Hz, 1H), 7.44 (t, J = 7.3 Hz, 1H),
 7.42 – 7.34 (m, 4H), 7.29 – 7.25 (m, 1H), 7.04 – 6.96 (m, 2H), 5.52 (br s, 1H), 3.47 – 3.22
 (m, 2H), 2.76 (s, 3H), 2.14 – 1.92 (m, 2H), 1.24 (s, 9H).

¹³C NMR (DMSO-d₆), δ: 155.50 (C=O), 155.18, 141.01 (C, C-aromatic), 134.29, 129.15,
 128.26 (CH, C-aromatic), 127.26 (q, J = 5.2 Hz, CH, C-aromatic), 126.23 (CH, C-aromatic),
 124.39 (q, J = 272.2 Hz, C, C-aromatic), 120.54 (CH, C-aromatic), 117.84 (q, J = 29.9 Hz,
 C, C-aromatic), 114.85 (CH, C-aromatic), 78.82 (C), 76.70 (CH), 45.30, 36.75 (CH₂), 34.10,
 28.35 (CH₃).

¹⁹F NMR (DMSO-d₆), δ: -60.77 (s, 3F).

tert-butyl methyl(3-phenyl-3-(3-(trifluoromethyl)phenoxy)propyl)carbamate (242)**(C₂₂H₂₆F₃NO₃; M.W.: 409.45)**

General procedure 11

pale yellow oil;

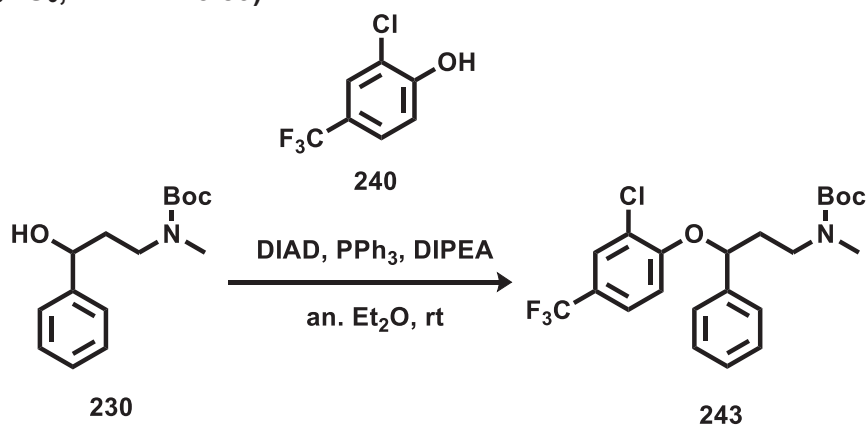
T.L.C. System: n-Hexane/EtOAc 8:2 v/v, R_f: 0.47

Purification: Purification: Flash column chromatography Isolera One system (Biotage),
 Cartridge: SNAP KP-Sil 50g (n-hexane-DCM 100:0 v/v increasing to 0:100 v/v in 5 CV and
 DCM/MeOH 100:0 v/v increasing to 97:3 v/v in 10 CV).

Yield: 280 mg (61%)

Purity: >90%

¹H NMR (DMSO-d₆), δ: 7.46 – 7.39 (m, 3H), 7.36 (t, *J* = 7.6 Hz, 2H), 7.30 – 7.25 (m, 1H),
 7.20 (d, *J* = 7.8 Hz, 1H), 7.17 (m, 2H), 5.41 (dd, *J* = 8.4, 4.2 Hz, 1H), 3.51 – 3.19 (m, 2H),
 2.78 (s, 3H), 2.17 – 1.90 (m, 2H), 1.29 (br s, 9H).

tert-butyl (3-(2-chloro-4-(trifluoromethyl)phenoxy)-3-phenylpropyl)(methyl) carbamate (243)**(C₂₂H₂₅ClF₃NO₃; M.W.: 443.89)**

General procedure 11

Colorless oil

T.L.C. System: n-Hexane/EtOAc 8:2 v/v, R_f: 0.5

Purification: Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: SNAP KP-Sil 50g (n-hexane-DCM 100:0 v/v increasing to 0:100 v/v in 5 CV DCM/MeOH 100:0 v/v increasing to 97:3 v/v in 10 CV).

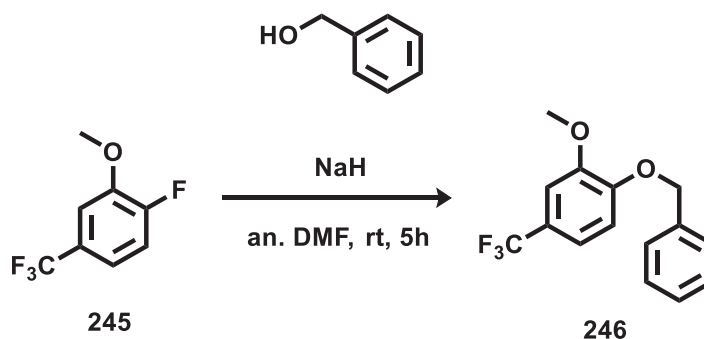
Yield: 215 mg (48%)

Purity: >90%

¹H NMR (DMSO-d₆), δ: 7.82 (s, 1H), 7.53 (d, *J* = 8.3 Hz, 1H), 7.43 – 7.34 (m, 4H), 7.31 – 7.26 (m, 1H), 7.18 – 7.07 (m, 1H), 5.57 (br s, 1H), 3.53 – 3.22 (m, 2H), 2.78 (s, 3H), 2.24 – 1.94 (m, 2H), 1.27 (br s, 9H).

1-(benzyloxy)-2-methoxy-4-(trifluoromethyl)benzene (246)

(C₁₅H₁₃F₃O₂; M.W.: 282.26)



Sodium hydride (2.36 mmol) was added to a solution of benzyl alcohol (2.15 mmol) in DMF (10 mL). After thirty minutes, 1-fluoro-2-methoxy-4-(trifluoromethyl) benzene **245** (2.58 mmol) was added dropwise to the mixture. The reaction was stirred at room temperature for five hours. The reaction mixture was diluted in ethyl acetate (20 ml) and washed with water (3x20 ml), brine (3x20 ml). The organic layer was dried over sodium sulphate, filtrated and evaporate at reduced pressure. The residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP-Sil 25g, n-Hexane-DCM 100:0 v/v increasing to 50:50 v/v in 15 CV) to afford 1-(benzyloxy)-2-methoxy-4-(trifluoromethyl) benzene **246** as white powder.

T.L.C. System: n-Hex/EtOAc 8:2v/v, R_f: 0,73

Yield: 516 mg (85%)

Purity: >95%

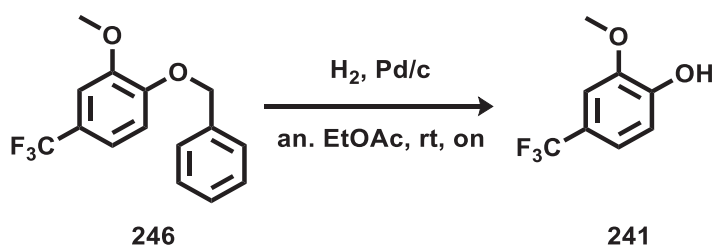
¹H NMR (DMSO-d₆), δ: 7.48 – 7.44 (m, 2H), 7.44 – 7.38 (m, 2H), 7.35 (td, ³*J* = 8.6, 4.5 Hz, ⁴*J* = 1.4 Hz, 1H), 7.29 – 7.19 (m, 3H), 5.18 (s, 1H), 3.84 (s, 2H).

¹³C NMR (DMSO-d₆), δ: 151.21, 149.72, 136.93 (C, C-aromatic), 128.95, 128.51, 128.34 (CH, C-aromatic), 124.89 (q, *J* = 271.4 Hz, CF₃), 121.82 (q, *J* = 32.1 Hz, C, C-aromatic), 118.54 (q, *J* = 3.5 Hz, CH, C-aromatic), 113.48 (CH, C-aromatic), 108.89 (q, *J* = 3.5 Hz, CH, C-aromatic), 70.40 (CH₂), 56.34 (CH₃).

^{19}F NMR (DMSO- d_6), δ : -59.45 (s, 3F).

2-methoxy-4-(trifluoromethyl)phenol (241)

($\text{C}_8\text{H}_7\text{F}_3\text{O}_2$; M.W.: 192.14)



A solution of 1-(benzyloxy)-2-methoxy-4-(trifluoromethyl) benzene **194** (1.59 mmol) in anhydrous EtOAc (0.15 M) was stirred under H_2 atmosphere in presence of Pd/C (15%) at room temperature overnight. The reaction mixture was filtered on Celite. The filtered solution was concentrated under reduced pressure to give 2-methoxy-4-(trifluoromethyl)phenol (**191**) as a colorless oil which was used without further purification.

T.L.C. System: n-Hex/EtOAc 8:2 v/v, Rf: 0.65

Yield: 257 mg (84%)

Purity: >95%

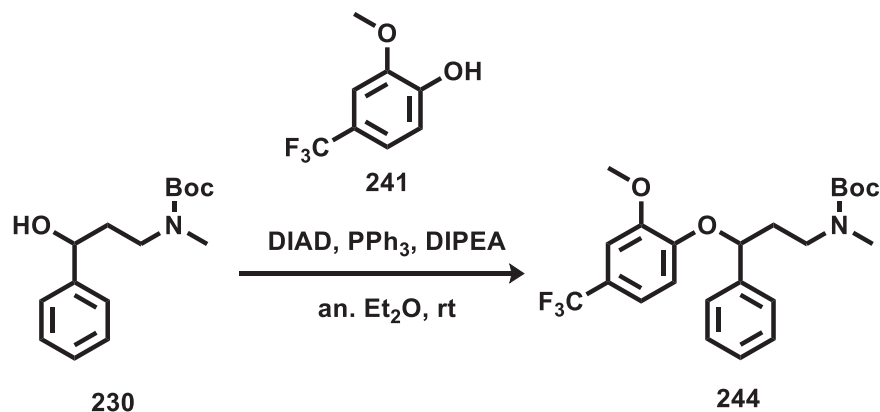
^1H NMR (DMSO- d_6), δ : 9.88 (s, 1H), 7.17 (s, 1H), 7.14 (d, $J = 8.2$ Hz, 1H), 6.93 (d, $J = 8.2$ Hz, 1H), 3.83 (s, 3H).

^{13}C NMR (DMSO- d_6), δ : 150.58, 148.30 (C, C-aromatic), 125.10 (q, $J = 270.9$ Hz, CF_3), 120.11 (q, $J = 32.0$ Hz, C, C-aromatic), 118.83 (q, $J = 4.2$ Hz, CH, C-aromatic), 115.88 (CH, C-aromatic), 109.27 (q, $J = 3.5$ Hz, CH, C-aromatic), 56.30 (CH_3).

^{19}F NMR (DMSO- d_6), δ : -59.45 (s, 3F).

tert-butyl (3-(2-methoxy-4-(trifluoromethyl)phenoxy)-3-phenylpropyl)(methyl) carbamate (244)

(C₂₃H₂₈F₃NO₄; M.W.: 439.48)



General procedure 11

pale yellow oil

T.L.C. System: n-Hexane/EtOAc 8:2 v/v, R_f: 0.49

Purification: Purification: Flash column chromatography Isolera One system (Biotage), Cartridge: SNAP KP-Sil 50g (n-hexane-DCM 100:0 v/v increasing to 0:100 v/v in 5 CV DCM/MeOH 100:0 v/v increasing to 97:3 v/v in 10 CV).

Yield: 325 mg (64%)

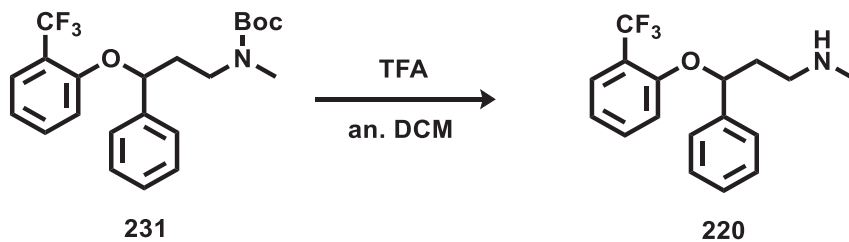
Purity: >95%

UPLC-MS method C: Rt: 2.272, MS (ESI)⁺: 462.3[M+Na]⁺, 192.1[C₁₁H₁₄NO₂]⁺

¹H NMR (DMSO-d₆), δ: 7.41 – 7.31 (m, 4H), 7.26 (t, *J* = 7.2 Hz, 1H), 7.20 (d, ⁴*J* = 2.0 Hz, 1H), 7.10 (d, *J* = 8.4 Hz, 1H), 6.93 (d, *J* = 8.5 Hz, 1H), 5.44 – 5.32 (m, 1H), 3.88 (s, 3H), 3.46 – 3.18 (m, 12H), 2.77 (s, 3H), 2.22 – 2.07 (m, 1H), 2.07 – 1.91 (m, 1H), 1.41 – 1.16 (m, 9H).

¹³C NMR (DMSO-d₆), δ: 155.19 (C=O), 150.33, 150.10, 141.19 (C, C-aromatic), 129.08, 128.22, 126.36 (CH, C-aromatic), 124.79 (q, *J* = 271.4 Hz, CF₃), 121.75 (q, *J* = 31.9 Hz, C, C-aromatic), 118.28 (q, *J* = 4.2 Hz, CH, C-aromatic), 114.76 (CH, C-aromatic), 109.11 (CH, C-aromatic), 78.80 (C), 77.39 (CH), 56.47 (CH₃), 45.37, 36.58 (CH₂), 34.05, 28.31 (CH₃).

¹⁹F NMR (DMSO-d₆), δ: -59.90 (s, 3F).

N*-methyl-3-phenyl-3-(2-(trifluoromethyl)phenoxy)propan-1-amine (220)*(C₁₇H₁₈F₃NO; M.W.: 309.33)**

General procedure 12

yellow oil

T.L.C. System: DCM/MeOH 95: v/v, R_f: 0.21

Purification: Purification: Flash column chromatography Isolera One system (Biotage),

Cartridge: ZIP KP-Sil 5g (DCM/MeOH 100:0 v/v increasing to 90:10 v/v in 24 CV).

Yield: 59 mg (83%)

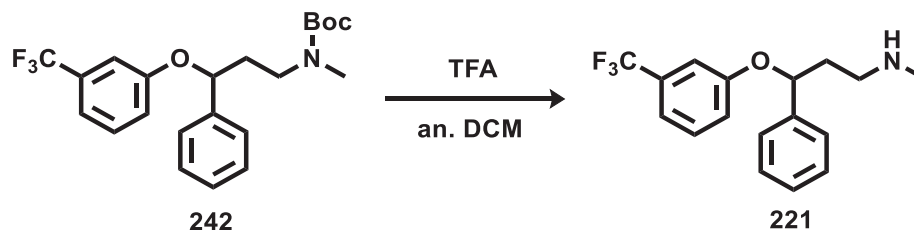
Purity: >95%

UPLC-MS method C: Rt: 1.733, MS (ESI)⁺: 310[M+1]⁺

¹H NMR (DMSO-d₆), δ: 8.88 (s, 2H), 7.61 (d, *J* = 6.8 Hz, 1H), 7.48 (t, *J* = 7.4 Hz, 1H), 7.44 – 7.37 (m, 4H), 7.34 – 7.28 (m, 1H), 7.06 – 7.00 (m, 2H), 5.81 – 5.73 (m, 1H), 3.05 – 2.92 (m, 2H), 2.55 (s, 3H), 2.34 – 2.15 (m, 2H).

¹³C NMR (DMSO-d₆), δ: ¹³C NMR (126 MHz, DMSO) δ 155.15 (C, C-aromatic), 140.20, 134.36, 129.34, 128.66 (CH, C-aromatic), 127.32 (q, *J* = 5.0 Hz, CH, C-aromatic), 126.28 (CH, C-aromatic), 124.36 (q, *J* = 272.4 Hz, CF₃), 120.89 (CH, C-aromatic), 117.80 (q, *J* = 30.0 Hz, C, C-aromatic), 115.14 (CH, C-aromatic), 76.57 (CH), 45.45, 34.50 (CH₂), 32.89 (CH₃).

¹⁹F NMR (DMSO-d₆), δ: -60.77 (s, 3F).

N*-methyl-3-phenyl-3-(3-(trifluoromethyl)phenoxy)propan-1-amine (221)*(C₁₇H₁₈F₃NO; M.W.= 309.33)**

General procedure 12

yellow oil

T.L.C. System: DCM/MeOH 95: v/v, Rf: 0.33

Purification: Purification: Flash column chromatography Isolera One system (Biotage),

Cartridge: ZIP KP-Sil 5g (DCM/MeOH 100:0 v/v increasing to 90:10 v/v in 24 CV).

Yield: 46 mg (62%)

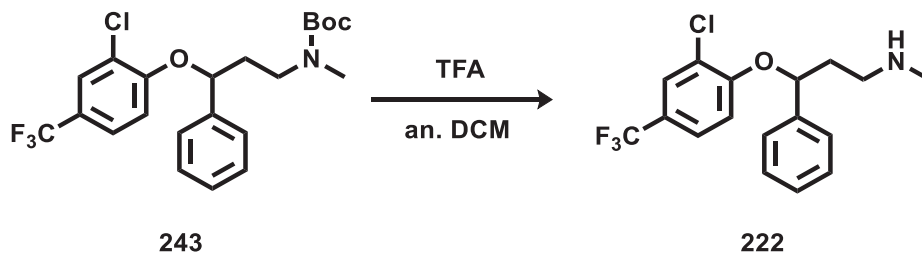
Purity: >95%

UPLC-MS method C: Rt: 1.77, MS (ESI)⁺: 310[M+1]⁺

¹H NMR (DMSO-d₆), δ: 7.46 – 7.38 (m, 3H), 7.35 (t, *J* = 7.6 Hz, 2H), 7.27 (t, *J* = 7.3 Hz, 1H), 7.22 – 7.16 (m, 3H), 5.55 (dd, *J* = 7.7, 5.4 Hz, 1H), 2.62 – 2.54 (m, 2H), 2.28 (s, 3H), 2.15 – 2.05 (m, 1H), 1.95 – 1.85 (m, 1H).

