UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Università degli Studi di Padova

Dipartimento di Scienze Chimiche

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE MOLECOLARI

INDIRIZZO SCIENZE FARMACEUTICHE

CICLO XXVIII

DESIGN, SYNTHESIS AND BIOCHEMICAL CHARACTERIZATION OF FINGOLIMOD ANALOGS FOR TARGETING PP2A

Direttore della Scuola: Ch.mo Prof. Antonino Polimeno

Coordinatore d'indirizzo: Ch.mo Prof. Alessandro Dolmella

Supervisore: Ch.mo Prof. Giuseppe Zagotto

Dottoranda: Valeria Pavan

People talking without speaking
People hearing without listening
(cit.)

CONTENTS

ABSTRACT (ITALIANO)

PP2A (Protein Fosfatasi 2A) è una serin/treonin fosfatasi espressa in maniera ubiquitaria e altamente conservata tra le specie (eucarioti *in primis*). Essa è coinvolta in delicati *pathway* di segnalazione cellulare, ed ha come effetto complessivo quello di promuovere eventi pro-apoptotici ed anti-proliferativi; una sua diminuita attività comporta un aumento di segnali anti-apoptotici e proliferativi. La sua attività, come quella di ogni fosfatasi, è controbilanciata dalle protein chinasi; mentre queste ultime sono state ampiamente indagate nel corso degli anni, il ruolo di PP2A è stato preso in considerazione solo molto più di recente, e insieme ad esso è stata esplorata la possibilità di considerare l'enzima un *target* per farmaci. Infatti, è stato dimostrato che l'attività di PP2A è ridotta in svariate forme tumorali, in particolare leucemie e ancor più nel dettaglio nei casi di B-CLL (leucemia linfatica cronica a cellule B). Al momento, la principale ipotesi per spiegare questo fenomeno è che PP2A sia inattivata non tanto in maniera diretta, quanto piuttosto per aumentata espressione della proteina SET, suo inibitore endogeno. In questi casi, la capacità della cellula di entrare in apoptosi è fortemente compromessa, quindi la possibilità di ripristino dell'attività di PP2A è attualmente presa in considerazione come nuova, possibile terapia anti tumorale.

Fingolimod (nome IUPAC 2-ammino-2-[2-(4-ottilfenil)etil]propan-1,3-diolo) è stato approvato nel 2010 per il trattamento della sclerosi multipla in forma recidivante-remitente. La sua attività è legata al fatto che, una volta internalizzato dalla cellula, esso viene fosforilato e diventa strutturalmente simile alla sfingosina-1-fosfato (S1P), mimandone l'azione sui linfociti T: sovrastimolando i recettori per S1P, infatti, Fingolimod può portare alla loro degradazione, evitando che i linfociti T possano uscire dai linfonodi e perpetrare il danneggiamento degli assoni neuronali. Più di recente, è stato dimostrato che Fingolimod può, come effetto collaterale, ristabilire nei tumori l'attività di PP2A; il meccanismo sottostante a questa attività non è stato ancora del tutto chiarito, ma sembra legato all'interazione tra Fingolimod e il complesso PP2A/SET. Curiosamente, questa funzione richiede che Fingolimod non sia fosforilato.

Questo lavoro di tesi presenta la sintesi e la caratterizzazione di una libreria di analoghi di Fingolimod che vogliono isolare l'effetto pro-apoptotico da quello immunosoppressivo. Tenendo a mente che Fingolimod agisce su PP2A solo se non fosforilato, ciò significa progettare e creare molecole con caratteristiche strutturali modulate rispetto al composto preso a modello, ma con un'attenzione particolare alla loro

capacità di subire fosforilazione. Questo approccio ha lo scopo di ottenere nuove, potenziali *drug entities* anti-cancro che possano ripristinare la capacità delle cellule interessate di subire apoptosi. I composti sono stati testati nel laboratorio della Prof. Brunati (Dipartimento di Medicina Molecolare, Università di Padova) per valutare la loro attività verso PP2A in cellule di B-CLL. In questo senso, sono stati ottenuti alcuni composti meritevoli di interesse e vengono qui riportati i risultati dei saggi biologici sul composto più attivo.

Inoltre, è stato perseguito lo scopo di produrre le proteine PP2A (subunità catalitica) e SET tramite tecnica del DNA ricombinante al fine di ottenere dettagli strutturali sull'interazione tra proteine, e sulla eventuale modalità di legame tra queste e i composti sintetizzati; questa parte del progetto è stata svolta presso Diamond Light Source (UK). Il lavoro svolto getta le basi per un approfondito studio strutturale dell'interazione tra PP2A e SET (che, ad oggi, è piuttosto lacunosa), sia alla razionalizzazione delle relazioni struttura-attività tra le proteine di interesse e composti anti-neoplastici; questo potrà aiutare a chiarire le caratteristiche strutturali di cui necessitano le molecole per essere attive, e di conseguenza a progettare in modo razionalizzato analoghi di Fingolimod nuovi e migliorati.

Bibliografia

Bononi, A. et al. Protein Kinases and Phosphatases in the Control of Cell Fate. *Enzyme Res.* **2011**, 1–26 (2011).

Ciccone, M., Calin, G. A. & Perrotti, D. From the biology of PP2A to the PADs for therapy of hematologic malignancies. *Front. Oncol*. **5**, 1–10 (2015).

Brinkmann, V., Billich, A., Baumruker, T. & Heining, P. Fingolimod (FTY720): discovery and development of an oral drug to treat multiple sclerosis. *Nat. Publ. Gr.* **9**, 883–897 (2010).

Perrotti, D. & Neviani, P. Protein phosphatase 2A : a target for anticancer therapy. *Lancet Oncol.* **14**, e229–e238 (2013).

ABSTRACT (ENGLISH)

PP2A (Protein Phosphatase 2A) is a Ser/Thr phosphatase ubiquitously expressed and highly conserved among the species (especially eukariotes). It is involved in delicate signaling pathways, and overall it has the effect to enhance pro-apoptotic and anti-proliferative events; a decrease of its activity makes anti-apoptotic and proliferative signals prevail. Its activity, as the activity of any other phosphatase, is counterbalanced by protein kinases; while the latter have been widely studied through the years, the role of PP2A was taken into account much more recently, and together with that the possibility to consider it as a drug target has been exploited. Indeed, PP2A is demonstrated to be down-regulated in several forms of tumors, in particular leukemias with a peculiar focus on B-CLL (B cell chronic lymphocytic leukemia). To date, the main hypothesis to explain this is that PP2A is inactivated not directly, but because of the over-expression of the protein SET, its endogenous inhibitor. In these cases, the capability of the cell to die is strongly impaired, so trying to restore PP2A activity is nowadays being exploited as a new, possible anti-cancer therapy.

Fingolimod (IUPAC name 2-amino-2-[2-(4-octylphenyl)ethyl]propan-1,3-diol) has been approved in 2010 for the treatment of relapsing-remitting forms of multiple sclerosis. Its activity is due to the fact that, once internalized by the cell, it gets phosphorylated and reaches a structural similarity with sphingosine-1-phosphate (S1P), mimicking it on T cells: by overstimulating S1P receptors, indeed, Fingolimod can lead to their degradation, avoiding the possibility for T cells to egress secondary lymphoid organs and to keep on damaging neural axons. In more recent years, it has been shown that Fingolimod can restore PP2A activity in tumors as a side effect; the mechanism underlying this event is still not clear, but it is probably due to the interaction Fingolimod has with the PP2A/SET complex. Curiously, this activity needs Fingolimod not to be phosphorylated.

This thesis work presents the synthesis and characterization of a library of Fingolimod analogs which are meant to isolate the pro-apoptotic effect from the immunosuppressive one. Keeping in mind that Fingolimod acts on PP2A only if unphosphorylated, this means designing and synthesizing molecules with modulated structural characteristics, but with a peculiar focus on their phosphorylatability. This approach is meant to obtain new potential anti-cancer drug entities that could restore the capability of the involved cells to die for apoptosis. The compounds were assayed in the laboratory of Prof. Brunati

(Department of Molecular Medicine, University of Padova) for their capability to restore PP2A activity in B-CLL cells. In this sense, some promising compounds were obtained and the results of the biochemical assays on the most performing compound have been here reported.

In addition to this, the aim of producing PP2A (catalytic subunit) and SET proteins through recombinant DNA technique was pursued in order to obtain more details about the structural features involved in the interactions between the proteins themselves, and also the eventual ones between the proteins and the synthesized compounds; this part of the project was performed at Diamond Light Source (UK). Combining all the pieces of information obtained from these aims will lead both to an enhanced knowledge about the structural interactions between PP2A and SET (which to date is lacking) and to the possibility of rationalize eventual structure-activity relationships between the proteins of interest and anti-cancer compounds; this will help in elucidating the structural characteristics which are needed for the molecules to be active, and as a consequence in designing more, improved, rationalized Fingolimod analogs.

References

Bononi, A. et al. Protein Kinases and Phosphatases in the Control of Cell Fate. *Enzyme Res.* **2011**, 1–26 (2011).

Ciccone, M., Calin, G. A. & Perrotti, D. From the biology of PP2A to the PADs for therapy of hematologic malignancies. *Front. Oncol*. **5**, 1–10 (2015).

Brinkmann, V., Billich, A., Baumruker, T. & Heining, P. Fingolimod (FTY720): discovery and development of an oral drug to treat multiple sclerosis. *Nat. Publ. Gr.* **9**, 883–897 (2010).

Perrotti, D. & Neviani, P. Protein phosphatase 2A : a target for anticancer therapy. *Lancet Oncol.* **14**, e229–e238 (2013).

1. INTRODUCTION

1.1 CELLULAR SIGNALING

Signaling is a complex and delicate *ensemble* of processes within the cell and among cells; the robustness of the signaling pathway performances and their repeatability are mandatory in order to elicit the appropriate physiological response to given *stimuli*. Cells use signaling pathways for obtaining a series of effects such as modulation of enzyme activity, regulation of proteolysis, assembly of macro-molecular machines, intracellular localization, creation of protein docking sites, cell division, proliferation, differentiation and apoptosis, in order to modulate all the aspects of cell life. Because of this, each signaling pathway needs to be submitted to a fast and fine regulation through proper effectors and a proper control on the effectors themselves[1].

Effectors are often represented by molecules subject to phosphorylation and dephosphorylation events which control their activation or inactivation. Reversible phosphorylation represents for a cell a fast and easy way to transmit and amplify signals of various types; because of this, it is easy to understand the fact that it plays key roles in several physiological processes. Phosphorylation state unbalances can be dangerous because they can affect the normal life of the cell in many ways, and indeed they have often been found in pathological conditions (among the others, hematological malignancies)[1,2].

It is well known that, while phosphorylation of nucleophilic aminoacid side-chains in protein is catalyzed by protein kinases (PKs), dephosphorylation is controlled by protein phosphatases (PPs). The role of PKs has been widely investigated, and it has been proved that the dysregulation of the relative oncogenes is responsible for cancer initiation and maintenance[2,3]; by counterbalancing the activity of kinases, phosphatases started in recent years to be considered as tumor suppressors[3,4] as demonstrated by the fact that the loss of function of various phosphatases has been detected in many cancers[5].

The human genome encodes for 518 protein-kinases; among them, 428 are Ser/Thr kinases and 90 are Tyr kinases. Actually, PPs are not so abundant: the human genome encodes for 107 Tyr phosphatases and only 40 Ser/Thr phosphatases, for a total amount which is around one fifth of total one of the kinases. Apart from this discrepancy, Ser and Thr account together for 98% of the phosphoprotein targets (86% and 12% respectively), but the total amount of PKs and PPs specific for these residues is much lower than

the number of Tyr specific kinases and phosphatases[6].

Probably due to their abundance, PKs have been much studied through the years, and now they represent one of the main drug targets for pharmaceutical industry; imatinib mesylate, the first commercial Bcr-Abl1 Tyr kinase inhibitor developed for the treatment of chronic myeloid leukemia (CML), was commercialized starting from 1998[7]. The appreciation of the potential role of PPs was made only much later and only a few phosphatases were characterized until recent years, because of their small amount (at least when compared to kinases) and of the objective limitations of the available tools for targeting – and, as a consequence, analyzing - them (above all, the lack of specificity of PPs inhibitors)[8].

1.1.1 PROTEIN PHOSPHATASES

Several classifications have been made for PPs through the years, but the most recent one divides these phosphoproteins in three main classes basing on sequence, structure and catalytic mechanism[8]:

- Ser/Thr phosphatases, which are divided between
    - Phosphoprotein phosphatases or PPPs (PP1, PP2A, PP2B, PP4, PP5, PP6 and PP7)
    - Mg^{2+} or Mn^{2+} dependent phosphatases or PPMs (PP2C).

  Most of the Ser/Thr phosphatase *in vivo* activity is performed by PP1, PP2A, PP2B and PP2C. The heterogeneity of this family has to be addressed to the way the subunits assemble, since these phosphatases are usually composed of hetero-oligomeric complexes made up by relatively simple subunits, that can interact with a number of regulators affecting targets and activity[6];
- Tyr phosphatases (PTPs), which are divided in classes I, II and III. PTPs are multidomain proteins presenting the peculiar feature of having the catalytic and regulatory domains combined in one single chain. They are responsible for a variety of signals: cell adhesion, insulin signaling, immune response, cell cycle control[9]. Among them, a relevant role is played by SHP1 (Src-homology 1 domain phosphatase), a PTP which is shown to dephosphorylate Bcr–Abl oncoprotein: its activity, also due to the downstream signals that involve other phosphatases and kinases as PP2A and Lyn, leads to the production of pro-apoptotic factors and because of this it is often down-regulated in imatinib-resistant CML forms[10];
- Asp-based protein phosphatases, less abundant and less known; they use a peculiar mechanism for the dephosphorylation which involves an Asp residue, and comprise
    - Dual specificity phosphatases (DUSPs), that can act on all of the three mentioned aa residues
    - FCP (TFIIF-associating C-terminal domain phosphatase)/SCP (Small C-terminal domain phosphatase) family
    - HAD (haloacid dehalogenase) family.

To date, the Asp-based PP category is the less investigated and known one, and also the classification here reported is not univocally recognized (*e.g.*, some authors consider DUSPs to be belonging to PTPs)[3,6,8,11].

In recent years, because of its role and its cellular abundance, one of the most pharmaceutically promising phosphatases has been considered to be PP2A (Protein Phosphatase 2A).

1.2 PROTEIN PHOSPHATASE 2A

Protein Phosphatase 2A (PP2A) is a term that comprises a family of Ser/Thr phosphatases, which are ubiquitously expressed and highly conserved among the species (mostly eukariotes)[1]; this family makes up 1% of all cellular proteins and together with PP1 realizes over 90% of all Ser/Thr phosphatase activity in the cell[12]. PP2A is involved in many cellular functions: cell metabolism, cell cycle, DNA replication, transcription and translation, cell proliferation, cell mobility and, above all the others, apoptosis[13]. Its importance in oncogenesis was shown first when okadaic acid (OA), an already known carcinogen, was demonstrated to be able to selectively inhibit PP2A. This evidence was followed by others, for example that PP2A can also be inhibited by the oncogenic simian vacuolating virus 40 (SV40) small T-antigen (which probably displaces one of the subunits from the complex PP2A is formed by). Furthermore, different components of the PP2A holoenzyme were shown to be mutated in various tumor forms, *e.g.* the already cited CML, acute myeloid leukemia (AML) or lung cancer; the restoration of a normal PP2A activity induces instead cell death in cancer cells (both *in vitro* and *in vivo*) and counteracts uncontrolled proliferation[12,14].

1.2.1 PROTEIN PHOSPHATASE 2A: STRUCTURE

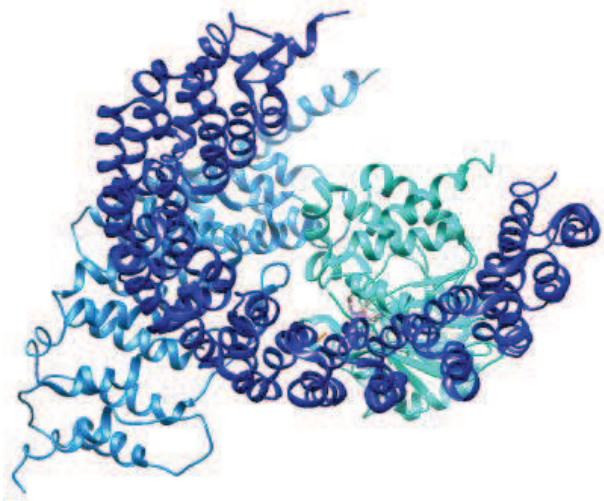


Figure 1. PP2A structure; modified from PDB ID: 2NYM.

Subunit	Gene	Isoform	Other name
Catalytic (C)	PPP2CA	α	PPP2Cα
Catalytic (C)	PPP2CB	β	PPP2Acβ
Structural (A)	PPP2R1A	α	PR65α, PP2A-Aα
Structural (A)	PPP2R1B	β	PR65β, PP2A-Aβ
Regulatory (B)	PPP2R2A	α	B55α, PR55α, PP2ABα
Regulatory (B)	PPP2R2B	β	B55β, PR55β, PP2ABβ
Regulatory (B)	PPP2R2C	γ	B55γ, PR55γ, PP2ABγ
Regulatory (B)	PPP2R2D	δ	B55δ, PR55δ, PP2ABδ
Regulatory (B')	PPP2R5A	α	B56α, PR56/61α, PP2AB'α
Regulatory (B')	PPP2R5B	β	B56β, PR56/61β, PP2AB'β
Regulatory (B')	PPP2R5C	γ1,2,3	B56γ, PR56/61γ, PP2AB'γ
Regulatory (B')	PPP2R5D	δ	B56δ, PR56/61δ, PP2AB'δ
Regulatory (B')	PPP2R5E	ε	B56ε, PR56/61ε, PP2AB'ε
Regulatory (B'')	PPP2R3A	α	PR130, B''α1
Regulatory (B'')	PPP2R3A	α	PR72, B''α2
Regulatory (B'')	PPP2R3B	β	PR70, PR48, B''β
Regulatory (B'')	PPP2R3C	γ	G5PR, G4-1
Regulatory (B'')	PPP2R3D	δ	PR59, B''δ
Regulatory (B''')	STRN		Striatin, PR110
Regulatory (B''')	STRN3		SGN2A
Regulatory (B''')	STRN4		
Regulatory (B''')	PPP2R4		PTPA, PRS3

Table 1.Possible isoforms of PP2A subunits[15].

Active PP2A holoenzyme is composed of three main subunits, which can all exist in various isoforms (as shown in the previous table):

- $PP2A_C$, the catalytic subunit (coloured with bright azure in figure 1)
- $PP2A_A$, the structural subunit (coloured with dark blue in figure 1)
- $PP2A_B$, the regulatory subunit (coloured with light blue in figure 1).

The core of the enzyme ($PP2A_D$) is constituted by $PP2A_C$ and $PP2A_A$, and can exist as it is in a dimeric form (one third of the total cellular PP2A) or can be associated to one of the types of B subunits. The numerous types of genes and isoforms which are responsible for PP2A subunits, together with their genetic redundancy (at least two genes encoding for each subunit), lead to a large heterogeneity of PP2A holoenzymes[16]. Taking into account all the possible dimers and trimers, nearly 100 PP2A holoenzymes can be found in the cell.[17]

C subunit

The catalytic subunit of PP2A, also called $PP2A_C$, hosts the dephopshorylative activity of the whole enzyme. $PP2A_C$ has a MW of 36 kDa and it is encoded by two genes (both made up by seven exons and six introns) that produce two isoforms, Cα and Cβ, with 97% identity in their primary sequence[13,18]. This high similarity, however, does not lead to redundancy in the roles of the two isoforms; it has been shown that loss in $PP2A_C\alpha$ activity cannot be replaced by $PP2A_C\beta$ in yeast and mice, and this loss leads to cell death[12,17]. These two isoforms are ubiquitously expressed (with peaks in brain and the heart), being Cα predominant in the plasma membrane and Cβ in the cytoplasm and nucleus; their expression levels are different, since Cα is 10 times more abundant than Cβ. This is probably due to some differences which involve the activities of the promoters and the turnover of the corresponding mRNAs[18].

Molecular cloning and overexpression of $PP2A_C$ has been achieved in lower organisms as insects thanks to baculovirus technology, and here it showed to have had its structure conserved through the evolution, being maybe the most conserved among the known enzymes; unluckily, until now it was not possible to clone it in mammal cells[18].

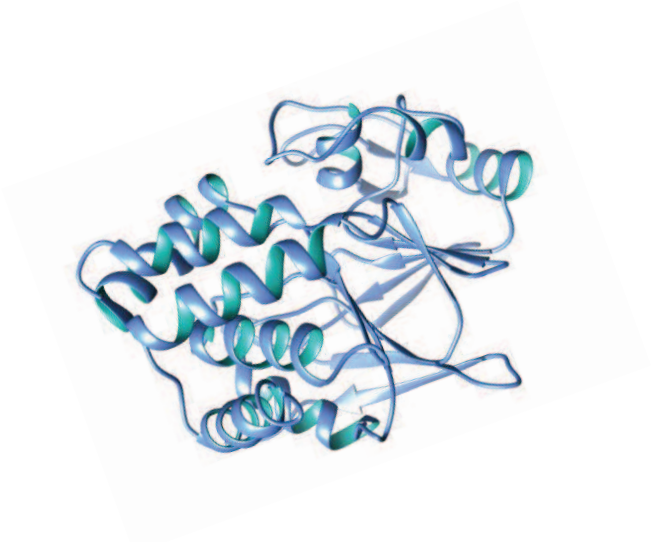


Figure 2. C subunit; isolated from PDB ID 2NYM.

Apart from the fundamental catalytic role it has, this subunit also shows another peculiarity since it results to be the mainly involved one in enzyme regulation; so, in $PP2A_C$ reside both the dephosphorylative activity and the regulation of the activity itself. It has been demonstrated that this subunit can be subject to post-translational modifications as well as to the modulation from a series of interactors, which can either promote or inactivate the enzyme itself[19]. This point will be better elucidated in the "Protein phosphatase 2A: regulation" section.

A subunit

A subunit, also known as PR65 or structural subunit, is the structural docking platform for subunits B and C; in the absence of a B subunit, it can alter the substrate specificity of $PP2A_C$. The structure of $PP2A_A$ is made by 15 tandem repeats of a 39 aa sequence, termed HEAT (huntingtin/elongation/A subunit/TOR, where TOR is target of rapamycin) motif[18]. The reported crystal structure of PR65 (PDB ID: 1B3U) shows that each repeat is virtually the same, *i.e.* two superimposed α-helices stacked in a stable, elongated unit[20].

Figure 3. Crystal structure of A subunit. PDB ID: 1B3U.

As already seen for $PP2A_C$, also $PP2A_A$ is encoded by two genes which lead to the production of two protein isoforms, Aα and Aβ, which share around 87% of sequence identity and the MW of 65 kDa. Once again, however, this does not mean that they are interchangeable: it has been shown that loss of Aα activity in mice cannot be replaced by Aβ. In addition to this, 90% of PP2A holoenzymes contain the Aα isoform and 10% the Aβ, so it is reasonable to think that they do have different roles in the cell; it has been discovered that Aβ has an higher abundance in testis when compared to other tissues, suggesting a role for Aβ in spermatogenesis[15,21].

B subunit

The group of B subunits is the most heterogeneous, and the scaffolds belonging to it are coded by genes (15 were identified until now) that totally lack structural similarity[12,18]. The presence of this subunit results to be of extreme importance, since the subtype of B subunit is related both to the tissue it is expressed in, and to the substrate specificity it is meant to mediate; this distinction lead to the classification of four distinct families, which share almost no structural similarity[12].

B family (PR55)

This first family comprises four distinct subunits (called α, β, γ, δ) with a MW of 55 kDa each; while α and δ are present in all tissues in comparable amounts, β and γ are highly enriched in brain. One of the

most interesting features of this family is the presence of five degenerate Trp/Asp-40 (or WD-40) repeats, *i.e.* sequences composed by 40 aa (conserved just in a minimal part) and ending with Trp-Asp. These portions are thought to mediate the protein-protein interaction, for example the one with TGF- β receptors[18].

B' family (PR61)

These subunits have a MW of 61 kDa and are encoded by five genes which, thanks to their splicing variants, lead to the production of seven subunits. They show to have a peculiar and precise distribution in the cell: α, β, ε are mostly found in the cytoplasm and γ1, γ2 and γ3 are mostly found in the nucleus, while δ is expressed in both the nucleus and the cytoplasm[12]. All the PR61s have similar structure in the central part (80% of sequence identity), while they are different at N- and C- termini; this leads to different expression levels in tissues[13]. It has been shown that this type of subunits exposes a highly acidic surface, which is believed to be responsible of further protein recruitment[12]; the structure is mostly α helical and it is a potential substrate for phosphorylation[13].

B'' family (PR72)

The pieces of information about this family are not really abundant and also the number of members is uncertain: some authors claim it exists in five forms, while others believe they are four[18]. PR72 is found exclusively in heart and skeletal muscle, PR130 is ubiquitous (with highest levels in heart and muscle), PR59 in testis, kidney, liver, brain, heart and lung; PR48 and G5PR are still matter of discussion[13,18].

B''' family

Like the previous one, also this family is not extremely well known. These subunits have been shown to share a conserved structure which is the same as B' subunits, and also the WD-40 repeats which are typical for B family. They are supposed to be somehow involved in Ca^{2+}-dependent signaling[18].

1.2.2 PROTEIN PHOSPHATASE 2A: ROLE

Being a phosphate, PP2A has of course the main role of dephosphorylating substrates; however, it is worth considering the very final effects of this dephosphorylative activity. Since there are so many possible subunit combinations, it is reasonable to imagine that PP2A has an impressive number of functions: so it is, and they mainly depend on the specific assembly and localization of the subunit themselves, and of course on the specific substrates. Because of its profusion of roles, PP2A has been defined as a "social butterfly"[22]. Overall, it can mainly affect signal transduction, with influence in cell proliferation, differentiation, adhesion, migration, metabolism and, most of all, survival[16,17].

The following image shows only some of the signaling pathways regulated by PP2A.

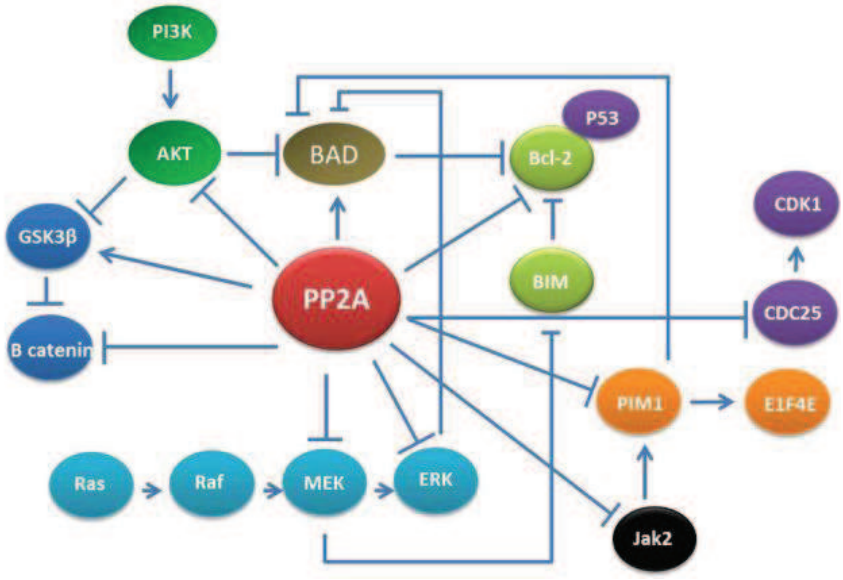


Figure 4. Biochemical pathways in which PP2A is involved[2].

PP2A and the cell cycle

As mentioned above, PP2A has a deep role concerning cell cycle: indeed, it can inhibit the phase transitions that characterize cell life. When duplicating, eukaryotic cells leave their latent phase (G_0) to go through a four stages cycle, tightly controlled by a number of checkpoint mechanisms:

- G_1 phase (gap 1), corresponding to the time gap between mitosis and initiation of DNA replication;
- S phase, where the S stands for synthesis: at this point DNA is replicated replicated to provide a copy of the genetic material to each of the daughter cells;
- G_2 phase (gap 2), where the cell grows and produces proteins in preparation of mitosis;
- M phase, and so mitosis (when the mother cell divides into two genetically and functionally equivalent cells), which is usually followed by cytokinesis[23].

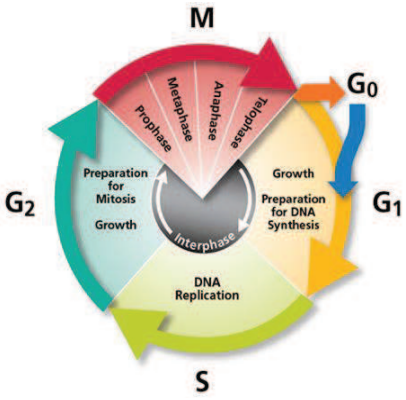


Figure 5. Schematic representation of the cell cycle[24].

Depending on its specific structural composition, PP2A is capable of controlling some of the phase transitions in the cell cycle, *i.e.* G_1-S and especially G_2-M[11]; the latter occurs thanks to the inhibition of some maturation-promoting factors (CDK1 and cyclin B) and it is more likely in the presence of a B55α-containing PP2A form. In the presence of B56δ, instead of B55α, PP2A catalyses the exit from mitosis thanks to the inactivation of CDC25 (Cell division cycle 25 phosphatase).

The activity of this enzyme is also exerted through the inactivation of MEK (Mitogen-activated protein kinase kinase) and ERK (Extracellular signal-regulated kinase): this leads to the decrease of the functionality of transcription factors as c-MYC and STAT5, leading to the inhibition of their mitogenic and antiapoptotic effects and to an overall negative regulation of the so-called MAPK (Mitogen activated protein kinase) signaling pathway [5].

PP2A and apoptosis

Apart from the effects on the inhibition of the cell cycle PP2A can act on many biochemical pathways, mainly exerting its effects on the promotion of the apoptosis. Apoptosis – or programmed cell death - is a crucial event for cellular turnover; a lacking or an excessive apoptotic signaling is a common hallmark of pathologies as neurodegenerative diseases, ischemia, autoimmune disorders and cancer. The capability of modulating cell life and death is an extremely desirable pharmaceutical aim because of its huge therapeutic potential[25].

The influence of PP2A on apoptotic pathways can be seen on many sides; one of them is the inhibitory effect it has on Jak2 (Janus kinase 2) Tyr kinase. PP2A and Jak2 have opposite effects, so the activation of PP2A induces inactivation of Jak2, resulting on the suppression of pro-survival signals[1].

β-catenin is a protein subunit belonging to the cadherin, and it is the central focus of Wnt (Wingless-type MMTV integration site)/β-catenin pathway. β-catenin activates growth and anti-apoptotic mechanisms thanks to the activation of c-MYC genes, and it has been shown that its levels are often increased in colorectal cancer. PP2A can inactivate this pathway and block these pro-survival signals, and also stem cells self renewal and embryonic development[5,14].

PP2A is also involved in one of the main apoptosis-modulating signaling cascades: the PI3K (phosphoinositide 3-kinase)/AKT (Protein kinase B) pathway. PP2A suppresses the signals coming from this pathway by directly dephosphorylating and inactivating AKT and by suppressing Akt-activation signals[2]. This is extremely important because AKT is a well-known kinase, involved in the regulation of cell survival and proliferation: it promotes key features as cell cycle progression, cell survival, growth, proliferation, angiogenesis, and self-renewal of stem cells[1]. The pro-survival activity of AKT is exploited in two main ways: the inactivation of pro-apoptotic signals and the parallel activation of anti-apoptotic signals. For example, part of its effects is due to the fact it inhibits the release of cytochrome C from mitochondria through phosphorylation (*i.e.* inactivation) of pro-apoptotic factor caspase-9[5]. The same pathway partly also involves the inactivation of Bcl-2 (B-cell lymphoma 2), an anti-apoptotic factor, and the activation of BAD (Bcl-2-associated death promoter) and Bim (Bcl-2-like protein 11). This is the result of the direct inactivation of AKT and the activation of BAD performed by PP2A which leads to the activation and translocation of BAD to the membrane of mitochondria, where it binds and inhibits Bcl-

2[26]. For being active, however, AKT needs to be phosphorylated; this can occur on a Tyr or a Ser residue. The activation leads to its translocation to various subcellular compartments, where it can phosphorylate its substrates and interfere with the activation of apoptotic signals. As a consequence, it is not surprising that it is considered to be a protooncoprotein and that aberrant Akt activity has been found in diseases as cancer or neurodegenerative disorders[1]. Having it dephosphorylated by PP2A means inactivating it, favouring cell death.

PP2A as a tumor suppressor

Many evidences support the idea that PP2A can be considered to be a tumor suppressor; the very first one was the demonstration it is inhibited by okadaic acid (OA), which is both a tumor promoter and a specific PP2A inhibitor[27]. However, the strongest evidence arose with the studies on eukaryotic cells, that showed how the alteration of balance in one or more of the subunits cause severe misregulation in normal cell life; in particular,

- overexpression of Aα subunit leads to the inhibition of cytokinesis; indeed, this event can lead to the disruption of the protein complexes that lead to cytokinesis and/or the sequestration of the catalytic subunit. As a result, it is possible to have inefficient dephosphorylation and activation of myosin light chain kinase, issue that leads at the very end to the inhibition of the contractile force in the cleavage furrow[28]. In addition to this, overexpression of Aα subunit can enhance the formation of bi- or multi-nucleated cells[28];

- truncation of PR61 leads to radio-resistance and metastasis; PP2A is in fact involved in the regulation of paxillin and Mdm2 (an inhibitor of p53 tumor suppressor). The dephosphorylation of the latter permits to Mdm2 itself to bind p53 and block its pro-transcription activity[29]. In addition to this, the decreased dephosphorylation of paxillin (a multidomain-containing adaptor molecule involved in cell migration), enhances cell motility promoting metastasis in case of cancer[30];

- mutations in the A subunit have been demonstrated to be present in a variety of human cancers, specially lung tumors, lung-derived tumors and colon tumors[31], while mutations of PPP2R1B gene are involved in the onset of almost all types of colorectal cancers[32];

- subversion of PP2A activity performed by SV40 ST (Small t antigen, a viral protein involved in the onset of cancer), which mainly targets B56γ-containing holoenzymes, contributes to malignant cell transformation[33]; in fact, suppressing the expression of the PP2A B56γ subunit in human embryonic kidney (HEK) epithelial cells expressing SV40 large T antigen, hTERT, and H-RAS lead to the inhibition of PP2A-specific phosphatase activity, to an anchorage-independent cell growth and to the formation of tumors. On the other hand, overexpression of PP2A B56γ3 in tumorigenic HEK cells expressing ST partially inverted the tendency towards tumorigenicity of these cells[33];

- deletion of PR55 in eukaryotes (mainly *Drosophila*) leads the activation of MAPK signaling, responsible for cell overgrowth, in a very short amount of time (5-15 min)[34]. The deletion of either A or C subunits provokes the disappearance of all PP2A subunits; ablation of total PP2A by using dsRNA against either A or C subunit enhances ERK activation, affecting as a consequence the MAPK cascade[34].

Taken together, all these pieces of information can lead to the reasonable statement that PP2A can be considered to be as a tumor suppressor[22]. As a direct consequence, a deep comprehension of the properties and the regulation of this enzyme could lead to a new approach in cancer treatment strategies.

1.2.3 PROTEIN PHOSPHATASE 2A: REGULATION

Given the important roles PP2A has, it is necessary to know how it can be regulated in order to understand where misregulations can occur and how to correct them. Because of the many possible isoform combinations, the modes of regulation of PP2A are quite heterogeneous, and can be divided mainly between post-translational modifications and interactors-mediated modulation.

Post-translational modifications

The main post-translational events that can affect PP2A activity are two: phosphorylation and methylation, both occurring on the catalytic subunit.

Plenty of kinases can phosphorylate $PP2A_C$, and this event is enhanced by the presence of OA[19]. Phosphorylation can occur on Tyr^{307} or Thr^{304} residues, and in both the cases this leads to the inactivation of the whole enzyme but in different ways: indeed, the phosphorylation of Thr^{304} gives enough hindrance for avoiding the recruitment of other subunits, while P-Tyr^{307} adds to this function the lack of possibility for the subunit to undergo methylation[5]. However, there is the chance for the enzyme to auto-dephosphorylate itself, thanks to its Tyr auto-dephosphorylative activity, and being restored[19].

At the C-terminal of $PP2A_C$, also resides the methylation substrate: it is represented by Leu^{309}, which can be methylated by a specific enzyme, called Leu-carboxyl methyltransferase of type I (LCMT1)[5,19]. So, in contrast to the previous case in which the post-translational modification could be performed by many kinases, here there is only one enzyme; its action can be reverted only trough a specific phosphatase methylesterase. The addition of a methyl residue to the Leu leads to the recruitment of B subunits in a preferential way than B' and B'', which instead prefer non-methylated $PP2A_C$[19].

Physiological modulators

Another level of PP2A regulation is represented by the interaction that can occur between the phosphatase and its endogenous inhibitors or enhancers.

Inhibitors

The inhibitor category comprises various macromolecules. Among them there is CIP2A (Cancerous inhibitor of PP2A), which prevents the PP2A-mediated c-Myc degradation; it is significantly up-regulated in some tumor cells, *e.g.* head, neck and breast cancer[35]. It is supposed to directly bind $PP2A_C$, but structural information is still lacking. Another one which is worthy to consider is ANP32A (acidic nuclear phosphoprotein 32 kDa A) or I_1PP2A (PP2A inhibitor 1): it has been found that this type of protein can bind sphingosine and inhibit PP2A as a consequence[36]. From a functional point of view, this can affect cell proliferation, but striking evidences are still lacking and the protein is still a matter of study.

Among the inhibitors of PP2A, probably the best known – and more widely studied – is SET (Suvar3-9, enhancer of zeste, trithorax), sometimes named after I_2PP2A or TAF-Iβ[5].

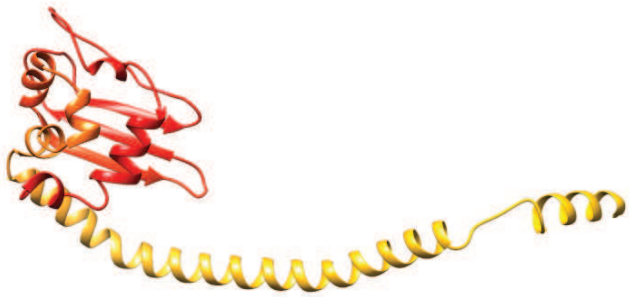


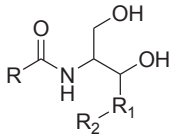
Figure 6. Crystal structure of SET protein, monomer (modified from PDB ID 2E50).

This potent and specific protooncoprotein exists in two splice variants (α and β, 290 and 277 aa respectively), and both of them can inhibit PP2A. In its active form, SET has an average MW of around 28 kDa, and a calculated pI around 4; the N-terminal globular domain is the part of the macromolecule responsible for PP2A interaction[37], while the C-terminal elongated α-helical domain is extremely acidic and seems to be responsible for histone binding[14]. SET can be located both in the cytoplasm and the nucleus[17], and has been shown to have several functions: it is involved in the promotion of chromatin remodeling[19,37] and histone acetylation[14]. Apart from these functions, which are not PP2A-inhibition-dependent, SET is believed to inhibit PP2A by the direct binding to $PP2A_C$; notwithstanding, the precise mechanism by which this happens has still to be cleared out[14]. In fact, the surfaces which are involved on this interaction are not well characerized, since to date there is only one crystal structure for a truncated dimer form of SET on the PDB protein structure databank (ID: 2E50), while no co-crystallized structures

between SET and PP2A are available. What is structurally known is that SET, for being active towards PP2A, has to be phosphorylated on two Ser residues, probably by PKC[18]. Given the binding to PP2A leads to phosphatase inactivation, it should not surprise that SET is found overexpressed in a plethora of tumors: CML and Ph1-ALL (both BCR/ABL driven leukemias), Wilms tumors, lung cancers, AML, B-cell non-Hodgkin lymphoma, B-CLL. SET gene was also found to be fused with the nucleoporin NU214 (CAN) in t(6;9) acute undifferentiated non-lymphocytic leukaemia, probably as a consequence of a chromosomal translocation[5,14,19]. Moreover, ectopic SET expression demonstrated the role it has in primary human NK (Natural Killer) cells, since it leads to the production of interferon-γ and to an increased cytotoxicity[5]. An indirect way in which SET can be upregulated is the overexpression of SETBP1 (SET binding protein 1), a regulator that prevents the inactivation of SET by protecting it from proteases, as found in the case of AML[5].

Activators

The category of PP2A activity enhancers has its main representing element in ceramide (**1**), a sphingolipid metabolite and a structural component of the cell membrane furnished with potent signaling properties[35]. Ceramide has been shown to activate PP2A and to restore its pro-apoptotic activity, but the mechanism by which this happens is poorly understood: indeed, it has been hypothesized to directly bind PP2A holoenzymes containing the PR61/B' subunit, *in vitro* and *ex vivo*[35]. More recently, however, the most acknowledged theory is that ceramide activates PP2A because it binds SET[38] and prevents its inhibiting function towards the phosphatase. . The mechanism by which ceramide can induce apoptosis in prostate-cancer cells involves the suppression of Akt signaling pathway, since the reactivation of PP2A it performs leads to a subsequent AKT dephosphorylation[5].

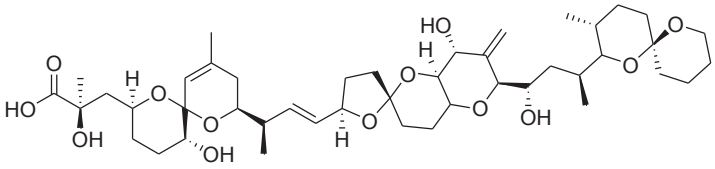


1

Non-physiological modulators

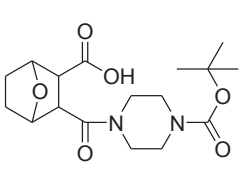
Inhibitors

PP2A has also plenty of non-physiological modulators, which were discovered through the years. The very first one, as already explained, was okadaic acid (OA, **2**), the toxin that brought to consider PP2A as a tumor suppressor for the first time. OA is a carboxylic acid derived from common black sponges, and which in the late 1980s was demonstrated to be a non competitive/mixed inhibitor for PP2A, even if the mechanism which underlies its activity is not totally clear. The inhibition it provides has been shown to be reversible on myosin light chain phosphorylation, so it has been used for a long time as an analysis tool for phosphorylation regulated systems[27].

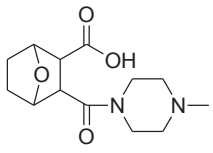


2

In more recent years, cantharidin (which origins from *Lytta vesicatoria* beetles) derivatives LB100 (**3**)[39] and LB102 (**4**)[40] have shown PP2A inhibiting activity as well. Both of these molecules displayed significant effects when co-administered to classical anti-tumor drugs in several forms of cancer (LB100 showed its effects on hepatocellular carcinoma, while LB102 on glioblastoma, neuroblastoma, and stem cell–derived aggressive sarcoma); indeed, their inhibitor activity towards PP2A alters cancer cell signaling leading them to avoid cell cycle checkpoints, fact that makes the cell more vulnerable towards external factors as anti-tumor drugs[39,40].

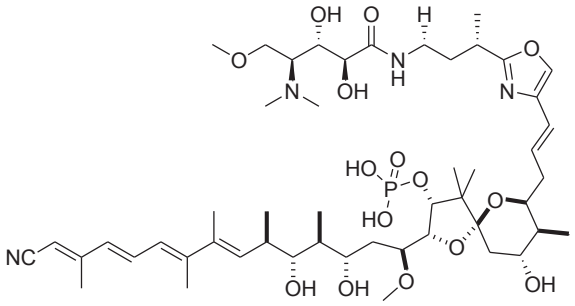


3



4

Another inhibitor coming from natural source is calyculin A (**5**), a spiro ketal skeletal molecule derived from a marine sponge (*Discodermia calyx*). Actually, this compound was first studied in comparison with OA: the two molecules displayed comparable results for what concerning the inhibitory activity on PP2A[41]. However, calyculin A has been shown not to be selective for PP2A, but rather a broad spectrum inhibitor of phosphatases[42].

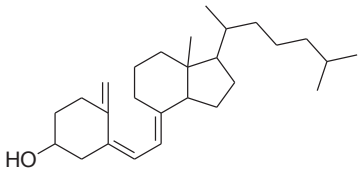


5

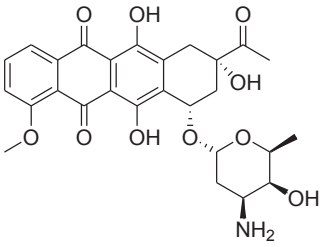
Activators

Given the role the activity of PP2A has in the correct apoptosis of the cells, and keeping in mind the fact that many tumor forms have PP2A activity impaired, it is not surprising that restoring the activity of PP2A has recently been considered as a potential new therapeutic approach against cancerous diseases. In this context, many compounds are currently being investigated, especially in the field of small molecules.

