



UNIVERSITAT DE
BARCELONA

Design and synthesis of new potentially inhibitors of PCSK9 and KRAS proteins

Marta Vilaplana Saiz

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.



UNIVERSITY OF BARCELONA

FACULTY OF PHARMACY AND FOOD SCIENCES

**DESIGN AND SYNTHESIS OF NEW POTENTIALLY INHIBITORS
OF PCSK9 AND KRAS PROTEINS**

MARTA VILAPALANA SAIZ

2021

UNIVERSITY OF BARCELONA

FACULTY OF PHARMACY AND FOOD SCIENCES

**DESIGN AND SYNTHESIS OF NEW POTENTIALLY INHIBITORS
OF PCSK9 AND KRAS PROTEINS**

ORGANIC CHEMISTRY DOCTORAL PROGRAM

Memory presented by Marta Vilaplana Saiz to apply for the Doctorate Degree (PhD)
from the University of Barcelona

Director

Tutor

Doctorand

PhD M^a Dolors Pujol Dilmé

PhD Jaime Rubio Martínez

Marta Vilaplana Saiz

MARTA VILAPLANA SAIZ, 2021



UNIVERSITAT DE
BARCELONA

DESIGN AND SYNTHESIS OF NEW POTENTIALLY INHIBITORS OF PCSK9 AND KRAS PROTEINS

Memòria presentada per **Marta Vilaplana Saiz**, graduada en farmàcia per la Universitat de Barcelona, per optar al grau de doctor per la Universitat de Barcelona.

El projecte de tesi doctoral es troba inscrit en el **Departament de Farmacologia, Toxicologia i Química Terapèutica** de la Facultat de Farmàcia i en el **Departament de Química Física i Ciències de Materials** de la Facultat de Química de la Universitat de Barcelona. El treball experimental i la redacció de la memòria que es presenta han estat dirigits per la doctora **M^a Dolors Pujol Dilmé** i el doctor **Jaime Rubio Martínez**.

Barcelona, 22 de Gener de 2021

Directora de la tesi

Tutor de la tesi

Doctorand

Dra. M^a Dolors Pujol Dilmé

Dr. Jaime Rubio Martínez

Marta Vilaplana Saiz

AGRAÏMENTS

Aquesta memòria és el resultat de quatre anys de dedicació a un projecte molt il·lusionant per mi, on no tan sols consisteix en l'adquisició de coneixements purament relacionats en l'àmbit científic, sinó també en l'adquisició de coneixements per la vida mateixa com són l'esforç, la constància, el temperament, gestió de nervis i aprendre que el fracàs també forma part del dia a dia i que els dies bons i dolents i són per tothom. És cert que hi ha hagut obstacles i moments de patiment, però els moments bons, l'alegria i la felicitat dels petits resultats han ajudat a portar a terme aquest treball. Destacar que aquesta experiència ha estat enriquidora i on l'alegria, els ànims i l'afecta de tots aquells que han col·laborat han predominat.

En primer lloc, a la persona a qui voldria donar-li els meus agraïments és a la meva directora de tesi, la doctora Maria Dolors Pujol Dilmé, sense la qual la realització d'aquest treball no hagués estat possible. La Dra. Pujol va ser la primera persona en dipositar tota la confiança en mi i la que em va ajudar incondicionalment des del primer moment. Gràcies a ella he pogut entrar al món de la investigació mitjançant la realització d'un projecte de química mèdica molt enriquidor, on he pogut gaudir d'un gran aprenentatge tan a nivell personal com professional. El més important que m'ha ensenyat a part dels seus coneixements científics que són molt amplis, és la perseverança en el treball, l'esforç, trobar l'equilibri intern, la calma i l'apaivagament dels nervis, aspectes que considero que m'han fet créixer a nivell personal de cara la vida mateixa.

Només li puc agrair tots els consells, l'ajuda, la dedicació, la paciència, la capacitat d'esforç insuperable, l'atenció i els ànims que m'ha fet arribar tan en els moments bons com en els moments més complicats. Ha set un honor tenir-la com a directora de tesi.

En segon lloc, m'agradaria agrair-li al meu director i tutor de tesi, el doctor Jaime Rubio Martínez, el qual sense ell tampoc hagués estat possible la realització de la tesi. La seva ajuda va ser immediata, on gràcies a ell he pogut gaudir de la beca Fi. El doctor Jaime Rubio sense coneixem personalment em va obrir les portes del seu grup de recerca, i sense tenir coneixements del seu àmbit científic em va introduir com una més al grup. Tal i com he dit anteriorment, gràcies a ell he pogut entrar al món de la química física, el qual el trobo apassionant i de gran interès.

Només li puc agrair l'ajuda, la paciència, la constància, l'esforç i totes les hores dedicades cap a mi. També agrair-li la seva forma de veure la vida on m'ha fet veure que el més important a part de la feina també és gaudir de la vida i dels petits moments. La seva frase estrella que sempre em quedarà gravada és: *"quien manda, manda, sabe más y siempre tiene razón"*. Aquesta frase em fa molta gràcia i també reconec que té part de raó.

Només agrair de tot cor la seva ajuda i confiança dipositada en mi tan en els moments difícils com en els bons.

També m'agradaria agrair-li a la Dra. Elisebetta Profumo l'oportunitat d'haver treballat al seu grup de recerca, a l'Istituto Superiore di Sanità de Roma, i posar al meu abast els mitjans necessaris per dur a terme l'avaluació biològica de dos compostos sintetitzats en aquest present treball. Així mateix, m'agradaria agrair a la Fundació Universitària Pere i Pons el seu ajut econòmic que m'ha facilitat la realització d'aquesta estada.

Aquests 4 anys de treball no haguéssin estat possible sense les persones que han passat pel grup de recerca tan de la doctora Maria Dolors Pujol com del doctor Jaime Rubio. Durant aquests anys he tingut el plaer de treballar amb persones que actualment considero amigues i que amb elles el dia a dia ha sigut sempre una aventura. M'agradaria primer de tot agrair a la Dra. Laura Grau, la qual va ser la primera persona amb qui vaig entrar en contacte al laboratori i la primera que em va ajudar en tot moment. La Dra. Grau a part d'aportar-me i ensenyar-me tot el seu coneixement científic també em va aportar moltes coses a nivell personal, actualment és una gran amiga que m'ajuda sempre que ho necessito, m'aporta dosis d'alegria i energia. La Dra. Lorena Navarro per ser un exemple de treball inigualable i per tots els consells que m'ha transmès el llarg dels anys. A la Dra. Lucia Acedo per la serenitat i la manera d'afrontar la vida i els reptes. La Dra. Patricia Mateo i el Dr. Enric Lizano els quals els considero un pilar ja que sense ells està clar que res hauria set el mateix. A la Patrícia només li puc donar les gràcies per ajudar-me i fer-me anar pel camí correcte en tot moment, i també per fer-me veure com són les coses i per ensenyar-me a veure la vida de forma diferent, que no és poc. A l'Enric també donar-li les gràcies per aguantar-me durant tots aquests anys.

Només li puc dir que és una de les persones més meravelloses que conec i que l'estimo molt com a persona i amic, no tinc cap dubte que el seu futur tan a nivell personal com professional serà immens. En Miquel Viñas i en Daniel Jahani per la seva forma d'afrontar la vida i la serenor que transmeten.

Per part del grup de recerca del Dr. Jaime Rubio, agrair en Jordi, en Cristian i en Guillem per l'ajuda que m'han brindat sempre que ho he necessitat i per l'alegria que m'han transmès cada dia.

Finalment, agrair a aquelles persones que en algun moment de la meva tesi doctoral s'han creuat en el meu camí i que han set partícips d'aquest treball com en Marc Luque, Clàudia de Diego, Clara Arainciburu, l'Anna Cesari, en Pere Girbau i l'Àlex Ramos, sense els quals no hauria obtingut alguns dels resultats que s'inclouen en aquesta memòria.

Donar les gràcies a la Dra. Laura Grau i la Dra. Vanessa Prior per haver realitzat el treball previ amb el qual s'ha pogut encaminar aquest projecte.

També especial menció als membre del tribunal, Dra. Marta Cascante, Dr. José Manuel Granadino i a la Dra. Marie-Claude Viaud-Massuard per el temps i la dedicació que els ha suposat formar-ne part.

Finalment, agrair a la meva família i amics per tot el suport que m'han donat durant aquests 4 anys tan intensos tan a nivell personal com professional. Gràcies per haver estat present en els moments més complicats i per seguir estant a la meva vida i ajudar-me a tirar endavant.

GRÀCIES!

"A person who never made a mistake never tried anything new"

Albert Einstein

ABBREVIATIONS, ACRONYMS AND FORMULAS

Å	<i>Angstrom</i>
AA	<i>Amino Acids</i>
ACN	<i>Acetonitrile</i>
Ac ₂ O	<i>Acetic Anhydride</i>
Akt	<i>Protein Kinase B</i>
ALT	<i>Alanine amino Transferase Levels</i>
AMBER	<i>Assisted Model Building with Energy Refinement</i>
AML	<i>Acute Myeloid Leukemia</i>
ApoB	<i>Apolipoprotein B</i>
ARH	<i>Autosomal Recessive Hypercholesterolemia</i>
Aspirine	<i>Acetylsalicylic acid</i>
AST	<i>Aspartate amino Transferase</i>
ATP	<i>Adenosine Triphosphate</i>
BAD	<i>B-Cell Lymphoma 2 associated Agonist of cell Death</i>
BCL2	<i>B-Cell Lymphoma 2</i>
BER	<i>Base Excision Repair</i>
BINAP	<i>(2,2'-Bis(diphenylphosphino)-1,1'-binaphthyl)</i>
BRCA1	<i>Breast Cancer type 1 susceptibility protein</i>
bs	<i>Broad Singlet</i>
C α	<i>Carbon alpha</i>
CA4	<i>Combretastatin A4</i>

CAD	<i>Coronary Artery Disease</i>
CCK-8	<i>Cell Counting Kit 8</i>
CD	<i>Catalytic Domain</i>
CDCl_3	<i>Deuterated Chloroform</i>
$(\text{CD}_3)_2\text{CO}$	<i>Deuterated Acetone</i>
CHD	<i>Cardiovascular Heart Disease</i>
cMD	<i>Conventional Molecular Dynamics</i>
CSs	<i>Consensus Sites</i>
CTX	<i>Cerebrotendinous Xanthomatosis</i>
CVD	<i>Cardiovascular Diseases</i>
CYP3A4	<i>Cytochrome P450 3A4</i>
d	<i>Days</i>
d	<i>Doublet</i>
DCM	<i>Dichloromethane</i>
DES	<i>Deep Eutectic Solvent</i>
DIBAL-H	<i>Diisobutylaluminium Hydride</i>
dL	<i>Deciliter</i>
DLD-1	<i>Human Duke's type C colorectal adenocarcinoma</i>
DMA	<i>N,N-Dimethylaniline</i>
DMF	<i>Dimethylformamide</i>
DMSO-d_6	<i>Deuterated Dimethyl Sulfoxide</i>
DNA	<i>Deoxyribonucleic acid</i>

dt	<i>Double Triplet</i>
4EBP1	<i>4E-Binding Protein 1</i>
EBM-2	<i>Endothelial Cell Growth Basal Medium 2</i>
EFG-(A)	<i>Epidermal Growth Factor-like repeat A</i>
EGM-2	<i>Endothelial Cell Growth Medium 2</i>
EKO	<i>Exploring Key Orientation</i>
ELK-1	<i>ETS Like-1 protein</i>
eq	<i>Equivalent</i>
ER	<i>Endoplasmic Reticulum</i>
ERK	<i>Extracellular signal-Regulated Kinase</i>
ESI	<i>Electrospray Ionization</i>
Et ₃ N	<i>Triethylamine</i>
F	<i>Farnesyl</i>
FDA	<i>Food and Drug Administration</i>
ff	<i>Force Field</i>
FGF-2	<i>Fibroblast Growth Factor</i>
FH	<i>Familial Hypercholesterolemia</i>
FOXO1	<i>Forkhead box protein O1</i>
FPP	<i>Farnesyl Pyrophosphate</i>
fs	<i>Femtosecond</i>
FTI	<i>Farnesyl Transferase Inhibitors</i>
g	<i>Grams</i>

gaff	<i>General Amber Force Field</i>
GaMD	<i>Gaussian Accelerated Molecular Dynamics</i>
GAP	<i>GTPase-activating protein</i>
GCO	<i>Global Cancer Observatory</i>
GDP	<i>Nucleotide Guanosine Diphosphate</i>
GEF	<i>Guanine Exchange Factor</i>
GF	<i>Growth Factor</i>
GG	<i>Geranylgeranyl</i>
GTP	<i>Nucleotide Guanosine Triphosphate</i>
h	<i>Hours</i>
HCC	<i>Hepatocellular Carcinoma</i>
HDL	<i>High-Density Lipoprotein</i>
HEAP	<i>2-Hydroxy-ethylammonium propionate</i>
HER2	<i>Human Epidermal growth factor Receptor 2</i>
HMG-CoA	<i>3-Hydroxy-3-methylglutaryl-CoA</i>
HRAS	<i>Harvey Rat Sarcoma</i>
HUVEC	<i>Human Umbilical Vein Endothelial Cells</i>
HVR	<i>Hypervariable Region</i>
HWE	<i>Horner-Wadsworth-Emmans</i>
IC ₅₀	<i>Half maximal Inhibitory Concentration</i>
ICAM-1	<i>Intracellular Adhesion Molecule 1</i>
ICMT	<i>Isoprenylcysteine carboxyl Methyltransferase</i>

IDL	<i>Intermediate-Density Lipoprotein</i>
IGF	<i>Insulin-like Growth Factor</i>
INE	<i>Instituto Nacional de Estadística</i>
J	<i>Coupling Constant</i>
K	<i>Kelvin</i>
Kb	<i>Kilobase</i>
Kcal	<i>Kilocalorie</i>
kDa	<i>Kilodaltons</i>
kPa	<i>Kilopascal</i>
KRAS	<i>Kirsten Rat Sarcoma</i>
L	<i>Loop</i>
LCPO	<i>Linear Combination of Pairwise Overlap</i>
LDL	<i>Low-Density Lipoprotein</i>
LDL-C	<i>Low-Density Lipoprotein Cholesterol</i>
LDL-R	<i>Low-Density Lipoprotein Receptor</i>
LDLRAP1	<i>Low-Density Lipoprotein Receptor Adaptor Protein 1</i>
m	<i>Multiplet</i>
MAPK	<i>Mitogen-Activated Protein Kinase</i>
MD	<i>Molecular Dynamics</i>
MDS	<i>Myelodysplastic Syndrome</i>
MEDPED	<i>Make Early Diagnoses Prevent Early Deaths Program Diagnostic Criteria</i>
MEK	<i>Mitogen-activated protein kinase kinase</i>

MeOH	<i>Methanol</i>
mg	<i>Milligrams</i>
MHz	<i>Megahertz</i>
mL	<i>Mililiters</i>
MM-GBSA	<i>Molecular Mechanics Generalized Born Surface Area</i>
mm Hg	<i>Millimetre of Mercury</i>
MM-PBSA	<i>Molecular Mechanics Poisson Boltzmann Surface Area</i>
MMR	<i>Deoxyribonucleic acid Mismatch repair</i>
m-RNA	<i>Messenger Ribonucleic Acid</i>
m-TOR	<i>Mammalian Target Of Rapamycin</i>
m-TORC1	<i>Mammalian Target Of Rapamycin Complex 1</i>
MTS	<i>Tetrazolium salt</i>
MTTP	<i>Microsomal Triglyceride Transfer Protein</i>
MW	<i>Microwave</i>
NBS	<i>N-Bromosuccinimide</i>
NER	<i>Nucleotide Excision Repair</i>
NIS	<i>N-Iodosuccinimide</i>
nM	<i>Nanomolar</i>
NMR	<i>Nuclear Magnetic Resonance</i>
NPT	<i>Isothermal-isobaric ensemble</i>
NRAS	<i>Neuroblastoma Rat Sarcoma</i>
ns	<i>Nanoseconds</i>

NSCLC	<i>Non-Small-Cell Lung Carcinoma</i>
NVT	<i>Canonical ensemble</i>
OBC	<i>Onufiev-Bashford-Case</i>
ρ	<i>Density</i>
p53	<i>Tumor suppressor p53</i>
p85	<i>Phosphatidylinositol 3-kinase regulatory kinase</i>
p110	<i>Phosphatidylinositol-4,5-bisphosphate 3-kinase</i>
PB	<i>Poisson-Boltzman</i>
PBC	<i>Periodic Boundary Condition</i>
PCA	<i>Principal Component Analysis</i>
PCs	<i>Proprotein Convertases</i>
PCSK9	<i>Proprotein Convertase Subtilisin/Kexin type 9</i>
PDAC	<i>Pancreatic Ductal Adenocarcinoma Cancer</i>
PDB	<i>Protein Data Bank</i>
PDGFR- β	<i>Platelet-Derived Growth Factor Receptor Beta</i>
PDK1	<i>Phosphoinositide-Dependent Kinase-1</i>
pH	<i>Power of Hydrogen</i>
PI3K	<i>Phosphatidylinositol 3-Kinase</i>
PIP ₂	<i>Phosphatidylinositol-4,5-trisphosphate</i>
PIP ₃	<i>Phosphatidylinositol-3,4,5-trisphosphate</i>
PME	<i>Particle-Mesh Ewald</i>
PMS	<i>Phenazine Methyl Sulfate</i>

POCl ₃	<i>Phosphoryl Chloride</i>
PPI	<i>Protein-Protein Interaction</i>
ppm	<i>Parts-per-million</i>
ps	<i>Picosecond</i>
PTEN	<i>Tumor-supressor Phosphate with Tensin Homology</i>
RAF	<i>Serine/threonine-specific protein kinase</i>
RAL	<i>Ras-related protein</i>
RCC	<i>Renal Cell Carcinoma</i>
RCE1	<i>Ras Converting Enzyme 1</i>
REDECAN	<i>Red Española de Registros de Cáncer</i>
REMS	<i>Restricted Risk Evaluation and Mitigation Strategy</i>
RESP	<i>Restrained Electrostatic Potential</i>
Rg	<i>Radius of Gyration</i>
RMSD	<i>Root-Main Square Deviation</i>
RMSF	<i>Root-Main Square Fluctuation</i>
RPE	<i>Retinal Pigment Epithelial</i>
r.t.	<i>Room Temperature</i>
RTK	<i>Receptor Tyrosine Kinase</i>
s	<i>Singlet</i>
S	<i>Strength</i>
SASA	<i>Solvent-Accessible Surface Area</i>
SERP-2	<i>Stress Associated Endoplasmic Reticulum Protein Family Member 2</i>

SH2	<i>Src Homology 2</i>
SH3	<i>Src Homology 3</i>
S6K1	<i>Ribosomal protein S6 Kinase beta-1</i>
S _N Ar	<i>Nucleophilic Aromatic Substitution</i>
SOS	<i>Son of Sevenless</i>
Sp1	<i>Transcription factor 1</i>
SREBP	<i>Sterol Regulatory Element-Binding Protein</i>
SwI	<i>Switch I</i>
SwII	<i>Switch II</i>
t	<i>Triplet</i>
t-BuOK	<i>Potassium tert-butoxide</i>
THF	<i>Tetrahydrofuran</i>
TNF- α	<i>Tumor Necrosis Factor Alpha</i>
TPP	<i>Triphenylphosphine</i>
μ M	<i>Micromolar</i>
UV	<i>Ultraviolet</i>
VdW	<i>Van der Waals</i>
VEGF	<i>Vascular Endotelial Growth Factor</i>
VEGFR-2	<i>Vascular Endotelial Growth Factor Receptor 2</i>
VLDL	<i>Very-Low-Density Lipoprotein</i>
WB	<i>Western Blot</i>
WST-8	<i>Water-Soluble Tetrazolium salt 8</i>

WT	<i>Wild Type</i>
°C	<i>Degrees Celsius</i>
ΔG	<i>Gibbs Free Energy</i>
ΔH	<i>Enthalpy</i>
ΔS	<i>Entropy</i>
ΔU	<i>Internal Energy</i>
%	<i>Percentage</i>

ABSTRACT

This thesis shows the research carried out on two specific diseases: familial hypercholesterolemia, which is included within the cardiovascular diseases group, and cancer. These diseases are the major cause of deaths in Spain and worldwide.

The study of familial hypercholesterolemia is of great interest, as it is mainly caused by a protein called PCSK9 whose atypical functioning leads to an increase of LDL-C in blood. Thus, the study of new diseases results in the study of new therapeutic targets. Therefore, a thorough investigation on the PCSK9 protein has been conducted by means of molecular modeling analysis. Also, the synthesis of new potentially PCSK9 inhibitors has been performed. The modeling studies have been based on the research of new allosteric binding sites and in *de novo* synthesis of new inhibitors. On the other side, the experimental synthesis of new compounds has been conducted by the preparation of pyrrolo[2,3-*d*]pyrimidines and finally, by the preparation of a series of compounds analogues of *Combretastatin A4* through modifications in the two aromatic rings and the fixation of its configuration with the introduction of a new cycle.

In the case of cancer, the study is based mainly on the research of a key protein in cell proliferation, growth and cell signalling. KRAS protein is either active or mutated in most common cancers, being the main cause of pancreas and colorectal cancer. Research on this target is of high chemical and biological relevance since no drugs have been found to act directly on said target. Molecular modeling studies are performed in order to find the binding site of a reference product whose inhibiting capacity over the protein is known. Furthermore, the synthesis of a series of symmetric [REDACTED] with different structural modifications has also been carried out.

Additionally, biological WST-8/CCK8 assays have been performed on some of the synthesized compounds, specifically on two *Combretastatin A4* analogues, in an endothelial cell line that allows studying both endothelial dysfunction and cardiovascular diseases. Research on the MAPK signalling pathway has been performed with [REDACTED] structure derivatives.

RESUM

La tesi que es presenta a continuació mostra la investigació realitzada sobre dues malalties concretes, la hipercolesterolèmia familiar, la qual es troba englobada dins les malalties cardiovasculars, i el càncer. Aquestes malalties són les principals causes de mort a Espanya així com també a la resta del món.

És de gran interès l'estudi de la hipercolesterolèmia familiar ja que és causada principalment per una proteïna anomenada PCSK9, el funcionament anòmal d'aquesta condueix a un augment de LDL-C en sang. L'estudi de noves malalties doncs, porta lloc a l'estudi de noves dianes terapèutiques. Tan mateix, s'ha profunditzat l'estudi de la proteïna PCSK9 mitjançant estudis de modelització molecular i també s'ha portat a terme la síntesi de nous compostos potencialment inhibidors d'aquesta proteïna. Els estudis de modelització s'han basat en la recerca de nous llocs d'unió al·lostèrics i la síntesi de *novo* de nous inhibidors. Per contra la síntesi experimental de nous compostos s'ha realitzat mitjançant la preparació de pirrolo[2,3-*d*]pirimidines i finalment en la preparació d'una sèrie de compostos anàlegs a la *Combretastatina A4* mitjançant modificacions als dos anells aromàtics i la fixació de la seva configuració mitjançant la introducció d'un cicle.

En aquest cas l'estudi del càncer s'ha basat principalment en la investigació d'una proteïna clau en la proliferació cel·lular, el creixement i la senyalització cel·lular. La proteïna KRAS es troba mutada o activa en la major part dels càncers que es coneixent, essent així la principal causa del càncer de pàncrees i colorectal. L'estudi d'aquesta diana és de gran interès químic i biològic ja que no s'ha trobat cap fàrmac que actuï directament sobre aquesta. Estudis de modelització molecular es porten a terme per tal de trobar el lloc d'unió d'un producte de referència del qual es coneix la seva capacitat inhibidora sobre la proteïna. Per altre banda, també s'ha sintetitzat una sèrie de ████████ simètrics amb diferents modificacions estructurals.

Adicionalment s'han realitzat assajos biològics de tipus WST-8/CCK8 d'alguns dels compostos sintetitzats, en concret de dos compostos anàlegs de la *Combretastatina A4* en una línia de cèl·lules endotelials, la qual permet realitzar estudis sobre la disfunció endotelial i l'estudi de malalties cardiovasculars. Estudis sobre la via de senyalització de les MAPK han set realitzats amb derivats d'estructura ████████.

INDEX

1. INTRODUCTION	1
1.1. Familial hypercholesterolemia (FH)	1
1.1.1. <i>Definition of FH</i>	1
1.1.2. <i>Phenotypes and genetic causes</i>	1
1.1.3. <i>Screening strategies</i>	5
1.1.4. <i>Diagnosis</i>	6
1.1.5. <i>Prognosis</i>	7
1.1.6. <i>Treatment</i>	7
1.1.6.1. <i>Heterozygous FH treatment</i>	7
1.1.6.2. <i>Homozygous FH treatment</i>	12
1.1.6.3. <i>Heterozygous and Homozygous FH treatment</i>	14
1.1.7. <i>Molecular target</i>	15
1.1.7.1. <i>PCSK9 protein</i>	15
1.2. Cancer	17
1.2.1. <i>Cancer introduction</i>	17
1.2.2. <i>Types of cancer</i>	18
1.2.3. <i>Epidemiology</i>	20
1.2.3.1. <i>Epidemiology in Spain</i>	21
1.2.4. <i>Molecular target</i>	22
1.2.4.1. <i>RAS</i>	22
1.2.4.2. <i>KRAS</i>	23
1.3. Background	39
1.3.1. <i>Preparation of new inhibitory agents of PCSK9 protein involved in cholesterol homeostasis</i>	39
1.3.1.1. <i>Design of new enzyme inhibitors through molecular modeling studies</i>	39
1.3.1.2. <i>Preparation of pyrrolo[2,3-d]pyrimidines</i>	44
1.3.1.3. <i>Preparation of Combretastatin A4 (CA4) derivatives</i>	46

1.3.2. Preparation of new KRAS protein inhibitory agents	48
1.3.2.1. Optimization of potential KRAS inhibitors or activators from lead P14	48
1.3.2.2. Preparation of KRAS inhibitors	49
2. OBJECTIVES	50
2.1. Preparation of new inhibitory agents of PCSK9 protein involved in cholesterol homeostasis	50
2.1.1. Design of new enzyme inhibitors through molecular modeling studies	50
2.1.1.1. Search for new allosteric binding sites (Pockets).....	50
2.1.1.2. Screening of libraries (Docking).....	50
2.1.1.3. Rational design of inhibitors based on the structure of the ligands	51
2.1.2. Preparation of pyrrolo[2,3-d]pyrimidines.....	51
2.1.3. Preparation of Combretastatin A4 (CA4) derivatives	53
2.2. Preparation of new KRAS protein inhibitory agents	55
2.2.1. Optimization of potential KRAS inhibitors or activators from lead P14	55
2.2.1.1. Search of the binding site of the lead P14 by performing gaussian accelerated molecular dynamics simulations (GaMD)	55
2.2.2. Preparation of KRAS inhibitors	56
3. THEORETICAL DISCUSSION	59
3.1. Preparation of new inhibitory agents of PCSK9 protein involved in cholesterol homeostasis	59
3.1.1. Design of new enzyme inhibitors through molecular modeling studies	59
3.1.1.1. Search for new allosteric binding sites (Pockets).....	59
3.1.1.2. Screening of libraries (Docking).....	70
3.1.1.3. Rational design of inhibitors based on the structure of the ligands	81
3.1.2. Preparation of pyrrolo[2,3-d]pyrimidines.....	94
3.1.2.1. Preparation of 4-amine-5-bromo-6-chloropyrimidine (55)	94
3.1.2.2. Preparation of 4-amine-6-chloro-5-(phenylethynyl)pyrimidine (53) and 4-amine-6- chloro-5-(4-chlorophenyl)ethynyl)pyrimidine (54)	96
3.1.2.3. Preparation of 4-chloro-6-phenyl-7H-pyrrolo[2,3-d]pyrimidine (1) and 4-chloro-6-(4- chlorophenyl)-7H-pyrrolo[2,3-d]pyrimidine (2)	98

3.1.2.4. Preparation of 4-chloro-6-(4-chlorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidine (3) and 4-chloro-6-(4-fluorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidine (4)	100
3.1.2.5. Preparation of 4-amine-6-(4-chlorophenyl)-7-methyl-N-(4-nitrophenyl)-7H-pyrrolo[2,3-d]pyrimidine (5) , 4-amine-6-(4-fluorophenyl)-7-methyl-N-(4-nitrophenyl)-7H-pyrrolo[2,3-d]pyrimidine (6) , 4-(4-aminophenylamine)-6-(4-chlorophenyl)-pyrrolo[2,3-d]pyrimidine (7) and 4-(4-aminophenylamine)-6-(4-fluorophenyl)-pyrrolo[2,3-d]pyrimidine (8)	102
3.1.2.6. Preparation of 1-(4-((6-(4-chlorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidin-4-yl)amino)phenyl)-3-(2-fluoro-5-(trifluoromethyl)phenyl)urea (9) and 1-(2-fluoro-5-(trifluoromethyl)phenyl)-3-(4-((6-(4-fluorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidin-4-yl)amino)phenyl)urea (10)	105
3.1.2.7. Preparation of 7-benzyl-4-chloro-6-phenyl-7H-pyrrolo[2,3-d]pyrimidine (11)	109
3.1.2.8. Preparation of 7-benzyl-6-phenyl-4-(p-tolyl)-7H-pyrrolo[2,3-d]pyrimidine (12) and 7-benzyl-6-phenyl-4-(m-tolyl)-7H-pyrrolo[2,3-d]pyrimidine (13)	111
3.1.3. Preparation of Combretastatin A4 (CA4) derivatives	113
3.1.3.1. Preparation of 15-31	114
3.1.3.2. Preparation of pyrazolidine-3-ones (32-38)	130
3.2. Preparation of new KRAS protein inhibitory agents	138
3.2.1. Optimization of potential KRAS inhibitors or activators from lead P14	138
3.2.1.1. Search of the binding site of the lead P14 by performing gaussian accelerated molecular dynamics simulations (GaMD)	138
3.2.2. Preparation of KRAS inhibitors	149
3.2.2.1. Preparation of [REDACTED] (41) , [REDACTED] (42) and [REDACTED] (43)	153
3.2.2.2. Preparation of [REDACTED] (44) , [REDACTED] (45) and [REDACTED] (46)	154
3.2.2.3. Preparation of [REDACTED] (47) , [REDACTED]	

[redacted] (48), [redacted] [redacted] (49) and [redacted] [redacted] (50)	156
3.2.2.4. Preparation of [redacted] (51)	159
3.2.2.5. Preparation of the [redacted] [redacted] (52).....	160
3.3. Biological results.....	161
3.3.1. <i>Biological results related to cardiovascular diseases</i>	161
3.3.1.1. <i>Biological activity evaluated at the Istituto Superiore di Sanità (Rome)</i>	161
3.3.2. <i>Biological results related to KRAS</i>	165
3.3.2.1. <i>Biological activity evaluated at the Hospital Clinic</i>	165
4. EXPERIMENTAL SECTION	170
4.1. Physical Chemistry laboratory materials and methods	170
4.1.1. <i>Preparation of new inhibitory agents of PCSK9 protein involved in cholesterol homeostasis</i>	170
4.1.1.1. <i>Design of new enzyme inhibitors through molecular modeling studies</i>	170
4.1.2. <i>Preparation of new KRAS protein inhibitory agents</i>	188
4.1.2.1. <i>Optimization of potential KRAS inhibitors or activators from lead P14</i>	188
4.2. Organic Chemistry laboratory materials and methods.....	191
Preparation of 4-chloro-6-phenyl-7H-pyrrolo[2,3-d]pyrimidine (1)	192
Preparation of 4-chloro-6-(4-chlorophenyl)-7H-pyrrolo[2,3-d]pyrimidine (2).....	193
Preparation of 4-chloro-6-(4-chlorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidine (3).....	195
Preparation of 4-chloro-6-(4-fluorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidine (4)	196
Preparation of 4-nitrophenylamino-6-(4-chlorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidine (5)	197
Preparation of 4-nitrophenyl-6-(4-fluorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidine (6).....	199
Preparation of 4-(4-aminophenylamine)-6-(4-chlorophenyl)-pyrrolo[2,3-d]pyrimidine (7).....	200
Preparation of 4-(4-aminophenylamine)-6-(4-fluorophenyl)-pyrrolo[2,3-d]pyrimidine (8)	202

Preparation of 1-(4-((6-(4-chlorophenyl)-7-methyl-7 <i>H</i> -pyrrolo[2,3- <i>d</i>]pyrimidin-4-yl)amino)phenyl)-3-(2-fluoro-5-(trifluoromethyl)phenyl)urea (9).....	203
Preparation of 1-(2-fluoro-5-(trifluoromethyl)phenyl)-3-(4-((6-(4-fluorophenyl)-7-methyl-7 <i>H</i> -pyrrolo[2,3- <i>d</i>]pyrimidin-4-yl)amino)phenyl)urea (10).....	205
Preparation of 7-benzyl-4-chloro-6-phenyl-7 <i>H</i> -pyrrolo[2,3- <i>d</i>]pyrimidine (11).....	206
Preparation of 7-benzyl-6-phenyl-4-(<i>p</i> -tolyl)-7 <i>H</i> -pyrrolo[2,3- <i>d</i>]pyrimidine (12).....	207
Preparation of 7-benzyl-6-phenyl-4-(<i>m</i> -tolyl)-7 <i>H</i> -pyrrolo[2,3- <i>d</i>]pyrimidine (13)	209
Preparation of (<i>E</i>)-3-(3,4-dihydroxyphenyl)-2-(3,4,5-trimethoxyphenyl)acrylic acid (15).....	211
Preparation of (<i>E</i>)-methyl 3-(3,4-dihydroxyphenyl)-2-(3,4,5-trimethoxyphenyl)acrylate (16).....	212
Preparation of (<i>E</i>)-2,3- <i>bis</i> (3,4,5-trimethoxyphenyl)acrylic acid (17).....	215
Preparation of (<i>E</i>)-methyl 2,3- <i>bis</i> (3,4,5-trimethoxyphenyl)acrylate (18)	216
Preparation of (<i>E</i>)-3-(4-hydroxy-3-nitrophenyl)-2-(3,4,5-trimethoxyphenyl)acrylic acid (19).....	218
Preparation of (<i>E</i>)-3-methyl (4-hydroxy-3-nitrophenyl)-2-(3,4,5-trimethoxyphenyl)acrylate (20).....	219
Preparation of (<i>E</i>)-3-(benzo[<i>d</i>][1,3]dioxol-5-yl)-2-(3,4,5-trimethoxyphenyl)acrylic acid (21).....	221
Preparation of (<i>E</i>)-methyl 3-(benzo[<i>d</i>][1,3]dioxol-5-yl)-2-(3,4,5-trimethoxyphenyl)acrylate (22).....	222
Preparation of (<i>E</i>)-3-(3-hydroxy-4-methoxyphenyl)-2-phenylacrylic acid (23)	223
Preparation of (<i>E</i>)-methyl 3-(3-hydroxy-4-methoxyphenyl)-2-phenylacrylate (24)	225
Preparation of (<i>E</i>)-3-(3-hydroxy-4-methoxyphenyl)-2-(4-pivaloylphenyl)acrylic acid (25).....	226
Preparation of (<i>E</i>)-3-(3-hydroxy-4-methoxyphenyl)-2-(4-(methylsulfonyl)phenyl)acrylic acid (26).....	228
Preparation of (<i>E</i>)-methyl 3-(3-hydroxy-4-methoxyphenyl)-2-(4-(methylsulfonyl)phenyl)acrylate (27)	229
Preparation of (<i>E</i>)-2-(2-((2,6-dichlorophenyl)amino)phenyl)-3-(3-hydroxy-4-methoxyphenyl)acrylic acid (28).....	231
Preparation of (<i>E</i>)-methyl 2-(2-((2,6-dichlorophenyl)amino)phenyl)-3-(3-hydroxy-4-methoxyphenyl)acrylate (29)	232
Preparation of (<i>E</i>)-methyl 2-(2-((2,6-dichlorophenyl)amino)phenyl)-3-(3-hydroxy-4-methoxyphenyl)acrylate (29)	233

Preparation of (<i>E</i>)-methyl 2-(2-((2,6-dichlorophenyl)amino)phenyl)-3-(3-hydroxy-4-methoxyphenyl)acrylate (29)	236
Preparation of (<i>E</i>)-3-(3-hydroxy-4-methoxyphenyl)-2-(1 <i>H</i> -indol-3-yl)acrylic acid (30)	239
Preparation of (<i>E</i>)-methyl 3-(3-hydroxy-4-methoxyphenyl)-2-(1 <i>H</i> -indol-3-yl)acrylate (31)	240
Preparation of 5-(3,4-dihydroxyphenyl)-4-(3,4,5-trimethoxyphenyl)pyrazolidin-3-one (32)	241
Preparation of 4,5- <i>bis</i> (3,4,5-trimethoxyphenyl)pyrazolidin-3-one (33).....	243
Preparation of 5-(4-hydroxy-3-nitrophenyl)-4-(3,4,5-trimethoxyphenyl)-1 <i>H</i> -pyrazol-3(2 <i>H</i>)-one (34)	244
Preparation of 5-(benzo[<i>d</i>][1,3]dioxol-5-yl)-4-(3,4,5-trimethoxyphenyl)-1 <i>H</i> -pyrazol-3(2 <i>H</i>)-one (35)	246
Preparation of 5-(3-hydroxy-4-methoxyphenyl)-4-phenylpyrazolidin-3-one (36)	247
Preparation of 5-(3-hydroxy-4-methoxyphenyl)-4-(4-(methylsulfonyl)phenyl)-1 <i>H</i> -pyrazol-3(2 <i>H</i>)-one (37).....	249
Preparation of 5-(3-hydroxy-4-methoxyphenyl)-4-(1 <i>H</i> -indol-3-yl)pyrazolidin-3-one (38)	251
Preparation of 4-acetyl-3-(3,4-dimethoxyphenyl)-5-methyl-2-(3,4,5-trimethoxyphenyl)-2,3-dihydrofuran (39)	253
Preparation of 4-acetyl-2-(3,4-dimethoxyphenyl)-5-methyl-3-(3,4,5-trimethoxyphenyl)-2,3-dihydrofuran (40)	254
Preparation of [REDACTED] (41).....	256
.....	256
Preparation of [REDACTED] (42). 257	
Preparation of [REDACTED] (43)	259
Preparation of [REDACTED] (44)	260
Preparation of [REDACTED] (45)	262
Preparation of [REDACTED] (46)	264

Preparation of [REDACTED]	
[REDACTED] (47)	266
Preparation of [REDACTED]	
[REDACTED] (48)	267
Preparation of [REDACTED]	
[REDACTED] (49)	269
Preparation of [REDACTED]	
[REDACTED] (50)	270
Preparation of [REDACTED] (51).....	272
Preparation of [REDACTED] (52)	273
Preparation of 4-amine-6-chloro-5-(phenylethynyl)pyrimidine (53)	274
Preparation of 4-amine-6-chloro-5-(4-chlorophenyl)ethynyl)pyrimidine (54).....	276
Preparation of 4-amine-5-bromo-6-chloropyrimidine (55).....	277
Preparation of 4-amine-6-chloropyrimidine (56).....	278
Preparation of 4,6-dichloropyrimidine (57)	279
Preparation of 4-amine-6-chloro-5-iodopyrimidine (59).....	281
Preparation of 4-amine-6-chloro-5-((4-chlorophenyl)ethynyl)- <i>N</i> -methylpyrimidine (60).....	282
Preparation of 4-amine-6-chloro-5-((4-fluorophenyl)ethynyl)- <i>N</i> -methylpyrimidine (61)	284
Preparation of 4-amine-6-chloro-5-iodo- <i>N</i> -methylpyrimidine (62).....	285
Preparation of 4-amine-6-chloro- <i>N</i> -methylpyrimidine (63)	286
Preparation of 4-amine- <i>N</i> -benzyl-6-chloro-5-(phenylethynyl)pyrimidine (64)	287
Preparation of 4-amine- <i>N</i> -benzyl-5-bromo-6-chloropyrimidine (65).....	289
Preparation of 4-amine- <i>N</i> -benzyl-6-chloropyrimidine (66).....	290
Preparation of 2-(4-pivaloylphenyl)acetic acid (70)	291
Preparation of ethyl 2-(4-pivalophenyl)acetate (79).....	292
Preparation of 1-(2-((2,6-dichlorophenyl)amino)phenyl)propan-2-one (81).....	293
Preparation of 1-acetyl-5-(4-hydroxy-3-nitrophenyl)-4-(3,4,5-trimethoxyphenyl)-1 <i>H</i> -pyrazol-3(2 <i>H</i>)-one (86).....	295
Preparation of 5-(benzo[<i>d</i>][1,3]dioxol-5-yl)-4-(3,4,5-trimethoxyphenyl)isoxazolidin-3-one (87). 296	

Preparation of methyl 4-acetyl-5-methyl-2,3-bis(3,4,5-trimethoxyphenyl)-2,3-dihydrofuran-2-carboxylate (88)	297
Preparation of 1-(4-(3-(benzyloxy)-4-methoxyphenyl)-2-methyl-5-(3,4,5-trimethoxyphenyl)-4,5-dihydrofuran-3-yl)ethanone (90).....	298
Preparation of (E)/(Z)-5-(3,4-dimethoxystyryl)-1,2,3-trimethoxybenzene (91)	299
Preparation of 2-((4-chlorobenzyl)amino)benzaldehyde (96)	301
Preparation of 2-((3-(benzyloxy)phenyl)amino)benzaldehyde (97).....	302
Preparation of 4-((3-(benzyloxy)phenyl)amino)benzaldehyde (98).....	303
Preparation of 4-methoxy-3-(4-nitrophenoxy)benzaldehyde (99)	305
Preparation of 5-methoxy-2-(4-nitrophenoxy)benzaldehyde (100)	306
Preparation of 5-methoxy-2-(4-nitrilephenoxy)benzaldehyde (101).....	307
Preparation of 2-ciano-N-(4-chlorobenzyl)-aniline (104)	309
Preparation of 4-((3-(benzyloxy)phenyl)amino)benzotrile (109).....	310
Preparation of [REDACTED] (114).....	311
4.3. Biochemical laboratory materials and methods.....	314
4.3.1. <i>Biological activity evaluated at the Istituto Superiore di Sanità (Rome)</i>	314
4.3.1.1. <i>Cell cultures</i>	314
4.3.1.2. <i>Cell viability</i>	314
4.3.2. <i>Biological activity evaluated at the Hospital Clinic</i>	315
4.3.2.1. <i>Cell cultures</i>	315
4.3.2.2. <i>Cell viability</i>	315
4.3.2.3. <i>Cell signalling pathway</i>	316
5. CONCLUSIONS.....	317
5.1. Preparation of new inhibitory agents of PCSK9 protein involved in cholesterol homeostasis	317
5.1.1. <i>Design of new enzyme inhibitors through molecular modeling studies</i>	317
5.1.1.1. <i>Search for new allosteric binding sites (Pockets)</i>	317
5.1.1.2. <i>Screening of libraries (Docking)</i>	317

5.1.1.3. Rational design of inhibitors based on the structure of the ligands	317
5.1.2. Preparation of pyrrolo[2,3-d]pyrimidines.....	318
5.1.3. Preparation of Combretastatin A4 (CA4) derivatives	319
5.2. Preparation of new KRAS protein inhibitory agents	319
5.2.1. Optimization of potential KRAS inhibitors or activators from lead P14	319
5.2.1.1. Search of the binding site of the lead P14 by performing gaussian accelerated molecular dynamics simulations (GaMD)	319
5.2.2. Preparation of KRAS inhibitors	320
5.3. Biological conclusions.....	320
5.3.1. Biological activity evaluated at the Istituto Superiore di Sanità (Rome).....	320
5.3.2. Biological activity evaluated at Hospital Clinic.....	320
6. BIBLIOGRAPHY.....	325

1. INTRODUCTION

1.1. Familial hypercholesterolemia (FH)

1.1.1. Definition of FH

Familial hypercholesterolemia is an inherited condition, involving a single gene, which produces an increase of low-density lipoprotein cholesterol (LDL-C) in plasma promoting deposition of cholesterol in the skin (xanthelasma), tendons (xanthomas) and coronary arteries (atherosclerosis).¹

Their inheritance may be autosomal dominant or recessive and not all dominant disorders have a gene-dose effect. The main defect lies in the LDL-receptor (LDL-R) which is encoded by a gene (45 Kb), located in the short arm of chromosome 19 (p13.1 – p13.3) with 18 exons and 17 introns (Figure 1).²

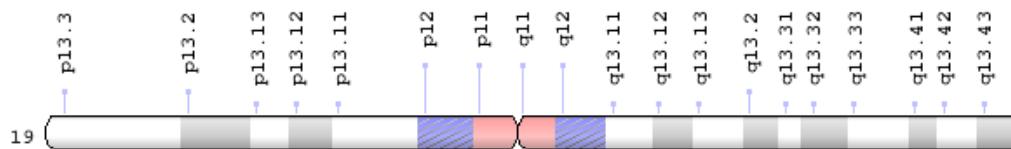


Figure 1. Chromosome 19³

1.1.2. Phenotypes and genetic causes

Familial hypercholesterolemia may be classified into heterozygous or homozygous clinical phenotype.