¹³C NMR (DMSO-d₆), δ: 158.46, 141.57 (C, C-aromatic), 131.07, 129.07, 128.16, 126.53, 120.30 (CH, C-aromatic), 117.56 (q, *J* = 4.0 Hz, CH, C-aromatic), 112.86 (q, *J* = 4.0 Hz, CH, C-aromatic), 77.95 (CH), 47.92, 38.08 (CH₂), 36.36 (CH₃).

¹⁹F NMR (DMSO-d₆), δ: -61.22 (s, 3F).

3-(2-chloro-4-(trifluoromethyl)phenoxy)-*N*-methyl-3-phenylpropan-1-amine (222)**(C₁₇H₁₇ClF₃NO; M.W.: 343.77)**

General procedure 12

Colourless oil

T.L.C. System: DCM/MeOH 95: v/v, R_f: 0.3Purification: Purification: Flash column chromatography Isolera One system (Biotage),
Cartridge: ZIP KP-Sil 5g (DCM/MeOH 100:0 v/v increasing to 90:10 v/v in 24 CV).

Yield: 53 mg (63%)

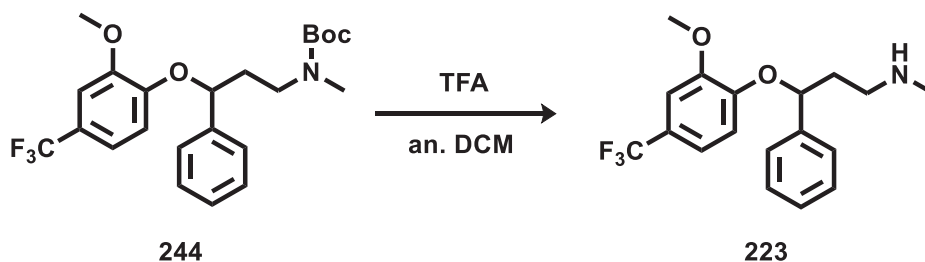
Purity: >95%

UPLC-MS method C: Rt: 1.631, MS (ESI)⁺: 344.2[M+1]⁺

¹H NMR (DMSO-d₆), δ: 7.81 (d, ⁴J = 2.0 Hz, 1H), 7.54 (dd, ³J = 8.7, ⁴J = 2.0 Hz, 1H), 7.42 – 7.34 (m, 4H), 7.31 – 7.26 (m, 1H), 7.16 (d, J = 8.7 Hz, 1H), 5.71 (dd, J = 7.9, 5.1 Hz, 1H), 2.69 – 2.58 (m, 2H), 2.30 (s, 3H), 2.21 – 2.10 (m, 1H), 2.00 – 1.96 (m, 1H).

¹³C NMR (DMSO-d₆), δ: 156.35, 140.78 (C, C-aromatic), 129.19, 128.41 (CH, C-aromatic), 127.49 (q, J = 3.7 Hz, CH, C-aromatic), 126.31 (CH, C-aromatic), 125.98 (q, J = 3.9 Hz, CH, C-aromatic), 124.12 (q, J = 256.5 Hz, C, C-aromatic), 122.36 (q, J = 32.9 Hz, C, C-aromatic), 115.99 (CH, C-aromatic), 78.95 (CH), 47.65, 37.77 (CH₂), 36.17 (CH₃).

¹⁹F NMR (DMSO-d₆), δ: -60.15 (s, 3F).

3-(2-methoxy-4-(trifluoromethyl)phenoxy)-*N*-methyl-3-phenylpropan-1-amine (223)**(C₁₈H₂₀F₃NO₂; M.W.: 339.36)**

General procedure 12

dark yellow oil

T.L.C. System: DCM/MeOH 95: v/v, R_f: 0.2

Purification: Purification: Flash column chromatography Isolera One system (Biotage),

Cartridge: ZIP KP-Sil 5g (DCM/MeOH 100:0 v/v increasing to 90:10 v/v in 24 CV).

Yield: 78 mg (64%)

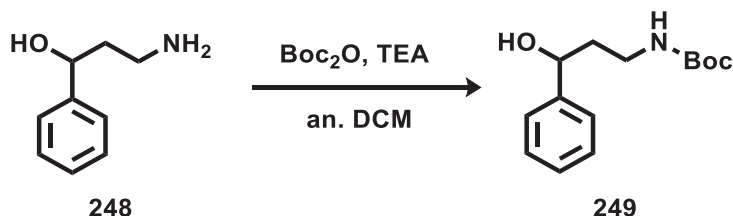
Purity: >95%

UPLC-MS method C: R_t: 1.585, MS (ESI)⁺: 340.2[M+1]⁺

¹H NMR (DMSO-d₆), δ: 7.41 – 7.33 (m, 4H), 7.29 – 7.24 (m, 1H), 7.20 (d, ⁴J = 2.0 Hz, 1H), 7.11 (dd, ³J = 8.4 Hz, ⁴J = 1.2 Hz, 1H), 6.94 (d, J = 8.4 Hz, 1H), 5.53 (dd, J = 7.8, 5.1 Hz, 1H), 3.88 (s, 3H), 2.63 (s, 2H), 2.31 (s, 3H), 2.19 – 2.08 (m, 1H), 1.99 – 1.90 (m, 1H).

¹³C NMR (DMSO-d₆), δ: 150.42, 150.04, 141.47 (C, C-aromatic), 129.06, 128.18, 126.38 (CH, C-aromatic), 124.80 (q, J = 271.3 Hz, CF₃), 121.71 (q, J = 32.2 Hz, C, C-aromatic), 118.31 (q, J = 4.2 Hz, CH, C-aromatic), 114.94 (CH, C-aromatic), 109.09 (q, J = 3.4 Hz, CH, C-aromatic), 78.41 (CH), 56.48 (CH₃), 47.76, 37.69 (CH₂), 35.97 (CH₃).

¹⁹F NMR (DMSO-d₆), δ: -59.88 (s, 3F).

Synthesis of *N*-(3-phenyl-3-(4-(trifluoromethyl)phenoxy)propyl)acetamide **247****tert-butyl (3-hydroxy-3-phenylpropyl)carbamate (**249**)****(C₁₄H₂₁NO₃; M.W.: 251.33)**

To a solution of 3-amino-1-phenylpropan-1-ol **248** (1.65 mmol) in DCM (7 mL), di-tert-butyl dicarbonate (1.8 mmol) was added, the reaction mixture was stirred at room temperature overnight. The reaction mixture was then washed with NH₄Cl (3x30 ml), brine (3x30 ml), dried over Na₂SO₄ and concentrated under reduce pressure to give tert-butyl (3-hydroxy-3-phenylpropyl) carbamate **249** as a yellow solid which, was used without further purification.

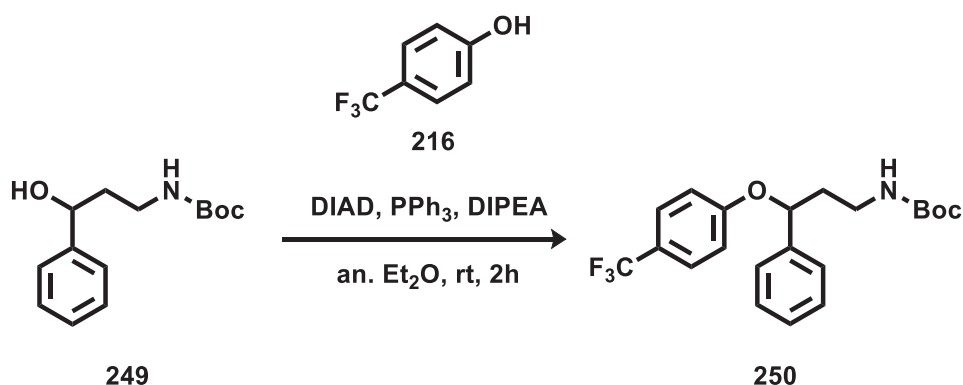
T.L.C. System: n-Hex/EtOAc 8:2x v/v, R_f: 0.18

Yield: 1.08 g (87%)

Purity: >95%

¹H NMR (DMSO-d₆), δ: 7.40 – 7.34 (m, 4H), 7.31 – 7.24 (m, 1H), 6.82 (t, *J* = 5.0 Hz, 1H), 5.25 (d, *J* = 4.5 Hz, 1H), 4.60 (dd, *J* = 11.0, 6.4 Hz, 1H), 3.03 (dd, *J* = 13.6, 6.3 Hz, 2H), 1.75 (dd, *J* = 14.2, 6.8 Hz, 2H), 1.43 (s, 9H).

¹³C NMR (DMSO-d₆), δ: 156.05 (C=O), 146.55 (C, C-aromatic), 128.45, 127.11, 126.14 (CH, C-aromatic), 77.87 (CH), 70.68 (C), 39.79, 37.70 (CH₂), 28.74 (CH₃).

tert-butyl (3-phenyl-3-(4-(trifluoromethyl)phenoxy)propyl)carbamate (250)**(C₂₁H₂₄F₃NO₃; M.W.: 395.42)**

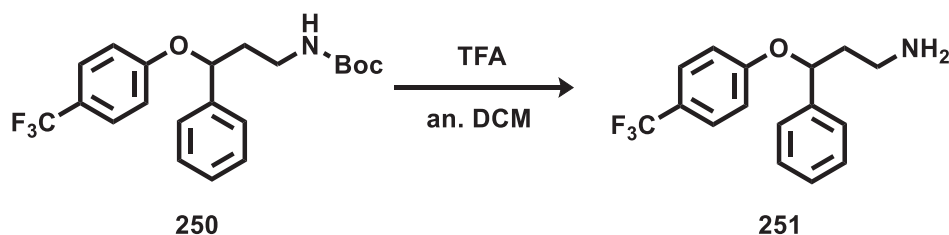
To a solution of triphenylphosphine (2.99 mmol) in diethyl ether (5.97 mL) cooled at 0°C, diisopropyl azodicarboxylate (2.01 mmol) in diethyl ether (1 mL) was added. After twenty minutes 4-(trifluoromethyl)phenol **216** (2.39 mmol) in diethyl ether (4 mL) was added before adding the solution of tert-butyl (3-hydroxy-3-phenylpropyl)(methyl)carbamate **249** (1.99 mmol) in diethyl ether (5 mL). The reaction was stirred at room temperature for two hours. The reaction mixture was concentrated under reduced pressure and the residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP-Sil 50g, n-hexane-DCM 100:0 v/v increasing to 0:100 v/v in 5 CV and DCM/MeOH 100:0 v/v increasing to 98:2 v/v in 10 CV) to afford tert-butyl (3-(3-(trifluoromethyl)phenoxy)-3-phenylpropyl)(methyl)carbamate **250** as a colourless oil.

T.L.C. System: n-Hex/EtOAc 8:2 v/v, R_f: 0.4

Yield: 193 mg

Two species observed. Major/minor species ratio: 4/1.

¹H NMR (DMSO-d₆), δ (major species): 7.56 (d, *J* = 8.8 Hz, 2H), 7.42 – 7.33 (m, 4H), 7.27 (t, *J* = 10.0, 4.3 Hz, 1H), 7.05 (d, *J* = 8.7 Hz, 2H), 6.93 (s, 1H), 5.47 (dd, *J* = 8.0, 4.9 Hz, 1H), 3.11 – 3.02 (m, 2H), 2.12 – 2.01 (m, *J* = 13.8, 7.6 Hz, 1H), 1.95 – 1.85 (m, 1H), 1.31 (s, 9H). **(minor species):** 10.28 (s, 1H), 7.53 (d, *J* = 8.6 Hz, 2H), 6.92 (d, *J* = 8.6 Hz, 2H).

3-phenyl-3-(4-(trifluoromethyl)phenoxy)propan-1-amine (251)**(C₁₆H₁₆F₃NO; M.W.: 295.31)**

To a solution of tert-butyl (3-phenyl-3-(4-(trifluoromethyl)phenoxy)propyl)carbamate **250** (0.84 mmol) in DCM (4.2 mL) at 0°C TFA (4.20 mmol) was added. The reaction was stirred at room temperature for two hours. The reaction mixture was diluted with DCM (8 mL) and was washed with sat. NaHCO₃ solution (3 x 10 mL) and brine (3 x 10 mL). The organic layers were then dried over Na₂SO₄ and concentrated under reduced pressure. The residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP-Sil 25g, n-Hexane-DCM 100:0 v/v increasing to 50:50 v/v in 15 CV) to afford 3-phenyl-3-(4-(trifluoromethyl)phenoxy)propan-1-amine **251** as a yellow oil.

T.L.C. System: DCM/MeOH 95:5 v/v, R_f: 0,15

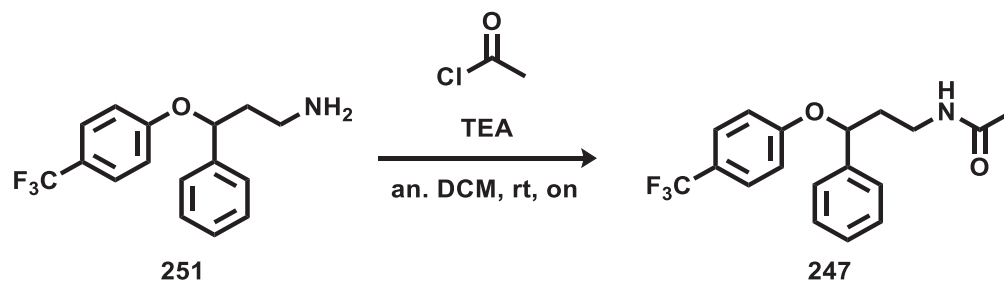
Yield: 160 mg (27% over two steps)

Purity: >95%

¹H NMR (DMSO-d₆), δ: 7.55 (d, *J* = 8.7 Hz, 2H), 7.42 – 7.32 (m, 4H), 7.26 (t, *J* = 7.2 Hz, 1H), 7.06 (d, *J* = 8.6 Hz, 2H), 5.61 – 5.55 (m, 1H), 2.69 (t, *J* = 6.8 Hz, 2H), 2.13 – 2.03 (m, 1H), 1.93 – 1.83 (m, 1H).

¹³C NMR (DMSO-d₆), δ: 160.99, 141.43 (C, C-aromatic), 129.10, 128.18 (CH, C-aromatic), 127.27 (q, *J* = 3.7 Hz, CH, C-aromatic), 126.46 (CH, C-aromatic), 124.92 (q, *J* = 271.1 Hz, CF₃), 121.54 (q, *J* = 32.1 Hz, C, C-aromatic), 116.61 (CH, C-aromatic), 77.48 (CH), 41.01, 37.95 (CH₂).

¹⁹F NMR (DMSO-d₆), δ: -59.88 (s, 3F).

N*-(3-phenyl-3-(4-(trifluoromethyl)phenoxy)propyl)acetamide (247)*(C₁₈H₁₈F₃NO₂; M.W.: 337.34)**

Acetyl chloride was added drop-wise to a solution, at 0°C, of 3-phenyl-3-(4-(trifluoromethyl)phenoxy)propan-1-amine **251** (0.2 mmol), triethylamine (0.38 mmol) in anhydrous DCM (0.4 M) under nitrogen atmosphere. The reaction was allowed to warm-up to room temperature and was stirred for 2 hours. The reaction mixture was diluted with DCM (8 ml) and extracted with 2N HCl solution (3x10ml). The organic layer was washed with brine (3 x 10 mL), dried over Na₂SO₄ and evaporated at reduced pressure. The residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP-Sil 5g, DCM-DCM/MeOH 100:0 v/v to 98:2 v/v) to afford pure *N*-(3-phenyl-3-(4-(trifluoromethyl)phenoxy) propyl) acetamide **247** as a yellow oil.

T.L.C. System: DCM/MeOH 95:5 v/v, R_f: 0.5

Yield: 63 mg (98%)

Purity: >95%

UPLC-MS method C: Rt: 1.923, MS (ESI)⁺: 338.2[M+1]⁺, 176.1[C₁₁H₁₄NO]⁺

¹H NMR (DMSO-d₆), δ: 7.93 (t, *J* = 5.0 Hz, 1H), 7.56 (d, *J* = 8.8 Hz, 2H), 7.41 – 7.32 (m, 5H), 7.27 (t, *J* = 7.2 Hz, 1H), 7.05 (d, *J* = 8.7 Hz, 2H), 5.47 (dd, *J* = 7.9, 5.0 Hz, 1H), 3.21 – 3.11 (m, 2H), 2.15 – 2.03 (m, 1H), 1.97 – 1.87 (m, 1H), 1.80 (s, 3H).