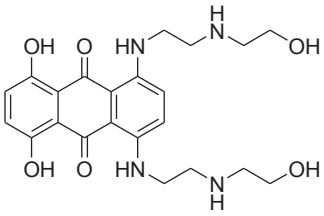
As previously described, ceramides are good PP2A enhancers; however, they cannot be considered as potential drugs because, if administered, they can be easily metabolized. A possible way to increase ceramide levels could be the treatment with ceramide biosynthesis activators; this would increase the physiologically produced quantity of ceramide itself and can be obtained through molecules of very different nature, for example vitamin D3 (**6**) and anthraquinones (daunorubicin, **7**; mitoxantrone, **8**)[43].



6

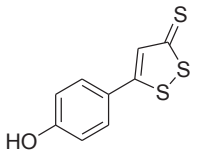


7



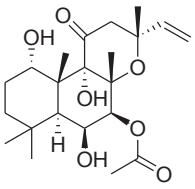
8

Apart from this approach, PP2A activation can be obtained thanks to compounds as dithiolethiones. These are known as chemopreventive molecules, and their activity has been attributed to their antioxidant responsive elements (AREs); however, it has been shown that some of them (for example ACS-1, **9**) target PP2A, activating it independently from *de novo* phosphatase synthesis and restoring its inhibitory activity towards the Akt pathway[44].



9

Forskolin (**10**) is mostly known for its activity of increasing cAMP levels; however, it has also been shown that it can restore PP2A activity in leukemic cells when administered *in vivo*. The involved mechanism has still to be clared out, but it does not seem to reside in a direct interaction with the phosphatase or with one of its inhibitors. Curiously, the capability of forskolin to activate PP2A does not involve cAMP[19].

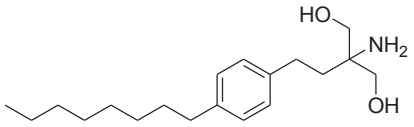


10

Growing interest was gained by OP449, a peptide which comes from Apolipoprotein E (ApoE), known for being responsible for the transport of cholesterol. The residues 133-149 of ApoE (called COG133) fused with the cell-penetrating antennapedia peptide, form a dimeric peptide which is the final

OP449[5,45,46]. Its activatory activity towards PP2A is believed to be caused by its binding to SET, which prevents it from sequestering PP2A; it has been shown to exert antineoplastic effects in some forms of leukemia[45].

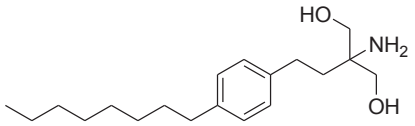
Anyway, the compound that recalled more attention in this sense is Fingolimod (**11**).



11

1.3 FINGOLIMOD

Fingolimod (FTY720), or 2-amino-2-[2-(4-octylphenyl)ethyl]propan-1,3-diol, is commercially known as Gilenya™ (Novartis); it was approved by FDA (*Food and Drug Administration*) in 2010 for the treatment of relapsing-remitting forms of Multiple sclerosis (MS)[47].



11

MS is a chronic disease of the central nervous system which has its main effect in the demyelination of nerves; it is believed to be driven by two main pathogenetic mechanisms, *i.e.* inflammation and degeneration. In this context, inflammation is related to the occurrence of acute episodes of neurologic dysfunction (relapses) and the formation of focal demyelinating lesions in the brain and the spine, while the degenerative effect would be responsible for progression of disability[48].

From the physio-pathological point of view, in MS central memory T cells become self-reactive, and they expand following low-intensity cross-reaction on self antigens in lymph nodes that drain the antigen from the central nervous system[49]. This provides axonal damages by progressively destroying the myelin coating axons in the central nervous system[47].

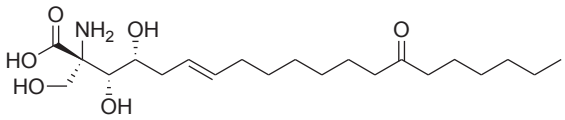
The disease occurs in different clinical forms, likely related to the varying interplay of the pathogenetic mechanisms:

- relapsing-remitting MS (85% of patients) is characterized by relapses followed by total or partial recovery;
- primary progressive MS (10% of patients), characterized by the progression of the degenerative mechanism but not of the inflammatory one;
- secondary progressive MS (30%–40% of patients), which usually follows the relapsing-remitting form after 10-15 years. It is characterized by the progression of degeneration with or without superimposed relapses, and it can be considered to be the overt form of the pathology;

- progressive-relapsing MS (5% of patients), characterized from a progressive onset with superimposed relapses, followed by partial/total recovery[48].

Fingolimod represented an innovation in the field of MS treatments because it was the first drug suitable for oral administration; indeed, until 2010 the only available treatments were injective, with the main part played by interferon β, a cytokine, and natalizumab, a monoclonal antibody. This kind of therapies was not lacking of adverse effects as liver and thyroid dysfunctions, leucopenia, leucoencephalopathy or mood disorders[48].

The structure of Fingolimod was first described in 1995 and was inspired to myriocin (**12**), a fungal metabolite coming from *Isaria sinclarii*[50].

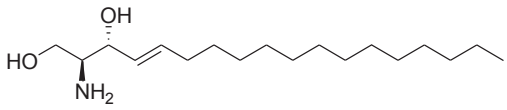


12

Myriocin, traditionally used in Chinese medicine as a drug for "eternal youth", was shown to have immunosuppressant activity. Its potent and selective inhibitory effects on SPT (Serine palmitoyltransferase) were determined in mammalian cells *in vitro* at a nanomolar range[50]. SPT is responsible for the first step of sphingolipid biosynthesis, and its inhibition through myriocin suppresses T cell proliferation thanks to the interference with sphingolipid metabolism. However, due to its severe gastrointestinal effects, it was not suitable as a drug[51]. Nonetheless, its scaffold was used for deriving less toxic analogs: among them, Fingolimod was found to be the best performing one, but its mechanism of action is completely different from that of its parent compound.

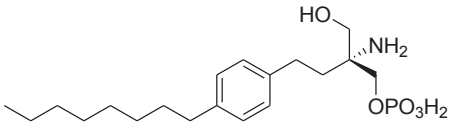
1.3.1 Fingolimod: immunosuppressive action

The mechanism of action that makes Fingolimod an immunosuppressive drug resides in its structural similarity with sphingosine (**13**, see next page).

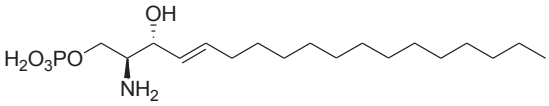


13

Once internalized, Fingolimod can be phosphorylated in vivo by Sphingosine kinase 2 (SK2) to S-Fingolimod-P (**14**), and because of this it reaches a structural similarity with sphingosine-1-P (S1P, **15**), the product of phosphorylation of sphingosine performed by SK1 (mainly) and SK2[49].



14



15

S1P is extremely important in MS; central memory T cells egress from lymph nodes and recirculate to blood in a S1P-dependent manner, and they invade the central nervous system following restimulation by self antigen presented on phagocytes. Here, this type of reactivated T cells causes proliferation and local differentiation of effector T cells[49].

The structural similarity between S-FTY720-P and S1P leads the phosphorylated form of Fingolimod to mimic the action of S1P itself on the same receptors[49]; S1P can interfere with several signaling pathways thanks to the interactions with plasma membrane localized GPCRs (G-protein coupled receptors), called S1PRs (sphingosine-1-phosphate receptors) [52]. Five receptors, named S1PR1-5, are currently known. S1PR1, S1PR2 and S1PR3 are ubiquitous (they are present in the immune, cardiovascular and central nervous systems), S1PR4 is selectively expressed in lymphoid tissues and lung, while S1PR5 in spleen and white matter. The expression of S1PR1 in CD4+ T cells is much higher than the other S1PRs, making of S1PR1 the main receptor on lymphocytes[53]. Indeed, the S1PR1 receptor plays a crucial role in

lymphocyte mobility, by selectively arresting the trafficking of naïve and central memory T and B lymphocytes in peripheral lymphoid organs without affecting the trafficking of effector memory populations[54].

Fingolimod-P (but not parent Fingolimod) can work as an agonist for S1PR1, 4 and 5 (EC_{50} ~0.3–0.6 nM) and for S1PR3 at higher concentrations (EC_{50} ~3 nM) *in vitro*, but not for S1PR2[49]. However, after the the initial stimulation of the receptors, the drug can work as a "functional antagonist" since S-FTY720-P bound to S1PR inhibits their function inducing the internalization of the receptors themselves when these are overstimulated[48,49]. As a consequence also the migration of T-cells, which is mediated by S1PR1 expressed on the membrane surface, is blocked and the blood lymphocyte count decreases[49].

Fingolimod has the great advantage, over other MS drugs, of being orally administrable, which is ideal for the long-term treatment a chronic disease requires; the classical injective therapy for MS lacks this compliance requirement. In addition to this, side effects of Fingolimod are limited to those typical for immunosuppressant drugs: leucopenia and increased risk of infections. In addition to this, however, since S1PRs are also expressed in the cardiac tissue, Fingolimod also leads to the increased risk of bradycardia[48].

1.3.2 Fingolimod and PP2A: why?

The interconnection between PP2A and Fingolimod is not trivial, because it means combining the properties of a pro-apoptotic enzyme with a molecule which is known to be immunosuppressive. After its discovery, Fingolimod showed unexpected anti-cancer functions in CML and Ph1 ALL as a side effect; this was verified *in vitro* and *in vivo* using model systems, but also *ex vivo* on cultured human cells. In these assays FTY720 showed to induce apoptosis in imatinib/dasatinib-sensitive and –resistant leukemia cells, leaving the healthy ones unaffected[55]. In lymphoid cell lines, FTY720 restores Bcl-2 dephosphorylation and apoptosis; primary CLL cell lines treated with FTY720 displayed down-regulation of Mcl-1 but not of Bcl-2, and concurrent activation of the dephosphorylation of ERK1/2[2]; these combined mechanisms are a way to promote apoptosis. Actually, these are not the only cases in which FTY720 showed pro-apoptotic effects, since similar tests were performed also on a number of solid tumors with comparable results. The detailed mechanism of action of Fingolimod is not clear yet; however, in this case for providing this sort of effect it has been shown that Fingolimod does not have to

be phosphorylated[55]. This implies the presence of another mechanism of action, alternative to the S1PRs-related one. Several *in vitro* studies were performed in order to elucidate this point, suggesting the effectiveness of FTY720 on tumor forms like multiple myeloma, hepatocellular, bladder, breast and prostate carcinomas: its pro-apoptotic activity and its ability to suppress mitogenic and survival signals resulted from the inhibition of ERK and PI3K/AKT pathways[5]. Since PP2A is involved in the inactivation of these pathways, pro-apoptotic effect of unphosphorylated FTY720 is thought to be due to the activation of PP2A, at least in leukemias[2,5].

Structural data that could elucidate the mechanism of action at a molecular level are still missing. The most exhaustive study (performed by Saddoughi *et al.*)[38] assessed the occurrence of a direct interaction between Fingolimod and SET, rather than PP2A; the activation of the phosphatase is supposed to be indirect, and mediated by the inhibitory action of Finoglimod on its physiological inhibitor. As already mentioned, SET is overexpressed in several forms of tumor, especially leukemia (considering, however, that the studies are still at a preclinical stage[2]), so the capability of FTY720 of restoring PP2A activity through the interaction with SET could represent a valuable tool in the treatment of leukemia.

Experimental data were obtained through SPR (Surface Plasmon Resonance) performed between SET and, alternatively, ceramide C18 (the most performing of the ceramides, when considering their binding to SET), FTY720 and S-FTY720-P. The response curve for unphosphorylated FTY720 is comparable to that of ceramide, while S-FTY720-P seems to bind quite inefficiently to SET[38]. This confirms that the phosphorylated form of Fingolimod does not bind to SET, while unphosphorylated Fingolimod does.

1.4 LEUKEMIA: AN OVERVIEW

Leukemia represents a group of heterogeneous blood cancers that belong to the group of hematologic disorders; they occur when immature or mature white blood cells multiply without control within the bone marrow[56]. This type of tumor affects men more than women (almost 2:1 ratio) and exists in a number of peculiar forms, which differ for the biological alterations occurring in the cell and, as a consequence, for their main morphological and molecular characteristics[57]. Types of leukemias are traditionally classified on the bases of:

- progression of the disease:
    - acute: fast growing cancer that affects mainly immature cells;
    - chronic: cancer with a slow growth and that affects mature cells;
- type of affected cells:
    - myeloid (involving cells of the innate immune system);
    - lymphocytic (involving cells of the adaptive immune system)[56,57].

The graphs in Figure 7 and 8 indicate the incidence and the mortality of leukemias are shown for the US in a period of 20 years.

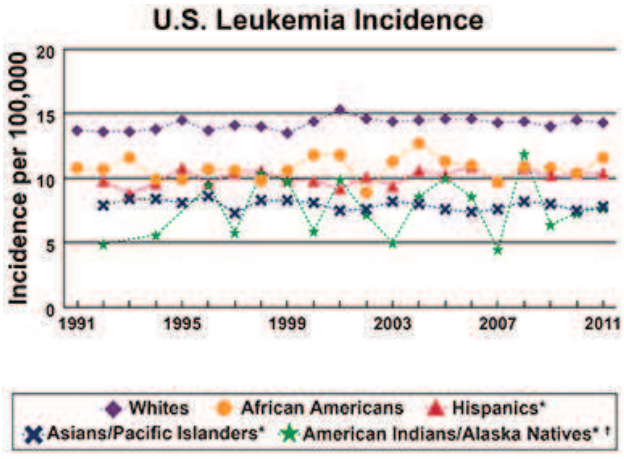


Figure 7. US leukemia incidence/100000 inhabitants [58].

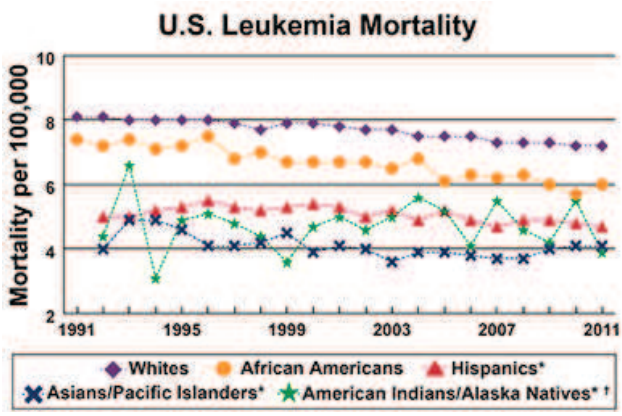


Figure 8.US leukemia mortality/100000 inhabitants[58].

During the considered period of time, the trend of the curves does not seem to change during the considered two decades. Ethnic origins differentiate the incidence of the disease, with Whites being the most affected, while the number of African Americans, Hispanics, Asians and Native Americans is sensibly lower. This could be linked to genetic and congenital factors, which are thought to be the key factor to the onset of the pathology (together with ionizing radiations or exposition to mutagenic chemicals); the details of this ethnical diversification, however, are still not clear [56].

1.4.1 B-CLL

Among the blood tumors, the most outstanding is B-CLL (B cell chronic lymphocytic leukemia), the most common type of leukemia in the Western world, and also the one with the most heterogeneous clinical course among patients[2]; the American Cancer Society estimates the diagnosis of about 14,620 new cases in the USA in 2015[59]. The incidence is higher in male population, with an average age of 65; it is diagnosed to 1 person/year/100000 inhabitants in the 40-50 years age range, and this number increases to 20 cases/year/100000 inhabitants in the 70-80 range. More than 40% of B-CLLs are diagnosed at 75 years old, while the diagnosis before 50 years old accounts for 10% of the whole[60].

B-CLL is a chronic blood disease characterized by the accumulation of monoclonal, mature, neoplastic B lymphocytes co-expressing CD5, CD23 and CD19 antigens[61]. In this kind of leukemia, the signaling pathway of BCR (B cell receptor) is impaired, which leads to the unusual expression of cell surface receptors on the disease-affected B cells. BCR-mediated stimulation has an important role in the

development of the disease, as confirmed by the impairment of the signal transmission they perform in CLL[62]. These abnormal B lymphocytes are found in the peripheral blood, bone marrow, lymph nodes and spleen, and as a consequence normal immunocompetent cells are reduced in their number and impaired in their function[62]. The affected subjects have a high prevalence of autoimmune phenomena, characterized by the production of polyclonal autoantibodies (autoAb) against autoantigenes (autoAg) which are expressed exclusively in blood cells[63].

Nowadays, the main hypothesis about the pathogenesis of B-CLL is that the cells involved in the disease might be selected by some sort of antigenic pressure resulting in a restricted gene repertoire for immunoglobulin heavy chain variable region (IGHV) gene repertoire compared to that found in normal adult B-cells[64]..

1.4.2 PP2A and Leukemia: why?

Even if the precise pathogenesis of leukemia is not completely clear, many considerations about typical pathogenic features can be made; in particular, in this section results of experimental evidences that correlate various forms of leukemia with PP2A activity alterations are reviewed.

In myeloid leukemia, even if genes encoding PP2A subunits are most of the times intact, the enzyme is inhibited in its pro-apoptotic activity because of two main reasons: post-translational modifications of C the subunit, and/or interactions with negative modulators on B and A subunits[16]. Focusing peculiarly on AML, PP2A dysregulation is characterized by the silencing of A subunit and the overexpression of B subunit[16]; in addition to this, SET was demonstrated to be expressed in a markedly increased manner[65]. AML can also result from genetic translocations leading to the overexpression of SETBP1; by protecting SET from protease cleavage, SETBP1 indirectly inhibits PP2A and leads to an enhanced proliferation of leukemic cells[66]. Another study examinated hematopoietic stem cells (HSCs) from CML patients and revealed that PP2A activity was suppressed in CML compared with normal HSCs[67].

As for B-CLL, the most common type of leukemia, several considerations can be made on the role of PP2A. Indeed, some CLL result from 11q gene deletion, and as a consequence reduced $PP2A_A\beta$ subunit transcript levels can be observed. On the other hand, the expression of alternative splicing of $PP2A_A\beta$ (PP2R1B) transcripts (skipping of exons 2/3, 3, 9) was associated with a reduced activity of PP2A[68].

In order to evaluate PP2A most classical counterpart, also SET expression and activity were assessed in B-CLL. Basically, SET (both α and β isoforms) is found to be overexpressed in CLL cells when compared with healthy B cells. As a confirmation of this, elevated SET levels in patients were associated with the already acknowledged prognostic factors, like IGVH mutational status. To determine whether SET levels are indicative of more rapid CLL disease progression, cell extracts coming from 285 patients were evaluated; as a result, patients with high SETα, high total SET (α+β), and a high α/β ratio showed earlier need for therapeutic intervention, indicating that SET expression levels correlate with the progression of the pathology. In parallel to this, it was demonstrated that, in CLL cells, increased levels of SET correlated significantly with disease severity and a poorer outcome[69].

This effect can also be enhanced through the stabilization of PP2A/SET complex, which is shown to happen in CLL due to an aberrant activity of the Tyr kinase Lyn; this kinase is one of the most relevant enzymes in the onset of B-CLL, and it has been shown to exert its activity in phosphorylating Tyr^{307} of $PP2A_C$, stabilizing the interaction between PP2A and SET[70].

All of these data confirm a link between PP2A inactivation and the pathogenesis of leukemia, especially B-CLL; this feature can be extremely useful for a tumor cell, which has in immortalization one of its main developmental characteristics.

1.5 AIM

B-CLL is the most common type of leukemia in the Western world, and the American Cancer Society estimatd the diagnosis of about 14,620 new cases in the USA in 2015[59]. Since unphosphorylated FTY720 can affect leukemia by the inhibition/disruption of the interaction between PP2A and SET[38], it would be worth considering it as a lead compound for the development of new potential drugs against B-CLL. This project aims to develop new Fingolimod analogs that could reactivate PP2A in tumors, avoiding immunosuppressive effects. As already mentioned, these two activities are discriminated by the phosphorylation state of FTY720: the compound does not have to be phosphorylated for acting as a tumor suppressor, while it has immunosuppressive activity in its phosphorylated form. In the proposed approach, this was performed by modulating/limiting phosphorylatable sites; in addition to this, also further molecular features of Fingolimod were modified in order to elucidate the structural features required for the compounds to be active.

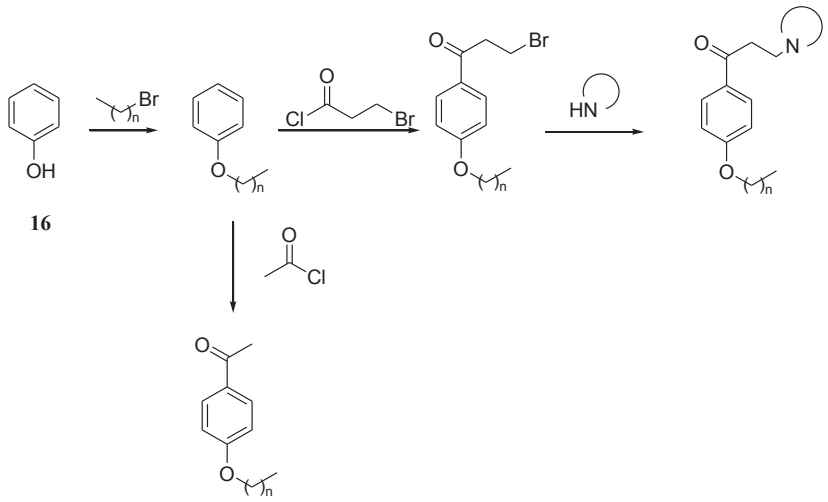
A second aim of this project has been the production of the recombinant forms of the proteins of interest: SET and PP2A (being more specific, the catalytic subunit of PP2A, which is the one that was demonstred to interact with SET[71]). This was pursued through a placement at Diamond Light Source and the Research Complex at Harwell, Oxfordshire (UK) thanks to the supervision of Dr. M. Mazzorana, Prof. T.S. Sorensen and Prof. R. Owens. In this way, the newly synthesized compounds could be directly tested on the recombinant proteins, leading to the development of biochemical assays for inhibitors of SET-$PP2A_C$ interaction. After an optimization phase, experiments were performed to evaluate and quantitate both the binding and the kinetic processes underlying the interactions between the considered proteins and the obtained small molecules, through techniques such as HRMS and NMR. All the structure/function data will allow the design of an effective structure-activity relationship and an aware application of *in silico* tools for better understanding the interactions occurring between the proteins and the active compounds. In perspective, this approach will allow the design of Fingolimod derivatives optimized for the treatment of neoplastic malignancies and with reduced side effects on the immune system.

2. RATIONALE AND RESULTS

2.1 RATIONALE OF THE SYNTHETIC PATHWAYS

The following synthetic schemes are meant to provide the synthesis of compounds inspired by FTY720: most of the times the para-disubstituted aromatic ring is maintained, but with structural characteristics which could modulate/limit the affinity for S1PRs (for example, the length of the "tail") and enhance the affinity towards PP2A/SET. In the explorative approach that has been decided to pursue, this meant starting from reactants readily available in commerce, for example phenol, p-hydroxybenzaldehyde and more, in order to create a small library of compounds that could provide the widest possible range of structural features. This kind of approach aims, after adequate biochemical and biophysical evaluations, to provide information in terms of structure-activity relationship for a deeper understanding of the structural features underling the potential interaction between PP2A/SET and the molecules themselves.

PATHWAY 1.1:

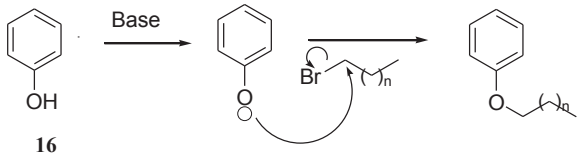


The first synthetic pathway comes from the idea of creating a set of molecules with two main characteristics: a modulated "tail" length and a "head" in some way comparable to Fingolimod one, since the aromatic scaffold is functionalized with a linker which displays (in all of the cases but one) a

morpholine or a piperidine ring. The latter mentioned compounds were used because Fingolimod has a primary amine and two hydroxyl groups in its "head"; keeping this in mind, morpholine was chosen because it owns both a nitrogen atom and an oxygen one, but comprised in a cycle. This means that the oxygen atom lacks of the capability of being phosphorylated by SKs, which is responsible for the activation of S1PRs pathway that is needed to be avoided. Because of this the importance of the oxygen presence could be better elucidated, evaluating it just for its physical-chemical characteristics and not for being a phosphorylation substrate. This is the reason why also piperidine was used, since it is extremely similar to morpholine but lacking the oxygen atom. As a sort of negative control, also a compound lacking rings on the "head" was synthesized.

***Williamson* reaction**

The first reaction shown here is a classical S_N2 *Williamson* etherification (through the procedure reported by *Gamliel*[72]) on the hydroxyl group of the phenol (1 equivalent), which is deprotonated by a base (in this case potassium carbonate, 2 equivalents) and in this way it can attack the haloalkane (0.9 equivalent) with the following mechanism:



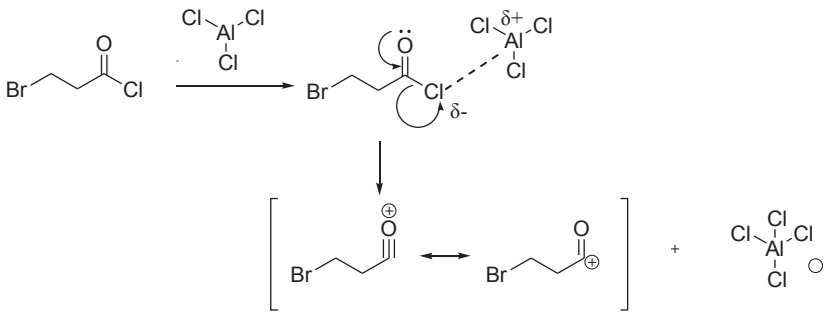
16

This reaction has been performed with a series of halo-alkenes for mimicking the "tail" of Fingolimod but varying its length.

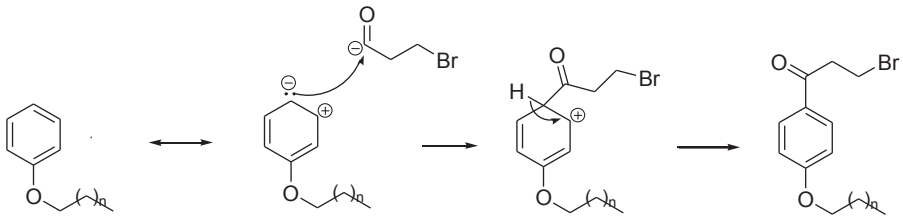
***Friedel-Crafts* acylation**

The second step of the path is represented by a *Friedel-Crafts* acylation through the procedure reported by *Kiuchi*[73], which inserts a linker in the para position of the molecule through a electrophilic aromatic addition. The reaction is carried out at 0°C in dichloroethane (DCE) in the presence of a *Lewis* acid, *i.e.* aluminum trichloride, in stoichiometric ratio (1.5 equivalents). The following scheme represents the

mechanism of activation of 3-bromopropyonil chloride (1 equivalent) in these conditions, in order to create an acylium cation thanks the formation of a complex between $AlCl_3$ and the acyl chloride:



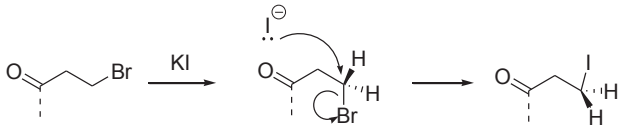
This step has also been carried out with acetyl chloride. The carbocation is stabilized by resonance[74] and can react easily with the aromatic system of the alkylated phenol (1 equivalent); actually, the presence of the ether activates the ring on ortho and para positions, but the acylation occurs preferentially in para. This probably happens because the alkyl chain of the ether provides a hindrance which makes the reaction more difficult to occur in ortho. Since this reaction introduces an acceptor group, multiple substitutions are unlikely to occur[75].



The presence of the bromine atom permits in this case a further substitution on it, in order to insert either morpholine or piperidine; of course this cannot be done with the derivative obtained with acetyl chloride.

Halide exchange and nucleophilic substitution: *Finkelstein* procedure

The nucleophilic substitution occurs through the *Finkelstein* procedure[76], which showed to be extremely efficient for the insertion of the heterocycles of interest. Basically, the final aim here is to substitute the bromine atom of the "head" linker with a nucleophile but this, due to the *Finkelstein* protocol, occurs in two steps: indeed, first of all the bromine atom is substituted with a iodine atom, which is more reactive and represents an extremely good leaving group; the iodine atom itself is then easily substituted with the nucleophile of interest (in this case, either morpholine or piperidine, 1.5 equivalents with respect of 1 equivalent of acylated substrate). The halogen exchange is permitted by the addition to the reaction mixture of potassium iodide (1.5 equivalents), while the addition of calcium carbonate (1.5 equivalents) buffers the acid which is produced as a byproduct.



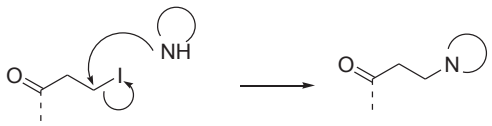
As shown by the mechanism, the reaction occurs with an inversion of the configuration on the involved carbon atom, and it is a S_N2. The reaction can be considered to be driven by the different solubility of the involved salts, since the byproduct KBr is poorly soluble. Apart from this, I^- is a good leaving group for two main reasons: the strength of the C-X bond and the stability on solution of X^-, once the C-X bond is broken. Both the characteristics can be seen in the following table[77], since it reports the values of the C-X bond strengths (kJ/mol) and the pK_a of the HX acids, which is inversely proportional to the degree of dissociation of the acid itself (and so to the stability of X^- in solution).

Halide (X)	Strength of C–X bond, kJ mol^{-}	pK_a of HX
fluorine	118	+3
chlorine	81	–7
bromine	67	–9
iodine	54	–10

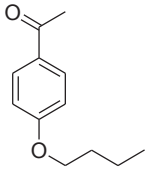
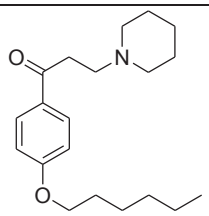
Table 2. Comparison among halides as leaving groups[77].

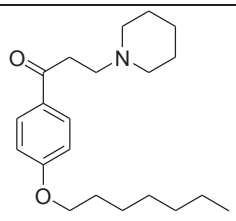
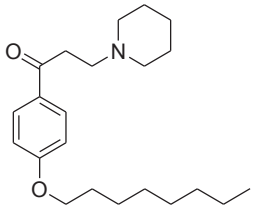
Compared to the other halogens, iodine has the lower bond energy and the lower pK_a, and so I^- can be considered to be the best leaving group.

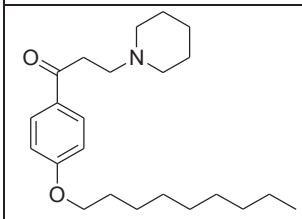
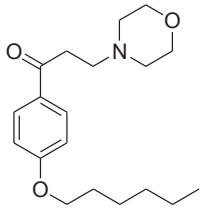
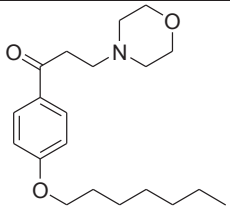
After the halide exchange, the nucleophilic substitution can easily occur:



In the following tables, the obtained compounds are reported.

Compound structure and number	IUPAC name
 **25**	1-(4-butoxyphenyl)ethanone
 **30**	1-(4-(hexyloxy)phenyl)-3-(piperidin-1-yl)propan-1-one

Compound structure and number	IUPAC name
 **31**	1-(4-(heptyloxy)phenyl)-3-(piperidin-1-yl)propan-1-one
 **32**	1-(4-(octyloxy)phenyl)-3-(piperidin-1-yl)propan-1-one

Compound structure and number	IUPAC name
 **33**	1-(4-(nonyloxy)phenyl)-3-(piperidin-1-yl)propan-1-one
 **34**	1-(4-(hexyloxy)phenyl)-3-morpholinopropan-1-one
 **35**	1-(4-(heptyloxy)phenyl)-3-morpholinopropan-1-one

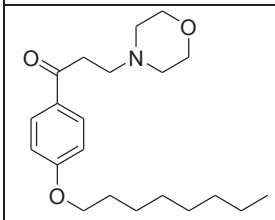
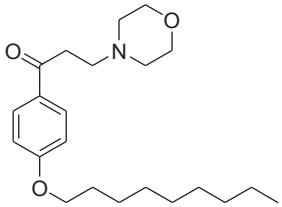
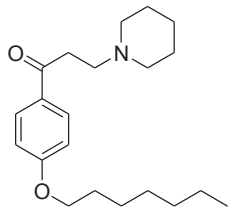
Compound structure and number	IUPAC name
 **36**	3-morpholino-1-(4-(octyloxy)phenyl)propan-1-one
 **37**	3-morpholino-1-(4-(nonyloxy)phenyl)propan-1-one

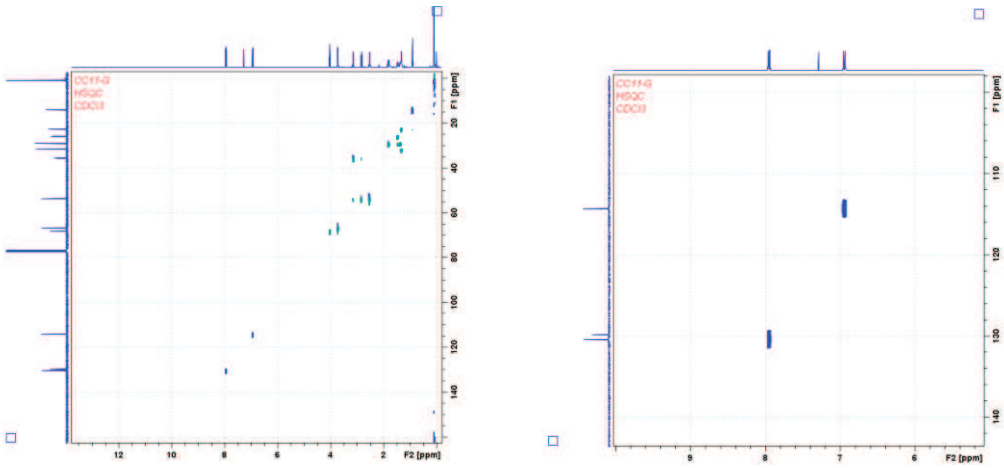
Table 3. Compounds obtained from pathway 1.1.

Characterization: a focus

One of the compounds synthesized through this pathway (**31**) was taken as an example for further characterizations, mainly performed by NMR; being more specific, 2D heterocorrelated NMRs as HSQC (Heteronuclear Single Quantum Coherence, for detecting ^{1}J) and HMBC (Heteronuclear Multiple Bond Coherence, for detecting ^{2}J and ^{3}J) are provided here.



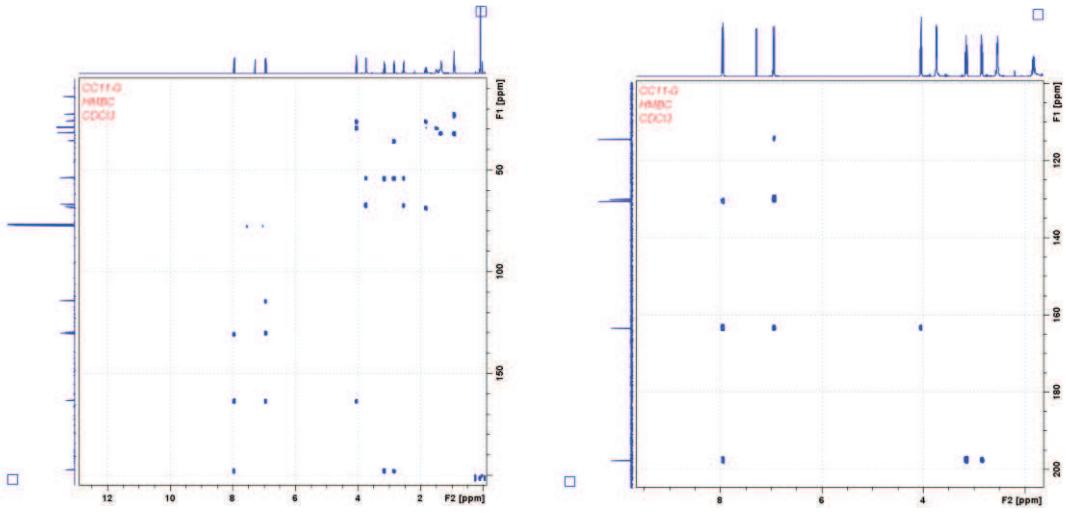
31



(a) (b)

Figure 9. HSQC spectrum of **31** (a) and a zoom on the aromatic part (b).

From the zooming on the aromatic part of HSQC (Figure 9b), it is possible to see that the ^{13}C at δ 129.7 ppm is directly bonded to the proton at δ 7.95 ppm (2H, ArH), while the carbon at δ 114.2 ppm is linked to the proton signal at δ 6.94 ppm (2H, ArH).



(a) (b)

Figure 10. HMBC spectrum of **31** (a) and a zoom of it (b).

From the zooming on the HMBC spectrum (Figure 10b) it is possible to find out the coupling between the ^{13}C at δ 197.5 ppm and the proton signal at δ 7.95 ppm (2H, ArH). As it can be seen from the zooming on the combined HSQC (blue)-HMBC (red) spectra (Figure 11), the carbon signal at 197.5 ppm just gives an

HMBC signal but not an HSQC one; because of this, and of its chemical shift too, it is possible to assign the ^{13}C signal at 197.5 to the carbonyl carbon.

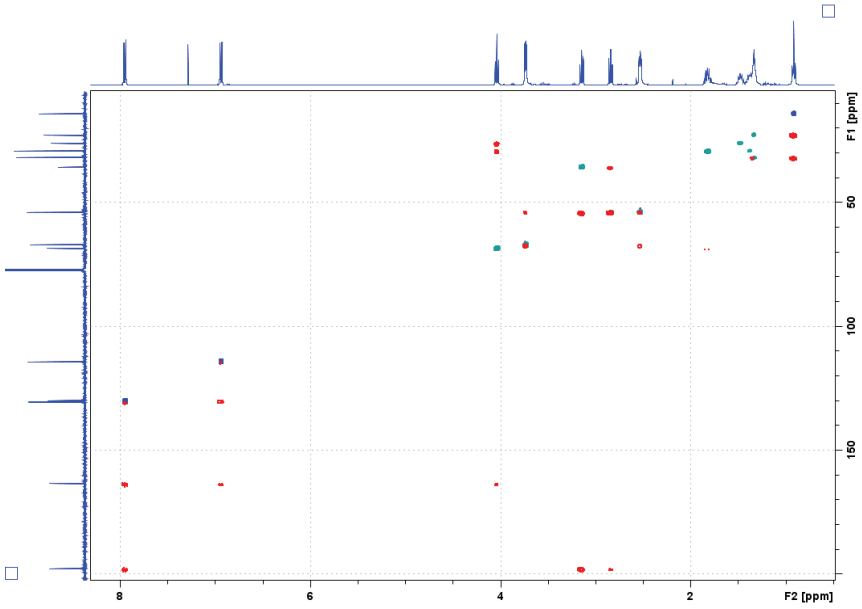
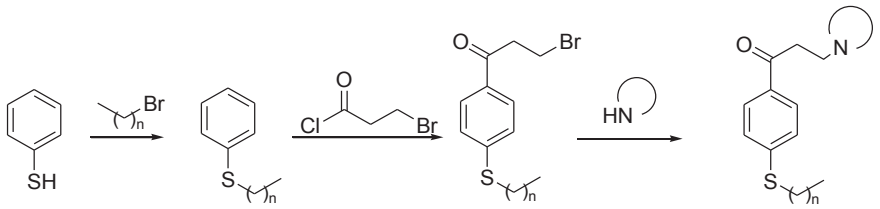


Figure 11. Combined HSQC (blue)-HMBC (red) spectra.

PATHWAY 1.2:

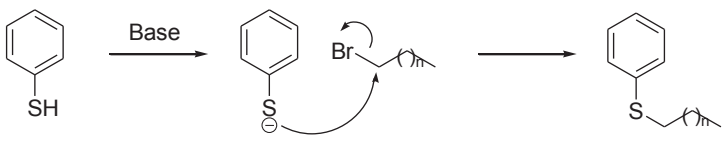


38

This pathway originates from the first presented one but with a substantial variation: instead of the phenol, here thiophenol is used. Both the elements belong to the VI group but to two subsequent periods, so they have similar chemical characteristics but different atomic radius and, connected to this, electronegativity. This specific change could tune the eventual affinity of the molecule for the substrate in a more fine way, providing more clues about the interactions they can have.

Thiol alkylation

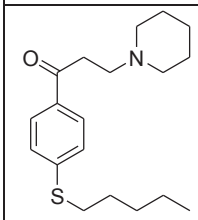
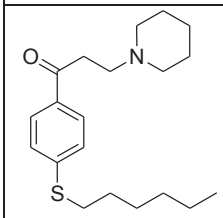
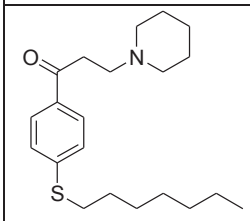
The first step is again an alkylation on the heteroatom (in this case, sulfur) through a procedure quite similar to the previously seen *Williamson* reaction: the S_N2 reaction between thiophenol (1 equivalent) and the chosen bromo-alkane (1 equivalent) is carried out in the presence of potassium hydroxyde (1 equivalent) as a base, and ethanol as a solvent[78]. As shown, the reaction mechanism is totally identical to the previous one: the base deprotonates the thiol from its proton, and in this way the S^- atom can attack the bromoalkane of interest.



38

The synthetic steps which follow are performed in the same manner as previously reported: a *Friedel-Crafts* acylation is carried out with 3-bromo propyonil chloride, followed by a nucleophilic substitution with morpholine or piperidine through the *Finkelstein* procedure.

The following table shows the obtained compounds:

Compound structure and number	IUPAC name
 **46**	1-(4-(pentylthio)phenyl)-3-(piperidin-1-yl)propan-1-one
 **47**	1-(4-(hexylthio)phenyl)-3-(piperidin-1-yl)propan-1-one
 **48**	1-(4-(heptylthio)phenyl)-3-(piperidin-1-yl)propan-1-one

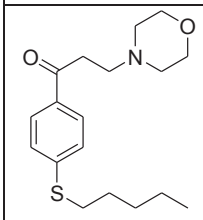
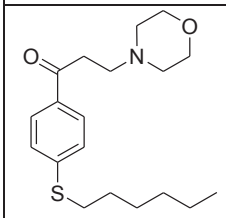
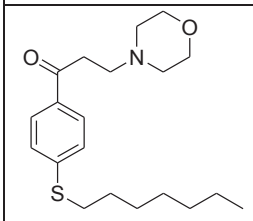
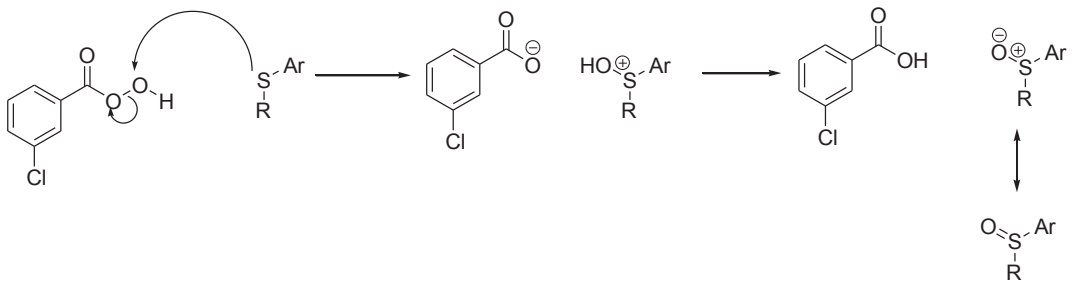
Compound structure and number	IUPAC name
 **49**	3-morpholino-1-(4-(pentylthio)phenyl)propan-1-one
 **50**	1-(4-(hexylthio)phenyl)-3-morpholinopropan-1-one
 **51**	1-(4-(heptylthio)phenyl)-3-morpholinopropan-1-one

Table 4. Compounds obtained from the pathway 1.2.

One of the advantages of this pathway is the possibility to insert further modifications: on compound **51**, in fact, an oxidation on the S atom was performed in order to produce a sulfoxide from the thioether group.



The oxidation of the thioether (1 equivalent) is performed through the use of m-CPBA (meta-chloroperbenzoic acid, 1 equivalent) in DCM.

This promptness of the sulfur atom in being further modified makes this pathway extremely suitable for the explorative approach which is pursued in this work.

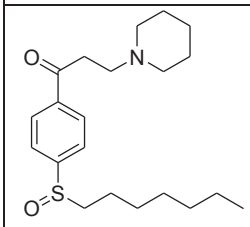
Compound structure and number	IUPAC name
 **52**	1-(4-(heptylsulfinyl)phenyl)-3-morpholinopropan-1-one

Table 4. Compound obtained from the prosecution of pathway 1.2.