- Heterozygous

It is the most common form. Historically, the prevalence was 1 in 500 persons, but now is suggested to be 1 in 200 to 250 persons. Heterozygotes arise when a mutation is inherited from one parent only.^{4,5,6,7,8}

This is a monogenic autosomal dominant disorder, where the number of LDL-R drops into 50% being enough to bind the cell, but decreasing the LDL-R capacity to clear LDL from the circulation.^{4,5,6}

¹ L. Ose. *Tidsskr Nor Laegeforen* **2002**, 122, 924-925.

² M. Pocovi, S. Castillo. *Cardiovascular Risk Factors* **2002**, 11, 144156. DOI: [10.3305/nh.2015.32.6.9885](https://doi.org/10.3305/nh.2015.32.6.9885).

³ <https://ghr.nlm.gov/chromosome/19#resources> (03/04/2019)

⁴ P. Hopkins, P. Toth. *J. Clin. Lipidol.* **2011**, 5, 9-17.

⁵ B. Nordestgaard, M. Chapman, S. Humphries. *Eur. Heart J.* **2013**, 34, 3478-3490a.

⁶ M. Cuchel, E. Bruckert, H. Ginsberg. *Eur. Heart J.* **2014**, 35, 2146-2157.

⁷ M. Benn, G. F. Watts, A. Tybjaerg-Hansen. *J. Clin. Endocrinol. Metab.* **2012**, 97, 3956-3964.

⁸ B. Sjouke, D. Kusters, I. Kindt. *Eur. Heart J.* **2015**, 35, 560-565.

The main cause is a mutation in the LDL-R. There are six different types of mutations:²

- Mutation type 1: it is the most common. The alleles are null preventing the synthesis of any receptor. Caused by deletions of LDL-R promoter and also by generating an abnormal m-RNA. This leaves no detectable m-RNA and no detectable LDL-R.
- Mutation type 2: defective alleles for the transport due to wrong receptor tridimensional structure of encoding proteins, being completely or impartially blocked during the transport between endoplasmic reticulum (ER) and Golgi apparatus.
- Mutation type 3: defective alleles for the binding, the receptors encoded proteins to be normally transported to the cell surface, lacking capacity of binding to LDL particles.
- Mutation type 4: defective alleles for internalization. Proteins are encoded to be transported into the cell surface, binding to LDL but being incapable to be regrouped in clathrin vesicles, so, preventing internalization of LDL. Typically, due to mutations in the cytoplasmic tail.
- Mutation type 5: defective alleles producing improper recycling. The receptors are encoded to bind and internalize the ligand in clathrin vesicles but without recycling the ligand and the receptor in the cell surface.
- Mutation type 6: failure of directing LDL-R to the basolateral surface of polarized cells. Therefore these mutations lead to an increase of blood cholesterol, facilitating the formation of xanthelasma, tendon xanthomas, and presence of corneal arcus, hypothyroidism and nephrotic syndrome.

According to recent studies it has been observed that some genes play a very important role in the development of the disease. These genes are listed below:

- ApoB (Apolipoprotein B) gene: is encoded on chromosome 2, located in 2p24.1, which is the short (p) arm of chromosome 2. ApoB gene provides instructions for making two different isoforms: ApoB-48 and ApoB-100.

ApoB-48 is a building block of a lipoprotein type called chylomicron. The main function is to carry fat and cholesterol from the intestine to the bloodstream.^{9,10,11}

ApoB-100 is a building block of very-low-density lipoprotein (VLDL), LDL and intermediate-density lipoprotein (IDL). They all transport fats and cholesterol into the bloodstream. ApoB-100 allows LDL to attach to specific receptors on the cell surface. When ApoB-100 is altered, prevents LDL from effectively binding to their LDL-R on the cell surface. As a result, cholesterol levels are much higher and deposited mainly in skin, tendons and arteries; which increases the risk of developing cardiovascular heart diseases (CHD).¹²

- SREBP (Sterol Regulatory Element-Binding Protein) isoforms have been identified and differ in relative abundance in different tissues. SREBP-1c selectively activates genes involved in fatty acid synthesis, SREBP-1a is a potent activator and maintains basal levels of cholesterol and fatty acids, while SERP-2 (Stress Associated Endoplasmic Reticulum 2) regulates important genes for cholesterol homeostasis by activating the transcription of PCSK9 (Proprotein Convertase Subtilisin/Kexin type 9).^{13,14,15}
- PCSK9 gene: is encoded on chromosome 1 located in 1p32.3, which is the short arm of chromosome 1. PCSK9 controls the number of LDL-R, which are found on the cell surface. In normal conditions, the receptors bind to LDL and are internalized and degraded in the lysosomes, while the receptors are recycled into the cell surface. PCSK9 mutations, leads a gain of function of the protein giving an atypical function. As a result, PCSK9

⁹ T. L. Innerarity, K. H. Weisgraber, K. S. Arnold, R. W. Mahley, R. M. Krauss, G. L. Vega, S. M. Grundy. *Proceed. Nat. Aca. Sci. USA*. **1987**, *84*, 6919-6923.

¹⁰ L. F. Soria, E. H. Ludwig, H. R. Clarke, G. L. Vega, S. M. Grundy, B. J. McCarthy. *Proceed. Nat. Aca. Sci. US A*. **1989**, *86*, 587-591.

¹¹ N. B. Myant. *Atherosclerosis* **1993**, *104*, 1-18.

¹² J. R. Schaefer, H. Scharnagl, M. W. Baumstark, H. Schweer, L. A. Zech, H. Seyberth, K. Winkler, A. Steinmetz, W. März. *Arterioscler. Thromb. Vasc. Biol.* **1997**, *17*, 348-353.

¹³ X. Wang, X. Chen, X. Zhang, C. Su, M. Yang, W. He, Y. Du, S. Si, L. Wang, B. Hong. *EBioMedicine* **2020**, *52*, 102650. DOI: [10.1016/j.ebiom.2020.102650](https://doi.org/10.1016/j.ebiom.2020.102650).

¹⁴ A. N. Tsouka, C. C. Tellis, A. D. Tselepis. *Curr. Pharm. Des.* **2018**, *24*, 3622-3633.

¹⁵ T. F. Whayne. *Arch. Cardiol. Mexico* **2017**, *87*, 43-48.

binds to LDL-R, internalizing LDL and LDL-R. This event, causes the degradation of LDL and also his receptor, driving to an increase of LDL in blood circulation, due to the lack of LDL-R in the cell surface.^{16,17,18}

- **Homozygous**

It is rare and severe LDL hypercholesterolemia. In this case the prevalence is 1 in 1,000,000 persons. Recent studies suggests to be 1 in 250,000. Homozygous arise when includes the concurrence of two LDL-R defects.⁸

This is an autosomal recessive form caused by loss-of-function mutations in the LDL-R adaptor protein 1 (LDLRAP1), which encodes a protein required for clathrin-mediated internalization of LDL-R.⁶ There are also other sterol disorders as phytosterolaemia and cerebrotendinous xanthomatosis (CTX).

- LDLRAP1/ARH (Autosomal Recessive Hypercholesterolemia) gene: is encoded on chromosome 1 located in 1p36-35,¹⁹ related to a problem of LDL-R internalization. The disorder is due to mutations in the adaptor protein for endocytosis of the LDL-R, which are unable to remove LDL from the bloodstream efficiently.^{20,21}
- Phytosterolaemia: this disorder is due to an over absorption of plant sterols from enterocytes and has variable LDL hypercholesterolemia and high plasma plant sterol concentrations.²²
- CTX: is due to mutations in cholesterol 27 hydroxylase, an essential step in cholesterol oxidation in order to synthesize bile acids.²²

As can be seen in Table 1, all isoforms according to the type of gene lead to an increase of cholesterol, thus producing FH, from which other cardiovascular diseases (CVD) are derived.²³

¹⁶ Y.-W. Qian, R. J. Schmidt, Y. Zhang, S. Chu, A. Lin, H. Wang, X. Wang, T. P. Beyer, W. R. Bensch, W. Li. *J. Lipid Res.* **2007**, *48*, 1488-1498.

¹⁷ J. D. Horton, J. C. Choen, H. H. Hobbs. *J. Lipid Res.* **2009**, *50*, S172-S177.

¹⁸ N. G. Seidah. *Curr. Opin. Lipidol.* **2016**, *27*, 274-281.

¹⁹ E. R. Eden, R. P. Naoumova, J. J. Burden, M. I. McCarthy, A. K. Soutar. *Am. J. Hum. Genet.* **2001**, *68*, 653-660.

²⁰ C. K. Garcia, K. Wilund, M. Arca, G. Zuliani, R. Fellin, M. Maioli, S. Calandra, S. Bertolini, F. Cossu, N. Grishin, R. Barnes, J. C. Choen, H. H. Hobbs. *Science* **2001**, *292*, 1394-1398.

²¹ A. K. Soutar, R. P. Naoumova, L. M. Traub. *Arterioscler. Thromb. Vasc. Biol.* **2003**, *23*, 1963-1970.

²² M. H. Moghadasian, G. Salen, J. J. Frohlich, C. H. Scudamore. *Arch. Neurol.* **2002**, *59*, 527-529.

²³ D. Matías-Pérez, E. Pérez-Campos, I. A. García-Montalvo. *Nutr. Hosp.* **2015**, *32*, 2421-2426.

Table 1. Relationship between gene, isoform and derived phenotype²³

GENE	ISOFORM	PHENOTYPE
LDL-R	Heterozygous	Cholesterol deposition in blood and coronary disease
ApoB	Heterozygous	Accumulation of cholesterol in blood
PCSK9	Heterozygous	Related in FH and an increase of cholesterol levels in blood
LDLRAP1/ARH	Homozygous	LDL-R accumulation in membranes cells

1.1.3. Screening strategies

FH is usually diagnosed late and after a severe coronary event. Clinical symptoms may occur during the first decade of the illness, but is usually accompanied by atherosclerosis. For this reason, a guideline of different screening methods is needed in order to diagnose and identify individuals at risk.^{24,25,26,27}

These methods of screening are the following:

- Serum cholesterol²⁴
 - Adults (> 20 years): LDL-C levels \geq 160 mg/dL
 - Children/adolescents (< 20 years): LDL-C levels \geq 160 mg/dL
- Family history of high cholesterol and heart disease among first and second degree relatives of patients diagnosed with FH.^{24,25,26,27}
- Physical findings:²⁴
 - Tendon xanthomas at any age
 - Corneal arcus(< 45 years)
 - Tuberosus xanthomas (20 to 25 years)
- LDL-C levels with 80% of probability of FH in general population:²⁴
 - \geq 250 mg/dL in adults \geq 30 years
 - \geq 220 mg/dL in adults aged 20 to 29
 - \geq 190 mg/dL in patients under age 20

²⁴ A. C. Goldberg, P. N. Hopkins, P. P. Toth, C. M. Balantyne, D. J. Rader, J. G. Robinson, S. R. Daniels, S. S. Gidding, S. D. de Ferranti, M. K. Ito, M. P. McGowan, P. M. Moriarty, W. C. Cromwell, J. L. Ross, P. E. Ziajka. *J. Clin. Lipidol.* **2011**, *5*, S1-S8.

²⁵ P. J. Talmud, S. Shah, R. Whittall, M. Futema, P. Howard, J. A. Cooper, S. C. Harrison, K. Li, F. Drenos, F. Karpe, H. A. W. Neil, O. S. Descamps, C. Langenberg, N. Lench, M. Kivimaki, J. Whittaker, A. D. Hingorani, M. Kumari, S. E. Humphries. *Lancet.* **2013**, *381*, 1293-1301.

²⁶ G. F. Watts, S. Gidding, A. S. Wierzbicki, R. Alonso, W. V. Brown, E. Bruckert, K. K. Lin, P. Mata, K. G. Parhofer, F. J. Raal, R. D. Santos, W. G. Simpsons. *J. Clin. Lipidol.* Vol. 8, **2014**, 148-172.

²⁷ E. Youngblom, M. Pariani, J. W. Knowles. *Gene Rev.* **2014**, 4-7.

1.1.4. Diagnosis

The diagnosis is based on high lipid levels, specifically triglycerides; family history of FH, high levels in plasma of LDL-C, history of premature coronary artery disease (CAD), and physical findings, such as xanthomas and corneal arcus. In homozygous FH is also remarkable the type of mutation, which is important to confirm the diagnosis. There are two types of genetic analysis (Table 2), the serial single-gene testing (sequence analysis of LDL-R and followed by LDL-R deletion/duplication analysis if no pathogenic is found, and, sequence analysis of ApoB and PCSK9) and the multigene panel (ApoB, LDL-R, PCSK9 and others genes).²⁷

Table 2. Detection and proportion of pathogenic genes²⁷

Gene	Proportion attributed to pathogenic variants in this gene	Proportion of pathogenic variants detectable by method	
		Seq. analysis	Gene-targeted
ApoB	1-5%	>99%	1
LDL-R	60-80%	>90%	2.5-10%
PCSK9	0-3%	~100%	Non-reported

Three different clinical tools are developed to diagnose FH. The Make Early Diagnoses Prevent Early Deaths Program Diagnostic Criteria (MEDPED),^{28,29} which uses total and LDL cholesterol measurements and family history; The Dutch Lipid Clinic Network Diagnostic Criteria, which includes LDL cholesterol levels, physical findings and family history,^{5,28,29} and; The Simone Broome Register Diagnostic Criteria, which includes LDL cholesterol levels, family history and genetic mutations.³⁰

These criteria are typically used to diagnose heterozygous FH. To diagnose homozygous FH the criteria is based on the presence of untreated LDL > 500 mg/dL or treated LDL > 300 mg/dL plus, presence of cutaneous or tendon xanthomas < 10 years, both parents with evidence of heterozygous FH and genetic analysis showing mutations in LDL-R, ApoB, PCSK9, LDLRAP1 genes.^{6,31}

²⁸ M. A. Austin, C. M. Hutter, R. L. Zimmern, S. E. Humphries. *Am. J. Epidemiol.* **2004**, *160*, 407-420.

²⁹ A. Haase, A. C. Goldberg. *Curr. Opin. Lipidol.* **2012**, *23*, 282-289.

³⁰ D. From Marks, M. Thorogood, H. A. Neil, S. E. Humphries. *Atherosclerosis* **2003**, *168*, 1-14.

³¹ F. J. Raal, R. D. Santos. *Atherosclerosis* **2012**, *223*, 262-268.

1.1.5. Prognosis

Prognosis depends heavily on the extent to which LDL-C levels can be reduced.

- Heterozygous FH: asymptomatic in childhood and early adulthood (5% heart attacks < 60 years; 20% < 45 years).^{4,5}
- Homozygous FH: extreme hypercholesterolemia with accelerated atherosclerosis. CAD, are the most common and with premature death.^{6,31}

1.1.6. Treatment

Early treatment is beneficial and long term drug therapy can substantially reduce or eliminate the added lifetime risk of CHD from having FH and can lower the CHD event rate.^{5,32,33}

The lifestyle is a very important part of the treatment to reduce risk factors. There are a number of modifications that could be made to improve our lifestyle, such as modifications on the diet (improvement in lipid modifications), physical activity, avoidance of gain weight, avoid smoking and the treatment of diabetes and hypertension.^{26,34}

In homozygous patients, as soon as the diagnosis is made, treatment is rapidly required and changes in lifestyle are needed.^{6,35}

1.1.6.1. Heterozygous FH treatment

The main pharmacologic therapy are statins, considered the first-line treatment in heterozygous FH and may also benefit homozygous patients. Statins have shown to increase the expression of LDL-R by reducing HMG-CoA (3-hydroxy-3-methylglutaryl-CoA) reductase and also leads to a plethora of lipid-independent effects involving anti-inflammatory action and improvement in endothelial function.^{26,34}

³² A. Neil, J. Cooper, J. Betteridge, N. Capps, I. McDowell, P. Durrington, M. Seed, S. E. Humphries. *Eur. Heart J.* **2008**, *29*, 2625-2633.

³³ N. J. Stone, J. Robinson, A. H. Lichtenstein, C. N. B. Merz, C. B. Blum, R. H. Eckel, A. C. Goldberg, D. Gordon, D. Levy, D. M. Lloyd-Jones, P. McBride, J. S. Schwartz, S. T. Shero, S. C. Smith, K. Watson, P. W. F. Wilson. *J. Am. Coll. Cardiol.* **2014**, *63*, 2889-2934.

³⁴ M. Ito, M. McGowan, P. Moriarty. *J. Clin. Lipidol.* **2011**, *5*, S38-S45.

³⁵ F. J. Raal, G. J. Pilcher, V. R. Panz, H. E. van Deventer, B. C. Brice, D. J. Blom, A. D. Marais. *Circulation* **2011**, *124*, 2202-2207.

Moderate to high potency statins should be used to reach a reduction higher or equal to 50% LDL-C from baseline (Table 3). As is popularly known, statins lead to a large number of side effects such as the listed below:³⁶

- Rhabdomyolysis (dose dependent)
- Renal or hepatic impairment
- Hypothyroidism
- Vitamin D deficiency
- Rheumatologic disorders
- Increased levels of AST and ALT
- Drowsiness
- Dizziness
- Diarrhea
- Constipation
- Rash

Statin treatment requires a regular monitoring and possible drug interactions should be noticed due to CYP3A4 interactions.³⁶

Table 3. Reduction of LDL-C (%) for different statins depending on the dose³⁶

Statins	Dose range (mg)	Mean reduction LDL-C (%)
<i>Rosuvastatin</i>	5-40	46-55
<i>Atorvastatin</i>	10-80	37-51
<i>Simvastatin</i>	5-80	26-47
<i>Lovastatin</i>	10-80	21-40
<i>Pravastatin</i>	10-80	20-36
<i>Fluvastatin</i>	20-80	22-35
<i>Pitavastatin</i>	1-4	32-43

It is important to know that doubling the dosage only achieves a reduction of 6-7% more. As shown in Table 3, the LDL-C reduction is never higher than 50% and, patients with risk of CHD require more intensive drug therapy, known as combination therapy. This kind of therapy is given to patients with:^{24,26}

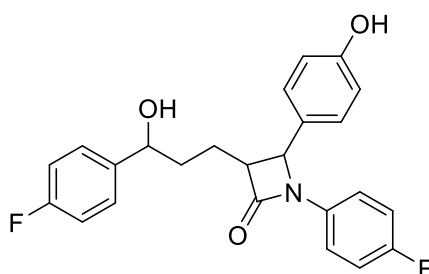
- Evidence with CHD or CAD
- Diabetes
- Family history with early CHD
- Smoking patients
- Two or more CHD risk factors
- High lipoprotein (≥ 50 mg/dL)

³⁶ V. E. Boucharie, A. C. Goldberg. *Cardiol. Clin.* **2015**, 33, 169-179.

Patients may require multiple medications depending on their LDL-C levels and tolerance to statins. This treatment, as said above, is combined with statins and depends on the responsiveness to the treatment.³⁶

The main drugs used in combination with statins are described further down.

- *Ezetimibe*: is an azetidinone derivative and a cholesterol absorption inhibitor. The mechanism of action resides in the interaction with cholesterol transporters at the brush border of the small intestine and inhibits absorption of cholesterol and phytosterols, reducing about 15-20% when is used alone, and provides 20% additional reduction in combination with statins. The main side effects include diarrhea, abdominal pain, dizziness and depressed mood.^{37,38}



Ezetimibe

- Bile acid sequestrants (*Colesevelam*): the mechanism of action resides in bonding to bile acids in the terminal ileum impeding their reabsorption. *Colesevelam* is indicated for co-administration with statins as adjunctive therapy to diet to provide an additional reduction of LDL-C levels. Monotherapy demonstrated to be effective with 15-18% of reduction levels and as adjunctive therapy with statins a reduction of 32-48% of LDL-C levels. As was observed in a study the main adverse effects were:^{39,40}
 - Constipation (7.9%)
 - Abdominal pain
 - Gastrointestinal effects
 - AST/ALT elevated
 - Myalgia (2.1%)

³⁷ <https://pubchem.ncbi.nlm.nih.gov/compound/Ezetimibe> (13/04/2019)

³⁸ C. Gagne, D. Gauder, E. Bruckert. *Circulation* **2002**, *105*, 2469-2475.

³⁹ R. Huijgen, E. Abbink, E. Bruckert, A. F. H. Stalenhoef, B. P. M. Imholz, P. N. Durrington, M. D. Trip, M. Eriksson, F. L. J. Visseren, J. R. Schaefer, J. J. P. Kastelein. *Clin. Ther.* **2010**, *32*, 615-625.

⁴⁰ F. Zieve, M. Kalin, S. Schwartz. *Clin. Ther.* **2007**, *29*, 74-83.

However, *colesevelam* has fewer gastrointestinal side effects and drug interactions compared to other bile acid sequestrants and it has also been approved for diabetes.^{39,40}

- *Niacin*: a water-soluble B vitamin, which lowers LDL-C levels (20%) and raises high-density lipoprotein (HDL) (15-30%). *Niacin* when is added to statins is effective in lowering LDL-C levels and with high doses showed different adverse events such as, headache, rash, dizziness and flushing.^{34,36}
- PCSK9 inhibitors (current treatment): the most promising new therapeutic approach for FH is a completely novel class of drugs called PCSK9 inhibitors, which involves monoclonal antibodies. Nowadays, *alirocumab* and *evolocumab* are two of the most promising drugs. In different studies was shown that therapy with *alirocumab* decreases LDL-C levels by 49% from baseline (Figure 2).^{34,36,41,42}

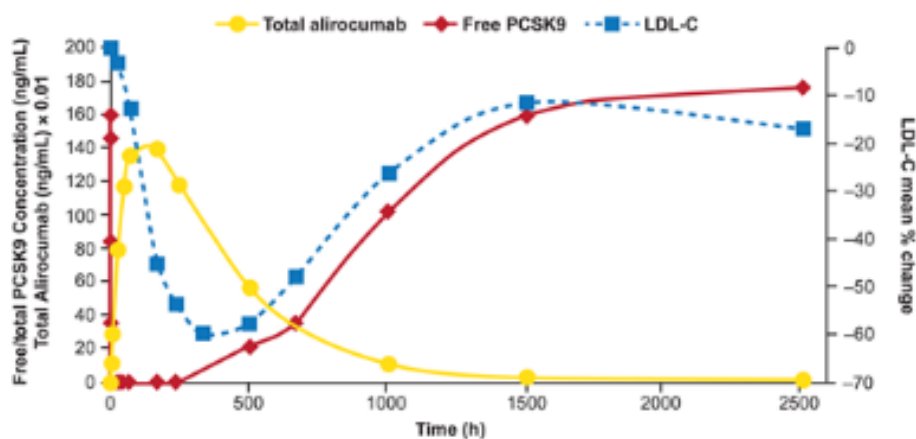


Figure 2. Relationship between *alirocumab*, PCSK9 and LDL-C⁴²

The mechanism of action is well known. Under normal physiological conditions, the LDL-C is bonded to LDL-R in the hepatocyte surface and then are internalized through clathrin-mediated endocytosis. The low pH (Power of Hydrogen) in the endosome leads to dissociation of the LDL from LDL-R, and LDL is degraded in the lysosome, while LDL-R is recycled to the cell surface (Figure 3a).^{41,42}

⁴¹ J. Malo, A. Parajuli, S. W. Walker. *Ann. Clin. Biochem.* **2020**, *1*, 7-25.

⁴² Y. J. Shimada, C. P. Cannon. *Eur. Heart J.* **2015**, *36*, 2415-2424.

However, in the presence of PCSK9 (mutated), the catalytic domain is bonded to the epidermal growth factor-like (EFG) repeat A of the LDL-R and the complex LDL/LDL-R/PCSK9 is internalized and the low pH of the endosome enhances the affinity of PCSK9 for the LDL-R, preventing the LDL-R being recycled to the cell surface. Instead, the LDL-R is degraded in the lysosome. As a result, increases LDL-C circulating in plasma (Figure 3b).^{41,42}

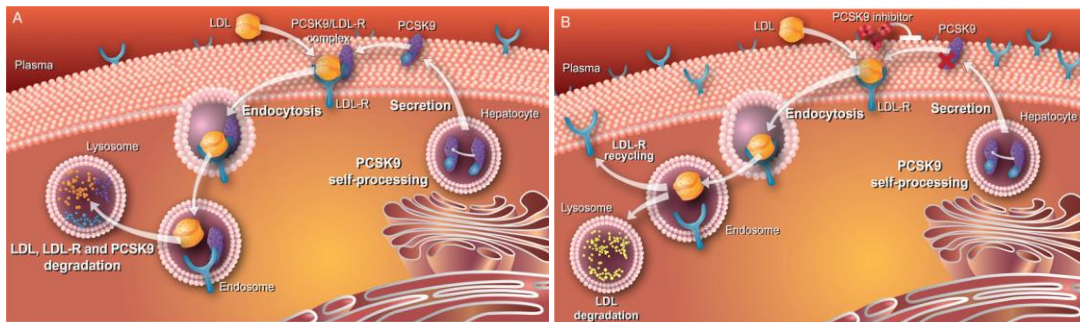


Figure 3. a) Regulation of LDL-R by PCSK9 under normal conditions; **b)** Mechanism of LDL-C reduction by PCSK9 inhibition⁴²

The presence of *alirocumab* or *evolocumab* produces the following result:

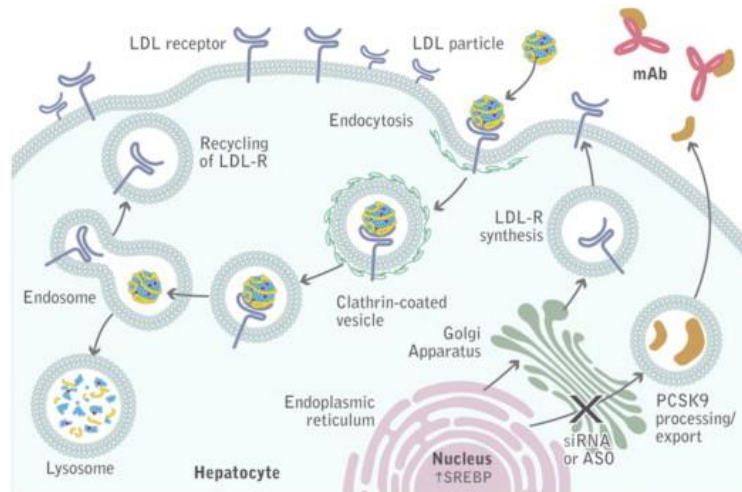


Figure 4. Mechanism of action of monoclonal antibodies⁴³

Both monoclonal antibodies as shown in Figure 4 binds to the PCSK9 protein and allows the LDL-R to bind the LDL and provides the recycling of the LDL-R on the cell surface.⁴³

⁴³ M. H. Davidson. *J. Clin. Lipidol.* **2013**, 3, S11-S15.

The use of monoclonal antibodies has different disadvantages, due to short shelf-life, elevated cost, lack of adherence to the drug due to parenteral administration, and different side effects (nasopharyngitis, respiratory tract infections and back pain).^{41,42,43}

Nowadays, at least nine different modalities of targeting PCSK9 have arisen and 3 strategies of targeting PCSK9 are known in terms of mechanism.⁴⁴

- Strategy 1: preventing the binding between PCSK9 to LDL-R on the cell surface
- Strategy 2: interfering with the maturation and secretion processing
- Strategy 3: developing inhibitors of PCSK9 synthesis and expression of mRNA levels

Other PCSK9 inhibitors:⁴⁴

- Blockers of the ribosome
- Capturers of PCSK9: heparin sulfate mimetics
- EFG-(A) mimetic drugs
- Disrupters of the protein-protein interaction (PPI): imidazole peptidomimetics
- PCSK9 inhibitors of maturation, expression or secretion: natural products, such as *barberine*, *corydaline*, *curcumin*, lignans, *moracin C*, *resveratrol*, *tanshinone II* among others

1.1.6.2. Homozygous FH treatment

In homozygous patients is essential to treat as soon as the diagnosis is made. The main treatment involves diet, statins, ezetimibe and often apheresis. Two novel treatments have been approved by the Food and Drug Administration (FDA), *lomitapide* and *mipomersen*.

- *Lomitapide*: binds and inhibits the microsomal triglyceride transfer protein (MTTP). The function of MTTP is to assist in the transfer of triglycerides to ApoB in the hepatocyte to form VDL and in the enterocyte to form chylomicrons.⁴⁵

⁴⁴ S. Xu, S. Luo, Z. Zhu, J. Xu. *Eur. J. Med. Chem.* **2019**, 162, 212-233.

⁴⁵ M. Cuchel, L. Bloedon, P. O. Szapary, D. M. Kolansky, M. L. Wolfe, A. Sarkis, J. S. Millar, K. Ikewaki, E. S. Siegelman, R. E. Gregg, D. J. Rader. *N. Engl. J. Med.* **2007**, 2, 148-156.

The main adverse effects are the following:⁴⁵

- High risk of hepatotoxicity
 - Gastrointestinal intolerance (diarrhea, flatulence)
 - Vitamin E deficiency
 - AST/ALT elevated levels
- *Mipomersen*: is an anti-sense oligonucleotide that encodes ApoB-100 produced by the liver. The delivery is by subcutaneous injection. The mechanism of action is characterized by binding m-RNA and inhibiting ApoB-100 synthesis causing LDL-C reduction. The main side effects include: ^{46,47}
 - Steatosis
 - Erythema
 - Flu-like symptoms

In both medications monitorization of liver function is needed (Table 4). The therapy may be considered only as adjunctive to diet and cholesterol-lowering drugs, since they are only available through the restricted risk evaluation and mitigation strategy (REMS) Program.^{36,46,47}

Table 4. Treatment for homozygous isoform³⁶

Treatment	Dossage	Mean reduction LDL-C (%)	Drug interactions	Safety issues	Side effects
<i>Lomitapide</i>	5-60 mg orally per day	38-50	<i>Atorvastatin</i>	Increased hepatic enzymes	Diarrhea, requires fat diet and fat soluble vitamins
<i>Mipomersen</i>	200 mg once per week	24-28	None	AST/ALT increases	Injection reactions, pyrexia, malise, fatigue, headache, nausea

⁴⁶ E. Stein, R. Dufour, C. Gagne. *Circulation* **2012**, *19*, 2283-2292.

⁴⁷ F. J. Raal, R. D. Santos, D. J. Blom. *Lancet*. **2010**, *375*, 998-1006.

1.1.6.3. Heterozygous and Homozygous FH treatment

When either of all the medications explained above do not achieve the treatment goals, different methods are used to remove LDL from circulation.

- LDL apheresis: is generally done every 1 to 2 weeks taking 3 hours per session. It consists in removing greater than 60% of ApoB containing lipoproteins. The reduction of LDL is temporary and associated with an elevated rebound in lipid levels. The efficacy can be enhanced by the addition of statin therapy achieving a 31% lowering LDL levels.⁴⁸
- Surgical therapy: partial ileum bypass and liver transplantation. The last one, produces a significant lowering of LDL-C by providing normal LDL-R. There is risk of transplant and the number of donors is limited. On the other hand, ileum bypass is rarely used and works by interrupting enterohepatic bile acid circulation.^{49,50,51}

After studying all the possible therapeutic options (Table 5), it was concluded that there is still a long way forward in the research into new FH inhibitors.

Table 5. LDL-C lowering response of the main drugs used in FH²⁷

Class	Mechanism of action	LDL-C lowering response (%)
Statins	↑ LDL-R activity	>35
<i>Ezetimibe</i>	↓ Cholesterol absorption	15
<i>Mipomersen</i>	Blocks ApoB production	50
<i>Lomitapide</i>	↓ MTTP activity	50
PCSK9 inhibitors	↓ LDL-R degradation	50
<i>Colesevelam</i>	↓ Bile acid re-absorption	15

The most promising drugs are currently PCSK9 inhibitors, but as seen, they have an elevated economic cost and in addition, their administration limits their adherence to the drug. For this reason, in this dissertation thesis, a complete study of the PCSK9 protein has been performed and new small organic molecules have been synthesized as potential PCSK9 inhibitors.

⁴⁸ G. R. Thompson, A. Catapano, S. Saheb. *Curr. Opin. Lipidol.* **2010**, *21*, 492-498.

⁴⁹ J. C. Choen, E. Boerwinkle, T. H. Mosley, H. H. Hobbs. *N. Eng. J. Med.* **2006**, *354*, 1264-1272.

⁵⁰ E. Stein, N. Honarpour, S. Wasserman. *Circulation* **2013**, *128*, 2113-2120.

⁵¹ E. Stein, S. Mellis, G. Yancopoulos. *N. Eng. J. Med.* **2012**, *366*, 1108-1118.

1.1.7. Molecular target

1.1.7.1. PCSK9 protein

1.1.7.1.1. PCSK9 family

PCSK9 is one of nine serine proteases. The first 8 proteases, proprotein convertases (PCs) 1-8, catalyze the proteolytic maturation of inactive secretory precursors to mature proteins, such as, hormones, enzymes and peptides. However, the PCSK9 is synthesized and secreted mainly in the liver as a 75-kDa zymogen consisting of a signal peptide, the prosegment, a catalytic domain (CD) and a cysteine-histidine rich C-terminal domain.^{52,53,54,55}

1.1.7.1.2. Maturation and secretion

Following its synthesis, PCSK9 undergoes an autocatalytic reaction required for the secretion and maturation from ER to Golgi. The co-translational autocatalytic cleavage takes place in the ER to yield two products, the prosegment and a 62-kDa mature PCSK9 protein containing the catalytic domain and the C-terminal domain.⁵²⁻⁵⁵

The continued attachment of the prodomain partially blocks the catalytic site and inactivates PCSK9. In the other PCs, a second cleavage is required to release the prodomain, which unmasks the catalytic site, resulting in an active protease. No secondary cleavage site has been identified that activates PCSK9.⁵²⁻⁵⁵

At the cell surface, PCSK9 binds to the EFG-(A) domain (residues from 293 to 338) of the LDL-R and it is consequently internalized as a complex in the endosomes, via clathrin-coated pits. Remarkably important, in an acidic pH the affinity of the PCSK9 for the LDL-R increases. The binding results into a redistribution of LDL-R to the lysosomes.⁵²⁻⁵⁵

1.1.7.1.3. PCSK9 structure

The PCSK9 is a 692 amino acid protein (Figure 5). The prodomain (residues from 31 to 152) serves a dual role as a chaperone for folding and as an inhibitor of its catalytic activity. The autocatalytic cleavage of the prodomain is produced between GLN152 and SER153.

⁵² R. M. DeVay, D. L. Shelton, H. Liang. *J. Biol. Chem.* **2013**, 10805-10818.

⁵³ S. Poirier, G. Mayer. *Drug Des. Devel. Ther.* **2013**, 7, 1135-1148.

⁵⁴ L. Wierod, J. Cameron, T. B. Strom, T. P. Leren. *Mol. Gen. Metab. Report* **2016**, 86-93.

⁵⁵ A. S. Peterson, L. G. Fong, S. G. Young. *J. Lipid Res.* **2008**, 6, 1152-1156.

The CD comprises residues from 153 to 454, while the C-terminal domain is composed of residues from 455 to 692.^{54,56}

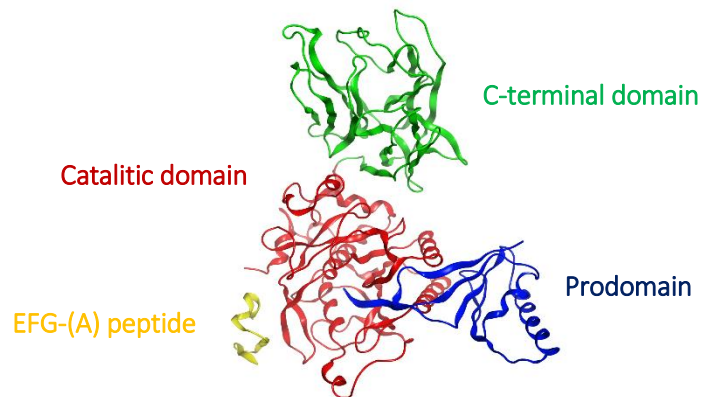


Figure 5. In blue the prodomain, in green the C-terminal domain, in red the catalytic domain and in yellow the EFG-(A) peptide

The prodomain is formed by two α -helices and four stranded antiparallel β -sheet as shown in Figure 6.⁵⁶

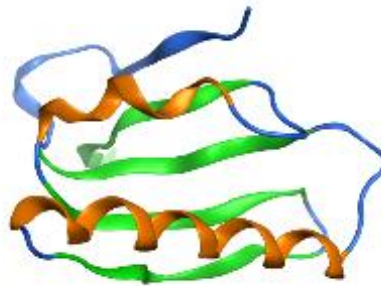


Figure 6. Prodomain structure; in green the β sheet and in orange the α -helices

⁵⁶ D. E. Piper, S. Jackson, Q. Liu, S. T. Thibault, B. Shan, N. P. C. Walker. *Structure* **2007**, 545-552.

The CD includes 7 stranded parallel β sheet flanked on both sides by α -helices (Figure 7a). The catalytic triad is made up of residues ASP186, HIS226 and SER386 (Figure 7b).⁵⁶

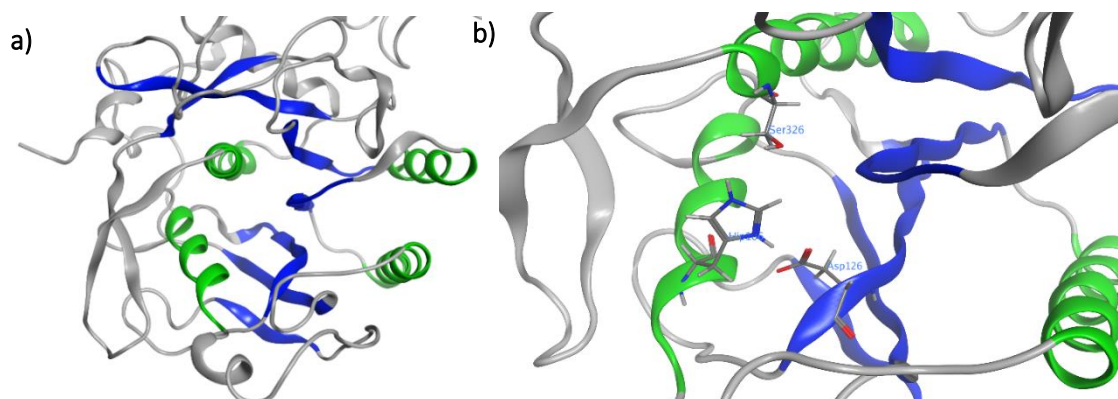


Figure 7. a) Catalytic domain; in blue the β sheet and in green the α -helices; b) Catalytic triad

The C-terminal domain contains six-stranded β -sheet subdomains arranged with quasi-threefold symmetry. Each β strand is held together by 3 internal disulfide bonds. Highly rich in cysteines and histidine's. The binding with the CD realizes in the 3 hydrogens bonds at the interface of this domain.⁵⁶

Surprisingly, the allosteric binding sites of the protein have not been reported yet in the literature. Therefore, the study of this protein is of great interest as it is a recently discovered target and has not been studied in depth by other research groups.