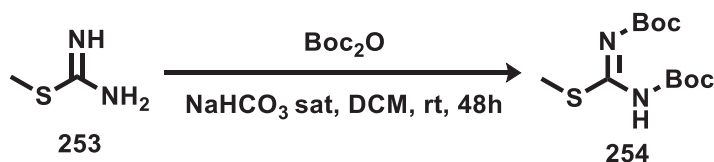
¹³C NMR (DMSO-d₆), δ: 169.64 (C=O), 160.91, 141.15 (C, C-aromatic), 129.13, 128.25 (CH, C-aromatic), 127.29 (q, *J* = 3.7 Hz, CH, C-aromatic), 126.45 (CH, C-aromatic), 124.92 (q, *J* = 271.1 Hz, CF₃), 121.60 (q, *J* = 32.1 Hz, C, C-aromatic), 116.64 (CH, C-aromatic), 77.53 (CH), 38.20, 35.74 (CH₂), 23.09 (CH₃).

¹⁹C NMR (DMSO-d₆), δ: -59.89 (s, 3F).

12.5 synthesis of Guanidine derivatives of fluoxetine

N,N'-Bis(tert-butoxycarbonyl)-*S*-methylisothiurea (**254**)

(C₁₂H₂₂N₂O₄S; M.W.: 290.38)



To a stirring mixture of methyl carbamimidothioate hemisulfate **253** (7.18 mmol) in sat. NaHCO₃ (8 mL) and DCM (16 mL) a solution of di-tert-butyl decarbonate (14.37 mmol) in DCM (12 mL) was added. The reaction was stirred at room temperature for 48 hours. The mixture was diluted in DCM (8 ml) and the organic layer was separated from the aqueous. The aqueous phase was extracted with DCM (2x20 ml). The combined organic layers were dried over Na₂SO₄, filtrated and concentrated under reduced pressure. The crude was stirred in EtOH/H₂O 9:1 for 1h, the resulting precipitate was filtrated under vacuum to give 1,14 g of **254** as a white powder.

T.L.C. System: n-Hex/EtOAc 8:2 v/v, R_f: 0.66

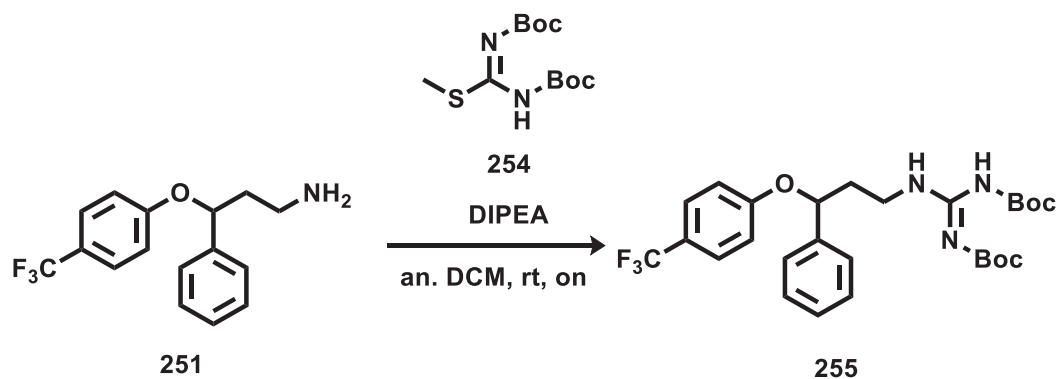
Yield: 1.4 g (55%)

Purity: >95%

UPLC-MS method C: Rt: 2.168, MS (ESI)⁺: 291.1[M+1]⁺, 313.1[M+Na]⁺,

¹H NMR (CDCl₃-d), δ: 11.60 (s, 1H), 2.40 (s, 3H), 1.53 (s, 9H), 1.51 (s, 9H).

¹³C NMR (CDCl₃-d), δ: 171.46 (C=N), 160.77, 150.78 (C=O), 83.24, 80.97 (C), 28.04, 14.41 (CH₃).

N,N'*-Bis(tert-butoxycarbonyl)-(1-(3-phenyl-3-(4-(trifluoromethyl)phenoxy)propyl))guanidine (255)*(C₂₇H₃₄F₃N₃O₅; M.W.: 537.58)**

A solution of *N, N'*-bis(tert-butoxycarbonyl)-*S*-methylisothiurea **254** (0.28 mmol) in DCM (0.5 M) was added dropwise to a solution of 3-phenyl-3-(4-(trifluoromethyl) phenoxy) propan-1-amine **251** (0.31 mmol) and DIPEA (0.34 mmol) in DCM (0.5 M) at 0°C. The reaction mixture was stirred at room temperature overnight. A stream of nitrogen gas was bubbled through the reaction mixture for 1 hour to purge the gaseous by-product CH₃SH. The residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP-Sil 10g, n-Hexane-EtOAc 100:0 v/v increasing to 80:20 v/v in 15 CV) to afford *N, N'*-bis(tert-butoxycarbonyl)-*N''*-3-phenyl-3-(4-(trifluoromethyl) phenoxy)propylguanidine **255** as a white solid.

T.L.C. System: n-Hex/EtOAc 8:2 v/v, R_f: 0.74

Yield: 134 mg (81%)

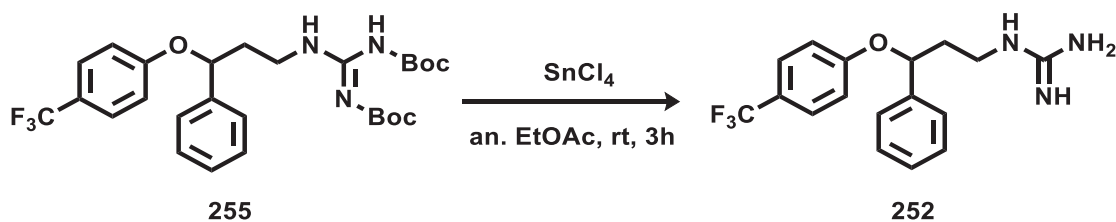
Purity: >95%

UPLC-MS method C: Rt: 2.43, MS (ESI)⁺: 538.2[M+1]⁺, 560.2[M+Na]⁺

¹H NMR (DMSO-*d*₆), δ: 11.46 (s, 1H), 8.49 (t, *J* = 5.4 Hz, 1H), 7.55 (d, *J* = 8.7 Hz, 2H), 7.41 (d, *J* = 7.2 Hz, 2H), 7.35 (t, *J* = 7.6 Hz, 2H), 7.27 (t, *J* = 7.2 Hz, 1H), 7.08 (d, *J* = 8.7 Hz, 2H), 5.58 (dd, *J* = 8.0, 4.3 Hz, 1H), 3.52 – 3.43 (m, 2H), 2.25 – 2.15 (m, 1H), 2.14 – 2.05 (m, 1H), 1.48 (s, 9H), 1.37 (s, 9H).

¹³C NMR (DMSO-*d*₆), δ: 163.50 (C=N), 160.73 (C=O), 155.69 (C=O), 152.44, 140.96 (C, C-aromatic), 129.11, 128.25 (CH, C-aromatic), 127.16 (q, *J* = 3.6 Hz, CH, C-aromatic), 126.49 (CH, C-aromatic), 121.60 (q, *J* = 31.9 Hz, C, C-aromatic), 116.64 (CH, C-aromatic), 83.32, 78.52 (CH₂), 78.33 (CH), 38.11, 37.34 (C), 28.42, 28.09 (CH₃).

¹⁹F NMR (DMSO-*d*₆), δ: -59.91 (s, 3F).

1-(3-phenyl-3-(4-(trifluoromethyl)phenoxy)propyl)guanidine (252)**(C₁₇H₁₈F₃N₃O; M.W.: 337.35)**

Tin(IV) chloride (0.86 mmol) was added to a solution of N',N''-Bis(tert-butoxycarbonyl)-1-(3-phenyl-3-(4-(trifluoromethyl) phenoxy)propyl)guanidine **255** (0.22 mmol) in anhydrous EtOAc (3 mL) at 0°C under a N₂ atmosphere . The reaction mixture was stirred at room temperature for three hours. The reaction mixture was diluted in EtOAc and washed with sat. NaHCO₃ solution, dried over Na₂SO₄, filtrate and concentrated under reduced pressure. The resulting solid is dissolved in EtOAc and the resulting formed precipitate was filtrated under vacuum to give 1-(3-phenyl-3-(4-(trifluoromethyl)phenoxy)propyl)guanidine **252** as a white powder.

Yield: 20 mg (28%)

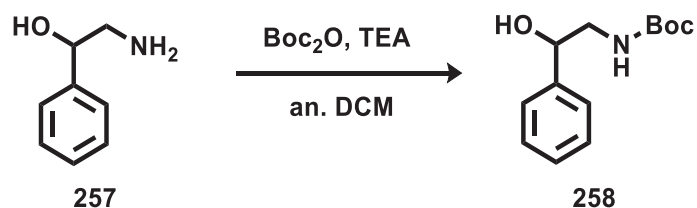
Purity: >95%

UPLC-MS method C: Rt: 1.673, MS (ESI)⁺: 338.2[M+1]⁺

¹H NMR (DMSO-d₆), δ: 7.56 (d, *J* = 8.7 Hz, 2H), 7.42 (d, *J* = 7.5 Hz, 2H), 7.36 (t, *J* = 7.5 Hz, 2H), 7.28 (t, *J* = 7.3 Hz, 1H), 7.07 (d, *J* = 8.6 Hz, 2H), 5.53 (dd, *J* = 8.3, 4.4 Hz, 1H), 3.28 – 3.23 (m, 2H), 2.21 – 2.09 (m, 1H), 2.07 – 1.95 (m, 1H).

¹³C NMR (DMSO-d₆), δ: 160.85 (C=N), 160.02, 140.92 (C, C-aromatic), 129.18, 128.33 (CH, C-aromatic), 127.30 (q, *J* = 3.7 Hz, CH, C-aromatic), 126.45 (CH, C-aromatic), 124.91 (q, *J* = 271.1 Hz, CF₃), 121.70 (q, *J* = 32.0 Hz, C, C-aromatic), 116.66 (CH, C-aromatic), 77.31 (CH), 38.24, 37.55 (CH₂).

¹⁹F NMR (DMSO-d₆), δ: -59.90 (s, 3F).

tert-butyl (2-hydroxy-2-phenylethyl)carbamate (258)**(C₁₃H₁₉NO₃; M.W.: 237.30)**

To a solution of 2-amino-1-phenylpropan-1-ol **257** (3.64 mmol) in anhydrous DCM (14 mL), di-tert-butyl dicarbonate (4 mmol) was added, the reaction mixture was stirred at room temperature overnight. The reaction mixture was then washed with 1M HCl solution (3x30 ml), brine (3x30 ml), dried over Na₂SO₄, filtered and concentrated under reduce pressure. The residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP-Sil 25g, n-Hexane-EtOAc 100:0 v/v increasing to 60:40 v/v in 11 CV) to afford to give tert-butyl (2-hydroxy-2-phenylethyl)carbamate **258** as a white solid.

T.L.C. System: n-Hex/EtOAc 8:2 v/v, Rf: 0.21

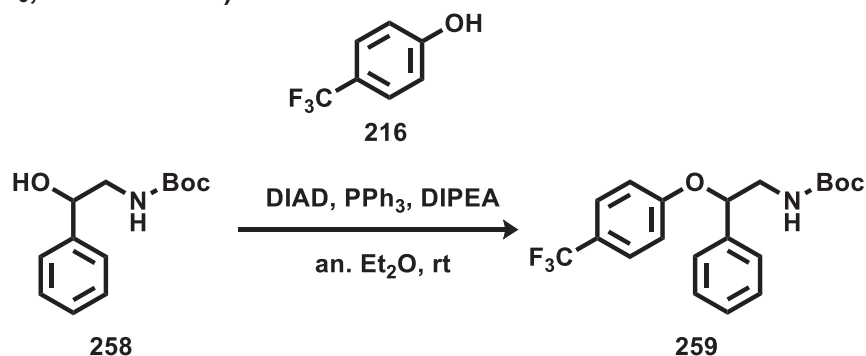
Yield: 808 mg (93%)

Purity: >95%

UPLC-MS method C: Rt: 1.668, MS (ESI)⁺: 164.0[C₉H₁₀NO₂]⁺

¹H NMR (DMSO-d₆), δ: 7.35 – 7.29 (m, 4H), 7.27 – 7.21 (m, 1H), 6.71 (t, *J* = 5.5 Hz, 1H), 5.34 (d, *J* = 4.5 Hz, 1H), 4.61 – 4.55 (m, 1H), 3.16 – 3.08 (m, 1H), 3.05 – 2.94 (m, 1H), 1.36 (s, 9H).

¹³C NMR (DMSO-d₆), δ: 156.09 (C=O), 144.16 (C, C-aromatic), 128.41, 127.39, 126.50 (CH, C-aromatic), 78.04 (C), 71.94 (CH), 48.63 (CH₂), 28.71 (CH₃).

tert-butyl (2-phenyl-2-(4-(trifluoromethyl)phenoxy)ethyl)carbamate (259)**(C₂₀H₂₂F₃NO₃; M.W.: 381.40)**

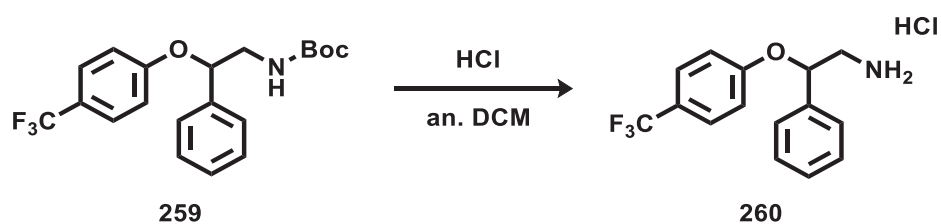
To a solution of triphenylphosphine (1.9 mmol) in diethyl ether (3.8 mL) cooled at 0°C, diisopropyl azodicarboxylate (1.28 mmol) in diethyl ether (0.6 mL) was added. After twenty minutes 4-(trifluoromethyl)phenol **216** (1.52 mmol) in diethyl ether (2.5 mL) was added before adding the solution of tert-butyl (2-hydroxy-2-phenylethyl)carbamate **258** (1.27 mmol) in diethyl ether (3.1 mL). The reaction was stirred overnight at room temperature. The reaction mixture was concentrated under reduced pressure and the residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP-Sil 10g, n-Hexane-DCM from 100:0 v/v increasing to 0:100 v/v in 5 CV and DCM-MeOH from 100:0 v/v increasing to 98:2 v/v in 8 CV) to afford tert-butyl (2-phenyl-2-(4-(trifluoromethyl)phenoxy)ethyl)carbamate **259** as a colourless oil.

T.L.C. System: n-Hex/EtOAc 8:2 v/v, R_f: 0.45

Yield: 338 mg

Purity: >95%

¹H NMR (DMSO-d₆), δ: (major species) 7.57 (d, *J* = 8.6 Hz, 2H), 7.41 – 7.27 (m, 5H), 7.16 (t, *J* = 5.4 Hz, 1H), 7.06 (d, *J* = 8.6 Hz, 2H), 5.43 (dd, *J* = 7.8, 4.4 Hz, 1H), 3.44 – 3.27 (m, 2H), 1.35 (s, 9H); **(minor species)** 10.28 (s, 1H), 7.53 (d, *J* = 8.6 Hz, 2H), 6.92 (d, *J* = 8.5 Hz, 2H).

2-phenyl-2-(4-(trifluoromethyl)phenoxy)ethan-1-amine hydrochloride (260)**(C₁₅H₁₄F₃NO; M.W.: 281.28)**

A solution of tert-butyl methyl(2-phenyl-2-(4-(trifluoromethyl) phenoxy) ethyl) carbamate **259** (0.89 mmol) in DCM (2.7 mL) was cooled to 0°C and HCl (2M) in diethyl ether (3.57 mmol) was added. The reaction was stirred overnight at room temperature. Formation of a precipitate was observed. The precipitate was filtrated under vacuum and washed with DCM to give 2-phenyl-2-(4-(trifluoromethyl) phenoxy) ethan-1-amine **259** as a white solid.

T.L.C. System: DCM/MeOH 95:5 v/v, R_f: 0.36

Yield: 132 mg (37% over 2 steps)

Purity: >95%

UPLC-MS method C: Rt: 1.7, MS (ESI)⁺: 282.0[M+1]⁺

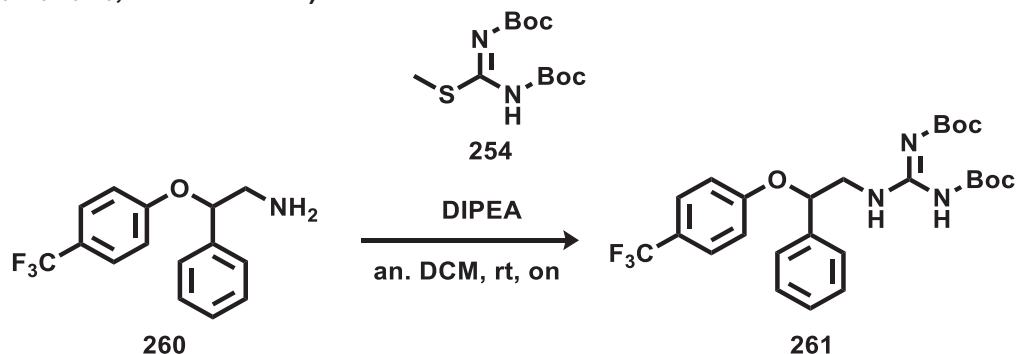
¹H NMR (DMSO-d₆), δ: 8.51 (s, 3H), 7.61 (d, *J* = 8.7 Hz, 2H), 7.48 – 7.39 (m, 4H), 7.35 (t, *J* = 7.2 Hz, 1H), 7.10 (d, *J* = 8.6 Hz, 2H), 5.76 (dd, *J* = 9.0, 3.7 Hz, 1H), 3.32 – 3.21 (m, 2H).