Characterization: a focus

In the following image, the ^{1}H-NMR of the intermediates and one of the final compounds (**49**) are reported (400 MHz, $CDCl_3$)

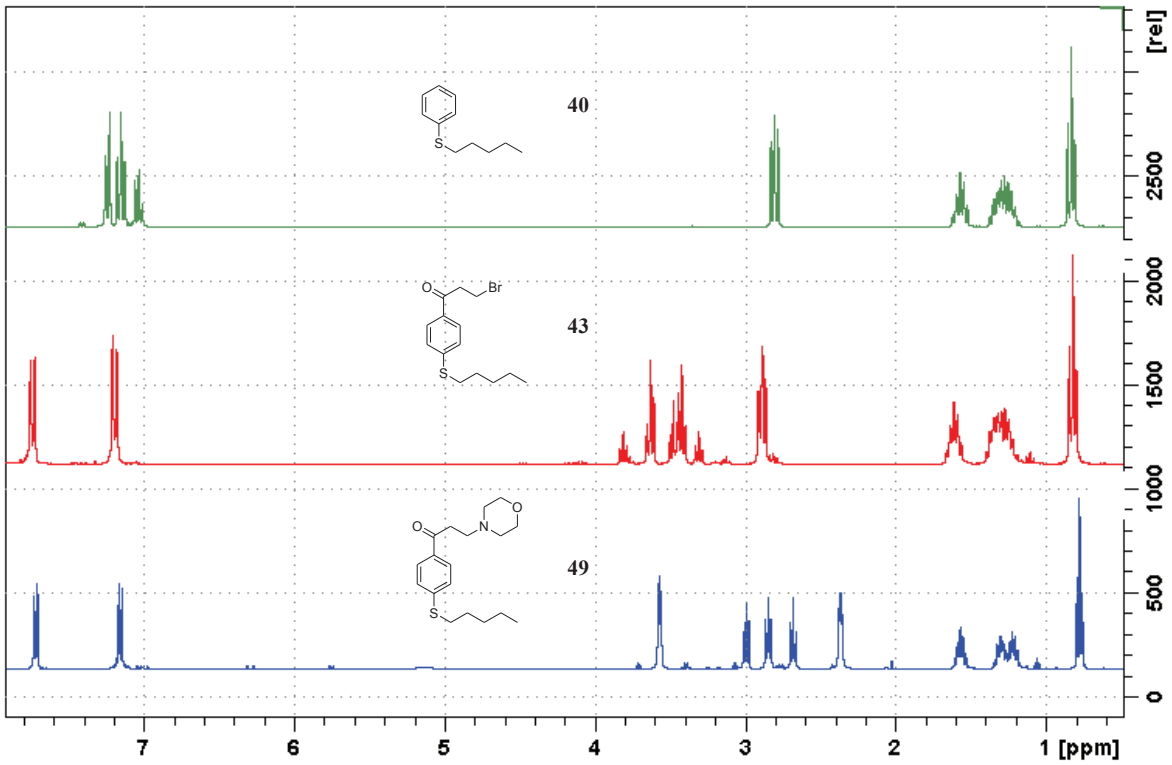
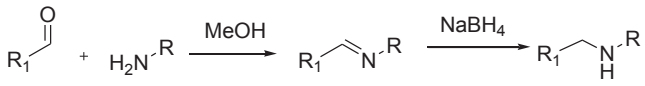


Figure 12. ^{1}H-NMR of compounds **40**, **43** and **49.**

Between the first two spectra, the change in the aromatic system is pretty clear: from a five-proton system, which is present in compound **40**, it was possible to obtain an AA'XX' system for compound **43** thanks to the *Friedel-Crafts* acylation, which clearly shows that the reaction occurs in the para position.

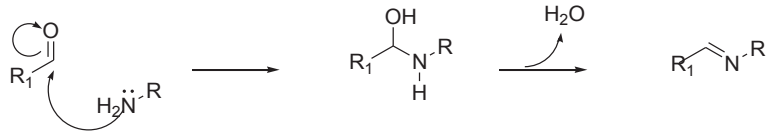
The main differences between the second and the third spectrum, instead, are the presence of the morpholine protons in compound **49** (δ 3.6 and 2.4 ppm, 4 H each) and the down-shift of the signals of the linker between the aromatic system and the morpholine itself. In fact, from two triplets at δ 3.6 and 3.4 ppm (2 H each) two triplets at δ 3.0 and 2.8 ppm were obtained.

PATHWAY 2.1:

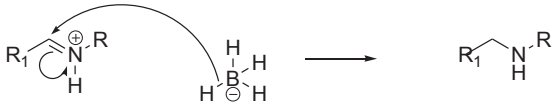


Reductive amination

This pathway was the most extensively used one, because it consists of an efficient and easy way to produce imines and amines; since Fingolimod owns a primary amine on the "head", the insertion of a nitrogen atom as an imine or an amine could help in detecting its role concerning the activity. This synthetic pathway, starting from a carbonyl group and a primary amine (1 equivalent each), leads to the production of a *Schiff* base through the attack of the electron-rich nitrogen atom towards the carbonyl carbon and its subsequent dehydration; the imine is then reduced to a secondary amine using a hydride (in this case sodium boron hydride, 2 equivalents).



After the attack of the nitrogen atom on the electron-poor carbon, the obtained N-substituted hemiaminal loses water to give a *Schiff* base. This reaction is possible both with aldehydes and with ketones, even if the latter react slowlier[76], and is carried out in methanol at 60°C. The addition of sodium boron hydride at 0°C provides the reduction of the double bond and the obtainment of the secondary amine:



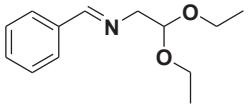
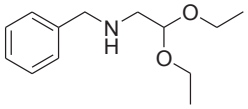
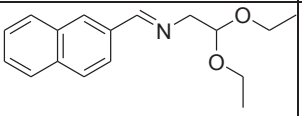
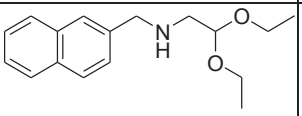
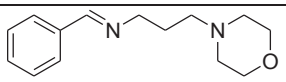
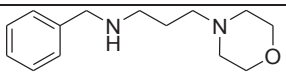
The reduction proceeds thanks to the transfer of a hydrogen atom of $NaBH_4$, with its electronic doublet, to the substrate that gets reduced. The reduction of the imine takes advantage of the fact that the imine is partly protonated as an iminium ion, that can be rapidly reduced to the corresponding amine[77]. The resulting BH_3 species is technically able to transfer other hydrogen atoms and electrons, and so to reduce more imines.

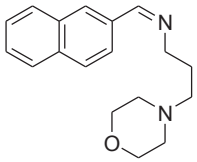
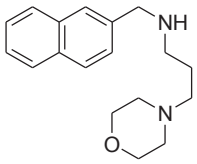
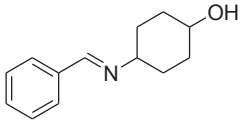
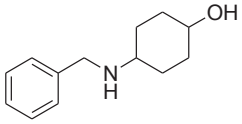
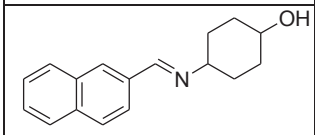
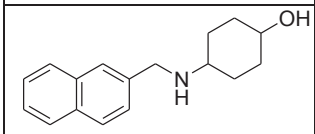
In some cases the isolation of the imine was possible, while sometimes it was not. The reasons behind this probably reside in the stability of the imine itself[77] and in the presence of undesired water in the reaction environment that also could decrease the quantity of final amine.

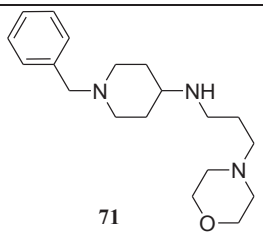
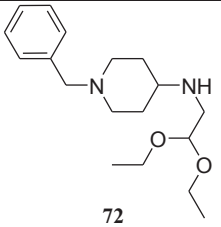
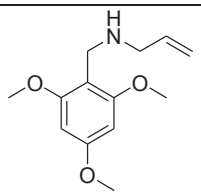
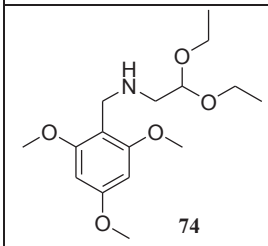
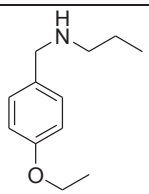
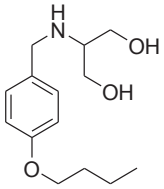
Several carbonyl substrates have been considered for this approach, for gaining as many pieces of structural information as possible: benzaldehyde, 2-naphthaldehyde, 2,4,6-trimethoxybenzaldehyde, N-benzyl piperidone. In addition to this, also p-hydroxybenzaldehyde and 3,5-dimethoxybenzaldehyde were used, but in this case the hydroxyl group/groups were previously alkylated through the already reported *Williamson* reaction, using bromoalkanes of various lengths. This was meant in order to obtain a variation on the "tails" of the molecules, which could be combined to the insertion of the imine/amine on the "head" of the molecule.

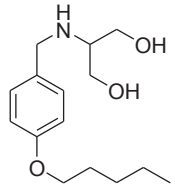
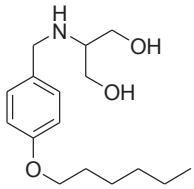
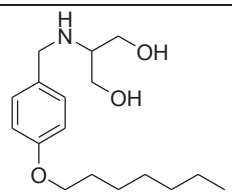
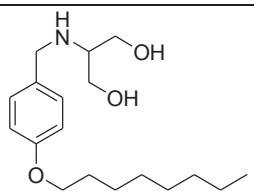
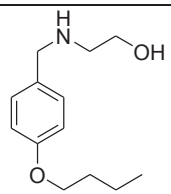
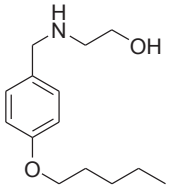
The considered primary amines are even more, and provide the insertion from none to two hydroxyl groups on the "head" of the molecule; the choice included also phosphorylatable substrates, which is supposed to be an unwanted phenomenon, but in this way the library which has been synthesized can be considered to be more complete and all the possibilities about structure-activity relationships can be explored.

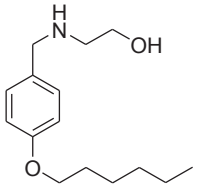
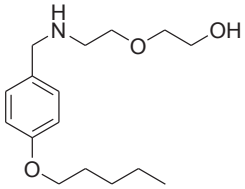
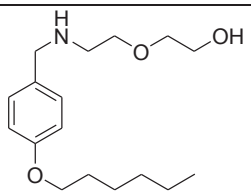
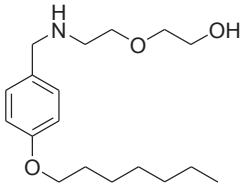
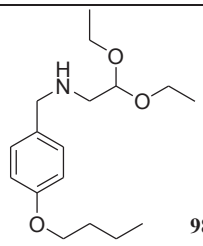
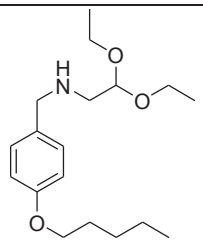
In the following tables, the compounds which were obtained following this synthetic pathway (both imines and secondary amines) are reported:

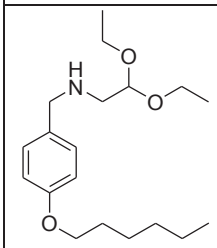
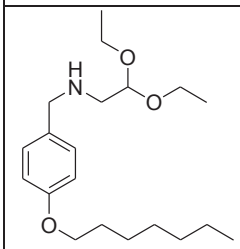
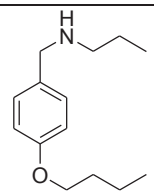
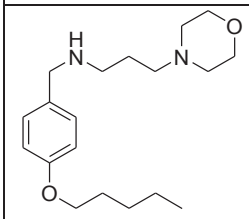
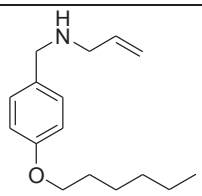
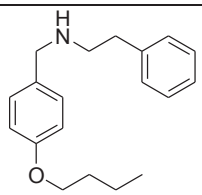
Compound structure and number	IUPAC name
 **56**	N-benzylidene-2,2-diethoxyethanamine
 **57**	N-benzyl-2,2-diethoxyethanamine
 **58**	2,2-diethoxy-N-((naphthalen-3-yl)methylene) ethanamine
 **59**	2,2-diethoxy-N-((naphthalen-3-yl)methyl)ethanamine
 **60**	N-benzylidene-3-morpholinopropan-1-amine
 **61**	N-benzyl-3-morpholinopropan-1-amine

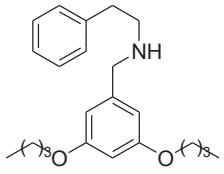
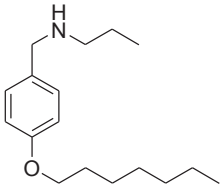
Compound structure and number	IUPAC name
 **62**	3-morpholino-N-((naphthalen-3-yl)methylene)propan-1-amine
 **63**	3-morpholino-N-((naphthalen-3-yl)methyl)propan-1-amine
 **64**	4-(benzylideneamino)cyclohexanol
 **65**	4-(benzylamino)cyclohexanol
 **66**	4-((naphthalen-3-yl)methyleneamino)cyclohexanol
 **67**	4-((naphthalen-3-yl)methylamino)cyclohexanol

Compound structure and number	IUPAC name
 **71**	1-benzyl-N-(3-morpholinopropyl)piperidin-4-amine
 **72**	1-benzyl-N-(2,2-diethoxyethyl)piperidin-4-amine
 **73**	N-(2,4,6-trimethoxybenzyl)prop-2-en-1-amine
 **74**	N-(2,4,6-trimethoxybenzyl)-2,2-diethoxy ethanamine
 **86**	N-(4-ethoxybenzyl)propan-1-amine
 **87**	2-(4-butoxybenzylamino)propane-1,3-diol

Compound structure and number	IUPAC name
 **88**	2-(4-pentyloxybenzylamino)propane-1,3-diol
 **89**	2-(4-(hexyloxy)benzylamino)propane-1,3-diol
 **90**	2-(4-(heptyloxy)benzylamino)propane-1,3-diol
 **91**	2-(4-(octyloxy)benzylamino)propane-1,3-diol
 **92**	2-(4-butoxybenzylamino)ethanol
 **93**	2-(4-(pentyloxy)benzylamino)ethanol

Compound structure and number	IUPAC name
 **94**	2-(4-(hexyloxy)benzylamino)ethanol
 **95**	2-(2-(4-(pentyloxy)benzylamino)ethoxy)ethanol
 **96**	2-(2-(4-(hexyloxy)benzylamino)ethoxy)ethanol
 **97**	2-(2-(4-(heptyloxy)benzylamino)ethoxy) ethanol
 **98**	N-(4-butoxybenzyl)-2,2-diethoxyethanamine
 **99**	N-(4-(pentyloxy)benzyl)-2,2-diethoxyethanamine

Compound structure and number	IUPAC name
 **100**	N-(4-(hexyloxy)benzyl)-2,2-diethoxyethanamine
 **101**	N-(4-(heptyloxy)benzyl)--2,2-diethoxyethanamine
 **102**	N-(4-butoxybenzyl)propan-1-amine
 **103**	N-(4-(pentyloxy) benzyl)-3-morpholino propan-1-amine
 **104**	N-(4-(hexyloxy)benzyl)prop-2-en-1-amine
 **105**	N-(4-butoxybenzyl)-2-phenyl ethanamine

Compound structure and number	IUPAC name
 **106**	N-(3,5-dibutoxybenzyl)-2-phenylethanamine
 **107**	N-(4-(heptyloxy)benzyl)propan-1-amine

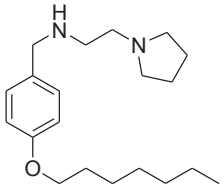
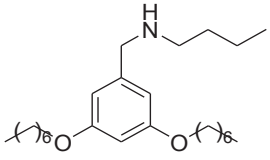
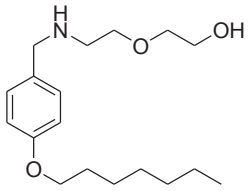
Compound structure and number	IUPAC name
 **108**	N-(4-(heptyloxy)benzyl)-2-(pyrrolidin-1-yl)ethanamine
 **109**	N-(3,5-bis(heptyloxy)benzyl)butan-1-amine

Table 5. Compounds obtained from the pathway 2.1.

Characterization: a focus

Some of the molecules belonging to this group were characterized in deeper detail, as can be seen for compound **97**:



97

This compound was characterized not only by HRMS and 1D-NMR (^{1}H and ^{13}C), but also through bidimensional homocorrelated NMR spectra as CoSY (Correlation Spectroscopy) and NOESY (Nuclear Overhauser Effect Spectroscopy). This was meant in order to better elucidate the attribution of the ^{1}H signals.

In the CoSY experiment, the scalar coupling of protons is shown through cross-peaks on a 2D surface. The following image represents a zoom on the aliphatic part of the spectrum

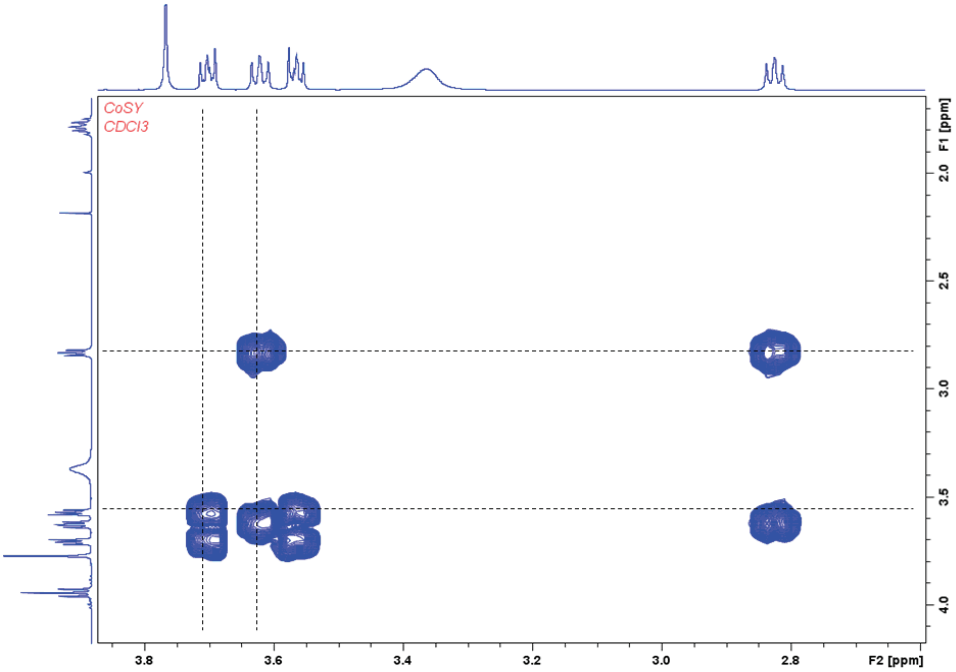


Figure 13. CoSY spectrum of compound **97**.

This portion of the spectrum makes possible to identify the scalar coupling between the signals at δ 3.72 ppm and the ones at δ 3.58 ppm, and between the signals at δ 3.62 ppm and δ 2.83 ppm. Since all of them are triplets and integrate for 2 H, they can be considered to be two distinct groups of CH_2-CH_2 protons. The only credible attribution that can be made is that these protons belong to the aminoethoxyethanol portion of the molecule.

For getting a clearer idea about the attributions, a NOESY experiment was performed; in this case the homonuclear bidimensional spectrum takes into account the space proximity of the considered groups of protons, instead of their couplings through bonds. In the following image, a zoom on the aliphatic portion of the spectrum is reported.

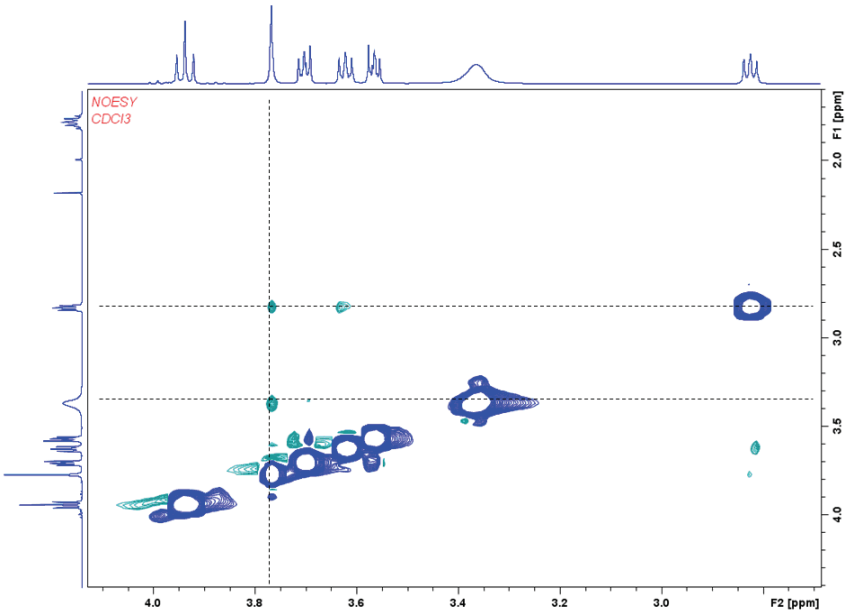


Figure 14. Zoom#1 of the NOESY spectrum of compound **97**.

Here the cross-peaks show space closeness between the protons responsible for the signal at δ 3.76 ppm (s, 2H) with the proton at δ 3.36 ppm (bs, 1H), and also between the same signal at δ 3.76 ppm and the group at δ 2.83 ppm (t, 2H), which was present also in the previously shown portion of CoSY and was said to belong to the aminoethoxyethanol "head". Because of the integrals, the molteplicity and the cross-peaks, the signal at δ 3.76 ppm can be considered to belong to the protons in the benzylic position, the one at δ 3.36 ppm to the amine and the one at δ 2.83 ppm to the portion of aminoethoxyethanol which is closer to the nitrogen atom.

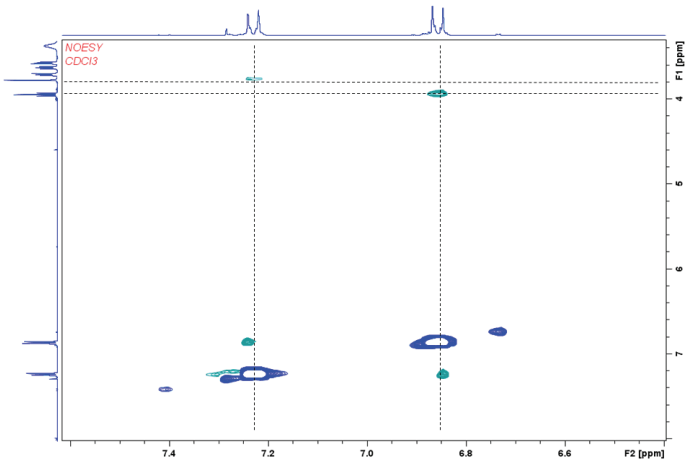


Figure 15. Zoom#2 of the NOESY spectrum of compound **97**.

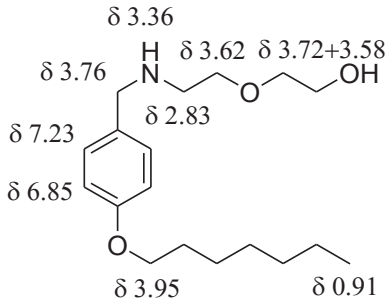
The second NOESY zoom shows that the aromatic protons responsible for the signals at δ 6.85 ppm (AA', 2H) are close to the protons at δ 3.95 ppm (t, 2H) while the aromatic protons at δ 7.23 ppm (XX', 2H) are close to the ones at δ 3.76 ppm, previously attributed to the benzylic group. The attribution of the aromatic protons can be explained considering the mesomeric and the inductive effects, keeping in mind that:

- mesomeric (or resonance) effects are defined as the redistribution of electrons which takes place in unsaturated (and preferably conjugated) systems via their π-orbitals; the delocalization of electrons through resonance is the most powerful factor that affects the stability of molecules. Mesomeric effects are permanent polarizations in the ground state of the molecule, and in neutral compounds there will always be a +M group (the one that donates electrons) and –M group (the one that accepts them)[79];

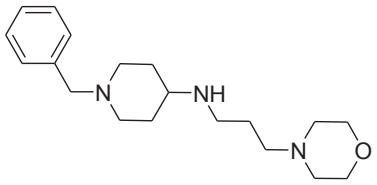
- inductive effects are defined as the polarization of the electron density of a σ-bond caused by the electronegativity of a nearby atom; it is a polarization effect that proceeds from atom to atom along a chain, and that decreases as it proceeds away from the atom/group which produces it. In this sense, the groups of a molecule can be classified as +I (electron donating) and –I (electron attracting)[79].

In the molecule taken into account both the inductive and the mesomeric effects are present, because of the presence of electronegativity differences among close atoms and of the high possibility of delocalization π of electrons, respectively. However, inductive effects would deshield more the part of the aromatic system which is closer to the ether group, because of the electronegativity of the O atom, which is definetly higher than the one of the C atom in para; this would lead the aromatic protons in ortho to O to be present at higher ppm. On the other hand, mesomeric effects lead to the opposite result, because of the way π electrons can delocalize in the system, involving both the aromatic system and the lone pair of the nitrogen atom. Anyway, it is known by the literature that mesomeric effects are usually stronger than inductive effects and are more effective over longer distances[79], so it is reasonable to see that the aromatic protons in meta to the O atom can be found at higher ppm than the ones in ortho.

Taken together, all of this information can lead to the following attribution of the signals for compound **97**:



Another of the compounds synthesized with this approach is **71**:



71

This compound was obtained by a reductive amination between 1-benzyl-4-piperidone and propylamino morpholine; it was chosen as an example compound for better elucidating structural and conformational characteristics because it owns structural features (*e.g.* propyl linker, morpholine, phenyl group …) which are detected also in other of the synthesized molecules, so it can be considered to be representative for almost all the structural features taken into account. The characterization of **71** was perfomed by NMR, and in particular ^{1}H, ^{13}C, CoSY and NOESY.

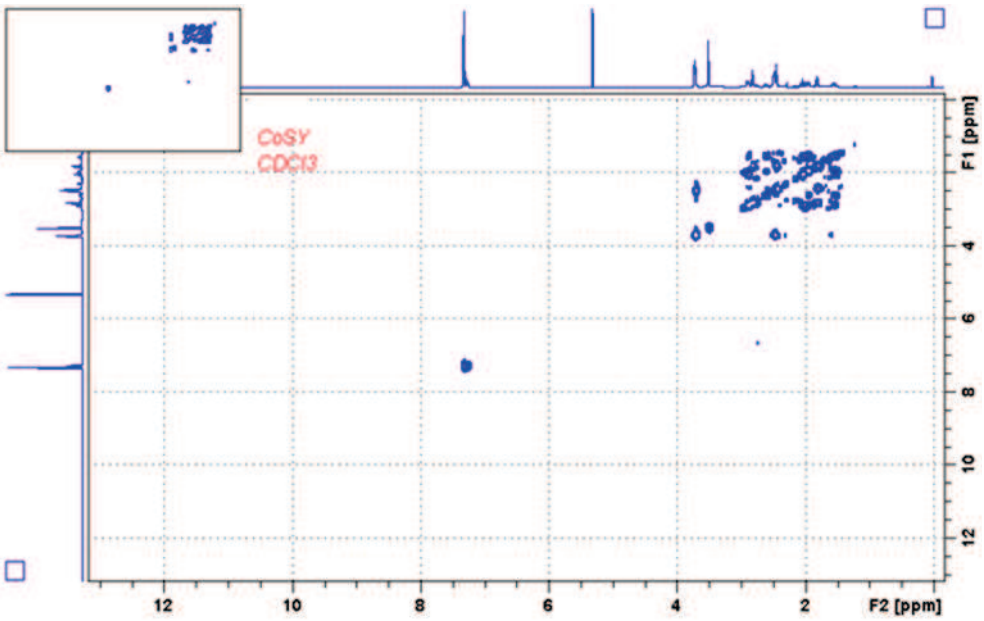


Figure 16. CoSY spectrum of compound **71**.

The experiment of Correlation Spectroscopy shows several cross-peaks (very crowded in the aliphatic part), relative to the scalar coupling of the protons belonging to the molecule; being more specific, the following signals are detected (values in δ, ppm):

- 2.89 – 2.04
- 2.81 – 1.81
- 2.61 – 1.55
- 2.61 – 1.95
- 2.45 – 1.81
- 2.45 – 3.71
- 2.04 – 1.55.

NOESY (Nuclear Overhauser Effect Spectroscopy) gave the spectrum which is shown in the following figure:

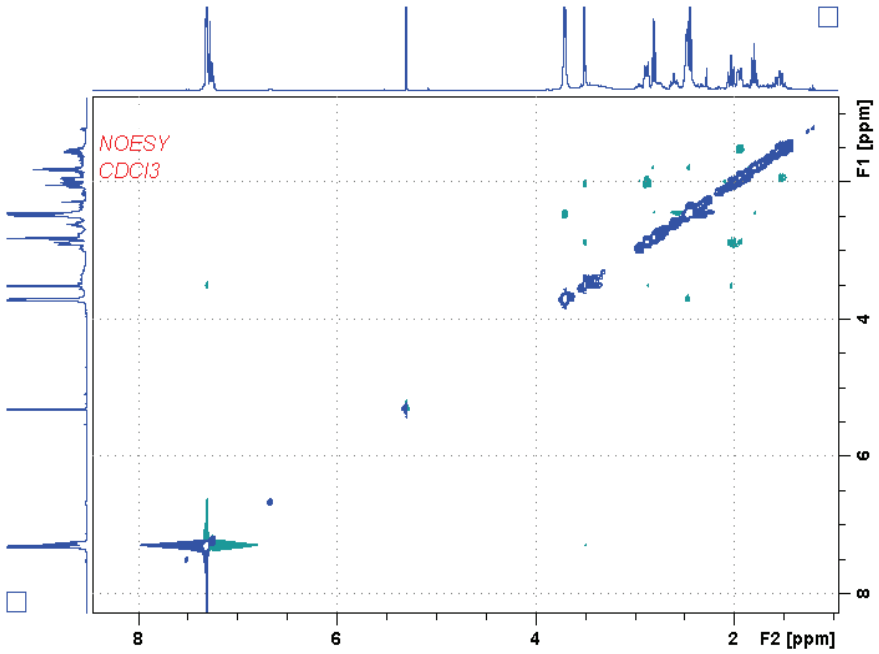


Figure 17. NOESY spectrum of compound **71**.

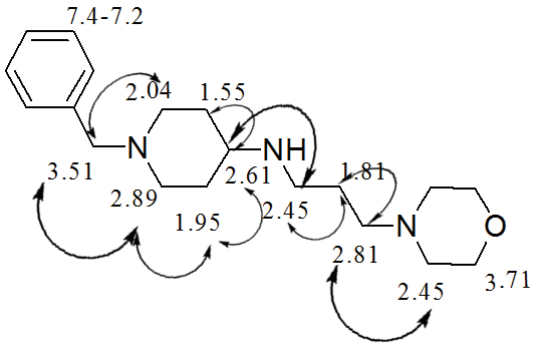
The following considerations can be made:

- The singlet at δ 3.51 ppm cross-peaks both with the methylene at δ 2.04 ppm and with the methylene at δ 2.89 ppm; for its molteplicity, integral and δ, the singlet can be considered to be the signal of the benzylic methylene. This means that the signals at δ 2.04 ppm and δ 2.89 ppm both belong to the piperidine ring, in the ortho positions with respect to the nitrogen atom of the ring itself.
- The signal at δ 2.89 ppm, apart from the previously described cross-peak, also gives a cross-peak with the methylene at δ 1.95 ppm; this means that the signal at δ 1.95 ppm still belongs to the piperidine ring.
- The signal at δ 2.61 ppm (1H) cross-peaks both with the signal at δ 1.95 ppm and with the one at δ 1.55 ppm; these protons are likely to belong to the piperidine ring, in the further position from the nitrogen atom.
- The signal at δ 2.45 ppm (6H, undefined molteplicity) cross-peaks with the signals at δ 1.81 ppm (which also gave a CoSY cross-peak) and with the signals at 2.81 ppm (both integrating for 2H each). These considerations make possible to attribute these six protons to the methylene of

the propyl region of the molecule and to the two methylenes of the morpholine which are closer to the nitrogen atom.

- The signal at δ 2.45 ppm also cross-peaks with the one at δ 2.61 ppm, which was previously hypotesized to belong to the piperidine.

Because of this, the signals can be attributed as follows (the arrows represent the NOESY detectable cross-peaks):



With the aim of enhancing the characterization on the synthesized compounds, and gaining structural pieces of information about the possible conformation of the molecule and the reliability of the predictions that can be made about it, a further study was performed; getting reliable predicted conformations may be relevant because they could be later combined to the pieces of information gained from the structural study of the target, aim which is pursued in this project as well. In this case, the evaluation was made on the proton-proton distances; in fact, for some functional groups, these lengths were measured thanks to UCSF Chimera[80] software in the here reported model, that was minimized in the MMFF94 force field through Avogadro[81] software.

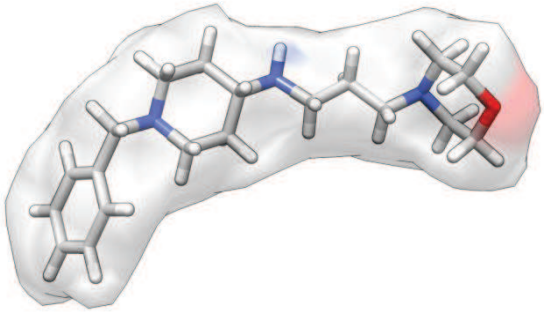


Figure 18. minimized conformation of **71**.

The calculated distances were then compared to the quantitative analysis of the NOESY spectrum signals, according to the formula[82]:

$$r_{ij} = r_{ref} \cdot \sqrt[6]{(I_{ref} / I_{ij})}$$

where

- r_{ij}: distance among two considered groups of protons;
- r_{ref}: reference distance;
- I_{ref}: reference cross-peak integral;
- I_{ij}: integral of the signal corresponding to the considered protons.

The reference protons need to be chosen: in this case, it was arbitrarily decided to consider the cross-peak between the ones at δ 3.51 ppm (benzylic position) and the aromatic ones in ortho position with respect to the benzylic protons themselves (the considered protons are shown in green in figure 19).

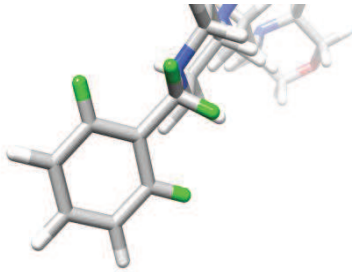


Figure 19. Reference protons for compound **71**.

Actually, it is needed to consider the average of the measure of the distances between the two groups of protons, which computationally gives a theorical value of 3.1 Å (r_{ref}); the value of the corresponding

integral in the spectrum is 76.25 (I_{ref}). The distances between proton groups were measured in the model using the tools that come with UCSF Chimera[80] software; the integrals for the signals of interest, instead, were calculated by TopSpin (Bruker corporation), and are reported in table 6:

Current data set:
Name = AF42 ExpNo = 5 ProcNo = 1
Disk = //LS-WXL91E/RoxyBackup/bkp test/Zagotto/AF/400
Baseline = 0 Noise = 0

#	SI_F1 SI_F2	row1 col1	row2 col2	row1(ppm) col1(ppm)	row2(ppm) col2(ppm)	abs. Int.	Integral	Mode
1	1024	721	736	3.63844	3.42075	-7.9957e+007	-76.253	a
	1024	468	472	7.35049	7.29564			
2	1024	767	782	2.96984	2.75215	-2.0389e+007	-19.445	a
	1024	469	472	7.33335	7.29564			
3	1024	828	838	2.06800	1.92806	-7.8115e+006	-7.4496	a
	1024	470	472	7.32992	7.29564			
4	1024	823	841	2.14680	1.88431	-8.373e+007	-79.851	a
	1024	729	732	3.52309	3.48012			
5	1024	823	845	2.15196	1.82404	-1.6963e+008	-161.77	a
	1024	789	795	2.65086	2.55377			
6	1024	860	873	1.60532	1.41586	-2.9857e+007	-28.474	a
	1024	787	796	2.67367	2.54431			
7	1024	766	778	2.97585	2.80198	-3.9343e+008	-375.2	a
	1024	828	834	2.07791	1.97837			
8	1024	800	805	2.48241	2.41604	-1.4773e+008	-140.89	a
	1024	789	797	2.64264	2.52468			
9	1024	799	807	2.49561	2.38492	-6.0141e+007	-57.355	a
	1024	777	780	2.82638	2.77867			
10	1024	765	781	2.99509	2.76351	-2.0137e+008	-192.04	a
	1024	834	839	1.98062	1.91669			

Table 6: TopSpin calculated integrals.

The proton-proton (non-bonding) interactions were measured *in silico* with Chimera and compared to the values coming from the previously reported formula (inserting the integrated areas of the cross-peaks of interest).

Many evaluations can be made about these values, which are reported in the following table. For the considered bonds, the values found *in silico* and the ones measured by the equation seem to be quite parallel; so, even if this is only one example and the results can still vary, it is probably worthy to consider this kind of approach for further analysis and characterizations.

Coupling signals (ppm)	Area	Calcd. from NOESY length (Å)	*In silico* calcd. lenght (Å)
2.04 - (7.4-7.2)	7.45	4.6	4.7
2.89 – (7.4-7.2)	19.45	3.8	3.7
3.51 – 2.89	61.48	2.8	2.7
3.51 – 2.04	41.41	3.0	3.2
2.89 – 1.95	192.04	2.7	2.6
2.89 – 1.55	176.08	2.6	2.6
1.95 – 2.61	161.77	2.7	2.7
2.61 – 2.45	140.89	2.8	2.8
2.45 – 2.81	57.36	3.3	3.4

Table 7. *In silico* and calculated values for the considered groups of protons.

The reported values also confirm some of the attributions that were already assumed: first of all, that the protons at δ 1.95 ppm have a distance to the ones at δ 2.89 ppm which is compatible with the attribution which has already been made, *i.e.* two of the methylenes belonging to piperidine (in figure 20, the protons of interest are orange).

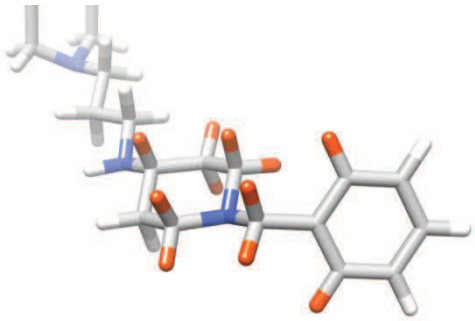


Figure 20. Coupling proton group #1.

In addition to this, it was possible to confirm the attribution of the signals at δ 2.45 ppm (6 H); the measured distances between it and the signals at δ 2.61 ppm (2.8 Å) and at δ 2.81 ppm (3.3 Å) is compatible with the attribution which has been previously made. Keeping in mind that the signal at δ 2.61

ppm was attributed to the only methylidene signal, and that the one at δ 2.81 ppm to the methylene belonging to the propyl chain in the closer position to the morpholine group, in fact, we can confirm that it is likely that the signal at δ 2.45 ppm comprises the methylenes of the morpholine which are closer to the nitrogen atom (cross-peak with the protons at 2.81 ppm), and also the methylene of the propyl chain bonded to the non-endocyclic nitrogen atom (cross-peak with the protons at 2.61 ppm) (in figure 21, the protons of interest are orange).

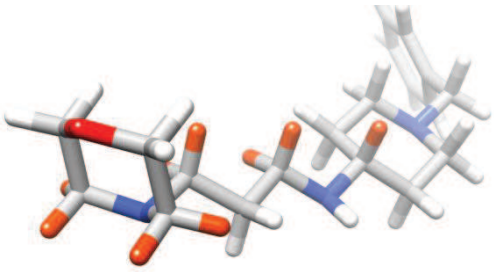
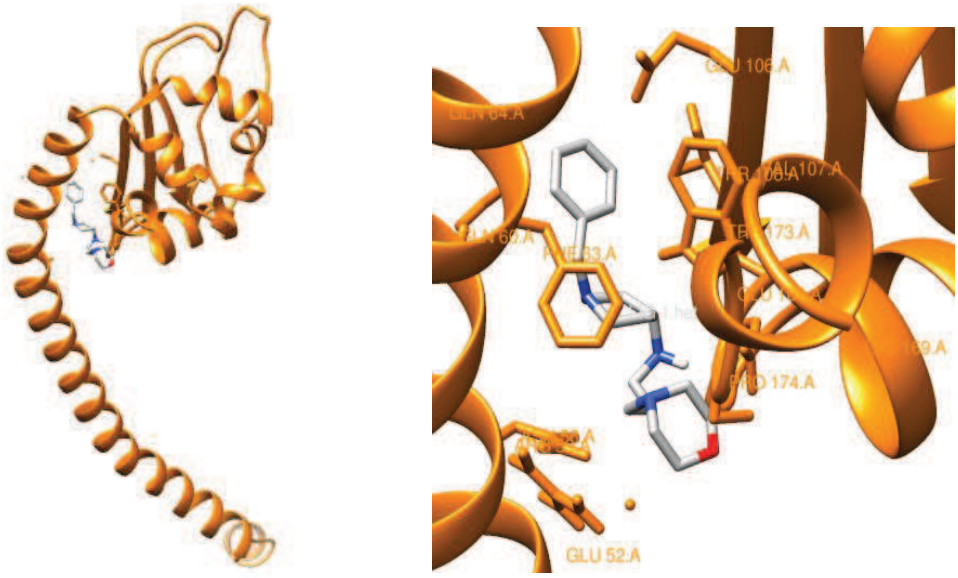


Figure 21. Coupling proton group #2.

A further step on the study of this compound was made through the use of docking tools: basically, the pursued attempt was to compare the computed structure with two other, but this time coming from molecular docking. Giving credit to the hypothesis that Fingolimod binds to SET[38], here the attempt was to figure out if it is reasonable to think that also the synthesized analogs can bind SET, and how they change their conformation after doing that. For pursuing this aim, compound **71** was docked to the monomer SET form (coming from PDB ID 2E50, modified through USCF Chimera[80] in order to remove water and ions) thanks to AutoDock[83], and some proton distances were measured on the docked compound; the lower energy docking outcome is shown in the next figures. The choice of docking the substrate on the monomer first comes both for the sake of simplicity, and for the fact that little is known about the interaction between Fingolimod (or its analogs) and SET, so it is reasonable to exploit both the docking with the monomer and the dimer form of SET. The grid box is placed, in both the cases, on the whole protein. The reported aa residues have a distance inferior than 5 Å from the compound.



(a) (b)

Figure 22. SET monomer/**71** docking (whole, a; zoom, b).

Explicating the hydrogen atoms and isolating the molecule from the protein, but maintaining its conformation, the following result is obtained:

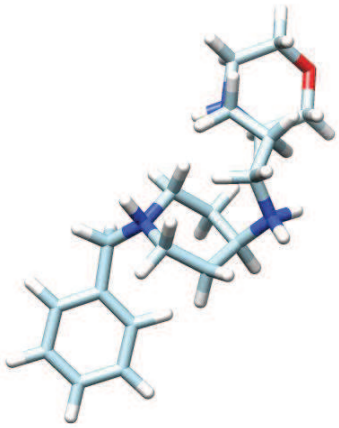
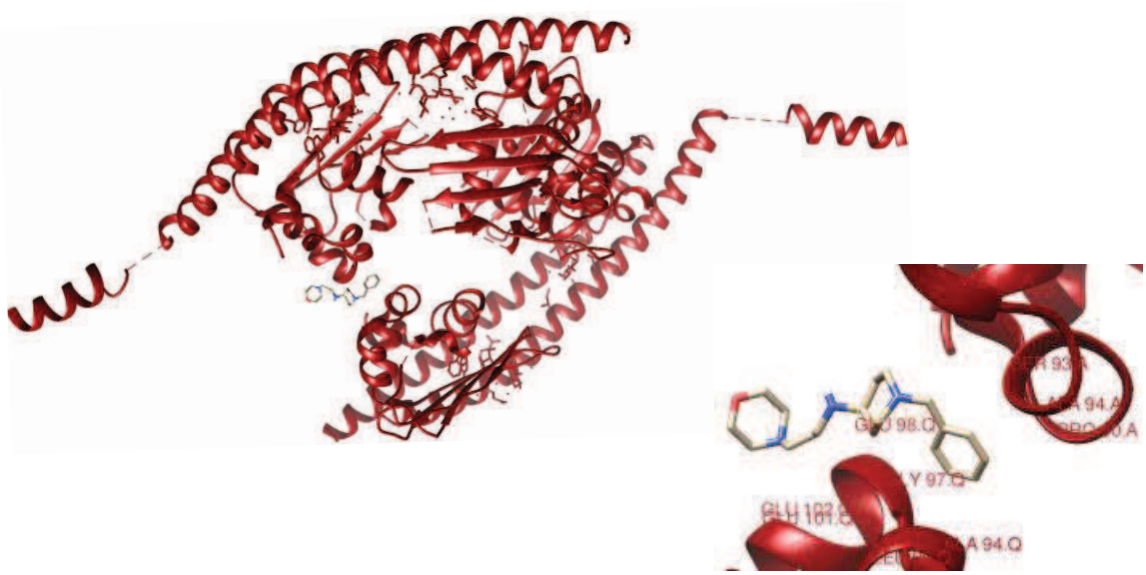


Figure 23. Compound **71** isolated from figure 22.