1.2. Cancer

1.2.1. Cancer introduction

Cancer is a group of diseases characterized by abnormal growth of cells with great potential for invasion and division beyond their usual boundaries, which invade and spread into other parts of the body through the blood or lymph system forming metastasis. It is noteworthy to emphasize that metastatic cancer has the same name and the same type of cancer cells as the original tumor.^{57,58}

Not every change in the body tissue or rapid cell formation and growth develops in cancer. In hyperplasia, for example, the tissue division is faster than normal cells and new

⁵⁷ <http://www.who.int> (07/03/2020)

⁵⁸ <http://www.cancer.gov> (07/03/2020)

extra cells are formed; and in dysplasia, an abnormal organization of the cells is observed and new extra cells are also built up.⁵⁸

It is well known that cancer is a genetic disease related to mutations in the genes that control cell division and proliferation. These mutations can be inherited or acquired during the lifetime due to cell division or environmental exposure (tobacco, UV). It is remarkable that each person has a particular combination of genetic changes and, as the cancer continues to develop, additional changes may occur. These genetic changes mainly affect four genes:⁵⁸

- Proto-oncogenes: involved in cell growth and division. The alteration of these genes leads to cell growth and increases the survival of the cells, thus promoting cancer (renamed as oncogenes). Ex: RAS, ERK.
- Tumor suppressor genes: involved in controlling cell division. Alterations may cause uncontrolled cell division. Ex: BRCA1, p53.
- DNA repair genes: involved in fixing damaged DNA. Alterations in these genes promotes the mutations in other genes, which may cause the cells to become cancerous. Ex: BER, NER, MMR.
- Apoptotic genes: when the DNA or the cell is too damaged to be repaired, the programmed cell death is activated and undergoes to apoptosis. Mutations in these genes can cause the cells to become cancerous.

1.2.2. Types of cancer

There are more than 100 types of cancer, many of them named according to the place where they are formed or by their tissue type. For example, lung cancer begins in the cells of the lung, while the cancer that begins in the colon cells is called colorectal cancer. The main categories of cancer are the following:^{58,59}

- Carcinoma: developed by epithelial cells. Concretely, it begins in the skin or in tissues that covers the inside or outside of internal organs. Carcinoma is the most common type of cancer. As there are different epithelial cell types, different subtypes of cancers can be developed.

⁵⁹ <http://www.cancerresearchuk.org> (08/03/2020)

- Adenocarcinoma: starts in the glandular cells which are called adenomatous cells. These cells are responsible of keeping the humidity of the tissues. Breast cancer, colorectal or prostate cancer are found in this group.
- Squamous cell carcinoma: cancer that starts in squamous cells, beneath the outer part of the surface skin.
- Basal cell carcinoma: begins in the lower layer of the epidermis, which is the outer layer of the skin.
- Transitional cell carcinoma: formed in the transitional epithelium or urothelium, which is formed by many layers of epithelium cells that can get bigger or smaller. Some of them are found in ureters and kidneys.
- Sarcoma: starts in the connective tissues, which include bone and soft tissues, responsible of supporting the body tissues. The most common cancers are osteosarcoma (bone cancer) and soft tissue sarcoma such as Kaposi sarcoma.
- Leukemia: begins in the marrow bone, where the blood cells are not fully formed and extra white blood cells are found. A lack of normal cells makes it harder for the body to get the appropriate oxygen to the tissues, control the bleeding and fight infections. This kind of cancer does not produce solid tumors.
- Lymphoma and myeloma: cancers of the lymphatic system. They are involved in filtering the fluids of the body and fighting infections.
 - Lymphoma: it is an uncontrolled division of lymphocytes before their mature, which causes the body to be unable to fight infections. There are two main types, Hodgkin lymphoma and Non-Hodgkin lymphoma.
 - Multiple myeloma: cancers that begin in plasma cells, which are found in the marrow bone. These cells produce many types of antibodies. When they are subject to undesired changes, they just produce one type of antibodies making the body weak to fight against infections.
- Melanoma: their origin is the melanocytes and their precursors. These cells are specialized in synthesizing melanin. The most common melanomas are in the skin but they can be formed in other organs such as the eyes.
- Brain and spinal cord tumors: tumors that start in the brain or spinal cord. The most common brain tumor is glioma, which is formed in glial cells.

1.2.3. Epidemiology

According to the last data published by GCO (Global Cancer Observatory) the total number of new cases worldwide in 2018 was 18,078,957 (8,622,534 in women and 9,456,418 in men). The mortality arise to 9,555,027 in 2018, the higher percentage of deaths being in men (Figure 8).⁶⁰

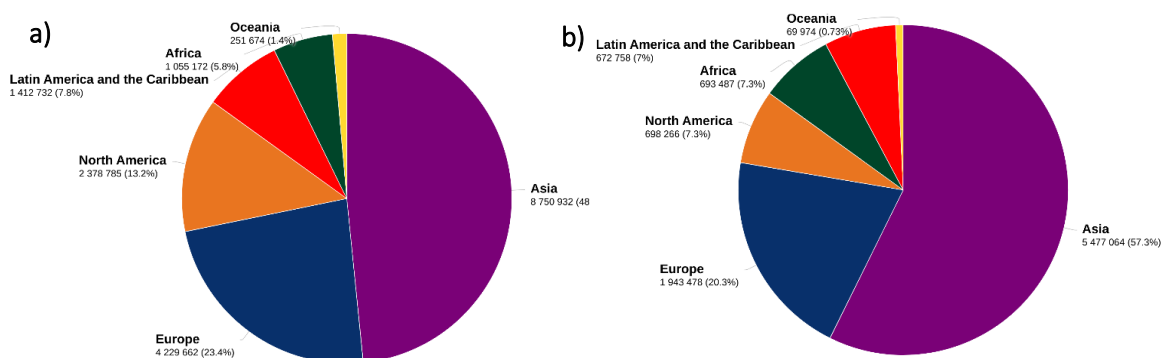


Figure 8. a) Estimated number of new cases in 2018, both sexes, all ages around the world; b) Estimated number of deaths in 2018, both sexes, all ages around the world⁶⁰

The types of cancer that caused most deaths worldwide in 2018 for both sexes were lung cancer (18.4%), colorectal cancer (9.2%) and stomach cancer (8.2%). On the other hand, the types of cancer with more incidence in both sexes worldwide in 2018 were lung cancer (11.6), breast cancer (11.6%) and colorectal cancer 10.2%). The incidence and mortality reported for men and women was slightly different as shown below (Table 6, Table 7).⁶⁰

Table 6. Global incidence of cancer in 2018⁶⁰

Men		Women	
Lung cancer	1,386,524 (14.5%)	Breast cancer	2,088,849 (24.2%)
Prostate cancer	1,276,106 (13.5%)	Colorectal cancer	823,303 (9.5%)
Colorectal cancer	1,026,215 (10.9%)	Lung cancer	725,352 (8.4)

⁶⁰ <http://www.gco.iarc.fr> (08/03/2020)

Table 7. Global mortality of cancer in 2018⁶⁰

Men		Women	
Lung cancer	1,184,947 (22%)	Breast cancer	626,679 (15%)
Liver cancer	548,375 (10.2%)	Lung cancer	576,060 (13.8%)
Stomach cancer	513,555 (9.5%)	Colorectal cancer	396,568 (9.5%)

1.2.3.1. Epidemiology in Spain

The data reported by INE (Instituto Nacional de Estadística) in 2017, indicated that the main causes of death in women were primarily due to diseases of the circulatory system (54%) and followed by cancer (40%). However, the first cause of death in men were tumors, while circulatory diseases were the second. Whereas diseases from the respiratory system were increased by 10% in both sexes.⁶¹

In 2018, new cases of cancer were estimated, in particular a total of 270,363. In Spain, the most diagnosed cancers were colorectal cancer (37,172), breast cancer (32,825) and prostate cancer (31,728). On the other hand, the cancer that caused the most deaths in 2018 were lung cancer (22,896), colorectal cancer (16,683) and pancreatic cancer (7,279).⁶⁰ However, the types of cancer that cause death are different if is differentiated by gender.

As shown in Figure 9, in 2018 the main mortal cancers for men were lung cancer (17,559; 25.5%), colorectal cancer (10,038; 14.6%) and prostate (5,793; 8.4%). Whereas it was colorectal cancer (6,645; 14.9%), breast cancer (6,421; 14.4%) and lung cancer (5,337; 11.9%) for woman.⁶⁰

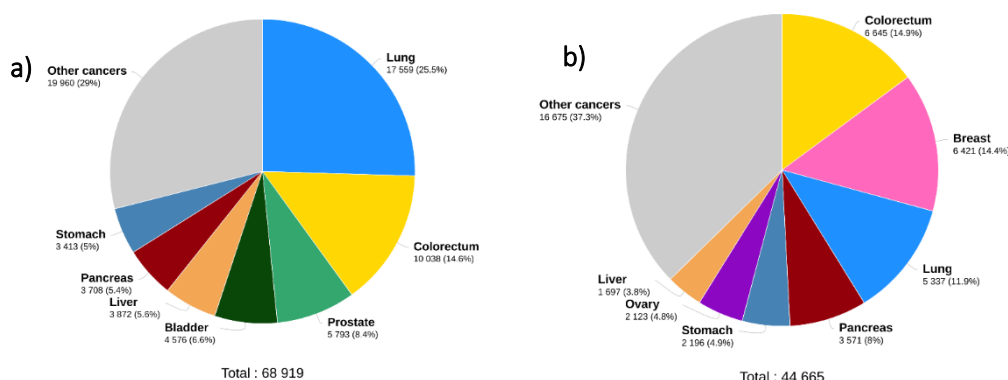


Figure 9. a) Estimated number of men deaths (2018, Spain); **b)** Estimated number of woman deaths (2018, Spain)⁶⁰

⁶¹ <http://www.ine.es> (08/03/2020)

As it can be seen, cancer is one of the leading causes of morbi-mortality. According to REDECAN (Red Española de Registros de Cáncer) the number of cases diagnosed in Spain is estimated to reach 277,394 by 2020. The estimations also indicate that the most commonly diagnosed cancers will be colorectal cancer (44,231 new cases), prostate cancer (35,126 new cases) and breast cancer (32,953 new cases). In men, the main diagnosed cancers will be colorectal and lung cancer, while in women, it will be breast and colorectal cancer.⁶²

The number of cases in Spain is increasing, probably due to population growth, as well as, due to exposure to risk factors and age. Nevertheless, in other cancers such as colorectal, the increase of early detection cases should reduce the mortality.⁶²

After analyzing all the data, colorectal cancer is one of the most diagnosed and with higher mortality rate. This is why in this dissertation thesis aims to study one of the main targets for this type of cancer, such as KRAS (Kristen Rat Sarcoma).

1.2.4. Molecular target

1.2.4.1. RAS

RAS proteins belong to the family of proteins that bind guanosine diphosphate (GDP) and triphosphate nucleotide (GTP), with intrinsic activity like GTPase. These proteins act like signal transducers. The best-studied members of RAS family proteins are KRAS, HRAS (Harvey Rat Sarcoma) and NRAS (Neuroblastoma Rat Sarcoma). These proteins play a very important role in signalling, proliferation, differentiation and cellular senescence.⁶³

Mutations in RAS gene are powerful proto-oncogenes, which are responsible for the development of certain types of cancers (Table 8).⁶³

⁶² <http://www.seom.org> (08/03/2020)

⁶³ M. Porru, L. Pompili, C. Caruso, A. Biroccio, C. Leonetti. *J. Exp. Clin. Cancer Res.* **2018**, 37-57.

Table 8. Main cancers produced by RAS^{63,64}

MUTATION	TYPES OF CANCER	FREQUENCY (%)
RAS mutation	Pancreatic	90 (K)
	Lung adenocarcinoma	35 (K)
	Colorectal	45 (K)
	Thyroid (follicular)	55 (H, K, N)
	Thyroid (papillary)	60 (H, K, N)
	Testicle cancer	45 (K, N)
	Melanoma	15 (N)
	Bladder	10 (H)
	Liver	30 (N)
	Kidney	10 (H)
	Myelodysplastic syndrome	40 (N, K)
	Acute myeloid leukemia	30 (N)

1.2.4.2. KRAS

1.2.4.2.1. KRAS structure

The KRAS gene encodes for KRAS protein, located on human chromosome 12. This chromosome contains 188 residues and has a molecular mass of 21.6 kD. The KRAS gene consists of six exons. Codons 12 and 13 encode for a glycine and they are found at exon 1. However, codon 61 encodes for a glutamine and is found at exon 2. In addition, there exists two splice variants of KRAS protein due to a variation at exon 4, which induces different C-terminal sequences: exons 4A and 4B encode 39 and 38 amino acids respectively, with 19 identical and 4 conserved substitutions. Of note, KRAS4A is predominant in viral KRAS while the predominant isoform in human cells is KRAS4B.^{63,65}

There are three main structural components in KRAS proteins:^{63,65}

- The catalytic domain (residues 1-166)
Highly similar within the three isoforms, this domain is comprised of six β -sheets (β 1- β 3 and β 4- β 6) and five α -helices (α 1- α 5). These are interconnected by 10 phosphate-binding loop (L1 (loop P)-L10) and two parts of the nucleotide-binding switch regions (Switch I (SwI) and Switch II (SwII) (Figure 10).

⁶⁴ Y. Pylayeva-Gupta, E. Grabocka, D. Bar-Segi. *Nat. Rev. Cancer* **2011**, *11*, 761-774.

⁶⁵ A. D. Cox, C. J. Der. *Small GTPases* **2010**, *1*, 2-27.

These motifs are located among the frequently mutated amino acid residues GLY13, GLY15 and GLN61. They are essential for RAS catalytic activity such as changing the conformation during GDP/GTP cycling and binding phosphate/Mg²⁺ and the guanine base of GDP and GTP.

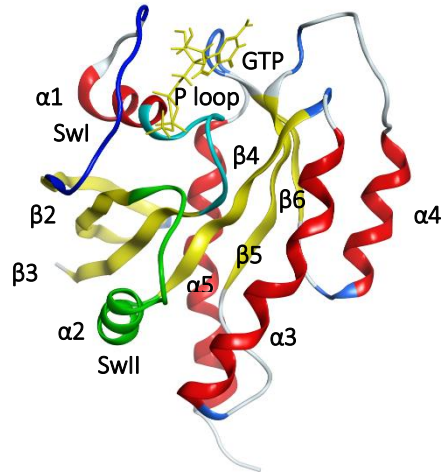


Figure 10. Functional sites of KRAS catalytic domain

- The C-terminal hypervariable region HVR (residues 167-188/189)
The HVR site is critical for RAS post-translational modifications to translocate it to the plasma membrane. Although the HVRs of the three isoforms share only 15% homology. The process of each RAS isoform post-translational modification is very similar. The responsible parts of this process are half of $\alpha 5$ C-terminal and the C-terminal of the protein.
- KRAS active site
The GTP and GDP are found in P loop and $\alpha 1$ and switch regions I and II (Figure 11). The binding of GTP or GDP produces a change of conformation of switch regions, which helps other proteins to recognize KRAS.

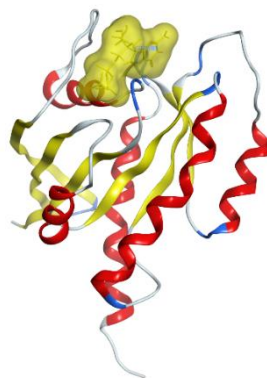


Figure 11. Molecular surface of GTP represented in yellow

As seen above, KRAS is a clinically relevant target for the treatment of colorectal cancer. KRAS is thought to be directly undruggable as no drugs have been seen to bind it specifically. For this reason, it is interesting to know what the major binding sites of the protein are, in order to synthesize a drug that could bind it in an allosteric site and inhibit the protein.

The allosteric binding sites were first discovered through theoretical techniques based on site prediction and later by resolution of crystalline structures.^{66,67,68}

Nowadays, different allosteric binding sites are well known. The main pockets associated to the allosteric binding sites are described below:^{66,68}

- Pocket P1

Pocket P1 is located behind SwII and with $\alpha 2$; and in front of beta sheets 1 to 3 as shown in Figure 12.

- Hydrophobic cavity: made by aromatic rings (indole or imidazole ring). The following amino acids are found: VAL7, LEU56 and MET67. This cavity is very important due to the interactions that establishes with KRAS substrates.
- Polar region: formatted by LYS5, ASP54, THR74 and TYR71.
- Electronegative cavity near GLU37 and ASP38.

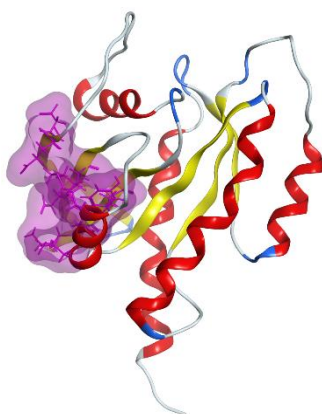


Figure 12. Molecular surface of pocket 1

⁶⁶ B. J. Grant, S. Lukman, M. J. Hocker, J. Sayyah, J. H. Brown, J. A. McCammon, A. A. Gorfe. *PLoS One* **2011**, *6*, 10.

⁶⁷ G. Buhrman, C. O'Connor, B. Zerbe, B. M. Kearney, R. Napoleon, E. A. Kovrigina, S. Vajda, D. Kozakov, E. L. Kovrigin, C. Mattos. *J. Mol. Biol.* **2011**, *4*, 773-789.

⁶⁸ P. Prakash, J. F. Hancock, A. A. Gorfe. *Proteins* **2015**, *83*, 898-909.

- Pocket P2

This pocket is found in SwII, β 1 sheet and P loop and helix α 3 (Figure 13). This pocket is only visible when GDP is bound to the protein, since GDP does not interact with GLY60 as GTP does. The main amino acids that constitute this pocket are VAL7, VAL9, GLY60, MET72, PHE78, GLN99 and ILE100.

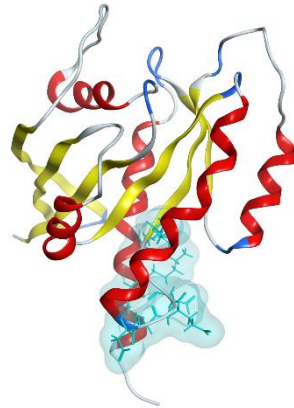


Figure 13. Molecular surface of pocket 2

- Pocket P3

Located in the C-terminus of helices α 5- α 3 (Figure 14). The residues involved in this cavity are ASP105, SER106, GLU107, ASP108, MET111, GLU162, LYS165 and HIS166.

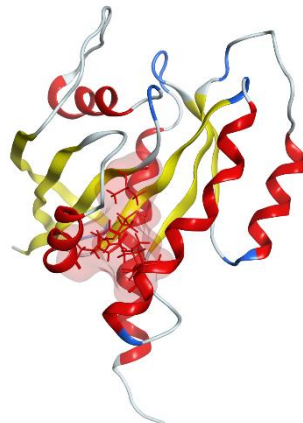


Figure 14. Molecular surface of pocket 3

- Pocket P4

Behind SWI region. The pocket is only visible when this region adopts an open conformation. No crystal structure of a ligand bound at this site is currently available. However, studies shown that Andrographolide derivatives suited into this pocket (Figure 15). Polar residues are found such as ASP30, ASP33, ASP38, SER39 and TYR40. A hydrophobic cavity is also found by ILE21 and ILE36.

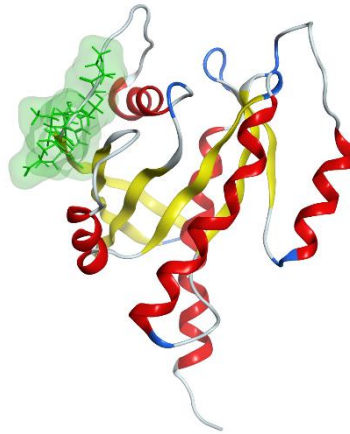


Figure 15. Molecular surface of pocket 4

The main possible allosteric sites of the protein are summed up in Figure 16. They all could be druggable and KRAS inhibitors could bind in these pockets.

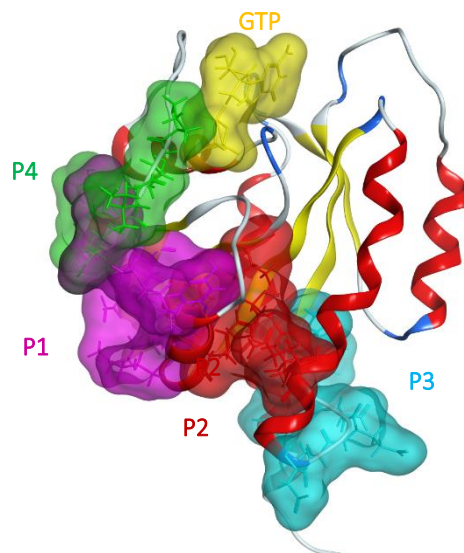


Figure 16. Molecular surface of pockets P1 (in pink), P2 (in red), P3 (in blue), P4 (green) and GTP molecule (in yellow)

1.2.4.2.2. Mechanism of activation

KRAS protein can be found in two different conformations, open or closed. In the closed conformation it is linked to GDP, while in the open conformation it is linked to GTP. Under physiological conditions, the switch from one conformation to the other is regulated by the guanine exchange factor (GEF), which promotes KRAS activation. Through the GDP/GTP switch and the GTPase-activating proteins (GAP), the hydrolysis of GTP is accelerated (Figure 17).^{58,69}

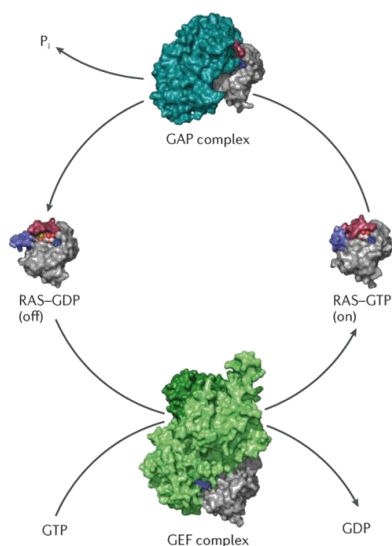


Figure 17. Activation mechanism for KRAS⁷⁰

1.2.4.2.3. Post-translational modifications

To become functionally active, it first must be in the cytosol and needs to be activated by means of anchoring to the plasma membrane. Once the RAS protein is translated, it suffers a series of post-translational modifications that lead to the incorporation of a farnesyl (F) and palmitoyl molecule at the C-terminal of the protein. The process by which the RAS protein is activated (Figure 18) is indicated below:⁷¹

- Introduction of a farnesyl pyrophosphate (FPP) molecule to the cysteine residue in the CAAX box sequence for farnesylation, which allows the anchorage of the protein to the microsomal membrane. It is important to note that KRAS4A, KRAS4B and NRAS can either be farnesylated or geranylgeranylated while HRAS can only be done by farnesylation.

⁶⁹ A. A. Samatar, P. I. Poulidakos. *Nat. Rev. Drug Discov.* **2014**, *13*, 928-942.

⁷⁰ A. Mullard. *Nat. Rev. Drug Discov.* **2019**, *18*, 887-891.

⁷¹ P. A. Konstantinopoulos, M. V. Karamouzis, A. G. Papavassiliou. *Nat. Rev. Drug Discov.* **2007**, *6*, 541-555.

- The three terminal amino acids (AAX) are eliminated by endoprotease RCE1 (Ras Converting Enzyme 1).
- The terminal carboxylic group of cysteine is methylated by the ICMT (Isoprenylcysteine carboxyl Methyltransferase) enzyme.
- Finally, once methylated, the RAS proteins NRAS, HRAS and KRAS4A are matured by a process of palmitoylation in the Golgi and then they are attached to the plasma membrane using their farnesyl or geranylgeranyl (GG), and palmitoyl moieties. However, protein KRAS4B, does not go through palmitoylation but it attaches to the plasma membrane through its farnesyl moiety and through a polybasic lysine-rich sequence located near the terminal cysteine.

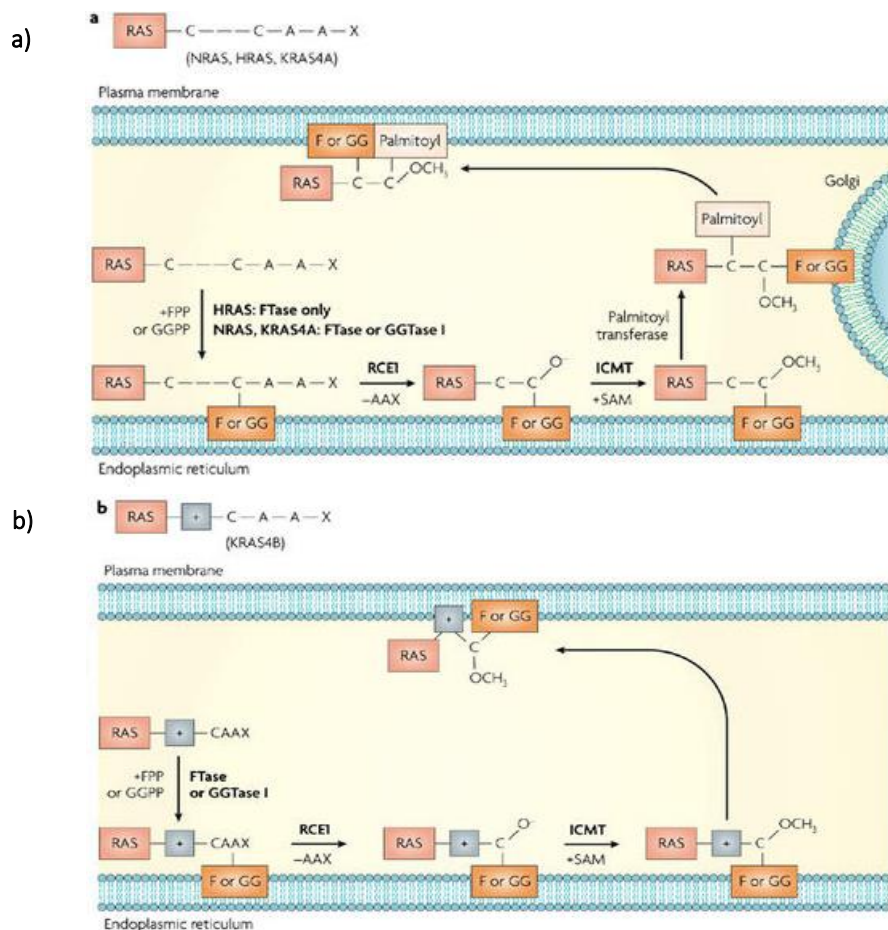


Figure 18. a) HRAS, NRAS and KRAS4A post-translational modifications; **b)** KRAS4B post-translational modifications⁷¹

Once KRAS is in the plasma membrane, it participates in different signal transduction pathways, starting with the union of growth factors (GF) in the tyrosine kinase receptor (RTK) which through dimerization, transphosphorylation and autophosphorylation of tyrosine residues, is activated.

These tyrosine residues are specifically recognized by SH2 (Src Homology 2) and SH3 proteins, which at the same time are recognized by SOS (Son of Sevenless). The SOS protein is a GEF responsible for the conformational change of KRAS, from RAS-GDP to RAS-GTP, the latest being the active form of RAS.⁷¹

The activation of the KRAS protein promotes the growth and proliferation of the cells. In many types of cancer, tumor cells have been shown to be capable of permanently releasing the growth factor, leading to the continuous activation of KRAS. As a final result, the growth and uncontrolled proliferation of the cell occurs.⁷¹

1.2.4.2.4. RAS downstream effectors signalling: primary pathways

The proximal downstream RAS effectors are defined as proteins that have a strong affinity to GTP-RAS and share a common characteristic of RAS-binding domain, which is the RAS core effector domain. The following are the three major RAS downstream signalling pathways:

- Mitogen-activated protein kinase/extracellular signal-regulated kinase (MAPK/ERK) pathway

This pathway begins with RAS activation through the binding of EGF to RTK, as described above. Once RAS is activated, successive steps of phosphorylation amplify the signal in MAPK/ERK pathway as follows (Figure 19):^{72,73}

- First, activated RAS recruits RAF (Serine/threonine-specific protein kinase) to the plasmatic membrane where it is phosphorylated.
- Then, RAF induces the phosphorylation of MEK (Mitogen-activated protein kinase kinase) leading to the activation of MEK1 and MEK2.
- Thereafter, the activated MEK1/2 phosphorylates the ERK protein, leading to the activation of the final effectors of the MAPK pathway, ERK1 and ERK2.
- Once phosphorylated, ERK regulates various cytosolic and nuclear transcription factors including Sp1, E2F, ELK-1 and AP-1.

⁷² M. Burotto, V. L. Chiou, J.-M. Lee, E. C. Kohn. *Cancer* **2014**, *120*, 3446-3456.

⁷³ E. Puxeddu, C. Durante, N. Avenia, S. Filetti, D. Russo. *Trends Endocrinol. Metab.* **2008**, *19*, 138-145.

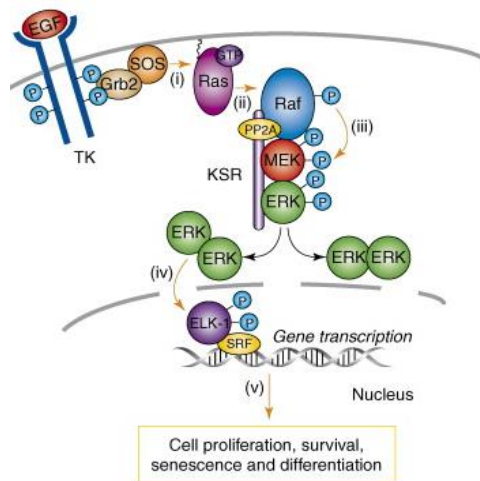


Figure 19. MAPK/ERK pathway⁷³

- Phosphatidylinositol 3-kinase (PI3K) pathways

PI3K proteins are classified in three classes according to their structural characteristics and substrate specificity. However, the most commonly studied is the class IA, which is activated directly by cell surface receptors including RTKs, GPCRs and certain oncogenes including the small G protein RAS.⁷⁴

Once growth factors have bound to RTKs, the following steps describe PI3K activation (Figure 20):⁷⁴

- In response to growth factor stimulation and the subsequent activation of RTKs, PI3K activation may be achieved by the binding of its p85 regulatory subunit to an activated receptor. Alternatively, EGF-RTK signalling may activate the small G protein RAS, which in turn, recruits PI3K to the plasma membrane and induces PI3K activation.
- Once PI3K is activated, the p110 catalytic subunit converts phosphatidylinositol-4,5-trisphosphate (PIP₂) to phosphatidylinositol-3,4,5-trisphosphate (PIP₃), which can serve as a docking site for signalling proteins that have pleckstrin homology domains. This domain can bind phosphatidylinositol lipids within biological membranes. This facilitates the recruitment of PDK1 (Phosphoinositide-Dependent Kinase-1) and Akt (Protein Kinase B) to the plasmatic membrane and their binding to PI3K, where Akt is activated via phosphorylation by PDK1. Then, the activated Akt phosphorylates many other proteins including fork head box protein

⁷⁴ P. Liu, H. Cheng, T. M. Roberts, J. J. Zhao. *Nat. Rev. Drug Discov.* **2009**, *8*, 627-644.

O1 (FOXO1), BCL2-associated agonist of cell death (BAD) and glycogen synthase kinase 3 β (GSK3 β).

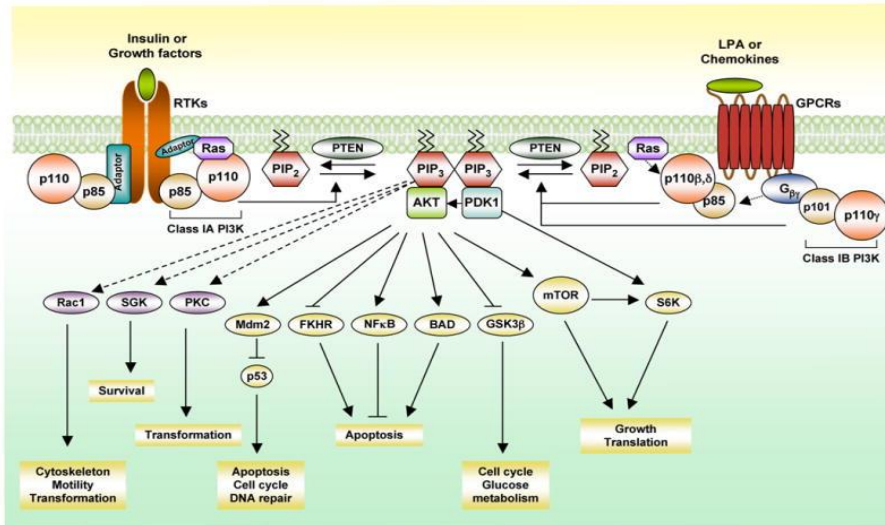


Figure 20. PI3K pathway⁷⁵

In addition, activated Akt can also phosphorylate and activate the mammalian target of rapamycin (m-TOR). Then, m-TOR promotes phosphorylation of eukaryotic translation initiation factor 4E-binding protein 1 (4EBP1) and ribosomal protein S6 kinase beta-1 (S6K1), which are two of the most extensively characterised downstream targets of m-TOR pathway. These two proteins play a crucial role in the regulation of cell growth and proliferation by monitoring nutrient availability, cellular energy levels, oxygen levels and mitogenic signals. Thus, the activation of m-TOR may provide tumour cells with a growth advantage by promoting protein synthesis.⁷⁵

The tumor-suppressor phosphatase with tensin homology (PTEN), is the most important negative regulator of the cell-survival signalling pathway initiated by PI3K. So, when PTEN is mutated or otherwise inactivated, the activation of PI3K effectors and particularly the activation of Akt, can occur in the absence of any exogenous stimulus, initiating tumorigenesis.⁷⁵

⁷⁵ M. Cully, H. You, A. Levine, T. W. Mak. *Nat. Rev. Cancer* **2006**, *6*, 184-192.

- Ral-GEF-Ral pathway

This RAS-related protein (Ral), commonly called Ral-GEFs, shares more than 50% homology with HRAS, KRAS and NRAS. Ral has two isoforms: Ral-A and Ral-B, which are involved in a variety of downstream effectors. They play distinct and key roles in both normal and neoplastic cell physiology including regulation of vesicular trafficking, migration and invasion, tumour formation, metastasis and gene expression. The Ral-GEFs-Ral signalling pathway comprises the third most commonly studied effector of RAS-dependent human oncogenesis. Consequently, Ral GTPases as well as their effectors are being explored as possible therapeutic targets in order to develop cancer treatment focussing on RAS mutation.⁶⁵

1.2.4.2.5. Treatment

In the current therapeutic arsenal, there are no drugs that can directly target KRAS oncogene even though there are some drugs that have been approved by the FDA, which can inhibit KRAS activity by targeting different points of its pathways. Mutated RAS can be inhibited either by directly targeting RAS proteins in order to stop their progression or by targeting the RAS effectors, especially from MAPK and PI3K pathways as illustrated in Figure 21. Nevertheless, due to directly targeting RAS G-domain is being a difficult challenge, the research and development of new drugs are more focused on targeting RAS indirectly.⁷⁶

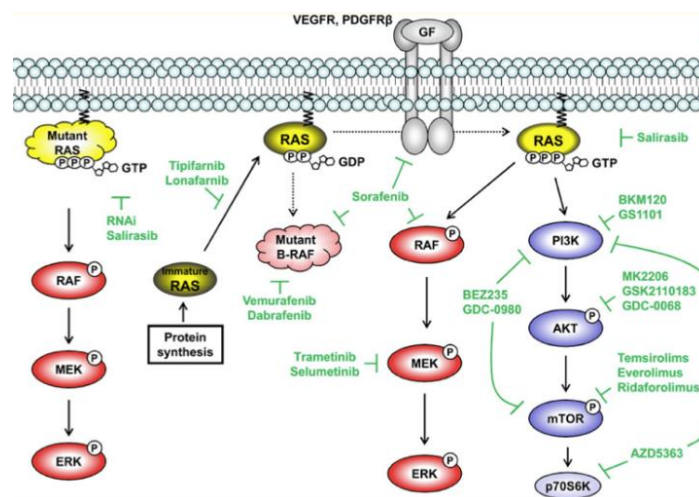


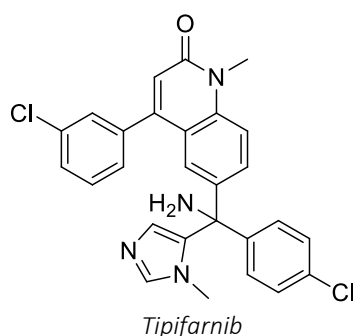
Figure 21. Targeting RAS pathways⁷⁶

⁷⁶ A. Takashima, D. V. Faller. *Expert. Opin. Ther. Targets* **2013**, *17*, 507-531.

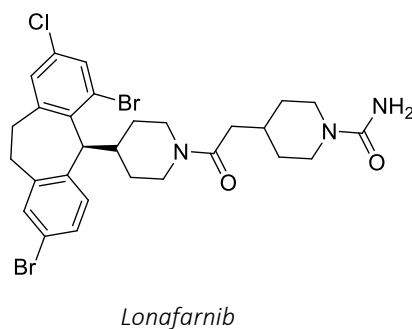
- Farnesyl transferase inhibitors (FTI)

There are two major FTIs, *tipifarnib* and *lonafarnib*, with potential antineoplastic activity by inhibiting RAS farnesylation. Thus, these compounds prevent the activation of RAS oncogenes, concretely, they inhibit cell growth, induce apoptosis, and inhibit angiogenesis. Both of them have shown to be efficient against HRAS and KRAS substrates and are advanced to phase III trials, but with little success so far.

- *Tipifarnib* (**Zarnestra®**): this drug is a quinolone compound, which showed some activity in advanced breast cancer and myelodysplastic syndrome (MDS) in phase II trials. Then, in multiple phase III trials, it did not produce improvements in the overall survival. Regardless of using *tipifarnib* as monotherapy in acute myeloid leukemia (AML) and in refractory advanced colorectal cancer or using it in combination with *gemcitabine* in advanced pancreatic cancer.^{58,77}



- *Lonafarnib* (**Sarasar®**): this compound is a synthetic tricyclic derivative of carboxamide. A phase III trial of *lonafarnib* in combination with *carboplatin/paclitaxel* failed to improve overall survival in advanced or metastatic non-small-cell lung carcinoma (NSCLC).⁵⁸



⁷⁷ S. Gysin, M. Salt, A. Young, F. McCormick. *Genes Cancer* **2011**, 2, 359-372.

- **Salirasib**: this compound is a salicylic acid derivative with potential antineoplastic activity. In contrast to FTIs, *salirasib* acts as a RAS farnesylcysteine mimetic to dislodge all RAS isoforms from their membrane-anchoring sites. Thus, it prevents RAS activation.

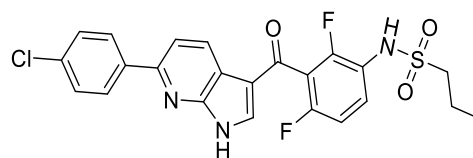
Salirasib was well tolerated in both monotherapy and in combination with *gemcitabine*. In addition, phase I and II trials of *salirasib* as a single agent, and in combination with *gemcitabine* in metastatic pancreatic and lung adenocarcinoma have been completed. However, to determine if there is a significant impact on tumour response and survival, extensive clinical testing will be required.⁷⁸

In addition, statins inhibits the formation of downstream isoprenoids F and GG, which are used as substrates for the enzymes farnesyl transferase and geranyl-geranyl transferase I. Combinations of statins with FTIs have been evaluated in several cancers because they are thought to induce a more effective inhibition of farnesylation.⁷¹

- RAF inhibitors

To date, two RAF inhibitors have been approved by FDA, *sorafenib* and *vemurafenib*. These act as antineoplastic agents by inhibiting RAF kinase, which is a critical component of the RAF/MEK/ERK signalling pathway, and by stopping the cell division and proliferation.⁷⁹

- **Vemurafenib (Zelboraf®)**: this antineoplastic agent is a selective inhibitor of the BRAF (V600E) isoform, which may result in the inhibition of an over-activated MAPK signalling pathway downstream in BRAF (V600E), hence, inducing a reduction in tumour cell proliferation. *Vemurafenib* was approved by FDA in 2011 for the treatment of patients with previously untreated metastatic or unresectable melanoma with the BRAF-V600E mutation.⁸⁰



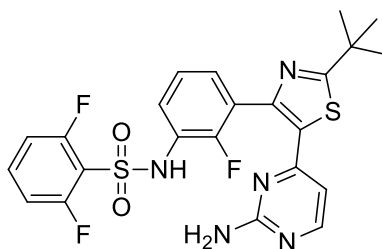
Vemurafenib

⁷⁸ I. R. Vetter, A. Wittinghofer. *Science* **2001**, 294, 1299-1304.

⁷⁹ A. Takashima, D. V. Faller. *Expert Opin. Ther. Targets* **2013**, 17, 507-531.