¹³C NMR (DMSO-d₆), δ: 160.21, 137.23 (C, C-aromatic), 129.44, 129.23 (CH, C-aromatic), 127.35 (q, *J* = 3.7 Hz, CH, C-aromatic), 126.82 (CH, C-aromatic), 124.84 (q, *J* = 271.3 Hz, CF₃), 122.34 (q, *J* = 32.1 Hz, C, C-aromatic), 117.01 (CH, C-aromatic), 76.83 (CH), 44.94 (CH₂).

¹⁹F NMR (DMSO-d₆), δ: -59.99 (s, 3F).

***N,N'*-Bis(tert-butoxycarbonyl)-(1-(2-phenyl-2-(4-(trifluoromethyl)phenoxy)ethyl)guanidine (261)**

(C₂₆H₃₂F₃N₃O₅; M.W.: 523.55)



A solution of *N,N'*-bis(tert-butoxycarbonyl)-*S*-methylisothiurea **254** (0.50 mmol) in anhydrous DCM (0.5 M) was added dropwise to a solution of 2-phenyl-2-(4-(trifluoromethyl)phenoxy)ethan-1-amine **260** (0.48 mmol) and DIPEA (0.86 mmol) in anhydrous DCM (0.5 M) at 0°C. The reaction mixture was stirred at room temperature overnight. A stream of nitrogen gas was bubbled through the reaction mixture for 1 hour to purge the gaseous by-product CH₃SH. The residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP-Sil 10g, hexane/ EtOAc 100:0 v/v increasing to 90:10 v/v in 16 CV) to afford *N,N'*-bis(tert-butoxycarbonyl)-*N''*-2-phenyl-2-(4-(trifluoromethyl)phenoxy)ethylguanidine **261** as a white solid.

T.L.C. System: n-Hex/EtOAc 8:2 v/v, R_f: 0.72

Yield: 90 mg (46%)

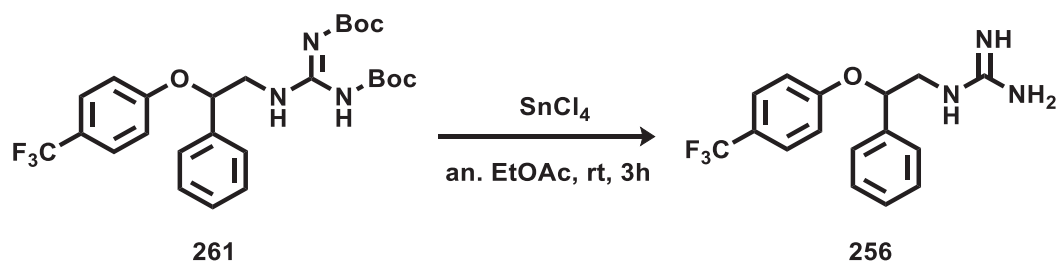
Purity: >95%

UPLC-MS method C: Rt: 2.466, MS (ESI)⁺: 524.2[M+1]⁺

¹H NMR (DMSO-*d*₆), δ: 11.44 (s, 1H), 8.55 (t, *J* = 5.4 Hz, 1H), 7.59 (d, *J* = 8.7 Hz, 2H), 7.45 (d, *J* = 7.2 Hz, 2H), 7.39 (t, *J* = 7.5 Hz, 2H), 7.32 (t, *J* = 7.3 Hz, 1H), 7.15 (d, *J* = 8.7 Hz, 2H), 5.71 (dd, *J* = 7.0, 5.0 Hz, 1H), 3.84 – 3.73 (m, 2H), 1.45 (s, 9H), 1.41 (s, 9H).

¹³C NMR (DMSO-*d*₆), δ: 163.36 (C=N), 160.66, 155.98 (C=O), 152.51, 138.33 (C, C-aromatic), 129.15, 128.82 (CH, C-aromatic), 127.40 (q, *J* = 3.7 Hz, CH, C-aromatic), 126.89 (CH, C-aromatic), 124.86 (q, *J* = 271.1 Hz, CF₃), 122.05 (q, *J* = 31.7 Hz, C, C-aromatic), 116.73 (CH, C-aromatic), 83.67, 78.83 (C), 77.85 (CH), 46.50 (CH₂), 28.42, 28.00 (CH₃).

¹⁹f NMR (DMSO-*d*₆), δ: -60.00 (s, 3F).

1-(2-phenyl-2-(4-(trifluoromethyl)phenoxy)ethyl)guanidine (256)**(C₁₆H₁₆F₃N₃O; M.W.: 323.32)**

Tin(IV) chloride (0.29 mmol) was added to a solution of N, N'-bis(tert-butoxycarbonyl)-N''-3-phenyl-3-(4-(trifluoromethyl) phenoxy)propylguanidine **261** (0.074 mmol) in anhydrous EtOAc (1 mL) at 0°C under a N₂ atmosphere. The reaction mixture was stirred at room temperature for three hours. The reaction mixture was diluted in EtOAc and washed with sat. NaHCO₃ solution, dried over Na₂SO₄, filtrated and concentrated under reduced pressure. The resulting solid was dissolved in EtOAc and the resulting formed precipitate was filtrated under vacuum to give 1-(3-phenyl-3-(4-(trifluoromethyl)phenoxy)propyl)guanidine **256** as a white powder.

T.L.C. System: DCM/MeOH 99:1 v/v, R_f: 0.22

Yield: 23 mg (95%)

Purity: >95%

UPLC-MS method C: Rt: 1.627, MS (ESI)⁺: 324.2[M+1]⁺

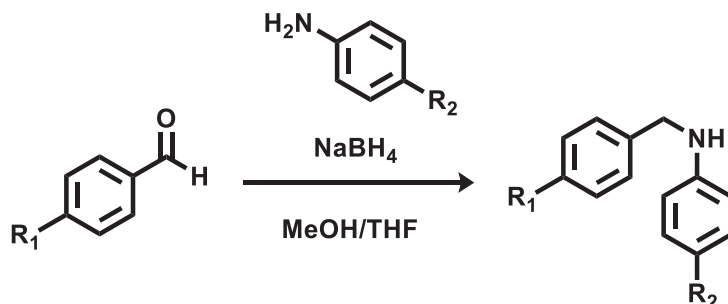
¹H NMR (DMSO-d₆), δ: 7.58 (d, *J* = 8.7 Hz, 2H), 7.47 (d, *J* = 7.3 Hz, 2H), 7.38 (t, *J* = 7.5 Hz, 2H), 7.31 (t, *J* = 7.3 Hz, 1H), 7.08 (d, *J* = 8.7 Hz, 2H), 5.60 – 5.52 (m, 1H), 3.62 – 3.47 (m, 2H).

¹³C NMR (DMSO-d₆), δ: 160.76 (C=N), 157.57 (s), 138.57 (C, C-aromatic), 129.08, 128.70 (CH, C-aromatic), 127.34 (q, *J* = 3.5 Hz, CH, C-aromatic), 127.01 (CH, C-aromatic), 124.89 (q, *J* = 269.3 Hz, CF₃), 121.88 (q, *J* = 32.8 Hz, C, C-aromatic), 116.69 (CH, C-aromatic), 78.76 (CH), 57.09 (CH₂).

¹⁹F NMR (DMSO-d₆), δ: -59.91 (s, 3F).

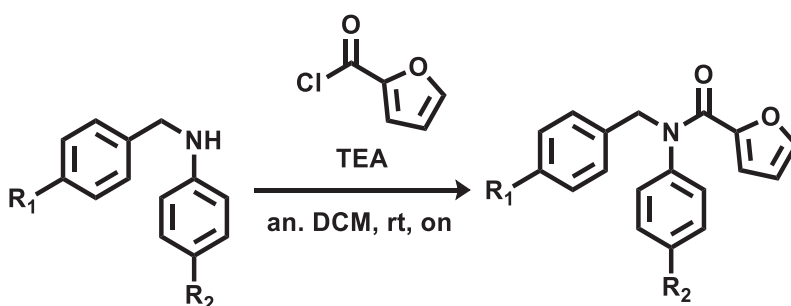
12.6 Synthesis of *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)furan-2-carboxamide and related analogues

General procedure 13: synthesis of *N*-benzylanilines

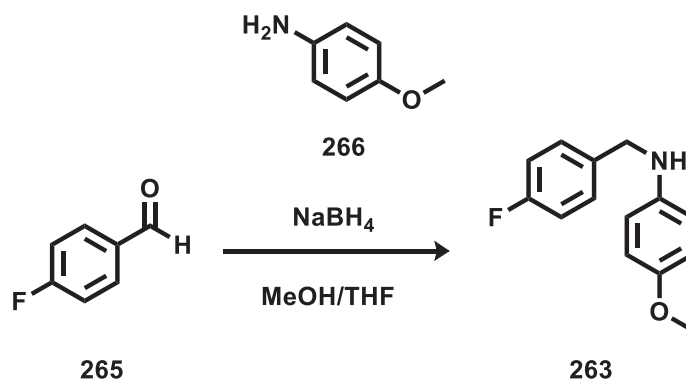


4-fluorobenzaldehyde (4.03 mmol) was dissolved in MeOH/THF (4:1, 25 ml) and 4-Methoxyaniline (4.43 mmol) was added. The reaction mixture was stirred at room temperature for 6. The reaction mixture was cooled 0°C before adding NaBH₄ (8.86 mmol). The reaction was allowed to warm-up to room temperature and stirred overnight. The reaction mixture was concentrated under reduced pressure. The resulting residue was suspended in water and extracted with DCM (3x20 ml). The combined organic layers were washed with brine (3x20 ml), dried over Na₂SO₄, filtered and concentrated under reduced pressure to give *N*-(4-fluorobenzyl)-4-methoxyaniline which, was used next without further purification.

General procedure 14: synthesis of *N*-benzyl-*N*-phenylfuran-2-carboxamides



N-(4-fluorobenzyl)-4-methoxyaniline (0.19 mmol) was dissolved in anhydrous DCM (0.2 M) and TEA (0.33 mmol) was added. The reaction was cooled to 0°C before adding furan-2-carbonyl chloride (0.17 mmol). The reaction was allowed to warm-up to room temperature and stirred for 2 hours. The reaction mixture was diluted in DCM (10mL) and washed with a sat. NaHCO₃ solution (3x8 mL), brine (2 x 20mL), dried over Na₂SO₄, filtered and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system) to give pure *N*-benzyl-*N*-phenylfuran-2-carboxamides.

N*-(4-fluorobenzyl)-4-methoxyaniline (263)*(C₁₄H₁₄FNO; M.W.= 231.27)**

general procedure 13

yellow oil

T.L.C. System: n-Hex/EtOAc 8:2 v/v, Rf: 0.61

Yield: 930 mg (99%)

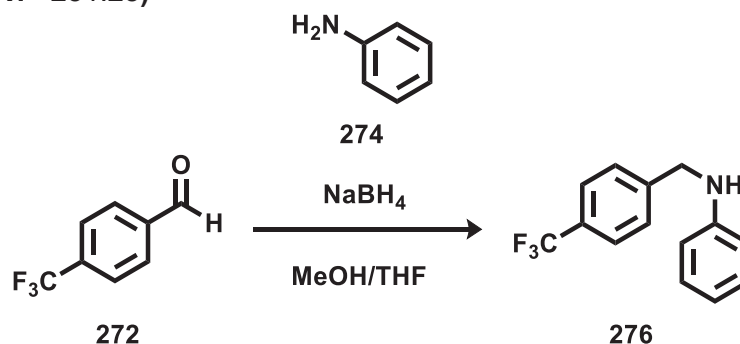
Purity: >95%

UPLC-MS method C: Rt: 1.568, MS (ESI)⁺: 232.2[M+1]⁺

¹H NMR (DMSO-d₆), δ: 7.38 (d, *J* = 8.1 Hz, 2H), 7.13 (t, *J* = 8.8 Hz, 2H), 6.68 (d, *J* = 8.8 Hz, 2H), 6.52 (d, *J* = 8.7 Hz, 2H), 5.81 (t, *J* = 6.0 Hz, 1H), 4.19 (d, *J* = 6.1 Hz, 2H), 3.61 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 162.46, 151.22, 143.17, 137.14 (C, C-aromatic), 129.50, 115.36 (CH, C-aromatic), 115.01, 113.83 (CH, C-aromatic), 55.72 (CH₃), 46.96 (CH₂).

¹⁹F NMR (DMSO-d₆), δ: -116.69 (s, F).

N*-(4-(trifluoromethyl)benzyl)aniline (276)*(C₁₄H₁₂F₃N; M.W.= 251.25)**

general procedure 13

yellow oil

T.L.C. System: n-Hex/DCM 1:1 v/v, Rf: 0.59

Yield: 300 mg (99%)

Purity: >95%

UPLC-MS method C: Rt: 2.091, MS (ESI)⁺: 252.1[M+1]⁺

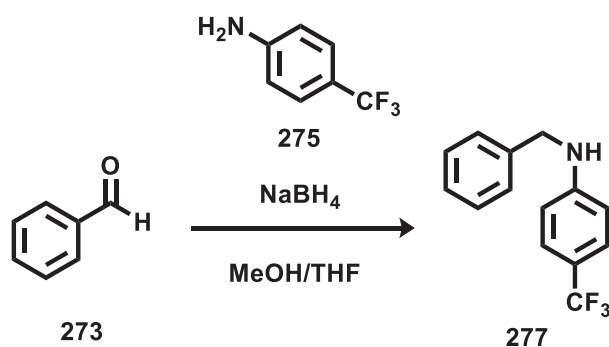
¹H NMR (DMSO-d₆), δ: 7.68 (d, *J* = 8.0 Hz, 2H), 7.59 – 7.54 (m, 2H), 7.04 (t, *J* = 7.7 Hz, 2H), 6.58 – 6.49 (m, 3H), 6.36 (t, *J* = 5.5 Hz, 1H), 4.37 (d, *J* = 5.1 Hz, 2H).

¹³C NMR (DMSO-d₆), δ: 148.77, 146.04 (C, C-aromatic), 129.36, 128.20 (CH, C-aromatic), 127.77 (q, *J* = 31.5 Hz, C, C-aromatic), 125.60 (q, *J* = 3.8 Hz, CH, C-aromatic), 124.86 (q, *J* = 271.9 Hz, CF₃), 116.46, 112.75 (CH, C-aromatic), 46.39 (CH₂).

¹⁹F NMR (DMSO-d₆), δ: -60.74 (s, 3F).

N-benzyl-4-(trifluoromethyl)aniline (277)

(C₁₄H₁₂F₃N; M.W. = 251.25)



general procedure 13

white solid

T.L.C. System: n-Hex/EtOAc 8:2 v/v, R_f: 0.59

Purification: Purification: Flash column chromatography Isolera One system (Biotage),

Cartridge: SNAP KP-Sil 10g (n-Hex-DCM 50:50 v/v in 6 CV).

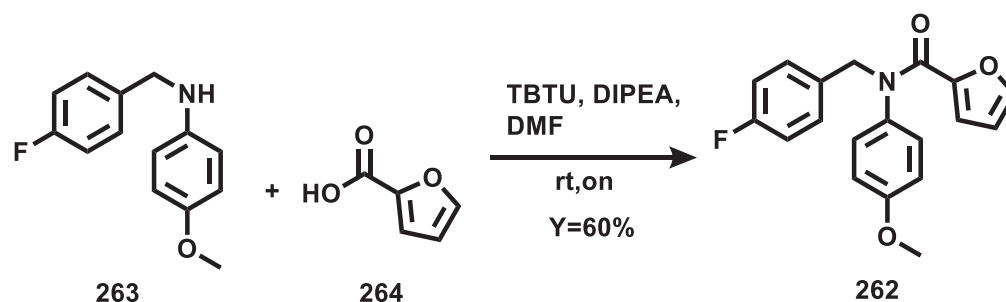
Yield: 102 mg (43%)

Purity: >95%

¹H NMR (DMSO-d₆), δ: 7.39 – 7.29 (m, 6H), 7.27 – 7.22 (m, 1H), 6.99 (t, *J* = 6.0 Hz, 1H), 6.67 (d, *J* = 8.6 Hz, 2H), 4.33 (d, *J* = 6.0 Hz, 2H).

¹³C NMR (DMSO-d₆), δ: 152.07, 139.85 (C, C-aromatic), 128.86, 127.59, 127.28 (CH, C-aromatic), 126.65 (q, *J* = 3.8 Hz, CH, C-aromatic), 112.06 (CH, C-aromatic), 46.33 (CH₂).

¹⁹F NMR (DMSO-d₆), δ: -58.88 (s, 3F).

N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)furan-2-carboxamide (**262**)*(C₁₉H₁₆FNO₃; M.W. = 325.34)**

A mixture of furan-2-carboxylic acid **264** (0.24mmol), *N*-(4-fluorobenzyl)-4-methoxyaniline **263** (0.29 mmol), TBTU (0.29mmol), and DIPEA (0.98 mmol) in anhydrous DMF (0.2 M) was stirred at room temperature overnight. The reaction mixture was diluted in EtOAc (10 mL) and was washed with a HCl (1N) solution (3 x 8 ml), a sat. NaHCO₃ solution (3 x 8 mL) and brine (3 x 8 mL). The organic layers were then dried over Na₂SO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP-Sil 5g, n-Hex - EtOAc 100:0 v/v increasing to 70:30 v/v in 24 CV) to give pure *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)furan-2-carboxamide **262** as white solid.