Since the very first glance, many differences are detectable in the disposition of the atoms in the space, when compared to the structure minimized through MMFF94 (figure 18). Then, the docking with the whole SET (*i.e.* the dimer form) was performed; reported aa residues have a distance inferior than 5 Å from the compound.



(a) (b)

Figure 24. SET dimer/**71** docking (whole, a; zoom, b).

The docking site is different than before, since the molecule is placed between the two monomers that compose SET instead of involving the sharp bend of the monomer. As before, the molecule was isolated from the protein:

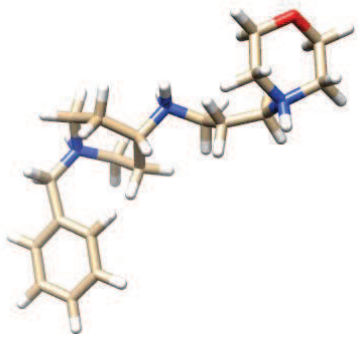


Figure 25. Compound **71** isolated from figure 24.

In both the models, the distances between some proton groups has been measured; this was made through UCSF Chimera, considering the average of the H-H distances for the atoms composing the considered groups themselves. The obtained values were then compared to the previously obtained ones. The

following table shows the outcome of these calculations, and it permits to make a few considerations about them.

Coupling signals (ppm)	Calcd. from NOESY length (Å)	*In silico* calcd. lenght (Å)	*In silico* calcd. lenght after docking (Å) with monomer	*In silico* calcd. lenght after docking (Å) with dimer
2.04 - (7.4-7.2)	4.6	4.7	6.5	6.5
2.89 – (7.4-7.2)	3.8	3.7	5.2	5.3
3.51 – 2.89	2.8	2.7	3.1	3.1
3.51 – 2.04	3.0	3.2	3.6	3.6
2.89 – 1.95	2.7	2.6	2.6	2.5
2.89 – 1.55	2.6	2.6	4.3	4.3
1.95 – 2.61	2.7	2.7	2.5	2.5
2.61 – 2.45	2.8	2.8	3.8	3.2
2.45 – 2.81	3.3	3.4	2.3	2.8

Table 8. Comparison among *in silico* distances in compound **71** groups of protons (MMFF94, docking with SET monomer, docking with SET dimer).

The distance between the reference protons is not changing in all of the three considered cases (it is computed to be 3.1 Å in all the simulations); however, the other distances seem to differ quite a lot between the MMFF94-minimized structure and the docked ones. Only a couple of values seem to correspond in the three columns, especially concerning the values corresponding to 2.89-1.95 and 1.95-2.61 ppm cross-peaks, all belonging to the piperidine ring. The last two columns, *i.e.* the distances calculated for the two docked structures, are instead extremely similar one to each other, especially in the area that comprises the aromatic and the piperidine ring; the propyl linker part shows more differences. Moreover, the calculated binding energies are exactly the same (-5.7 kcal/mol)[83]. Even if it is too early to make considerations about these dockings, also because of the kind of resulting interactions (all hydrophobic), it is worthy to consider this kind of approach for further evaluations. The limited

conformational changes in the two output docks, at least concerning the considered distances, is an aspect that should be evaluated in the future for its role and reliability.

These data need to be completed with experimental studies before considering them to be trustworthy; the most rational thing to do, once the protein is obtained in a sufficient amount, would probably be an NMR study of the hypothetical interaction between the compound and SET, in order to investigate whether this prediction in the variation of distances, *i.e.* this dramatic change in the structure, is affordable or not. The eventual confirmation of these data would be extremely helpful for improving the rationale of the design of further molecules.

In the following image, the three conformations of compound **71** are superimposed: in grey, the MMFF94-minimized structure; in orange, the compound docked to SET in its monomer form; in dark red, compound **71** docked to the dimer form.

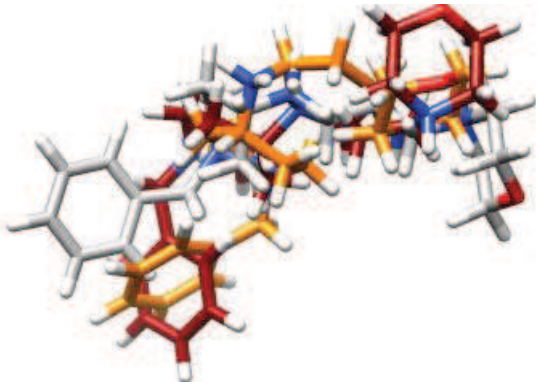


Figure 26. Superimposed **71** structures.

The grey one was not expected to share structural features with the other two, since the distances which were measured are pretty different, as already commented; in addition to this, from a very qualitative point of view, it is possible to see that the orange and the dark red structure develop in the same direction in a more similar manner than the grey one. However, they do not overlap precisely in the expected areas, even if they have comparable distances among H groups.

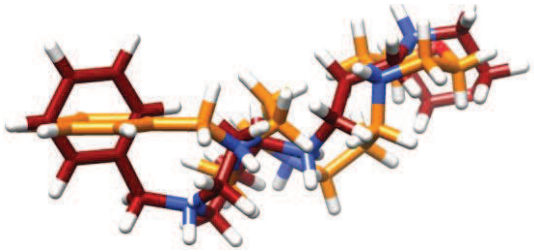
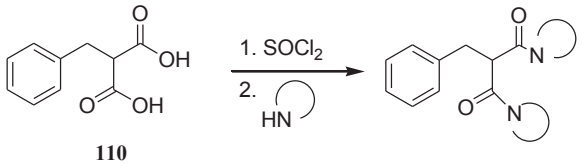


Figure 27. Superimposed **71** structures, excluding the MMFF94 obtained one.

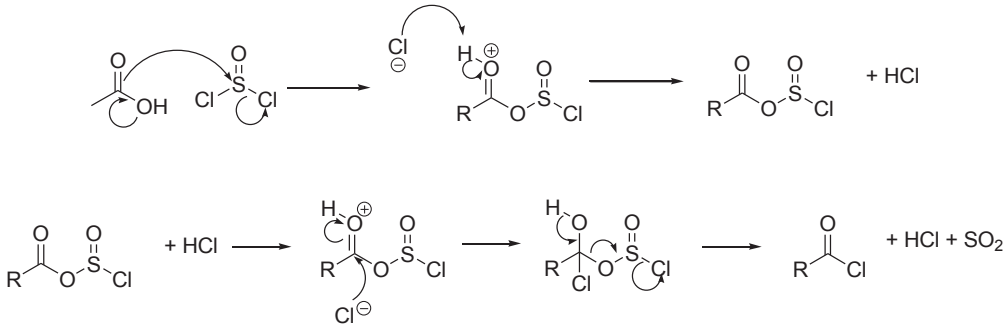
PATHWAY 3:



110

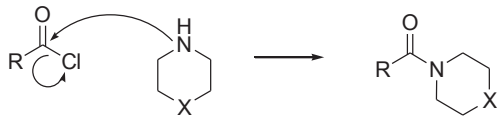
This pathway, which involves phenyl malonic acid, leads to the obtainment of compounds which seem to be slightly detached from Fingolimod structure. Since the main aim of the performed explorative approach presented in this thesis is to provide a library of FTY720 analogs with modulated structural features, the insertion of variables – even if extensive - is reasoned and performed in order to improve the knowledge about the actual groups of interest or the molecules to be active. Here, the effect of the total absence of the aliphatic "tail" in the molecule, while maintaining the aromatic ring which characterizes Fingolimod and most of the analogs reported in this thesis, was the subject of evaluation. In addition to this, the carboxyl acid residues are reacted with morpholine or piperidine for the same reasons explained for pathway 1.1: morpholine owns a non-phosphorylatable oxygen atom, which could help in elucidating the role (or the lack of it) of this atom *per se* and not because it is prompt to phosphorylation. Piperidine is used in the same way as a sort of negative control of this eventual effect.

First of all, the acid needs to be activated to the correspondent acyl chloride through the use of a large eccess of thionyl chloride, which also acts as a solvent for the reaction. The activation is carried on at 60°C and it arrives to completeness in around 1 h. The formation of the product is monitored by TLC, dropping a small amount of the reaction crude in 2 mL of EtOH; in this way the ester of the acid is formed, and it is easier to see it on a TLC plate (hexane/ethyl acetate 4:1) because of its higher R_f when compared to the acyl chloride as it is.



The reaction brings, at the beginning, to the formation of an acyl chloro sulphite: this converts the hydroxyl group, a poor leaving group, into a chloro sulphite group, a more stable anion[77]. After that, a chlorine atom (resulted as a byproduct) attacks the carbonyl carbon for giving an intermediate of reaction which rearranges and gives the acid chloride, sulphur dioxide and hydrochloric acid[77].

The solvent is then removed through a nitrogen flux and the crude can be used without further purifications. At this point, the acyl chloride can be easily attacked by the nitrogen atom of the morpholine (2 equivalents) or of the piperidine (2 equivalents).



The reaction is carried out in toluene in the presence of triethylamine (2 equivalents) in order to buffer the hydrochloric acid which is produced during the reaction.

The obtained compounds are shown in the following table:

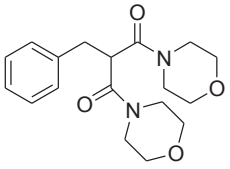
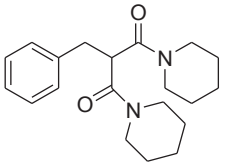
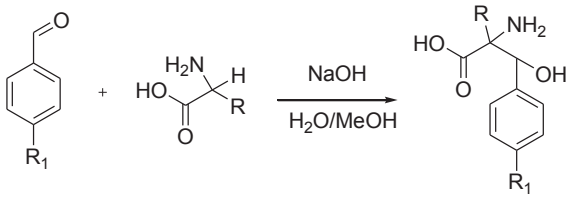
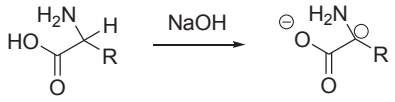
Compound structure and number	IUPAC name
 **112**	2-benzyl-1,3-dimorpholinopropane-1,3-dione
 **113**	2-benzyl-1,3-di(piperidin-1-yl)propane-1,3-dione

Table 9. Compounds obtained with the synthetis pathway 3.

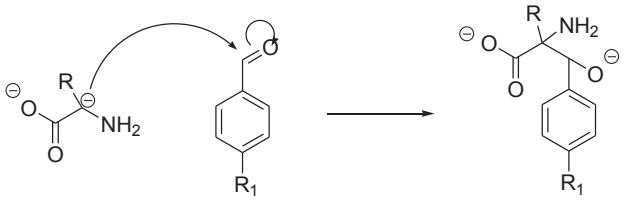
PATHWAY 4:



This reaction was used to insert amino acidic "heads" on the substrates of interest. Maintaining a scaffold which is pretty similar to the already seen ones, and always inspired by Fingolimod, a modified *Knoevenagel* reaction was performed following the procedure reported by *Portelli* [84]; in practical terms, this means that the reaction stops at the nucleophilic attack instead of prosecuting until the dehydration of the product, as the original *Knoevenagel* should provide[77]. Basically, the first step occurs in an aqueous solution of sodium hydroxide in order to deprotonate the α-carbon of the aminoacid of interest (and, of course, its carboxylic group which however is not involved in the reaction).



In this way, the carbanion which is originated can attack an electron-poor aldehydic carbon.

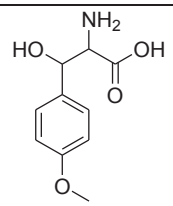
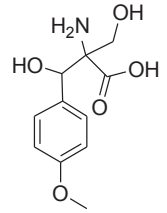
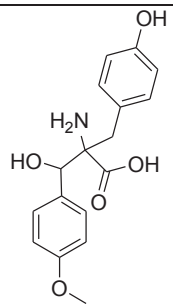
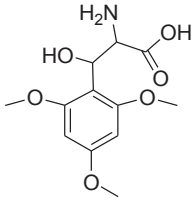


The subsequent addition of hydrochloric acid allows the re-protonation of the hydroxyl groups. It is not probable that the newly inserted hydroxyl groups could be a substrate for SKs, even if placed in a similar position of Fingolimod: in fact, the carboxylic acid is likely to be deprotonated at physiological pH, so the addition of a phosphate group (which carries another negative charge) is not probable.

The first aldehyde taken into account in this pathway was p-anisaldehyde, because the optimisation of the reaction requested a simple scaffold to start from. Since the main issue was to differentiate the phosphorylatability and the polarity of the compounds, the amino acids which have been chosen were glycine, tyrosine and serine; another point which was considered was the content of both heteroatoms and

aliphatic/aromatic portions, and these three amino acids are sufficiently different in all of these characteristics. This reaction was also performed on 2,4,6-trimethoxybenzaldehyde and on previously alkylated (for the procedure, see pathway 1.1) p-hydroxybenzaldehye, using glycine as an amino acid.

The following table shows the obtained compounds:

Compound structure and number	IUPAC name
 **116**	2-amino-3-hydroxy-3-(4-methoxyphenyl) propanoic acid
 **117**	2-amino-3-hydroxy-2-(hydroxymethyl)-3-(4-methoxyphenyl) propanoic acid
 **118**	2-(4-hydroxybenzyl)-2-amino-3-hydroxy-3-(4-methoxyphenyl) propanoic acid
 **119**	2-amino-3-hydroxy-3-(2,4,6-trimethoxyphenyl) propanoic acid

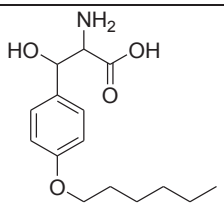
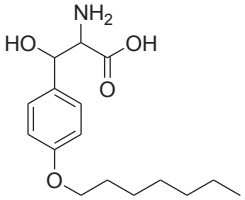
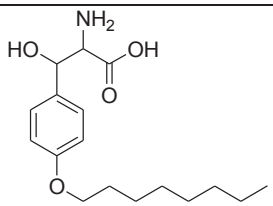
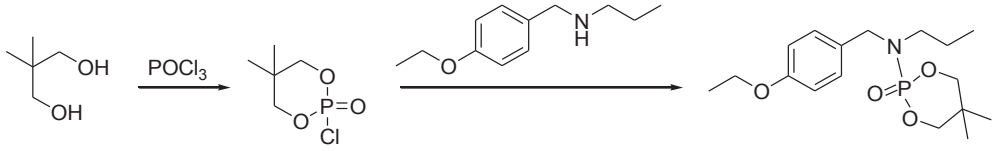
Compound structure and number	IUPAC name
 **120**	2-amino-3-(4-(hexyloxy)phenyl)-3-hydroxy propanoic acid
 **121**	2-amino-3-(4-(heptyloxy)phenyl)-3-hydroxy propanoic acid
 **122**	2-amino-3-hydroxy-3-(4-(octyloxy)phenyl) propanoic acid

Table 10. Compounds obtained with pathway 4.

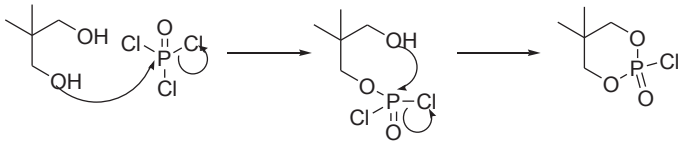
PATHWAY 5:



123 **125** **126**

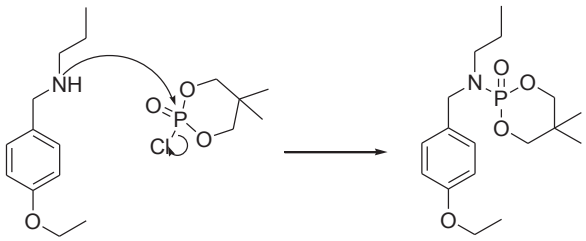
This last pathway provides the synthesis of a cyclic phosphodiester, which is subsequently reacted with one of the products which were already obtained with a previous synthetic scheme (pathway 2.1). This compound can be considered to be a model for further derivatives. In this case, the decision of inserting a phosphorus atom was taken, even if the main aim of these synthetic schemes is to have compounds which are not likely to be phosphorylated. The insertion of this group by purpose is performed because in this way the role of a phosphorus atom could be evaluated *per se*, avoiding the negative charge. Basically, not having the negative charge of the phosphate could be a *pro* because both $PP2A_C$ and SET have an acidic pI (calculated around 5 and 4, respectively[85]), which is not likely to provide a good interaction with a negatively charged molecule (as Fingolimod-P is); to prove this hypothesis, it is worthy to consider the presence of a phosphorus atom lacking the negative charge and to evaluate its activity.

2,2-dimethyl-1,3-propandiol (1 equivalent) is reacted with phosphorus (V) oxychloride (1 equivalent) in Et_2O as described by *Huttunen*[86] to obtain the cyclic phosphodiester through the following mechanism:



The hydroxyl groups can, one after the other, attack the phosphorus atom and lead to the exit of Cl^- ions. The presence of triethylamine (2 equivalents) in the reaction mixture buffers the pH, which could decrease because of the production of hydrochloric acid.

The remaining chlorine atom of the phosphodiester (2 equivalents) can be easily substituted by a secondary amine as the one which is present on compound **86** (1 equivalent), which was synthesized following the synthetic pathway 2.2.



Once again, Cl^- acts as a leaving group; the reaction is carried out in toluene, and the addition of calcium carbonate (2 equivalents) buffers the acid which is obtained as a byproduct.

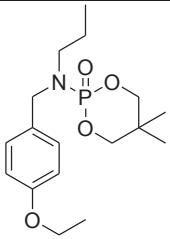
Compound structure and number	IUPAC name
 **126**	2-((4-ethoxybenzyl)(propyl)amino)-5,5-dimethyl-1,3,2-dioxaphosphinane 2-oxide

Table 11. Compound obtained with pathway 5.

2.2 RATIONALE OF THE RECOMBINANT PROTEIN PRODUCTION

This part of the project, performed at the laboratories of Diamond Light Source and the Research Complex at Harwell, Didcot, Oxfordshire (UK), under the supervision of Dr. Marco Mazzorana aimed at the production of recombinant proteins to be used for biochemical and biophysical assays as well as for structural characterization of the PP2A-SET interaction. Well-estabilished molecular biology protocols[87,88] were optimized for this specific case to obtain large quantities of purified proteins, especially needed for christallographic experiments.

2.2.1 High throughput cloning of $PP2A_C$ and SET

The first step for obtaining the recombinant forms of the proteins consisted in the amplification of cDNA encoding for SET or $PP2A_C$ through PCR (Polimerase Chain Reaction). The genes were inserted in pOPIN expression vectors available from OPPF-UK (Oxford protein production facility)[89]. This choice allows to perform all the phases of cloning, expression and purification tests with a high throughput format using robotic instrumentation and support. In addition, the selected vectors used for overexpression in *E.coli*, is also suitable for expression in insect cells with the baculovirus technique as well as for transfection in cultured human cells. Overexpression in bacteria or insect cells is required for structural investigation such as chrystallography, and the possibility of transfection of human cells with the same vectors opens the possibility of using the same constructs to gain a deeper understanding of the physiological role of the proteins as well as for *in vitro* PP2A activation studies. Most of the chosen expression tags can also be easily and precisely cleaved during the purification steps by treatment with a His_6-tagged human rhinovirus 3C protease, which is easy to produce and to remove once it is no more needed. The cloning strategy for SET, starting from the vector provided by the group of Prof. Masami Horikoshi, maintained the additional His_6 tag and thrombin cleavage site at the N-terminus of the protein sequence. As for $PP2A_C$, the cDNA sequence was designed applying a codon-optimization algorithm to avoid the usage of rare codons and the company Genscript synthesized the whole sequence ex novo within a carrier vector we used as a template.The chosen vectors provide the coding for the POI with six different tags, resulting in 12 constructs as shown below:

Construct A5 - pOPIN-F-PP2A$_C$ (His$_6$-3C-PP2A$_{C,1-245}$)

Construct B5 - pOPIN-J-PP2A$_C$ (His$_6$-GST-3C-PP2A$_{C,1-245}$)

Construct C5 - pOPIN-M-PP2A$_C$ (His$_6$-MBP-3C-PP2A$_{C,1-245}$)

Construct D5 - pOPIN-S3C-PP2A$_C$ (His$_6$-SUMO-3C-PP2A$_{C,1-245}$)

Construct E5 - pOPIN-E-PP2A$_C$ (PP2A$_{C,1-245}$-His$_6$)

Construct F5 - pOPIN-eGFP-PP2A$_C$ (PP2A$_{C,1-245}$-3C-eGFP-His$_6$)

Construct A6 - pOPIN-F-SET (His$_6$-3C-His$_6$-Thrombin-SET$_{1-225}$)

Construct B6 - pOPIN-J-SET (His$_6$-GST-3C-His$_6$-Thrombin-SET$_{1-225}$)

Construct C6 - pOPIN-M-SET (His$_6$-MBP-3C-His$_6$-Thrombin-SET$_{1-225}$)

Construct D6 - pOPIN-S3C-SET (His$_6$-SUMO-3C-His$_6$-Thrombin-SET$_{1-225}$)

Construct E6 - pOPIN-E-SET (SET$_{1-225}$-His$_6$)

Construct F6 - pOPIN-eGFP-SET (SET$_{1-225}$-eGFP-3C-His$_6$)

The choice of vectors allowed placing a hexa-histidine tag both at the N-terminus (pOPINF, -J, -M and S3C) and at the C-terminus for purification on immobilized metal affinity chromatography (IMAC). In addition to this, solubilizing tags such as glutathione S-transferase (GST), maltose binding protein (MBP) and SUMO (Small Ubiquitin-like Modifier) were added to improve expression and solubility, while eGFP (endothelial Green Fluorescent Protein) was designed to provide a fluorescent probe for *in vivo* experiments. The cDNA coding for PP2A$_C$ and SET was amplified by PCR and integrated in the host vectors (previously linearized with KpnI and HindIII restriction enzymes) using the InFusion technology (Clontech Laboratories, Inc.). The cloning mixtures were then used to transform One Shot OmnimaxII supercompetent *E.coli* cells (Invitrogen), a high competency strain which can yield above 10^9 colonies per μg of plasmid. Omnimax also contain a deletion for the endogenous □-galactosidase gene (lacZ), allowing for blue/white screening of colonies; in the absence of the inserted gene, pOPIN plasmids complement the lacZ gene, allowing the bacteria to degrade X-gal to a dark blue product and resulting in . blue colonies when bacteria are cultured on solid media selective for the vector (presence of carbenicillin) and X-gal. If the gene of interest has been inserted in the vector, this cannot complement lacZ and the colonies will remain white. In this way colonies can be isolated, amplified by growing the bacteria in liquid medium (usually LB or PB broths), and the plasmid can be purified on ion exchange magnetic DNA mini-prep columns using the Bio-Robot 8000 (Qiagen) of OPPF-UK equipped with magnetic beads (Agencourt AMPure XP, Beckman Coulter).

The presence of the desired construct can be verified through a hybrid PCR, which uses a forward primer specific for the vector (pOPIN-Forward) and a reverse primer specific of the cloning. Qualitative assessment is then possible by running the PCR products on 1% agarose-TAE gel electrophoresis to show the presence of an amplification product of the expected size (given by the sum of sizes of the tag and of the gene of interest). As shown by the following gel (Figure 28), all clonings gave satisfying results and the purified plasmids were sequenced to confirm that the insert was in a frame with the expression vector and did not contain any undesired mutations.

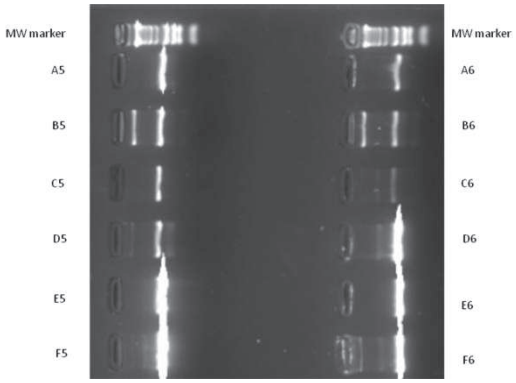


Figure 28. HTP mini-prep verification through agarose electrophoresis.

Stocks of the positive clones were stored in 100 µL aliquots at -80°C as glycerol stocks (30%w/v glycerol final) to conveniently archive a stable source of plasmid for future needs.

2.2.2 Recombinant protein expression

As mentioned above, the chosen vectors allow the expression of recombinant protein in the most commonly used bacterial system, *E. coli*. Once a suitable expression strain is identified from small-scale cultures, the production of the protein can be scaled up to litres of cultures, usually yielding several milligrams of protein. Some targets, however, might not get expressed in prokaryotic cells or need post-translational modifications which are only available in higher complexity systems. The vectors of the pOPIN suite have been chosen to expand the possibility of expression systems to insect and human cells in culture. This choice has revealed to be particularly important for $PP2A_C$, which has been shown to express using the baculovirus technology in Hi-5 cells[90].

The purified plasmids containing the genes of interest were then used to transform two strains of *E. coli*, optimized for the expression of proteins: Rosetta2 (DE3), especially suitable for the expression of human proteins, and Lemo21 (DE3), a BL21 (DE3) analogue providing fine tuning of expression and recommended for proteins showing toxic effects in overexpression. Colonies obtained were grown in high-throughput format and the protein expression was either induced by addition of IPTG (in PB media) or auto-induced (AI medium) overnight. These cells wereused for small-scale IMAC purification on Ni-NTA magnetic agarose beads (Qiagen) through the Bio-Robot 8000; once again, a library of glycerol stocks was produced.

The presence of the proteins in the purified soluble fraction of bacterial lysates was assessed by SDS-PAGE, which showed that two of the constructs for SET were obtainable in good quantities: C6 (pOPIN-M construct, carrying the solubilizing His_6-MBP tag upstream of the protein sequence) and D6 (pOPIN-S3C construct, with a solubilizing His_6 SUMO tag at the N-terminus). The results were satisfying both for Lemo21 and for Rosetta2 cells, and both for the cells grown in PB and AI media. In the case of construct C6, which is supposed to have a MW of about 70 kDa (SET+MBP), the gel shows a single band just above 66 kDa; D6 shows a band between 45 and 55 kDa, confirming the presence of the construct should have a MW around 43 kDa (SET+SUMO). The electrophoretic mobility of this last construct is decreased due to the high content of acidic residuesof both SET and SUMO (pI~4), resulting in a higher apparent MW for D6. This characteristic is not so evident for C6.

For what concerning the expression of $PP2A_C$, no samples gave satisfying results in *E. coli*: this is not totally a surprise, since the structures reported in literature for $PP2A_C$ [91,92] have been obtained from samples expressed by baculovirus. In addition to this, many of them are co-crystallized at least with a regulatory subunit of the B family (*e.g.* PDB ID 2NPP, 3C5W, 3DW8, 3FGA and more) which is likely to provide some degree of stability to the whole structure. The choice of pOPIN vectors shows here its rationale, of opening to further screening of the construct in insect cells, which will be described in more detail later.

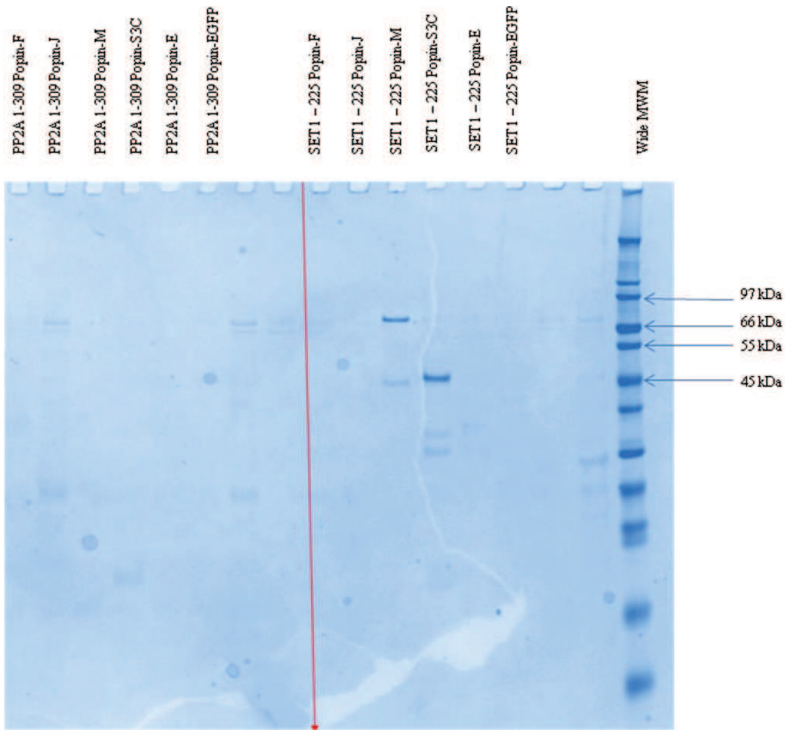


Figure 29. HTP by Ni^{2+}-NTA of the obtained proteins (Rosetta2, AI medium).

2.2.3 Scale up of expression

These small scale results, verified by SDS-PAGE, lead to the selection of the samples for the scaling-up; in the preliminary phase, few mL of cell culture are lysed and protein purified by affinity only. Only the products showing high concentration, the correct MW and the absence of degradation are considered for scale-up usually at the 1÷10 L culture scale. After evaluating these factors, Rosetta2 cells transformed with clones C6 and D6 were considered for the scale-up, since they displayed the best overall expression levels. The corresponding samples from glycerol stocks were cultured overnight in medium scale to provide a pre-culture suitable as a starter for the expression of 12 x 1 L of autoinduction medium[93]. Once large-scale cultures were harvested, cells were stored at -80°C until the day of purification and overexpression of the desired was verified via SDS-PAGE on the total cell content, compared to that of cells not harbouring the plasmid for the protein of interest.

2.2.4 Purification strategy

The purification protocol applied here can be considered a general method for protein purification, and it involves three main steps:

1. IMAC purification: this exploits the capacity of polyhistidine tags to coordinate divalent cations. Usually the protein carrying a His_6-tag is bound to a polysaccharide matrix functionalized with nitrilotriacetic acid and chelating Ni^{2+}. Elution is usually performed by competition with imidazole in high concentrations.
2. Tag cleavage and removal: the His_6-tag used to purify the sample, sometimes also supplemented by a bulky to solubilize the protein of interest, needs to be removed. The constructs have been designed to integrate a cleavage site for thrombin or for the rhinovirus 3C protease. These enzymes recognize specific sequences within the protein and cleave precisely at that level. A mild overnight treatment of the fusion protein is sufficient to convert this into the untagged protein of interest, separated from the tag. This can be removed by a further reverse-affinity step or exploiting the difference in size with the protein of interest.
3. Size-exclusion (or gel filtration) chromatography: partially purified protein mixtures can be conveniently separated in their key components due to their differences in size. The retention is longer for the smaller molecules, which elute slower, than for the bigger ones, which remain in the column for shorter times. Given the difference in size between SET or $PP2A_C$ and the tags provided by the cloning vector, a convenient separation strategy would be that of gelfiltration. This also allows to remove uncleaved proteins, whose MW results from the sum of that of the protein of interest and of the tag, as well as to exchange the buffer to the one selected for protein storage.

Proteins are thus concentrated to 10÷100 mg/mL, divided in small volume aliquots (25 µL), flash frozen in liquid nitrogen and stored at -80°C.

2.2.5 Purification of Sumo-tagged SET (D6 deriving samples)

The purification protocol for His_6-SUMO tagged SET from Rosetta2 cells was estabilished in three major steps:

- Affinity chromatography on HisTrap column (GE Healthcare);
- Dialysis and tag cleavage (3C protease treatment);

- Size exclusion chromatography.

The affinity chromatography purification exploited the hexa-histidine stretch placed at the N-terminus of the SUMO tag itself, conferring to the protein affinity for immobilized Ni^{2+}. The purificationccolumn is a prepacked matrix of Ni^{2+}-immobilized Sepharose, which binds the His-tagged proteins selectively, allowing non-tagged proteins to be washed off. The His_6-tagged proteinis conveniently detachted by addingto the buffer a sufficient amount of imidazole to compete for the Ni^{2+}coordination sites (250 mM). During the chromatographic steps, performed on an automatic FPLC system (GE healthcare) the presence of protein was indicated by the level of absorbance at 280 nm as shown in Fig. 30. In the case of the purification reported, washes were performed by stepwise increases of imidazole concentration in the buffer and elution of the protein was obtained at 250 mM imidazole.

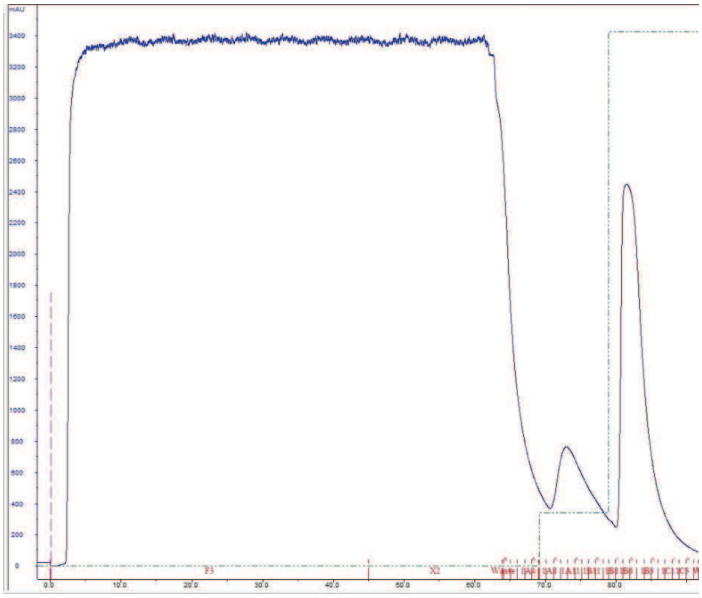


Figure 30. Ni^{2+}-NTA purification of D6 samples. In blue, the absorbance; dotted in green, the percentage of NPI-250.

The promising fractions were then analyzed by SDS-PAGE as reported in Fig. 31. A protein displaying the expected MW for the protein of interest is clearly present throughout the elution peak. The same protein also elutes in the wash step, probably because the capacity of the column was not sufficient to retain all molecules of SET present in the lysate.

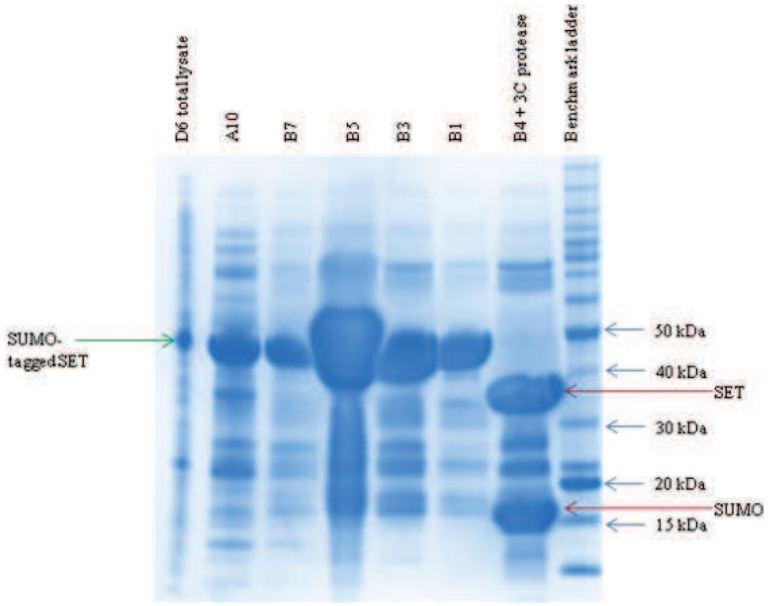


Figure 31. Purification result and cleavage attempt.

The fractions from the peak were treated overnight with 3C protease at 4°C to remove the tag. This resulted in the cleavage of the protein into smaller ones which in SDS-PAGE displayed a lower MW and correspond respectively to SET (28 kDa, apparent MW 35 kDa due to the acidic pI) and SUMO (15 kDa), (as shown in Fig. 31 – lane "B4 + 3C protease").

Tag cleavage was performed while dialyzing the protein against a buffer without imidazole, in order to reduce the concentration of the eluting agent which often decrease protein stability in the long term.

The last purification step consisted in a size-exclusion chromatography on a crosslinked dextran matrix.

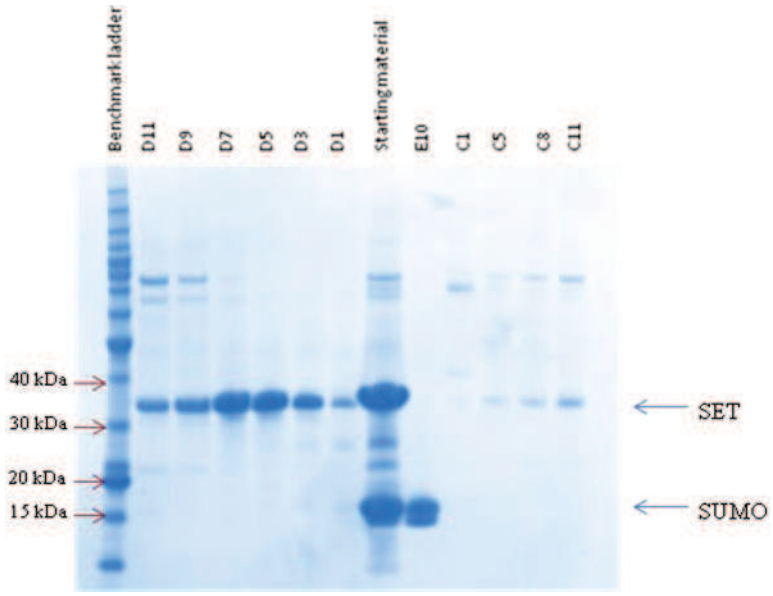


Figure 32. Gel filtration result.

SDS-PAGE clearly shows that the protein had been purified (D7 to D1 in Fig. 32), and SUMO was efficiently separated (E10 in Fig. 32). Higher molecular weight bands were visible in the gels for the less pure fractions, discarded as not sufficiently pure although the bands are possibly due to the spontaneous formation of SET multimers. The retention volume of SET was used to estimate the molecular weight from the gelfiltration column calibration profile and strongly suggests that the protein in fractions D7-D1 is in a dimeric state. This is confirmed by previous literature[94] that describes SET as a protein which exists in solution, and crystallizes (PDB ID: 2E50), as a dimer. The central fractions of the peak were estimated to be >95% pure from band densitometry and the total concentration of SET was measured by its absorbance at 280 nm by NanoDrop in 0.35 mg/mL (12 μM). This corresponds to 2.5 mg of clean SET in the 7 mL of purified sample, which was concentrated to approx. 10 mg/ml, aliquotedto 25 μL and stored at -80°C.

To assess the real MW of the protein, the stored sample was subjected to HRMS: being more specific, SET was concentrated at 25 μM in a solution of ammonium acetate 150 mM. In this case, the mass spectrum was run using positive ionization in denaturant conditions (acetonitrile/0.1% aqueous formic acid, 1:1 ratio). The obtained spectrum was automatically deconvoluted and shows that the main signal is the one corresponding to SET (28599 Da).

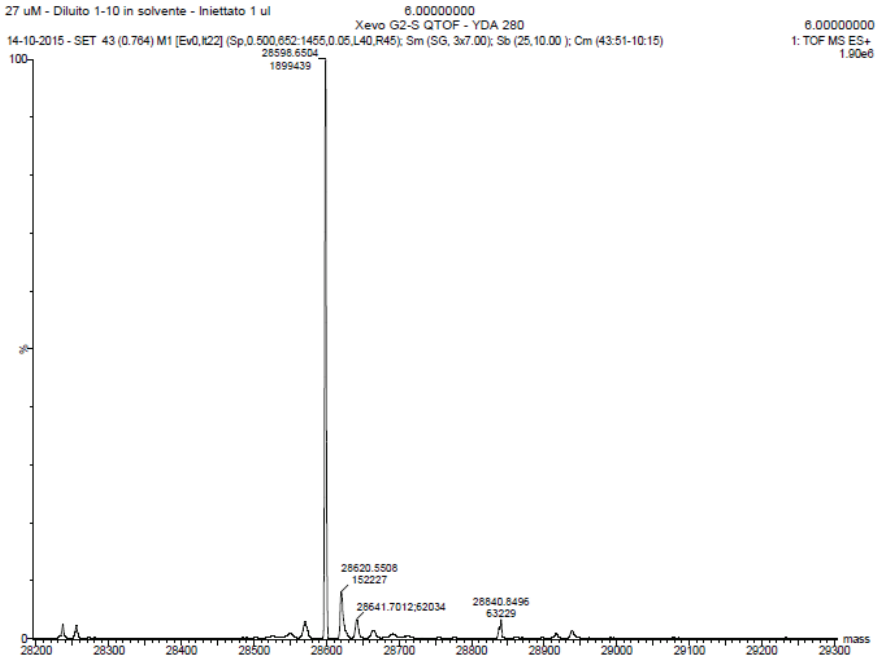


Figure 33. Deconvoluted mass spectrum for D6-derived SET.

2.2.6 Purification of MBP-tagged SET (C6 deriving samples)

The purification of protein expressed by the C6 clones Rosetta2 cells, was initially performed as described for the SUMO-tagged version of the protein. Despite the large amounts of initial material (band densitometry of total lysate expression gels shows C6 expressed at least 3 times more protein than D6), the same purity was hardly achieved. This is particularly clear from the gel filtration step: after tag cleavage, both the chromatograms and the SDS-PAGE analysis show co-elution of SET with MBP. The reason for this is clear if we consider again the spontaneous dimerization of SET, whose dimer has comparable MW to that of the MBP tag.

As shown in Fig. 34, SET elutes earlier than MBP. This supports the idea of its dimeric state: given that SET dimer should have a MW of around 56 kDa (2x28 kDa), this elutes earlier than MBP, which has a MW of 42 kDa. To confirm that SET behaves in solution as a dimer, analytical gelfiltration on a column calibrated with MW standards and coupled with a MALS detector, showed that the peak of a purified protein corresponds to an object of approximately 55 kDa.

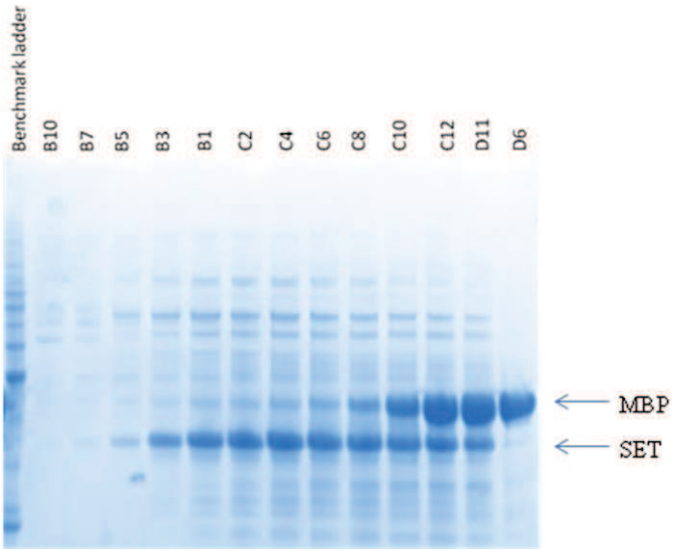


Figure 34. Gel filtration result on C6 samples.

The sample purified following the strategy designed for SUMO-tagged SET remains heavily contaminated with what is believed to be proteins aspecifically bound to the HisTrap column. Because of this, purifications were repeated in different conditions to try improving the purity of the final sample but with no success. Increasing of the salt concentration in the affinity or gel filtration buffers to contrast electrostatic interactions between proteins and resin, as well as a second affinity column for MBP did not produce better quality samples. Also also a Co^{2+} affinity chromatography was attemped, exploiting the lower affinity of Co^{2+} for His-tags (when compared to Ni^{2+}) in the attempt of getting rid of aspecifically binding proteins, but also this approach was not successful.

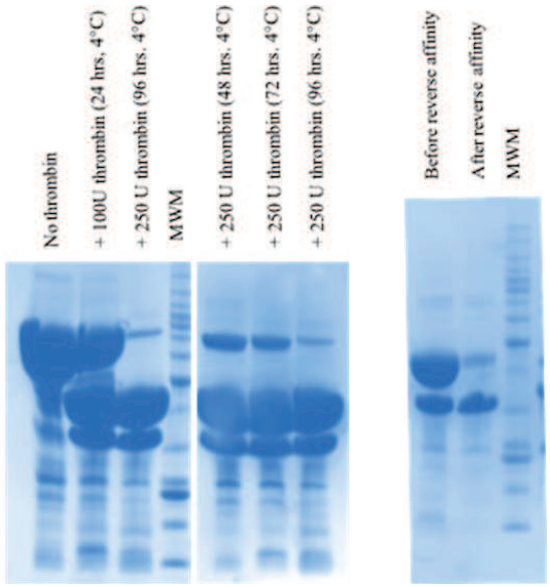
The final breakthrough was achieved by two modifications to the protocol:

- Increasing the starting material to limit aspecific binding at early stage;
- Tag cleavage with thrombin, followed by reverse affinity purification.