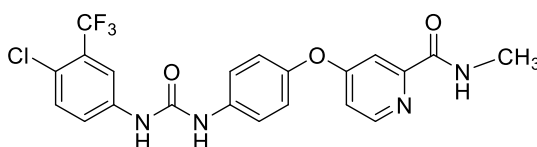
⁸⁰ R. Nazarian, H. Shi, Q. Wang, X. Kong, R. C. Koya, H. Lee, Z. Chen, M. K. Lee, N. Attar, H. Sazegar, T. Chodon, S. F. Nelson, G. McArthur, J. A. Sosman, A. Ribas, R. S. Lo. *Nature* **2010**, 468, 973-977.

- *Dabrafenib (Tafinlar®)*: this antineoplastic agent has a similar preclinical profile to *vemurafenib* but with higher specificity against mutant BRAF. A phase II trial confirmed a 59% response rate to *dabrafenib* in melanoma. Then, several phase III trials are currently on-going. However, it seems that tumour cells can develop resistance to this drug. In order to avoid the resistance of the compound, it has to be combined with *trametinib* which is a MEK inhibitor.⁵⁸



Dabrafenib

- *Sorafenib (Nexavar®)*: is an inhibitor of RAF protein kinase, the vascular endothelial growth factor receptor 2 (VEGFR-2) and the platelet-derived growth factor receptor (PDGFR-β) signalling pathway. Thus, this compound inhibits cell division, cell growth, and angiogenesis. The FDA approved *sorafenib* in 2005 for the treatment of advanced renal and unresectable liver cancer. However, currently *sorafenib* is used in different phases in more than 150 clinical trials in various cancers, either as monotherapy or as combination with other conventional chemotherapy drugs.⁸⁰



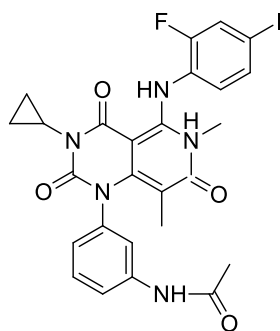
Sorafenib

- MEK inhibitors

In contrast to RAF inhibitors which have a specific target, the mutant BRAF; the MEK kinases are rarely mutated, thereby, do not provide a tumour-specific target. However, MEK inhibitors bind to MEK 1 and/or MEK 2, which play a key role in the activation of the RAS/RAF/MEK/ERK signalling pathway, thereby resulting in an

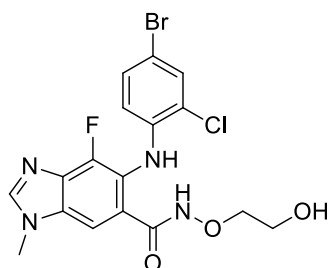
inhibition of growth factor-mediated cell signalling and cellular proliferation in various cancers.⁸¹

- **Trametinib (Mekinist®)**: the most promising investigated MEK inhibitor is a highly-selective allosteric inhibitor of MEK1/2. *In vivo*, the efficacy was also observed in models with activating mutations in BRAF or KRAS. In addition, on-going phase III trials are exploring the safety and efficacy of combination therapy of *trametinib* with the mutant BRAF inhibitor *dabrafenib*.⁸¹



Trametinib

- **Selumetinib**: this compound is an adenosine triphosphate (ATP) independent inhibitor of MEK 1 and 2. Currently, *selumetinib* is undergoing multiple trials in phases I, I/II and II. In phase II trials, *selumetinib* as monotherapy did not show clinical activity in hepatocellular carcinoma (HCC), advanced melanoma, or advanced pancreatic cancer compared to conventional chemotherapies treatment. However, the pharmacodynamic marker carried on the assessment of *selumetinib*, despite the lack of clinical response, suggested this agent could provide additional activity if combined with a BRAF inhibitor.⁸¹

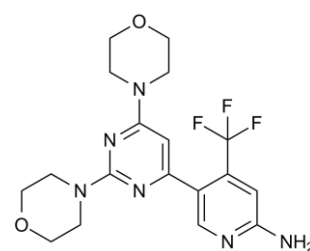


Selumetinib

⁸¹ <http://www.chemdrug.com> (05/05/2016)

- PI3K inhibitors

Currently, there are many investigational drugs undergoing clinical trials, but among the investigational drugs in this class, the most advanced at this time is the oral class I PI3K inhibitor called **BKM120**. This compound inhibits the mutant p85, which generates p110 protein.^{74,75}



BKM120

Surprisingly, cell with mutated p110 was more sensitive to **BKM120** than P110 α wild-type lines. Moreover, on-going phase III trials are being conducted either as a single agent or in combination with *fulvestrant*, in patients with previously treated locally or metastatic breast cancer with positive oestrogen receptor and negative human epidermal growth factor receptor 2 (HER2).^{74,75}

- Akt inhibitors

There are very few Akt inhibitors in human testing. The most advanced is **MK2206** and **AZD5363**. However, the first compound prevents Akt proteins translocation to the plasma membrane and the second compound possesses inhibitory activity against S6K1. These agents are in clinical stages, where it has been seen that they could be used in cancers with PTEN mutations.^{74,75}

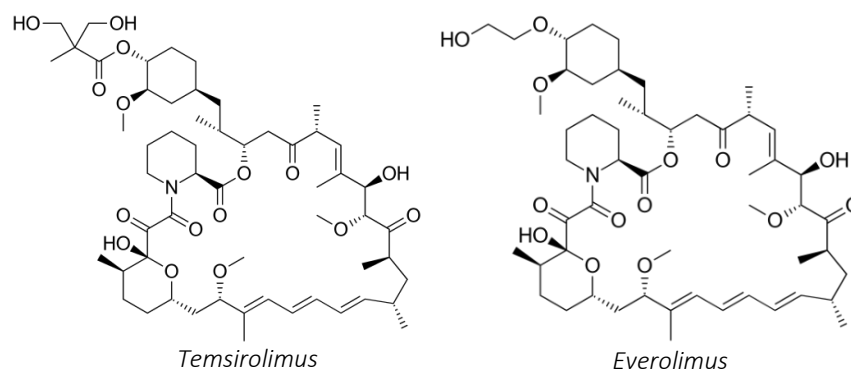
- m-TOR inhibitors

Currently there are different drugs, which are in experimental phases that inhibit m-TOR, but only two m-TOR inhibitors have been approved by FDA, *temsirolimus* and *everolimus*. Those two drugs bind and inhibit a component of mammalian target of rapamycin complex 1 (m-TORC1). This restriction may result in the inhibition of the PI3K pathway and an inhibition in the expression of vascular endothelial cell growth factor (VEGF). Consequently, there is a decrease in tumour cell proliferation and tumour angiogenesis.⁸²

- *Temsirolimus* (**Torisel**[®]): this agent is an ester analogue of *rapamycin*. It is used as renal cell carcinoma (RCC) treatment and in adults for the lymphoma treatment.⁸²

⁸² R. Dienstrmann, J. Rodon, V. Serra, J. Tabernero. *Mol. Cancer Ther.* **2014**, *13*, 1021-1031.

- *Everolimus*: this compound is a derivative of the natural macrocyclic lactone *rapamycin* with immunosuppressive and antineoplastic activities. Thus, it is also used in RCC. Later on, *everolimus* was approved for three more indications: subependymal giant cell astrocytoma, metastatic pancreatic neuroendocrine tumours and in advanced breast cancer.⁸²



After a deeply study of KRAS protein and discovering that there are no treatments that could directly inhibit KRAS,⁸³ different compounds that could directly inhibit the protein were synthesized in this dissertation thesis. Studies of molecular dynamics were performed to find out how the lead compound (**P14**) interacts with KRAS, as well as biological studies of the compounds synthesized were completed in the Hospital Clinic of Barcelona.

1.3. Background

1.3.1. Preparation of new inhibitory agents of PCSK9 protein involved in cholesterol homeostasis

1.3.1.1. Design of new enzyme inhibitors through molecular modeling studies

1.3.1.1.1. Search for new allosteric binding sites (Pockets)

The search for new PCSK9 protein binding sites is of great interest because as seen above an alteration of this protein leads to familial hypercholesterolemia. Currently, the allosteric pockets of the protein have not been described yet in the literature. There are no studies carried out by our research group and, for this motive the present thesis aims to identify the main binding sites of the protein.

⁸³ A. D. Cox, S. W. Fesik, A. C. Kimmelman, J. Luo, C. J. Der. *Nat. Rev. Drug Discov.* **2018**, *11*, 828-851.

It is well known that the PCSK9 binds to the LDL-R which is composed by 3 domains, EFG-(A), EFG-(B) and EFG-(C). The largest interaction by surface is between the catalytic domain of the PCSK9 and the EFG-(A) domain.⁸⁴

For that reason, a possible binding site of the protein is where the EFG-(A) domain binds to the protein. The key interactions between the PCSK9 and the EFG-(A) domain are made up by residues 367-381, being also important ARG194 and ASP238 of the PCSK9^{85,86} and residues ASN295, GLU296, CYS308, ASP310, TYR315 and LEU318 of the EFG-(A) domain.⁸⁷

Another possible binding site of the protein is the catalytic domain of the protein, which is comprised by residues 153-454.

It is intended to find these two binding sites and also more possible allosteric binding sites (Figure 22), since the interactions between the LDL-R and the protein are very strong and, therefore, it would be difficult to displace it.

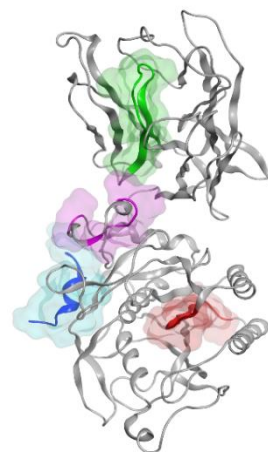


Figure 22. In blue, the EFG-(A) binding site; in red the CD; and, in pink and green two possible allosteric binding sites

1.3.1.1.2. Screening of libraries (Docking)

As previously explained in this thesis, there are two drugs on the market that inhibit the PCSK9 protein, *alirocumab* and *evolocumab*, both monoclonal antibodies. At this time, however, no small organic molecules capable of inhibiting the protein are known. That is why different research groups have conducted docking and PPI studies in order to identify new small organic molecules as inhibitors.

⁸⁴ J. Malo, A. Parajuli, S. W. Walker. *Ann. Clin. Biochem.* **2020**, *1*, 7-25.

⁸⁵ T. Yamamoto, C. Lu, R. O. Ryan. *J. Biol. Chem.* **2011**, 5464-5470.

⁸⁶ S. Xu, S. Luo, Z. Zhu, J. Xu. *Eur. J. Med. Chem.* **2019**, 212-233.

⁸⁷ H.-M. Gu, A. Adijiang, M. Mah, D.-W. Zhang. *J. Lipid Res.* **2013**, *12*, 3345-3357.

Docking studies have been carried out by different research groups and, each used different strategies to obtain small organic molecules as PCSK9 inhibitors. Some of this studies are explained below.

- S. W. Park et al.⁸⁸ considered the EFG-(A) domain of the LDL-R to be the target of the inhibitors. As the EFG-(A) domain interacts with residues 367-381 of the PCSK9 protein, these were used in the docking process. The study consisted of screening an available ChemBrigde Express Collection (450000 chemicals) using the GOLD algorithm and, it was found that a selected chemicals may inhibit the PPI. A maximum of ten docked poses were calculated for each chemical. The top 100 chemicals with the highest chemscore were selected for further calculations. Three-dimensional analogues were obtained as PCSK9 inhibitors, which are decapited in Figure 23.⁸⁸

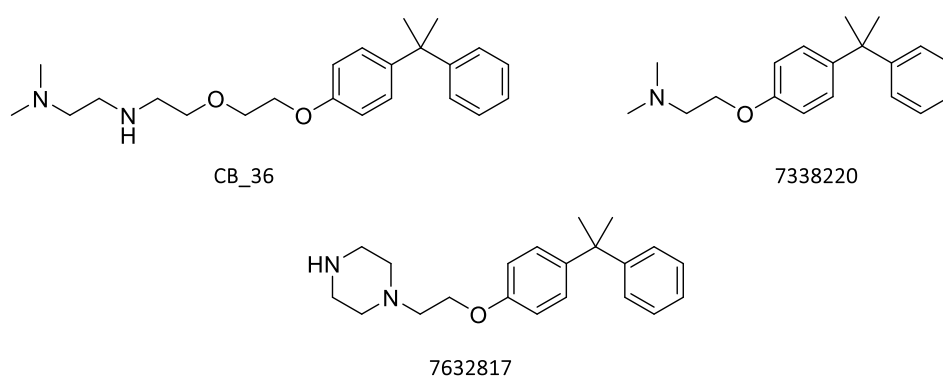


Figure 23. Chemical tructures obtained from the virtual screening⁸⁸

- J. Bonnar et al.⁸⁹ targeted the catalytic domain. It was noted that a groove normally occupied by the C-terminus of the prodomain and defined by residues HIS226, LEU230, LEU253, THR264, LEU286, LEU289, ALA290, LEU297, ASN317 and SER386 may satisfy the requirements to inhibit the PPI. In order to obtain a compound that could inhibit the PCSK9, an *in silico* screen of FDA-approved small molecules collection in BindingDB was deployed.

⁸⁸ D.-K. Min, H.-S. Lee, N. Lee, C. J. Lee, H. J. Song, G. E. Yang, D. Yoon, S. W. Park. *Yonsei Med. J.* **2015**, *5*, 1251-1257. DOI: [10.3349/ymj.2015.56.5.1251](https://doi.org/10.3349/ymj.2015.56.5.1251).

⁸⁹ B. J. Evison, J. T. Palmer, G. Lambert, H. Treutlein, J. Zeng, B. Nativel, K. Chemello, Q. Zhu, J. Wang, Y. Teng, W. Tang, Y. Xu, A. K. Rathis, S. Kumar, A. K. Suchowerska, J. Parmer, I. Dixon, G. E. Kelly, J. Bonnar. *Bioorg. Med. Chem.* **2020**, *28*, 115344. DOI: [10.1016/j.bmc.2020.115344](https://doi.org/10.1016/j.bmc.2020.115344).

Ten theoretical candidates were obtained, in which was included *nilotinib*. The last one seemed to bound within the binding groove of the PCSK9 and the EFG-(A) domain. For this reason, *nilotinib* was then modified in order to increase the PPI. Some *nilotinib* derivatives are represented in the following Figure 24.⁸⁹

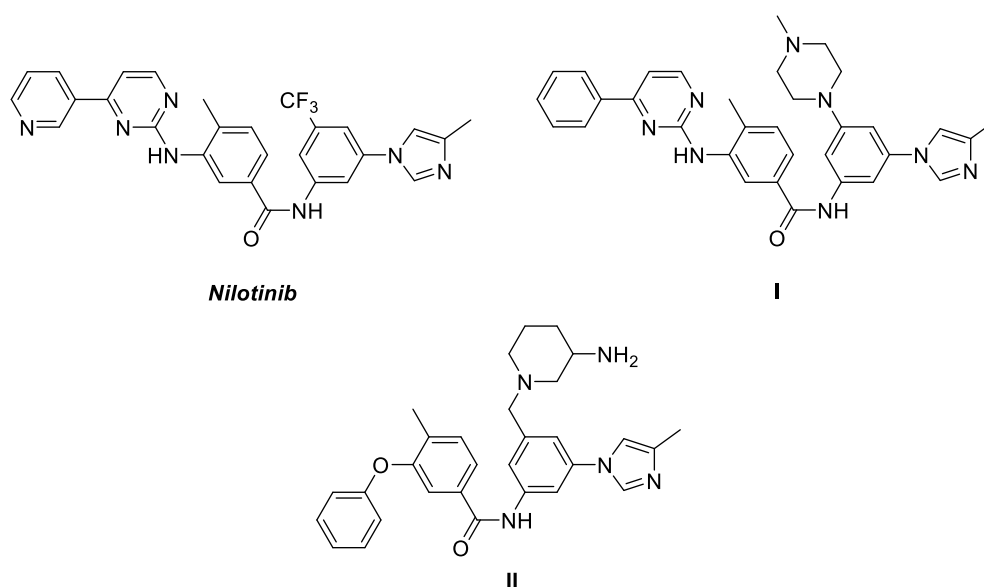


Figure 24. Structures proposed as PCSK9 inhibitors after performing virtual screening⁸⁹

Rational target-based computational approach have been used to identify the main PPI between the EFG-(A) and the PCSK9 in order to synthesize peptidomimetics of the EFG-(A) domain, which in turn are modified in order to increase the established interactions and thus obtain potential PCSK9 inhibitors.

There are several research groups that followed this strategy, some studies of them are described below.

- G. Grazioso et al.⁹⁰ synthesized an *N*-methyl tetraimidazole derivative (**Melm**) as β -strand peptidomimetic of the EFG-(A) domain. The binding mechanism of that imidazole was predicted by molecular dynamics calculations. Then, new imidazole derived compounds were synthesized by introducing new substituents in the imidazole rings (Figure 25). The strategy consisted of introducing new moieties in order to improve the PPI: (1) bearing a *para*-methoxy group on the phenyl ring; (2) showing different chemical properties from those of an aliphatic group of the first imidazole; (3) addition of moieties in the second imidazole ring; (4)

⁹⁰ C. Lammi, J. Sgrignani, A. Arnoldi, G. Lesma, C. Spatti, A. Silvani, G. Grazioso. *J. Med. Chem.* **2019**, *62*, 6163-6174.

substituents to create hydrogen bonds with the third imidazole ring; (5) addition of selected substituents to improve basicity of the amino group.⁹⁰

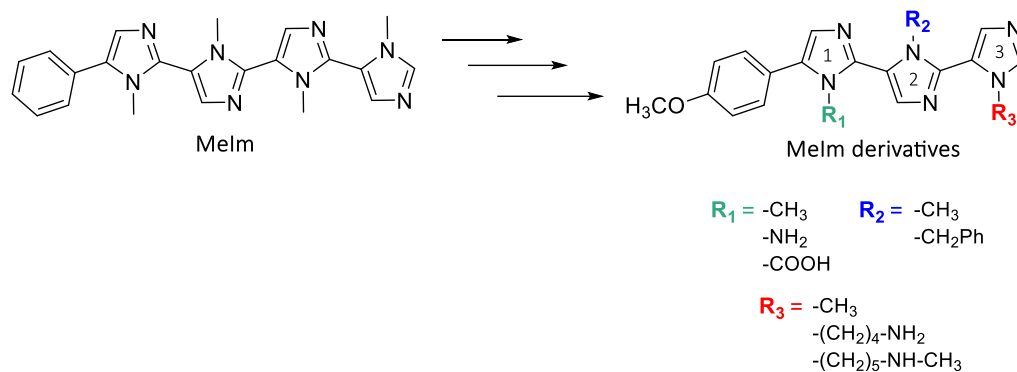


Figure 25. Synthesized imidazole derivatives as peptidomimetics of the EFG-(A)⁹⁰

- K. Burgess et al.⁹¹ built up an EFG-(A) peptidomimetic compound by developing a data mining algorithm (Exploring Key Orientation (EKO)) to find such chemotypes that fit PPI interfaces. They found molecule III (Figure 26) that specifically match to LDL-R. Consequently, they decided to improve the affinity of this compound and docking was then performed. They used compound III as pharmacophore. Finally, three hits were selected as possible inhibitors.^{44,91}

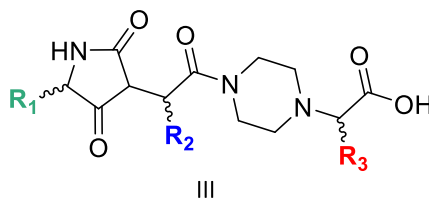


Figure 26. Structure of molecule III^{44,91}

1.3.1.1.3. Rational design of inhibitors based on the structure of the ligands

In the PhD thesis of Vanessa Prieur (2015),⁹² which was developed in our group, a series of pyrrolo[2,3-*d*]pyrimidines were synthesized, while, their biological activity as PCSK9 inhibitors was evaluated by Elli Lilly company.

In this present thesis, the scaffold of these compounds are used to design new fragments as possible PCSK9 inhibitors (Figure 27). It is remarkably to point that these fragments are small organic molecules that combine hydrophobic and polar groups.

⁹¹ J. Taechalestpaisarn, B. Zhao, X. Liang, K. Burgess. *J. Am. Chem. Soc.* **2018**, *140*, 3242-3249.

⁹² V. Prieur. Pyrrolo[2,3-*d*]pyrimidines: Conception, Synthèse, Fonctionnalisation. *Doctoral Thesis*. **December 2015**. University of Barcelona (Spain) and University of Orleans (France).

It is important to mention that said fragments are built up in order to increase its size depending on the pocket that will bind.

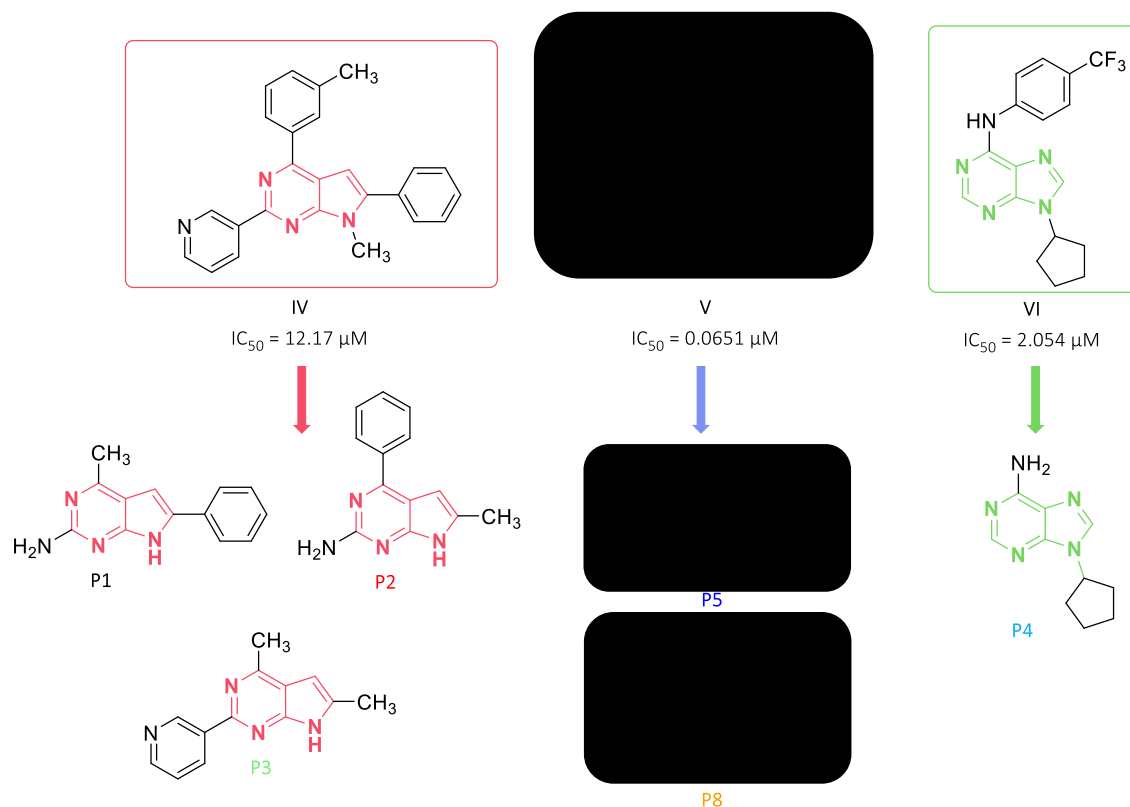


Figure 27. Structures of the fragments designed as possible PCSK9 inhibitors⁹²

1.3.1.2. Preparation of pyrrolo[2,3-d]pyrimidines

Purines are nitrogenized compounds of DNA (adenosine and guanosine bases). These are biochemical components found in the microenvironment of tumors and regulate the immune response as well as cell growth. Most purines are used as antineoplastic agents. The first purine used to treat cancer was 6-mercaptopurine (Figure 28), which was approved in 1953 by the FDA to treat acute lymphoblastic leukemia.⁹³

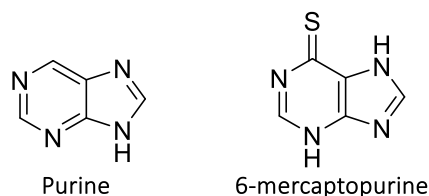


Figure 28. Purine and 6-mercaptopurine structures⁹³

⁹³ R. Vardanyan, V. Hruby. *Synthesis of Best-Seller Drugs*. Academic Press. 1st Edition. 2016, 495-547. ISBN: 9780124115248.

Modifications of such purines leads to different structures such as pyrrolo[2,3-*d*]pyrimidines (Figure 29) which have different therapeutic activities.

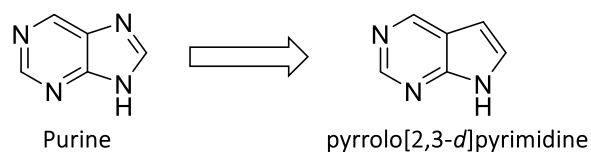


Figure 29. Obtaining of pyrrolo[2,3-*d*]pyrimidines through slight modifications⁹⁴

These compounds exhibit different types of biological activity such as antitumor, anti-tubercular, anti-convulsants, anti-diarrhea, antibacterial, antimicrobial, tyrosine kinase inhibitor, anti-inflammatory and analgesic among others.⁹⁴

As previously mentioned, 3 different series of pyrrolopyrimidines with potential PCSK9 inhibitory activity were prepared in the experimental work of the PhD thesis of Vanessa Prieur (2015).⁹² Modifications in positions 4, 5, 6 or 4, 6 and 2, 4, 6 (Figure 30) were carried out in order to obtain new compounds which biological activity was evaluated by Elly Lilly Company.

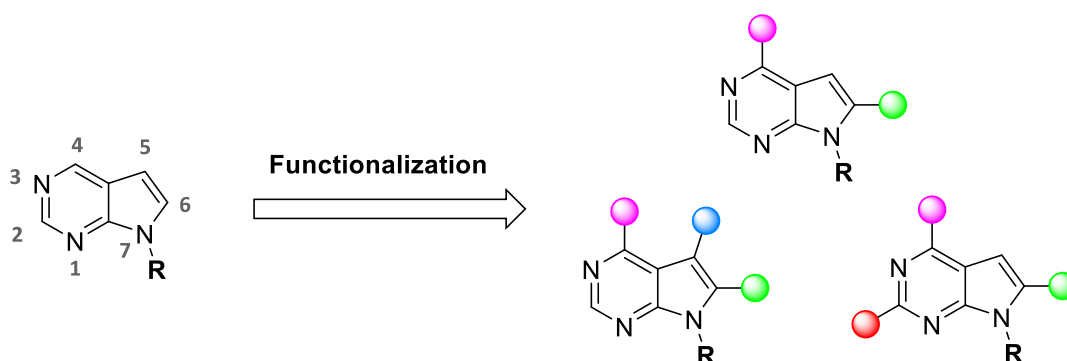


Figure 30. General structures of pyrrolo[2,3-*d*]pyrimidines prepared in Vanessa Prieur doctoral thesis⁹²

The evaluated compounds showed biological activity as antitumor and other types highlighting a new therapeutic activity as PCSK9 inhibitors. Some of the compounds with this therapeutic activity are shown below (Figure 31).

⁹⁴ M. Asif. *Org. Med. Chem.* **2017**, *2*, 555598. DOI: [10.19080/OMCIJ.2017.02.555598](https://doi.org/10.19080/OMCIJ.2017.02.555598).

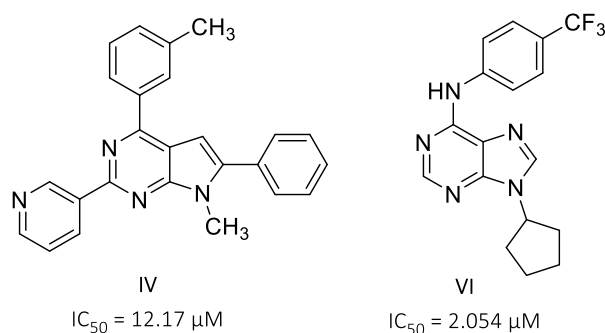


Figure 31. Structures with activity against PCSK9 protein⁹²

In order to continue the development of a new family of pyrrolo[2,3-*d*]pyrimidines with improved activity as PCSK9 inhibitors, the synthesis of several analogues and derivatives are considered as an objective of this present thesis.

1.3.1.3. Preparation of Combretastatin A4 (CA4) derivatives

Combretastatin disodium phosphate (Figure 32) is a natural derivative of the stilbene type from a bark of the South African bush willow tree called *Combretum caffrum*, with activity as a vascular and antineoplastic disruptor. After being administrated, the prodrug is dephosphorylated to CA4, which acts as a microtubule-depolarizing agent producing a mitotic and apoptotic arrest of endothelial cells. In addition, this agent interrupts the association between the endothelial-cadherin vascular molecule to the specific binding of endothelial cells (VE-cadherin) and thus the activity of VE-cadherin/ β -catenin/Akt signalling pathway, producing inhibition of endothelial cell migration and capillary tube formation. Consequently, the vascular tumor collapses, and as a result there is a decrease in blood flow and ischemic necrosis of the tumor tissue.⁹⁵ S. Rafii et al.⁹⁶ proposed a model to inhibit CA4-mediated angiogenesis. Tumor cells form new vessels and CA4 destabilizes these vessels by inducing the separation of VE-cadherin, thereby increasing antineogenic effects through formation of antibodies against VE-cadherin without increasing vascular toxicity.^{96,97}

⁹⁵ L. M. Greene, M. J. Meegan, D. M. Zisterer. *J. Pharmacol. Exp. Ther.* **2015**, 355, 212-222.

⁹⁶ L. Vincent, P. Kermani, L. M. Young, J. Cheng, F. Zhang, K. Shido, G. Lam, H. Bompais-Vincent, Z. Zhu, D. J. Hicklin, P. Bohlen, D. J. Chaplin, C. May, S. Rafii. *J. Clin. Invest.* **2005**, 115, 2992-3006.

⁹⁷ R. Singh, H. Kaur. *Synthesis* **2009**, 15, 2471-2491.

CA4 is a (*Z*)-2-methoxy-5-(3,4,5-trimethoxystyryl)phenol (Figure 32), a simple structure containing two substituted aromatic rings linked by a double bond bridge in the *cis*-configuration from which a large number of analogues and derivatives can be derived.⁹⁸

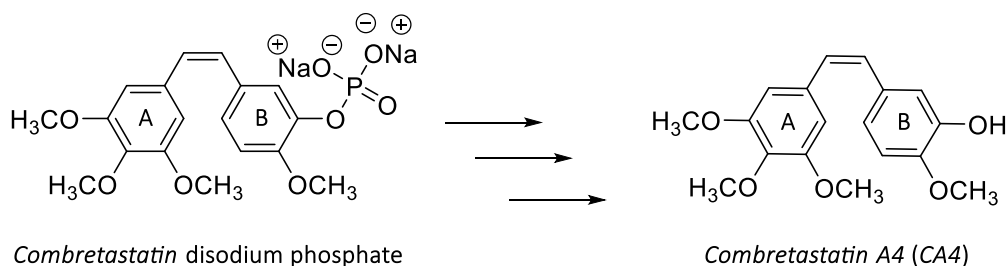


Figure 32. *Combretastatin* disodium phosphate and *Combretastatin A4* (CA4) structures⁹⁸

As shown in the study realized by D. C. Myles et al.⁹⁸ in the **A** ring, the presence of the 3,4,5-trimethoxy groups are essential for its biological activity. However, an increase in the size of the methoxy groups or exchanges of a methoxy for a phenol group leads to a loss of potency. On the other hand, the **B** ring allows to introduce modifications in the hydroxyl position. Nevertheless, removing or replacing the methoxy group lowers the activity severally.⁹⁸ The activity of the *trans*-isomer is much lower than in *cis*-configuration. The IC₅₀ value of CA4 was determined by a wide range of laboratories and ranged from 0.53 to 2.4 nM, whereas *trans*-isomer ranged in the order of μM.⁹⁹

It is well known that CA4 is toxic and that is why analogues and derivatives are synthesized in order to decrease cytotoxicity. In this way, new compounds can be used as anti-tumor agents.

In the PhD thesis of Laura Grau Valls (2017),¹⁰⁰ which was developed in our group, a series of CA4 derivatives were prepared as anti-tumor agents. Different compounds were evaluated by Elly Lilly Company and highlighted the new therapeutic activity discovered for such compounds. Several of them showed activity as PCSK9 inhibitors (Figure 33), opening a window into a new line of research.

⁹⁸ B. R. Hearn, S. J. Shaw, D. C. Myles. *Comprehensive Medicinal Chemistry* Vol. 6. Ed. Elsevier **2007**, 81-110.

⁹⁹ R. Mikstacka, T. Stefanski, J. Rózanski. *Cel. Mol. Biol. Lett.* **2013**, *18*, 368-397.

¹⁰⁰ L. Grau Valls. Disseny, síntesi i evaluació biològica de nous compostos potencialment antitumorals per inhibició enzimàtica. *Doctoral Thesis*. **June 2017**. University of Barcelona (Spain).

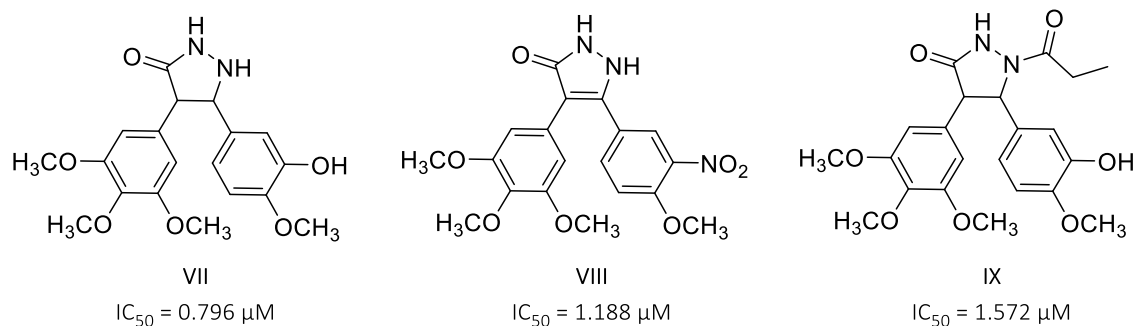


Figure 33. PCSK9 inhibitors synthesized in the doctoral thesis of Laura Grau Valls¹⁰⁰

With this in mind, in this present thesis, a new series of CA4 derivatives are synthesized as potential inhibitors against the PCSK9 protein.

1.3.2. Preparation of new KRAS protein inhibitory agents

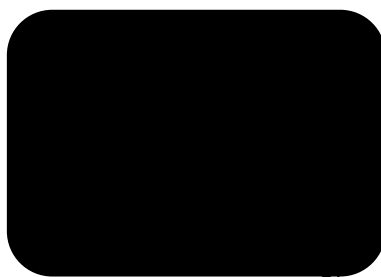
1.3.2.1. Optimization of potential KRAS inhibitors or activators from lead **P14**

1.3.2.1.1. Search of the binding site of the lead P14 by performing gaussian accelerated molecular dynamics simulations (GaMD)

Searching for drugs capable of inhibiting any of the hallmarks of cancer is today around the world one of the main goals of the researchers. That is why, the study of the main targets is of great interest. As explained above there is no drug on the market capable of directly inhibiting KRAS protein. For this reason, in the PhD thesis of Eduardo Garrido (2017),¹⁰¹ which was developed in our group, the exploration for new KRAS inhibitors was performed. The KRAS protein was docked into two different databases, ZINC Lead-Like database and MOE Lead-Like database. The best 100 compounds of each database were selected and a refined docking was performed. When the docking process was completed, the best 200 compounds were selected to carry out MM-GBSA (Molecular Mechanic-Generalized Born Surface Area) calculations. The best 10 compounds were selected for further calculations.

¹⁰¹ E. Garrido. New strategies of inhibition of K-Ras and study of the K-Ras/CaM interaction. *Doctoral Thesis*. September 2017. University of Barcelona (Spain).

Finally, compound **P14** (Figure 34) was selected as promising lead to inhibit KRAS.



P14

Figure 34. P14 structure

The biological activity of **P14** was evaluated by a group of the Department of Biomedicine of the Faculty of Medicine from the Hospital Clinic (Barcelona), in which the **P14** cell viability and the effect of the compound on the signalling pathway of the MAPK was determined.

Due to **P14** is a promising lead, the determination of its protein binding site is of great interest in this present work.

1.3.2.2. Preparation of KRAS inhibitors

As explained above, **P14** compound is a promising lead to inhibit KRAS. For this reason, analogues of **P14** are synthesized through a multicomponent reaction (Described in section 2 and 3 of the present thesis).

2. OBJECTIVES

2.1. Preparation of new inhibitory agents of PCSK9 protein involved in cholesterol homeostasis

2.1.1. Design of new enzyme inhibitors through molecular modeling studies

2.1.1.1. Search for new allosteric binding sites (Pockets)

The main goal is the identification of prospective allosteric binding sites for the PCSK9 protein. It is remarkable to point that only the catalytic and C-terminal domain of the protein were used.

For this purpose a detailed analysis of the flexibility of the protein was carried out by calculating the root-mean square deviation (RMSD), root-mean square fluctuation (RMSF), radius of gyration (Rg) among others types of analysis (Described in section 3 and 4). Conventional and gaussian accelerated dynamics (cMD, GaMD) calculations were performed in order to identify as much as possible the conformational space that the protein could acquire over the time. Finally, different programs were used to discover the most promising allosteric binding sites. In addition a score number was calculated for each new allosteric binding site to qualify its druggability.

2.1.1.2. Screening of libraries (Docking)

The aim of this study is to obtain new small organic molecules with potential inhibitory activity against PCSK9. To accomplish this goal, two different databases (Selleck Drug® and Fragments®) were docked in two different binding sites. The first one corresponded to the place where the LDL-R interacts to the protein and, the second was the catalytic domain of the protein. The structure 4NE9¹⁰² from the protein data bank (PDB) was used for the docking process. This structure corresponds to a dimer each one of them in complex with the EFG-(A5),¹⁰² which is a truncated peptide from the wild-type EFG-(A), which is part of the LDL-R. In this case, the EFG-(A5) has a similar activity to the wild-type EFG-(A). It was also chosen because the protein was almost completed with all residues.

¹⁰² C. I. Schroeder, J. E. Swedberg, J. M. Withka, J. Rosengren, M. Akcan, D. J. Clayton, N. L. Daly, O. Cheneval, K. A. Borzilleri, M. Griffor, I. Stock, B. Colless, P. Walsh, P. Sunderland, A. Reyes, R. Dullea, M. Ammirati, S. Liu, K. F. McClure, M. Tu, S. K. Bhattacharya, S. Liras, D. A. Price, D. J. Craik. *Chem. Biol.* **2014**, *21*, 284-294.

2.1.1.3. Rational design of inhibitors based on the structure of the ligands

To carry out the work shown later in this thesis, the scaffold of the mentioned compounds in Section 1 were used as reference to create new fragments. Modifications of the scaffold were performed in order to include hydrophobic and polar groups that could fit within the pocket. Once, the ligands were built up, the main goal was to find out where the ligands bound into the protein in order to identify new allosteric binding sites of the PCSK9 (no prodomain). This work was carried out through gaussian accelerated and conventional molecular dynamics calculations. Once the binding site was identified new modifications of these ligands were performed in order to improve the main interactions established with the protein. Also, the energy analysis of each new generated ligand was realized by performing MM-GBSA calculations. These $\Delta G_{\text{binding}}$ values were used to select the best modifications and discover new potential PCSK9 inhibitors.

2.1.2. Preparation of pyrrolo[2,3-*d*]pyrimidines

As described above, one of the objectives of this thesis is the synthesis of 7-deazapurines, which were obtained from a modification of the purine nucleus. In this case a bioisosteric change such as a nitrogen by a carbon (N/C) in the pyrrole ring was considered.

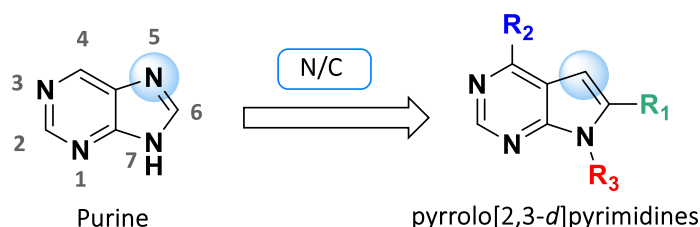


Figure 35. Bioisosteric change from a purine to pyrrolo[2,3-*d*]pyrimidines

From the scaffold of pyrrolo[2,3-*d*]pyrimidines different modifications were studied in order to synthesize chemical compounds, **1-14** (Figure 36, Table 9).