T.L.C. System: n-Hex/EtOAc 8:2 v/v, Rf: 0.2

Yield: 58 mg (60%)

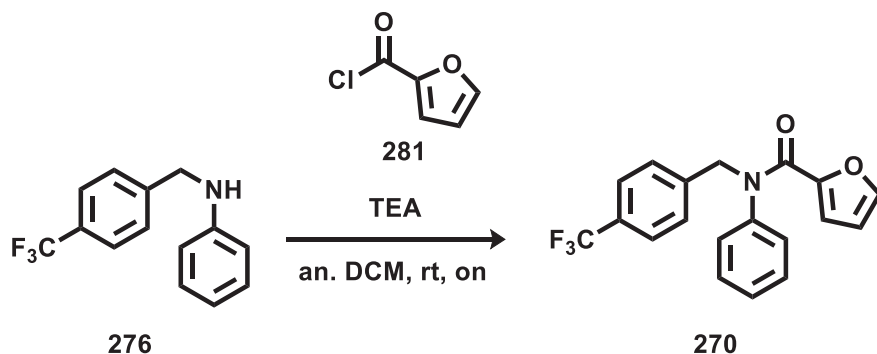
Purity: >95%

UPLC-MS method C: Rt: 1.92, MS (ESI)⁺: 326.2[M+1]⁺, 348.2[M+Na]⁺

¹H NMR (DMSO-d₆), δ: 7.71 – 7.67 (m, 1H), 7.27 (dd, ³J = 8.7 Hz, ⁴J = 5.6 Hz, 2H), 7.12 (t, J = 8.9 Hz, 2H), 7.03 (d, J = 8.9 Hz, 2H), 6.91 (d, J = 8.9 Hz, 2H), 6.39 (dd, ³J = 3.5 Hz, ⁴J = 1.7 Hz, 1H), 5.72 (s, J = 17.1 Hz, 1H), 4.93 (s, 2H), 3.74 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 162.80 (C=O), 159.04, 158.99 (C, C-aromatic), 146.97 (CH, C-aromatic), 145.64, 134.91, 133.93 (C, C-aromatic), 130.80, 129.85, 116.63, 115.59, 115.02, 111.78 (CH, C-aromatic), 55.75 (CH₃), 52.83 (CH₂).

¹⁹F NMR (DMSO-d₆), δ: -115.39 (s, F).

N*-phenyl-*N*-(4-(trifluoromethyl)benzyl)furan-2-carboxamide (270)*(C₁₉H₁₄F₃NO₂; M.W.= 345.32)**

general procedure 14

pale yellow solid

T.L.C. System: n-Hex/EtOAc 8:2 v/v, R_f: 0.32Purification: Purification: Flash column chromatography Isolera One system (Biotage),
Cartridge: SNAP KP-Sil 10g (n-Hex -EtOAc 100:0 v/v increasing to 70:30 v/v in 18 CV).

Yield: 70 mg (88%)

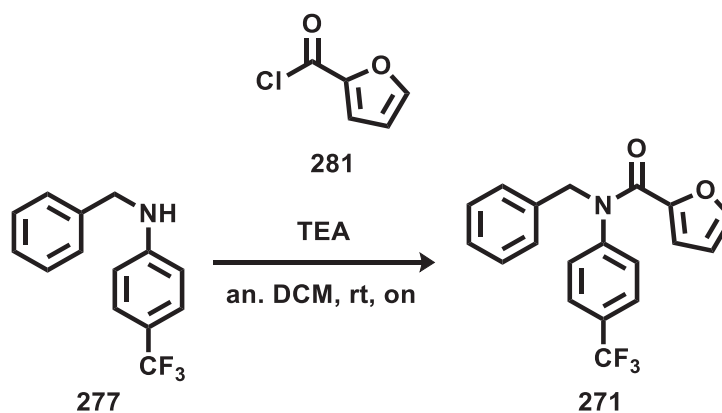
Purity: >95%

UPLC-MS method C: Rt: 2.031, MS (ESI)⁺: 346.2[M+1]⁺

¹H NMR (DMSO-*d*₆), δ: 7.72 – 7.64 (m, 3H), 7.50 (d, *J* = 8.0 Hz, 2H), 7.41 – 7.30 (m, 3H),
7.20 (d, *J* = 7.0 Hz, 2H), 6.40 (dd, ³*J* = 3.5 Hz, ⁴*J* = 1.7 Hz, 1H), 5.89 (d, *J* = 3.3 Hz, 1H),
5.11 (s, 2H).

¹³C NMR (DMSO-*d*₆), δ: 159.13 (C=O), 146.87, 142.53 (C, C-aromatic), 129.97, 129.17,
128.37, 128.35 (CH, C-aromatic), 125.75 (q, *J* = 3.7 Hz, CH, C-aromatic), 117.01, 111.83
(CH, C-aromatic), 53.18 (CH₂).

¹⁹F NMR (DMSO-*d*₆), δ: -60.86 (s, 3F).

N*-benzyl-*N*-(4-(trifluoromethyl)phenyl)furan-2-carboxamide (271)*(C₁₉H₁₄F₃NO₂; M.W.= 345.32)**

general procedure 14

yellow solid

T.L.C. System: n-Hex/EtOAc 8:2 v/v, Rf: 0.33

Purification: Purification: Flash column chromatography Isolera One system (Biotage),
Cartridge: SNAP KP-Sil 10g (n-Hex - DCM 100:0 v/v increasing to 0:100 v/v in 18 CV).

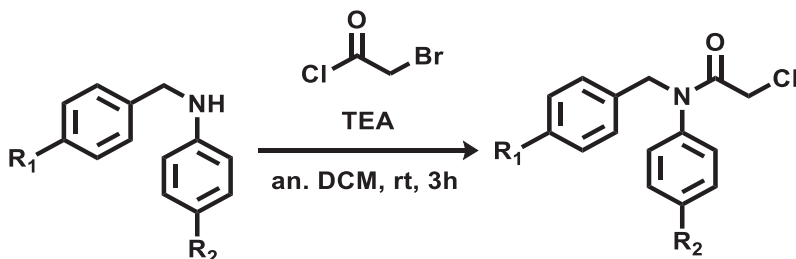
Yield: 80 mg (99%)

Purity: >95%

UPLC-MS method C: Rt: 2.059, MS (ESI)⁺: 346.2[M+1]⁺**¹H NMR (DMSO-d₆), δ:** 7.70 (d, *J* = 8.4 Hz, 2H), 7.66 (dd, *J* = 1.7, 0.8 Hz, 1H), 7.38 (d, *J* = 8.2 Hz, 2H), 7.33 – 7.20 (m, 5H), 6.47 (dd, ³*J* = 3.5 Hz, ⁴*J* = 1.7 Hz, 1H), 6.35 (dd, ³*J* = 3.5 Hz, ⁴*J* = 0.7 Hz, 1H), 5.10 (s, 2H).**¹³C NMR (DMSO-d₆), δ:** 159.17 (C=O), 147.03, 146.47 (C, C-aromatic), 145.94 (CH, C-aromatic), 137.31 (C, C-aromatic), 128.94, 128.58, 128.34 (CH, C-aromatic), 127.82 (q, *J* = 32.1 Hz, C, C-aromatic), 127.80 (CH, C-aromatic), 126.78 (q, *J* = 3.7 Hz, CH, C-aromatic), 124.42 (q, *J* = 272.2 Hz, CF₃), 117.52, 112.05 (CH, C-aromatic), 53.15 (CH₂).**¹⁹F NMR (DMSO-d₆), δ:** -60.81 (s, 3F).

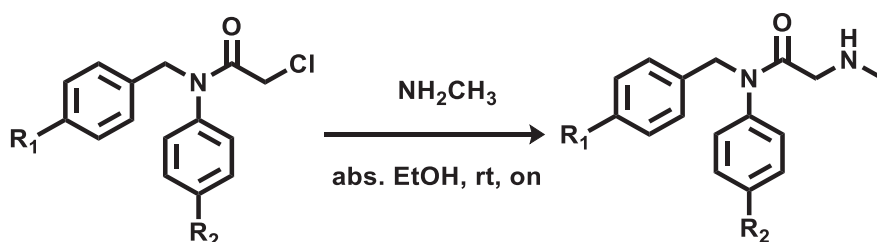
12.6.1 Synthesis of *N*-benzyl-2-(methylamino)-*N*-phenyl acetamides

General procedure 15: synthesis of *N*-benzyl-2-chloro-*N*-phenylacetamides

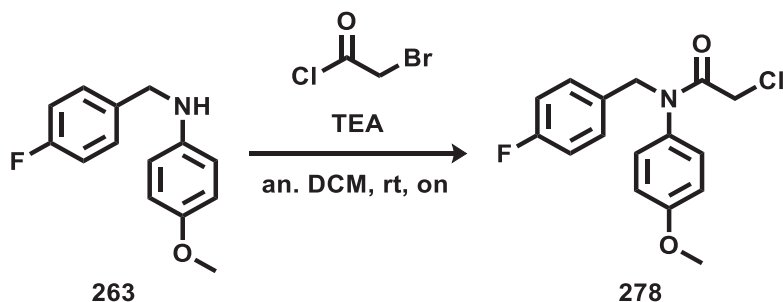


2-bromoacetyl chloride (0.31 mmol) was added drop-wise to a solution of *N*-(4-fluorobenzyl)-4-methoxyaniline (0.37 mmol) and TEA (0.62 mmol) in anhydrous DCM (0.2 M) at 0°C. The resulting mixture was stirred at room temperature for 3 hours. The reaction mixture was diluted in DCM (10mL) and washed with a 1M HCl solution (3x8 mL), brine (2 x 15mL), dried over Na₂SO₄, filtered and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system) to give pure *N*-benzyl-2-chloro-*N*-phenylacetamide.

General procedure 16: synthesis of *N*-benzyl-2-(methylamino)-*N*-phenylacetamides



N-benzyl-2-chloro-*N*-phenylacetamide (mmol) was dissolved in absolute EtOH and methylamine in EtOH (mmol) was added. The reaction mixture was stirred at room temperature overnight. The solvent was evaporated under reduced pressure and the resulting residue was purified by flash column chromatography (Biotage Isolera One system) to give the desired *N*-benzyl-2-(methylamino)-*N*-phenylacetamide.

2-chloro-N-(4-fluorobenzyl)-N-(4-methoxyphenyl)acetamide (278)**(C₁₆H₁₅ClFNO₂; M.W.: 307.75)**

general procedure 15

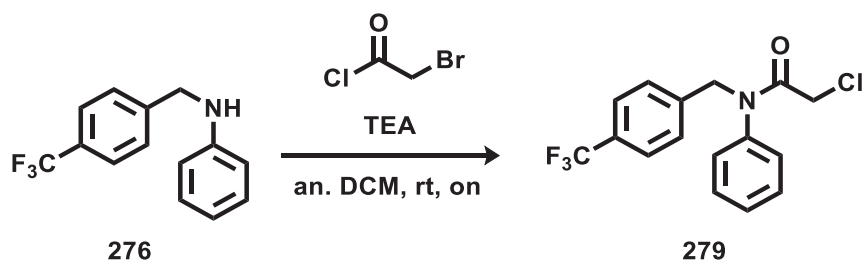
yellow oil

T.L.C. System: DCM, Rf: 0.6

Purification: Purification: Flash column chromatography Isolera One system (Biotage),
Cartridge: SNAP KP-Sil 10g (DCM - MeOH 100:0 v/v increasing to 99:1 v/v in 8 CV).

Yield: 47 mg (87%)

Purity: >90%

¹H NMR (DMSO-d₆), δ: 7.25 – 7.18 (m, 2H), 7.18 – 7.08 (m, 4H), 6.95 – 6.84 (m, 2H), 4.81 (s, 2H), 4.04 (s, 2H), 3.74 (s, 3H).**2-chloro-N-phenyl-N-(4-(trifluoromethyl)benzyl)acetamide (279)****(C₁₆H₁₃ClF₃NO; M.W.: 327.73)**

general procedure 15

yellow/orange oil

T.L.C. System: DCM, Rf: 0.59

Purification: Purification: Flash column chromatography Isolera One system (Biotage),
Cartridge: SNAP KP-Sil 25g (isocratic DCM in 5 CV).

Yield: 77 mg (68%)

Purity: >95%

UPLC-MS method C: Rt: 2.037, MS (ESI)⁺: 328.2[M+1]⁺

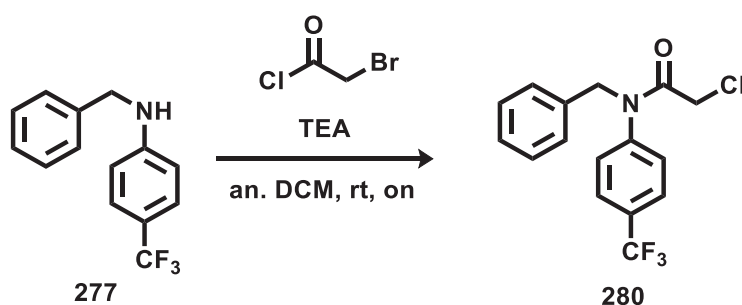
¹H NMR (DMSO-d₆), δ: 7.67 (d, *J* = 8.1 Hz, 2H), 7.49 – 7.39 (m, 4H), 7.35 (dd, ³*J* = 14.6 Hz, ⁴*J* = 7.7 Hz, 1H), 7.31 (d, *J* = 7.3 Hz, 2H), 4.99 (s, 2H), 4.11 (s, 1H).

¹³C NMR (DMSO-d₆), δ: 172.05, 142.70 (C, C-aromatic), 130.10, 129.06, 128.38 (CH, C-aromatic), 127.97 (q, *J* = 31.8 Hz, C, C-aromatic), 125.68 (q, *J* = 3.8 Hz, CH, C-aromatic), 122.56 (q, *J* = 272.2 Hz, CF₃), 60.62, 52.31 (CH₂).

¹⁹F NMR (DMSO-d₆), δ: -60.86 (s, 3F).

***N*-benzyl-2-chloro-*N*-(4-(trifluoromethyl)phenyl)acetamide(280)**

(C₁₆H₁₃ClF₃NO; M.W.: 327.73)



general procedure 15

yellow oil

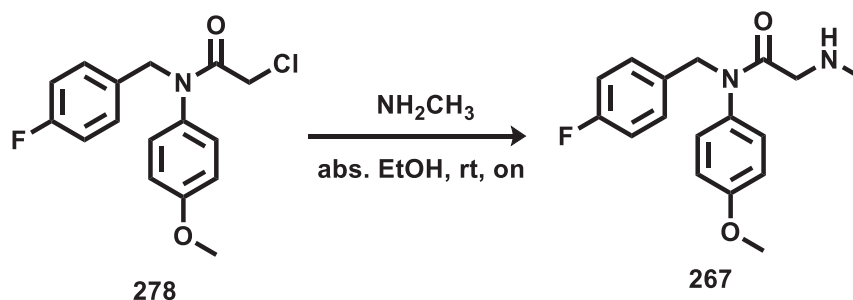
T.L.C. System: DCM, R_f: 0.56

Purification: Purification: Flash column chromatography Isolera One system (Biotage),
Cartridge: SNAP KP-Sil 10g (n-Hex-DCM 100:0 v/v increasing to 0:100 v/v in 16 CV).

Yield: 88 mg (76%)

Purity: >95%

¹H NMR (DMSO-d₆), δ: 7.78 (d, *J* = 8.3 Hz, 2H), 7.52 (d, *J* = 8.2 Hz, 2H), 7.34 – 7.16 (m, 5H), 4.96 (s, 2H), 4.19 (s, 2H).

N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)-2-(methylamino)acetamide (267)*(C₁₇H₁₉FN₂O₂; M.W.: 302.35)**

general procedure 16

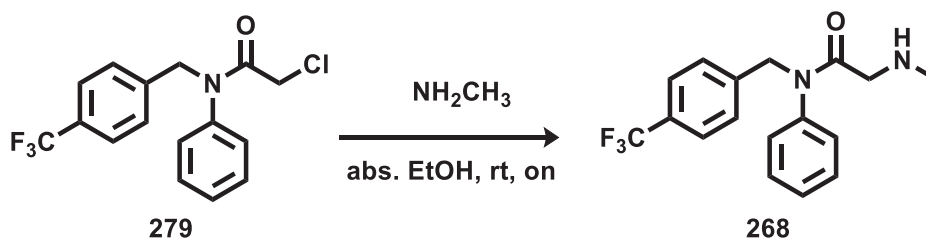
yellow oil

T.L.C. System: DCM/MeOH 95:5 v/v, Rf: 0.22

Purification: Purification: Flash column chromatography Isolera One system (Biotage),
Cartridge: SNAP KP-Sil 10g (DCM - MeOH 100:0 v/v increasing to 95:5 v/v in 18 CV).

Yield: 18 mg (66%)

Purity: >95%

UPLC-MS method C: Rt: 1.452, MS (ESI)⁺: 303.2[M+1]⁺**¹H NMR (DMSO-d₆), δ:** 7.21 (dd, ³J = 8.5 Hz, ⁴J = 5.7 Hz, 2H), 7.11 (t, J = 8.9 Hz, 2H),
7.05 (d, J = 8.9 Hz, 2H), 6.92 (d, J = 8.9 Hz, 2H), 4.81 (s, 2H), 3.74 (s, 3H), 3.03 (s, 2H),
2.22 (s, 3H).**¹³C NMR (DMSO-d₆), δ:** 170.70 (C=O), 162.75, 160.82, 159.04, 133.96 (C, C-aromatic),
130.61, 129.76, 115.55, 115.16 (CH, C-aromatic), 55.75 (CH₃), 52.51, 51.87 (CH₂), 35.98
(CH₃).**¹⁹F NMR (DMSO-d₆), δ:** -115.53 (s, F).**2-(methylamino)-*N*-phenyl-*N*-(4-(trifluoromethyl)benzyl)acetamide (268)****(C₁₇H₁₇F₃N₂O; M.W.: 322.33)**

general procedure 16

yellow oil

T.L.C. System: DCM/MeOH 95:5 v/v, Rf: 0.37

Purification: Purification: Flash column chromatography Isolera One system (Biotage),
Cartridge: SNAP KP-Sil 10g (DCM - MeOH 100:0 v/v increasing to 95:5 v/v in 18 CV).