Since SET has been reported to crystallize from solutions 70 mg/mL, the aim of obtaining at least 100 mg was pursued, so 12 x 1L cultures of Rosetta2 in autoinduction media were settled up and purified. This required scaling up the amount of resin used: 5 mL columns still showed lots of His_6-MBP-SET in the flow through, and for this reason the flow through was loaded repeatedly on re-equilibrated 5 mL HisTrap columns and eluted with imidazole until the SDS-PAGE analysis showed no more protein in the unbound fraction. This procedure was completed with 8 passages and the total volume of eluted protein was approximately 300 mL with an average concentration estimated in 5 mg/mL (1.5 g protein overall).

In order to remove the tag efficiently, the procedure was modified and rhinovirus 3C protease cleavage was modified to thrombin treatment. In this way, the construct His_6-MBP-3C-His_6-thrombin-SET_{1-225} could effectively be separated from both polyhistidine tags, eventually removed by means of reverse affinity, passing the proteolysed mixture on the same HisTrap affinity resin used at the beginning of the purification. For this step, however, the competing imidazole in the elution buffer needs to be removed from the protein solution to allow His-tagged constructs to bind the resin.

Given the large amount of eluted protein, the most efficient way was to dyalise the fractions eluted from HisTrap against a buffer containing no imidazole, while performing the cleavage with thrombin. The sample was pooled and dialysed twice overnight against 7 volumes of buffer to reduce the imidazole concentration to approx. 5 mM (7x7 folds dilution from initial 250 mM). As shown in Fig. 35a, the cleavage with thrombin was monitored by SDS-PAGE and required a very large amount of enzymes (over 200 units added also at the end of the dyalisis process). Finally, the completely proteolysed sample was passed on a 40 mL HisTrap column refluxing for 24 hours to bind all the tagged proteins (Fig. 35b).



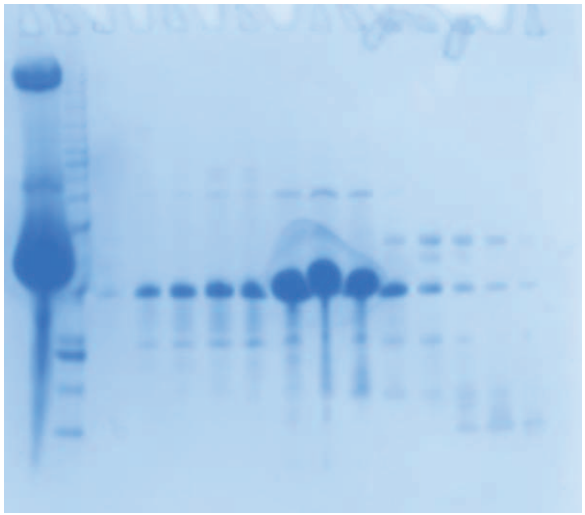
(a) (b)

Figure 35. SDS-PAGE analysis of the cleavage of the obtained construct with thrombin at different concentrations (a), and reverse affinity outcome (b).

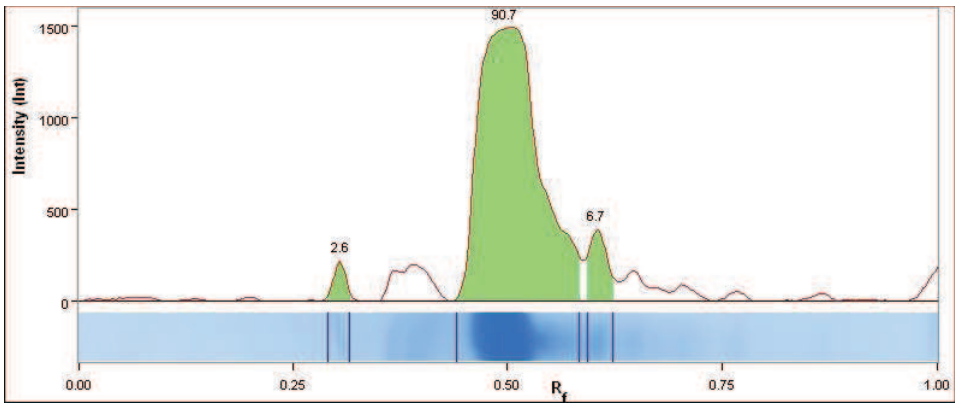
The sample obtained through this procedure was much purer than in the previous attempts of purification and required only a final polishing step by gelfiltration to remove small amounts of impurities including thrombin, which not being His_6-tagged cannot be eliminated through reverse affinity chromatography.

Once again, the large amount of sample could not be loaded in the size-exclusion column as such, but needed to be concentrated at least 10 times. Given the limited capacity of the column used (max 5 mL of protein solution and less than 100 mg), the sample was split in 6 aliquots each one purified in successive gel filtration runs. The purity of the sample, quantified by SDS-PAGE band densitometry is >90% (Fig. 36) and perfectly comparable to the previous report by Muto and co-workers[95].

The purified sample was pooled and concentrated to approximatively 4 mL and the concentration of the protein was estimated in 84 mg/mL from measurements of absorbance at 280 nm. The sample, divided in 42 x 100 µL aliquots was frozen in liquid nitrogen and stored at -80°C for structural and functional analysis.



(a)



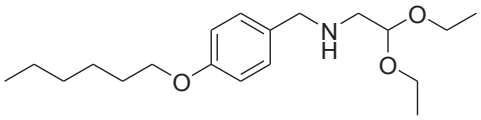
(b)

Figure 36. SDS-PAGE of gel filtration fractions (a) and quantitative evaluation of the resulting bands (b).

2.3 RESULTS AND DISCUSSION

2.3.1 Biochemical assays on the synthesized compounds and perspectives

During these years a library of compounds was synthesized and characterized. The molecules have been designed keeping in mind Fingolimod structure, with the aim to isolate its pro-apoptotic and potentially anti-tumor effect from the immunosuppressive one. This meant maintaining most of the times the aromatic di-substituted ring, modulating structural characteristics as the lenght of the "tail" (or even its presence/absence), or the polarity, hindrance and, above all, phosphorylatability of the "head". The variations taken into account brought to the creation of a panel of molecules, sometimes quite dissimilar one to the other, for obtaining from them the most in terms of structure-activity relashionships (SAR). Many of the compounds were assayed at the Department of Molecular Medicine, University of Padova, in the laboratory of Prof. Brunati. The most promising results were obtained with the compound **100**, published as MP07-66 (N-(4-(hexyloxy)benzyl)-2,2-diethoxyethanamine)[70]:



MP07-66

PP2A/SET complex was isolated from B cells from 42 untreated CLL patients and 6 normal volunteers, through anti-PP2A polyclonal antibody precipitation. The complex was then treated with increasing concentrations of different compounds, and the dissociation between PP2A and SET was evaluated through immunoprecipitation. As can be seen (figure 37), this compound can lead to the disruption of the complex.

Figure 37. Evaluation of the disruption of PP2A/SET complex.

The following step was the evaluation of the restoration of PP2A activity after the disruption of the complex. For doing this, the PP2A/SET complex treated with the compounds was evaluated for the capability of dephosphorylating the anti-apoptotic enzyme p-Akt, one of the main substrates of PP2A; once dephosphorylated, the enzyme is inactive and cannot perform its anti-apoptotic effect. As shown by densitometry on Western blot (figure 38), the increasing of the concentration of compound MP07-66 leads to the dephosphorylation of Akt and this can be considered an indirect proof of the reactivation of PP2A.

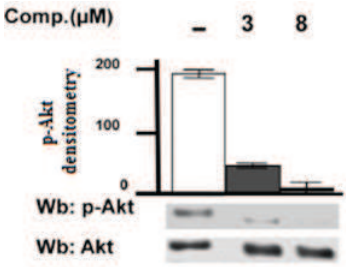


Figure 38. Evaluation of the reactivation of PP2A through Western blot.

Finally, the factual effect on the whole cell was tested in order to see if its capability to die is restored after the treatment with the compound of interest. So it is, since the percentage of apoptotic cells increases with the increasing of the concentration of MP07-66 (figure 39). All of this makes possible to say that this compound can destroy the complex between PP2A and its inhibitor SET, and this fact provokes the reactivation of PP2A and ends with the death of the cell.

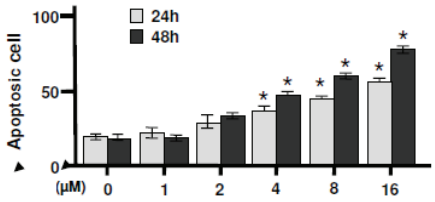
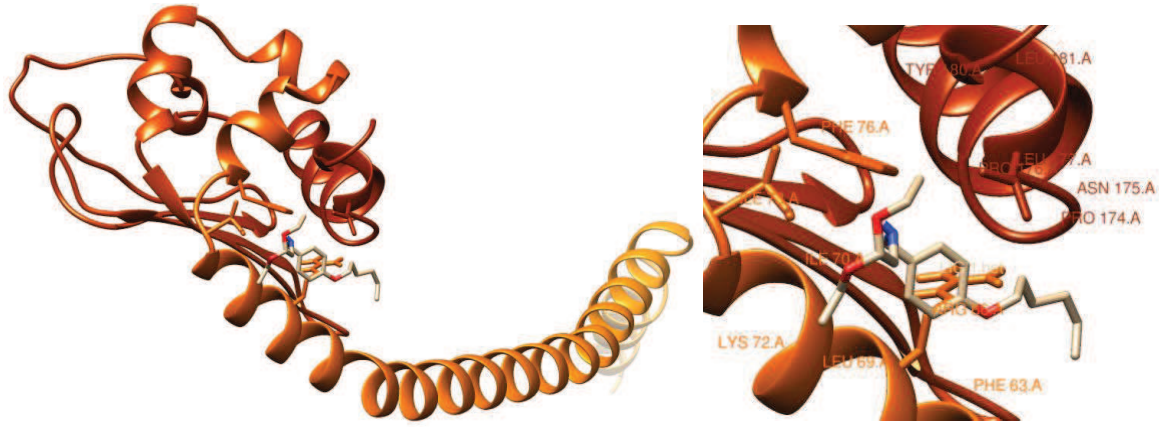


Figure 39. Percentage of apoptotic cells vs MP07-66 concentration.

In order to integrate the pieces of information obtained through these assays, the co-crystallization between the compound and the target would be a real improvement; basically, the main hypothesis about

the interaction between Fingolimod and PP2A/SET complex is that Fingolimod acts on SET and binds it[38], but it still has to be confirmed. In addition to this, the analogs described in this thesis could act in a different manner than the molecule they are inspired to. Obtaining the POIs in a sufficient amount to try to co-crystallize them with the small molecules, could help in elucidating this point.

The obtainment of the actual structures could also have another consequence, *i.e.* approaching molecular modeling tools for the design of future compounds. The lack of structural data about PP2A/SET complex and the way it potentially interacts with other substrates, indeed, renders the molecular docking quite difficult to perform, and in these conditions its reliability could be a matter of discussion. In the beginning, the dock between SET (PDB ID: 2E50) and MP07-66 was performed through the use of AutoDock[83] and elaborated with UCSF Chimera[80]. SET protein structure is registered as a dimer on PDB, but for the sake of simplicity at the very beginning the monomeric form of SET was considered. The grid box was placed on the whole monomer, since the docking site is not known yet.

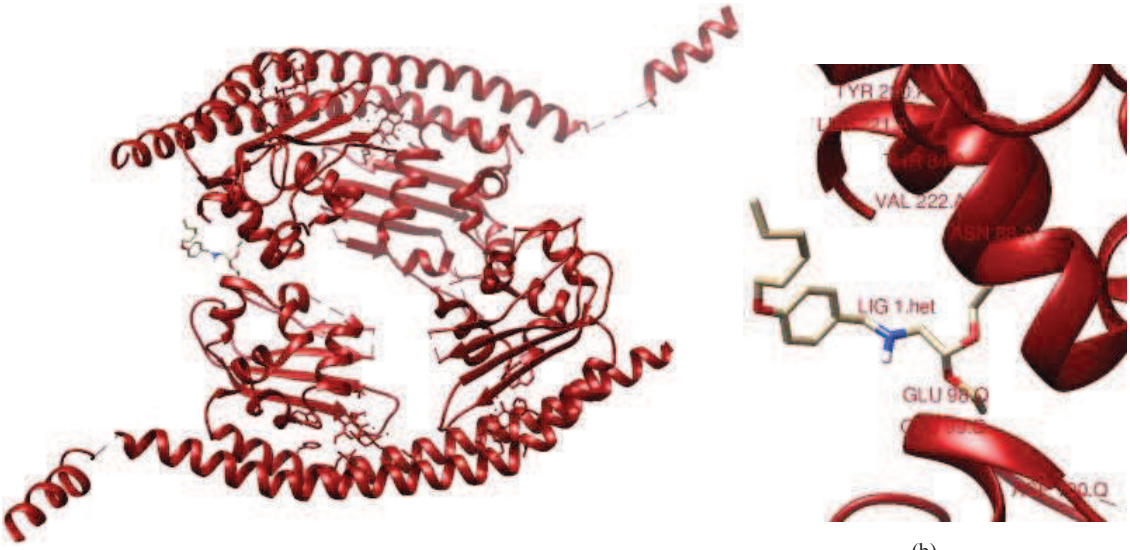


(a) (b)

Figure 40. SET monomer/MP07-66 docking (whole, a; zoom, b).

As shown in the images above, the molecule seems to allocate in the sharp-bend of the protein; the labeled aa residues are the ones that have a distance <5Å from MP07-66 (Phe[63], Arg[66], Leu[69], Ile[70], Lys[72], Phe[76], Leu[161], Pro[174], Asn[175], Pro[176], Leu[177], Tyr[180], Leu[181]). Many of these 13 residues are hydrophobic, with only two charged aa (Arg, Lys) and two polar aa (Asn, Tyr).

However, SET protein is registered on PDB as a dimer, so the docking was then performed on the whole reported structure (placing again the grid on the whole macromolecular structure). The docking output with lower energy is shown in the next figure.



(a) (b)

Figure 41. SET dimer/MP07-66 docking (whole, a; zoom, b).

Differently than before, and as shown by the reported dock, the compound seems to interact with both of the subunits taken into account, on the interface between one and the other. The labeled aa residues are the ones that have a distance <5Å from MP07-66, and they result to be present on both the monomers that compose SET (Thr^{84}, Asn^{88}, Tyr^{220}, Leu^{221}, Val^{222} on one of the subunits; Glu^{98}, Glu^{99}, Asp^{100} on the other). This situation is much different from the previous one, also because of the polarity of the involved residues: here only two residues are hydrophobic out of 8 (Leu, Val), while the most of them is polar (Thr, Asn, Tyr). On one of the two subunits there are even only negatively charged residues (Glu, Asp). At this point, only experimental data could say which the most likely situation between the two is.

In addition to this discrepancy, another relevant concern about these dockings is that the obtained outputs might represent non-specific interactions: given that MP07-66 is extremely lypophilic (ClogP: 4.70[96]), the main interactions it can give are supposed to be non-polar, and indeed the docking reveals the presence of no H bonds. Lacking this kind of interactions makes the predictions difficult to evaluate for their

reliability; because of this, defining the actual structure through experimental techniques as X-ray crystallography is once again needed. Structural details would be of a great help for fully understanding the mechanism of action of the compound itself, revealing the involved interactors on both MP07-66 and the protein. As already said for compound **71**, another possibility to better elucidate the interactions between the small molecule and the POI could be NMR spectroscopy; this technique offers several chances (all very well reviewed[97]) in terms of detecting and quantifying the interaction between a small molecule and a protein. This kind of approach, once optimized, permits to identify binding events either by evaluating the signals of the ligand or of the POI. The experiments more often performed in this kind of situation usually involve chemical-shift changes due to the interaction between the protein and the ligand, but this is just one of the options. The important point is to consider parameters (as the already mentioned chemical-shift changes, but also, for instance, changes of NOEs) that are obtainable with high sensitivity and in an easy way [97]. A few examples of the possible experiments, together with the main characteristics and comments about, are reported in the following table.

Table 2: NMR spectroscopy techniques for the identification and characterization of binding of ligands to proteins

	SAR by NMR	STD NMR	Spin labeling	Diffusion editing	Inverse NOE Pumping	Water-LOGSY
Large protein (>30 kDa)	limited[a]	yes	yes	no	yes	yes
small protein (<10 kDa)	yes	no	yes	yes	no	no
Isotope-labeled protein required	yes	no	no	no	no	no
Binding epitope on protein	yes[b]	no	no	no	no	no
Binding epitope on ligand	no	yes	no	no	yes[d]	yes[e]
Amount of protein [nmol] at 500 MHz	25	0.1	~1	~100	~25	~25
K_D tight binding K_D weak binding	no limit ~1 mM	100 pM ~10 mM	100 pM ~10 mM	~100 nM ~1 mM	1 nM ~1 mM	100 pM ~10 mM
Identification of ligand	no	yes	yes	yes	yes	yes
Comments	robust method	robust method	sensitive method, but results ambiguous if lysine positions unknown	relatively insensitive method	stable method, but ligand excess and mixing time need to be optimized	good for very hydrophilic targets and/or ligands

[a] TROSY necessary. [b] But chemical modification. [c] If protein is assigned. [d] Not realized. [e] But water contact surface.

Table 12. Main NMR techniques for detecting ligand-protein binding (SAR: structure-activity relationships; STD: saturation-transfer-difference; NOE: nuclear Overhauser effect; Water- LOGSY: water-ligand observation with gradient spectroscopy)[97].

2.3.2 Molecular biology achievements

The part of the work performed at the laboratories of Diamond Light Source and Research Complex at Harwell aimed at the production of large amounts of SET and $PP2A_C$ for structural and functional characterization. The high-throughput facilities of OPPF-UK were made available to clone and express in small scale a large number of constructs designed to express the protein in multiple expression systems.

An initial screening of conditions identified a good number of vectors as suitable for the expression of recombinant SET in *E. coli*. Two vectors expressing MBP and SUMO at the N-terminus of the protein SET_{1-225}, previously reported by Muto and co-workers[95], were successfully scaled up for large scale production. Initially from 2 x 1 L of culture of *E. coli* harboring the construct D6, the fusion protein His_6-SUMO-3C-His_6-Thrombin-SET_{1-225} was expressed and 2.5 mg of purified SET were obtained after a three-step purification process. Despite the high purity of the protein, the amounts are largely insufficient for structural studies since previous reports described the conditions for crystallization of SET starting from a 70 mg/mL of the protein.

It has to be noted that the construct used for this thesis work is the one used in the initial work from Muto and co-workers, rather than the one associated to the PDB entry of SET (PDB ID: 2E50). In fact, both works are based on N-terminally truncated forms of SET (aminoacids 1-225), but the last includes three point mutations to methionine to allow experimental phasing of the protein with the Se-Met method. Once a suitable model is available on the PDB, the structure solution can directly start from that and the three substitutions are no longer required. For this reason we decided to use the wild-type gene, which produces a protein whose structure is more similar to the naturally occurring one, thus avoiding artifacts that might pollute further structural and biochemical assays.

Large amounts of pure SET_{1-225} were obtained from the construct C6, which expressed much more protein fused to a maltose-binding protein solubilizing agent (His_6-MBP-3C-His_6-Thrombin-SET_{1-225}). After a redefinition of the purification strategy, the yields were increased by starting from a larger pool of cells. This implied the use of much more affinity resin and the volumes of sample produced were particularly difficult to work with. Given the large amount of protein eluted from the affinity resin, both proteolysis required large amounts of protease and repeated runs of size-exclusion chromatography. This lengthy process has been well tolerated by the protein, which did not show any propensity to degradation or precipitation. Analytical gel filtration tests coupled with MALS confirmed the protein is in its dimeric

form and MS analysis identified the correct MW for the protein, pure to >90% as from densitometric analysis of the SDS-PAGE. The overall yield was extremely satisfying: from 12 x 1 L cultures, approximately 340 mg of protein were produced.

Crystallization experiments are currently undergoing in the laboratories of Diamond Light Source.

Concerning $PP2A_C$, *E. coli* did not produce any of the protein in small-scale tests and the production was not scaled up. However, as for the production of the holoenzyme using Hi-5 insect cells, preliminary results show that both constructs A5 (His_6-3C-$PP2A_{C,1-245}$) and E5 ($PP2A_{C,1-245}$-His_6) could produce large amounts of the protein in soluble form. The absence of degradation bands is also an excellent confirmation that the protein can be obtained separately from the other components of the holoenzyme.

Based on the promising preliminary expression tests, the expression of construct A5 (preferred to E5 for the presence of a cleavable tag) is currently being expressed and purified in large scale in collaboration with Dr. Joanne Nettleship (OPPF-UK) and Dr. Marco Mazzorana (Diamond Light Source). The first small-scale purification test is very promising and indicates a good amount of soluble protein was expressed in a 2 L culture, as shown in figure 42.

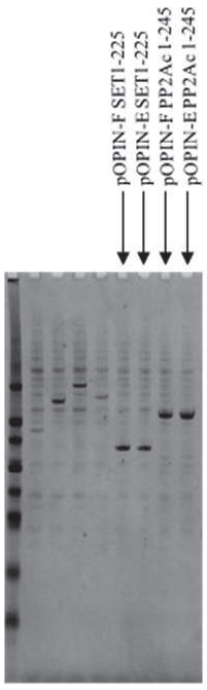


Figure 42. A6, D6, A5 and D5 construct expression in insect cells (3 µL virus/3 mL culture).

Purification of the protein will be performed and the sample will be used to try the reconstitution of the complex *in vitro*, followed by structural investigation, mainly through crystallography and small-angle X-ray scattering. This technique, despite the lower resolution, does not need the protein complex in a crystallized form and will be integrated with high-resolution X-ray crystallography data to describe atomic details. This is of fundamental importance to generate new compounds with increased potency and specificity towards this peculiar target.

With the quantity of SET so far obtained, HRMS experiments will also be performed to estimate the extent of binding between SET and the compounds. This is done by evaluating the shift of the mass peak corresponding to the protein itself when bonded to the analyzed compound. The optimization of the experiment (currently in progress) is, however, a lengthy process requiring tuning of the conditions for which the protein remains in its native conformation (*i.e.* not denaturated). For the data to be reliable and comparable to *in vitro* and *in vivo* behavior of SET, one must avoid denaturation, which would expose to the surface (and to the interaction with the ligand) residues that probably have nothing to deal with the binding studied. Avoiding denaturation, which would also probably inactivate the physiological function of SET, is not trivial: the best conditions for the protein to fly in the detector are often denaturing. At present, tuning experiments are carried out using SET by itself without any potential ligand, acquiring spectra in negative mode due to the low pI of the protein.

3. MATERIALS AND METHODS

3.1 ABBREVIATIONS

- Å: Angstrom/angstroms
- aa: Aminoacid/aminoacids/aminoacidic
- ACN: acetonitrile
- Akt: Protein kinase B
- AI: Autoinduction medium
- AML: Acute myeloid leukemia
- Asp: Aspartate
- B-CLL: B-cell chronic lymphocytic leukemia
- bs: Broad singlet
- °C: Celsius degrees
- Calcd: Calculated
- cAMP: Cyclic adenosin mono phosphate
- $CDCl_3$: Deuterated chloroform
- CIP2A: Cancerous inhibitor of PP2A
- CML: Chronic myeloid leukemia
- Concd: Concentrated
- CoSY: Correlation spectroscopy
- CV: Column volume
- δ: Delta, ppm
- d: Doublet
- D_2O: Deuterated water
- dd: Double doublet
- Da: Dalton
- DCE: 1,2-Dichloroethane
- DCM: Dichloromethane
- DMSO: Dimethylsulfoxide
- DUSP: Dual specificity phosphatase
- EB: Elution buffer

- EDTA: Ethylendiamino tetraacetic acid
- ESI: Electrospray ionization
- Et_2O: Diethylether
- EtOH: Ethanol
- Exper: Experimental
- FCP: TFIIF-associating C-terminal domain phosphatase
- FDA: Food and drug administration
- FPLC: Fast protein liquid chromatography
- g: Gram/grams
- *g*: Gravity acceleration
- GPCR: G-protein coupled receptor
- h: Hour/hours
- HAD: Haloacid dehalogenase
- HEAT: Huntingtin/elongation/A subunit/TOR
- His: Histidine
- HMBC: Heteronuclear multiple bond correlation
- HRMS: High resolution mass spectrometry
- HSC: Hematopoietic stem cells
- HSQC: Heteronuclear single quantum coherence
- HTP: High throughput purification
- I_2PP2A: PP2A inhibitor 2
- IMAC: Immobilized metal ion affinity chromatography
- IPTG: Isopropyl β-D-1-thiogalactopyranoside
- Jak2: Janus kinase 2
- k: Kilo
- KPSI: Kilopounds per square inch
- L: Liters
- LB: Lysogeny broth
- LCMT1: Leu-carboxyl methyltransferase
- Leu: Leucine

- µ: Micro
- m: Milli
- M: Molarity
- MALS: multi-angle light scattering
- MBP: Maltose binding protein
- MCP: Multi channel pipette
- MeOH: Methanol
- MES: 2-(N-morpholino)ethanesulfonic acid
- Met: Methionin
- m-CPBA: meta-Chloroperbenzoic acid
- MeOD: Deuterated methanol
- MeOH: Methanol
- mg: Milligrams
- MHz: Megahertz
- mL: Milliliter/milliliters
- mmol: Millimole/millimoles
- mol: Mole/moles
- MS: Multiple sclerosis
- MW: Molecular weight
- MWM: Molecular weight marker
- Ni-NTA: Ni^{2+}-nitrilotriacetate
- nM: Nanomolar
- NMR: Nuclear magnetic resonance
- NOESY: Nuclear Overhauser effect spectroscopy
- NPI: Sodium chloride, phosphate, imidazole
- OA: Okadaic acid
- OD: Optical density
- o/n: Overnight
- OPPF-UK: Oxford protein production facility-UK
- PAGE: Polyacrylamide gel electrophoresis

- PB: Power broth
- PCR: Polymerase chain reaction
- pI: Isoelectric point
- PI3K: Phosphoinositide 3-kinase
- PK: Protein kinase
- PMSF: Phenyl methyl sulfonyl fluoride
- PP: Protein phosphatase
- PPM: Mg^{2+}/Mn^{2+} dependent protein phosphatase
- ppm: Parts per million
- PPP: Phosphoprotein phosphatase
- POI: Protein of interest
- pOPIN: Plasmids for OPPF In-Fusion
- PP2A: Protein phosphatase 2A
- PTP: PhosphoTyr phosphatase
- R_f: Retention factor
- rpm: Roots per minute
- rt: Room temperature
- s: Singlet
- S1P: Sphingosine-1-phosphate
- S1PR: Sphingosine-1-phosphate receptor
- SCP: Small C-terminal domain phosphatase
- SDS: Sodium dodecyl sulfate
- SET: Suvar3-9, enhancer of zeste, trithorax
- SHP1: Src-homology 1 domain phosphatase
- SK: Sphingosine kinase
- S_N2: Bimolecular nucleophilic substitution
- SUMO: Small ubiquitin-related modifier
- SV40: Simian vacuolating virus 40
- t: Triplet
- TAF-Iβ: Template-activating factor-Iβ

- TB: Terrific broth
- TBE: Tris borate EDTA
- TEA: Triethylamine
- TGF: Transforming growth factor
- TLC: Thin layer chromatography
- TOF: Time of flight
- TOR: Target of rapamycin
- Tris: Tris(hydroxymethyl)aminomethane
- Trp: Tryptophan
- Wnt: Wingless-type MMTV integration site
- X-gal: 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside.

3.2 CONSUMABLES

Chemicals

All the chemicals were purchased by Agentcourt, Alfa-Aesar, Fluka, Sigma-Aldrich, Thermoscientific, Novagen, Qiagen or Invitrogen and used without further purifications.

Solvents

All the solvents were purchased by Sigma-Aldrich, Carlo Erba and VWR.

Deuterated solvents

All the deuterated solvents were purchased by Sigma-Aldrich and EurisoTop.

TLC plates

Analytical TLC was performed on 60F245 precoated silica gel plates (Merck); spots were visualized with UV light at 254 nm.

Columns for Flash chromatography

The utilized columns were SNAP cartridges 10g and 25g by Biotage.

Vectors

The pET-14b template vector containing the full-length SET has been kindly sent us from the research group of Prof. Masami Horikoshi (Institute of Molecular and Cellular Biosciences, University of Tokyo, Japan). The cDNA encoded by the vector translates in the following aa sequence:

MGSSHHHHHHSSGLVPRGSHMSAQAAKVSKKELNSNHDGADETSEKEQQEAIEHIDEVQNEIDR
LNEQASEEILKVEQKYNKLRQPFFQKRSELIAKIPNFWVTTFVNHPQVSALLGEEDEEALHYLTR
VEVTEFEDIKSGYRIDFYFDENPYFENKVLSKEFHLNESGDPSSKSTEIKWKSGKDLTKRSSQTQN
KASRKRQHEEPESFFTWFTDHSDAGADELGEVIKDDIWPNPLQYYLVPDMDDEEGEGEEDDDD
DEEEEGLEDIDEEGDEDEGEEDEDDDEGEEGEEDEGEDD.

The $PP2A_C$ coding plasmid pUC57 was purchased by GenScript Gene Synthesis service. The codon optimized vector was desigigned to express the protein sequence:

MDEKVFTKELDQWIEQLNECKQLSESQVKSLCEKAKEILTKESNVQEVRCPVTVCGDVHGQFHD
LMELFRIGGKSPDTNYLFMGDYVDRGYYSVETVTLLVALKVRYRERITILRGNHESRQITQVYGF
YDECLRKYGNANVWKYFTDLFDYLPLTALVDGQIFCLHGGLSPSIDTLDHIRALDRLQEVPHEGP
MCDLLWSDPDDRGGWGISPRGAGYTFGQDISETFNHANGLTLVSRAHQLVMEGYNWCHDRNV
VTIFSAPNYCYRCGNQAAIMELDDTLKYSFLQFDPAPRENLYFQGWSHPQFEK.

Both these plasmids were treated in order to insert the coding sequences in six of the pOPIN (plasmids for OPPF In-Fusion) expression suite vectors[89], which were kindly provided by OPPF-UK (Oxford Protein Production Facility); in particular, the following pOPINs were used:

pOPIN-F: His6-3C-POI

pOPIN-J: His6-GST-3C-POI

pOPIN-M: His6-MBP-3C-POI

pOPIN-S3C: His6-SUMO-3C-POI

pOPIN-E: POI-KHis6

pOPIN-EGFP: POI-3C-eGFP-KHis6

(where POI is the protein of interest).

Gels for electrophoresis

Gels for DNA revealing were prepared *in situ* dissolving electrophoresis-grade agarose to 1.6% in a TAE buffer; for SDS-PAGE, precast Nu-PAGE 4-12% BIS-TRIS GEL (Novex by Life technologies) were used.

Cells

Competent *E.coli* were provided by OPPF-UK in three main strains: OmniMaxII (Invitrogen) as cloning strain, Lemo21 (DE3) and Rosetta2 (DE3) as expression strains.

Buffers and additives

5052 50x: glycerol 2.7 M, glucose 140 mM, α-lactose 2.9 M.

DNA loading buffer: 0.25%w/v bromophenol blue in 30%v/v glycerol.

Dyalisis buffer: NaH_2PO_4 pH 8.0 50 mM, NaCl 300 mM, β-mercaptoethanol 7 mM.

NEBuffer 3 10x: NaCl 1 M, tris-HCl 500mM, $MgCl_2$ 100 mM, DTT 10 mM, pH 7.9 at 25°C.

NPI-10: NaH_2PO_4 pH 8.0 50 mM, NaCl 300 mM, imidazole 10 mM, tween 20 1%, β-mercaptoethanol 7 mM.

NPI-20: NaH_2PO_4 pH 8.0 50 mM, NaCl 300 mM, imidazole 20 mM, tween 20 0.05%, β-mercaptoethanol 7 mM.

NPI-250: NaH_2PO_4 pH 8.0 50 mM, NaCl 300 mM, imidazole 250 mM, tween 20 0.05%, β-mercaptoethanol 7 mM.

NPS 20x: $(NH)_4SO_4$ 0.5 M, KH_2PO_4 1 M, Na_2HPO_4 1 M, pH 6.75 at 20°C.

SET gel filtration buffer: Tris pH 8.0 20 mM, NaCl 100 mM, β-mercaptoethanol 10 mM.

TAE: Tris 40mM, EDTA 1mM, pH 8.3

3.3 INSTRUMENTS

Nuclear Magnetic Resonance (NMR)

NMR spectra were obtained with a NMR Bruker AVANCE III 400 MHz spectrometer and a NMR Bruker AMX 300 MHz spectrometer.

High Resolution Mass Spectrometry (HRMS)

Mass spectra were obtained with a Mariner ESI-TOF (Perceptive Biosystems Inc.) spectrometer, and with a Xevo G2-XS QTof mass spectrometer (Waters, Manchester, UK).

Flash Chromatography

A chromatographic Isolera One by Biotage unit has been used.

High throughput purification

A Bio-Robot 8000 (Qiagen) was used, equipped with QIA soft 4.2 software.

Fast Protein Liquid Chromatography (FPLC)

An AKTA Purifier System was utilized for performing FPLC purifications; the instrument was controlled by the System Control AKTA Purifier software and data were analyzed by the Unicorn Manager 5.31 software. The system was equipped with HisTrap FF (GE Healthcare) for IMAC purifications and a HiLoad 16/600 Superdex 200 prep grade (GE Healthcare) for gel filtration.

Gel electrophoresis devices

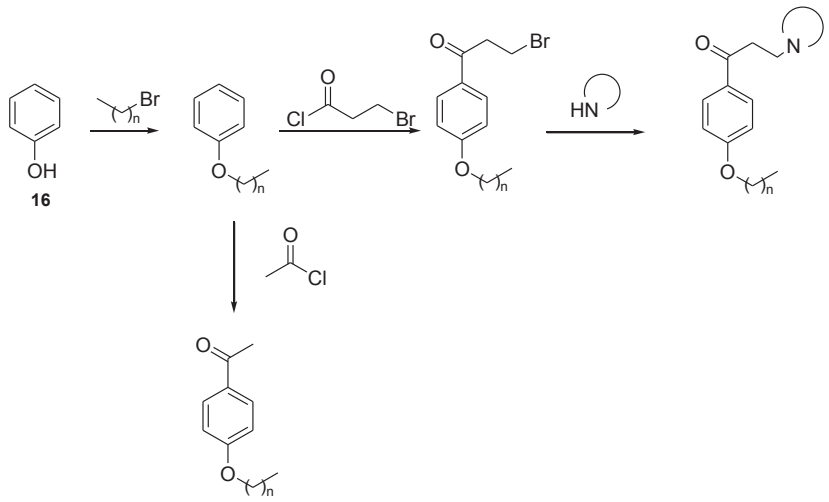
X cell sure lock system (Novex by Life technologies) and a Fisher scientific power station were used; gels were revealed with Instant Blue (Expedeon) and acquired with a GelDoc EZ Imager (Biorad).

UV spectrophotometric analysis

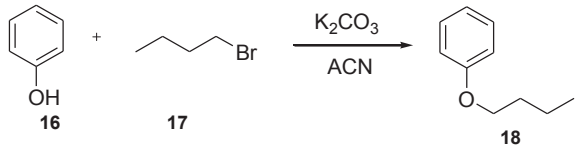
A NanoDrop 2000c spectrophotometer (Thermoscientific) was used, equipped with NanoDrop software.

3.4 SYNTHETIC SCHEMES

PATHWAY 1.1:



Synthesis of compound 18 (butoxybenzene)



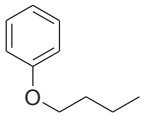
Phenol (**16**, 1.000 g, 10.64 mmol) was dissolved in 20 mL of acetonitrile in the presence of potassium carbonate (2.936 g, 21.28 mmol); the mixture was kept under magnetic stirring at rt for 1 h. Then, 1-bromobutane (**17**, 1.307 g, 1.029 mL, 9.54 mmol) was added and the reaction was heated at 60°C for 24 h and monitored by TLC (hexane/ethyl acetate 4:1).

After that, the reaction was filtered to remove the salts, and evaporated; the residue was dissolved in diethyl ether (30 mL) and washed with an aqueous solution of sodium hydroxyde 10% (30 mL, 3 times) and brine (20 mL). The organic phase was dried with magnesium sulfate and evaporated to give a colorless oil as a product.

Yield: 52%.

With the same procedure, the synthesis of more ethers was performed using other bromoalkanes, *i.e.* 1-bromohexane, 1-bromoheptane, 1-bromooctane, 1-bromononane. For all of these reactions, the obtained yields were between 50% and 60%.

Compound **18** – butoxybenzene



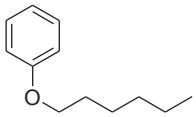
$C_{10}H_{14}O$

MW: 150.2176

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.32 (2H, m, ArH), 6.96 (3H, m, ArH), 4.00 (2H, m, OCH_2), 1.82 (2H, m, CH_2), 1.55 (2H, m, CH_2), 1.03 (3H, t, CH_3).

HRMS (ESI): calcd. $(M+H)^+$ 151.1045, exper. 151.1268.

Compound **19** – hexyloxybenzene



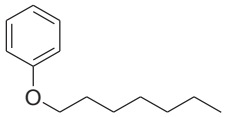
$C_{12}H_{18}O$

MW: 178.2707

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.30 (m, 2H, ArH), 6.93 (m, 3H, ArH), 3.98 (t, 2H, OCH_2), 1.81 (m, 2H, CH_2), 1.48 (m, 2H, CH_2), 1.37 (m, 4H, 2x CH_2), 0.93 (t, 3H, CH_3).

HRMS (ESI): calcd. $(M+H)^+$ 179.1358, exper. 179.1549.

Compound **20** – heptyloxybenzene



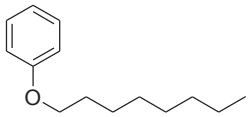
$C_{13}H_{20}O$

MW: 192.2973

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.34 (m, 2H, ArH), 6.98 (m, 3H, ArH), 4.02 (t, 2H, OCH_2), 1.86 (m, 2H, CH_2), 1.58-1.35 (m, 8H, 4x CH_2), 0.99 (t, 3H, CH_3).

HRMS (ESI): calcd. $(M+H)^+$ 193.1514, exper. 193.1540.

Compound **21** – octyloxybenzene



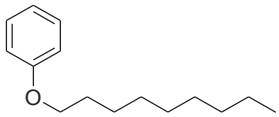
$C_{14}H_{22}O$

MW: 206.3239

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.25 (2H, ArH), 6.89 (3H, ArH), 3.93 (t, 2H, OCH_2), 1.76 (m, 4H, 2x CH_2), 1.5-1-2 (m, 8H, 4x CH_2), 0.88 (t, 3H, CH_3).

HRMS (ESI): calcd. $(M+H)^+$ 207.1671, exper. 207.3490.

Compound **22** – nonyloxybenzene



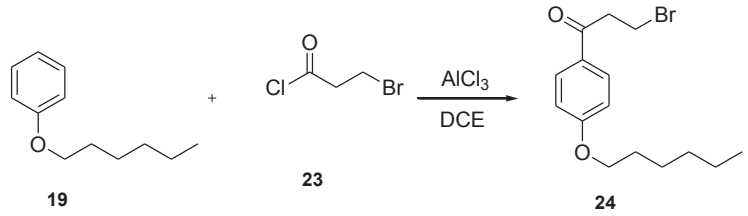
$C_{15}H_{24}O$

MW: 220.3505

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.31 (m, 2H, ArH), 6.95 (m, 3H, ArH), 3.99 (t, 2H, OCH_2), 3.44 (t, 2H, CH2), 1.80 (m, 2H, CH_2), 1.82 (m, 2H, CH_2), 1.6-1.4 (m, 4H, 2x CH_2), 1.0-0.9 (m, 7H, 2x CH_2, CH_3).

HRMS (ESI): calcd. $(M+H)^+$ 221.1827, exper. 221.1815.

Synthesis of compound 24 (3-bromo-1-(4-(hexyloxy)phenyl)propan-1-one)



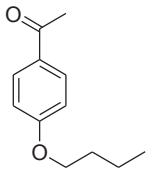
19 **23** **24**

Aluminum trichloride (450 mg, 3.36 mmol) in DCE (10 mL) was placed under magnetic stirring in an ice bath. Another solution was prepared with compound **19** (299 mg, 1.68 mmol) and 3-bromopropyonil chloride (**23**, 289 mg, 169 μL, 1.68 mmol) in DCE (5 mL); this second solution was slowly dropped in the first one, and the reaction was allowed to stir at rt for 3h (time by which the reaction, starting from a bright yellow color, becomes of a brownish orange). After the completion of the reaction, verified by TLC (hexane/ethyl acetate 4:1) the reaction mixture was poured in water (100 mL) and added with 20 mL of diethyl ether. The phases were then separated through a separation funnel, the organic phase was dried with $MgSO_4$ and the solvent was evaporated to give a yellow oil as a product.

Yield: 90%.

The same reaction was repeated on compounds **20**, **21** and **22**, while on compound **18** the acylation was performed with acetyl chloride instead of 3-bromopropyonil chloride. For all of these reactions, the yields are between 80 and 90%.

Compound **25** - 1-(4-butoxyphenyl)ethanone



$C_{12}H_{16}O_2$

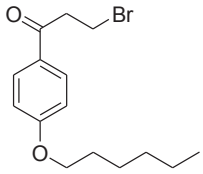
MW: 192.2542

^{1}H-NMR (300MHz, $CDCl_3$) δ: 7.94 (2H, AA'), 6.93 (2H, XX'), 4.04 (2H, t, OCH_2), 2.57 (3H, s, $COCH_3$), 1.81 (2H, m, CH_2), 1.52 (2H, m, CH_2), 1.00 (3H, t, CH_3).

^{13}C-NMR (75MHz, $CDCl_3$) δ: 196.9, 163.2, 130.6, 130.1, 114.1, 67.9, 31.1, 26.3, 19.3, 13.8.

HRMS (ESI): calcd. $(M+H)^+$ 193.1150, exper. 193.1258.

Compound **24** - 3-bromo-1-(4-(hexyloxy)phenyl)propan-1-one



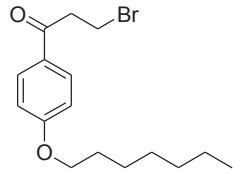
$C_{15}H_{21}BrO_2$

MW: 313.2300

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.95 (2H, AA'), 6.96 (2H, XX'), 4.05 (t, 2H, CH_2, J=6.7 Hz), 3.76 (t, 2H, OCH_2), 3.54 (t, 2H, CH_2, J=6.7 Hz), 1.9-1.8 (m, 2H, CH_2), 1.5-1.4 (m, 2H, CH_2), 1.38 (m, 4H, 2x CH_2), 0.93 (t, 3H, CH_3).

HRMS (ESI): calcd. $(M+H)^+$ 313.0725, exper. 313.1027 ($C_{15}H_{21}{}^{79}BrO_2$); calcd. $(M+H)^+$ 315.0725, exper. 315.1041 ($C_{15}H_{21}{}^{81}BrO_2$).

Compound **26** - 3-bromo-1-(4-(heptyloxy)phenyl)propan-1-one



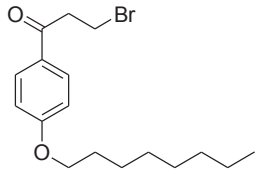
$C_{16}H_{23}BrO_2$

MW: 327.2566

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.94 (2H, AA'), 6.95 (2H, XX'), 4.04 (2H, t, J=6.8Hz, OCH_2), 3.76 (2H, t, J=6.8Hz, CH_2), 3.54 (2H, t, J=6.8Hz, $COCH_2$), 1.83 (2H, m, CH_2), 1.48 (2H, m, CH_2), 1.4-1.3 (6H, m, 3xCH_2), 0.92 (3H, t, J=6.8Hz, CH_3).

HRMS (ESI): calcd. $(M+H)^+$ 327.0881, exper. 327.1252 ($C_{15}H_{21}{}^{79}BrO_2$); calcd. $(M+H)^+$ 329.0881, exper. 329.1207 ($C_{15}H_{21}{}^{81}BrO_2$).

Compound **27** - 3-bromo-1-(4-(octyloxy)phenyl)propan-1-one



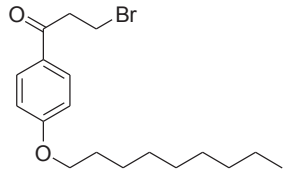
$C_{17}H_{25}BrO_2$

MW: 341.2832

^{1}H-NMR (400MHz, MeOD) δ: 7.99 (2H, AA'), 7.03 (2H, XX'), 4.09 (t, 2H, CH_2Br), 3.76 (t, 2H, CH_2), 3.60 (t, 2H, CH_2), 1.82 (m, 2H, CH_2), 1.51 (m, 2H, CH_2), 1.5-1.3 (m, 8H, 4x CH_2), 0.93 (t, 3H, CH_3).