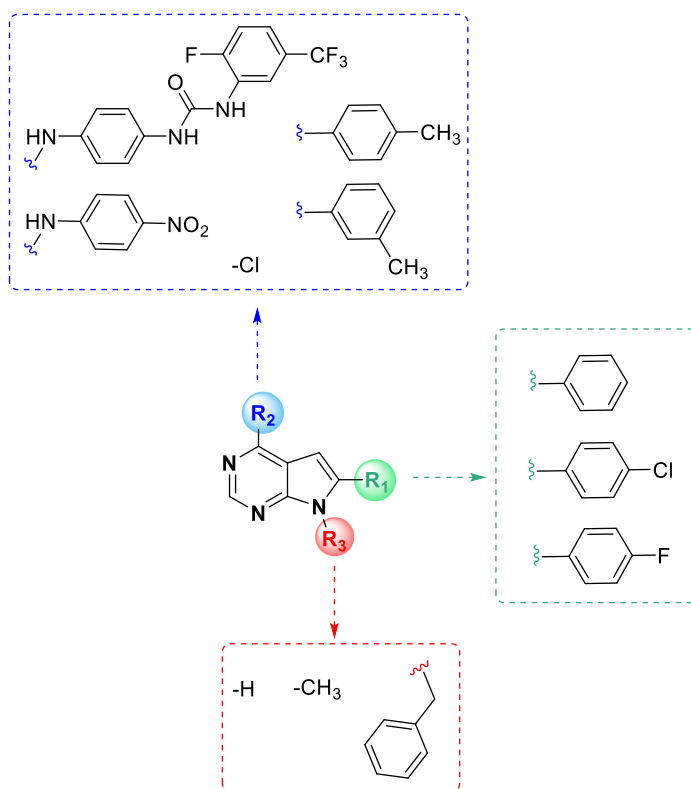


Figure 36. Synthesized bioactive compounds

Table 9. Different structures synthesized according to the type of substituent

Compounds	R ₁	R ₂	R ₃
1	-C ₆ H ₅	-Cl	-H
2	<i>p</i> -ClC ₆ H ₄ -	-Cl	-H
3	<i>p</i> -ClC ₆ H ₄ -	-Cl	-CH ₃
4	<i>p</i> -FC ₆ H ₄ -	-Cl	-CH ₃
5	<i>p</i> -ClC ₆ H ₄ -	<i>p</i> -nitrophenylamino-	-CH ₃
6	<i>p</i> -FC ₆ H ₄ -	<i>p</i> -nitrophenylamino-	-CH ₃
7	<i>p</i> -ClC ₆ H ₄ -	<i>p</i> -aminophenylamino-	-CH ₃
8	<i>p</i> -FC ₆ H ₄ -	<i>p</i> -aminophenylamino-	-CH ₃
9	<i>p</i> -ClC ₆ H ₄ -	-(2-fluoro-5-trifluoro-methylaniline)carbonylamino-4-phenylamino	-CH ₃
10	<i>p</i> -FC ₆ H ₄ -	-(2-fluoro-5-trifluoro-methylaniline)carbonylamino-4-phenylamino	-CH ₃
11	-C ₆ H ₅	-Cl	C ₆ H ₅ -CH ₂ -
12	-C ₆ H ₅	<i>p</i> -tolyl	C ₆ H ₅ -CH ₂ -
13	-C ₆ H ₅	<i>m</i> -tolyl	C ₆ H ₅ -CH ₂ -
14	4-chloro-5-(4-(4-chloro-7H-pyrrolo[2,3-d]pyrimidine-6-yl)phenyl)-6-(4-chlorophenyl)-7H-pyrrolo[2,3-d]pyrimidine		

2.1.3. Preparation of Combretastatin A4 (CA4) derivatives

As explained before in the background section, CA4 by itself is cytotoxic and for this reason, modifications in the structure are considered. In order to maintain the activity of the compounds, in **A** ring no modifications should be performed whereas in **B** ring only the hydroxyl group could be substituted. It is remarkably important to preserve the *cis*-configuration otherwise the compounds will not be active (Figure 37).

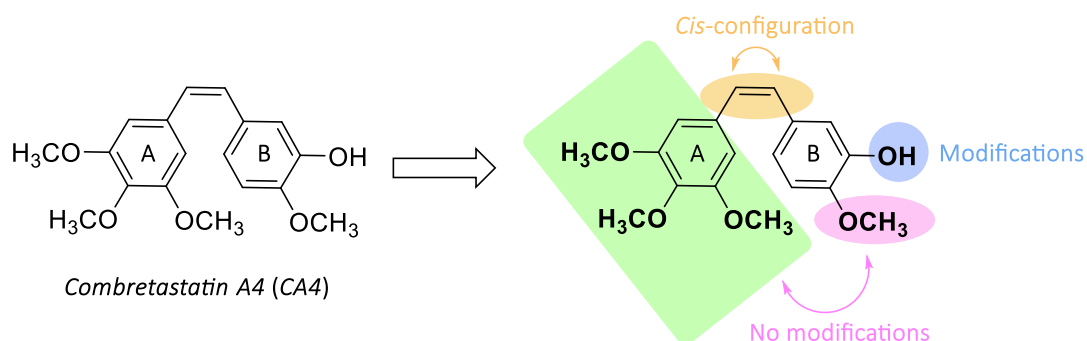


Figure 37. Optimal changes introduced in CA4 in order to maintain the biological activity

The aim of this current thesis is to synthesize structures in which they present a pyrazolidin-3-one ring, in order to obtain compounds with potential PCSK9 inhibitory activity.

To achieve the structure proposed, acrylic acid derivatives and methyl acrylates must be previously synthesized. With this in mind, the following derivatives are proposed (Figure 38, Table 10).

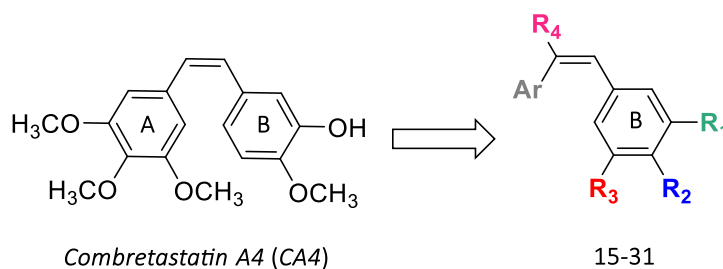


Figure 38. Synthesized acrylic acids and methyl acrylates

Table 10. *Combretastatin A4* derivatives

Compounds	Ar	R ₁	R ₂	R ₃	R ₄
15	3,4,5-trimethoxyphenyl	-OH	-OH	-H	-COOH
16	3,4,5-trimethoxyphenyl	-OH	-OH	-H	-COOCH ₃
17	3,4,5-trimethoxyphenyl	-OCH ₃	-OCH ₃	-OCH ₃	-COOH
18	3,4,5-trimethoxyphenyl	-OCH ₃	-OCH ₃	-OCH ₃	-COOCH ₃
19	3,4,5-trimethoxyphenyl	-NO ₂	-OH	-H	-COOH
20	3,4,5-trimethoxyphenyl	-NO ₂	-OH	-H	-COOCH ₃
21	3,4,5-trimethoxyphenyl	-O-CH ₂ -O-		-H	-COOH
22	3,4,5-trimethoxyphenyl	-O-CH ₂ -O-		-H	-COOCH ₃
23	phenyl	-OH	-OCH ₃	-H	-COOH
24	phenyl	-OH	-OCH ₃	-H	-COOCH ₃
25	2,2-dimethylpropionylphenyl	-OH	-OCH ₃	-H	-COOH
26	methylsulfonylphenyl	-OH	-OCH ₃	-H	-COOH
27	methylsulfonylphenyl	-OH	-OCH ₃	-H	-COOCH ₃
28	2-(2,6-dichlorophenylamino)phenyl	-OH	-OCH ₃	-H	-COOH
29	2-(2,6-dichlorophenylamino)phenyl	-OH	-OCH ₃	-H	-COOCH ₃
30	3-indolyl	-OH	-OCH ₃	-H	-COOH
31	3-indolyl	-OH	-OCH ₃	-H	-COOCH ₃

Modulative modifications were introduced into the methyl acrylate derivative in order to obtain the proposed pyrazolidin-3-one structures (**32-38**). However, compounds with a completely different nucleus were also synthesized to spot whether they potentiated the activity as PCSK9 inhibitors (**39-40**).

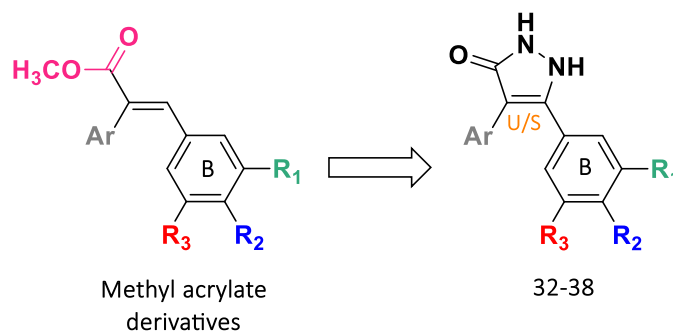
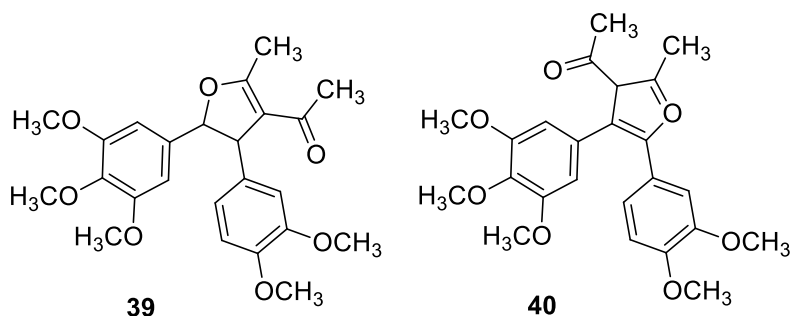


Figure 39. *Cyclic Combretastatin A4* derivatives

Table 11. Synthesized pirazolidin-3-one compounds

Compounds	Ar	R ₁	R ₂	R ₃	U/S
32	3,4,5-trimethoxyphenyl	-OH	-OH	-H	saturated
33	3,4,5-trimethoxyphenyl	-OCH ₃	-OCH ₃	-OCH ₃	saturated
34	3,4,5-trimethoxyphenyl	-NO ₂	-OH	-H	unsaturated
35	3,4,5-trimethoxyphenyl	-O-CH ₂ -O-		-H	unsaturated
36	phenyl	-OH	-OCH ₃	-H	saturated
37	methylsulfonylphenyl	-OH	-OCH ₃	-H	unsaturated
38	3-indolyl	-OH	-OCH ₃	-H	saturated



2.2. Preparation of new KRAS protein inhibitory agents

2.2.1. Optimization of potential KRAS inhibitors or activators from lead **P14**

2.2.1.1. Search of the binding site of the lead **P14** by performing gaussian accelerated molecular dynamics simulations (GaMD)

In this present molecular modeling study of KRAS protein, the main goal is to describe the interactions established by **P14** and KRAS protein. To carry out this work two different approximations to molecular dynamics were used, specifically, gaussian accelerated and conventional molecular dynamics. With this in mind, the characterization and description of the **P14** binding site was crucial to achieve the objective proposed.

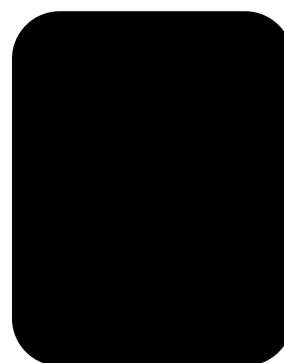


Figure 40. Structure of **P14**

2.2.2. Preparation of KRAS inhibitors

As explained earlier in this work, the lead **P14** is a potent KRAS inhibitor. For this reason, the synthesis of different analogues of **P14** is proposed, with the aim of obtaining new and more selective compounds for the treatment of colorectal cancer (Figure 41).

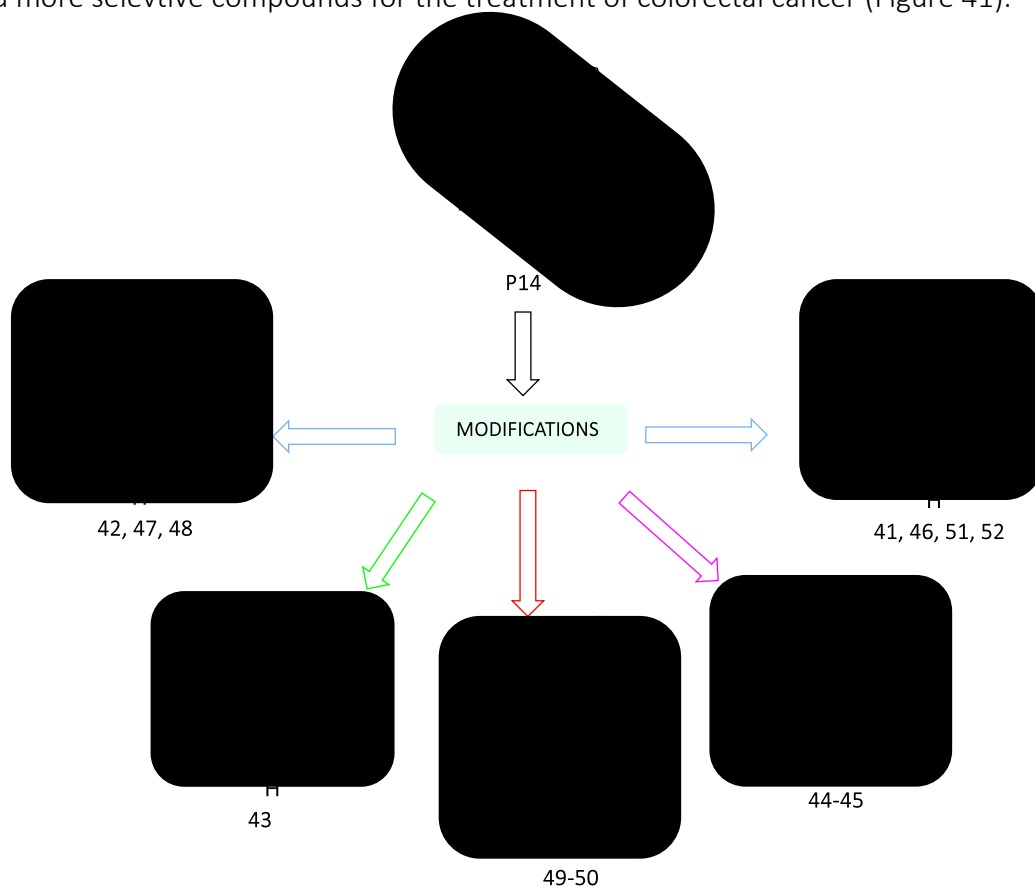


Figure 41. P14 derivatives

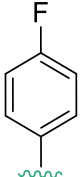
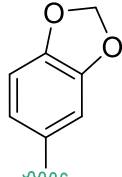
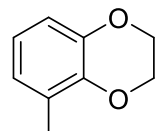
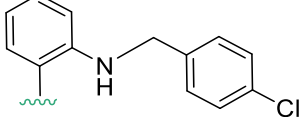
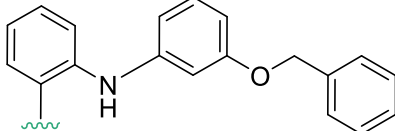
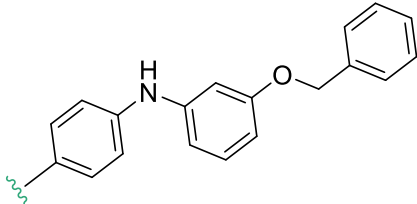
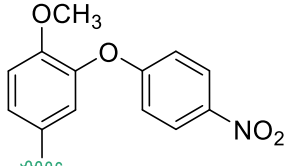
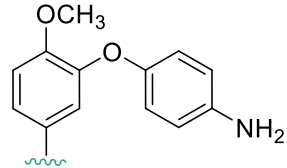
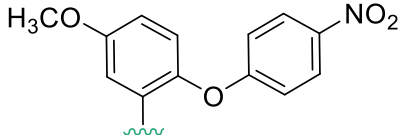
As it can be seen in Figure 41 the structure is comprised by a polar part, where the hydroxyl groups are found; and, by a hydrophobic part where the benzyl group is located, which is largely replaced.

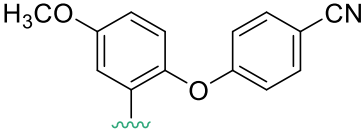
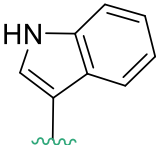
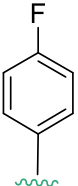
In the following Table 12 the synthesized analogues are indicated.



Figure 42. Synthesized structures

Table 12. Structural P14 derivatives

Compounds	R ₁	R ₂
41		-H
42		-H
43		-H
44		-H
45		-H
46		-H
47		-H
48		-H
49		-H

50	 <p>Chemical structure of 4-(4-methoxyphenyl)benzophenone. The structure shows a central carbonyl group (C=O) bonded to two phenyl rings. The left phenyl ring has a methoxy group (H₃CO) at the para position. The right phenyl ring has a cyano group (CN) at the para position. A wavy line is drawn below the carbonyl carbon atom.</p>	-H
51	 <p>Chemical structure of indole. It consists of a benzene ring fused to a pyrrole ring. A wavy line is drawn below the carbon atom at the 3-position of the pyrrole ring.</p>	-H
52	 <p>Chemical structure of a 4-fluorophenyl group. It consists of a benzene ring with a fluorine atom (F) at the para position. A wavy line is drawn below the carbon atom at the other para position.</p>	-CH ₃

Once the compounds with a sufficient degree of purity within these objectives have been prepared, biological tests are performed in order to know qualitatively and quantitatively the biological activity and to establish the corresponding structural requirements.

3. THEORETICAL DISCUSSION

3.1. Preparation of new inhibitory agents of PCSK9 protein involved in cholesterol homeostasis

3.1.1. Design of new enzyme inhibitors through molecular modeling studies

3.1.1.1. Search for new allosteric binding sites (Pockets)

For the four systems, differentiated each one by their initial temperature (molecular dynamic calculation (dyn)), the analysis of the flexibility of the PCSK9 protein was carried out using 15,000 structures from the 300 ns of GaMD calculation. The RMSF of the C α (Carbon Alpha) residues of the protein were determined as shown in Figure 43.

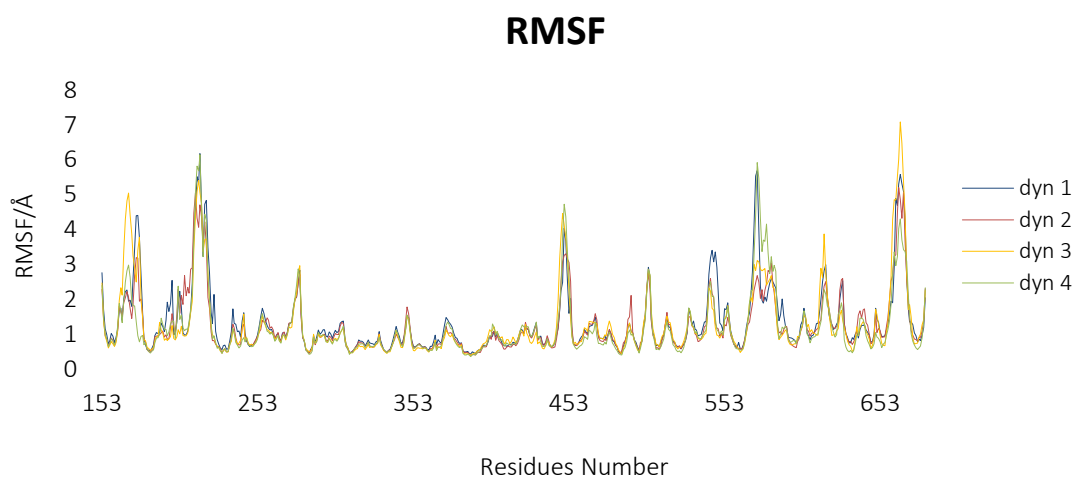


Figure 43. RMSF of all residues of the protein

As shown pictorially in Figure 44 the main fluctuations correspond to the loops residues of the protein pictured in color. On the other hand, the invariant nucleus of the protein is depicted in gray.

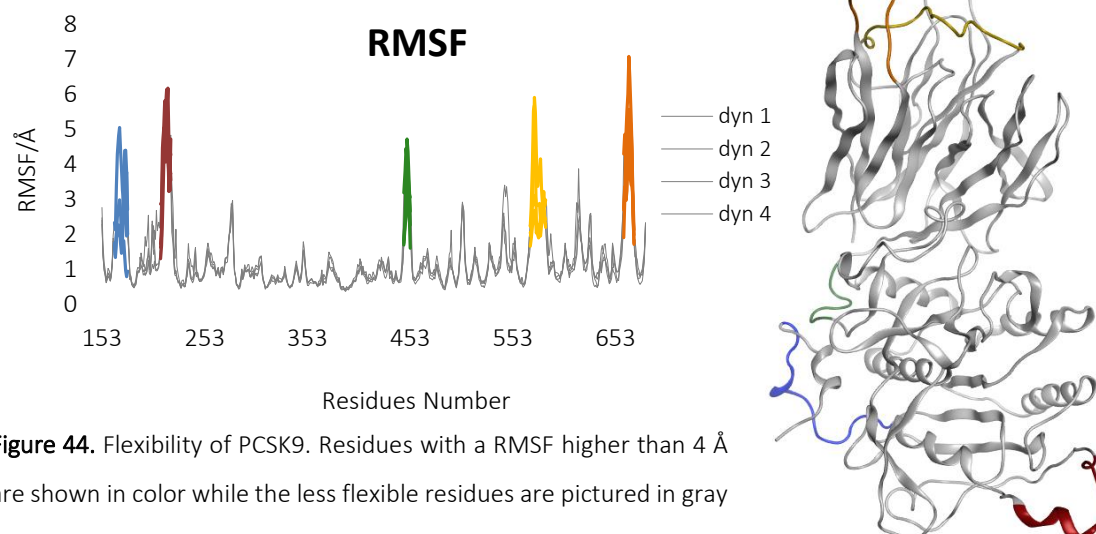


Figure 44. Flexibility of PCSK9. Residues with a RMSF higher than 4 Å are shown in color while the less flexible residues are pictured in gray

In order to measure the similarity between structures, the RMSD was calculated for each system (molecular dynamic calculation). The starting structure was superimposed with all the structures from the 300 ns of GaMD calculation. In this case, the residues of the C α were superimposed for each system, in which the residues picked were different.

For system 1 (dyn 1) the residues selected were the C α of 154-166, 179-194, 196-197, 199-202, 204-210, 223-224, 226-276, 281-447, 453-502, 506-542, 550-559, 577-580, 587-589, 591-615, 619-627, 630-661, 671-681.

For system 2 (dyn 2) the residues taken were the C α of 154-167, 171-173, 179-201, 222-277, 281-447, 455-502, 506-542, 546-571, 586-616, 669-626, 640-659, 671-681.

For system 3 (dyn 3) the residues picked out were the C α of 154-164, 179-211, 222-276, 281-446, 454-502, 506-542, 545-571, 587-614, 619-659, 671-681.

And finally, for system 4 (dyn 4) the residues elected were the C α of 154-167, 172-201, 203-210, 222-276, 281-447, 454-502, 506-542, 545-569, 587-616, 619-660, 672-681.

As shown in Figure 45, system 1 and 2 were similar to each other, as were 3 and 4. However, systems 1 and 2 were slightly different from 3 and 4. Therefore, the lower the RMSD, the better the model is compared to the target structure, in this case system 4.

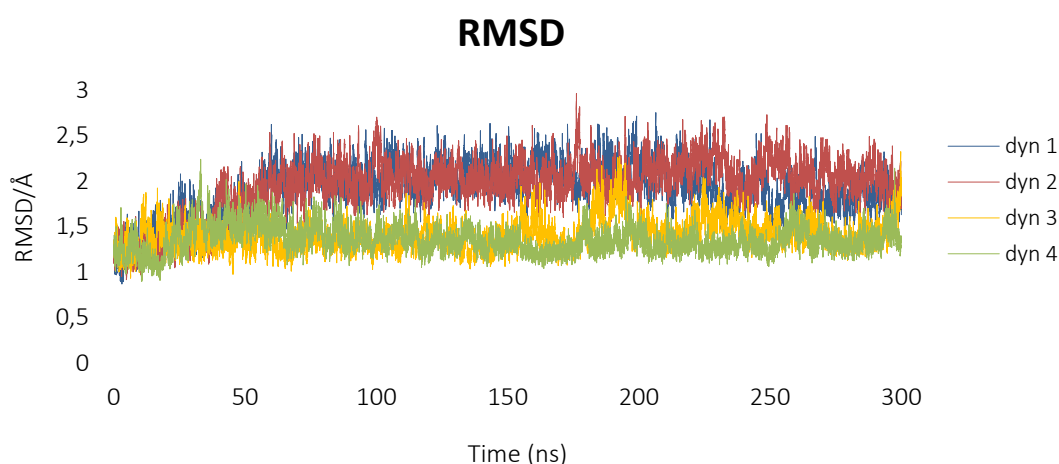


Figure 45. The RMSD values as function of the simulation time

The compactness of a protein has a direct relationship with the rate of folding, which can be studied with GaMD simulation as a computational method to calculate the radius of gyration. It is also helpful to study the static of the protein and its conformational state.

The radius of gyration (Figure 46) was calculated for each system after 300 ns of gaussian accelerated molecular dynamics calculation. As can be seen from Figure 45 the radius of gyration for the different systems has the same behavior as the RMSD.

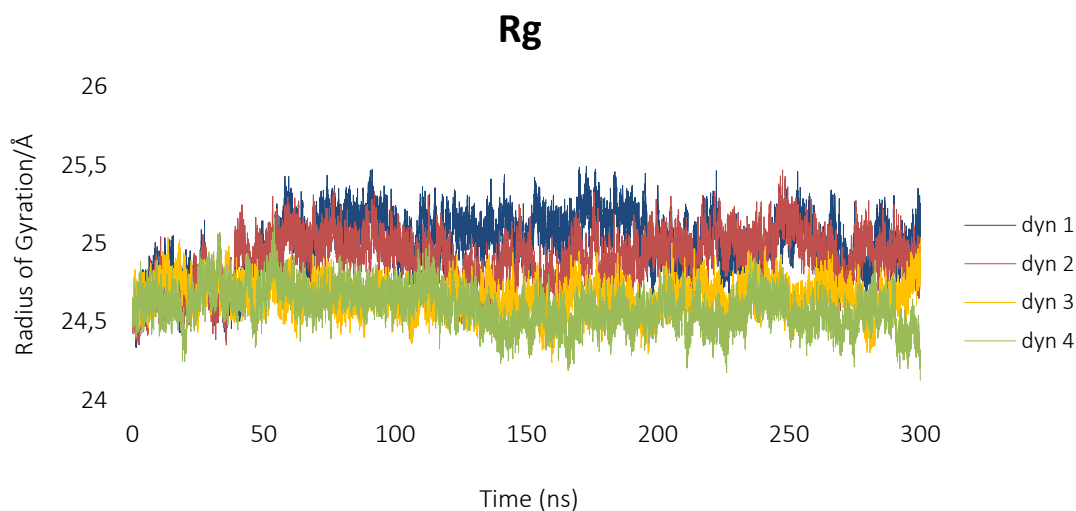


Figure 46. Radius of gyration after 300 ns of accelerated molecular dynamics calculation

After structural analysis of the PCSK9 protein, clustering was performed. At first, the 4 GaMD were joined, thus obtaining a 1200 ns of GaMD simulation. The clustering was done with 30,000 frames of the 60,000 contained in the 1200 ns length of GaMD for computational reasons. Secondly, the clustering was performed using the $C\alpha$ atoms of the secondary structure of the protein. Such residues were 156-160, 168-170, 181-185, 215-220, 225-235, 246-251, 261-277, 283-287, 291-292, 295-308, 311-315, 323-327, 335-340, 347-348, 351-352, 361-364, 368-371, 380-382, 386-402, 408-417, 456-462, 472-475, 485-489, 497-503, 507-512, 521-528, 533-539, 561-565, 595-617, 621-666, 634-638, 646-649, 653-658, 671-680.

As a result, 10 clusters were obtained, being cluster 0 the most populated with 21% of structures; for cluster 1 the population was 20%; the percentage of population for cluster 2 was 17; cluster 3 populated with 15% while cluster 4 with less population had 9%.

The conformation of the protein for each cluster was slightly different. The clustering allows to identify new conformational structures, being the least populated in time the structure most likely to be different and new, while the most populated correspond to the standard conformation. To help selecting the most important representative conformations a Principal Component Analysis (PCA) was carried out.

First, a superposition was made between a selected reference structure and the structures obtained from the four trajectories of GaMD, the same C α atoms used in clustering, were superposed. This process enabled the recomputation of atom fluctuations to build a covariance matrix that was diagonalized in order to obtain its principal component and the eigenvalues, focusing on the differences in structures between each of the captured states.

Figure 47 shows the projection of 60,000 snapshots defined by PC1 and PC2 after reweighting. The PCA analysis identified five different clusters, two of which had low free energy. The lowest minimum was located at coordinates (-3,-11), the second one at (11,-5). A third minimum was found at coordinates (-8,0), and the fourth and fifth at (-12,-1) and (5,12) respectively.

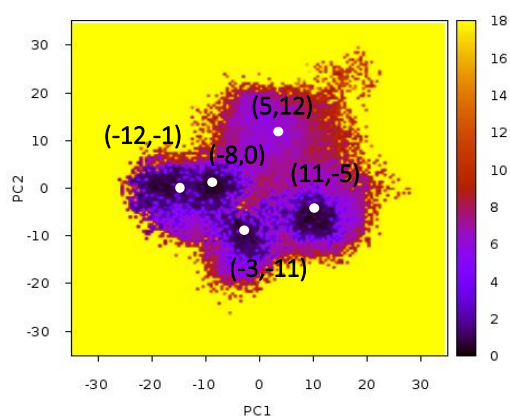


Figure 47. Free energy map calculated from the projection of the structures extracted from the GaMD of the main components PC1 and PC2

The comparison between PCA and the clustering permitted to obtain the 5 main conformations of the protein over the 1200 ns of GaMD. The coordinates obtained from the PCA corresponded to the centroids of the clusters, where a centroid is the point with the shortest distance to all other points belonging to that cluster. In this way, it was possible to compare both results and relate each minimum with the corresponding cluster (Figure 48).

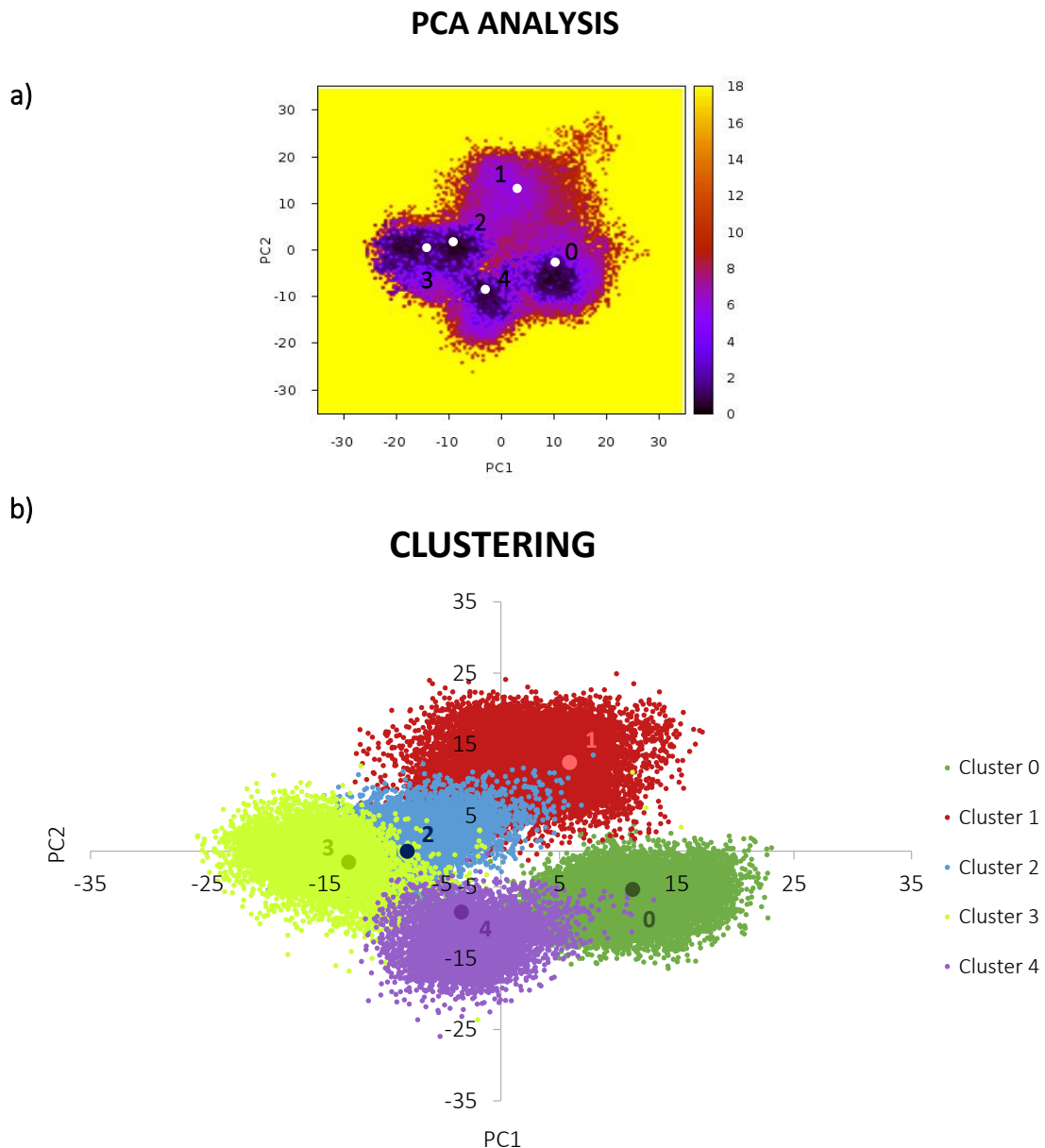


Figure 48. Comparison between PCA and clustering: **a)** Representation of PCA analysis; **b)** Clusters obtained from the clustering process

The lowest minimum with PCA coordinates $(-3,-11)$ permitted to identify cluster 4 in which the coordinates of the centroids were the same achieved in the PCA analysis, corresponding to $(-3,-11)$. The same procedure was used to identify the other clusters. In this case, the second minimum with coordinates $(11,-5)$ corresponded to cluster 0, which was the most populated. The other minimums were together, located at coordinates $(-8,0)$ corresponding to cluster 2, and coordinates $(-12,-1)$ matching to cluster 3. The cluster with less energy was located at $(5,12)$ resulting in cluster 1 (Figure 48).

The main differences between clusters were in the most variable loops while the α -helices and β -sheets were structurally very similar. Once the major conformations were established, the identification of prospective allosteric binding sites was carried out.

Surprisingly, the identification of the pockets of PCSK9 protein have not been described yet in the literature. For this reason, it is of great interest targeting the PCSK9 as well as to look for new druggable binding sites. With this purpose in mind, the main goal was to describe all possible binding sites (pockets) of the protein using three different methods, such as FTMap,^{103,104} Deepsite¹⁰⁵ and Pockdrug^{106,107,108,109,110}.

The representative clusters used to identify the pockets with all three programs were those obtained from the clustering process and the PCA analysis.

These clusters presented the most populated structural conformations during the GaMD simulations. The clusters used were the following in decreasing order of population: cluster 0, cluster 1, cluster 2, cluster 3 and cluster 4.

The FTMap program define the regions that bind several different probes (small organic molecules) as consensus sites (CSs), and the site containing the largest number of probes is considered the main hot spot. All other consensus sites are secondary hot spots. The most populated cluster (cluster 0) showed six different possible allosteric binding sites. The pocket with the maximum number of probes was described as P1 where residues found within a distance of 4.5 Å of said probes were LEU185, ASP186, VAL252, LEU253, GLY259, THR260, VAL261, THR264, LEU287, PRO288, LEU289, ALA290, GLY291, LEU297, SER326 and PRO327. This pocket features a catalytic triad and, therefore, could be a druggable site. The second hot spot was described as pocket P6 where 13 probes were identified near residues GLY356, ARG357, CYS358, VAL359, ASP360, LEU361, ILE416,

¹⁰³ D. Kozakoz, E. L. Grove, D. R. Hall, T. Bohnuud, S. Mottarella, L. Luo, B. Xia, D. Beglov, S. Vajda. *Nat. Protoc.* **2015**, *10*, 733-755.

¹⁰⁴ C. H. Ngan, T. Bohnuud, S. E. Mottarella, D. Beglov, E. A. Villar, D. R. Hall, D. Kozakov, S. Vajda. *Nucleic Acids Res.* **2012**, 271-275.

¹⁰⁵ J. Jimenez, S. Doerr, G. Martínez-Rosell, A. S. Rose, G. de Fabritiis. *Bioinformatics* **2017**, *19*, 3036-3042.

¹⁰⁶ H. A. Mussein, A. Borrel, C. Geneix, M. Petitjean, L. Regad, A.-C. Camproux. *Nucleic Acids Res.* **2015**, *43*, 436-442.

¹⁰⁷ A. Borrel, L. Regad, H. G. Xhaard, M. Petitjean, A.-C. Camproux. *J. Chem. Inf. Model.* **2015**, *55*, 882-895.

¹⁰⁸ T. Krotzky, T. T. Rickmeyer, T. Fober, G. Klebe. *J. Chem. Inf. Model.* **2014**, *54*, 3229-3237.

¹⁰⁹ A. Krasowski, D. Muthas, A. Sarkar, S. Schmitt, R. Brenk. *J. Chem. Inf. Model.* **2011**, *51*, 2829-2842.

¹¹⁰ P. Schmidtke, X. Barril. *J. Med. Chem.* **2010**, *53*, 5858-5867.

THR417, PRO418, LEU420, ARG458, THR459, VAL460 and TRP461. Another consensus site described was included within pocket P6 and contained 11 probes placed near residues ALA328, SER329, ALA330, PRO331, GLU332, VAL333, ILE334, ARG357, CYS358, VAL359, ASP360, ARG412, VAL460 and TRP461. Next cluster was described as P6 pocket and the probes were located near residues VAL435, LEU436, THR437, PRO438, ASN439, THR459, TYR648, ALA649 and VAL650. The third pocket, described as P18, was located near residues CYS562, ARG582, GLN587, CYS588, VAL589, VAL590, GLU592, SER636, ALA637, LEU638, PRO639, THR673 and ILE675, and had 10 probes. The fourth pocket reported was P14, with 9 probes which were placed near residues HIE449, ALA451, GLY452, TRP453, GLN454, LEU455, PHE456, LEU528, LEU529 and PRO530. The following 2 clusters reported contained 9 probes and formed pocket P13 which, in turn, was comprised by residues SER188, ILE189, VAL200, MET201, VAL202, THR203, ASP204, PHE205, GLU206, ASN207, MET247, ARG248, SER249, LEU250 and ARG251. Finally, the last pocket obtained consisted of 2 probes and was described as pocket P12. These last 2 probes were near residues THR187, LYS222, CYS223, ASP224, SER225, HID226, GLY227, ASN256, GLN258 and GLY259.

The same type of procedure was performed with clusters 1, 2, 3, and 4. In cluster 1, three more pockets were discovered: P5, P11 and P15. Pocket P10 was identified in cluster 2, whereas in cluster 3 no new pockets were found. However, in cluster 4 two new pockets were obtained, named as pocket P2 and P3. Pocket P3 was located at the same position where the EFG-(A) domain is found.

All pockets discovered through the FTMap program are represented in Figure 49.

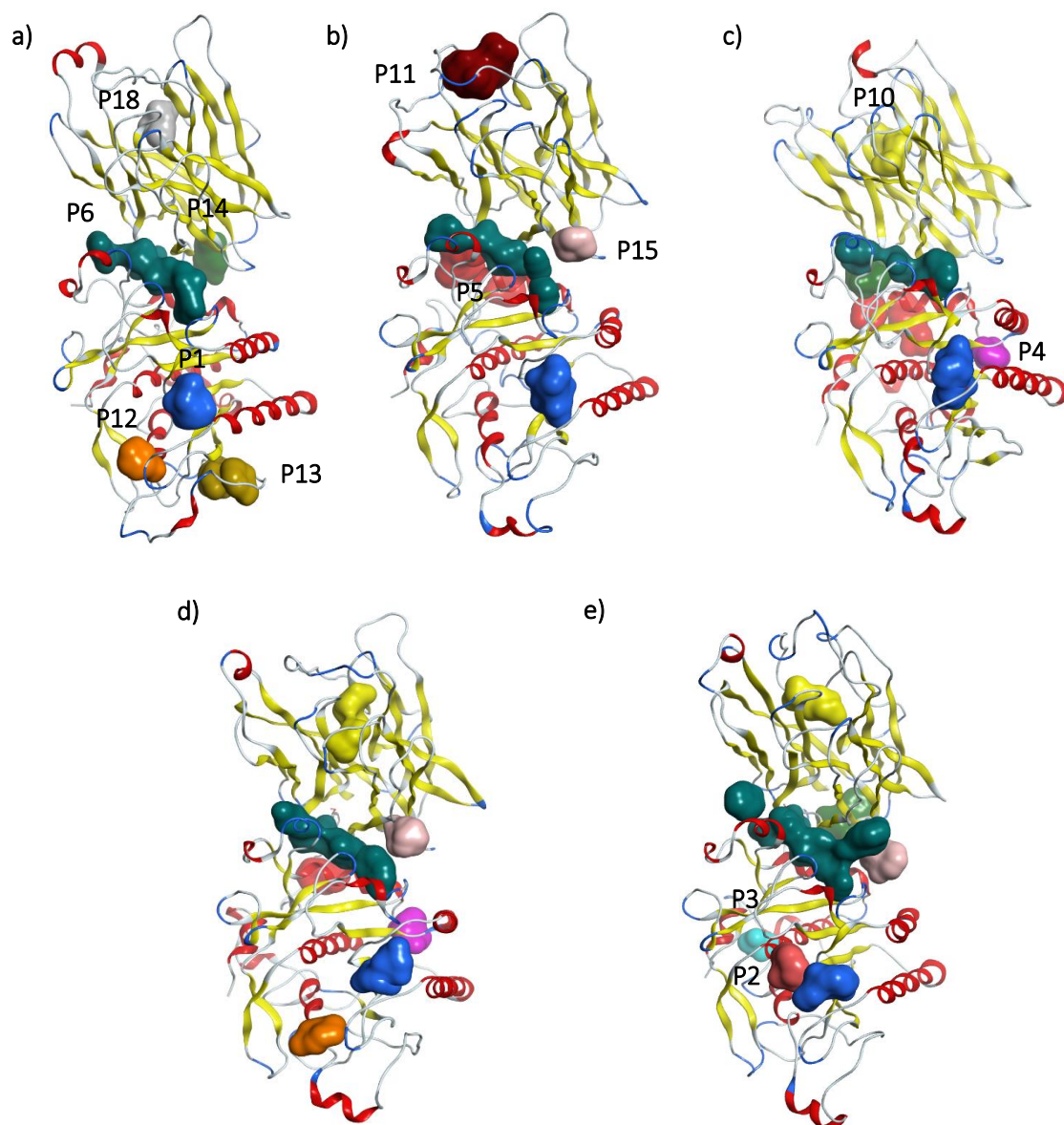


Figure 49. FTMap. **a)** Pockets of cluster 0; **b)** Pockets of cluster 1; **c)** Pockets of cluster 2; **d)** Pockets of cluster 3; **e)** Pockets of cluster 4

A further step was to determine new pockets by using the DeepSite program which only allows obtaining the three most important hot spots. The pockets discovered will be described only for cluster 0, as it was the most populated cluster.