Yield: 62 mg (64%)

Purity: >95%

UPLC-MS method C: Rt: 1.522, MS (ESI)⁺: 323.3[M+1]⁺

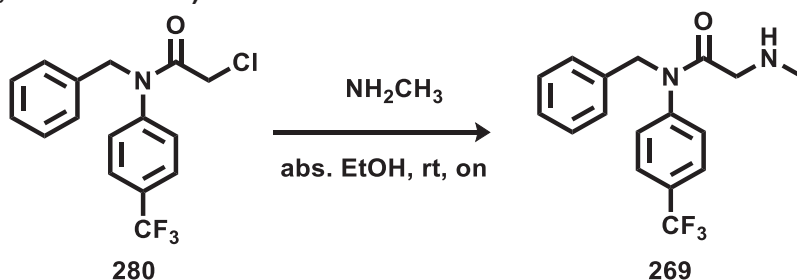
¹H NMR (DMSO-d₆), δ: 7.66 (d, *J* = 8.1 Hz, 2H), 7.41 (m, 4H), 7.37 – 7.31 (m, 1H), 7.24 (d, *J* = 7.2 Hz, 2H), 4.98 (s, 2H), 3.08 (s, 2H), 2.22 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 170.94 (C=O), 142.77, 141.45, 130.15 (C, C-aromatic), 129.05, 128.55, 128.43 (CH, C-aromatic), 128.21 (q, *J* = 31.6 Hz, C, C-aromatic), 125.69 (q, *J* = 3.7 Hz, CH, C-aromatic), 122.56 (q, *J* = 272.2 Hz, CF₃), 52.69, 52.27 (CH₂), 36.09 (CH₃).

¹⁹F NMR (DMSO-d₆), δ: -115.53 (s, 3F).

***N*-benzyl-2-(methylamino)-*N*-(4-(trifluoromethyl)phenyl)acetamide (269)**

(C₁₇H₁₇F₃N₂O; M.W.: 322.33)



general procedure 16

yellow oil

T.L.C. System: DCM/MeOH 95:5 v/v, Rf: 0.37

Purification: Purification: Flash column chromatography Isolera One system (Biotage),
Cartridge: ZIP KP Sil 5g (DCM - MeOH 100:0 v/v increasing to 97:3 v/v in 11 CV).

Yield: 52 mg (67%)

Purity: >95%

UPLC-MS method C: Rt: 1.565, MS (ESI)⁺: 323.3[M+1]⁺

¹H NMR (DMSO-d₆), δ: 7.75 (d, *J* = 8.4 Hz, 2H), 7.46 (d, *J* = 8.2 Hz, 2H), 7.34 – 7.16 (m, 5H), 4.95 (s, 2H), 3.11 (s, 2H), 2.22 (s, 3H).

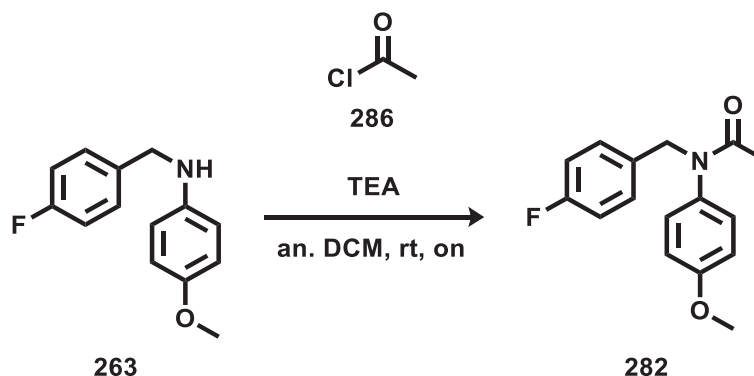
¹³C NMR (DMSO-d₆), δ: 176.62 (C=O), 167.48, 137.58 (C, C-aromatic), 128.91, 128.19, 127.71, 126.99 (q, *J* = 4.5 Hz, CH, C-aromatic), 53.04, 52.29 (CH₂), 36.24 (CH₃).

¹⁹F NMR (DMSO-d₆), δ: -60.94 (s, 3F).

12.6.2 Synthesis of furamide analogues

N-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)acetamide (**282**)

(C₁₆H₁₆FNO₂; M.W.: 273.31)



Acetyl chloride **286** (0.35 mmol) was added drop-wise to a solution, at 0°C, of *N*-(4-fluorobenzyl)-4-methoxyaniline **263** (0.22 mmol), triethylamine (0.43 mmol) in anhydrous DCM (0.1 M) under nitrogen atmosphere. The reaction was allowed to warm-up to rt and was stirred overnight. The reaction mixture was concentrated at reduced pressure and the crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 5g, n-hex - EtOAc 100:0 v/v increasing to 60:40 v/v in 17 CV) to afford *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)acetamide **282** as a colorless oil.

T.L.C. System: n-Hex/EtOAc 7:3 v/v, R_f: 0.26

Yield: 49 mg (83%)

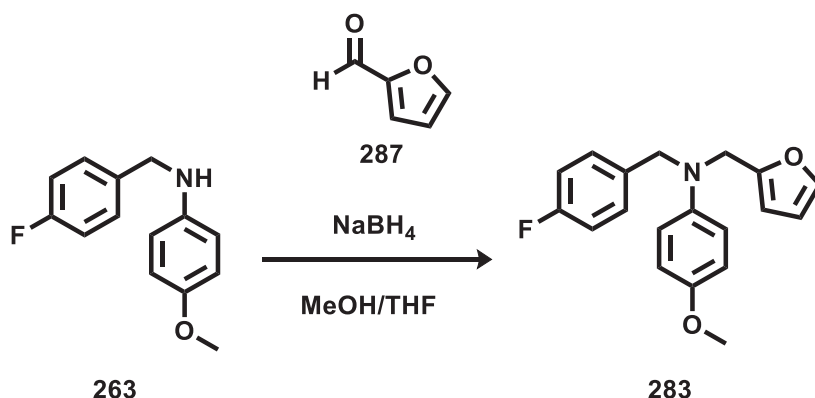
Purity: >95%

UPLC-MS method C: Rt: 1.813, MS (ESI)⁺: 274.1[M+1]⁺

¹H NMR (DMSO-d₆), δ: 7.20 (dd, ³J = 8.5 Hz, ⁴J = 5.7 Hz, 2H), 7.10 (t, J = 8.9 Hz, 2H), 7.06 (d, J = 8.9 Hz, 2H), 6.91 (d, J = 8.9 Hz, 2H), 4.78 (s, 2H), 3.73 (s, 3H), 1.79 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 170.07, 162.69, 158.73, 135.65, 134.47 (C, C-aromatic), 130.51, 129.61, 115.48, 115.00 (CH, C-aromatic), 55.71 (CH₃), 51.50 (CH₂), 22.77 (CH₃).

¹⁹F NMR (DMSO-d₆), δ: -115.74 (s, F).

N*-(4-fluorobenzyl)-*N*-(furan-2-ylmethyl)-4-methoxyaniline (283)*(C₁₉H₁₈FNO₂; M.W.: 311.36)**

N-(4-fluorobenzyl)-4-methoxyaniline **263** (3.59 mmol) was dissolved in MeOH/THF (4:1, 2 ml) and furan-2-carbaldehyde **287** (0.33 mmol) was added. The reaction mixture was stirred at room temperature for 3 hours. The reaction mixture was cooled 0°C before adding Sodium triacetoxyborohydride (0.65 mmol). The reaction was allowed to warm-up to room temperature and stirred overnight. The reaction mixture was concentrated under reduced pressure. The resulting residue was suspended in water and extracted with DCM (3x20 ml). The combined organic layers were washed with brine (3x20 ml), dried over Na₂SO₄, filtered and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP Sil 5g, n-hex - EtOAc 100:0 v/v increasing to 60:40 v/v in 17 CV) to afford *N*-(4-fluorobenzyl)-*N*-(furan-2-ylmethyl)-4-methoxyaniline **283** as a yellow oil.

T.L.C. System: n-Hex/EtOAc 9:1 v/v, R_f: 0.32

Yield: 31 mg (30%)

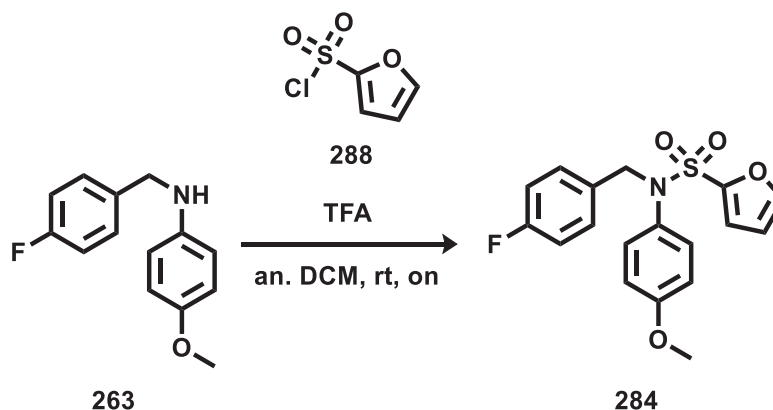
Purity: >95%

UPLC-MS method C: Rt: 2.13, MS (ESI)⁺: 312.1[M+1]⁺

¹H NMR (DMSO-d₆), δ: 7.57 (dd, ³J = 1.8 Hz, ⁴J = 0.8 Hz, 1H), 7.29 – 7.23 (m, 2H), 7.11 (ddd, ³J = 9.6, 5.9 Hz, ⁴J = 2.6 Hz, 2H), 6.74 (s, 4H), 6.37 (dd, ³J = 3.2 Hz, ⁴J = 1.8 Hz, 1H), 6.25 (dd, ³J = 3.2 Hz, ⁴J = 0.7 Hz, 1H), 4.49 (s, 2H), 4.47 (s, 2H), 3.63 (s, 3H).

¹³C NMR (DMSO-d₆), δ: 162.49, 152.92, 151.95, 142.79 (C, C-aromatic), 142.69 (CH, C-aromatic), 135.77 (C, C-aromatic), 129.23, 115.64, 115.51, 114.84, 110.79, 108.26 (CH, C-aromatic), 55.65 (CH₃), 54.32, 48.69 (CH₂).

¹⁹F NMR (DMSO-d₆), δ: -116.47 (s, F).

N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)furan-2-sulfonamide (**284**)*(C₁₈H₁₆FNO₄S; M.W.: 361.39)**

furan-2-sulfonyl chloride **288** (0.49 mmol) was added drop-wise to a solution, at 0°C, of *N*-(4-fluorobenzyl)-4-methoxyaniline **263** (0.43 mmol), triethylamine (0.87 mmol) in anhydrous DCM (0.1 M) under nitrogen atmosphere. The reaction was allowed to warm-up to rt and was stirred overnight. The reaction mixture was diluted with DCM and extracted with 1N HCl solution (3x10ml). The organic layer was washed with brine (3 x 10 mL), dried over MgSO₄ and evaporated at reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP-Sil 10g, n-Hex - EtOAc 100:0 v/v increasing to 70:30 v/v in 18 CV) to *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)furan-2-sulfonamide **284** as a white powder.

T.L.C. System: n-Hex/EtOAc 8:2v/v, R_f: 0.3

Yield: 153 mg (99%)

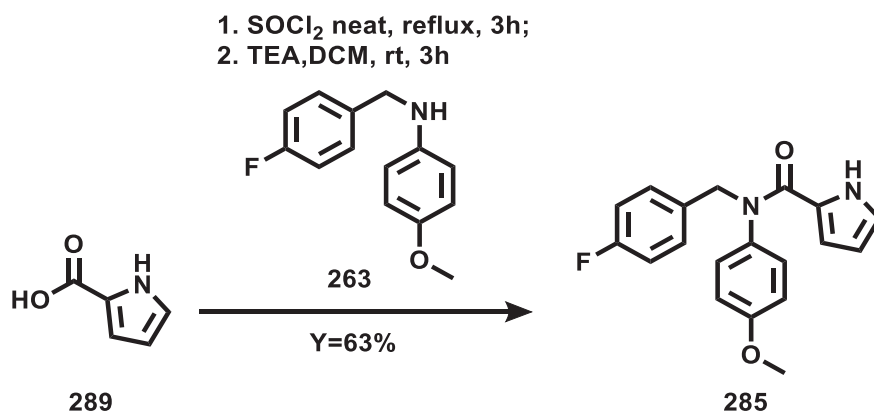
Purity: >95%

UPLC-MS method C: Rt: 1.981, MS (ESI)⁺: 362.3[M+1]⁺

¹H NMR (DMSO-*d*₆), δ: 8.09 (dd, ³*J* = 1.8 Hz, ⁴*J* = 0.9 Hz, 1H), 7.29 – 7.24 (m, 2H), 7.15 – 7.08 (m, 3H), 6.97 – 6.93 (m, 2H), 6.85 – 6.81 (m, 2H), 6.75 (dd, ³*J* = 3.5 Hz, ⁴*J* = 1.8 Hz, 1H), 4.81 (s, 2H), 3.70 (s, 3H).

¹³C NMR (DMSO-*d*₆), δ: 162.96, 161.02, 159.19 (C, C-aromatic), 148.29 (CH, C-aromatic), 147.41, 132.81, 130.86 (C, C-aromatic), 130.71, 130.51, 118.10, 115.67, 114.68, 112.20 (CH, C-aromatic), 55.70 (CH₃), 54.25 (CH₂).

¹⁹F NMR (DMSO-*d*₆), δ: -114.88 (s, F).

N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)-1*H*-pyrrole-2-carboxamide (**285**)*(C₁₉H₁₇FN₂O₂; M.W.: 324.36)**

1*H*-pyrrole-2-carboxylic acid **289** (0.43 mmol) in Thionyl chloride (0.5 ml) was refluxed for 2 hours. SOCl₂ was eliminated from the reaction mixture under reduced pressure. The resulting residue was dissolved in anhydrous DCM (2 ml) and added drop-wise to a solution, at 0°C, of *N*-(4-fluorobenzyl)-4-methoxyaniline **263** (0.43 mmol), triethylamine (0.43 mmol) in anhydrous DCM (4 ml) under nitrogen atmosphere. The reaction was allowed to warm-up to room temperature and was stirred overnight. The reaction mixture was diluted with DCM and extracted with 1*N* HCl solution (3x10ml). The organic layer was washed with brine (3 x 10 mL), dried over Na₂SO₄, filtered and evaporated at reduced pressure. The obtained crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: SNAP KP-Sil 10g, n-Hex - EtOAc 100:0 v/v increasing to 70:30 v/v in 19 CV) to afford *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)-1*H*-pyrrole-2-carboxamide **285** as a white solid.

T.L.C. System: n-Hex/EtOAc 8:2 v/v, R_f: 0.3

Yield: 89 mg (63%)

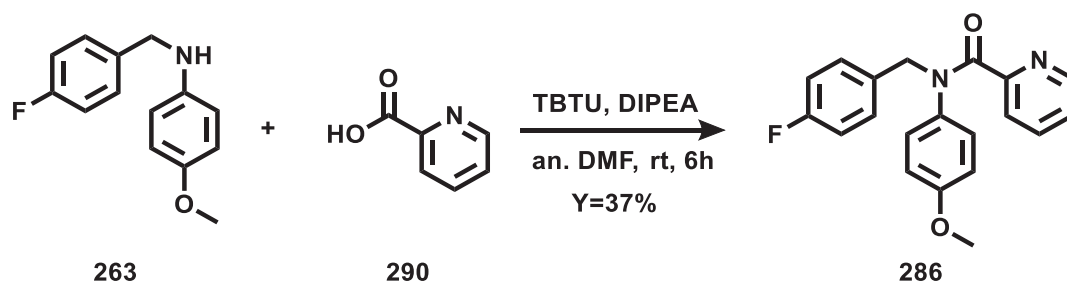
Purity: >95%

UPLC-MS method C: R_t: 1.924, MS (ESI)⁺: 325.2[M+1]⁺

¹H NMR (DMSO-*d*₆), δ: 11.53 (s, 1H), 7.28 (dd, ³*J* = 8.7 Hz, ⁴*J* = 5.6 Hz, 2H), 7.12 (t, *J* = 8.9 Hz, 2H), 7.02 (d, *J* = 9.0 Hz, 2H), 6.93 (d, *J* = 9.0 Hz, 2H), 6.80 (td, ³*J* = 2.8 Hz, ⁴*J* = 1.4 Hz, 1H), 5.81 (dt, ³*J* = 3.8 Hz, ⁴*J* = 2.5 Hz, 1H), 4.92 (s, 2H), 4.80 (ddd, *J* = 3.7, 2.5, 1.4 Hz, 1H), 3.76 (s, 3H).

¹³C NMR (DMSO-*d*₆), δ: 162.73, 161.42, 159.12, 135.65, 134.57 (C, C-aromatic), 130.73, 130.43 (CH, C-aromatic), 124.97 (C, C-aromatic), 122.02, 115.51, 115.03, 113.76, 109.21 (CH, C-aromatic), 55.76 (CH₃), 52.91 (CH₂).

¹⁹F NMR (DMSO-*d*₆), δ: -115.63 (s, F).