HRMS (ESI): calcd. $(M+H)^+$ 341.1038, exper. 341.1236 ($C_{15}H_{21}{}^{79}BrO_2$); calcd. $(M+H)^+$ 343.1038, exper. 343.3102 ($C_{15}H_{21}{}^{81}BrO_2$).

Compound **28** - 3-bromo-1-(4-(nonyloxy)phenyl)propan-1-one



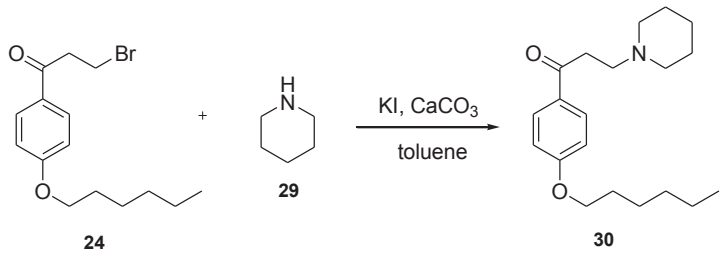
$C_{18}H_{27}BrO_2$

MW: 355.3098

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.95 (2H, AA'), 6.95 (2H, XX'), 4.04 (t, 2H, CH_2), 3.76 (t, 2H, OCH_2), 3.54 (t, 2H, CH_2), 3.43 (m, 2H, CH_2) 1.85 (m, 4H, 2x CH_2), 1.46 (m, 6H, 3x CH_2), 0.90 (m, 5H, CH_2-CH_3).

HRMS (ESI): calcd. $(M+H)^+$ 355.1194, exper. 355.1243 ($C_{15}H_{21}{}^{79}BrO_2$); calcd. $(M+H)^+$ 357.1174, exper. 357.0838 ($C_{15}H_{21}{}^{81}BrO_2$).

Synthesis of compound 30 (1-(4-(hexyloxy)phenyl)-3-(piperidin-1-yl)propan-1-one)



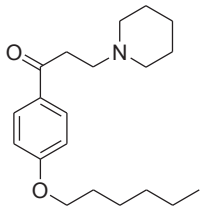
In a 100 mL round bottom flask containing 15 mL of toluene, were placed compound **24** (100 mg, 0.32 mmol), piperidine (**29**, 40 mg, 46 μL, 0.48 mmol), potassium iodide (79 mg, 0.48 mmol) and calcium carbonate (48 mg, 0.48 mmol). The reaction was stirred o/n heating it at 110°C; the completion of the reaction was verified by TLC (hexane/ethyl acetate 4:1).

The mixture was then cooled to rt and filtered, and the solid residue was washed with toluene; the filtrate was then evaporated to remove the solvent. The residue was dissolved in DCM (20 mL) and washed in a separation funnel with sodium hydroxyde 10% in water (20 mL, 3 times); the collected organic phase was dried with magnesium sulfate and the solvent was removed. The resulting product was a brownish oil.

Yield: 85%.

The same reaction was repeated also on compounds **26**, **27** and **28** in the same way. In addition to this, compounds **24**, **26**, **27** and **28** were also reacted with morpholine instead of piperidine. All the yields were between 75 and 85%.

Compound **30** - 1-(4-(hexyloxy)phenyl)-3-(piperidin-1-yl)propan-1-one



$C_{20}H_{31}NO_2$

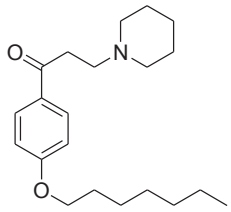
MW: 317.4656

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.80 (2H, AA'), 6.94 (2H, XX'), 4.04 (t, 2H, CH_2), 3.18 (t, 2H, O CH_2), 2.82 (t, 2H, CH_2), 1.82 (m, 2H, CH_2), 1.63 (m, 8H, 4x CH_2), 1.49 (m, 4H, 2x CH_2), 1.37 (m, 4H 2x CH_2), 0.93 (t, 3H, CH_3).

^{13}C-NMR (100MHz, $CDCl_3$) δ: 195.5, 161.8, 129.0, 128.5, 112.9, 75.9, 66.9, 53.3, 52.8, 30.3, 27.8, 24.6, 24.7, 21.3, 12.8.

HRMS (ESI): calcd. $(M+H)^+$ 318.2355, exper. 318.2690.

Compound **31** - 1-(4-(heptyloxy)phenyl)-3-(piperidin-1-yl)propan-1-one



$C_{21}H_{33}NO_2$

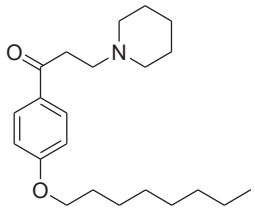
MW: 331.4922

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.95 (2H, AA'), 6.93 (2H, XX'), 4.04 (t, 2H, OCH_2), 2.80 (m, 2H, CH_2), 2.47 (m, 4H, 2x CH_2N), 1.82 (m, 2H, CH_2), 1.62 (m, 4H, 2x CH_2), 1.47 (m, 4H, 2x CH_2), 1.4-1.2 (m, 8H, 4x CH_2), 0.92 (t, 3H, CH_3).

^{13}C-NMR (100MHz, $CDCl_3$) δ: 198.0, 163.1, 130.3, 129.8, 114.2, 68.2, 54.6, 54.2, 36.1, 31.7, 29.1, 29.0, 26.0, 25.9, 24.3, 22.6, 14.1.

HRMS (ESI): calcd. $(M+H)^+$ 332.2511, exper. 332.2903.

Compound **32** - 1-(4-(octyloxy)phenyl)-3-(piperidin-1-yl)propan-1-one



$C_{22}H_{35}NO_2$

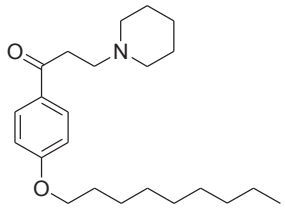
MW: 345.5188

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.85 (2H, AA'), 6.84 (2H, XX'), 3.92 (t, 2H, OCH_2), 3.62 (t, 2H, CH_2), 3.04 (t, 2H, CH_2), 2.73 (m, 2H, CH_2), 2.42 (m, 4H, 2x CH_2), 2.23 (m, 2H, CH_2), 1.72 (m, 2H, CH_2), 1.37 (m, 2H, CH_2), 1.3-1.1 (m, 10H, 5xCH_2), 0.79 (m, 5H, CH_2CH_3).

^{13}C-NMR (100MHz, $CDCl_3$) δ: 195.6, 159.5, 134.6, 130.9, 66.7, 59.6, 51.4, 40.2, 33.1, 30.7, 26.5, 25.2, 22.6, 14.0, 10.8.

HRMS (ESI): calcd. $(M+H)^+$ 346.2668, exper. 346.4092.

Compound **33** - 1-(4-(nonyloxy)phenyl)-3-(piperidin-1-yl)propan-1-one



$C_{23}H_{37}NO_2$

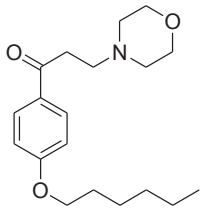
MW: 359.5454

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.95 (2H, AA'), 6.93 (2H, XX'), 4.04 (t, 2H, OCH_2), 3.16 (t, 2H, CH_2), 2.80 (t, 2H, CH_2), 2.5-2.4 (m, 4H, 2x NCH_2), 1.82 (m, 4H, 2x CH_2), 1.7-1.6 (m, 6H, 3x CH_2), 1.48 (m, 4H, 2x CH_2), 0.91 (m, 5H, CH_2CH_3).

^{13}C-NMR (100MHz, $CDCl_3$) δ: 198.0, 161.2, 135.2, 133.7, 120.4, 70.5, 51.4, 49.9, 35.3, 33.6, 29.4, 28.9, 26.2, 24.1, 22.6, 14.0.

HRMS (ESI): calcd. $(M+H)^+$ 360.2824, exper. 360.2878.

Compound **34** - 1-(4-(hexyloxy)phenyl)-3-morpholinopropan-1-one



$C_{19}H_{29}NO_3$

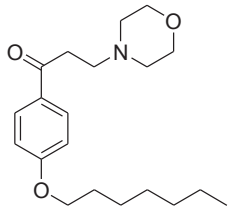
MW: 319.4385

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.95 (2H, AA'), 6.95 (2H, XX'), 4.04 (t, 2H, OCH_2), 3.74 (m, 4H, 2x CH_2), 3.15 (t, 2H, CH_2, J=7.7 Hz), 2.84 (t, 2H, CH_2, J=7.7 Hz), 2.53 (m, 4H, 2x CH_2), 1.83 (m, 2H, CH_2), 1.49 (m, 2H, CH_2), 4.38 (m, 4H, 2x CH_2), 0.93 (t, 3H, CH_3).

^{13}C-NMR (100MHz, $CDCl_3$) δ: 196.5, 162.2, 129.3, 113.2, 76.2, 67.3, 65.9, 52.8, 52.7, 34.5, 30.5, 28.0, 24.6, 21.6, 13.0.

HRMS (ESI): calcd. $(M+H)^+$ 320.2147, exper. 320.2304.

Compound **35** - 1-(4-(heptyloxy)phenyl)-3-morpholinopropan-1-one



$C_{20}H_{31}NO_3$

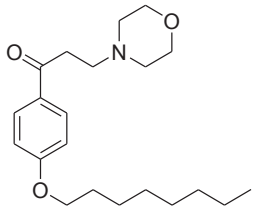
MW: 333.4650

^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.95 (2H, AA'), 6.94 (2H, XX'), 4.04 (t, 2H, OCH_2) , 3.74 (m, 4H, 2xCH_2), 3.16 (t, 2H, CH_2), 2.85 (t, 2H, CH_2), 2.54 (m, 4H, 2x CH_2N), 1.83 (m, 2H, CH_2), 1.48 (m, 2H, CH_2), 1.4-1.3 (m, 6H, 3x CH_2), 0.92 (t, 3H, CH_3) .

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 197.5, 163.2, 130.3, 129.7, 114.2, 68.3, 67.0, 53.8, 53.7, 35.6, 31.7, 29.1, 29.0, 25.9, 22.6, 14.1.

HRMS (ESI): calcd. $(M+H)^+$ 334.2304, exper. 334.2690.

Compound **36** - 3-morpholino-1-(4-(octyloxy)phenyl)propan-1-one



$C_{21}H_{33}NO_3$

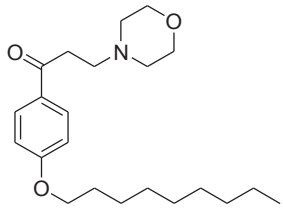
MW: 347.4916

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.86 (2H, AA'), 6.84 (2H, XX'), 3.93 (t, 2H, OCH_2), 3.64 (t, 2H, CH_2), 3.06 (t, 2H, CH_2), 2.74 (t, 2H, CH_2), 2.43 (m, 2H, CH_2), 1.72 (m, 4H, 2xCH_2), 1.38 (m, 2H, CH_2), 1.3-1.1 (m, 8H, 4xCH_2), 0.82 (m, 5H, CH_2CH_3).

^{13}C-NMR (100MHz, $CDCl_3$) δ: 197.5, 160.5, 135.5, 130.3, 114.6, 68.3, 66.3, 54.0, 49.9, 36.5, 33.4, 29.7, 26.2, 22.6, 12.8.

HRMS (ESI): calcd. $(M+H)^+$ 348.2460, exper. 348.4092.

Compound **37** - 3-morpholino-1-(4-(nonyloxy)phenyl)propan-1-one



$C_{22}H_{35}NO_3$

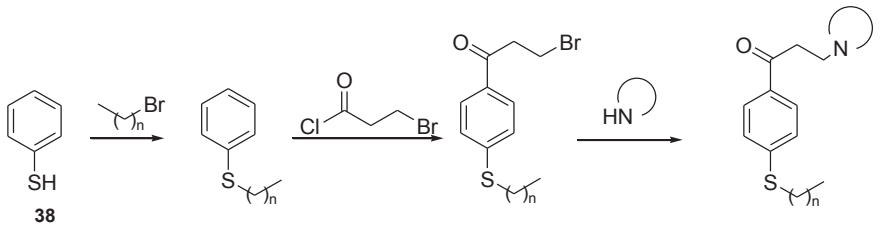
MW: 361.5182

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.81 (2H, AA'), 6.80 (2H, XX'), 3.90 (t, 2H, OCH_2), 3.60 (m, 4H, 2xCH_2), 3.01 (t, 2H, CH_2), 2,70 (t, 2H, CH_2), 2.39 (m, 4H, 2xCH_2), 1.68 (m, 2H, CH_2), 1.4-1.1 (m, 10H, 5xCH_2), 0.8-0.7 (m, 5H, CH_2CH_3).

^{13}C-NMR (100MHz, $CDCl_3$) δ: 197.6, 160.5, 134.9, 130.3, 114.6, 68.2, 66.3, 54.0, 49.6, 32.4, 29.3, 26.4, 22.6.

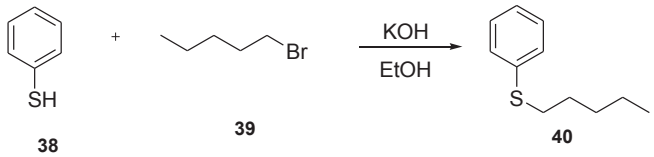
HRMS (ESI): calcd. $(M+H)^+$ 362.2617, exper. 362.2598.

PATHWAY 1.2:



38

Synthesis of compound 40 (pentyl(phenyl)sulfane)



38 **39** **40**

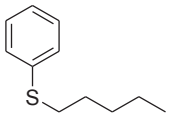
In a 100 mL round-bottom flask were placed thiophenol (**38**, 1.000 g, 0.932 mL, 9.07 mmol) and potassium hydroxide (509 mg, 9.07 mmol) in 20 mL of ethanol; in the stirring solution, 1-bromopentane was added dropwise (**39**, 1.369 g, 1.125 mL, 9.07 mmol). The solution was stirred at rt for 5 h, monitoring it by TLC (hexane/ethyl acetate 4:1); then, the mixture was filtered and the solvent was removed from the filtrate to give a yellow very dense oil.

Yield: 85%.

With the same procedure, the synthesis of other thioethers was performed using 1-bromohexane and 1-bromoheptane. In these cases, the yields were 81 and 75% respectively.

The subsequent steps (*Friedel-Crafts* acylation and *Finkelstein* substitution) have been performed in the same way which was described for compounds **24** and **30**. The only difference was that, at the very end, most of them needed to be purified (apart from **49**) through flash chromatography, using DCM/MeOH (100:0 to 95:5) as eluents. The yields for the *Friedel-Crafts* acylation were between 53 and 62%, while the ones for the *Finkelstein* substitution were between 60 and 74%.

Compound **40** - pentyl(phenyl)sulfane

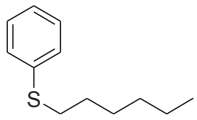


$C_{11}H_{16}S$

MW: 180.3097

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.38 (2H, m, ArH), 7.23 (3H, m, ArH), 2.99 (2H, t, SCH_2), 1.75 (2H, m, CH_2), 1.46 (4H, m, 2xCH_2), 1.00 (3H, m, CH_3).

Compound **41** – hexyl(phenyl)sulfane

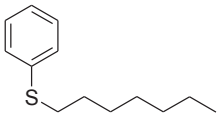


$C_{13}H_{18}S$

MW: 194.3363

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.5-7.3 (4H, m, ArH), 7.24 (1H, m, ArH), 3.01 (2H, t, J=7.30 Hz, SCH_2), 1.76 (2H, m, CH_2), 1.53 (2H, m, CH_2), 1.42 (4H, m, 2xCH_2), 1.02 (3H, t, J=7.00 Hz, CH_3).

Compound **42** - heptyl(phenyl)sulfane

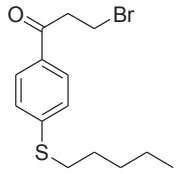


$C_{13}H_{20}S$

MW: 208.3629

^{1}H-NMR (300 MHz, $CDCl_3$) δ: 7.4-7.3 (4H, m, ArH), 7.18 (1H, m, ArH), 2.93 (2H, t, J=7.32 Hz, SCH_2), 1.67 (2H, m, CH_2), 1.5-1.2 (8H, m, CH_2x4), 1.00 (3H, t, J= 6.78, CH_3).

Compound **43** - 3-bromo-1-(4-(pentylthio)phenyl)propan-1-one

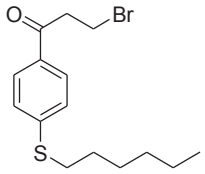


$C_{14}H_{19}BrOS$

MW: 315.2691

^{1}H-NMR (300 MHz, $CDCl_3$) δ: 7.72 (2H, AA'), 7.19 (2H, XX'), 3.64 (2H, t, CH_2), 3.42 (2H, t, CH_2), 2.90 (2H, t, SCH_2), 1.62 (2H, m, CH_2), 1.31 (4H, m, CH_2x2), 0.82 (3H, t, CH_3).

Compound **44** - 3-bromo-1-(4-(hexylthio)phenyl)propan-1-one

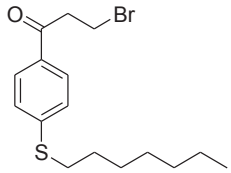


$C_{15}H_{21}BrOS$

MW: 329.2956

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.84 (2H, AA'), 7.30 (2H, XX'), 3.73 (2H, t, J=7.03 Hz, CH_2), 3.52 (2H, t, J=7.03 Hz, CH_2), 2.99 (2H, t, J=7.28 Hz, SCH_2), 1.71 (2H, m, CH_2), 1.47 (2H, m, CH_2), 1.33 (4H, m, CH_2x2), 0.91 (3H, t, CH_3).

Compound **45** - 3-bromo-1-(4-(heptylthio)phenyl)propan-1-one

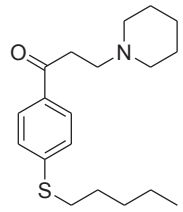


$C_{16}H_{23}BrOS$

MW: 343.3222

^{1}H-NMR (300 MHz, $CDCl_3$) δ: 7.84 (2H, AA'), 7.31 (2H, XX'), 3.73 (2H, t, J= 6.89 Hz, CH_2), 3.53 (2H, t, J=6.89 Hz, CH_2), 2.99 (2H, t, J=7.10, SCH_2), 1.68 (2H, m, CH_2), 1.5-1.2 (8H, m, CH_2x4), 0.88 (3H, t, CH_3).

Compound **46** – 1-(4-(pentylthio)phenyl)-3-(piperidin-1-yl)propan-1-one



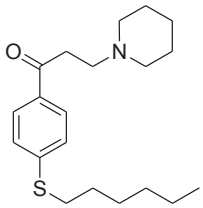
$C_{19}H_{29}NOS$

MW: 319.5047

^{1}H-NMR (300 MHz, $CDCl_3$) δ: 7.85 (AA', 2H), 7.26 (XX', 2H), 3.13 (2H, t, SCH_2), 2.96 (2H, t, CH_2), 2.75 (2H, t, CH_2), 2.91 (2H, m, CH_2), 1.67 (4H, m, CH_2x2), 1.58 (4H, m, CH_2x2), 1.42 (4H, m, 2xCH_2), 1.32 (2H, m, CH_2), 0.98 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 197.8, 145.1, 135.9, 133.0, 128.7, 69.7, 54.6, 36.5, 35.2, 33.5, 31.9, 30.1, 29.4, 22.3.

Compound **47** - 1-(4-(hexylthio)phenyl)-3-(piperidin-1-yl)propan-1-one



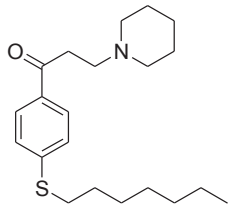
$C_{20}H_{31}NOS$

MW: 333.5312

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.86 (2H, AA'), 7.29 (2H, XX'), 3.20 (2H, t, CH_2), 3.13 (4H, bt, CH_2x2), 2.99 (2H, t, SCH_2), 2.85 (2H, t, CH_2), 2.51 (2H, m, CH_2), 1.87 (4H, m, CH_2x2), 1.7-1.6 (4H, m, CH_2x2), 1.46 (2H, m, CH_2), 1.31 (2H, m, CH_2), 0.89 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 198.2, 145.1, 133.4, 128.8, 126.3, 54.5, 53.9, 45.1, 36.0, 33.6, 31.9, 31.3, 29.1, 28.5, 25.7, 22.5.

Compound **48** - 1-(4-(heptylthio)phenyl)-3-(piperidin-1-yl)propan-1-one



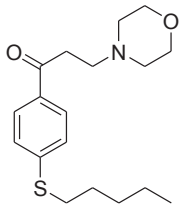
$C_{21}H_{33}NOS$

MW: 347.5578

H-NMR (300 MHz, $CDCl_3$) δ: 7.88 (AA', 2H), 7.30 (XX', 2H), 3.17 (2H, t, J=7.86 Hz, CH_2), 3.01 (2H, t, S-CH_2), 2.80 (2H, t, J=7.86 Hz, CH_2), 2.47 (2H, t, CH_2), 2.75 (4H, t, CH_2x2), 2.91 (2H, m, CH_2), 1.8-1.6 (8H, m, CH_2x4), 1.5-1.2 (8H, m, CH_2x4), 0.89 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 197.9, 145.0, 136.0, 133.1, 128.8, 69.9, 62.3, 54.7, 36.2, 35.8, 33.5, 31.8, 31.1, 30.2, 29.3, 28.2, 22.4.

Compound **49** - 3-morpholino-1-(4-(pentylthio)phenyl)propan-1-one



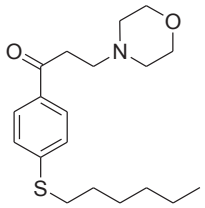
$C_{18}H_{27}NO_2S$

MW: 321.4775

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.83 (2H, AA'), 7.28 (2H, XX'), 3.69 (4H, t, CH_2x2), 3.12 (2H, t, CH_2), 2.97 (2H, t, CH_2), 2.81 (2H, t, SCH_2), 2.49 (4H, m, CH_2x2), 1.69 (4H, m, CH_2x2), 1.37 (2H, m, CH_2), 0.90 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 197.9, 145.1, 133.5, 128.4, 126.3, 66.8, 53.7, 53.6, 35.8, 31.8, 31.0, 28.4, 22.1, 13.8.

Compound **50** - 1-(4-(hexylthio)phenyl)-3-morpholinopropan-1-one



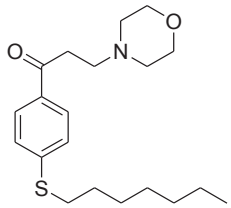
$C_{19}H_{29}NO_2S$

MW: 335.5041

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.84 (2H, AA'), 7.29 (2H, XX'), 3.71 (4H, t, CH_2x2), 3.15 (2H, t, CH_2), 2.99 (2H, t, J=7.40 Hz, SCH_2), 2.83 (2H, t, CH_2), 2.50 (4H, t, CH_2x2), 1.69 (2H, m, CH_2), 1.44 (2H, m, CH_2), 1.30 (4H, m, CH_2x2), 0.89 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 197.9, 145.2, 137.1, 133.4, 128.8, 68.9, 53.7, 35.7, 33.6, 31.9, 31.3, 30.5, 29.1, 28.5, 22.5.

Compound **51** - 1-(4-(heptylthio)phenyl)-3-morpholinopropan-1-one



$C_{20}H_{31}NO_2S$

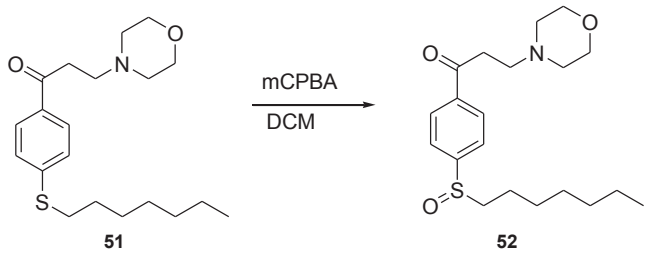
MW: 349.5306

^{1}H-NMR (300 MHz, $CDCl_3$) δ: 7.88 (2H, AA'), 7.30 (2H, XX'), 3.73 (4H, t, CH_2x2), 3.15 (2H, t, CH_2), 3.00 (2H, t, J=7.20 Hz, SCH_2), 2.84 (2H, t, CH_2), 2.52 (4H, m, CH_2x2), 1.69 (2H, m, CH_2), 1.5-1.2 (8H, m, CH_2x4), 0.90 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 197.7, 145.1, 136.1, 133.2, 128.9, 69.9, 54.7, 36.7, 35.4, 33.6, 31.9, 31.0, 30.1, 29.4, 28.4, 22.5.

PATHWAY 1.2.1

Synthesis of compound 52 (1-(4-(heptylsulfinyl)phenyl)-3-morpholinopropan-1-one)

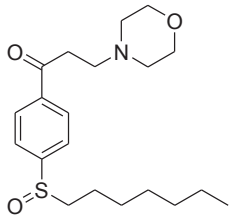


51 **52**

Compound **51** (50 mg, 0.14 mmol) was dissolved in dichloromethane (10 mL) at 0°C, and a solution of meta-chloroperbenzoic acid (24 mg, 0.14 mmol) in 5 mL of dichloromethane was added dropwise. The solution was stirred at 0°C for 2 h and monitored by TLC (hexane/ethyl acetate 4:1). The reaction was then added with an aqueous solution of sodium hydroxyde 1N (30 mL) and extracted with DCM; the organic phase was washed with brine (2x20 mL) and dried over magnesium sulfate. The organic phase was the concentrated and purified by flash chromatography, eluting the product through a gradient of DCM/MeOH (from 95:5 to 90:10) to give a yellow oil.

Yield: 30%.

Compound **52**– 1-(4-(heptylsulfinyl)phenyl)-3-morpholinopropan-1-one



$C_{20}H_{31}NO_3S$

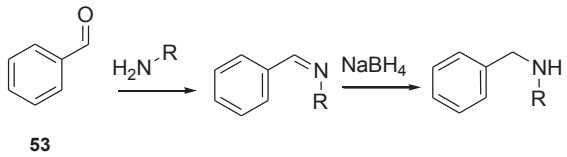
MW: 365.5300

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 8.11 (2H, AA'), 7.74 (2H, XX'), 3.74 (4H, t, CH_2x2), 3.24 (2H, t, CH_2), 2.9-2.8 (4H, t, CH_2x2), 2.55 (4H, m, CH_2x2), 1.80 (2H, m, CH_2), 1.63 (2H, m, CH_2), 1.5-1.2 (6H, m, CH_2x3), 0.89 (3H, t, CH_3).

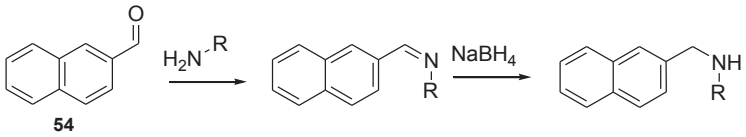
^{13}C-NMR (75 MHz, $CDCl_3$) δ: 198.1, 149.6, 128.8, 124.5, 66.8, 57.1, 53.7, 53.3, 36.2, 31.5, 28.8, 28.6, 22.5, 22.0, 14.0.

HRMS (ESI): calcd. $(M+H)^+$ 366.2025, exper. 366.2225.

PATHWAY 2.1.1:

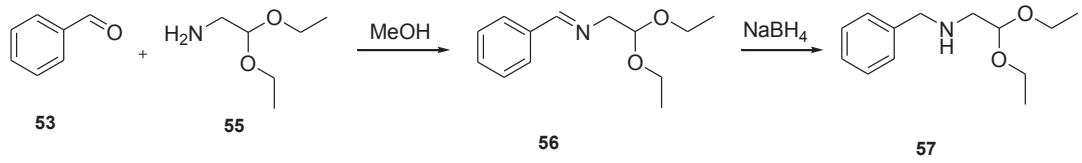


53



54

Synthesis of compounds 56 (N-benzylidene-2,2-diethoxyethanamine) and 57 (N-benzyl-2,2-diethoxyethanamine)



53 55 56 57

To a stirring solution of benzaldehyde (**53**, 500 mg, 0.477 mL, 4.72 mmol) in methanol (15 mL), 2,2-diethoxyethanamine (**55**, 628 mg, 0.682 mL, 4.72 mmol) was added. The mixture was heated at 60°C for 3 h, monitoring the formation of the imine by TLC (hexane/ethyl acetate 4:1).

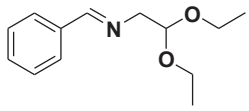
The solution was then divided in two halves: one of them was dried with magnesium sulfate and the solvent removed, in order to obtain the desired imine (**56**), while the second half was used for synthesizing the amine (**57**). In the latter case, the stirring solution was placed in an ice bath and carefully added with sodium boron hydride (269 mg, 7.08 mmol) and maintained at rt o/n. After the completion of the reaction, verified by TLC (hexane/ethyl acetate 4:1), the stirring solution was quenched at 0°C through the careful addition of a saturated aqueous solution of sodium hydroxy carbonate; the organic phase was then removed under reduced pressure, and the aqueous phase was extracted with dichloromethane (30 mL, 3 times). The combined organic phases were then washed with brine (20 mL, twice), dried with magnesium sulfate and the solvent was removed under reduced pressure to give a yellow oily product.

Yield for the imine (**56**): 81%.

Yield for the amine (**57**): 35%.

The same reaction was repeated with 2-naphthaldehyde (**54**) instead of benzaldehyde, and also using different primary amines (instead of 2,2-diethoxyethanamine) as 3-morpholinopropan-1-amine or 4-aminocyclohexanol in the same way. All the yields for the imines were between 73 and 85%, while for the amines were between 32 and 36%.

Compound **56** – N-benzylidene-2,2-diethoxyethanamine



$C_{13}H_{19}NO_2$

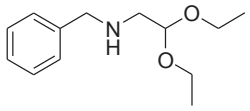
MW: 221.2955

^{1}H-NMR (300 MHz, $CDCl_3$) δ: 8.21 (1H, s, CH=N), 7.66 (2H, m, ArH), 7.33 (3H, m, ArH), 4.73 (1H, t, J=5.40 Hz, CH), 3.71 (2H, dd, J=5.40, J=1.45, $CHC\underline{H_2}$), 3.65 (2H, m, OCH_2), 3.51 (2H, m, OCH_2), 1.12 (6H, t, J=6.96 Hz, CH_3x2).

^{13}C-NMR (75 MHz, $CDCl_3$) δ: 163.4, 136.2, 130.7, 129.7, 129.0, 128.6, 128.2, 102.1, 64.7, 62.8, 15.4.

HRMS (ESI): calcd. $(M+H)^+$ 222.1416, exper. 222.1856.

Compound **57** - N-benzyl-2,2-diethoxyethanamine



$C_{13}H_{21}NO_2$

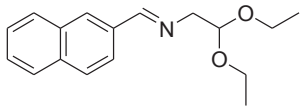
MW: 223.3113

^{1}H-NMR (300 MHz, $CDCl_3$) δ: 7.35 (4H, m, ArH), 7.24 (1H, m, ArH), 4.65 (1H, t, J=5.50 Hz, CH), 3.83 (2H, s, $ArCH_2$), 3.71 (2H, m, OCH_2), 3.55 (2H, m, OCH_2), 2.77 (2H, d, J=5.50 Hz, $CHCH_2$), 1.23 (6H, t, J=7.00 Hz, CH_3x2).

^{13}C-NMR (75 MHz, $CDCl_3$) δ: 140.1, 128.4, 128.2, 127.0, 102.2, 64.2, 53.9, 51.6, 15.4.

HRMS (ESI): calcd. $(M+H)^+$ 224.1572, exper. 224.1667.

Compound **58** - 2,2-diethoxy-N-((naphthalen-3-yl)methylene)ethanamine



$C_{17}H_{21}NO_2$

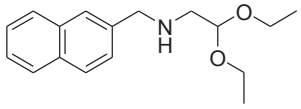
MW: 271.3541

^{1}H-NMR (300 MHz, $CDCl_3$) δ: 8.47 (1H, s, CH=N), 8.1-7.9 (5H, m, ArH), 7.53 (2H, m, ArH), 4.88 (1H, t, J=5.41 Hz, CH), 3.87 (2H, dd, J=5.50 Hz, J=1.45 Hz, $CHC\underline{H_2}$), 3.78 (2H, m, OCH_2), 3.64 (2H, m, OCH_2), 1.24 (6H, t, J=6.91 Hz, CH_3x2).

^{13}C-NMR (75 MHz, $CDCl_3$) δ: 192.3, 163.6, 134.7, 133.9, 133.1, 130.1, 128.6, 127.7, 127.2, 126.5, 123.8, 102.2, 64.8, 62.8, 15.4.

HRMS (ESI): calcd. $(M+H)^+$ 272.1572, exper. 272.1686.

Compound **59** - 2,2-diethoxy-N-((naphthalen-3-yl)methyl)ethanamine



$C_{17}H_{23}NO_2$

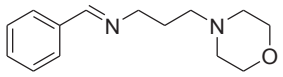
MW: 273.3700

^{1}H-NMR (300 MHz, $CDCl_3$) δ: 7.9-7.7 (4H, m, ArH), 7.5-7.4 (3H, m, ArH), 4.67 (1H, t, J=5.47 Hz, CH), 4.00 (2H, s, CH_2-Ar), 3.71 (2H, m, OCH_2), 3.56 (2H, m, OCH_2), 2.81 (2H, d, J=5.47 Hz, $CHC\underline{H_2}$), 1.23 (6H, t, J=6.91 Hz, CH_3x2).

^{13}C-NMR (75 MHz, $CDCl_3$) δ: 137.7, 133.4, 132.7, 128.1, 127.7, 127.6, 126.6, 126.5, 125.9, 125.5, 102.2, 62.4, 54.0, 51.6, 15.4.

HRMS (ESI): calcd. $(M+H)^+$ 274.1729, exper. 274.1623.

Compound **60** - N-benzylidene-3-morpholinopropan-1-amine



$C_{14}H_{20}N_2O$

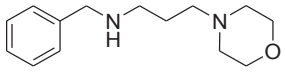
MW: 232.3214

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 8.30 (1H, s, CH=N), 7.73 (2H, m, ArH), 7.42 (3H, m, ArH), 3.72 (4H, m, CH_2x2), 3.66 (2H, t, CH_2), 2.46 (6H, m, CH_2x3), 1.91 (2H, m, CH_2).

^{13}C-NMR (75 MHz, $CDCl_3$) δ: 161.3, 136.2, 130.6, 128.6, 128.1, 67.0, 59.6, 56.7, 53.7, 27.7.

HRMS (ESI): calcd. $(M+H)^+$ 233.1576, exper. 233.1466.

Compound **61** - N-benzyl-3-morpholinopropan-1-amine



$C_{14}H_{22}N_2O$

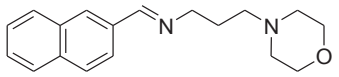
MW: 234.3373

^{1}H-NMR (300 MHz, $CDCl_3$) δ: 7.4-7.3 (5H, m, ArH), 3.78 (2H, s, CH_2Ar) 3.67 (4H, t, CH_2x2), 2.69 (2H, t, CH_2), 2.4-2.3 (6H, m, CH_2x3), 1.71 (2H, t, CH_2).

^{13}C-NMR (75 MHz, $CDCl_3$) δ: 140.0, 129.7, 128.8, 128.1, 127.3, 126.8, 66.9, 57.3, 53.7, 47.9, 26.4, 20.5.

HRMS (ESI): calcd. $(M+H)^+$ 235.1732, exper. 235.1632.

Compound **62** –3-morpholino-N-((naphthalen-3-yl)methylene)propan-1-amine



$C_{18}H_{22}N_2O$

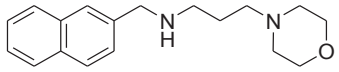
MW: 282.3801

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 8.47 (1H, s, CH=N), 8.1-7.9 (5H, m, ArH), 7.54 (2H, m, ArH), 3.8-3.7 (6H, m, CH_2x3), 2.49 (6H, m, CH_2x3), 1.97 (2H, m, CH_2).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 161.3, 134.7, 133.9, 133.1, 129.7, 128.9, 128.4, 127.9, 127.1, 126.5, 123.8, 67.0, 59.7, 56.7, 53.8, 27.8.

HRMS (ESI): calcd. $(M+H)^+$ 283.1732, exper. 283.1672.

Compound **63** - 3-morpholino-N-((naphthalen-3-yl)methyl)propan-1-amine



$C_{18}H_{24}N_2O$

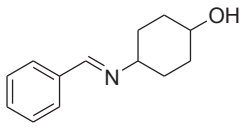
MW: 284.3960

^{1}H-NMR (300 MHz, $CDCl_3$) δ: 7.9-7.7 (4H, m, ArH), 7.5-7.4 (3H, m, ArH), 3.97 (2H, s, CH_2), 3.71 (4H, t, CH_2x2), 2.74 (2H, t, CH_2), 2.43 (6H, m, CH_2x3), 1.74 (2H, m, CH_2).

^{13}C-NMR (75 MHz, $CDCl_3$) δ: 137.9, 133.4, 132.6, 128.1, 127.7, 126.5, 126.4, 126.0, 125.6, 123.5, 102.0, 67.0, 57.4, 54.1, 53.8, 53.5, 48.0, 26.7.

HRMS (ESI): calcd. $(M+H)^+$ 285.1889, exper. 285.1988.

Compound **64** - 4-(benzylideneamino)cyclohexanol



$C_{13}H_{17}NO$

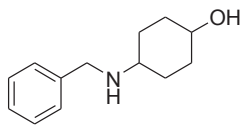
MW: 203.2802

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 8.36 (1H, s, CH=N), 7.74 (2H, m, ArH), 7.43 (3H, m, ArH), 3.75 (1H, m, CH), 3.23 (1H, m, CH), 2.09 (2H, m, CH_2), 1.9-1.7 (4H, m, CH_2x2), 1.54 (1H, bs, OH), 1.45 (2H, m, CH_2).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 159.4, 136.4, 130.5, 128.6, 128.1, 70.1, 68.9, 33.8, 32.2.

HRMS (ESI): calcd. $(M+H)^+$ 204.1310, exper. 204.1321.

Compound **65** - 4-(benzylamino)cyclohexanol



$C_{13}H_{19}NO$

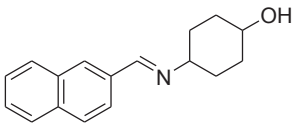
MW: 205.2961

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.4-7.3 (5H, m, ArH), 3.83 (2H, s, CH_2Ar), 3.64 (1H, m, CH), 2.53 (1H, m, CH), 2.00 (4H, m, CH_2x2), 1.28 (4H, m, CH_2x2).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 140.5, 128.5, 128.1, 126.1, 70.5, 55.4, 51.4, 33.9, 31.2.

HRMS (ESI): calcd. $(M+H)^+$ 206.1467, exper. 206.1362.

Compound **66** - 4-((naphthalen-3-yl)methyleneamino)cyclohexanol



$C_{17}H_{19}NO$

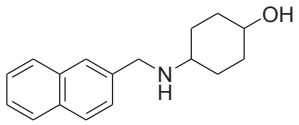
MW: 253.3389

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 8.50 (1H, s, CH=N), 8.1-7.8 (5H, m, ArH), 7.53 (2H, m, ArH), 3.77 (1H, m, CH), 3.28 (1H, m, CH), 2.12 (2H, m, CH_2), 1.9-1.7 (4H, m, CH_2x2), 1.48 (2H, m, CH_2).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 159.6, 134.7, 134.1, 133.1, 129.7, 128.6, 128.4, 127.9, 127.1, 126.4, 124.0, 70.2, 70.0, 69.1, 34.2, 33.8, 32.2.

HRMS (ESI): calcd. $(M+H)^+$ 254.1467, exper. 254.1565.

Compound **67** - 4-((naphthalen-3-yl)methylamino)cyclohexanol



$C_{17}H_{21}NO$

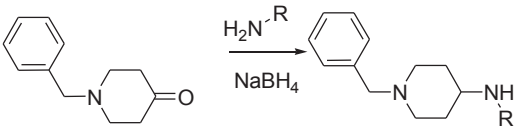
MW: 255.3547

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.9-7.8 (4H, m, ArH), 7.5-7.4 (3H, m, ArH), 3.99 (2H, s, CH_2Ar), 3.65 (1H, m, CH), 2.57 (1H, m, CH), 2.1-1.9 (4H, m, CH_2x2), 1.4-1.2 (4H, m, CH_2x2).

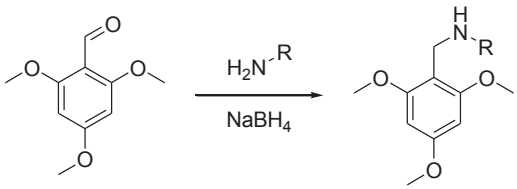
^{13}C-NMR (100 MHz, $CDCl_3$) δ: 138.1, 133.5, 132.6, 128.1, 127.7, 127.6, 126.5, 126.4, 126.0, 125.6, 70.5, 55.5, 51.5, 34.0, 31.3.

HRMS (ESI): calcd. $(M+H)^+$ 256.1623, exper. 256.1532.

PATHWAY 2.1.2:

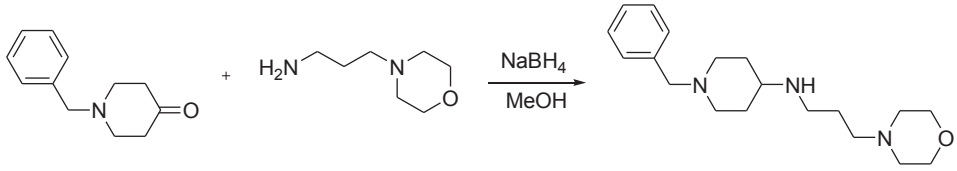


68



69

Synthesis of compound 71 (1-benzyl-N-(3-morpholinopropyl)piperidin-4-amine)



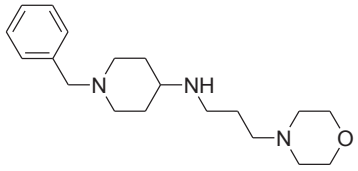
68 **70** **71**

To a stirring solution of N-benzyl piperidone (**68**, 500 mg, 0.490 mL, 2.65 mmol) in methanol (15 mL), 3-morpholinopropylamine (**70**, 380 mg, 0.386 mL, 2.65 mmol) was added. The reaction was heated at 60°C for 3 h, then cooled to 0°C in an ice bath and carefully added with sodium boron hydride (201 mg, 5.30 mmol). The reaction, monitored by TLC (hexane/ethyl acetate 4:1) was allowed to stir o/n at rt, then quenched at 0°C through the careful addition of a saturated aqueous solution of sodium hydroxy carbonate; the organic phase was then removed under reduced pressure, and the aqueous phase was extracted with dichloromethane (30 mL, 3 times). The combined organic phases were then washed with brine (20 mL, twice), dried with magnesium sulfate and the solvent was removed under reduced pressure to give a very dense yellow liquid as a product.

Yield: 78%.

With the same procedure, N-benzyl piperidine was reacted also with 2,2-diethoxyethanamine, with a yield of 71%. 2,4,6-trimethoxybenzaldehyde was reacted with allylamine and 2,2-diethoxyethanamine, to give the corresponding secondary amines. Yields were 60% and 67%, respectively.

Compound **71** - 1-benzyl-N-(3-morpholinopropyl)piperidin-4-amine



$C_{18}H_{30}N_2O_2$

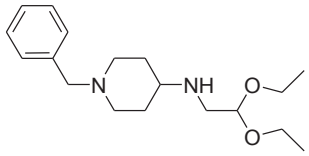
MW: 317.4689

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.4-7.2 (5H, m, ArH), 3.71 (4H, t, CH_2x2), 3.51 (2H, s, $ArCH_2$), 2.96 (1H, m, NH), 2.89 (2H, m, CH_2), 2.81 (2H, t, CH_2), 2.61 (1H, m, CH), 2.5-2.4 (6H, m, CH_2x3), 2.03 (2H, m, CH_2), 1.95 (2H, m, CH_2), 1.81 (2H, m, CH_2), 1.55 (2H, m, CH_2).

^{13}C-NMR (100MHz, $CDCl_3$) δ: 138.3, 129.0, 128.2, 127.0, 67.0, 62.9, 57.6, 55.0, 53.7, 53.4, 52.1, 45.3, 31.6.

HRMS (ESI): calcd. $(M+H)^+$ 318.2467, exper. 318.2448.

Compound **72** – 1-benzyl-N-(2,2-diethoxyethyl)piperidin-4-amine



$C_{18}H_{30}N_2O_2$

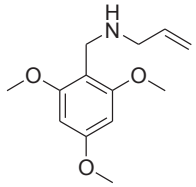
MW: 306.4430

^{1}H-NMR (300MHz, DMSO-d6) δ: 7.29 (5H, m, ArH), 4.46 (1H, t, CH), 3.58 (2H, m, CH_2), 3.46 (2H, m, CH_2), 3.41 (2H, s, CH_2Ar), 2.71 (2H, d, CH_2), 2.59 (1H, m, CH), 1.92 (2H, m, CH_2), 1.71 (2H, m, CH_2), 1.39 (2H, m, CH_2), 1.21 (2H, m, CH_2), 1.10 (6H, t, CH_3x2).