The pockets achieved were **P1**, **P3** and **P6**. In this case, **P1** was made up of the residues LEU185, ASP186, VAL252 and LEU253. The residues of **pocket 3** were ILE154, LEU158 and ILE161. Finally, **pocket P6** showed the residues PRO331, GLU332, VAL333, ILE334 and THR335. On the other hand, the pockets found in cluster 1 were **P1**, **P6** and **P11**. Pockets **4**, **5** and **12** were placed in cluster 2; in cluster 3 the pockets obtained were **P1** and **P6**;

and finally, for cluster 4 the pockets discovered were P1, P6 and P18. All these pockets are represented in Figure 50.

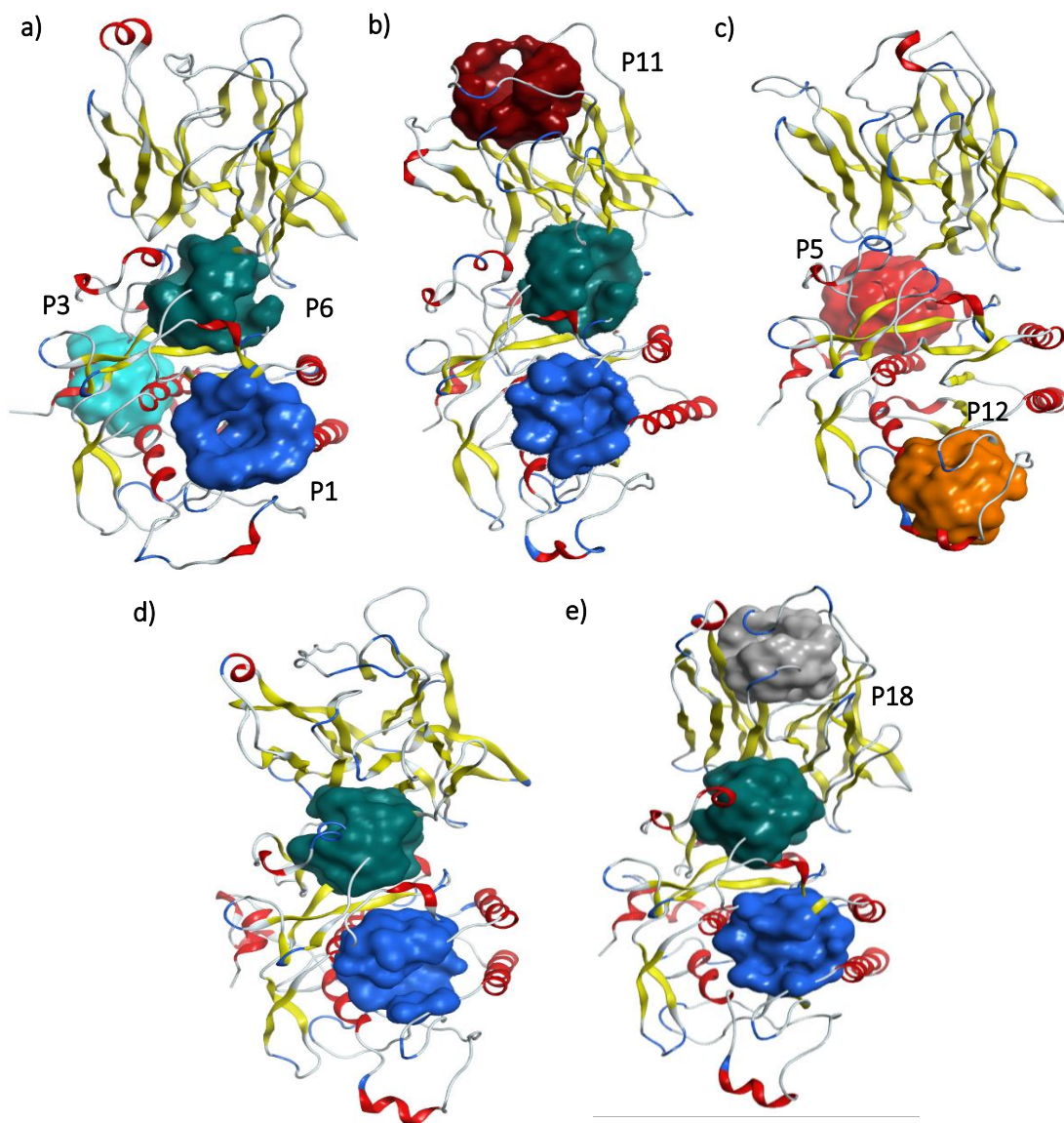


Figure 50. DeepSite. **a)** Pockets of cluster 0; **b)** Pockets of cluster 1; **c)** Pockets of cluster 2; **d)** Pockets of cluster 3; **e)** Pockets of cluster 4

A new program was used to confirm the results obtained so far. This time, the PockDrug program gave similar results to those detailed above.

The pockets found were the same as those described above and no new ones were discovered. Thus, in cluster 0 the pockets P1/P3/P4/P5/P6/P10/P13/P14/P18 were observed; cluster 1 showed pockets P1/P3/P4/P6/P10/P13/P14/P15/P18; pockets obtained in cluster 2 were P1/P3/P4/P5/P6/P10/P11/P15/P18; and for cluster 3, P1/P3/P4/P5/P6/P10/P11 were reported. Finally, fewer pockets were found in cluster 4, such as P2/P6/P10/P15/P18 as can be seen in Figure 51.

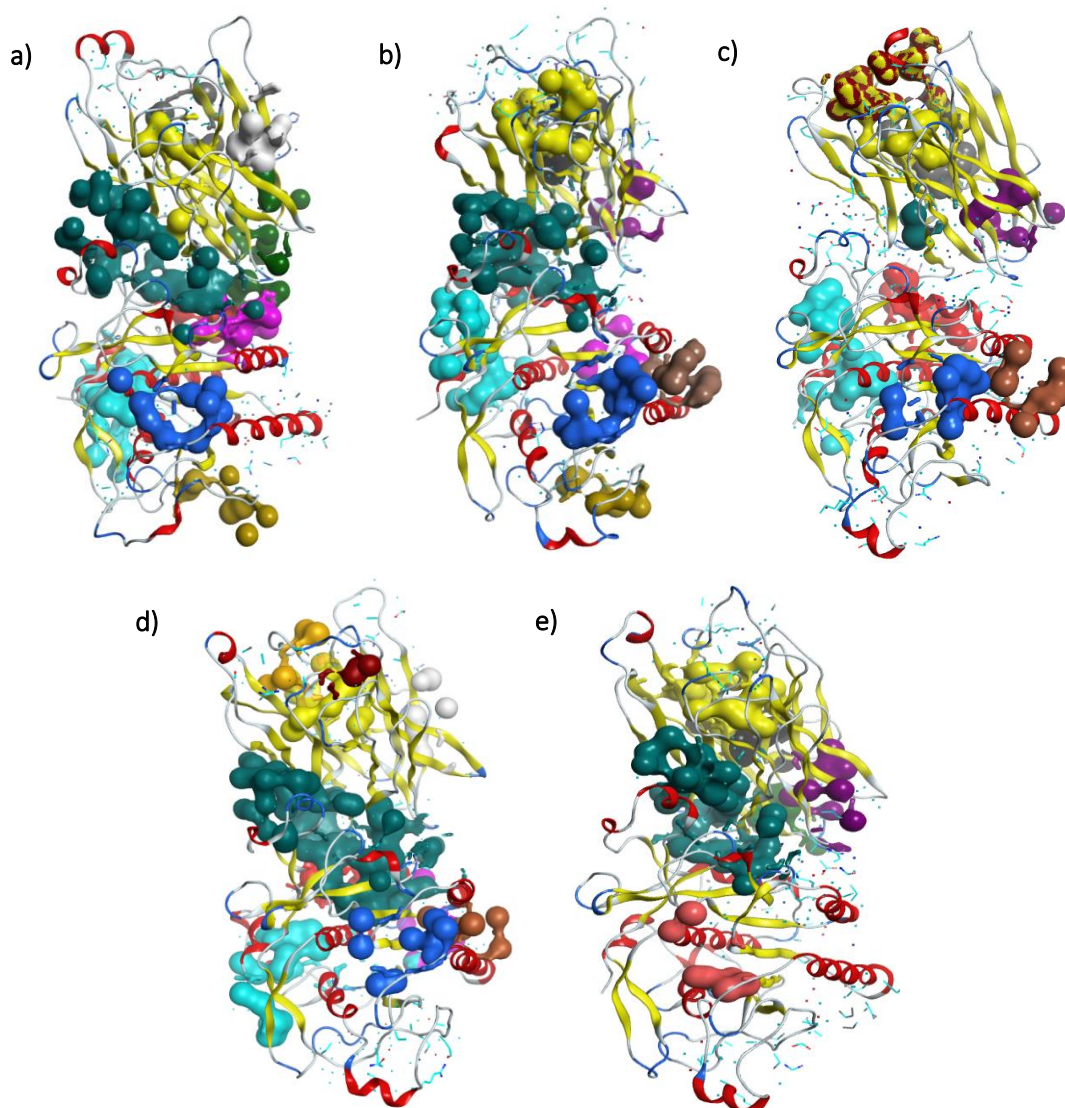


Figure 51. Pockdrug. **a)** Pockets of cluster 0; **b)** Pockets of cluster 1; **c)** Pockets of cluster 2; **d)** Pockets of cluster 3; **e)** Pockets of cluster 4

All three programs enabled to describe the pockets. No differences were found between programs, since the pockets discovered were the same. It should be noted that the residues obtained for each pocket could differ slightly depending on the program.

To reinforce the idea of which pockets were more druggable, a score drug (> 0.7) was calculated by Fpocket^{111,112} and the programs used previously. As shown in Table 13 according to the FTMap program, the pockets with the highest binding affinity for a potential PCSK9 inhibitor were P1, P6, P10 and P13. Conversely, Deepsite program determined a score drug of 1 for all the pockets discovered, having all the pockets the same binding affinity for a drug. PockDrug program set a score drug of nearly 1 for pockets P1, P6 and P10. However, the other pockets were set with a lower score drug. Finally, the Fpocket program determined P1 as the most druggable pocket, followed by P3, P6 and P10.

Table 13. Determination of the drug score for all pockets by using different programs

Methods	Clusters	Pockets													
		P1	P2	P3	P4	P5	P6	P10	P11	P12	P13	P14	P15	P18	
FTMAP Probes	0	21					41			2	9	9		11	
	1	30				11	27		22				2		
	2	25			2	14	32	12				10			
	3	17			10	2	37	11		2			7		
	4	4	8	2			52	15			9		3		
Deepsite Score drug >0.7	0	1		0.99			1								
	1	1					1		1						
	2					1				1					
	3	1					1								
	4	1					1							1	

¹¹¹ V. L. Guilloux, P. Schmidtke, P. Tuffery. *Bioinformatics* **2009**, *10*, 168.

¹¹² P. Schmidtke, V. L. Guilloux, J. Maupetit, P. Tuffery. *Nucleic Acids Res.* Vol. 38, **2010**, *2*, 582-589.

PockDrug Score drug >0.7	0	0.95		0.90	0.75	0.77	0.77	0.92			0.70	0.84		0.95
	1	0.96		0.91	0.97		0.93	0.81				0.84	0.94	
	2	0.96		0.88		0.70	0.90	0.90	0.90				0.92	0.90
	3						0.80							
	4		0.76				0.90	0.82			0.85		0.81	
Fpocket Score drug >0.7	0	0.96		0.96	0.76		0.77	0.92			0.82	0.84		0.86
	1	0.96		0.91	0.97		0.93	0.81				0.84	0.94	
	2	0.96		0.88		0.70	0.90	0.90	0.90				0.92	0.90
	3	0.95		0.88	0.85	0.76	0.93	0.92	0.71					
	4		0.76				0.90	0.83			0.85		0.81	

3.1.1.2. Screening of libraries (Docking)

One of the major challenges is the inhibition of the PCSK9 protein. Therefore, the strategy was focused on identifying small organic molecules (hits) capable of binding to PCSK9. To this end, virtual screening using a receptor-based pharmacophore approach was performed, and two logical binding sites were selected to carry out this process. The first position to dock was the EFG-(A5) domain of the LDL-R and the second position was the catalytic site of the protein by obtaining the interaction points between the prodomain and the CD.

The MM-GBSA enabled to confirm the stability of **System 1-EFG-(A5) domain** (Figure 52a) and **System 2-prodomain** (Figure 52b) after performing 100 ns of conventional molecular dynamics simulation.

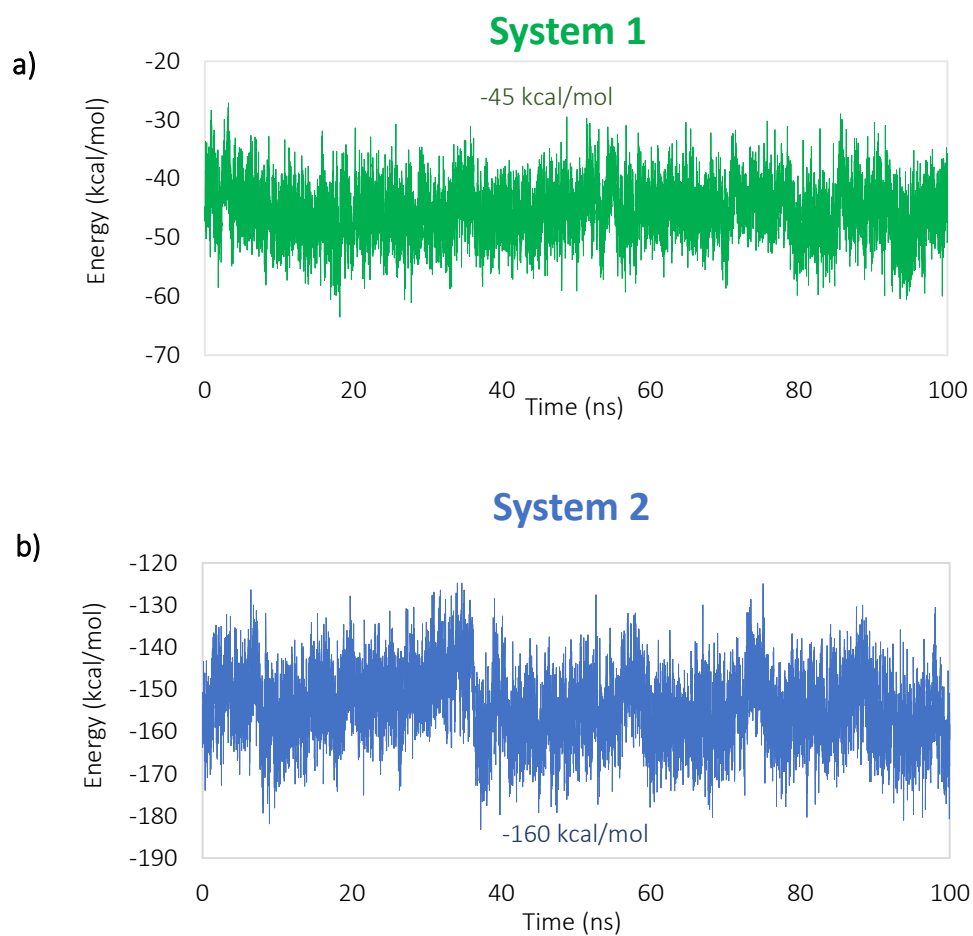


Figure 52. a), b) MM-GBSA calculated for System 1 and 2

As can be seen in both graphs after performing 100 ns of cMD, **System 1** fully converged with an average energy of -45.0 kcal/mol, whereas **System 2** converged after 40 ns with an average energy of -160.0 kcal/mol.

In a further step, MM-GBSA decomposition approach of these complexes was performed in order to determine which residues were the most important. For each system, those with highest energy contribution (≥ 1.0 kcal/mol) to the total binding free energy were considered as hot spot residues (Figure 53).

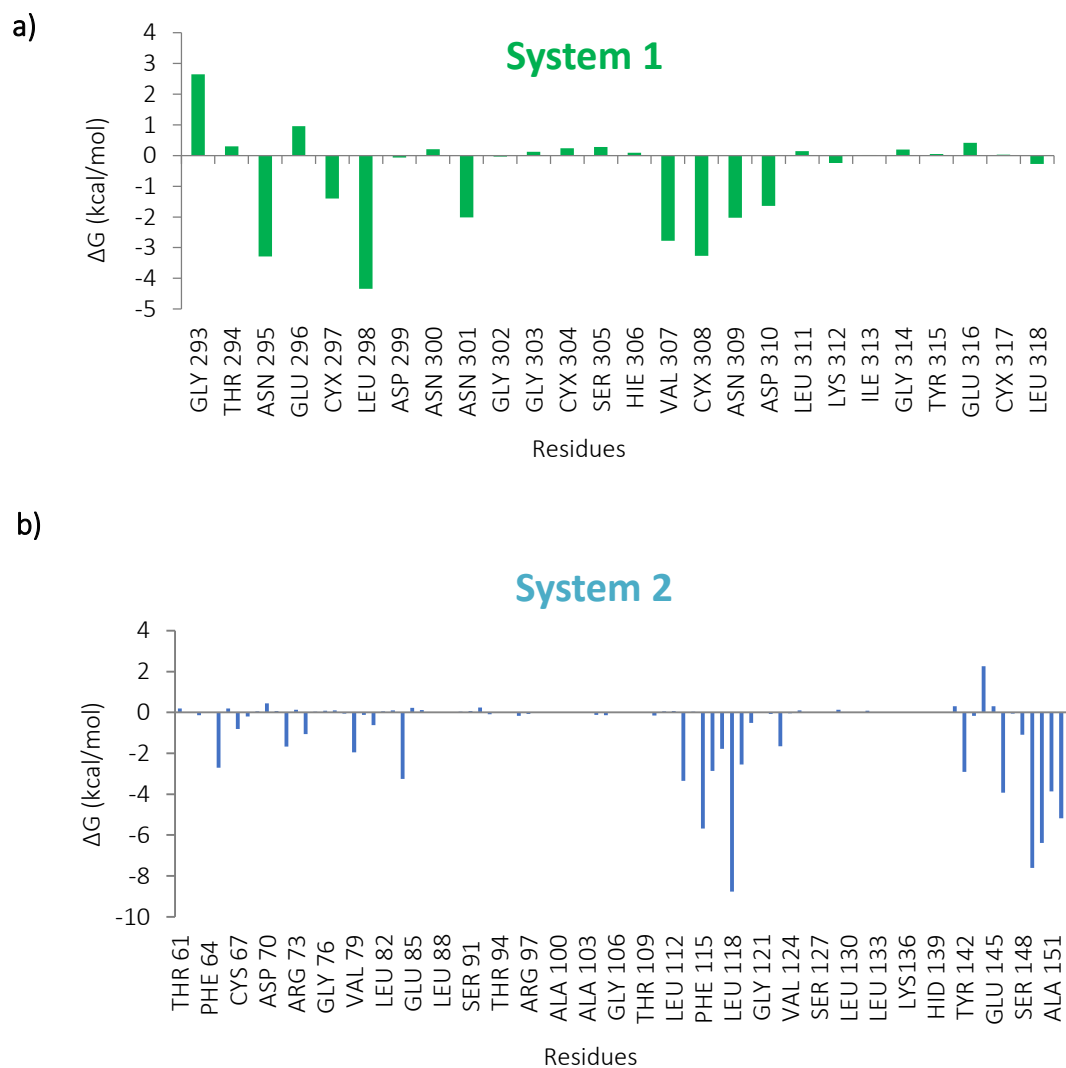


Figure 53. a), b) MM-GBSA decomposition approach for System 1 and 2

As shown in Figure 53a, for **System 1-EFG-(A5) domain** the main residues with the highest energy contribution were ASN295, CYS297, LEU298, ASN301, VAL307, CYX308, ASN309 and ASP310. Residues ASN295, ASN309 and ASP310 highly contribute to the binding free energy by forming hydrogen bond interactions with the protein (Table 14). However, the CYS297, LEU298, ASN301, VAL307 and CYX308 contributed to the total binding free energy by establishing Van der Waals (VdW) interactions.

Table 14. Parameters established between the acceptor-donor hydrogen bond

System 1	Donor	Acceptor	Occupancy (%)	Distance (Å)	Angle (°)
ASN295	HD22/ND2	ASP238 OD2	86	2.900	164.3
ASN309	THR377 HG1/OG1	OD1	88	2.718	158.6
ASP310	ARG194 HH22/NH2	OD2	98	2.777	155.6
	H/N	THR377 O	97	3.057	164.5

On the other hand, for **System 2-prodomain** the highest contributing residues to the binding free energy (Figure 53b) were HID65, TRP72, GLU84, HIE113, PHE115, HID116, GLY117, LEU118, LEU119, TYR142, ASP146, VAL149, PHE150, ALA151 and GLN152. The residues with highest energy contribution to the binding free energy by forming hydrogen bond interactions were HID65, GLU84, HID116, GLY117, LEU118, ASP146, VAL149, PHE150, ALA151 and GLN152; whereas residues TRP72, PHE 115, LEU119 and TYR142 established VdW interactions with the protein (Table 15).

Table 15. Parameters established between the acceptor-donor hydrogen bond

System 2	Donor	Acceptor	Occupancy (%)	Distance (Å)	Angle (°)
HID65	ARG295 H11/NH1	NE2	95	2.986	158.8
GLU84	ARG303 HH12/NH1	OE2	60	2.909	151.7
HIE113	HE2/NE2	GLU269 OE2	60	2.866	158.6
HID116	N/H	GLU269 OE1	60	2.848	162.2

GLY117	N/H	GLU269 OE1	60	2.925	157.9
LEU118	ARG303 HH11/NH1	O	75	2.885	158.1
ASP146	SER262 H/N	OD1	99	2.944	163.4
VAL149	H/N	GLY259 O	95	3.027	161.2
PHE150	H/N	ALA299 O	99	2.916	163.4
	ALA290 H/N	O	98	3.065	158.5
ALA151	H/N	GLY257 O	92	2.95	157.6
GLN152	ASN317 H/N	OE1	97	2.860	159.4

Then, a cluster analysis algorithm was applied to the attained trajectory frames. The representatives of the most populated clusters were used to construct the pharmacophore points of each system, based on the results obtained.

For [System 1](#), the following clusters were obtained: cluster 0 having a 61% of occupancy and, being the most populated; and cluster 1, being the second most populated with 21% of occupancy (Figure 54).

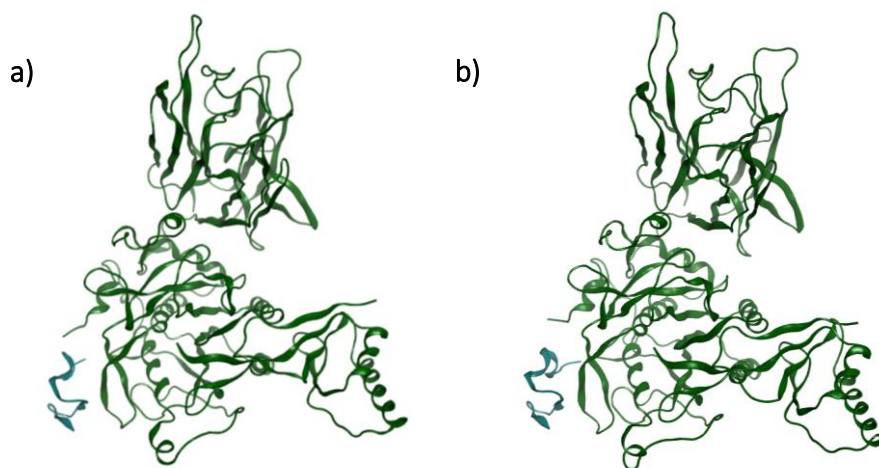


Figure 54. a) Representation of cluster 0; b) Representation of cluster 1

Regarding [System 2](#), only one cluster (Figure 55) was obtained with a 94% of occupancy for the last 40 ns of conventional molecular dynamic calculation.

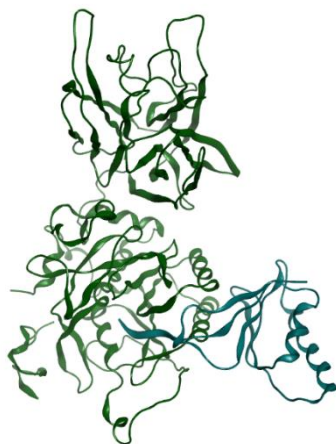


Figure 55. Cluster representation for [System 2](#)

After obtaining the principal interactions that were established by each receptor with its ligand and the representatives of the most populated clusters, the pharmacophore was constructed.

The construction of the pharmacophore for each system was based on the interactions established in the per-residue decomposition approach, giving to each residue its pharmacophore property, as donor or acceptor of hydrogen bond or hydrophobic properties (Figure 56).



Figure 56. Convention for the representation of the pharmacophore points

- Pharmacophore of [System 1](#)

In the clustering, the two most populated clusters were obtained, for which two pharmacophores were built, one for each representative.

As shown in Figure 57, each pharmacophore point corresponded to each of the residues studied in the per-residue energy decomposition. Although both representatives contained the same number of pharmacophoric points, only their spatial distribution differ.

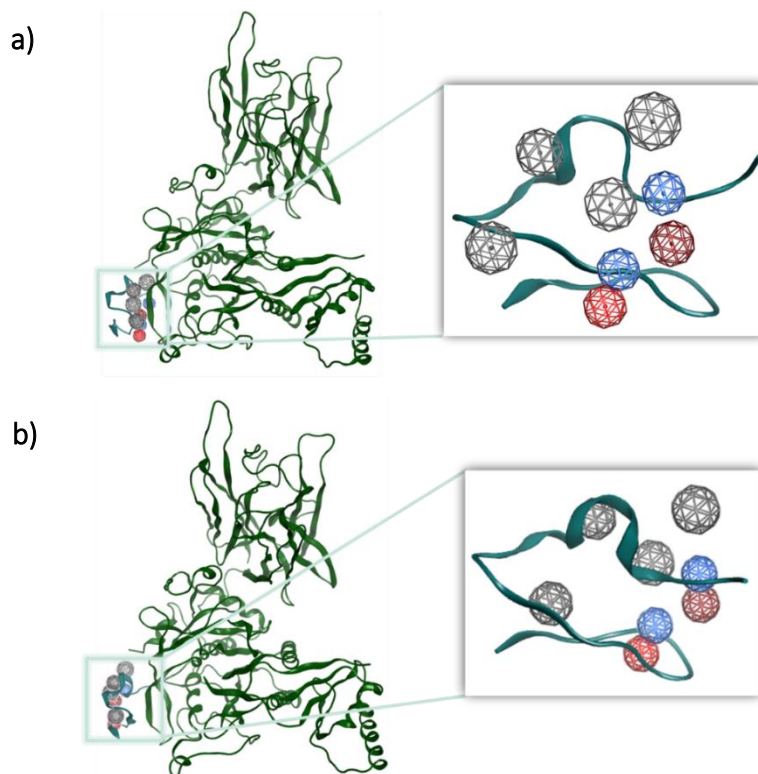


Figure 57. a) Pharmacophore for cluster 0; b) Pharmacophore for cluster 1

- Pharmacophore of [System 2](#)

Only the most populated cluster was selected, for this reason just one pharmacophore was build.

As shown in Figure 58, the pharmacophore was constructed for cluster 0, which was the most populated during the last 40 ns of cMD. The pharmacophore points were established according to the per-residue energy decomposition results.

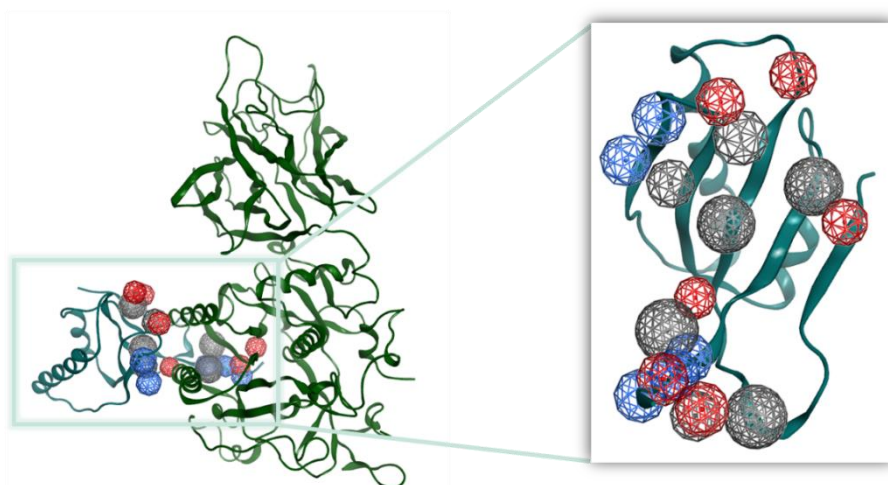


Figure 58. Representation of the pharmacophore points for [System 2](#)

Next step consisted of performing docking based on the pharmacophore points, by using two different databases Fragments[®] and Selleck Drug[®].

In order to know which compounds (hits) from the database had the best poses or configurations, two docking steps were carried out. First, the receptor was kept fixed and the different pharmacophores were used to direct the docking process. As a result, a new database was obtained containing the poses of the compounds that coped to join in the binding site and that met the pharmacophore restrictions, this time with a more manageable size. Then, the new database was used to conduct a second docking where the side-chains near the residues of the binding site were allowed to move (induced fit). A complete conformational freedom for the compounds of the database was always used in order to sample as much as possible the energetically accessible conformations. After performing the induced fit docking, a new database of the best energetically compounds that fit in the pharmacophore points was achieved; a selection of the 500 first complexes (protein+compound pose) with the best binding free energy was acquired for subsequent minimization step.

The MM-GBSA and the MM-PBSA (Molecular Mechanics Poisson Boltzmann Surface Area) were performed for each complex in order to select the best 100 compounds energetically. For [System 1](#), a consensus was required so as to discard the same compounds poses appearing in both approaches. Of the initial 1488 initial compounds from Fragments[®] Database, 63 hits were selected, whereas from the Selleck Drug[®] Database 74 hits of the initial 915 compounds were chosen. However, for [System 2](#) no consensus was needed since only one cluster was attained. In this case, 77 and 64 hits from Fragments[®] Database and Selleck Drug[®] Database were respectively chosen.

Once the compounds were selected, a production of 10 ns of cMD simulation was performed. The MM-GBSA and MM-PBSA were carried out again, this time for the full trajectory, in order to determine the energetic stability of each selected complex over time. Thus, for [System 1](#) four hits were selected from Fragments[®] Database, while no hits were chosen from Selleck Drug[®], since none of the complexes showed good energetic convergence over the time. On the other hand, for [System 2](#), seven hits showed good energetic convergence over the 10 ns of cMD trajectory for each database.

Finally, 40 ns of cMD simulation with its respective MM-GBSA and MM-PBSA calculations were performed by using the selected complexes aiming to identify which ones were better energetically and also stable during the new longer trajectory.

In the light of the results, **System 1-EFG-(A5) domain** showed only one reasonable result. The hit assigned in the position **1356_1** of the Fragments® Database was selected from the 4 best compounds obtained after calculating the free energy of binding. The hit structure is represented in Figure 59.

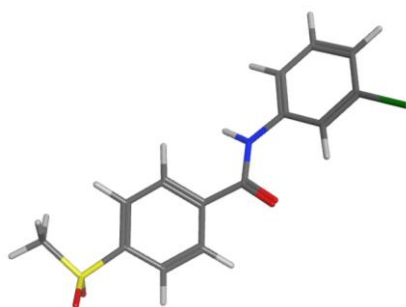


Figure 59. Structure of hit **1356_1**

The complex (protein+hit) was stabilized over the last 30 ns of cMD although the energy not completely converged. The average energy (last 30 ns of cMD) for GB was -15.0 kcal/mol and -13.0 kcal/mol for PB as shown in Figure 60a. The main interactions established with the protein were practically hydrophobic (benzene-hydrogen) as shown in Figure 60b.

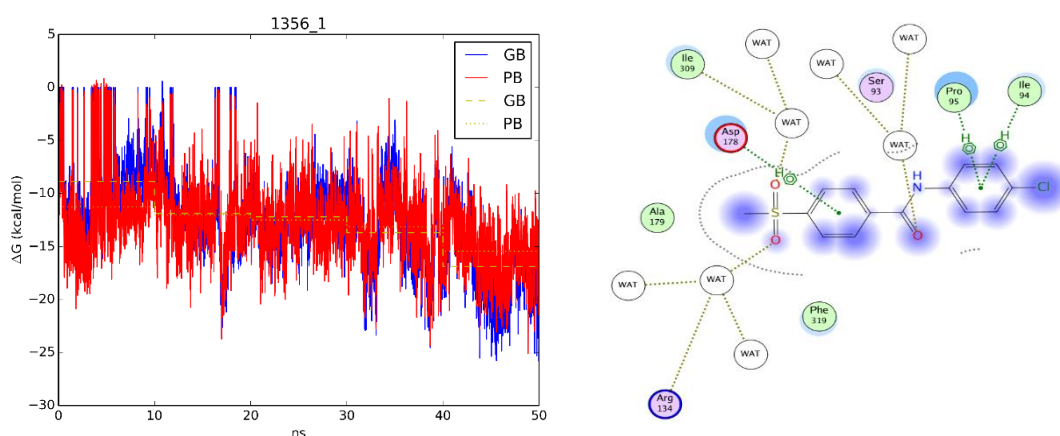


Figure 60. a) MM-GBSA/PBSA of hit **1356_1**; b) Interactions established with the protein

As regards to [System 2-prodomain](#), seven hits were obtained for each database as explained above, and therefore the most significant results are shown. From Fragments® Database, the hit assigned in position **1330_1** (Figure 61a) showed fully stability during the 50 ns of cMD, being -35.0 kcal/mol the energy obtained for GB, and slightly different for PB, with an average energy of -32.0 kcal/mol, as shown in Figure 61b.

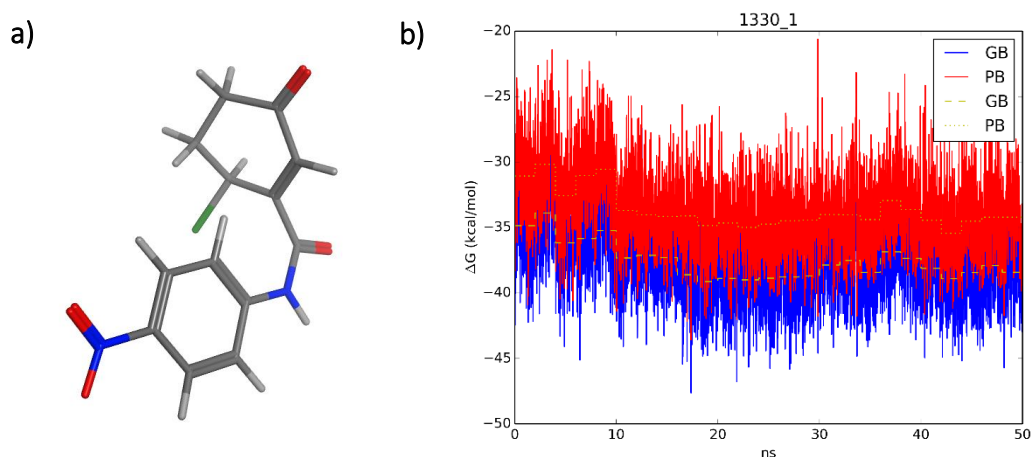


Figure 61. a) Structure of hit **1330_1**; b) MM-GBSA/PSA of hit **1330_1**

The hit totally converged due to the interactions established with the protein. The hit could perfectly match with the pocket, and the interactions established were basically hydrophobic, where HID134 established a π -stacking interaction with the nitrobenzene of compound **1330_1**. Only one hydrogen bond was formed between SER (*H*, 294) and the *O* of the hit as shown in Figure 62.

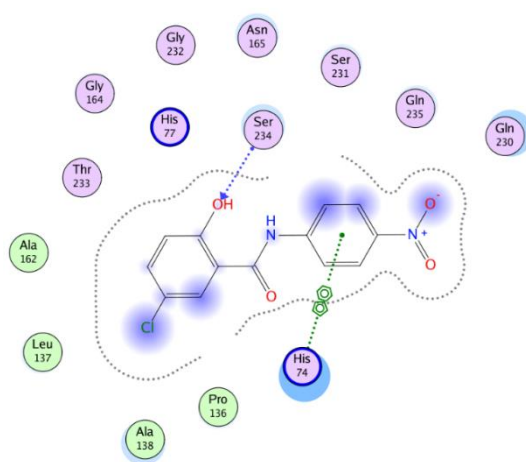


Figure 62. Main interactions established between hit **1330_1** and the protein

Two hits with remarkable energetic convergence over time were obtained from Selleck Drug[®] Database. The hit located in position **827_3** (Figure 63a) showed good stability, being the GB average energy of -45.0 kcal/mol, and the PB, with an average energy of -35.0 kcal/mol (Figure 63b). The compound fit inside the pocket and the main interactions were formed by 3 hydrogen bonds, specifically between the *NH* group with ALA (*O*, 198), the *O* with ALA (*NH*, 198) and the last one between an *O* and GLY (*NH*, 167). However, only one hydrophobic bond (π -stacking) was observed as shown in Figure 63c.

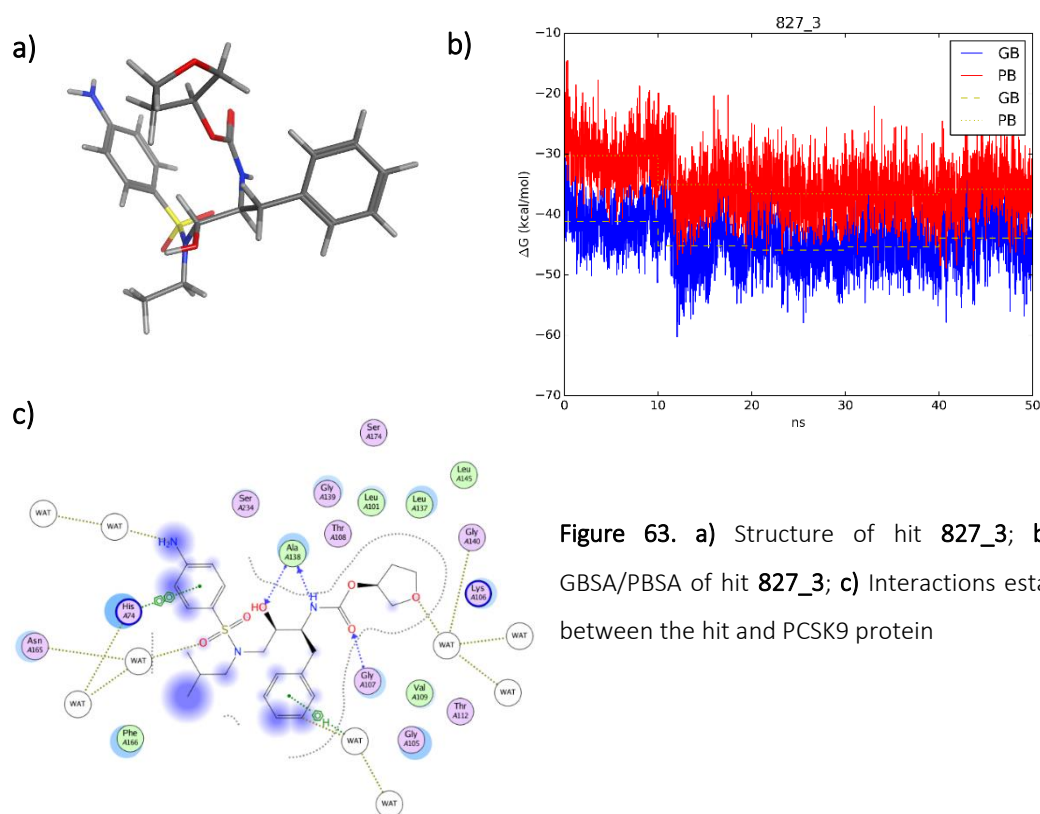


Figure 63. a) Structure of hit **827_3**; b) MM-GBSA/PBSA of hit **827_3**; c) Interactions established between the hit and PCSK9 protein

The last possible PCSK9 inhibitor was positioned in **896_2** (Figure 64a) from the database. As shown in Figure 64b the hit was fully energy converged. The average energy for GB method was approximately of -45.0 kcal/mol, and for PB method -32.0 kcal/mol. The highly contribution to the binding free energy was attributed to the interactions established, where 3 hydrogen bonds were observed. The first hydrogen bond was with the *NH* and GLY (*O*, 167), the second one was observed between the *N* and HID (*H*, 134), and the last bond was established between the *O* (acting as acceptor hydrogen bond) and ALA (*H*, 198) as shown in Figure 64c.