N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)picolinamide (286)*(C₂₀H₁₇FN₂O₂; M.W.: 336.37)**

A mixture of picolinic acid **290** (0.18 mmol), *N*-(4-fluorobenzyl)-4-methoxyaniline **263** (0.22 mmol), TBTU (0.22 mmol), and DIPEA (0.72 mmol) in anhydrous DMF (0.2 M) was stirred at room temperature overnight. The reaction mixture was diluted in EtOAc (20 mL) and was washed with sat. NaHCO₃ solution (3 x 20 mL) and brine (3 x 20 mL). The organic layers were then dried over Na₂SO₄ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (Biotage Isolera One system, Cartridge: ZIP KP-Sil 5g, DCM - MeOH 100:0 v/v increasing to 99:1 v/v in 11 CV) to give pure *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)picolinamide **286** as white solid.

T.L.C. System: DCM/MeOH 98:2 v/v, R_f: 0.29

Yield: 26 mg (37%)

Purity: >95%

UPLC-MS method C: Rt: 1.806, MS (ESI)⁺: 337.3[M+1]⁺, 359.3[M+Na]⁺

¹H NMR (DMSO-*d*₆), δ: 8.32 (s, 1H), 7.71 (t, *J* = 6.9 Hz, 1H), 7.45 (d, *J* = 7.4 Hz, 1H), 7.33 (dd, ³*J* = 8.4 Hz, ⁴*J* = 5.7 Hz, 2H), 7.23 (s, 1H), 7.15 (t, *J* = 8.5 Hz, 2H), 6.89 (d, *J* = 8.0 Hz, 2H), 6.67 (d, *J* = 8.0 Hz, 2H), 5.03 (s, 2H), 3.61 (s, 3H).

¹³C NMR (DMSO-*d*₆), δ: 168.85, 162.79, 157.93, 155.00 (C, C-aromatic), 148.79, 136.92 (CH, C-aromatic), 135.08, 133.96 (C, C-aromatic), 130.50, 129.48, 124.29, 123.50, 115.65, 114.28 (CH, C-aromatic), 55.55 (CH₃), 52.04 (CH₂).

¹⁹F NMR (DMSO-*d*₆), δ: -115.44 (s, F).

12.7 MD simulations of the ATPase pocket

Molecular Dynamics Simulations

Chain A and chain F of the EV-A71 2C protein crystal structure (PDB ID:5GRB) were used to prepare the system for the MD simulations of the ATPase pocket. OPLS3 was used as a force field. The protein-ligand complex system was prepared, placing the protein-ligand complex in a cubic box (buffer 10Å) and the system was solvated with water molecules, using the TIP3P model as a solvent model. The negative charges on the protein were neutralized, adding Na⁺ atoms to the system. Magnesium chloride (10 mM) was added to the box to simulate physiological conditions.

The protein-ligand complex system was then equilibrated for 112ps at 10 K in an NVT ensemble and then simulated for 48 ps at a constant pressure of 1 atm using the NPT ensemble. The equilibrated systems were then MD simulated for 500 ns at constant temperature (300 K) and pressure recording snapshots every 20 ps. The RMSD values of C α atoms during the simulations were used as an indication of structural stability and simulation integrity (see figure 44 in the appendix).

RMSD Clustering analysis

An RMSD conformational clustering of the MD simulations was performed to obtain a reduced number of representative structures for the ATPase pocket of the 2C protein. For the clustering analysis of the MD simulations were considered only the last 200ns of each trajectory. The structure of the dimer of 2C protein was extracted every 20 ps. The resulting 10000 trajectory structures for each simulation were superimposed using all C α atoms. The RMSD-clustering was performed on the subset of atoms of the binding pocket surrounding the ATP- γ -S (cutoff of 5Å). These atoms were clustered using the atom-positional RMSD of all atoms as the similarity criterion with the DBSCAN algorithm. Eight clusters were obtained considering all the frames with a neighbouring distance of 1Å ($\epsilon=1\text{\AA}$) with 30 as a minimum number of neighbouring frames. The cluster centroids were chosen as the representative structure of the ATPase pocket.

References

- 1 Duintjer Tebbens RJ, Pallansch MA, Chumakov KM, Halsey NA, Hovi T, Minor PD, et al. Expert review on poliovirus immunity and transmission. *Risk Anal.* **2013**, 33, 544–605. <https://doi.org/10.1111/j.1539-6924.2012.01864.x>
- 2 Tapparel C, Siegrist F, Petty TJ, Kaiser L. Picornavirus and enterovirus diversity with associated human diseases. *Infect. Genet. Evol.* **2013**, 14, 282–93. <https://doi.org/10.1016/j.meegid.2012.10.016>
- 3 Pons-Salort M, Parker EPK, Grassly NC. The epidemiology of non-polio enteroviruses: Recent advances and outstanding questions. *Curr. Opin. Infect. Dis.* **2015**, 28, 479–87. <https://doi.org/10.1097/QCO.000000000000187>
- 4 Baggen J, Thibaut HJ, Strating JRPM, Van Kuppeveld FJM. The life cycle of non-polio enteroviruses and how to target it. *Nature Reviews Microbiology.* **2018**, 16, 368–381. <https://doi.org/10.1038/s41579-018-0005-4>
- 5 Pilipenko EV., Gmyl AP, Maslova SV, Svitkin YV, Sinyakov AN, Agol VI. Prokaryotic-like cis elements in the cap-independent internal initiation of translation on picornavirus RNA. *Cell* **1992**, 68, 119–131. [https://doi.org/10.1016/0092-8674\(92\)90211-t](https://doi.org/10.1016/0092-8674(92)90211-t)
- 6 Benschop KSM, Van Der Avoort HG, Duizer E, Koopmans MPG. Antivirals against enteroviruses: A critical review from a public-health perspective. *Antivir. Ther.* **2015**, 20, 121–130. <https://doi.org/10.3851/IMP2939>
- 7 Bauer L, Lyoo H, van der Schaar HM, Strating JR, van Kuppeveld FJ. Direct-acting antivirals and host-targeting strategies to combat enterovirus infections. *Curr. Opin. Virol.* **2017**, 24, 1–8. <https://doi.org/10.1016/j.coviro.2017.03.009>
- 8 Echeverri A, Banerjee R, Dasgupta A. Amino-terminal region of poliovirus 2C protein is sufficient for membrane binding. *Virus Res.* **1998**, 54, 217–223. [https://doi.org/10.1016/s0168-1702\(98\)00016-1](https://doi.org/10.1016/s0168-1702(98)00016-1)
- 9 Echeverri AC, Dasgupta A. Amino Terminal Regions of Poliovirus 2C Protein Mediate Membrane Binding. *Virology* **1995**, 208, 540–553. <https://doi.org/10.1006/viro.1995.1185>

- 10** Teterina NL, Gorbalenya AE, Egger D, Bienz K, Ehrenfeld E. Poliovirus 2C protein determinants of membrane binding and rearrangements in mammalian cells. *J. Virol.* **1997**, 71, 8962–8972. <https://jvi.asm.org/content/71/12/8962/article-info>
- 11** Banerjee R, Dasgupta A. Interaction of picornavirus 2C polypeptide with the viral negative-strand RNA. *Journal of General Virology* **2001**, 82, 2621–2627. <https://doi.org/10.1099/0022-1317-82-11-2621>
- 12** Rodríguez PL, Carrasco L. Poliovirus protein 2C contains two regions involved in RNA binding activity. *J. Biol. Chem.* **1995**, 270, 10105–10112. <https://doi.org/10.1074/jbc.270.17.10105>
- 13** Pfister T, Jones KW, Wimmer E. A cysteine-rich motif in poliovirus protein 2C(ATPase) is involved in RNA replication and binds zinc in vitro. *J. Virol.* **2000**, 74, 334–43. <https://doi.org/10.1128/jvi.74.1.334-343.2000>
- 14** Rodriguez PL, Carrasco L. Poliovirus protein 2C has ATPase and GTPase activities. *J. Biol. Chem.* **1993**, 268, 8105–8110. <http://www.jbc.org/content/268/11/8105.long>
- 15** Tolskaya EA, Romanova LI, Kolesnikova MS, Gmyl AP, Gorbalenya AE, Agol VI. Genetic studies on the poliovirus 2C protein, an NTPase A plausible mechanism of guanidine effect on the 2C function and evidence for the importance of 2C oligomerization. *J. Mol. Biol.* **1994**, 236, 1310–1323. [https://doi.org/10.1016/0022-2836\(94\)90060-4](https://doi.org/10.1016/0022-2836(94)90060-4)
- 16** Mirzayan C, Wimmer E. Biochemical Studies on Poliovirus Polypeptide 2C: Evidence for ATPase Activity. *Virology* 1994, 199, 176–187. <https://doi.org/10.1006/viro.1994.1110>
- 17** Mirzayan C, Wimmer E. Genetic analysis of an NTP-binding motif in poliovirus polypeptide 2C. *Virology* **1992**, 189, 547–555. [https://doi.org/10.1016/0042-6822\(92\)90578-d](https://doi.org/10.1016/0042-6822(92)90578-d)
- 18** De Palma AM, Heggermont W, Lanke K, Coutard B, Bergmann M, Monforte A-M, et al. The Thiazolobenzimidazole TBZE-029 Inhibits Enterovirus Replication by Targeting a Short Region Immediately Downstream from Motif C in the Nonstructural Protein 2C. *J. Virol.* **2008**, 82, 4720–4730. <https://doi.org/10.1128/JVI.01338-07>
- 19** Sweeney TR, Cisnetto V, Bose D, Bailey M, Wilson JR, Zhang X, et al. Foot-and-mouth disease virus 2C is a hexameric AAA+ protein with a coordinated ATP hydrolysis

- mechanism. *J. Biol. Chem.* **2010**, 285, 24347–24359. <https://doi.org/10.1074/jbc.M110.129940>
- 20** Papageorgiou N, Lantez V, Chauvet O, Alexander E. The 2C putative helicase of echovirus 30 adopts a hexameric ring-shaped structure. *Biological Crystallography*, **2010**, 1116–1120. <https://doi.org/10.1107/S090744491002809X>
- 21** Xia H, Wang P, Wang GC, Yang J, Sun X, Wu W, et al. Human Enterovirus Nonstructural Protein 2CATPase Functions as Both an RNA Helicase and ATP-Independent RNA Chaperone. *PLoS Pathog.* **2015**, 11, 1–29. <https://doi.org/10.1371/journal.ppat.1005067>
- 22** Bienz K, Egger D, Pfister T, Troxler M. Structural and functional characterization of the poliovirus replication complex. *J. Virol.* **1992**, 66, 2740–2747. <http://jvi.asm.org/cgi/pmidlookup?view=long&pmid=1313898>
- 23** Adams P, Kandiah E, Effantin G, Steven AC, Ehrenfeld E. Poliovirus 2C protein forms homo-oligomeric structures required for ATPase activity. *J. Biol. Chem.* **2009**, 284, 22012–22021. <https://doi.org/10.1074/jbc.M109.031807>
- 24** Teterina NL, Gorbalenya AE, Egger D, Bienz K, Ehrenfeld E. Poliovirus 2C protein determinants of membrane binding and rearrangements in mammalian cells. *J. Virol.* **1997**, 71, 8962–8972. <https://jvi.asm.org/content/71/12/8962.long>
- 25** Cho MW, Teterina N, Egger D, Bienz K, Ehrenfeld E. Membrane Rearrangement and Vesicle Induction by Recombinant Poliovirus 2C and 2BC in Human Cells. *Virology* **1994**, 202, 129–145. <https://doi.org/10.1006/viro.1994.1329>
- 26** Liu Y, Wang C, Mueller S, Paul A V., Wimmer E, Jiang P. Direct Interaction between Two Viral Proteins, the Nonstructural Protein 2C ATPase and the Capsid Protein VP3, Is Required for Enterovirus Morphogenesis. *PLoS Pathog.* **2010**, 6, e1001066. <https://doi.org/10.1371/journal.ppat.1001066>
- 27** Vance LM, Moscufo N, Chow M, Heinz BA. Poliovirus 2C region functions during encapsidation of viral RNA. *J. Virol.* **1997**, 71, 8759–8765. <https://jvi.asm.org/content/71/11/8759.long>
- 28** Wang C, Jiang P, Sand C, Paul A V, Wimmer E. Alanine Scanning of Poliovirus 2C ATPase Reveals New Genetic Evidence that Capsid Protein/2C ATPase Interactions Are

Essential for Morphogenesis. *J. Virol.* **2012**, 86, 9964–9975.
<https://doi.org/10.1128/JVI.00914-12>

29 Guan H, Tian J, Qin B, Wojdyla JA, Wang B, Zhao Z, et al. Crystal structure of 2C helicase from enterovirus 71. *Sci. Adv.* **2017**, 3, e1602573
<https://doi.org/10.1126/sciadv.1602573>

30 Pincus SE, Diamond DC, Emini EA, Wimmer E. Guanidine-selected mutants of poliovirus: mapping of point mutations to polypeptide 2C. *J. Virol.* **1986**, 57, 638–646.
<https://jvi.asm.org/content/57/2/638.long>

31 Hadaschik D, Klein M, Zimmermann H, Eggers HJ, Nelsen-Salz B. Dependence of echovirus 9 on the enterovirus RNA replication inhibitor 2-(alpha-Hydroxybenzyl)-benzimidazole maps to nonstructural protein 2C. *J. Virol.* **1999**, 73, 10536–10539.
<https://jvi.asm.org/content/73/12/10536>

32 Pincus SE, Rohl H, Wimmer E. Guanidine-dependent mutants of poliovirus: Identification of three classes with different growth requirements. *Virology.* **1987**, 157, 83–88. [https://doi.org/10.1016/0042-6822\(87\)90316-3](https://doi.org/10.1016/0042-6822(87)90316-3)

33 Baltera RF, Tershak DR. Guanidine-resistant mutants of poliovirus have distinct mutations in peptide 2C. *J. Virol.* **1989**, 63, 4441–4444.
<https://jvi.asm.org/content/63/10/4441.long>

34 Loddo B, Ferrari W, Brotzu G, Spanedda A. In vitro Inhibition of Infectivity of Polio Viruses by Guanidine. *Nature* **1962**, 193, 97–98. <https://doi.org/10.1038/193097a0>

35 Rightsel WA, Dice Jr, Mcalpine RI, Timm EA, Mclean IW, Dixon GI, et al. Antiviral effect of guanidine. *Science* **1961**, 134, 558–559.
<https://doi.org/10.1126/science.134.3478.558>

36 Barton DJ, Flanagan JB. Synchronous replication of poliovirus RNA: initiation of negative-strand RNA synthesis requires the guanidine-inhibited activity of protein 2C. *J. Virol.* **1997**, 71, 8482–8489. <https://jvi.asm.org/content/71/11/8482.long>

37 Caliguirri LA, Tamm I. Action of guanidine on the replication of poliovirus RNA. *Virology* **1968**, 35, 408–417. [https://doi.org/10.1016/0042-6822\(68\)90219-5](https://doi.org/10.1016/0042-6822(68)90219-5)

- 38** Pfister T, Wimmer E. Characterization of the nucleoside triphosphatase activity of poliovirus protein 2C reveals a mechanism by which guanidine inhibits poliovirus replication. *J. Biol. Chem.* **1999**, 274, 6992–7001. <https://doi.org/10.1074/jbc.274.11.6992>
- 39** Bienz K, Egger D, Troxler M, Pasamontes L. Structural organization of poliovirus RNA replication is mediated by viral proteins of the P2 genomic region. *J. Virol.* **1990**, 64, 1156–1163. <https://jvi.asm.org/content/64/3/1156.long>
- 40** Ikegami N, Eggers HJ, Tamm I, Math JC, Math JL, Ahlberg JH, et al. Rescue of Drug-Requiring and Drug-Inhibited Enteroviruses. *PNAS.* **1964**, 52, 1419–1426. <https://doi.org/10.1073/pnas.52.6.1419>
- 41** Shimizu H, Agoh M, Agoh Y, Yoshida H, Yoshii K, Yoneyama T, et al. Mutations in the 2C region of poliovirus responsible for altered sensitivity to benzimidazole derivatives. *J. Virol.* **2000**, 74, 4146–4154. <https://doi.org/10.1128/JVI.74.9.4146-4154.2000>
- 42** Crowther D, Melnick JL. Studies of the inhibitory action of guanidine on poliovirus multiplication in cell cultures. *Virology* **1961**, 15, 65–74. [https://doi.org/10.1016/0042-6822\(61\)90078-2](https://doi.org/10.1016/0042-6822(61)90078-2)
- 43** Pincus SE, Wimmer E. Production of Guanidine-Resistant and -Dependent Poliovirus Mutants from Cloned cDNA: Mutations in Polypeptide 2C Are Directly Responsible for Altered Guanidine Sensitivity. *J. Virol.* **1986**, 60, 793-796 <https://jvi.asm.org/content/60/2/793.long>
- 44** Zuo J, Quinn KK, Kye S, Cooper P, Damoiseaux R, Krogstad P. Fluoxetine Is a Potent Inhibitor of Coxsackievirus Replication. *Antimicrob Agents Chemother.* **2012**, 56, 4838–4844. <https://doi.org/10.1128/AAC.00983-12>
- 45** Ulferts R, Van Der Linden L, Thibaut HJ, Lanke KHW, Leyssen P, Coutard B, et al. Selective serotonin reuptake inhibitor fluoxetine inhibits replication of human enteroviruses B and D by targeting viral protein 2C. *Antimicrob. Agents Chemother.* **2013**, 57, 1952–1956. <https://doi.org/10.1128/AAC.02084-12>
- 46** Sadeghipour S, Bek EJ, McMinn PC. Selection and characterisation of guanidine-resistant mutants of human enterovirus 71. *Virus Res.* **2012**, 169, 72–79. <https://doi.org/10.1016/j.virusres.2012.07.005>
- 47** Ulferts R, de Boer SM, van der Linden L, Bauer L, Lyoo HR, Maté MJ, et al. Screening of a Library of FDA-Approved Drugs Identifies Several Enterovirus Replication