^{13}C-NMR (100MHz, DMSO-d6) δ: 138.7, 138.6, 128.6, 128.0, 126.7, 101.9, 62.3, 61.1, 51.8, 48.8, 34.4, 32.3, 15.3.

HRMS (ESI): calcd. $(M+H)^+$ 307.2307, exper. 307.2553.

Compound **73** - N-(2,4,6-trimethoxybenzyl)prop-2-en-1-amine



$C_{13}H_{19}NO_3$

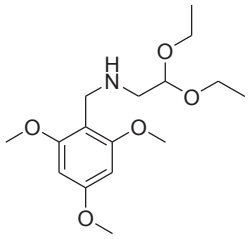
MW: 237.2949

^{1}H-NMR (300MHz, $CDCl_3$) δ: 6.04 (2H, "s", ArH), 5.9-5.8 (1H, m, CH), 5.1-4.9 (2H, m, CH$\underline{CH_2}$), 3.7-3.6 (11H, m, OCH_3x3 + $ArCH_2$), 3.12 (2H, m, NH$\underline{CH_2}$CH).

^{13}C-NMR (75 MHz, $CDCl_3$) δ: 160.3, 159.3, 137.3, 115.4, 108.8, 90.3, 55.6, 55.3, 51.5, 40.7.

HRMS (ESI): calcd. $(M+H)^+$ 238.1635, exper. 238.1495.

Compound **74** - N-(2,4,6-trimethoxybenzyl)-2,2-diethoxyethanamine



$C_{16}H_{27}NO_5$

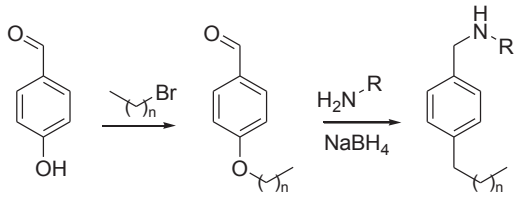
MW: 313.3893

^{1}H-NMR (400MHz, $CDCl_3$) δ: 5.91 (2H, s, ArH), 4.40 (1H, t, CH), 3.59 (2H, s, $ArCH_2$), 3.57 (6H, s, CH_3x2), 3.56 (3H, s, CH_3), 3.43 (2H, m, CH_2), 3.29 (2H, m, CH_2), 2.47 (2H, d, CH_2), 0.98 (6H, t, CH_3x2).

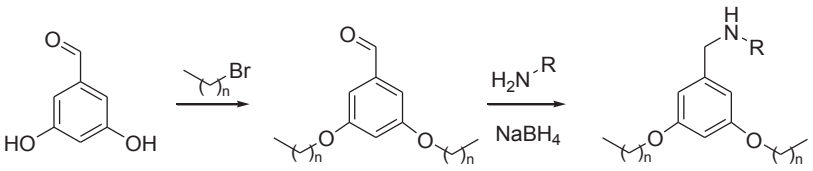
^{13}C-NMR (100MHz, $CDCl_3$) δ: 166.7, 155.8, 130.7, 102.3, 99.6, 61.9, 55.4, 50.1, 48.7, 15.2.

HRMS (ESI): calcd. $(M+H)^+$ 314.1889, exper. 314.2896.

PATHWAY 2.2:

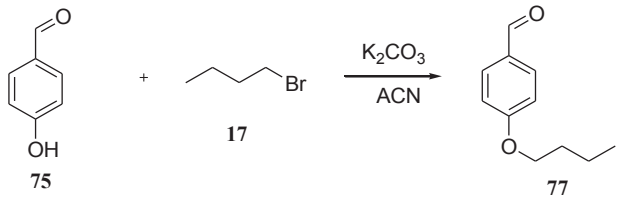


75



76

Synthesis of compound 77 (4-butoxybenzaldehyde)



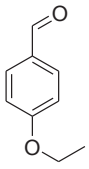
4-hydroxybenzaldehyde (**75**, 1.000 g, 8.19 mmol) was dissolved in 20 mL of acetonitrile in the presence of potassium carbonate (2.260 g, 16.38 mmol); the mixture was kept under magnetic stirring at rt for 1 h. Then, 1-bromobutane (**17**, 1.009 g, 0.791 mL, 7.37 mmol) was added and the reaction was heated at 60°C for 24 h and monitored by TLC (hexane/ethyl acetate 4:1).

After that, the reaction was filtered to remove the salts, and evaporated; the residue was dissolved in diethyl ether (30 mL) and washed with a solution of NaOH 10% (30 mL, 3 times) and brine (20 mL). The organic phase was dried with $MgSO_4$ and evaporated to give a colorless oily product.

Yield: 72%.

With the same procedure, the synthesis of more ethers was performed using other aloalkanes, *i.e.* 1-iodoethane, 1-bromopentane, 1-bromohexane, 1-bromoheptane and 1-bromooctane. For all of these reactions, the obtained yields were between 55 and 75%.

Compound **78** – 4-ethoxybenzaldehyde



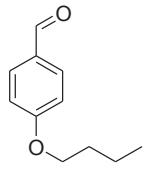
$C_9H_{10}O_2$

MW: 150.1745.

^{1}H-NMR (300MHz, $CDCl_3$) δ: 9.87 (1H, s, CHO), 7.82 (2H, AA'), 6.99 (2H, XX'), 4.10 (2H, q, OCH_2), 1.44 (3H, t, CH_3).

HRMS (ESI): calcd. $(M+H)^+$ 151.0681, exper. 151.1047.

Compound **77** - 4-butoxybenzaldehyde



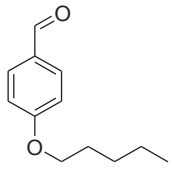
$C_{11}H_{14}O_2$

MW: 178.2277.

^{1}H NMR (400MHz, $CDCl_3$) δ: 9.89 (1H, s, CHO), 7.84 (2H, AA'), 7.01 (2H, XX'), 4.06 (2H, t, OCH_2), 1.81 (2H, m, CH_2), 1.52 (2H, m, CH_2), 1.00 (3H, t, CH_3).

HRMS (ESI): calcd. $(M+H)^+$ 179.0994, exper. 179.1127.

Compound **79** – 4-pentyloxybenzaldehyde



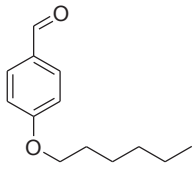
$C_{12}H_{16}O_2$

MW: 192.2542

^{1}H NMR (400MHz, $CDCl_3$) δ: 9.85 (1H, s, CHO), 7.79 (2H, AA'), 6.97 (2H, XX'), 4.01 (2H, t, OCH_2), 1.79 (2H, m, CH_2), 1.39 (4H, m, 2xCH_2), 0.92 (3H, t, CH_3).

HRMS (ESI): calcd. $(M+H)^+$ 193.1150, exper. 193.1250.

Compound **80** – 4-(hexyloxy)benzaldehyde



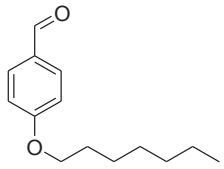
$C_{13}H_{18}O_2$

MW: 206.2808

^{1}H NMR (400MHz, $CDCl_3$) δ: 9.89 (1H, s, CHO), 7.84 (2H, AA'), 7.01 (2H, XX'), 4.02 (2H, t, OCH_2), 1.79 (2H, m, CH_2), 1.46 (4H, m, 2xCH_2), 1.34 (2H, m, CH_2), 0.91 (3H, t, CH_3).

HRMS (ESI): calcd. $(M+1)^+$ 207.1307, exper. 207.1439.

Compound **81** - 4-(heptyloxy)benzaldehyde



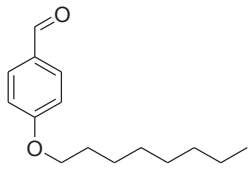
$C_{14}H_{20}O_2$

MW: 220.3074

^{1}H NMR (400MHz, $CDCl_3$) δ: 9.86 (1H, s, CHO), 7.81 (2H, AA'), 6.97 (2H, XX'), 4.02 (2H, t, OCH_2), 1.80 (2H, m, CH_2), 1.46 (2H, m, CH_2), 1.39-1.27 (6H, m, 3xCH_2), 0.89 (3H, t, CH_3).

HRMS (ESI): calcd. $(M+1)^+$ 221.1463, exper. 221.1531.

Compound **82** – 4-(octyloxy)benzaldehyde



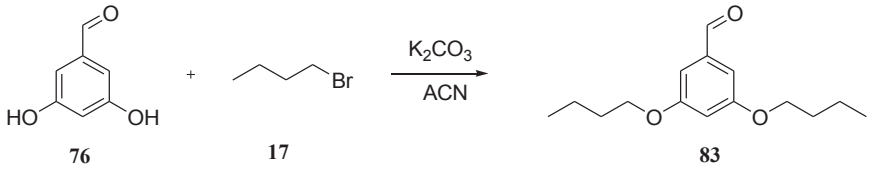
$C_{15}H_{22}O_2$

MW: 234.3339

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 9.90 (1H, s, CHO), 7.82 (2H, AA'), 7.03 (2H, XX'), 4.05 (2H, t, OCH_2), 1.81 (2H, m, CH_2), 1.50 (4H, m, CH_2x2), 1.32 (4H, m, CH_2x2), 1,20 (2H, m, CH_2), 0.91 (3H, t, CH_3).

HRMS (ESI): calcd. $(M+1)^+$ 235.1619, exper. 235.1741.

Synthesis of compound 83 (3,5-dibutoxybenzaldehyde)



76 **17** **83**

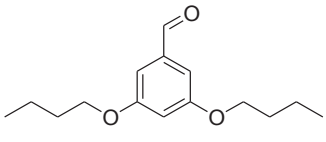
3,5-dihydroxybenzaldehyde (**76**, 100 mg, 0.72 mmol) was dissolved in 15 mL of acetonitrile in the presence of potassium carbonate (199 mg, 1.44 mmol); the mixture was kept under magnetic stirring at rt for 1 h. Then, 1-bromobutane (**17**, 178 mg, 0.140 mL, 1.30 mmol) was added and the reaction was heated at 60°C for 24 h and monitored by TLC (hexane/ethyl acetate 4:1).

Then, the reaction was filtered and the solvent evaporated; the residue was dissolved in diethyl ether (30 mL) and washed with a solution of NaOH 10% (30 mL, 3 times) and brine (20 mL). The organic phase was dried with $MgSO_4$ and evaporated to give a colorless oil as a product.

Yield: 34%.

With the same procedure, the synthesis was repeated on 3,5-dihydroxybenzaldehyde with 1-bromoheptane; in this case, the product had to be purified through flash chromatography (hexane/ethyl acetate 3:1; R_f=0.4). The final yield was 21%.

Compound **83** - 3,5-dibutoxybenzaldehyde



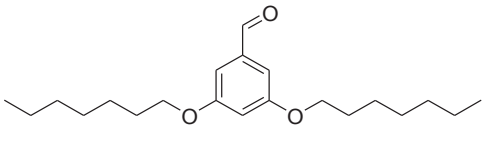
$C_{15}H_{22}O_3$

MW: 250.3334

^{1}H NMR (400 MHz, MeOD) δ: 9.87 (1H, s, CHO), 7.02 (2H, d, ArH), 6.73 (1H, t, ArH), 4.00 (4H, t, OCH_2x2), 1.7-1.8 (4H, m, CH_2x2), 1.6-1.4 (4H, m, CH_2x2); 1.00 (6H, t, CH_3x2).

HRMS (ESI): calcd. $(M+H)^+$ 251.1569, exper. 251.1686.

Compound **84** - 3,5-bis(heptyloxy)benzaldehyde



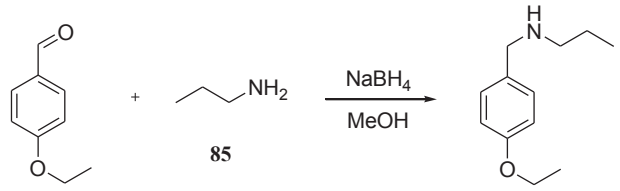
$C_{21}H_{34}O_3$

MW: 334.4929

^{1}H NMR (400 MHz, MeOD) δ: 9.88 (1H, s, CHO), 7.04 (2H, d, ArH), 6.76 (1H, t, ArH), 4.04 (4H, t, OCH_2x2), 1.8-1.7 (4H, m, CH_2x2), 1.5-1.3 (16H, m, CH_2x8), 0.94 (6H, t, CH_3x2).

HRMS (ESI): calcd. $(M+H)^+$ 335.2508, exper. 335.2651.

Synthesis of compound 86 (N-(4-ethoxybenzyl)propan-1-amine)



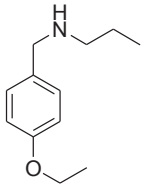
78 **86**

To a stirring solution of compound **78** (300 mg, 2 mmol) in methanol (15 mL), propylamine (**85**, 118 mg, 0.164 mL, 2 mmol) was added. The reaction was heated at 60°C for 3 h, then cooled to 0°C in an ice bath and carefully added with sodium boron hydride (152 mg, 4 mmol). The reaction was allowed to stir o/n at rt; after monitoring it by TLC (hexane/ethyl acetate 4:1), the morning after the stirring solution was quenched at 0°C through the careful addition of a saturated aqueous solution of sodium hydroxy carbonate; the organic phase was removed under reduced pressure, and the aqueous phase was extracted with dichloromethane (30 mL, 3 times). The combined organic phases were then washed with brine (20 mL, twice), dried with $MgSO_4$ and the solvent was removed under reduced pressure to give the product as a light yellow oil.

Yield: 63%.

With the same procedure also compounds **77**, **79, 80**, **81** and **82** were reacted with other amines, as ethanolamine, butylamine, serinol, 2,2-diethoxyethanamine, phenylethylamine, allylamine, 2-(pyrrolidin-1-yl)ethanamine and 2-(2-aminoethoxy)ethanol, to give the corresponding secondary amines. Yields were between 60 and 83%.

Compound **86** - N-(4-ethoxybenzyl)propan-1-amine



$C_{12}H_{19}NO$

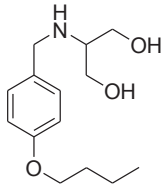
MW: 193.2854

^{1}H-NMR (400MHz, DMSO-d6) δ: 7.21 (2H, AA'), 6.84 (2H, XX'), 3.99 (2H, q, OCH_2), 3.60 (2H, s, CH_2Ar), 2.43 (2H, t, $NC\underline{H_2}CH_2$), 1.42 (2H, m, CH_2), 1.31 (3H, t, CH_3), 0.85 (3H, t, CH_3).

^{13}C-NMR (100 MHz, MeOD) δ: 158.7, 129.8, 122.6, 110.1, 102.2, 68.0, 66.4, 56.2, 28.4, 14.0.

HRMS (ESI): calcd. $(M+H)^+$ 194.1467, exper. 194.1526.

Compound **87** - 2-(4-butoxybenzylamino)propane-1,3-diol



$C_{14}H_{23}NO_3$

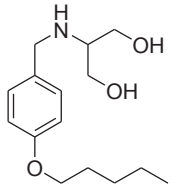
MW: 253.3373

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.24 (2H, AA'), 6.87 (2H, XX'), 3.96 (2H, t, OCH_2), 3.77 (2H, s, $C\underline{H_2}NH$), 3.73 (2H, m, $C\underline{H_2}OH$), 3.60 (2H, m, $C\underline{H_2}OH$), 2.79 (1H, m, CH), 1.78 (2H, m, CH_2), 1.59 (2H, m, CH_2), 0.99 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 158.6, 129.0, 127.8, 114.3, 66.1, 61.1, 58.3, 51.2, 31.4, 23.6, 13.8.

HRMS (ESI): calcd. $(M+H)^+$ 254.1678, exper. 254.1768.

Compound **88** - 2-(4-pentyloxybenzylamino)propane-1,3-diol



$C_{15}H_{25}NO_3$

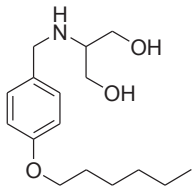
MW: 267.3639

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.21 (2H, AA'), 6.86 (2H, XX'), 3.94 (2H, t, OCH_2), 3.72 (2H, s, $C\underline{H_2}NH$), 3.67 (2H, m, $C\underline{H_2}OH$), 3.58 (2H, m, $C\underline{H_2}OH$), 2.72 (1H, m, CH), 1.79 (2H, m, CH_2), 1.43 (4H, m, CH_2x2), 0.96 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 158.1, 129.0, 127.9, 114.0, 66.8, 61.2, 58.4, 50.0, 31.0, 22.1, 19.0, 13.7.

HRMS (ESI): calcd. $(M+H)^+$ 268.1834, exper. 268.1752.

Compound **89** - 2-(4-(hexyloxy)benzylamino)propane-1,3-diol



$C_{16}H_{27}NO_3$

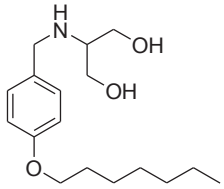
MW: 281.3905

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.23 (2H, AA'), 6.86 (2H, XX'), 3.94 (2H, t, OCH_2), 3.75 (2H, s, $C\underline{H_2}NH$), 3.70 (2H, m, $C\underline{H_2}OH$), 3.59 (2H, m, $C\underline{H_2}OH$), 2.67 (1H, m, CH), 1.78 (2H, m, CH_2), 1.47 (2H, m, CH_2), 1.36 (4H, m, CH_2x2), 0.93 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 158.3, 129.2, 128.1, 114.2, 66.4, 61.0, 58.3, 50.1, 31.2, 26.0, 22.7, 19.0, 13.8.

HRMS (ESI): calcd. $(M+1)^+$ 282.1991, exper. 282.2165.

Compound **90** - 2-(4-(heptyloxy)benzylamino)propane-1,3-diol



$C_{17}H_{29}NO_3$

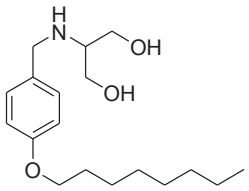
MW: 295.4171

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.21 (2H, AA'), 6.85 (2H, XX'), 3.93 (2H, t, OCH_2), 3.72 (2H, s, $C\underline{H_2}NH$), 3.67 (2H, m, $C\underline{H_2}OH$), 3.58 (2H, m, $C\underline{H_2}OH$), 2.71 (1H, m, CH), 1.78 (2H, m, CH_2), 1.46 (2H, m, CH_2), 1.4-1.2 (6H, m, CH_2x3), 0.92 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 158.1, 129.2, 128.3, 114.4, 66.7, 61.1, 58.1, 50.2, 31.4, 26.1, 22.9, 19.0, 18.9, 13.7.

HRMS (ESI): calcd. $(M+1)^+$ 296.2147, exper. 296.2299.

Compound **91** - 2-(4-(octyloxy)benzylamino)propane-1,3-diol



$C_{18}H_{31}NO_3$

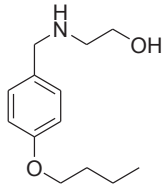
MW: 309.4436

^{1}H-NMR (400 MHz. $CDCl_3$) δ: 7.24 (2H, m, AA'). 6.87 (2H, m, XX'), 3.95 (2H, t, OCH_2), 3.78 (2H, s, $C\underline{H_2}NH$), 3.74 (2H, m, $C\underline{H_2}OH$), 3.60 (2H, m, $C\underline{H_2}OH$), 2.82 (1H, m, CH), 1.80 (2H, m, CH_2), 1.33 (10H, m, CH_2x5), 0.92 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 158.2, 129.4, 128.5, 114.3, 67.7, 62.1, 58.2, 50.3, 31.3, 29.7, 26.0, 22.8, 19.1, 19.0, 13.8.

HRMS (ESI): calcd. $(M+1)^+$ 310.2303, exper. 310.2431.

Compound **92** - 2-(4-butoxybenzylamino)ethanol



$C_{13}H_{21}NO_2$

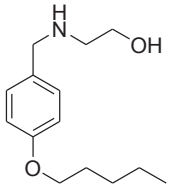
MW: 223.3113

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.23 (2H, AA'), 6.88 (2H, XX'), 3.97 (2H, t, OCH_2), 3.76 (2H, s, CH_2Ar), 3.67 (2H, t, CH_2), 2.82 (2H, t, CH_2), 1.78 (2H, m, CH_2), 1.51 (2H, m, CH_2), 0.99 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 159.1, 131.1, 128.9, 114.0, 103.2, 77.1, 76.1, 68.2, 52.7, 28.1, 22.3, 14.0.

HRMS (ESI): calcd. $(M+H)^+$ 224.1572, exper. 224.1632.

Compound **93** - 2-(4-(pentyloxy)benzylamino)ethanol



$C_{14}H_{23}NO_2$

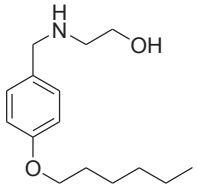
MW: 237.3379

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.23 (2H, AA'), 6.87 (2H, XX'), 3.95 (2H, t, OCH_2), 3.75 (2H, s, CH_2Ar), 3.67 (2H, t, CH_2), 2.80 (2H, t, CH_2), 1.80 (2H, m, CH_2), 1.43 (4H, m, CH_2x2), 0.95 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 159.3, 130.1, 127.9, 114.1, 103.3, 77.3, 77.0, 76.7, 68.0, 52.6, 28.2, 22.4, 14.0.

HRMS (ESI): calcd. $(M+H)^+$ 238.1729, exper. 238.1893.

Compound **94** - 2-(4-(hexyloxy)benzylamino)ethanol



$C_{15}H_{25}NO_2$

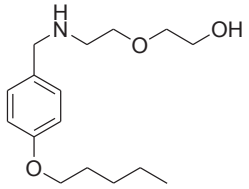
MW: 251.3645

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.22 (2H, AA'), 6.86 (2H, XX'), 3.96 (2H, t, OCH_2), 3.76 (2H, s, CH_2Ar), 3.68 (2H, t, CH_2), 2.81 (2H, t, CH_2), 1.79 (2H, m, CH_2), 1.48 (2H, m, CH_2), 1.36 (4H, m, CH_2x2) 0.93 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 159.0, 130.3, 128.2, 114.3, 103.2, 77.2, 77.0, 76.6, 74.8, 68.2, 52.5, 28.1, 22.2, 14.1.

HRMS (ESI): calcd. $(M+H)^+$ 252.1885, exper. 252.1985.

Compound **95** - 2-(2-(4-(pentyloxy)benzylamino)ethoxy)ethanol



$C_{16}H_{27}NO_3$

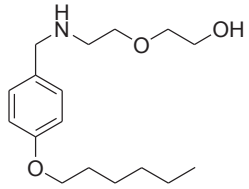
MW: 281.3905

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.23 (2H, AA'), 6.86 (2H, XX'), 3.95 (2H, t, OCH_2), 3.75 (2H, s, CH_2Ar), 3.71 (2H, t, CH_2), 3.62 (2H, t, CH_2), 3.58 (2H, t, CH_2), 2.82 (2H, t, CH_2), 1.79 (2H, m, CH_2), 1.43 (4H, m, CH_2x2), 0.95 (3H, t, CH_3).

^{13}C-NMR (100MHz, $CDCl_3$) δ: 158.2, 130.0, 129.1, 114.1, 72.7, 69.8, 68.9, 61.7, 52.8, 48.3, 29.5, 26.3, 22.4, 14.0.

HRMS (ESI): calcd. $(M+H)^+$ 282.1991, exper. 282.1897.

Compound **96** - 2-(2-(4-(hexyloxy)benzylamino)ethoxy)ethanol



$C_{17}H_{29}NO_3$

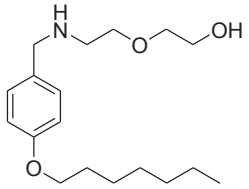
MW: 295.4171

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.21 (2H, AA'), 6.84 (2H, XX'), 3.93 (2H, t, OCH_2), 3.72 (2H, s, CH_2Ar), 3.68 (2H, t, CH_2), 3.59 (2H, t, CH_2), 3.54 (2H, t, CH_2), 2.79 (2H, t, CH_2), 1.77 (2H, m, CH_2), 1.45 (2H, m, CH_2), 1.34 (4H, m, CH_2x2), 0.91 (3H, t, CH_3).

^{13}C-NMR (100MHz, $CDCl_3$) δ: 158.4, 130.3, 129.2, 114.4, 72.6, 69.5, 68.1, 61.5, 52.9, 48.0, 31.6, 29.1, 26.2, 22.2, 14.1.

HRMS (ESI): calcd. $(M+H)^+$ 296.2147, exper. 296.2136.

Compound **97** - 2-(2-(4-(heptyloxy)benzylamino)ethoxy)ethanol



$C_{18}H_{31}NO_3$

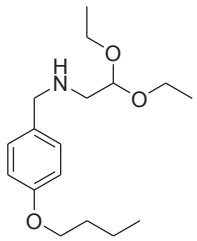
MW: 309.4436

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.23 (2H, AA'), 6.85 (2H, XX'), 3.95 (2H, t, CH_2), 3.76 (2H, s, CH_2Ar), 3.72 (2H, t, OCH_2), 3.63 (2H, t, OCH_2), 3.58 (2H, t, OCH_2), 3.36 (1H, bs, NH), 2.83 (2H, t, CH_2), 1.78 (2H, m, CH_2), 1.46 (2H, m, CH_2), 1.4-1.3 (6H, m, CH_2x3), 0.91 (3H, t, CH_3) .

^{13}C-NMR (100MHz, $CDCl_3$) δ: 158.5, 130.6, 129.6, 114.5, 72.5, 69.6, 68.0, 61.6, 52.8, 48.2, 31.8, 29.3, 29.0, 26.0, 22.6, 14.1.

HRMS (ESI): calcd. $(M+H)^+$ 310.2304, exper. 310.2488.

Compound **98** - N-(4-butoxybenzyl)-2,2-diethoxyethanamine



$C_{17}H_{29}NO_3$

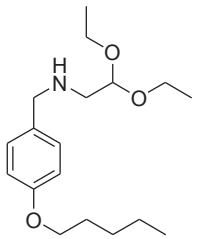
MW: 295.4171

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.20 (2H, AA'), 6.83 (2H, XX'), 4.60 (1H, t, C<u>H</u>CH_2), 3.92 (2H, t, OCH_2), 3.72 (2H, s, CH_2Ar), 3.66 (2H, m, OCH_2), 3.51 (2H, m, OCH_2), 2.72 (2H, d, CHC<u>H_2</u>), 1.74 (2H, m, CH_2) 1.48 (2H, m, CH_2), 1.24 (6H, t, J=14.0 Hz, CH_3x2), 0.96 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 158.3, 131.9, 129.3, 114.4, 102.1, 67.6, 62.3, 53.3, 51.4, 31.8, 19.3, 15.4, 13.9.

HRMS (ESI): calcd. $(M+H)^+$ 296.2147, exper. 296.2304.

Compound **99** - N-(4-(pentyloxy)benzyl)-2,2-diethoxyethanamine



$C_{18}H_{31}NO_3$

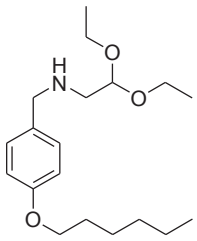
MW: 309.4436

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.22 (2H, AA'), 6.85 (2H, XX'), 4.62 (1H, t, C<u>H</u>CH_2), 3.94 (2H, t, OCH_2), 3.74 (2H, s, CH_2Ar), 3.68 (2H, m, OCH_2), 3.53 (2H, m, OCH_2), 2.74 (2H, d, CHC<u>H_2</u>), 1.78 (2H, m, CH_2), 1.42 (4H, m, CH_2x2), 1.21 (6H, t, CH_3x2), 0.94 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 158.2, 131.8, 129.1, 114.2, 102.0, 67.3, 62.3, 53.1, 51.6, 31.7, 22.4, 19.0, 15.2, 13.8.

HRMS (ESI): calcd. $(M+H)^+$ 310.2304, exper. 310.2430.

Compound **100** - N-(4-(hexyloxy)benzyl)-2,2-diethoxyethanamine



$C_{19}H_{33}NO_3$

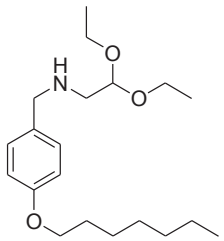
MW: 323.4702

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.22 (2H, AA'), 6.86 (2H, XX'), 4.63 (1H, t, C<u>H</u>CH_2), 3.95 (2H, t, OCH_2), 3.75 (2H, s, CH_2Ar), 3.69 (2H, m, OCH_2), 3.54 (2H, m, OCH_2), 2.75 (2H, d, CHC<u>H</u>$_2$), 1.78 (2H, m, CH_2), 1.47 (2H, m, CH_2), 1.36 (4H, m, CH_2x2), 1.21 (6H, t, CH_3x2), 0.92 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 158.6, 131.5, 129.2, 114.3, 101.9, 67.2, 62.5, 53.4, 51.5, 31.8, 27.3, 21.4, 19.5, 15.1, 13.9.

HRMS (ESI): calcd. $(M+H)^+$ 324.2460, exper. 324.2321.

Compound **101** - N-(4-(heptyloxy)benzyl)--2,2-diethoxyethanamine



$C_{20}H_{35}NO_3$

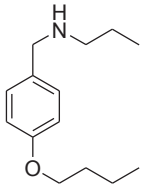
MW: 337.4968

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.21 (2H, AA'), 6.84 (2H, XX'), 4.64 (1H, t, C<u>H</u>CH_2), 3.96 (2H, t, OCH_2), 3.77 (2H, s, CH_2Ar), 3.66 (2H, m, OCH_2), 3.51 (2H, m, OCH_2), 2.74 (2H, d, CHC<u>H$_2$</u>), 1.77 (2H, m, CH_2), 1.42 (2H, m, CH_2), 1.30 (6H, m, CH_2x3), 1.20 (6H, t, CH_3x2), 0.90 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 158.3, 131.9, 129.4, 114.4, 102.1, 68.0, 62.4, 53.2, 51.4, 31.8, 29.3, 29.1, 26.0, 22.6, 15.4, 14.1.

HRMS (ESI): calcd. $(M+H)^+$ 338.2617, exper. 338.2812.

Compound **102** - N-(4-butoxybenzyl)propan-1-amine



$C_{14}H_{23}NO$

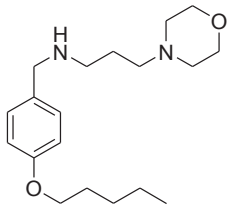
MW: 221.3385

^{1}H-NMR (300MHz, MeOD): δ: 7.22 (2H, AA'), 6.83 (2H, XX'), 3.91 (2H, t, OCH_2), 2.52 (2H, t, CH_2Ar), 2.25 (2H, m, CH_2), 1.54 (6H, m, CH_2x3), 0.99 (6H, m, CH_3x2).

^{13}C-NMR (100 MHz, MeOD) δ: 157.4, 127.2, 123.4, 111.9, 102.0, 69.0, 63.4, 55.2, 31.7, 29.6, 22.1, 15.0.

HRMS (ESI): calcd. $(M+H)^+$ 222.1780, exper. 222.1796.

Compound **103** – N-(4-(pentyloxy)benzyl)-3-morpholinopropan-1-amine



$C_{19}H_{32}N_2O_2$

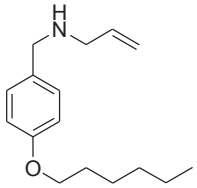
MW: 320.4696

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.20 (2H, AA'), 6.83 (2H, XX'), 3.90 (2H, m, CH_2), 3.68 (6H, m, CH_2x3), 2.65 (2H, m, CH_2), 2.38 (6H, m, CH_2x3) 1.68 (2H, m, CH_2), 1.78 (2H, m, CH_2), 1.40 (4H, m, CH_2x2); 0.92 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 158.1, 132.1, 129.1, 114.3, 67.9, 66.9, 57.3, 53.7, 53.3, 47.8, 28.9, 28.1, 26.5, 22.4, 13.9.

HRMS (ESI): calcd. $(M+H)^+$ 321.2464, exper. 321.2694.

Compound **104** – N-(4-(hexyloxy)benzyl)prop-2-en-1-amine



$C_{16}H_{25}NO$

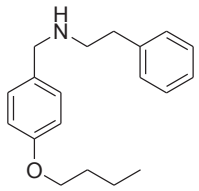
MW: 247.3758

^{1}H-NMR (300MHz, MeOD): δ: 7.08 (AA', 2H), 6.71 (XX', 2H), 5.79 (1H, m, CH), 5.07 (1H, dd, CH), 4.98 (1H, dd, CH), 3.78 (2H, t, OCH_2), 3.57 (2H, s, CH_2Ar), 3.12 (2H, d, CH_2), 1.63 (2H, m, CH_2), 1.10 (6H, m, CH_2x3), 0.30 (3H, t, CH_3).

^{13}C-NMR (100MHz, MeOD):158.3, 136.8, 136.7, 132.0, 129.3, 114.3, 68.2, 52.7, 51.5, 31.4, 29.2, 25.9, 22.7, 14.4.

HRMS (ESI): calcd. $(M+H)^+$ 248.1936, exper. 248.1943.

Compound **105** - N-(4-butoxybenzyl)-2-phenylethanamine



$C_{19}H_{25}NO$

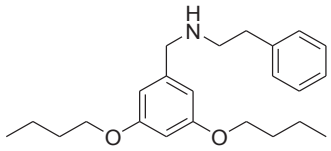
MW: 283.4079

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.4-7.2 (7H, m, ArH), 6.91 (2H, m, ArH), 4.00 (2H, m, CH_2), 3.77 (2H, s, CH_2Ar), 3.0-2.8 (4H, m, CH_2x2), 1.81 (2H, m, CH_2), 1.55 (2H, m, CH_2), 1.02 (3H, t, CH_3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 159.1, 143.1, 140.9, 130.7, 130.1, 127.0, 107.1, 102.4, 70.2, 55.4, 50.2, 37.2, 30.4, 15.8, 11.1.

HRMS (ESI): calcd. $(M+H)^+$ 284.1936, exper. 284.2113.

Compound **106** - N-(3,5-dibutoxybenzyl)-2-phenylethanamine



$C_{23}H_{33}NO_2$

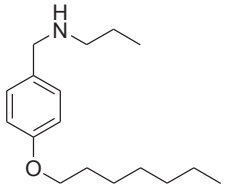
MW: 355.5136

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.3-7.2 (5H, m, ArH), 6.44 (2H, d, J=2.41 Hz, ArH), 6.36 (1H, t, J=4.42 Hz, ArH), 3.92 (4H, t, OCH_2x2), 3.68 (2H, s, CH_2Ar), 2.82 (4H, m, CH_2x2), 1.8-1.7 (4H, m, CH_2x2), 1.5-1.4 (4H, m, CH_2x2), 0.98 (6H, t, CH_3x2).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 160.4, 141.1, 139.5, 128.3, 128.1, 125.9, 106.4, 100.0, 67.3, 52.8, 49.6, 35.2, 31.0, 18.8, 12.9.

HRMS (ESI): calcd. $(M+H)^+$ 356.2511, exper. 356.2625.

Compound **107** - N-(4-(heptyloxy)benzyl)propan-1-amine



$C_{17}H_{29}NO$

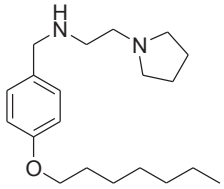
MW: 263.4183

^{1}H-NMR (300MHz, MeOD): δ: 7.20 (2H, AA'), 6.83 (2H, XX'), 3.91 (2H, t, OCH_2), 3.71 (2H, s, CH_2), 3.61 (2H, s, CH_2Ar), 2.26 (2H, t, NCH_2), 1.78 (2H, m, CH_2), 1.47 (4H, m, CH_2x2), 1.32 (8H, m, CH_2x4), 0.89 (6H, m, CH_3x2).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 158.4, 139.9, 107.7, 99.8, 69.6, 56.2, 48.8, 31.2, 30.9, 28.3, 28.0, 25.4, 23.3, 21.3, 14.0.

HRMS (ESI): calcd. $(M+H)^+$ 264.2249, exper. 264.2254.

Compound **108** - N-(4-(heptyloxy)benzyl)-2-(pyrrolidin-1-yl)ethanamine



$C_{20}H_{34}N_2O$

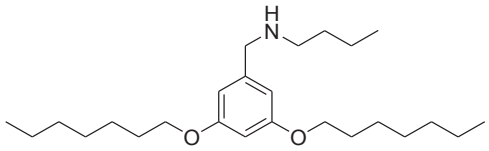
MW: 318.4968

^{1}H-NMR (400 MHz, acetone-d6) δ: 7.32 (2H, AA'), 6.86 (2H, XX'), 3.97 (2H, t, OCH_2), 3.72 (2H, s, NCH_2), 2.69 (2H, m, CH_2), 2.62 (2H, m, CH_2), 2.5-2.4 (4H, m, CH_2x2), 2.08 (2H, m, CH_2), 1.8-1.7 (m, 4H, CH_2x2), 1.48 (2H, m, CH_2), 1.4-1.2 (6H, m, CH_2x3), 0.92 (3H, t, CH_3).

^{13}C-NMR (75 MHz, $CDCl_3$) δ: 158.2, 129.9, 129.4, 114.3, 67.9, 55.5, 54.3, 54.1, 53.2, 47.2, 31.8, 29.2, 29.0, 23.4, 22.6, 14.1

HRMS (ESI): calcd. $(M+H)^+$ 319.2671, exper. 319.2770.

Compound **109** - N-(3,5-bis(heptyloxy)benzyl)butan-1-amine



$C_{25}H_{45}NO_2$

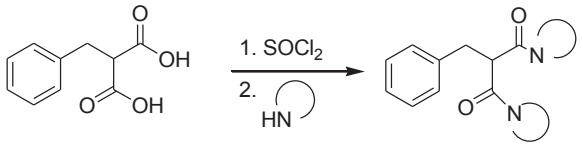
MW: 391.6303

^{1}H-NMR (400 MHz, $CDCl_3$): 6.45 (2H, d, J= 2.02 Hz, ArH), 6.34 (1H, t, J=4.43 Hz, ArH), 3.92 (4H, t, OCH_2x2), 3.71 (2H, s, CH_2Ar), 2.62 (2H, t, NCH_2), 1.75 (4H, m, CH_2x2), 1.5-1.3 (20H, m, CH_2x10), 0.91 (9H, m, CH_3x3).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 160.4, 142.9, 106.4, 99.9, 68.0, 54.2, 49.2, 32.3, 31.8, 29.3, 29.1, 26.0, 22.6, 20.5, 14.1, 14.0.

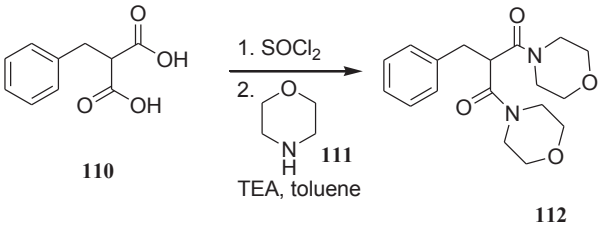
HRMS (ESI): calcd. $(M+H)^+$ 392.3450, exper. 392.3650.

PATHWAY 3:



110

Synthesis of compound 112 (2-benzyl-1,3-dimorpholinopropane-1,3-dione)



110 **112**

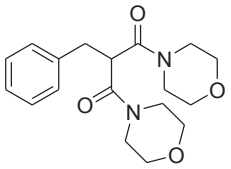
In a round bottom flask benzylmalonic acid (**110**, 100 mg, 0.52 mmol) is placed together with thyonil chloride (3 mL); the mixture is heated at 60°C and monitored by TLC (hexane/ethyl acetate 4:1). After 1 h, the reaction is cooled and the solvent evaporated; the obtained chloride is used for the following step without further purifications.

Morpholine (**111**, 90 mg, 0.089 mL, 1.04 mmol) is dissolved in anhydrous toluene (10 mL), and the freshly prepared chloride (dissolved in toluene, 5 mL) is subsequently added. After 1 h of stirring at rt, triethylamine (105 mg, 0.144 mL, 1.04 mmol) is added and the mixture is heated o/n at 60°C. The reaction is monitored by TLC (hexane/ethyl acetate 4:1) and, once completed, is cooled to rt and added with DCM (20 mL). The mixture is washed with water (20 mL, 3 times), dried with magnesium sulfate and the solvent evaporated under reduced pressure.

Yield: 53%.

The same reaction was repeated on benzylmalonic acid using piperidine instead of morpholine; in this case, the yield was 28%.

Compound **112** - 2-benzyl-1,3-dimorpholinopropane-1,3-dione



$C_{18}H_{24}N_2O_4$

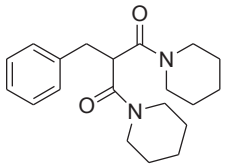
MW: 332.3942

^{1}H-NMR (400 MHz, acetone-d6) δ: 7.3-7.2 (5H, m, ArH), 4.27 (1H, t, C<u>H</u>CH_2), 3.6-3.2 (16H, m, CH_2x8), 3.16 (2H, d, C<u>H_2</u>CH).

^{13}C-NMR (100 MHz, acetone-d6) δ: 167.8, 139.5, 129.4, 128.3, 126.5, 66.1, 47.8, 45.7, 35.8.

HRMS (ESI): calcd. $(M+H)^+$ 333.3924, exper. 333.2037.

Compound **113** - 2-benzyl-1,3-di(piperidin-1-yl)propane-1,3-dione



$C_{20}H_{28}N_2O_2$

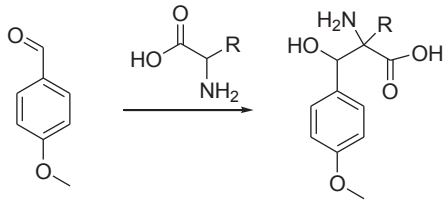
MW: 328.4485

^{1}H-NMR (400 MHz, acetone-d6) δ: 7.3-7.2 (5H, m, ArH), 3.95 (1H, t, C<u>H</u>CH_2), 3.7-3.6 (2H, m, CH_2), 3.5-3.3 (4H, m, CH_2x2), 3.26 (2H, d, C<u>H_2</u>CH), 3.2-3.1 (2H, m, CH_2), 1.7-1.3 (12H, m, CH_2x6).

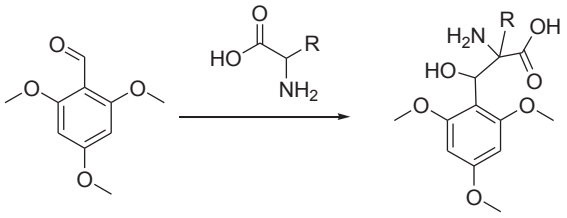
^{13}C-NMR (100 MHz, acetone-d6) δ: 167.2, 137.3, 129.1, 125.3, 123.5, 66.3, 48.8, 45.0, 42.2, 33.9.

HRMS (ESI): calcd. $(M+H)^+$ 329.2151, exper. 329.2274.

PATHWAY 4.1:

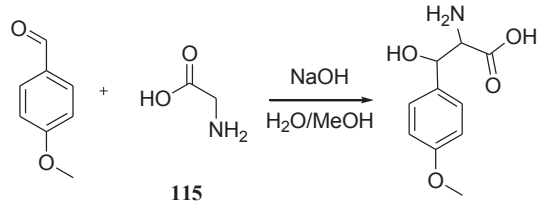


114



69

Synthesis of compound 116 (2-amino-3-hydroxy-3-(4-methoxyphenyl)propanoic acid)



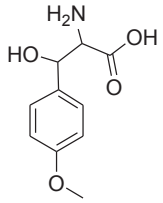
114 **115** **116**

In a round bottom flask containing water and methanol (1.5 and 2.5 mL respectively), were placed glycine (**115**, 50 mg, 0.67 mmol) and sodium hydroxide (80 mg, 2 mmol); then, p-anisaldehyde was added (**114**, 182 mg, 0.163 mL, 1.34 mmol) and the reaction was allowed to stir o/n at rt. The solution was then acidified with concd. hydrochloric acid, and allowed to stir 60 more minutes. The solvent was evaporated with the help of subsequent additions of ethanol, then the residue was added with an aqueous solution of sodium hydroxide, 10%. The reaction was stirred for 2 more h, until it cooled. The suspension was then filtered through a Buchner funnel and the solid white residue dried in a heater for 24 h.

Yield: 63%.

The same reaction has been repeated on p-anisaldehyde also with serine and threonine, and on 2,4,6-trimethoxybenzaldehyde with glycine. All the yields were between 49 and 65%.

Compound **116** – 2-amino-3-hydroxy-3-(4-methoxyphenyl)propanoic acid



$C_{10}H_{13}NO_4$

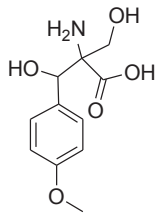
MW: 211.2145

^{1}H-NMR (300MHz, DMSO-d6) δ: 8.10 (1H, s, OH), 7.67 (2H, AA'), 6.97 (2H, XX'), 3.95 (2H, "d", NH_2), 3.87 (1H, s, OH), 3.79 (3H, s, OCH_3), 3.77 (1H, d, CH), 3.66 (1H, d, CH).

^{13}C-NMR (100MHz, DMSO-d6) δ: 161.3, 160.9, 130.1, 115.3, 114.4, 66.8, 55.6.

HRMS (ESI): calcd. $(M+H)^+$ 212.0845, exper. 212.0763.