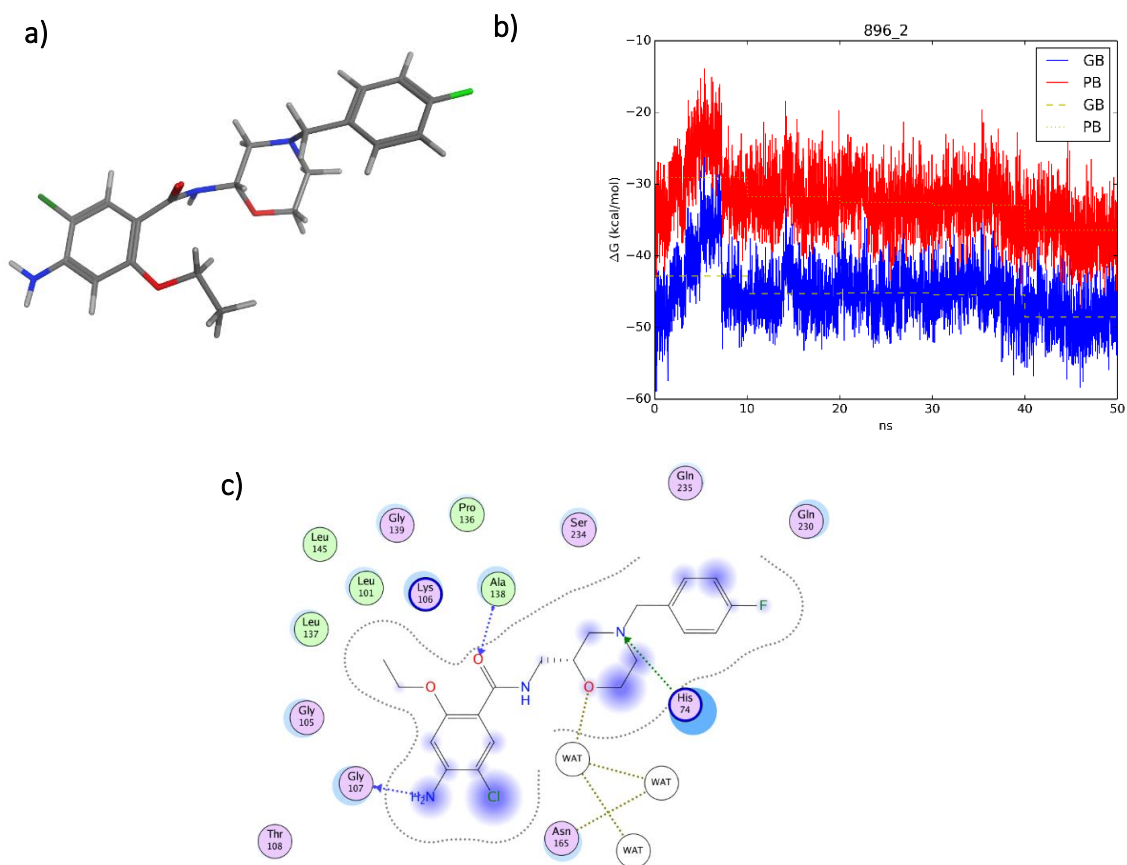


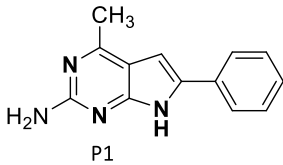
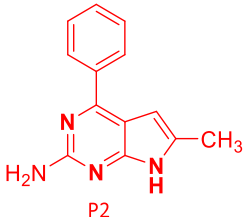
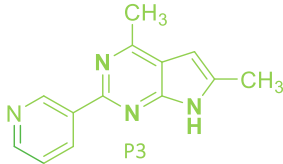
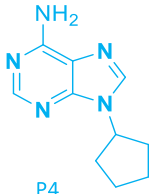


Figure 64. a) Structure of hit **896_2**; b) MM-GBSA/PBSA of hit **896_2**; c) Interactions established between the hit and PCSK9 protein

3.1.1.3. Rational design of inhibitors based on the structure of the ligands

A long process was carried out to obtain small organic molecules as PCSK9 inhibitors. That is why the strategy followed is summarized below.

First of all, the protein was solvated with the ligand, which was previously prepared as a BOX containing one ligand and a variable number of water molecules. This ligand-BOX was equilibrated with a short MD. In this way, 8 different ligands-BOXes were solvated with the protein and, as a result, 8 different systems were obtained, each one differentiated by the number of ligand-BOX that could fit inside the protein-BOX, as shown in Table 16.

Table 16. Number of ligands for each system

System	Ligand	Ligand Number
1	 <p>P1</p>	34
2	 <p>P2</p>	34
3	 <p>P3</p>	22
4	 <p>P4</p>	34
5	 <p>P5</p>	25
6	 <p>P8</p>	34

Then, 4 different GaMD, each one of 300 ns length, were performed for each system. The difference was in the starting temperature, which enabled to start with a different velocity in every molecular dynamic calculation.

Next, each GaMD trajectory was striped to obtain N trajectories, being N the number of ligands in the system, each one of them without waters and containing only one ligand. Then, the MM-GBSA was performed for each of the N trajectories and reactive ligands were selected using the average of $\Delta G_{\text{binding}}$ obtained. The software fdMD was used for the global process.

The selected ligands were used for further molecular dynamics calculations. The single trajectories of the ligand stracted from those containing multiple ligands were used; in particular, the last point of the GaMD was used as starting point for a 100 ns of conventional molecular dynamics calculations. The MM-GBSA was calculated for the full trajectory and the ligands with better energetic stability were chosen to further extend their cMD 100 ns.

Ligands with higher energetic stability along the full trajectory and good receptor-ligand interactions were chosen for possible ligand modifications. With this in mind, the main goal was to introduce modifications in the ligand in order to improve the binding free energy and the interactions established with the receptor. Therefore, the last position of 200 ns of cMD was used as a starting point for the 100 or 300 ns of cMD simulation for each implemented modification. However, no modifications of ligand **P5** and **P8** were performed.

System 1

From the 200 ns lenght of cMD three different ligands were selected (**lp1_531_din3**, **lp1_551_din1** and **lp1_557_din3**). The lp1 nomenclature refers to the ligand and din refers to the GaMD simulations. This nomenclature can be extrapolated to the other ligands.

Only ligands **lp1_551_din1** and **lp1_557_din3** were selected in order to improve the ligand interactions with the protein through slight modifications of the ligand **P1**. For compounds **lp1_551_din1** and **lp1_557_din3**, 300 and 100 ns of cMD were respectively performed. Moreover, their energetic convergence was studied trough MM-GBSA calculations. For **lp1_551_din1** two modifications were done and, in this case, only the

best one is explained. On the other hand, for **lp1_557_din3** four different modifications were carried out, being modification 1 the best compared to the other four.

Lp1_551_din1: Modification 2

The modified structure (Figure 65a) differed from **P1** by incorporating a chlorine at the *para* position of the benzene and by the addition of a polar substituted benzyl in the position 4 of the [2,3-*d*]pyrrolopyrimidine. The MM-GBSA was calculated for the 300 ns of conventional molecular dynamics calculations. The average of the energy of binding was -42.5 kcal/mol (Figure 65b). The energy values in comparison with ligand **P1** were clearly much better in the modification made, where the energy was stabilized for the last 100 ns. The ligand was positioned at **pocket P6**. In this case, the ligand was surrounded by protein residues (Figure 65c). Two hydrogen bonds, plus one hydrophobic interaction were observed. The hydrogen bonds were between the *H* of the ligand and the *O* of ASN439 and the *H* with ASP422.

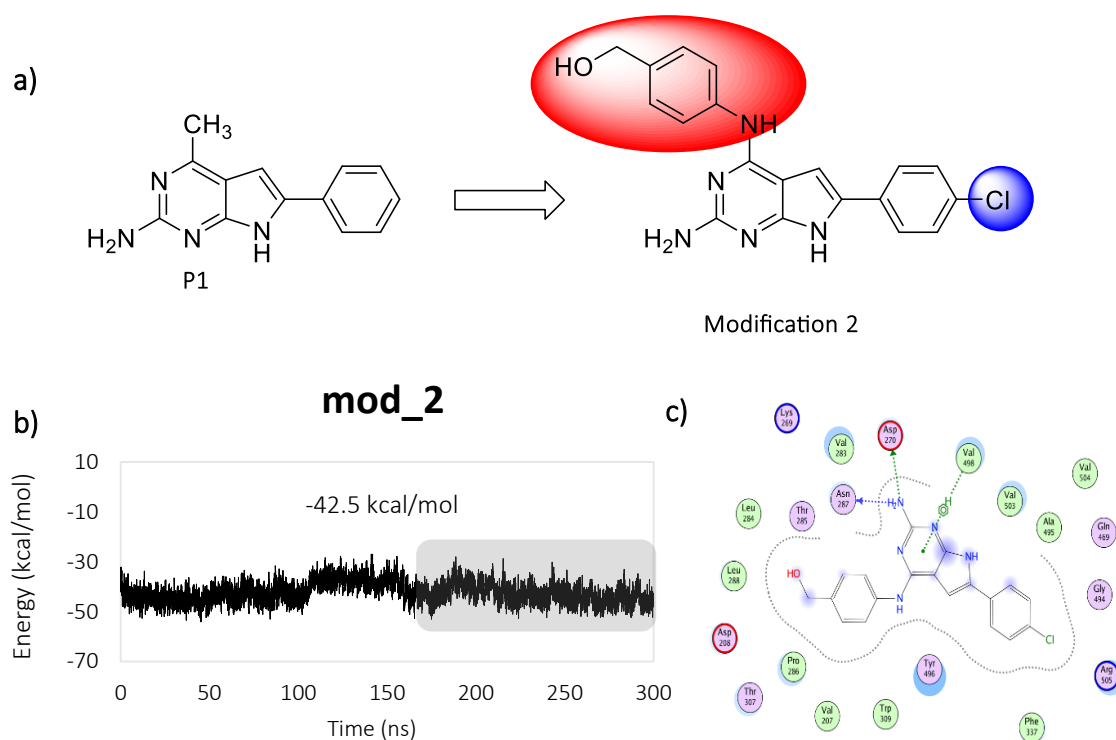


Figure 65. a) Structure of modification 2; b) MM-GBSA calculation; c) Interactions established with the protein

The modifications incorporated in **P1** were better because hydrogen bonds could be performed as well as benzene-hydrogen hydrophobic interactions.

Lp1 557 din3: Modification 1

The modified structure (Figure 66) differed from **P1** by switching the $-NH$ for a methyl group in which became a tertiary amine; the second modification was the addition of a chlorine group in the position 4 of the [2,3-*d*]pyrrolopyrimidine; and the last change consisted of the addition of an ethanol in the amino group.

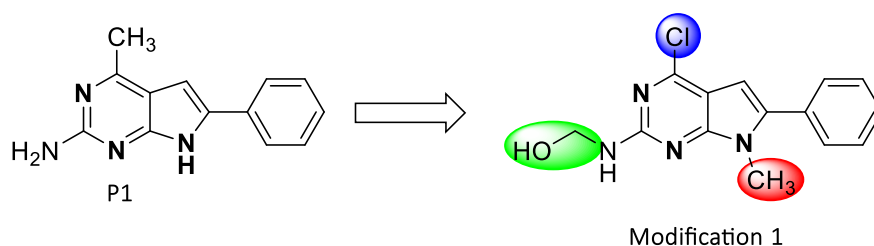


Figure 66. Modifications realized in **P1** ligand

The MM-GBSA was calculated for 100 ns of cMD. The average binding free energy was -29.0 kcal/mol for the last 40 ns of cMD trajectory and -31.0 kcal/mol for the initial 60 ns (Figure 67a). The ligand was positioned at **pocket P15**. In this case, the ligand established one more hydrogen bond compared to ligand **P1**. The *OH* group provided the ligand the possibility of establishing two hydrogen bonds, in this case ARG458 made a hydrogen bond with the *O*, and PRO331 made a hydrogen bond with the *H*. On the contrary, in **P1** ligand ARG458 made a hydrogen bond with the *N* of the ligand. Hydrophobic interactions were also observed, which gave stability to the ligand inside the pocket (Figure 67b).

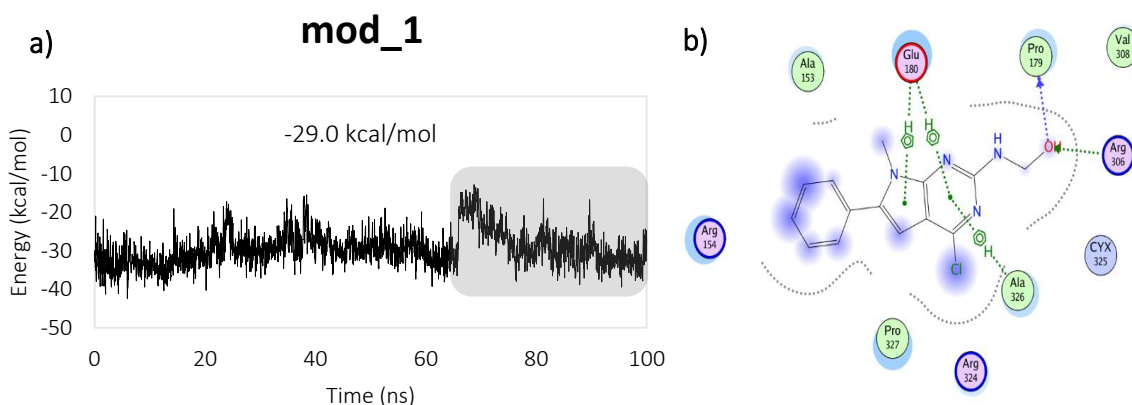


Figure 67. a) MM-GBSA calculated for modification 1; b) Interactions established

System 2

Three different ligands were selected from the last 200 ns of cMD trajectory, **lp2_546_din3**, **lp2_548_din4** and **lp2_558_din3**. Specifically, only **lp2_548_din4** was selected to improve the interactions with the protein through slight modifications of **P2** ligand. A modification was performed to this ligand, which is described below.

lp2_548_din4: Modification 1

The structure of this modification (Figure 68) was slight different from **P2** ligand. The main goal was to keep the ligand with bigger hydrophobic groups and to give polarity to the structure so as to increase the chance of creating new hydrogen bonds with the protein. The ligand modified was placed in **pocket P10**.

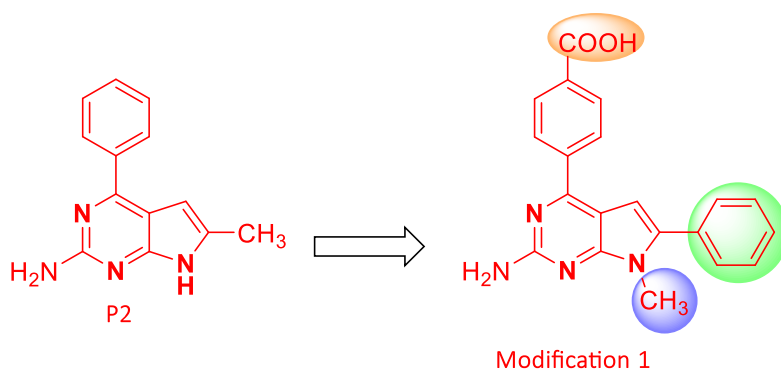


Figure 68. Structure of modification 1

When comparing the binding free energy between **P2** and modification 1, the latter was energetically better with an average energy of -45.6 kcal/mol being more stable over the 100 ns (Figure 69a). The interactions established were better in modification 1 than **P2** ligand. The two benzenes gave hydrophobicity and ligand stability making higher the energy values, while the methyl and the amino group seemed to be surrounded by water molecules. No hydrogen bonds were established (Figure 69b).

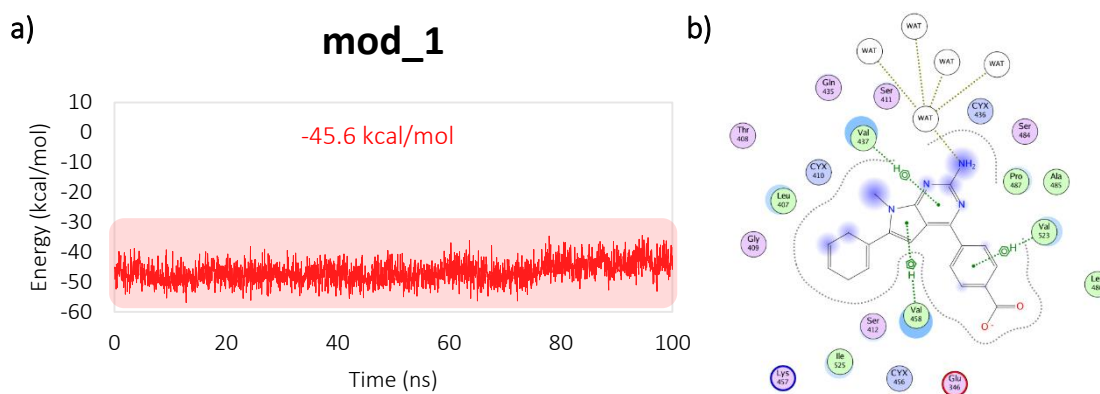


Figure 69. a) MM-GBSA calculated for **lp2_548_din4** modification 1; b) Interactions established with the protein

System 3

After performing 200 ns of cMD only one ligand was selected to carry out modifications of **P3** ligand. In this case, **lp3_552_din4** was modified in order to improve the interactions established with the protein, as well as to increase the energy values.

lp3_552_din4: Modification 1

The **P3** structure was modified as follows: at position 4, a benzoic acid was introduced and the methyl group was replaced for a benzyl group (Figure 70).

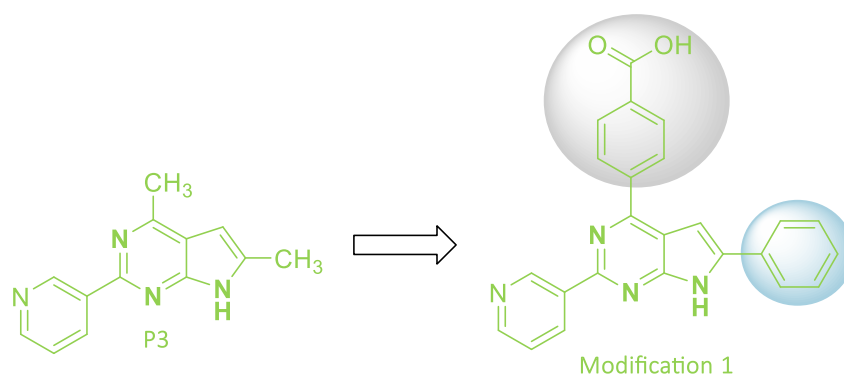


Figure 70. Ligand **P3** modified

This ligand was placed near pocket **P5** and **P3**, between two α -helices. The MM-GBSA was calculated. In view of the results, it is remarkable the improvement of binding free energy achieved by the modification, where the energy went from -25.0 to -44.0 kcal/mol (Figure 71a). Two hydrogen bonds were established between the *NH* and ARG167, as well as

between the *O* and HIS417. Hydrophobic interactions were also observed with the benzyl groups (Figure 71b).

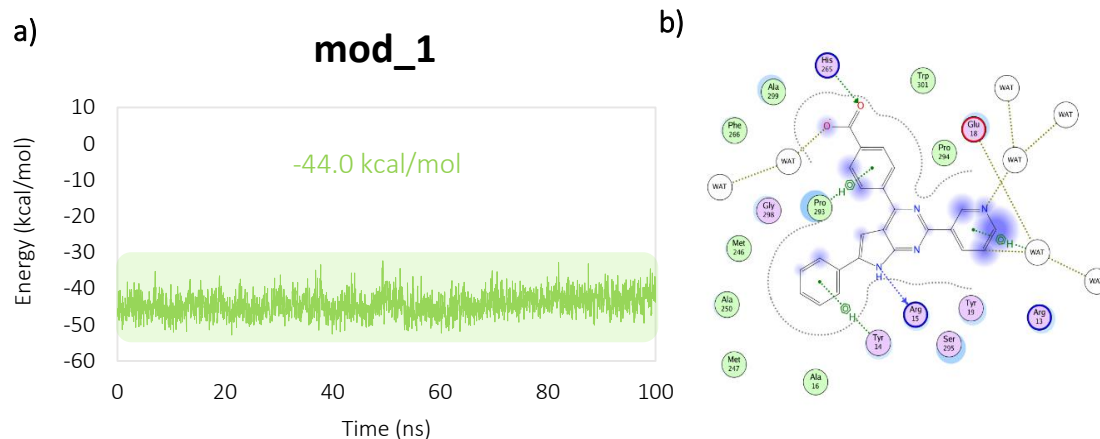


Figure 71. a) MM-GBSA calculated for modification 1; b) Interactions established with the protein

System 4

After performing 200 ns of cMD the ligands selected to carry out modifications were **ligand_534_din2**, **ligand_542_din4** and **ligand_543_din1** due to the good average of binding free energy respect the other ligands and, also for the good location of the selected ligands. **Ligand_542_din4** and **ligand_543_din1** were placed in the same **pocket P1**. In this case the best modification was **ligand_543_din1** modification 3 in which the ligand was introduced inside the **pocket P1**, where the catalytic domain is found. On the other hand, **ligand_534_din2** was finally discarded because the modifications made did not improve the energy values and the interactions established by the **P4** itself. As said above only the best modification is explained below.

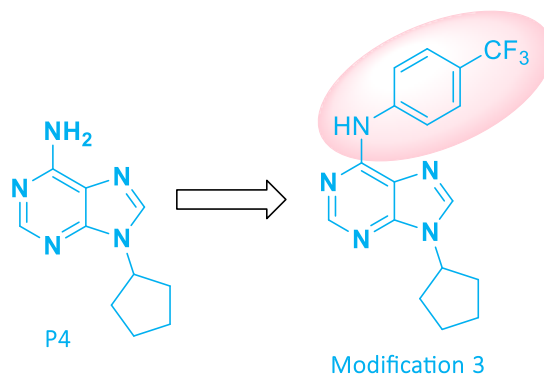
Lp4 543 din1: Modification 3

The only modification performed (Figure 72a) was the introduction of a 4-(trifluoromethyl)benzylamine group in position 4 of the [2,3-*d*]pyrrolopyrimidine.

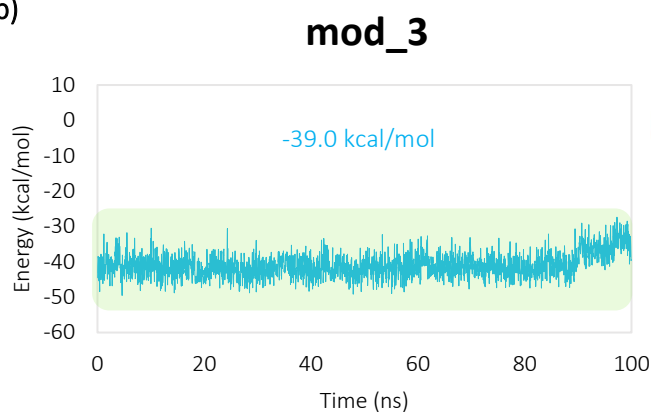
The MM-GBSA was calculated for the 100 ns of trajectory. The average of binding free energy was around -39.0 kcal/mol. The energy was stabilized for the entire cMD trajectory. However, for the last 10 ns there was an increase of energy as seen in Figure 72b. In comparison to **P4** the energy values were better. Modification 3 was placed inside the protein over the 100 ns of trajectory.

The energy values were higher due to the hydrophobic interactions established with the protein and, in no case the ligand was solvated by water molecules. The main interactions established were basically hydrophobic, such as benzene-H bonds, which provided the ligand the sufficient stability to reside inside the pocket. Only the cyclopenthyl was exposed to the solvent (Figure 72c).

a)



b)



c)

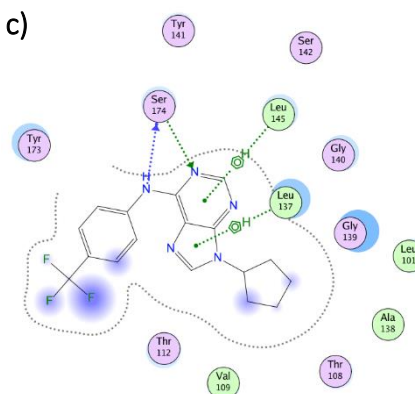


Figure 72. a) Structure of Ip4_543_din1; b) MM-GBSA of modification 3; c) Interactions established with the PCSK9

System 5

After carrying out 200 ns of cMD two ligands were selected. In this case, the P5 ligand was not modified and for this reason no more calculations were performed since 200 ns were enough. The selected ligands were Ip5_532_din2 and Ip5_553_din2.

Lp5 532 din2

This ligand was located in **pocket P15**. The MM-GBSA was calculated for the full 200 ns of cMD. It was found that the binding free energy was converged for the full trajectory with an average energy of -28.3 kcal/mol (Figure 73a). As the ligand had [REDACTED], 6 [REDACTED] bonds were observed and no water molecules were surrounding the ligand (Figure 73b).

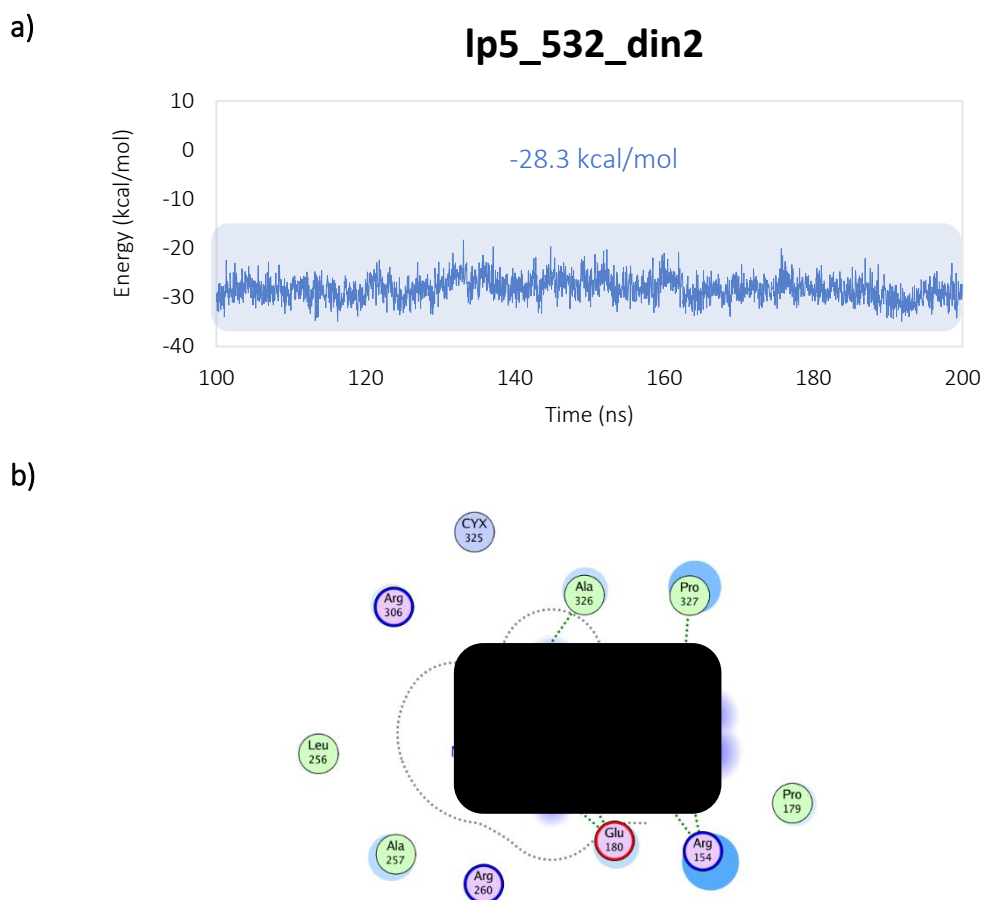


Figure 73. a) MM-GBSA; **b)** Interactions established between the ligand and the protein

Lp5 553 din2

The ligand was placed in **pocket 18**. The average of the binding free energy was approximately -29.9 kcal/mol for the 200 ns of trajectory (Figure 74a). Three hydrophobic interactions, particularly benzene-hydrogen interactions were made by SER545, GLY547 and ARG582. This made it possible for the ligand to be retained inside the protein (Figure 74b).

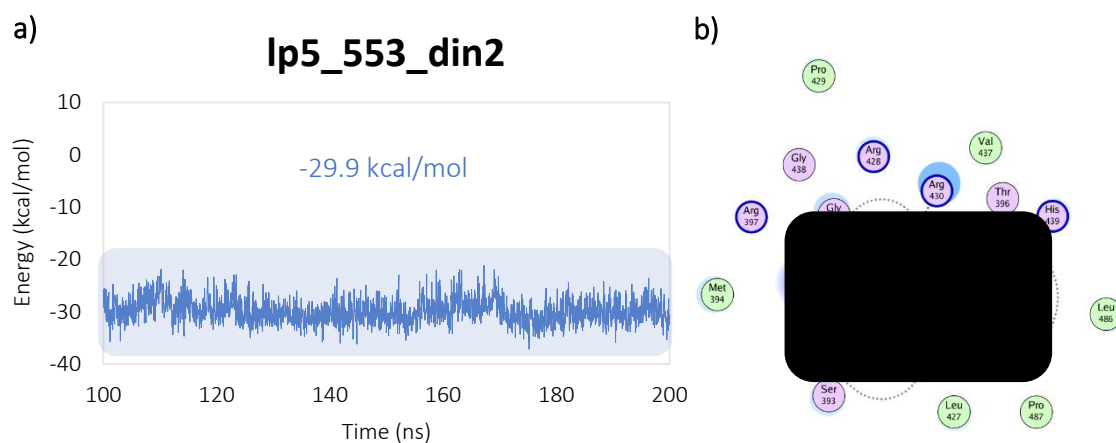


Figure 74. a) MM-GBSA of ligand **lp5_553_din2**; b) Interactions established with the protein

System 8

After 200 ns of cMD simulation, two ligands were selected. In this case the **P8** ligand was not modified and for this reason no more calculations were performed since 200 ns were enough. The selected ligands were **lp8_550_din1** and **lp8_564_din3**.

lp8_550_din1

This ligand was found in **pocket P15**. The binding free energy corresponding to the ligand was converged for all MD calculations, being -27.5 kcal/mol the average energy value.

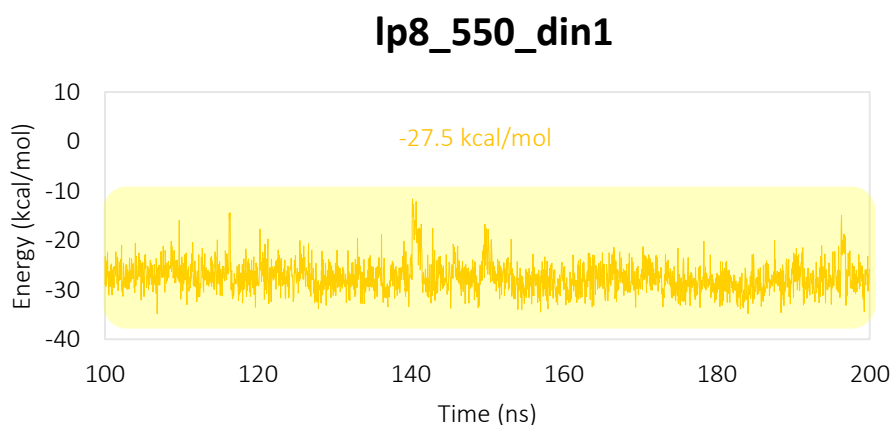


Figure 75. MM-GBSA of **lp8_550_din1**

In this pocket, two other ligands were placed (**lp8_536_din2** and **lp8_559_din4**). In the three cases, one hydrogen bond was observed between the [REDACTED] found in the ligand and ARG (*H*, 458). Also in all cases the [REDACTED] was surrounded by water molecules. The interactions established differed according to the position in which the ligand was found. Interactions with the protein increase when the ligand was positioned completely horizontal, whereas if it was upright, the interactions with the protein were less. In this case, the position of **lp8_536_din2** and **lp8_559_din4** were found in a vertical position and the interactions established were poor with respect to **lp8_550_din1** which was in a horizontal position (Figure 76).

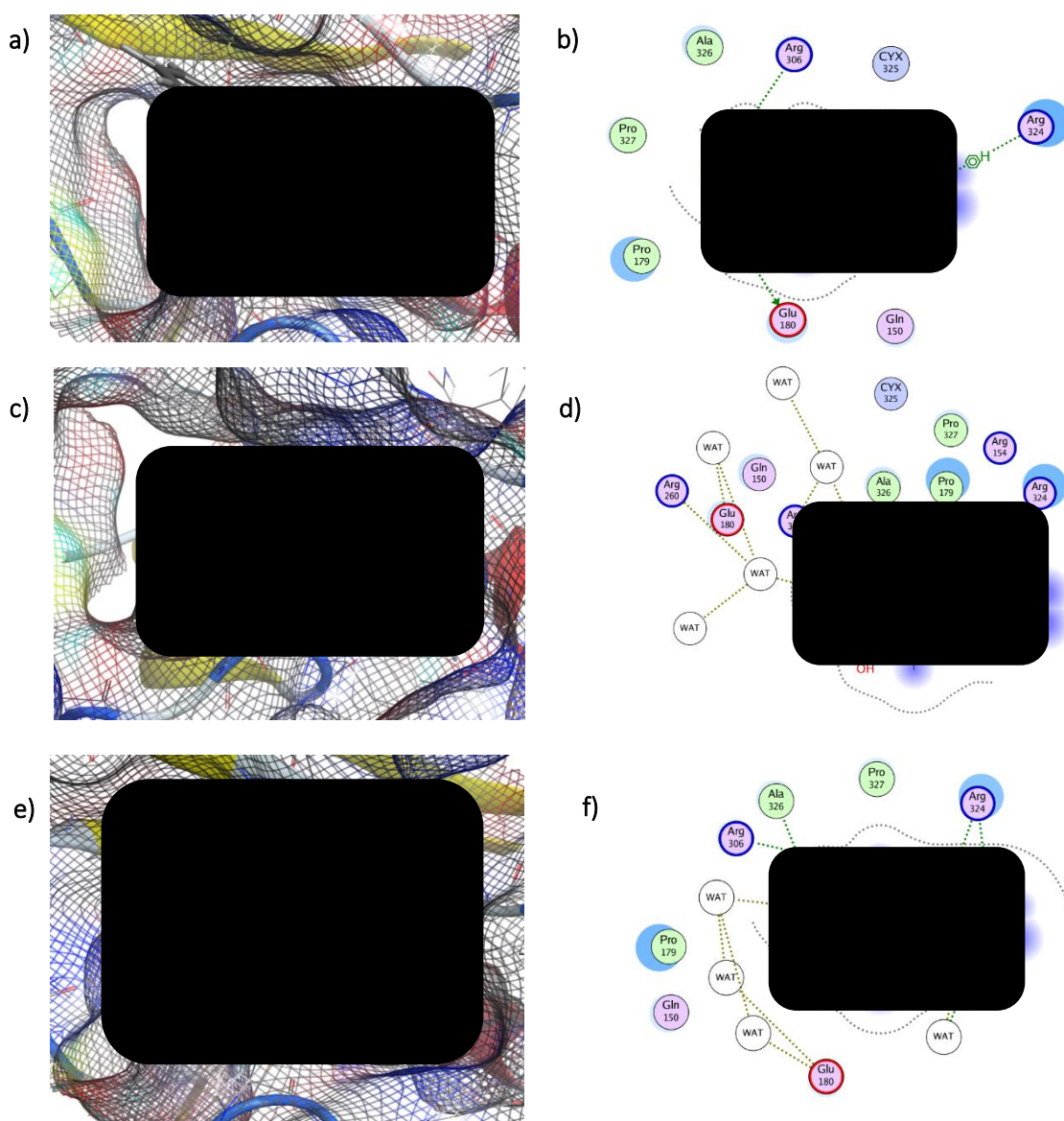


Figure 76. a) Position of **lp8_536_din2**; b) Interactions established by **lp8_536_din2**; c) Position of **lp8_550_din1**; d) Interactions established by **lp8_550_din1**; e) Position of **lp8_559_din4**; f) Interactions established by **lp8_559_din4**

The main differences lie in the hydrophobic interactions, being **lp8_550_din1** the one with more [REDACTED] interactions, specifically four (Figure 76d).

lp8_564_din3

This ligand was found in **pocket 11**. The binding free energy calculated was approximately of -34.9 kcal/mol for the whole 200 ns of cMD trajectory (Figure 77). **lp8_564_din3** was very stable over time, due to the interactions established with the protein.

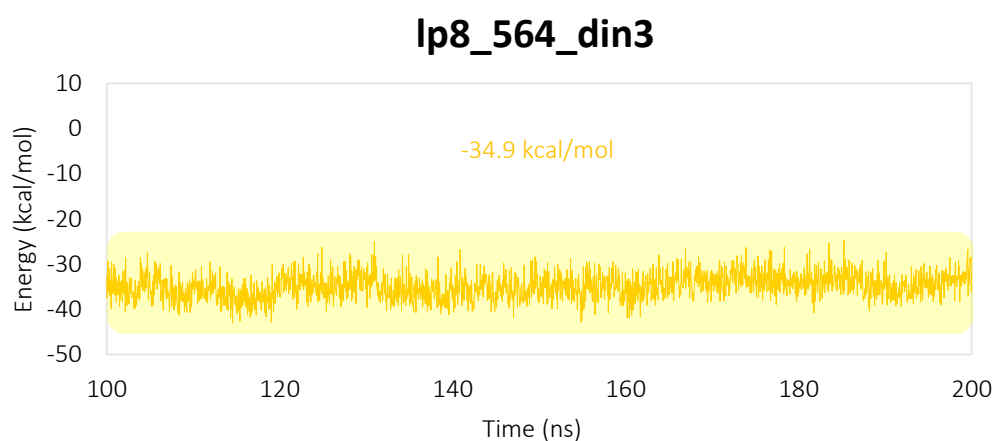


Figure 77. MM-GBSA of ligand **lp8_564_din3**

As shown in Figure 78, two hydrogen bonds were observed, one between ASP590 and the [REDACTED], and one between the [REDACTED] and GLY640. The hydrophobic bonds were majority between [REDACTED], and just one π -stacking was observed.

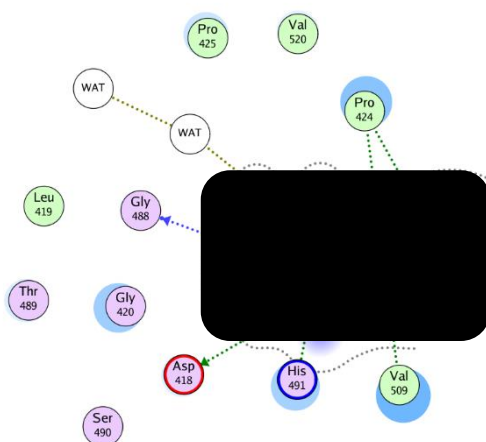
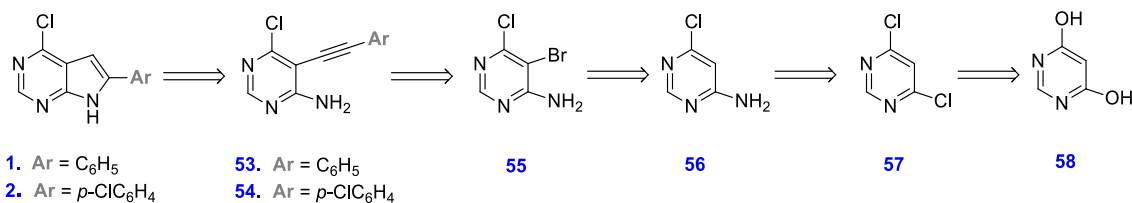


Figure 78. Main interactions established between the protein and **lp8_564_din3**

3.1.2. Preparation of pyrrolo[2,3-d]pyrimidines

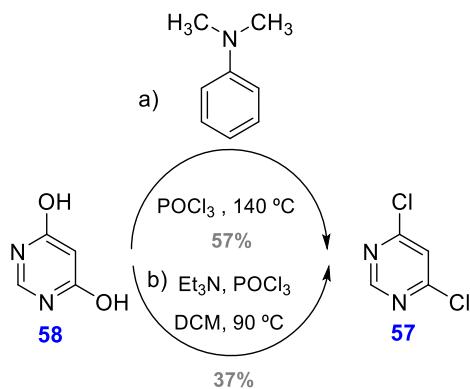
Pyrrolo[2,3-d]pyrimidines **1-14** were prepared from the nucleus of 6-aryl-4-chloropyrrolo[2,3-d]pyrimidines, whereas the pyrrolopyrimidine system was obtained from 4,6-dihydroxypyrimidine commercially available according to the retrosynthetic scheme shown below (Scheme 1).



Scheme 1

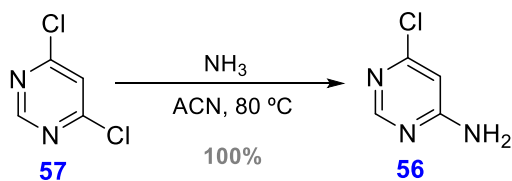
3.1.2.1. Preparation of 4-amine-5-bromo-6-chloropyrimidine (**55**)

The 4,6-dihydroxypyrimidine (**58**) was treated with an excess of POCl₃ in the presence of dimethylaniline (DMA) at 140 °C thus, the expected compound was obtained with a yield of **57%** (Scheme 2). Alternatively, chlorination with POCl₃ using triethylamine as assisting base in dichloromethane was performed, but in these conditions only the expected pyrimidine **57** with **37%** yield was obtained (Scheme 2).



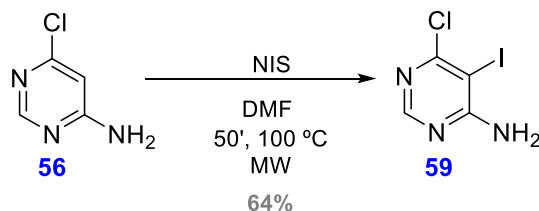
Scheme 2

Next, mono-amination of **57** was carried out with an excess of ammonium hydroxide. The transformation was conducted in a sealed Pyrex tube heated at 80 °C in acetonitrile. A **100%** yield was obtained (Scheme 3).