Inhibitors That Target Viral Protein 2C. *Antimicrob. Agents Chemother.* **2016**, 60, 2627–2638. <https://doi.org/10.1128/AAC.02182-15>

48 Zuo J, Kye S, Quinn KK, Cooper P, Damoiseaux R, Krogstad P. Discovery of structurally diverse small-molecule compounds with broad antiviral activity against enteroviruses. *Antimicrob. Agents Chemother.* **2016**, 60, 1615–1626. <https://doi.org/10.1128/AAC.02646-15>

49 Wenthur CJ. Classics in Chemical Neuroscience: Methylphenidate. *ACS Chem Neurosci.* **2016**, 7, 1030–1040. <https://doi.org/10.1021/acscchemneuro.6b00199>

50 Ulferts R, Van Der Linden L, Thibaut HJ, Lanke KHW, Leyssen P, Coutard B, et al. Selective serotonin reuptake inhibitor fluoxetine inhibits replication of human enteroviruses B and D by targeting viral protein 2C. *Antimicrob. Agents Chemother.* **2013**, 57, 1952–1956. <https://doi.org/10.1128/AAC.02084-12>

51 Ulferts R, De Boer SM, Van Der Linden L, Bauer L, Lyoo HR, Maté MJ, et al. Screening of a library of FDA-approved drugs identifies several enterovirus replication inhibitors that target viral protein 2C. *Antimicrob Agents Chemother.* **2016**, 60, 2627–2638. <https://doi.org/10.1128/AAC.02182-15>

52 Gofshteyn J, Cárdenas AM, Bearden D. Treatment of Chronic Enterovirus Encephalitis With Fluoxetine in a Patient With X-Linked Agammaglobulinemia. *Pediatr. Neurol.* **2016**, 64, 94–98. <https://doi.org/10.1016/j.pediatrneurol.2016.06.014>

53 Alderman CP, Moritz CK, Ben-Tovim DI. Abnormal Platelet Aggregation Associated with Fluoxetine Therapy. *Ann. Pharmacother.* **1992**, 26, 1517–1579. <https://doi.org/10.1177/106002809202601205>

54 Dalton SO, Johansen C, Mellekjær L, Sørensen HT, Nørgård B, Olsen JH. Use of Selective Serotonin Reuptake Inhibitors and Risk of Upper Gastrointestinal Tract Bleeding. *Arch. Intern. Med.* **2003**, 163, 59–64. <https://doi.org/10.1001/archinte.163.1.59>

55 Labos C, Dasgupta K, Nedjar H, Turecki G, Rahme E. Risk of bleeding associated with combined use of selective serotonin reuptake inhibitors and antiplatelet therapy following acute myocardial infarction. *CMAJ* **2011**, 183, 1835–1843. <https://doi.org/10.1503/cmaj.100912>

- 56** Robertson DW, Krushinski J, Fuller RW, David Leander J. Absolute Configurations and Pharmacological Activities of the Optical Isomers of Fluoxetine, a Selective Serotonin-Uptake Inhibitor. *J. Med. Chem.* **1988**, 31, 1412-1417 <https://doi.org/10.1021/jm00402a027>
- 57** Velázquez-Campoy A, Ohtaka H, Nezami A, Muzammil S, Freire E. Isothermal Titration Calorimetry. *Current Protocols in Cell Biology* **2004**, 17, 1-24. <https://doi.org/10.1002/0471143030.cb1708s23>
- 58** Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **2015**, 10, 449–461. <https://doi.org/10.1517/17460441.2015.1032936>
- 59** van der Schaar HM, Leyssen P, Thibaut HJ, de Palma A, van der Linden L, Lanke KHW, et al. A Novel, Broad-Spectrum Inhibitor of Enterovirus Replication That Targets Host Cell Factor Phosphatidylinositol 4-Kinase III β . *Antimicrob. Agents Chemother.* **2013**, 57, 4971–4981. <https://doi.org/10.1128/AAC.01175-13>
- 60** Lepore SD, He Y. Use of sonication for the coupling of sterically hindered substrates in the phenolic Mitsunobu reaction. *J. Org. Chem.* **2003**, 68, 8261–8263. <https://doi.org/10.1021/jo0345751>
- 61** Prous JR, Serradell N, Flores R, Garcia-Delgado. WO2014096377A1.
- 62** Huiban M, Tredwell M, Mizuta S, Wan Z, Zhang X, Collier TL, et al. A broadly applicable [^{18}F]trifluoromethylation of aryl and heteroaryl iodides for PET imaging. *Nat. Chem.* **2013**, 5, 941-944 <https://doi.org/10.1038/nchem.1756>
- 63** Mitsunobu O use of DA and T in S and T of NP. The use of Diethyl Azodicarboxylate and Triphenylphosphine in Synthesis and Transformation of Natural Products. *Synthesis* **1981**, 1, 1–28. <https://doi.org/10.1016/B978-0-08-023969-9.50025-1>
- 64** Approach M, Education D. Facile Synthesis of Guanidine Functionalised Building Blocks. *Asian J. Org. Chem.* **2011**, 9, 72–76. <https://doi.org/10.1002/ajoc.201402242>
- 65** Chin W, Zhong G, Pu Q, Yang C, Lou W, De Sessions PF, et al. A macromolecular approach to eradicate multidrug resistant bacterial infections while mitigating drug resistance onset. *Nat. Commun.* **2018**, 9, 1-14. <https://doi.org/10.1038/s41467-018-03325-6>

- 66** Miel H, Rault S. Total Deprotection of *N,N'*-bis(tert-butoxycarbonyl)Guanidines Using SnCl₄. *Tetrahedron Lett.* **1997**, 38, 7865–7866. [https://doi.org/10.1016/S0040-4039\(97\)10148-4](https://doi.org/10.1016/S0040-4039(97)10148-4)
- 67** Frank R, Schutkowski M. Extremely mild reagent for Boc deprotection applicable to the synthesis of peptides with thioamide linkages. *Chem. Commun.* **1996**, 2509–2510. <https://doi.org/10.1039/CC9960002509>
- 68** Krogstad P, Zuo J, Kye SH, Damoiseaux RD. US20150164910A1.
- 69** Enemark EJ, Joshua-Tor L. Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* **2006**, 442, 270–275. <https://doi.org/10.1038/nature04943>
- 70** Singleton MR, Dillingham MS, Wigley DB. Structure and Mechanism of Helicases and Nucleic Acid Translocases. *Annu. Rev. Biochem.* **2007**, 76, 23–50. <https://doi.org/10.1146/annurev.biochem.76.052305.115300>
- 71** Mccammon JA, Miao Y, Amaro RE, Baudry J, Chodera J, Smith JC. Ensemble Docking in Drug Discovery. *Biophys. Perspect.* **2018**, 114, 1–8. <https://doi.org/10.1016/j.bpj.2018.02.038>
- 72** Swift R V, Jusoh SA, Li ES, Amaro RE. Knowledge-Based Methods To Train and Optimize Virtual Screening Ensembles. *J. Chem. Inf. Model.* **2016**, 56, 830–842. <https://doi.org/10.1021/acs.jcim.5b00684>
- 73** Amaro RE, Baron AR, Mccammon JA. An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. *J. Comput. Aided Mol. Des.* **2008**, 22, 693–705. <https://doi.org/10.1007/s10822-007-9159-2>
- 74** Peng J-H, Wang W, Yu Y-Q, Gu H-L, Huang X. Clustering Algorithms to Analyze Molecular Dynamics Simulation Trajectories for Complex Chemical and Biological Systems *J. Chem. Phys.* **2018**, 31, 404-420 <https://doi.org/10.1063/1674-0068/31/cjcp1806147>
- 75** Ester M, Kriegel H-P, Sander J, Xu X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *KDD-96*, **1996**, 226-231 <https://doi.org/10.1109/ICSMC.2006.384769>
- 76** Bowers KJ, Sacerdoti FD, Salmon JK, Shan Y, Shaw DE, Chow E, et al. Scalable algorithms for molecular dynamics simulations on commodity clusters. *Proceedings of the*

2006 *ACM/IEEE conference on Supercomputing - SC '06*, **2006**, p. 84.
<https://doi.org/10.1109/SC.2006.54>

77 Richard A. Friesner, Robert B. Murphy, Matthew P. Repasky, Leah L. Frye, Jeremy R. Greenwood, Thomas A. Halgren, et al. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein–Ligand Complexes. *J. Med. Chem.* **2006**, 49, 6177-61796. <https://doi.org/10.1021/jm051256o>

78 Thomas A. Halgren, Robert B. Murphy, Richard A. Friesner, Hege S. Beard, Leah L. Frye, W. Thomas Pollard and, et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, 47, 1750-1759 <https://doi.org/10.1021/jm030644s>

79 Richard A. Friesner, Jay L. Banks, Robert B. Murphy, Thomas A. Halgren, Jasna J. Klicic, Daniel T. Mainz, et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, 47, 1739-1749 <https://doi.org/10.1021/jm0306430>

80 Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, 14, 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5)

81 Kluyver T, Ragan-kelley B, Pérez F, Granger B, Bussonnier M, Frederic J, et al. Jupyter Notebooks—a publishing format for reproducible computational workflows. *Position Power Acad. Publ. Play Agents Agendas* **2016**, 87–90. <https://doi.org/10.3233/978-1-61499-649-1-87>

82 Schrödinger Release 2014-4: Maestro, version 10.0, Schrödinger, LLC, New York, NY, 2014.

Appendix

RMSD analysis of the MD simulations of the fluoxetine enantiomers in site A and site B:

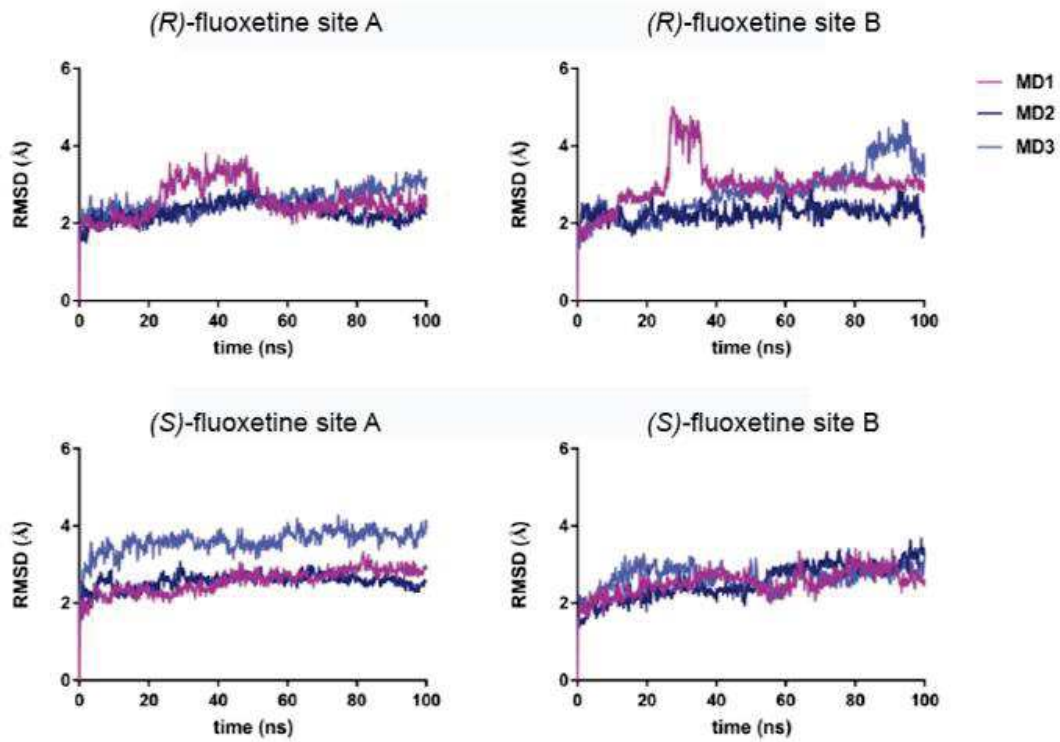


Figure 43: RMSD of the simulations

RMSD analysis of the MD simulations of 2C dimer:

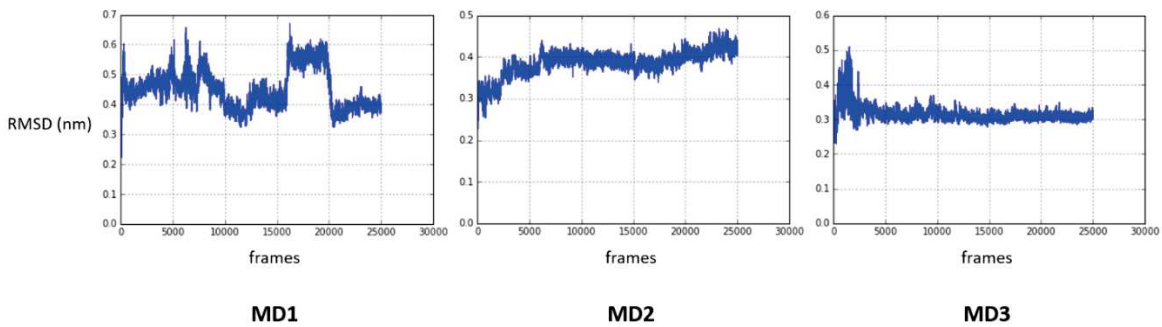


Figure 44: RMSD of the simulations

The compounds were purchased from SPECS

Compound number	Chemical structure	Compound number	Chemical structure
292		301	
293		302	
294		303	
295		304	
296		305	
297		306	
298		307	
299		308	
300			

Chapter 13: Conclusions

During this project, different molecular modelling techniques were applied, to identify new potential antiviral compounds able to inhibit RSV and enterovirus replication.

Two virtual screenings studies were performed on a reported druggable pocket of the RSV N protein, among the selected compounds, four hits were able to inhibit the RSV replication in the cell-based assay. Two of the four hits, the 2-phenylquinoline-4-carboxylic acid (**1**) and 2-(2-(*N*-phenyl-*N*-mesyl)acetamido)benzoic acid (**2**), were further investigated. Both hits were synthesised, but the antiviral activity was confirmed only for compound **1**.

Several derivatives were synthesised to explore the importance of the different chemical features of the 2-phenylquinoline-4-carboxylic acid. Among the synthesised compounds, few compounds showed antiviral activity against both RSV strains in the range of 3-11 μ M for RSV-A2 and in a range of 11-23 μ M for RSV-B. For the second hit compound, the antiviral activity was not confirmed. Among the small number of analogues that were synthesised the methyl ester derivative of compound **2** showed a comparable antiviral activity on both RSV stains to compound **1**.

For the F protein of RSV, several chemoinformatic techniques were used to generate and evaluate a focussed virtual library of α -helix mimic compounds for the inhibition of the rearrangement of the F protein during the membrane fusion process. One compound was selected and synthesis through the development of a highly versatile synthetic route.

The selected α -helix mimic compound was synthesised, together with a series of analogues. Among the synthesised compounds, only the selected compound showed a weak antiviral activity against both RSV stains.

For the enterovirus project on the non-structural 2C protein, the already reported antiviral activity of the SSRI fluoxetine was further investigated. The enantiospecific antiviral activity of (*S*)-fluoxetine on CVB3 was discovered, and TSA and ITC assays confirmed the enantiospecific binding to the 2C protein. Starting from the information acquired on the (*S*)-fluoxetine antiviral activity, molecular modelling techniques were used to elucidate the binding mode of fluoxetine on 2C protein. Two potential pockets near the ²²⁴AGSINA²²⁹ loop were explored with MD simulations, and the results obtained suggested that only site A was able to make a stable interaction with the *S*-enantiomer, but unfortunately, this was not confirmed by mutations study. Different fragments and analogues were synthesised and biologically evaluated to explore the chemical features of fluoxetine.

Two novel guanidine analogues of fluoxetine, 1-(3-phenyl-3-(4-(trifluoromethyl)phenoxy)propyl)guanidine **252** and 1-(2-phenyl-2-(4-(trifluoromethyl)phenoxy)ethyl)guanidine **256**, were designed and synthesised, starting from the information available on the antiviral activity of guanidine, and the stereospecific antiviral activity of fluoxetine.

Both compounds were found to be active on CVB3 and to have the same trend of antiviral activity against other enteroviruses. Compound **252**, which was synthesised as a racemic mixture, was found to be equipotent against CVB3 in comparison with the (*S*)-fluoxetine.

The scaffold of the *N*-(4-fluorobenzyl)-*N*-(4-methoxyphenyl)furan-2-carboxamide **262**, which was reported to inhibit the viral replication of CVB3 by targeting the non-structural protein 2C, was chosen to avoid the chiral centre on fluoxetine for further development.

Different analogues of compound **262** were designed to introduce the relevant features of fluoxetine for the antiviral activity into compound **262** and to explore the importance of the furamide group for the antiviral activity.

Interestingly, the replacement of the fluoro group on the benzylic ring of compound **262** by the trifluoromethyl group, was associated with pan-inhibitor antiviral activity, suggesting that the scaffold of **262** is suitable for the development of a pan-enterovirus inhibitor of the viral replication.

In the last part of the enterovirus project, MD simulations were performed to address the flexibility related to the ATPase pocket, and an ensemble docking virtual screening was performed. Among the selected compounds, one hit was found activity against CVB3 and EV71.