Compound **117** - 2-amino-3-hydroxy-2-(hydroxymethyl)-3-(4-methoxyphenyl)propanoic acid



$C_{11}H_{15}NO_5$

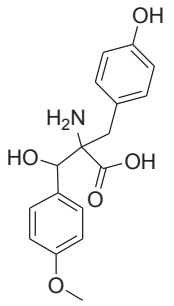
MW: 241.2405

^{1}H-NMR (300MHz, DMSO-d6) δ: 8.10 (1H, s, OH), 7.67 (2H, AA'), 6.98 (2H, XX'), 5.79 (bs, OH), 3.80 (3H, s, CH_3), 3.73 (1H, "t", CHH), 3.60 (1H, "t", CHH), 3.49 (1H, "s", CH).

^{13}C-NMR (100MHz, DMSO-d6) δ:206.9, 174.9, 161.3, 129.8, 114.3, 75.8, 64.8, 55.8, 31.2.

HRMS (ESI): calcd. $(M+H)^+$ 242.0950, exper. 242.1135.

Compound **118** - 2-(4-hydroxybenzyl)-2-amino-3-hydroxy-3-(4-methoxyphenyl)propanoic acid



$C_{17}H_{19}NO_5$

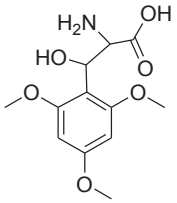
MW: 317.3365

^{1}H-NMR (300MHz, DMSO-d6) δ: 7.83 (1H, s, OH), 7.60 (2H, AA'), 6.94 (2H, XX'), 6.69 (2H, AA'), 6.28 (2H, XX'), 3.78 (3H, s, OCH_3), 3.66 (1H, dd, CH, J=4.5 Hz, J=4.5 Hz), 3.01 (dd, 1H, C<u>H</u>H, J=13.0 Hz, J=4.5 Hz), 2.69 (dd, 1H, CH<u>H</u>, J=13.0 Hz, J=4.5 Hz).

^{13}C-NMR (100MHz, DMSO-d6) δ: 174.9, 160.7, 158.4, 155.1, 130.1, 129.3, 114.6, 113.8, 79.2, 56.0, 55.2, 18.5.

HRMS (ESI): calcd. $(M+H)^+$ 318.1263, exper. 318.1158.

Compound **119** - 2-amino-3-hydroxy-3-(2,4,6-trimethoxyphenyl)propanoic acid



$C_{12}H_{17}NO_6$

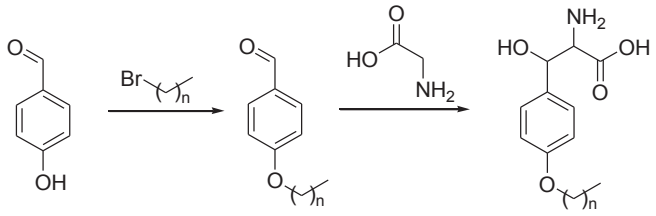
MW: 271.2665

^{1}H-NMR (300MHz, DMSO-d6) δ: 8.41 (1H, s, OH), 6.30 (2H, s, ArH), 3.91 (1H, d, CH), 3.83 (9H, s, CH_3x3), 3.18 (1H, d, CH).

^{13}C-NMR (100MHz, $CDCl_3$) δ: 175.2, 165.6, 108.6, 99.6, 73.6, 61.5, 55.9.

HRMS (ESI): calcd. $(M+H)^+$ 272.1056, exper. 272.3012.

PATHWAY 4.2:

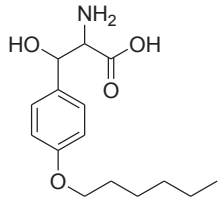


75

The synthesis of these compounds is accordant to the procedures already described. The procedure explained for compound **77** (*Williamson* reaction) was performed to obtain compounds **80**, **81** and **82**, the characterization of which have already been reported in this thesis. The procedure described for compound **116** (*Knoevenagel* reaction modified by *Portelli*) was applied to obtain compounds **120**, **121** and **122**.

Yields: between 50 and 55%.

Compound **120** - 2-amino-3-(4-(hexyloxy)phenyl)-3-hydroxypropanoic acid



$C_{15}H_{23}NO_4$

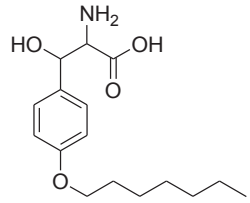
MW: 281.3474

^{1}H-NMR (400MHz, D_2O) δ: 7.81 (2H, AA'), 6.98 (2H, XX'), 4.10 (2H, t, OCH_2), 3.57 (1H, d, CH), 2.19 (1H, d, CH), 1.74 (2H, m, CH_2), 1.41 (2H, m, CH_2), 1.3-1.2 (m, 4H, CH_2x2), 0.83 (3H, t, CH_3).

^{13}C-NMR (100MHz, D_2O) δ: 158.5, 131.0, 130.6, 55.6, 31.2, 29.3, 29.0, 26.0, 22.6, 14.1.

HRMS (ESI): calcd. $(M+H)^+$ 282.1627, exper. 282.2008.

Compound **121** - 2-amino-3-(4-(heptyloxy)phenyl)-3-hydroxypropanoic acid



$C_{16}H_{25}NO_4$

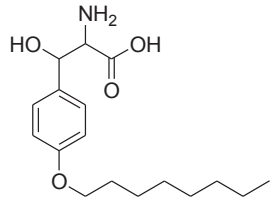
MW: 295.3740

^{1}H-NMR (400MHz, D_2O) δ: 7.79 (2H, AA'), 6.97 (2H, XX'), 4.08 (2H, t, OCH_2), 3.87 (1H, d, CH), 3.62 (1H, d, CH), 1.74 (2H, m, CH_2), 1.4-0.9 (6H, m, CH_2x3), 0.81 (5H, m, CH_2CH_3).

^{13}C-NMR (100MHz, D_2O) δ: 158.7, 130.8, 130.4, 55.6, 35.5, 31.2, 29.3, 29.0, 26.8, 24.6, 18.1.

HRMS (ESI): calcd. $(M+H)^+$ 296.3740, exper. 296.3539.

Compound **122** - 2-amino-3-hydroxy-3-(4-(octyloxy)phenyl)propanoic acid



$C_{17}H_{27}NO_4$

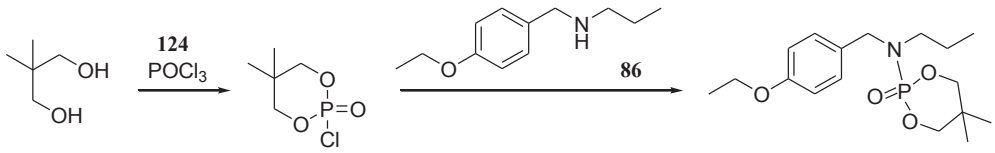
MW: 309.4006

^{1}H-NMR (400MHz, D_2O+NaOH) δ: 7.77 (2H, AA'), 6.95 (2H, XX'), 4.06 (2H, t, OCH_2), 3.52 (1H, d, CH), 3.08 (1H, d, CH), 2.11 (2H, m, CH_2), 1.70 (2H, m, CH_2), 1.4-1.2 (8H, m, CH_2x4), 1.09 (3H, t, CH_3).

^{13}C-NMR (100MHz, D_2O) δ: 155.4, 127.9, 120.4, 48.3, 35.5, 31.2, 29.3, 29.0, 26.8, 24.6, 18.1, 14.1.

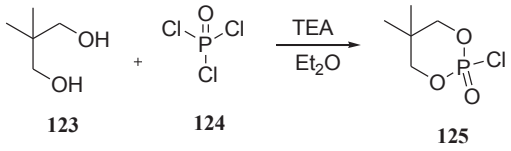
HRMS (ESI): calcd. $(M+H)^+$ 310.1940, exper. 310.1926.

PATHWAY 5:



123 **125** **126**

Synthesis of compound 125 (2-chloro-5,5-dimethyl-1,3,2-dioxaphosphinane 2-oxide)

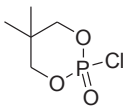


123 **124** **125**

Phosphorus (V) oxychloride (**124**, 1650 mg, 1.000 mL, 11 mmol) was dissolved in cold diethylether (5 mL), and added dropwise to a stirring solution of 2,2-dimethyl-1,3-propandiol (**123**, 1150 mg, 11 mmol) in diethylether (5 mL) and triethylamine (2226 mg, 3.050 mL, 22 mmol) at 0°C under nitrogen atmosphere. The solution was stirred for 1 h, then filtered; the solid was discarded, and the filtrate was evaporated under reduced pressure to obtain the product as a white solid.

Yield: 86%.

Compound **125** - 2-chloro-5,5-dimethyl-1,3,2-dioxaphosphinane 2-oxide



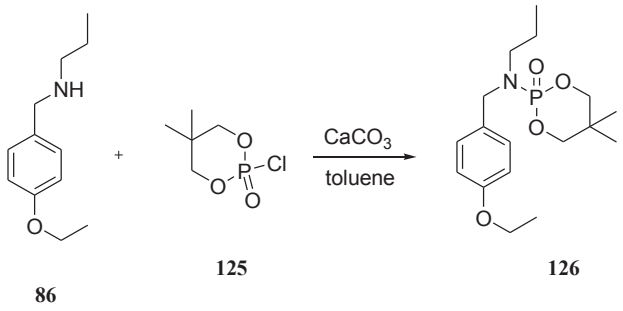
$C_5H_{10}ClO_3P$

MW: 184.5579

^{1}H-NMR (400MHz, $CDCl_3$) δ: 4.22 (2H, d, J=11.1, OCH_2,), 3.98 (2H, dd, J=11.1, OCH_2), 1.28 (3H, s, CH_3), 0.89 (3H, s, CH_3).

HRMS (ESI): calcd. $(M+H)^+$ 185.0056, exper. 185.0131 ($C_5H_{11}{}^{35}ClO_3P$); calcd. $(M+H)^+$ 187.0027, exper. 187.0098 ($C_5H_{11}{}^{37}ClO_3P$).

Synthesis of compound 126 (2-((4-ethoxybenzyl)(propyl)amino)-5,5-dimethyl-1,3,2-dioxaphosphinane 2-oxide)



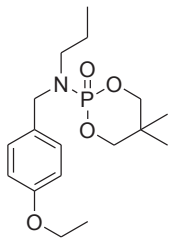
86 **125** **126**

In a round bottom flask all the reactants were placed: compound **125** (300 mg, 1.62 mmol), compound **86** (145 mg, 0.81 mmol), calcium carbonate (162 mg, 1.62 mmol) in toluene (20 mL). The mixture was allowed to stir at 110°C o/n; the completion of the reaction was verified through TLC (hexane/ethyl acetate 4:1).

The mixture was allowed to cool to rt, then it was filtered and the precipitate was washed with toluene. The filtrate was evaporated to remove the solvent, and the residue was dissolved with DCM and washed with sodium hydroxyde 10% (20 mL, 3 times) in a separation funnel. The organic phase was dried with magnesium sulfate and evaporated to give the product as a colorless oil.

Yield: 78%.

Compound **126** – 2-((4-ethoxybenzyl)(propyl)amino)-5,5-dimethyl-1,3,2-dioxaphosphinane 2-oxide



$C_{17}H_{28}NO_4P$

MW: 341.3823

^{1}H-NMR (400MHz, $CDCl_3$) δ: 7.24 (2H, AA'), 6.86 (2H, XX'), 4.49 (2H, d, CH_2), 4.00 (4H, m, CH_2x2), 3.74 (2H, s, CH_2Ar), 2.60 (2H, t, NCH_2), 1.54 (2H, m, CH_2), 1.41 (3H, t, CH_3), 1.33 (3H, s, CH_3), 0.92 (3H, t, CH_3), 0.91 (3H, s, CH_3).

^{13}C-NMR (100MHz, $CDCl_3$) δ: 158.3, 158.0, 129.4, 129.3, 114.4, 79.1, 63.4, 32.2, 21.9, 19.9, 14.9, 11.8.

HRMS (ESI): calcd. $(M+H)^+$ 342.1756, exper. 342.1854.

3.5 VECTOR CONSTRUCTION AND AMPLIFICATION

The design of constructs was performed inserting the template sequences for SET and PP2AC in the Opiner tool from OPPF-UK website[98].

The following constructs were designed:

Construct A5 - pOPIN-F-PP2A$_C$ (His_6-3C-PP2A$_{C,1-245}$)

Construct B5 - pOPIN-J-PP2A$_C$ (His_6-GST-3C-PP2A$_{C,1-245}$)

Construct C5 - pOPIN-M-PP2A$_C$ (His_6-MBP-3C-PP2A $_{C,1-245}$)

Construct D5 - pOPIN-S3C-PP2A$_C$ (His_6-SUMO-3C -PP2A $_{C,1-245}$)

Construct E5 - pOPIN-E-PP2A$_C$ (PP2A$_{C,1-245}$-His_6)

Construct F5 - pOPIN-eGFP-PP2A $_C$ (PP2A $_{C,1-245}$-3C-eGFP-His_6)

Construct A6 - pOPIN-F-SET (His_6-3C-His_6-Thrombin-SET$_{1-225}$)

Construct B6 - pOPIN-J-SET (His_6-GST-3C-His_6-Thrombin-SET$_{1-225}$)

Construct C6 - pOPIN-M-SET (His_6-MBP-3C-His_6-Thrombin-SET$_{1-225}$)

Construct D6 - pOPIN-S3C-SET (His_6-SUMO-3C-His_6-Thrombin-SET$_{1-225}$)

Construct E6 - pOPIN-E-SET (SET$_{1-225}$-His_6)

Construct F6 - pOPIN-eGFP-SET (SET$_{1-225}$-eGFP-3C-His_6).

On the basis of the design above, the following oligonucleotides were purchased from Sigma-Aldrich and dissolved to 10 μM:

PP2A_1Fw	aagttctgtttcagggcccgATGGATGAAAAAGTCTTTACCAAAGAA
PP2A_309Rv	atggtctagaaagctttaTTTTTCGAACTGCGGATGCGACC
PP2A_1FwE	aggagatataccatgATGGATGAAAAAGTCTTTACCAAAGAA
PP2A_309RvE	gtgatggtgatgtttTTTTTCGAACTGCGGATGCGACC

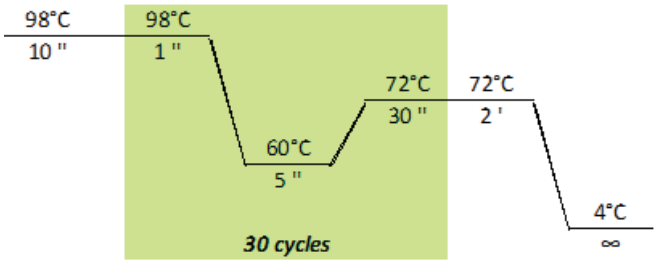
PP2A_309RvEGFP	cagaacttccagtttTTTTTCGAACTGCGGATGCGACC
SET_1Fw	aagttctgtttcagggcccgATGGGCAGCAGCCATCATCATCA
SET_245Rv	atggtctagaaagctttaCATATCGGGAACCAAGTAGTACTG
SET_24Fw	aggagatataccatgCAGGCGGCCAAAGTCAGTAAAAAG
SET_245RvE	gtgatggtgatgtttCATATCGGGAACCAAGTAGTACTG
SET_245RvEGFP	cagaacttccagtttCATATCGGGAACCAAGTAGTACTG

3.5.1 PCR cloning and purification

A PCR was run to amplify the sequences of interest. This step was meant to provide DNA fragments, bearing ends that could be compatible with pOPIN series vectors, which could codify for SET and $PP2A_C$. Each reaction was set up as follows:

- Phusion flash master mix (Thermoscientific F-548L) 2x, 25 µL
- Sterile water, 17 µL
- Appropriate forward primer 10 µM, 3 µL
- Appropriate reverse primer 10 µM, 3 µL
- Template plasmid 10 µM, 2 µL.

With the following thermal cycles:



* Phusion flash polymerase has an extension time of 15 sec/kbp

The products of the PCR reactions were checked by a verification agarose gel (1.6% agarose TAE, 100V, 30 min).

To each reaction, 5 µL of DpnI 1x in NE Buffer were added and the plate was incubated at 37°C for 30 min to digest the template cDNA.

After that, the PCR products were purified by adding 90 µL of AMPure XP magnetic beads; the plate was then incubated for 5 min at rt, then placed onto a SPRIPlate 96R magnet for 5 min to separate the beads from the solution; the cleared solutions were aspired and discarded. After that, 200 µL of EtOH 70% were added to wash the beads, incubated for 30 sec and discarded; this passage was repeated twice. The plate, left to air-dry at rt for 10 min, was subsequently removed from the magnet and 30 µL of buffer (EB 10 mM Tris pH 8.0) were added to elute the PCR products. The beads were removed by placing the plate again on the magnet and the sample contained in each well was collected in a clean PCR plate.

Also in this case, in order to control the efficiency of the purification, a verification gel was run (1.6% agarose-TAE, 100V, 30 min).

3.5.2 In-Fusion^{TM}, HTP transformation into cloning-grade E.coli and HTP mini-preps

The obtained inserts and the pOPIN vectors (linearized with KpnI and HindIII and concentrated to 100 ng/µL) were mixed in molar ratio 2:1 and supplemented with water to a final volume of 10 µL. These solutions were transferred to a dry-down In-Fusion^{TM} plate and incubated for 30 min at 42°C. The products were then added with 40 µL of TE buffer and 5 µL of this diluted reaction were used immediately to transform 50 µL OneShot OmniMaxII cells (cloning efficiency >10^9 cfu/µg plasmid).

The cells were incubated in ice for 30 min, heat shocked for 30 sec at 42°C and placed back in ice for 2 min. PB (300 µL) was then added and the cells were incubated for 1 h at 37°C.

The transformed cells were then plated on LB agar plates added with carbenicillin 50 mg/mL (1:1000), 20% X-Gal (1:1000), IPTG 1 M (1:1000); more specifically, each experiment was repeated twice, plating in one case 5 µL of cells added with 20 µL of PB, and in the other 25 µL of cells. The plates were incubated o/n at 37°C.

The day after, two colonies per well were picked and transferred into deep-well blocks containing 1.2 mL of PB supplemented with carbenicillin 50 mg/mL (1:1000) and chloramphenicol 35 mg/mL (1:1000). The plates were sealed and grown o/n at 37°C with vigorous agitation (600 rpm).

The subsequent day, an aliquot of each o/n culture was added with the same amount of filter sterilized LB/30%v/v glycerol and kept at -80°C as a stock.

The remaining samples were centrifuged at 6000 x *g* for 10 min and the medium was discarded. On the collected cells, an HTP mini plasmid preparation was performed by a Bio-Robot 8000 (Qiagen) on Wizard SV96 purification plates, following the instructions of the manufacturer. The presence of the insert in the purified vectors was verified through a PCR screening, using 0.15 µL of pOPIN forward primer 100 µM for each construct, added with 1.5 µL of the proper reverse primer 10 µM. Since the pOPIN forward primer is present in most of the pOPIN vectors, it can be used for all the constructs; its addition, together with the appropriate reverse primer, could reveal if the nucleic acids had been integrated and purified correctly. An agarose gel was run to verify the nature of the PCR products (1.6% agarose-TAE, 100V, 30 min).

3.6 PROTEIN EXPRESSION

3.6.1 *E.coli* expression strain transformation, protein expression and HTP purification

The purified plasmids (3 µL each) were used to transform 50 µL aliquots of expression *E.coli* (Lemo21 and Rosetta2); after the addition of the plasmids to the cells, they were incubated in an ice bath for 30 min. The cells were then heat shocked for 30 sec at 42°C, then returned back to ice for 2 min. 300 µL of PB per each aliquot were added and the cells were incubated for 1 h at 37°C. After that, 30 µL of cells were plated on LB agar added with carbenicillin 50 mg/mL (1:1000) and chloramphenicol 35 mg/mL (1:1000), and incubated o/n at 37°C.

The day after, one colony for each o/n culture was picked and transferred to a deep-well block containing 0.7 mL of PB added with carbenicillin 50 mg/mL (1:1000) and chloramphenicol 35 mg/mL (1:1000), and incubated o/n at 37°C, 600 rpm.

The following day, 100 µL of each culture were added with the same amount of filter-sterilised LB/30%v/v glycerol and collected in a plate stored at -80°C as a glycerol stock. Another aliquot of each colony (150 µL for Lemo21, 250 µL for Rosetta2) was scaled up both in PB and in AI; in both the cases, the samples were added with carbenicillin 50 mg/mL (1:1000) and chloramphenicol 35 mg/mL (1:1000), and incubated at 37°C, 600 rpm until they reached an OD of 0.5 (approximately 5 h). Cells growing in AI medium were simply switched to 25°C, while the ones which grew in PB were turned at 20°C and added with IPTG 0.5 M (1:500); these conditions were maintained for the whole night after.

An aliquot of 1 mL of each well was transferred in a 96 well deep-well block; this step was made twice. The cells were centrifuged at 6000 x *g* for 10 min, the media discarded and the plates stored at -80°C. The cells were then resuspended in 210 µL of NPI-10 supplemented with 1 mg/mL lysozyme and 400 units/mL of DNAse I, and then the high throughput purification was performed thanks to the QIAGEN Bio Robot 8000 equipped with Ni^{2+}-NTA beads, following the instructions of the constructor regarding soluble proteins. The buffers used in this case were NPI-10 (lysis buffer), NPI-20 (wash buffer) and NPI-250 (elution buffer). For verifying the expression and the purification of each protein, 10% SDS-PAGE gels were performed in MES 1x buffer.

3.6.2 Culture scale up

Basing on the results of the SDS-PAGE gels, some of the cultures were scaled up. In particular, 1 µL of the expression plasmid was added to an aliquot of the chosen cells; here pOPIN-M and pOPIN-S3C t vectors containing SET gene were used (*i.e.* C6 and D6). The aliquots were kept in an ice bath for 30 min, heat-shocked at 42°C for 30 sec and placed back in ice for 2 min. 450 µL of PB were added in each tube and the cells were incubated for 1 h at 37°C. After that, 50 µL of culture from the bottom of the tubes were plated in 5 mL of medium, made up by LB agar added with carbenicillin 50 mg/mL (1:1000) and chloramphenicol 35 mg/mL (1:1000); the plates were incubated at 37°C o/n.

The day after, a single colony was inoculated in 100 mL of PB, added with glucose 0.1% and with carbenicillin 50 mg/mL (1:1000) and chloramphenicol 35 mg/mL (1:1000); this step was made in duplicates. The incubation was carried on at 37°C, 225 rpm, o/n.

The following day, 10 mL of the o/n cultures (coming from just one of the two duplicates) were inoculated in 1 L of medium, either PB or AI (both added with chloramphenicol and carbenicillin, in the quantities which have already been described), as follows:

- PB: the flasks were incubated at 240 rpm, 37°C until the OD reached a value around 0.5 (approximately 5 h); then, IPTG 0.5 M (1:500) was added to promote the induction of the proteins and the temperature switched to 20°C for the o/n expression;
- AI: 5052 (50x), NPS (20x), $MgSO_4$ 2 M (1000x) were added at the very beginning and the flasks were incubated in the same way as the PB ones (240 rpm, 37°C) until the OD reached a value around 0.5 (approximately 5 h), then then the temperature was decreased to 25°C o/n.

Cells were harvested after approximately 18 h by centrifugation at 5000 rpm for 15 min; the supernatant was discarded, the pellet was transferred in 50 mL falcon tubes using a spatula and the tubes were stored at -80°C until the purification. On the average, 1 L of culture yielded 15 g of cells.

3.7 PROTEIN PURIFICATION

3.7.1 Cell lysis and Affinity chromatography

The collected cells were suspended in a proper amount of NPI-10 buffer (typically 15 g of cells in 50 mL of buffer), and added with an aliquot of DNAse I, $MgCl_2$ 1M (5 µL for one aliquot of DNAse), one aliquot of cOmplete EDTA free protease inhibitor cocktail and lysozyme 100 mg/mL (1000x); the sample was then put in an ice bath for 10 min. The cells were then added with PMSF 1M in DMSO (1000x) and disrupted at 4°C, 30 KPSI. The collected sample was ultracentrifuged (vacuum, 200000 x *g*, 4°C, 1 h). The obtained supernatant was filtered under vacuum and loaded on an affinity column (HisTrap FF, GE Healthcare) for the first purification step, collecting the flowthrough in a bottle. For the elution, NPI-10 and NPI-250 were used as solvents. Being more specific, the purification was run as follows:

- 10 CV NPI-250 0%
- 5 CV NPI-250 10%
- 10 CV NPI-250 100%.

The elute was collected in 1 mL fractions; 40-µL aliquots of the fractions themselves were analyzed by a 4-12% SDS-PAGE.

3.7.2 Proteolytic cleavage and Dialysis

The fractions obtained from the affinity chromatography, once united, were added with His-tagged 3C protease (~50 µL *per* estimated mg of protein) and dialyzed against dialysis buffer (stir o/n 4°C) through a dialysis tubing cellulose membrane, in order to both eliminate the imidazole and cleave the tag. The day after, the sample was collected in a bottle and a gel was run in order to verify if the cleavage was performed efficiently. Then, the sample was concentrated to 4-5 mL with MW concentration tubes (cut-off: 10 kDa), centrifuging at 4000 x *g* in 10-min slots. Between one slot and the other, the sample was carefully resuspended.

3.7.3 Gel filtration

The concentrated sample was injected, through a 5 mL loop, on the gel filtration column previously equilibrated with 10 CV of SET gel filtration buffer. 1 mL fractions were collected in a deep well 96-well plate with a flow of 0.3 mL/min, and 40-µL aliquots were analyzed by 4-12% SDS-PAGE to identify the presence of the cleaved protein.

3.8 *IN SILICO* STUDIES

3.8.1 Molecular modeling

Compound **71** was evaluated *in silico* for the comparison of the predicted conformation with the quantitative analysis of NOESY spectrum. For the modeling part, the compound was drawn in Avogadro[81] and minimized with a MMFF94 force field. The PDB outcome was then analized through UCSF Chimera[80] for calculating the distances between proton groups.

3.8.2 Docking

Some of the obtained compounds (**71** and **100**, or MP07-66) were written as PDB files using ChemDraw[96] and Avogadro[81], and then imported on AutoDockTools 1.5.6 furnished with AutoGrid 4 and AutoDock 4.2.6[83] for docking them with SET. The POI (PDB ID: 2E50) was modified through UCSF Chimera[80] in order to eliminate water, ions and other eventual co-crystallized ligands, and to obtain both its dimeric and monomeric form; protons and partial charges were added. The docking grid was placed on the whole macromolecule, since the docking site is not known yet. 10 genetic algorithm runs were performed with a population size of 15000. AutoDock gave 10 output files for each docking, and the one at lower energy was considered for each compound and evaluated through UCSF Chimera.

REFERENCES

1. Bononi, A. *et al.* Protein Kinases and Phosphatases in the Control of Cell Fate. *Enzyme Res.* **2011,** 1–26 (2011).

2. Ciccone, M., Calin, G. A. & Perrotti, D. From the biology of PP2A to the PADs for therapy of hematologic malignancies. *Front. Oncol.* **5,** 1–10 (2015).

3. Hunter, T. Protein Kinases and Phosphatases : The Yin and Yang of Protein Phosphorylation and Signaling. *Cell* **80,** 225–236 (1995).

4. Virshup, D. M. Protein phosphatase 2A : a panoply of enzymes. *Curr. Opin. Cell Biol.* **12,** 180–185 (2000).

5. Perrotti, D. & Neviani, P. Protein phosphatase 2A : a target for anticancer therapy. *Lancet Oncol.* **14,** e229–e238 (2013).

6. Smith, A. M., Roberts, K. G. & Verrills, N. M. in *Myeloid Leuk. - Basic Mech. Leukomiogenes.* 123–148 (2011).

7. Agrawal, M., Garg, R. J., Cortes, J. & Quintás-Cardama, A. Tyrosine kinase inhibitors: The first decade. *Curr. Hematol. Malig. Rep.* **5,** 70–80 (2010).

8. Moorhead, G. B. G., Trinkle-Mulcahy, L. & Ulke-Lemee, A. Emerging roles of nuclear protein phosphatases. *Nat. Rev. Mol. Cell Biol.* **8,** 234–244 (2007).

9. Marks, F., Klingmuller, U. & Muller-Decker, K. *Cellular Signal Processing: An Introduction to the Molecular Mechanisms of Signal Transduction*. (Garland Science, Taylor and Francis group, 2009).

10. Samanta, A. *et al.* Jak2 inhibition dactivates Lyn kinase through the SET-PP2A-SHP1 pathway, causing apoptosis in drug-resistant cells from chronic myelogenous leukenia patients. *Oncogene* **28,** 1669–1681 (2009).

11. Zolnierowicz, S. Type 2A Protein Phosphatase , the Complex Regulator of Numerous Signaling Pathways. *Biochem. Pharmacol.* **60,** 1225–1235 (2000).

12. Eichhorn, P. J. A., Creyghton, M. P. & Bernards, R. Protein phosphatase 2A regulatory subunits and cancer. *BBA - Rev. Cancer* **1795,** 1–15 (2009).

13. Seshacharyulu, P., Pandey, P., Datta, K. & Batra, S. K. Phosphatase : PP2A structural importance , regulation and its aberrant expression in cancer. *Cancer Lett.* **335,** 9–18 (2013).

14. Oaks, J. & Ogretmen, B. Regulation of PP2A by sphingolipid metabolism and signaling. *Front. Oncol.* **4,** 1–7 (2015).

15. Kiely, M. & Kiely, P. PP2A: The Wolf in Sheep's Clothing? *Cancers (Basel).* **7,** 648–669 (2015).

16. Ramaswamy, K., Spitzer, B. & Kentsis, A. Therapeutic Re-Activation of Protein Phosphatase 2A in Acute Myeloid Leukemia. *Front. Oncol.* **5,** 1–5 (2015).

17. Haesen, D., Sents, W., Lemaire, K., Hoorne, Y. & Janssens, V. The Basic Biology of PP2A in Hematologic Cells and Malignancies. *Front. Oncol.* **4,** 1–11 (2014).

18. Janssens, V. & Goris, J. Protein phosphatase 2A: a highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling. *Biochem. J.* **353,** 417–439 (2001).

19. Perrotti, D. & Neviani, P. Protein phosphatase 2A (PP2A), a drugable tumor suppressor in Ph1(+) leukemias. *Cancer Metastasis Rev.* **27,** 159–168 (2008).

20. Groves, M. R. *et al.* The Structure of the Protein Phosphatase 2A PR65 / A Subunit Reveals the Conformation of Its 15 Tandemly Repeated HEAT Motifs. *Cell* **96,** 99–110 (1999).

21. Zhou, J., Pham, H., Ruediger, R. & Walter, G. Characterization of the Aalpha and Abeta subunit isoforms of protein phosphatases 2A : differences in expression, subunit interaction, and evolution. *Biochem. J.* **369,** 387–398 (2003).

22. Janssens, V., Goris, J. & Van Hoof, C. PP2A: The expected tumor suppressor. *Curr. Opin. Genet. Dev.* **15,** 34–41 (2005).

23. Cooper, G. in *Cell A Mol. Approach* (Sunderland (MA): Sinauer Associates, 2000).

24. http://www.bdbiosciences.com. (2015).

25. Elmore, S. Apoptosis: a review of programmed cell death. *Toxicol. Pathol.* **35,** 495–516 (2007).

26. Perrotti, D. & Neviani, P. Protein phosphatase 2A: a target for anticancer therapy. *Lancet. Oncol.* **14,** e229–38 (2013).

27. Bialojan, C. & Takai, A. Inhibitory effect of a marine-sponge toxin, okadaic acid, on protein phosphatases. *Biochem. J.* **256,** 283–290 (1988).

28. Wera, S. *et al.* Deregulation of translational control of the 65-kDa regulatory subunit (PR65 alpha) of protein phosphatase 2A leads to multinucleated cells. *J. Biol. Chem.* **270,** 21374–21381 (1995).

29. Koma, Y., Ito, A., Watabe, K., Kimura, S. H. & Kitamura, Y. A truncated isoform of the PP2A B56 γ regulatory subunit reduces irradiation-induced Mdm2 phosphorylation and could contribute to metastatic melanoma cell radioresistance. *Histol. Histopathol.* **19,** 391–400 (2004).

30. Ito, A. *et al.* A truncated isoform of the PP2A B56 subunit promotes cell motility through paxillin phosphorylation. *EMBO J.* **19,** 562–571 (2000).

31. Wang, S. *et al.* Alterations of the PPP2R1B gene in human lung and colon cancer. *Science (80-.).* **282,** 284–287 (1998).

32. Takagi, Y. *et al.* Alterations of the PPP2R1B gene located at 11q23 in human colorectal cancers. *Gut* **47,** 268–71 (2000).

33. Chen, W. *et al.* Identification of specific PP2A complexes involved in human cell transformation. *Cancer Cell* **5,** 127–136 (2004).

34. Silverstein, A. M., Barrow, C. A., Davis, A. J. & Mumby, M. C. Actions of PP2A on the MAP kinase pathway and apoptosis are mediated by distinct regulatory subunits. *Proc. Natl. Acad. Sci.* **99,** 4221–4226 (2002).

35. Janssens, V. & Rebollo, A. The Role and Therapeutic Potential of Ser / Thr Phosphatase PP2A in Apoptotic Signalling Networks in Human Cancer Cells. *Curr. Mol. Med.* **12,** 1–19 (2012).

36. Reilly, P. T., Yu, Y., Hamiche, A. & Wang, L. Cracking the ANP32 whips: important functions, unequal requirement, and hints at disease implications. *Bioessays* **36,** 1062–1071 (2014).

37. Beresford, P. J. *et al.* Granzyme A Activates an Endoplasmic Reticulum-associated Caspase-

independent Nuclease to Induce Single-stranded DNA Nicks. *J. Biol. Chem.* **276,** 43285–43293 (2001).

38. Saddoughi, S. A. *et al.* Sphingosine analogue drug FTY720 targets I2PP2A / SET and mediates lung tumour suppression via activation of PP2A-RIPK1- dependent necroptosis. *EMBO Mol. Med.* **5,** 105–121 (2013).

39. Bai, X.-L. *et al.* Inhibition of protein phosphatase 2A enhances cytotoxicity and accessibility of chemotherapeutic drugs to hepatocellular carcinomas. *Mol. Cancer Ther.* **13,** 2062–2072 (2014).

40. Lu, J. *et al.* Inhibition of serine/threonine phosphatase PP2A enhances cancer chemotherapy by blocking DNA damage induced defence mechanisms. *Proc. Natl. Acad. Sci.* **106,** 11697–11702 (2009).

41. Suganuma, M. *et al.* Calyculin A, an inhibitor of protein phosphatases, a potent tumor promoter on CD-1 mouse skin. *Cancer Res.* **50,** 3521–3525 (1990).

42. Chatfield, K. & Eastman, A. Inhibitors of protein phosphatases 1 and 2A differentially prevent intrinsic and extrinsic apoptosis pathways. *Biochem. Biophys. Res. Commun.* **323,** 1313–1320 (2004).

43. Switzer, C. H. *et al.* Nitric oxide and protein phosphatase 2A provide novel therapeutic opportunities in ER-negative breast cancer. *Trends Pharmacol. Sci.* **32,** 644–651 (2011).

44. Switzer, C. H. *et al.* Dithiolethione compounds inhibit Akt signaling in human breast and lung cancer cells by increasing PP2A activity. *Oncogene* **28,** 3837–3846 (2009).

45. Neviani, P. & Perrotti, D. SETting OP449 into the PP2A-activating drug family. *Clin. Cancer Res.* **20,** 2026–8 (2014).

46. Vitek, M. P., Ribaudo, G. & Christensen, D. J. ApoE peptide dimers and uses thereof. WO2011/085110A1 (2011).

47. Cada, D., Levien, T. & Baker, D. Formulary Drug Reviews - Fingolimod. *Hosp. Pharm.* **46,** 122–129 (2011).

48. Tavazzi, E., Rovaris, M. & La Mantia, L. Drug therapy for multiple sclerosis. *Can. Med. Assoc. J.*

186, 833–840 (2014).

49. Brinkmann, V., Billich, A., Baumruker, T. & Heining, P. Fingolimod (FTY720): discovery and development of an oral drug to treat multiple sclerosis. *Nat. Publ. Gr.* **9,** 883–897 (2010).

50. Miyake, Y., Kozutsumi, Y., Nakamura, S., Fujita, T. & Kawasaki, T. Serine palmitoyltransferase is the primary target of sphingosine-like immunosuppressant, ISP-1/myriocin. *Biochem. Biophys. Res. Commun.* **211,** 396–403 (1995).

51. Hofrichter, M. in *Ind. Appl.* (Springer Science & Business Media, 2010).

52. Sanchez, T. & Hla, T. Structural and functional characteristics of S1P receptors. *J. Cell. Biochem.* **92,** 913–922 (2004).

53. Chiba, K. & Adachi, K. Sphingosine 1-Phosphate Receptor 1 as a Useful Target for Treatment of Multiple Sclerosis. *Pharmaceuticals* **1,** 514–528 (2012).

54. Hanson, M. A. & Peach, R. *Structural Biology of the S1P 1 Receptor*. (2014). doi:10.1007/978-3-319-05879-5

55. Neviani, P. *et al.* FTY720 , a new alternative for treating blast crisis chronic myelogenous leukemia and Philadelphia chromosome – positive acute lymphocytic leukemia. *J. Clin. Invest.* **117,** 2408–2421 (2007).

56. Leonard, B. *Leukemia: a research report*. (Diane Pub Co, 1993).

57. Guenova, M. & Balatzenko, G. *Leukemia*. (InTech, 2013). doi:10.13140/2.1.3583.4888

58. http://www.cancer.gov/research/progress/snapshots/leukemia.

59. http://www.cancer.org/cancer/leukemia-chroniclymphocyticcll/detailedguide/leukemia-chronic-lymphocytic-key-statistics.

60. Fecteau, J.-F. & Kipps, T. J. Structure and function of the hematopoietic cancer niche: focus on chronic lymphocytic leukemia. *Front Biosci (Schol Ed)* **4,** 61–73 (2012).

61. Chiorazzi, N., Rai, K. R. & Ferrarini, M. Chronic Lymphocytic Leukemia. *N. Engl. J. Med.* **352,** 804–815 (2005).

62. Antosz, H. *et al.* IL-6, IL-10, c-Jun and STAT3 expression in B-CLL. *Blood Cells. Mol. Dis.* **54,** 258–265 (2015).

63. Herishanu, Y., Katz, B.-Z., Lipsky, A. & Wiestner, A. Biology of chronic lymphocytic leukemia in different microenvironments: clinical and therapeutic implications. *Hematol. Oncol. Clin. North Am.* **27,** 173–206 (2013).

64. Caligaris-Cappio, F. & Ghia, P. Novel insights in chronic lymphocytic leukemia: Are we getting closer to understanding the pathogenesis of the disease? *J. Clin. Oncol.* **26,** 4497–4503 (2008).

65. Yang, Y., Huang, Q., Lu, Y., Li, X. & Huang, S. Reactivating PP2A by FTY720 as a novel therapy for AML with c-KIT tyrosine kinase domain mutation. *J. Cell. Biochem.* **113,** 1314–1322 (2012).

66. Cristòbal, I. *et al.* SETBP1 overexpression is a novel leukemogenic mechanism that predicts adverse outcome in elderly patients with acute myeloid leukemia. *Blood* **115,** 615–626 (2010).

67. Neviani, P. *et al.* PP2A-activating drugs selectively eradicate TKI-resistant chronic myeloid leukemic stem cells. **123,** (2013).

68. Kalla, C. *et al.* Analysis of 11q22 – q23 deletion target genes in B-cell chronic lymphocytic leukaemia : Evidence for a pathogenic role of NPAT, CUL5, and PPP2R1B. *Eur. J. Cancer* **3,** 1328–1335 (2007).

69. Christensen, D. J. *et al.* SET oncoprotein overexpression in B-cell chronic lymphocytic leukemia and non-Hodgkin lymphoma : a predictor of aggressive disease and a new treatment target. *Blood* **118,** 4150–4159 (2011).

70. Zonta, F. *et al.* Lyn sustains oncogenic signaling in chronic lymphocytic leukemia by strengthening SET-mediated inhibition of PP2A. *Blood* **125,** 3747–3755 (2015).

71. Sontag, J.-M. & Sontag, E. Protein phosphatase 2A dysfunction in Alzheimer's disease. *Front. Mol. Neurosci.* **7,** 16 (2014).

72. Gamliel, A., Afri, M. & Frimer, A. A. Determining radical penetration of lipid bilayers with new lipophilic spin traps. *Free Radic. Biol. Med.* **44,** 1394–1405 (2008).

73. Kiuchi, M. *et al.* Synthesis and Immunosuppressive Activity of 2-Substituted 2-Aminopropane-1,3-diols and 2-Aminoethanols. *J. Med. Chem.* **43,** 2946–2961 (2000).

74. Fox, M. A. & Whitesell, J. K. *Organic Chemistry - Third edition*. (Jonas and Barlett, 2004).

75. Bruckner, R. *Organic Mechanisms. Reactions, stereochemistry and synthesis*. *Eff. Br. mindfulness Interv. acute pain Exp. An Exam. Individ. Differ.* **1,** (Springer, 2010).

76. Smith, M. & March, J. *March's advanced organic chemistry: reactions, mechanisms and structures*. (Wiley Interscience, 2007).

77. Clayden, J. & Greeves, N. *Organic Chemistry*. (Oxford University Press, 2000).

78. Wnuk, S. F. & Robins, M. J. Antimony(III) chloride exerts potent catalysis of the conversion of sulfoxides to alpha-fluoro thioethers with (diethylamino)sulfur trifluoride. *J. Org. Chem.* **55,** 4757–4760 (1990).

79. Neuman, R. C. in *Org. Chem.* (2013).

80. http://www.cgl.ucsf.edu/chimera/.

81. http://avogadro.cc.

82. Neuhaus, D. & Williamson, M. *The Nuclear Overhauser Effect in structural and conformational analysis*. (Wiley-WHC, 2000).

83. http://autodock.scripps.edu/resources/references.

84. Portelli, M. On the synthesis of beta-phenylalanine. *Gazz. Chim. Ital.* **119,** 215–216 (1989).

85. http://web.expasy.org/compute_pi/.

86. Huttunen, K. *et al.* Novel cyclic phosphate prodrug approach for cytochrome P450-activated drugs containing an alcohol functionality. *Pharm. Res.* **24,** 679–687 (2007).

87. Bird, L. OPPF-UK Standard Protocols : Cloning and Expression Screening. (2012).

88. Nettleship, J. OPPF-UK Standard Protocols : Scale-up and Purification. (2015).

89. Berrow, N. S. *et al.* A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res.* **35,** e45 (2007).

90. Xing, Y. *et al.* Structure of Protein Phosphatase 2A Core Enzyme Bound to Tumor-Inducing Toxins. *Cell* **127,** 341–353 (2006).

91. Ikehara, T., Shinjo, F., Ikehara, S., Imamura, S. & Yasumoto, T. Baculovirus expression , purification , and characterization of human protein phosphatase 2A catalytic subunits alpha and beta. *Protein Expr. Purif.* **45,** 150–156 (2006).

92. Myles, T., Schmidt, K., Evans, D. R. H., Cron, P. & Hemmings, B. A. Active-site mutations impairing the catalytic function of the catalytic subunit of human protein phosphatase 2A permit baculovirus-mediated overexpression in insect cells. *Biochem. J.* **232,** 225–232 (2001).

93. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41,** 207–234 (2005).

94. Muto, S. *et al.* Relationship between the structure of SET/TAF-Ibeta/INHAT and its histone chaperone activity. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 4285–4290 (2007).

95. Muto, S. *et al.* Purification, crystallization and preliminary X-ray diffraction analysis of human oncoprotein SET/TAF-1β. *Acta Crystallogr. Sect. D - Biol. christallography* **60,** 712–714 (2004).

96. CambridgeSoft Corporation. ChemDraw Ultra 8.0. (2003).

97. Meyer, B. & Peters, T. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew. Chemie - Int. Ed.* **42,** 864–890 (2003).

98. https://www.oppf.rc-harwell.ac.uk/Opiner/AddTemplate.

RINGRAZIAMENTI

Voglio ringraziare tutti quelli che hanno, nel bene e nel male, in misura diversa e a diverso titolo, tra gli inevitabili alti e bassi, contribuito a questi miei tre anni: il mio supervisor, i miei colleghi, i miei amici e le loro famiglie, dei miei familiari, i miei vicini di laboratorio, gli studenti che ho visto passare negli anni, e le varie combinazioni possibili delle precedenti categorie. Spero che chi occupa un posto d'onore in questi ringraziamenti sappia di averlo, altrimenti mi premurerò di farglielo sapere. Ma non qui e non ora.

ACKNOWLEDGMENTS

I want to thank all those who, for better or for worse, with different degrees and for different reasons, between the inevitable ups and downs, contributed to these three years of mine: my supervisor, my colleagues, my friends and their families, my family, my laboratory neighbors, students who I saw through the years, and the various possible combinations of the above categories. I hope that those who occupy a place of honor in these acknowledgments know about it, otherwise I will make sure to let them know. But not here and not now.

Forever in debt to your priceless advice

(cit.)