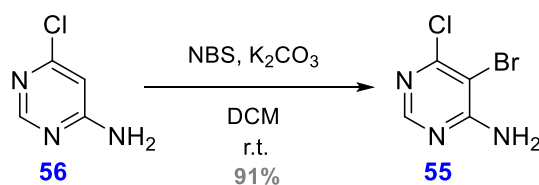
Scheme 3

Halogenation of amino-pyrimidine **56** was performed by treatment with *N*-iodosuccinimide (NIS) in DMF, and with the assistance of microwave irradiation, heated at 100 °C for 50 minutes. Under these conditions, 4-amino-6-chloro-5-iodopyrimidine (**59**) was obtained with **64%** yield (Scheme 4).



Scheme 4

Treatment of the aminopyrimidine **56** with *N*-bromosuccinimide (NBS) instead of NIS resulted in bromine-pyrimidine **55** with **91%** yield (Scheme 5).



Scheme 5

Shielding of the proton in position 2 in the halogenated derivatives, as well as the absence of signal in ¹H NMR spectrum for proton of position 5 allowed to configure the structures of pyrimidine **55** and **59** derivatives (Figure 79).

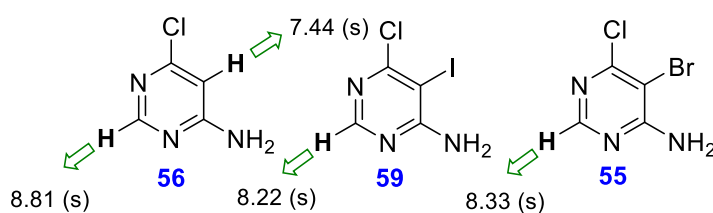
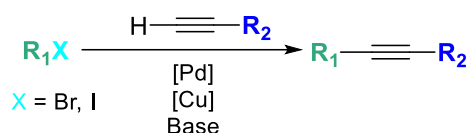


Figure 79. ¹H NMR spectral data of **56**, **59** and **55**

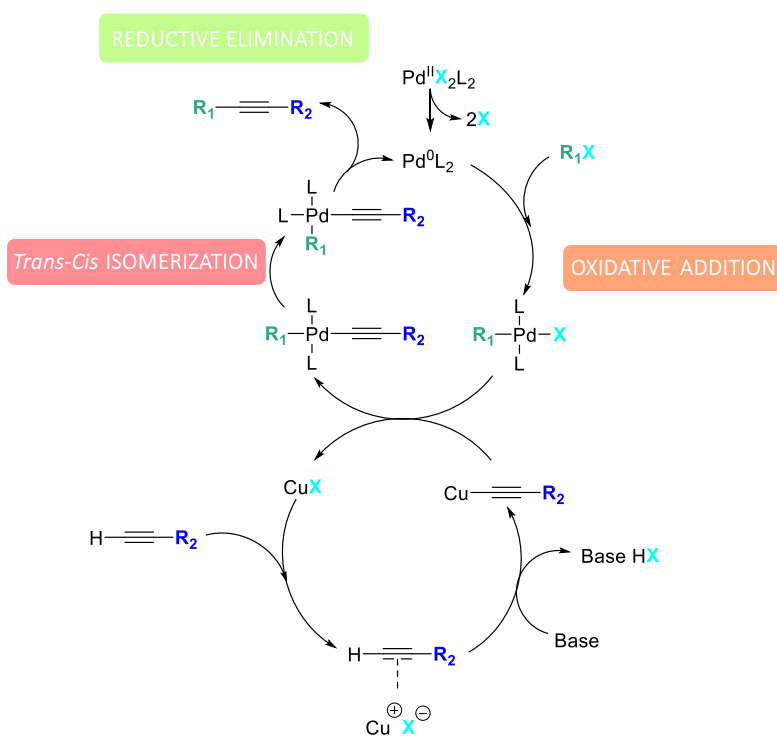
3.1.2.2. Preparation of 4-amine-6-chloro-5-(phenylethynyl)pyrimidine (**53**) and 4-amine-6-chloro-5-(4-chlorophenyl)ethynyl)pyrimidine (**54**)

The alkynylation of halogenated derivatives was carried out by applying the Sonogashira reaction conditions.^{113,114,115,116} This reaction employs acyl halides and acetylenes in the presence of palladium catalysers, copper salts and base (Scheme 6).



Scheme 6

A possible mechanism for this cross-coupling reaction is described below, where the involvement of the two types of catalysers, palladium and copper derivative is summarised (Scheme 7).^{113,114,115}



Scheme 7

¹¹³ K. Sonogashira, Y. Tohda, N. Hagihara. *Tetrahedron Lett.* **1975**, *16*, 4464-4470.

¹¹⁴ D. H. Lee, H. Qiu, M. H. Cho, I. M. Lee, M. J. Jin. *Synlett* **2008**, 1657-1660.

¹¹⁵ C. W. D. Gallop, M. T. Chen, O. Navarro. *Org. Lett.* **2014**, *16*, 3724-3727.

¹¹⁶ S. Handa, J. D. Smith, Y. Zhang, B. S. Takala, F. Gallau, B. H. Lipshutz. *Org. Lett.* **2018**, *20*, 542-545.

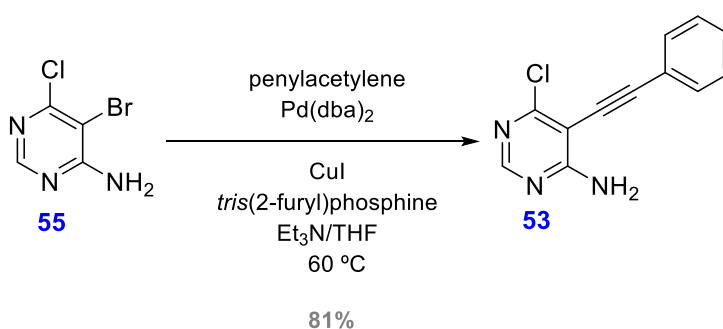
Regarding the palladium cycle it occurs as follows:

- Palladium (II) is reduced to palladium (0) by losing the two palladium substituents (halogens or acetates).
- Then, an oxidative addition of the corresponding aryl halide to the palladium-ligand complex occurs.
- Next step is transmetalation where the copper acetylide is inserted on the palladium complex.
- Organic ligands, that were initially *trans*-oriented, are isomerized to the corresponding *cis* isomers.
- The *cis* complex undergoes a reductive elimination, leads to the C-C bond. Palladium is generated once more which converts into palladium (0), thus the cycle begins again.

Regarding the copper cycle, the following transformation occurs:

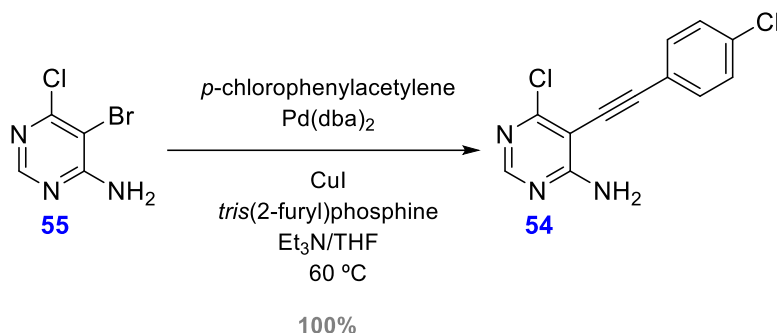
- The formation of a π -alkyl complex makes the proton of the acetylene derivative more labile, thus more likely to be ripped off by the added base (Et_3N).
- The alkyl-copper complex formed reacts with palladium complex (generated in the palladium cycle) and copper salt is regenerated.

In this case, the bromochloropyrimidine **55** was treated with phenylacetylene under the Sonogashira reaction conditions (Scheme 8), and so obtaining **53** with **81%** yield.



Scheme 8

These same conditions were applied in the cross-coupling reaction between bromopyrimidine **55** and *p*-chlorophenylacetylene leading to **54** with **100%** yield (Scheme 9).



Scheme 9

The presence in the ^1H NMR spectrum of a singlet at 8.14 and a doublet at 7.43 ppm (signals corresponding to protons in position 2 and positions 2' and 6') allowed to confirm the proposed structure for **53**, whereas the two doublets that integrated two protons each, which appeared in the proton spectrum of **54**, and were assigned to protons H-2', H-6' (7.45 ppm) and H-3', H-5' (7.30 ppm) allowed to confirm the structure of alkynyl **54** (Figure 80).

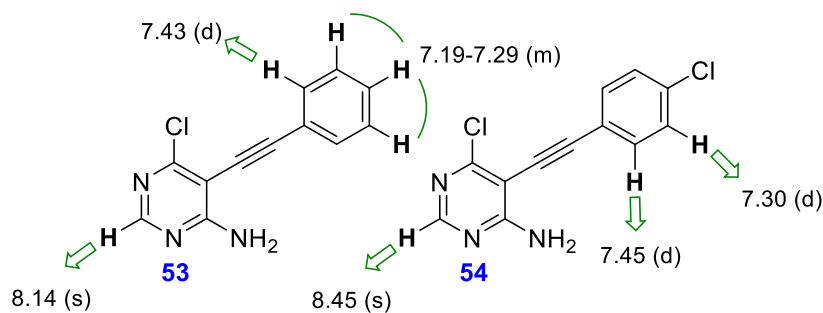
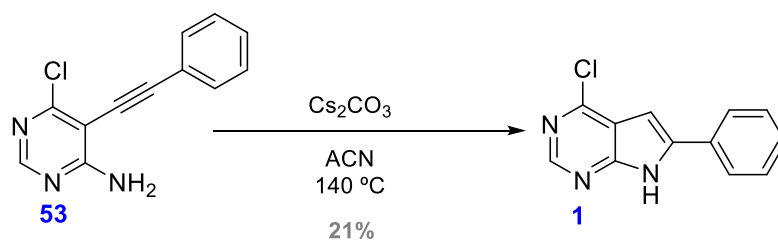


Figure 80. ^1H NMR spectral data of **53** and **54**

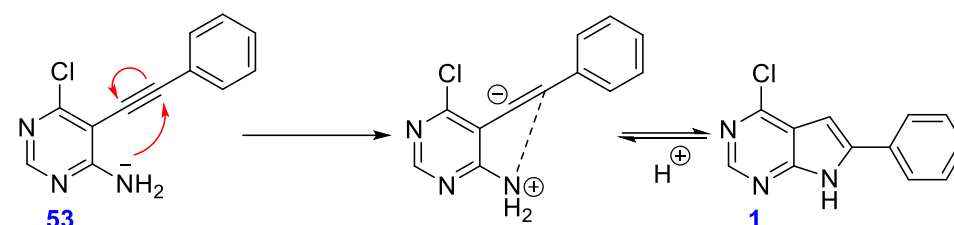
3.1.2.3. Preparation of 4-chloro-6-phenyl-7H-pyrrolo[2,3-d]pyrimidine (**1**) and 4-chloro-6-(4-chlorophenyl)-7H-pyrrolo[2,3-d]pyrimidine (**2**)

The treatment of the alkyne **53** with caesium carbonate in acetonitrile heated at 140 °C allowed the intramolecular cyclization and led to pyrrolo[2,3-d]pyrimidine **1** with **21%** yield after purification (Scheme 10).



Scheme 10

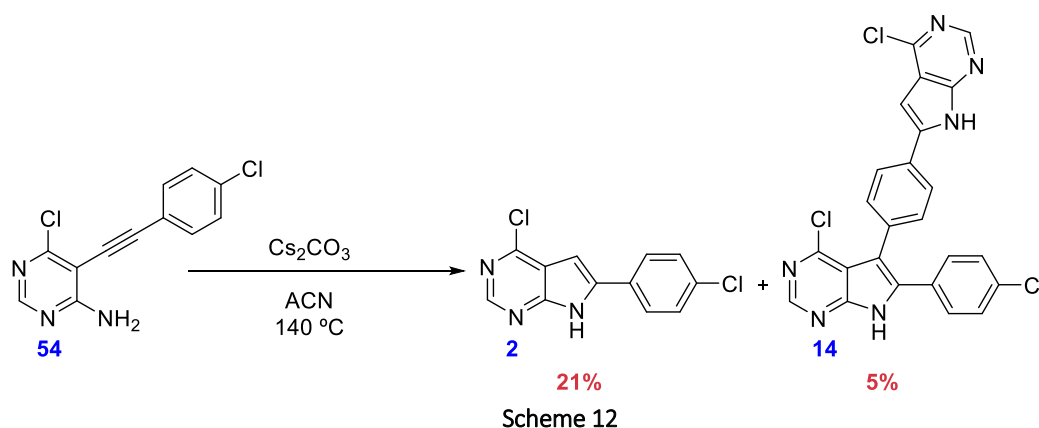
The addition of the amine to the alkyne occurred by means of intramolecular cyclization of a 5-*endo-dig* type (Baldwin's rules).¹¹⁷ The amino group of **53** was added to the triple bond, and there was no expulsion of any outgoing group (*endo*). In this case the electrophilic carbon was sp hybridised (*dig*: diagonal) (Scheme 11).



Scheme 11

This reaction was carried out without the presence of catalysts; only the addition of a base such as caesium carbonate was required. Initially CuI was added as a catalyst but, ultimately, it showed that without a catalyst the yields were even better.

Treatment of the alkyne **54** under the same conditions applied to **53** led to pyrolopyrimidine **2** with 21% yield (Scheme 12).



Scheme 12

¹¹⁷ J. E. Baldwin, *J. Chem. Soc. Chem. Commun.* **1976**, 734-736.

In this reaction in addition to pyrrolopyrimidine **2**, derivative **14** was obtained with **5%** yield. The mass found was 491.0358 and the calculated for this compound **14** was 491.0267.

The presence of singlets in the ^1H NMR spectrum at 7.10 and 8.15 ppm corresponding to protons at positions 2 and 5 allowed to confirm the structure of pyrrolopyrimidine (**1**), whereas in the ^1H NMR spectrum of **2**, the protons at positions 2 and 5 resonated at 8.47 and 6.28 ppm (Figure 81).

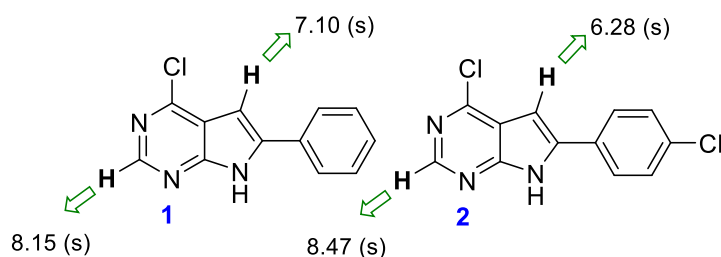
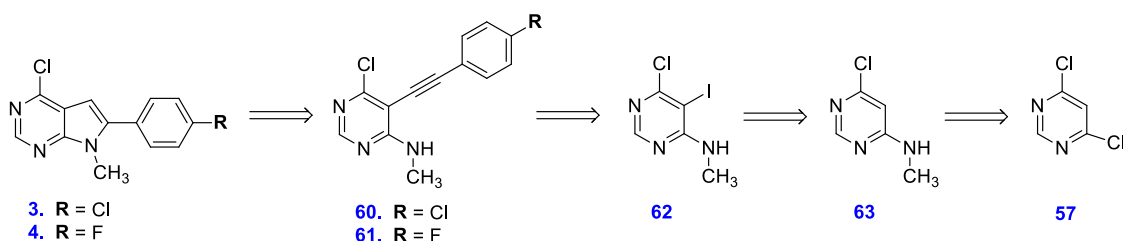


Figure 81. ^1H NMR spectral data of **1** and **2**

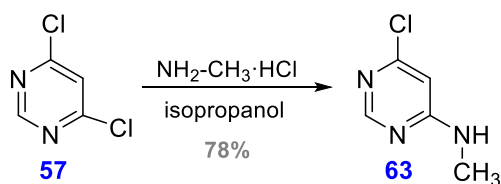
3.1.2.4. Preparation of 4-chloro-6-(4-chlorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidine (**3**) and 4-chloro-6-(4-fluorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidine (**4**)

The retrosynthetic analysis for pyrrolo[2,3-d]pyrimidines **3** and **4** is indicated below (Scheme 13).



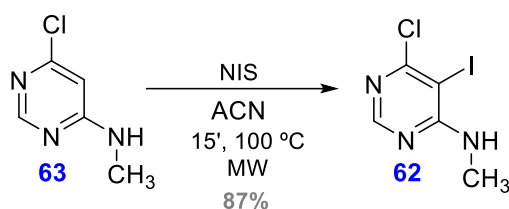
Scheme 13

Treatment of the dichloropyrimidine **57** with excess of methylamine in isopropanol, led to methylaminopyrimidine **63** with **78%** yield (Scheme 14).



Scheme 14

Then, pyrimidine **63** by NIS treatment in acetonitrile assisted by microwave irradiation, led to iodine-derivative **62** with **87%** yield (Scheme 15).



Scheme 15

The presence of a doublet in the ^1H NMR spectrum at 3.06 ppm (-NH-CH₃) and a singlet at 8.24 ppm (H-2) allowed to identify and confirm the proposed structure for compound **62** (Figure 82).

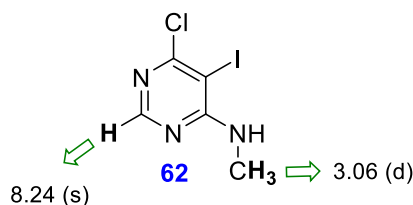
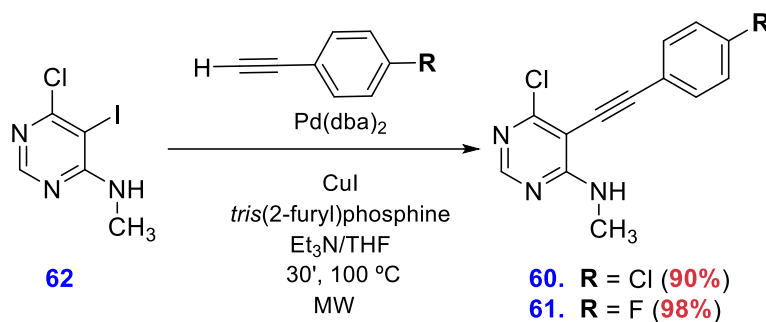


Figure 82. ^1H NMR spectral data of **62**

Then, the reaction between pyrimidine **62** and the corresponding phenylacetylene derivative using the Sonogashira cross-coupling reaction conditions, led to alkynes **60** and **61** (Scheme 16).



Scheme 16

The presence of signals in the ^1H NMR spectra of compounds **60** and **61** confirmed both structures 3.13 (s, -NH-CH₃); 7.37 ppm (d, $J = 8.8$ Hz, 2H, H-3', H-5'); 7.48 (d, $J = 8.8$ Hz, 2H, H-2', H-6') and 8.36 ppm (s, 1H, H-2) for compound **60** and 3.13 ppm (d, $J = 8.8$ Hz, 3H, -NH-CH₃); 7.09 ppm (d, $J = 8.8$ Hz, 2H, H-3', H-5'); 7.54 ppm (d, $J_1 = 8.8$ Hz, $J_2 = 5.3$ Hz, 2H, H-2', H-6') and 8.36 (s, 1H, H-2) for substituted pyrimidine **61** (Figure 83).

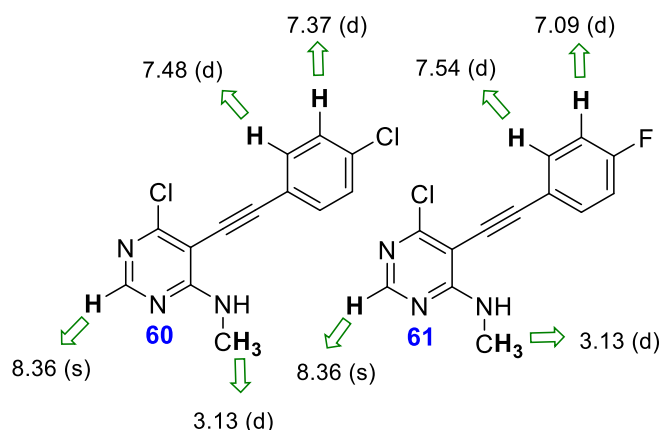
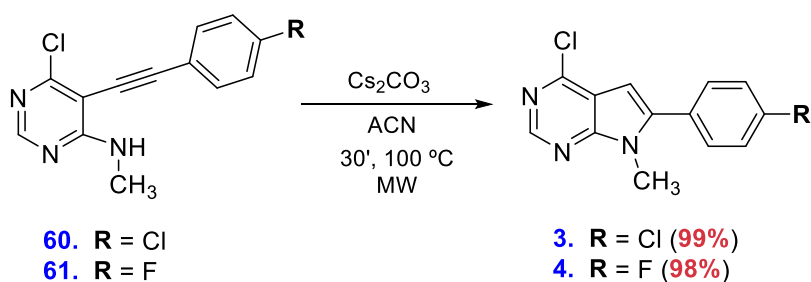


Figure 83. ^1H NMR spectral data of alkynes **60** and **61**

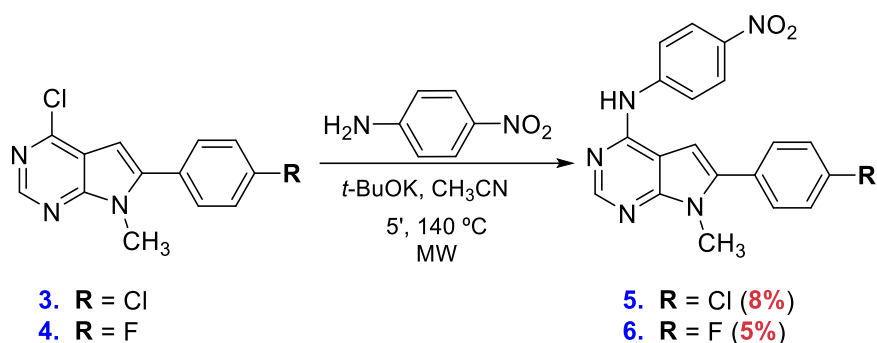
Following the methodology previously mentioned, the derivatives **60** and **61** by caesium carbonate treatment assisted by microwave irradiation led to the expected pyrrolo[2,3-*d*]pyrimidines **3** and **4** in high yields (Scheme 17).



Scheme 17

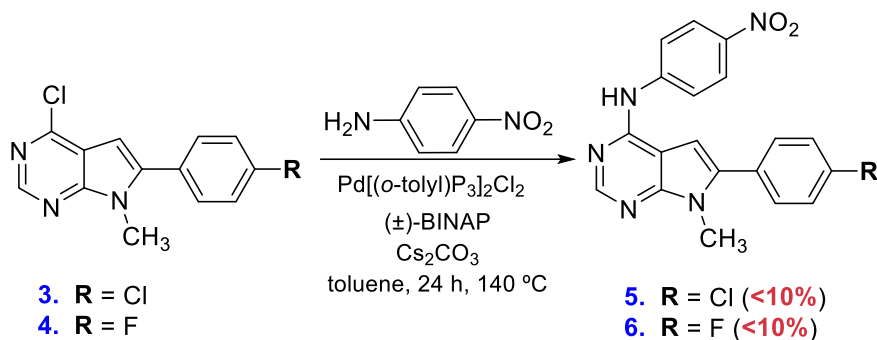
3.1.2.5. Preparation of 4-amine-6-(4-chlorophenyl)-7-methyl-N-(4-nitrophenyl)-7H-pyrrolo[2,3-*d*]pyrimidine (**5**), 4-amine-6-(4-fluorophenyl)-7-methyl-N-(4-nitrophenyl)-7H-pyrrolo[2,3-*d*]pyrimidine (**6**), 4-(4-aminophenylamine)-6-(4-chlorophenyl)-pyrrolo[2,3-*d*]pyrimidine (**7**) and 4-(4-aminophenylamine)-6-(4-fluorophenyl)-pyrrolo[2,3-*d*]pyrimidine (**8**)

Treatment of 4-chloropyrrolo[2,3-*d*]pyrimidines **3** and **4** with *p*-nitroaniline under classic $\text{S}_{\text{N}}\text{Ar}$ conditions, using *t*-BuOK as a base, in acetonitrile medium, and operating with microwave irradiation assistance, the expected products **5** and **6** were obtained respectively with very low yields (Scheme 18).



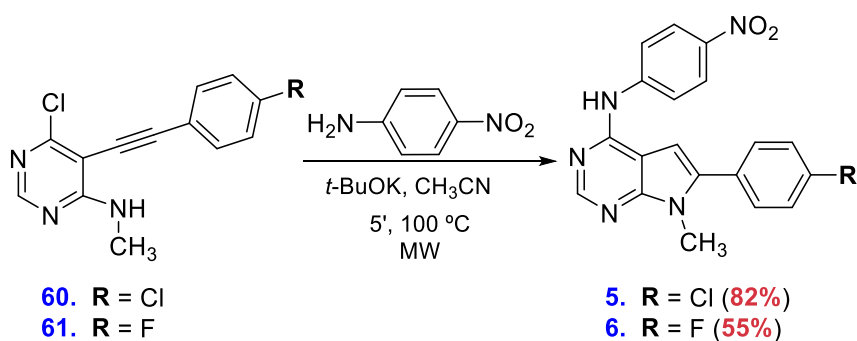
Scheme 18

The same amination was tested under cross-coupling conditions using palladium catalyst, BINAP as the ligand, caesium carbonate as the base, and heating the reaction mixture to 140 °C for 24 hours, but under these conditions, no better yields were obtained, and the purification of the crude reaction was even more difficult (Scheme 19).



Scheme 19

The above results lead to trying other conditions and, among them to highlight the addition of 4-nitroaniline to the **60** and **61** alkyne intermediates. This reaction was performed under microwave irradiation using *t*-BuOK as a base, and CH₃CN as a solvent. Under these reaction conditions, the pyrrolo[2,3-*d*]pyrimidines **5** and **6** were obtained with good yields of **82** and **55%** respectively (Scheme 20).



Scheme 20

In the ^1H NMR spectrum signals appeared in the form of multiplet (8.22-8.28 ppm) assignable to the aromatic protons of the nitrophenyl group, which allowed to confirm the proposed structure for **5**. Also, the signal at 8.25 ppm (s) attributable to the nitrophenyl protons indicated the incorporation of the *p*-nitrophenylamino fragment at position 4 of pyrrolopyrimidine **6** (Figure 84).

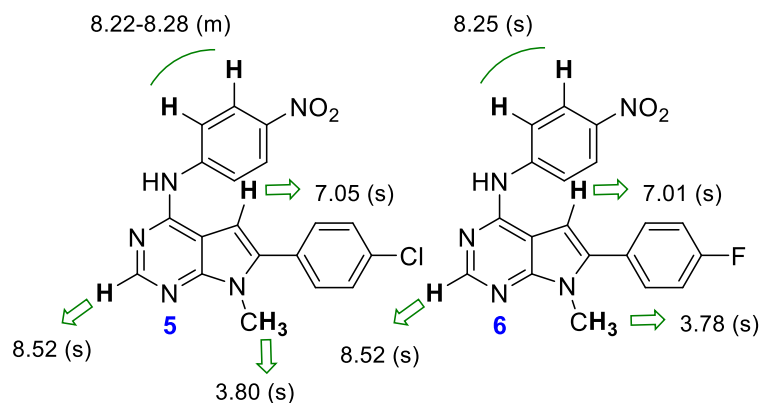
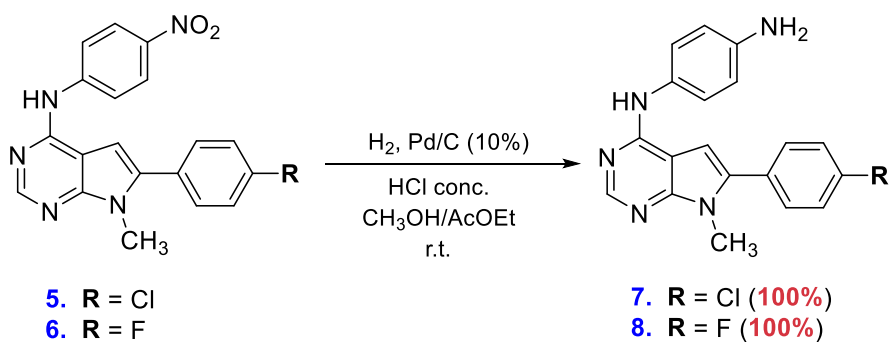


Figure 84. ^1H NMR (DMSO- d_6) spectral data of **5** and **6**

Reduction by catalytic hydrogenation of nitro-derivatives **5** and **6** allowed access to anilines **7** and **8** respectively with quantitative yields (Scheme 21).



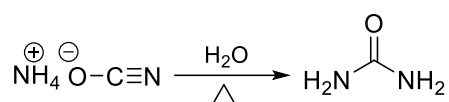
Scheme 21

3.1.2.6. Preparation of 1-(4-((6-(4-chlorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidin-4-yl)amino)phenyl)-3-(2-fluoro-5-(trifluoromethyl)phenyl)urea (**9**) and 1-(2-fluoro-5-(trifluoromethyl)phenyl)-3-(4-((6-(4-fluorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidin-4-yl)amino)phenyl)urea (**10**)

Ureas, also known as carbamide, are functions very much present in compounds possessing therapeutic activity^{118,119}, but also in industrial products since they confer an increased hydrophilicity and, at the same time, are stable products difficult to hydrolyse.

The preparation of urea by Friedrich Wöhler in 1828 from inorganic reagents was the gateway to modern chemical synthesis, and the synthesis of the first drug (Scheme 22).¹²⁰

Ammonium cyanate being heated in water, is transformed into urea by transposition.



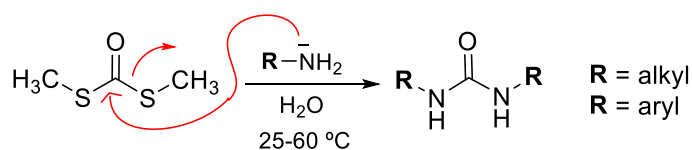
Scheme 22

Different synthetic methods for preparation of ureas are currently known. Here are some of the procedures most commonly used in the preparation of organic compounds.

- Ureas synthetic preparation

- By reaction of *S,S*-dimethyldithiocarbonate with amines

Adding amines or anilines to the *S,S*-dimethyldithiocarbonate leads to the corresponding synthetic urea, and methanethiol is generated (Scheme 23).¹²¹



Scheme 23

¹¹⁸ C. González Montero, F. J. Manso Platero, A. J. López Alba. *Avances en Odontología* ISSN, **2014**, 2340-3152.

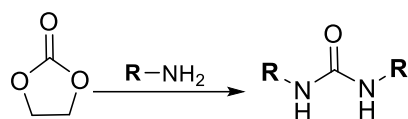
¹¹⁹ J. L. Vialuc, F. Calmel, D. Bigg, E. Carilla, A. Stenger, P. Chopin, M. Briley. *J. Med. Chem.* **1994**, 37, 689-695.

¹²⁰ P. S. Cohen, S. M. Cohen. *J. Chem. Educ.* **1996**, 73, 883.

¹²¹ M. Leung, J.-L. Lai, K.-H. Leu, H.-H. Yu, H.-J. Hsiao. *J. Org. Chem.* **1996**, 61, 4175-4179.

- By ethylencarbonate with amines

Addition of an excess of amines or anilines to the ethylencarbonate for access to symmetrical ureas. This method can lead to asymmetric ureas if the addition of the amine introduced in the first step is controlled (Scheme 24).¹²²

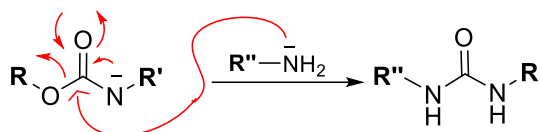


Scheme 24

- Preparation of non-synthetic ureas

- From carbonates and amines or anilines

The addition of amines or anilines to the corresponding carbonate allows the preparation of ureas.^{123,124} In many cases, having no adequate carbonate available, this preparation becomes multisteps, and requires the prior preparation of the carbonate by reaction of the amine or aniline with chloroformate in classical conditions (Scheme 25).¹²⁵



Scheme 25

- From 1,1'-carbonildiimidazole

Addition of the amine or aniline to the carbonildiimidazole leads to an intermediate that by reaction with an amine or aniline gives access to the expected urea.¹²⁶ The intermediate is stable under classical conditions and therefore allows obtaining symmetric ureas (treatment with an excess of the desired amine) or asymmetric when using two different amines (Scheme 26).

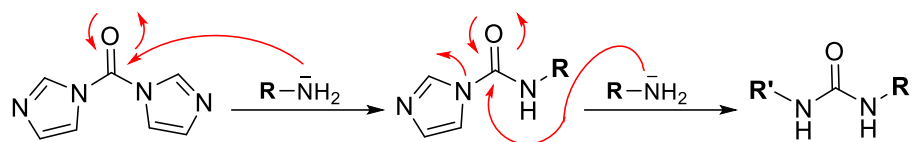
¹²² S. R. Jagtap, Y. P. Patil, A. Pauda, B. M. Bhanage. *Synth. Commun.* **2009**, *39*, 2093-2100.

¹²³ S. Kang, H.-K. Kim. *Tetrahedron* **2018**, *74*, 4036-4046.

¹²⁴ S. Hudson, M. Kiankarimi, M. W. Rowbottom, T. D. Vickers, D. Wu, J. Pontillo, B. Ching, W. Dwight, V. S. Goodfellow, D. Schwarz, C. E. Heise, A. Madan, J. Wen, W. Ban, H. Wang, W. S. Wade. *Biorg. Med. Chem.* **2006**, *16*, 4922-4930.

¹²⁵ a) Z. Hodnik, L. P. Masic, T. Tomasic. *J. Med. Chem.* **2014**, *57*, 4819-4833. b) B. Romano, D. Plano, I. Encio, J. A. Palop. *Biorg. Med. Chem.* **2015**, *23*, 1716-1727.

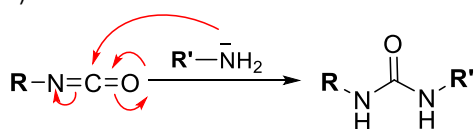
¹²⁶ a) D. Ke, C. Zhan, X. Li, A. D. Q. Li, J. Yao. *Tetrahedron* **2009**, *65*, 8269-8276. b) M. D. Mc Reynolds, K. T. Sprott, P. R. Hanson. *Org. Lett.* **2002**, *4*, 4673-4676.



Scheme 26

- From isocyanates and amines or anilines

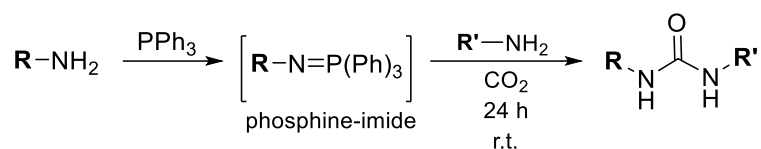
The addition of amines or anilines to isocyanates allows access to ureas in a single step.^{127,128} In general, it is a good method for ureas preparation, since it does not need any additives or catalysts, just a solvent such as THF or dichloromethane and the stirring of the reagents at room temperature (Scheme 27).



Scheme 27

- Through a phosphine-imide

The reaction between primary amines and triphenylphosphine (TPP) in basic medium (Et₃N), within dichloromethane, leads to a phosphine-imide intermediate that, by applying a carbon dioxide stream for 24 h leads to the expected urea (Scheme 28).¹²⁹



Scheme 28

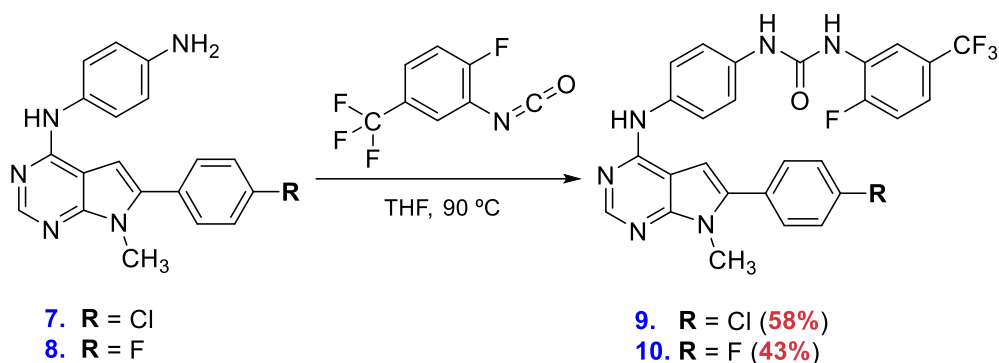
In this research work and for the preparation of ureas **9** and **10**, the above indicated methodology of adding amines to isocyanates was followed. In this case, aniline **7** was condensed with 2-fluoro-5-trifluorophenylisocyanate and the expected urea was obtained with **58%** yield (Scheme 29).

¹²⁷ E. L. Luzina, A. V. Popov. *Eur. J. Med. Chem.* **2009**, *44*, 4944-4953.

¹²⁸ J. Clayden, L. Lemiègre, M. Helliwell. *J. Org. Chem.* **2007**, *72*, 2302-2308.

¹²⁹ S. Porwanski, S. Menuel, X. Marsura, A. Marsura. *Tetrahedron Lett.* **2004**, *45*, 5027-5029 (and references cited therein).

Following the same procedure, the addition of aniline **8** to 2-fluoro-5-trifluorophenylisocyanate, the urea **10** was successfully isolated with **43%** yield (Scheme 29).



Scheme 29

The ^1H NMR spectrum signals of urea **9** (7.36-7.38 ppm (m, 2H, H-3''', H-4''') and 7.40 (s, 1H, H-2''')) corresponding to the phenyl group incorporated in the formation of the urea allowed to confirm the proposed structure for **9**, whereas the signals (7.22-7.26 (m, 1H, H-3'''), 7.33 (d, 1H, H-4''') and 7.27 (s, 1H, H-2''')) attributable to the 2-fluoro-5-trifluoromethylphenyl group from the isocyanate confirmed the formation of urea **10** (Figure 85).

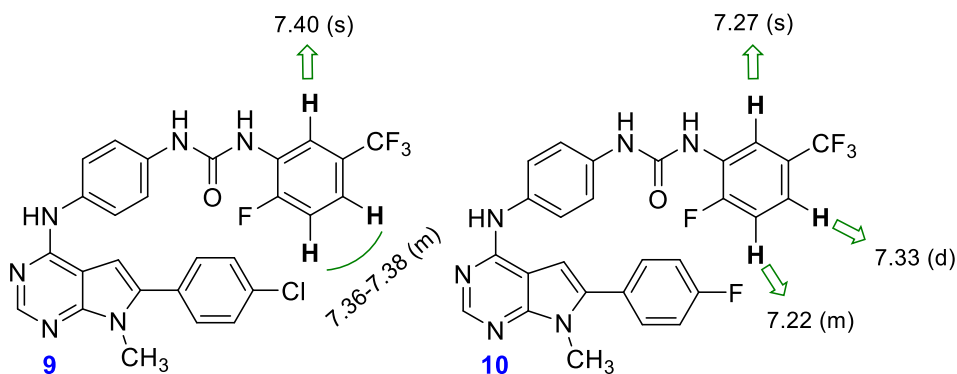
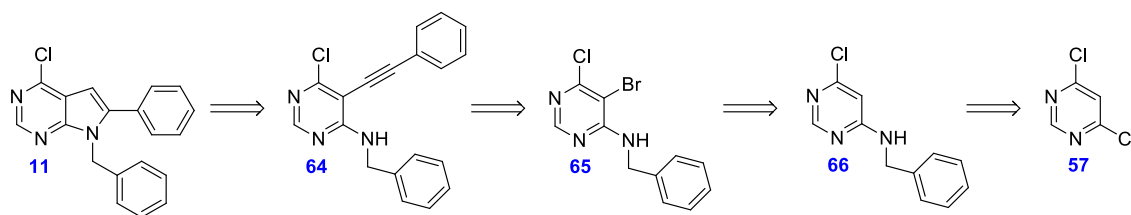


Figure 85. ^1H NMR spectroscopic data of **9** and **10**

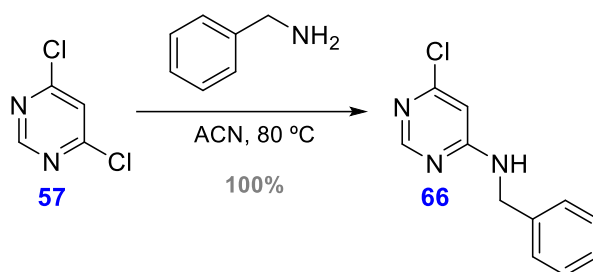
3.1.2.7. Preparation of 7-benzyl-4-chloro-6-phenyl-7H-pyrrolo[2,3-d]pyrimidine (**11**)

The retrosynthetic analysis for pyrrolo[2,3-d]pyrimidine **11** is indicated below (Scheme 30).



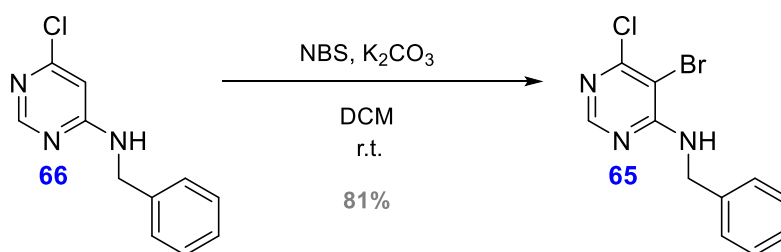
Scheme 30

Treatment of 4,6-dichloropyrimidine (**57**) with benzylamine led to monochloropyrimidine **66** with **100%** yield (Scheme 31).



Scheme 31

Then, the selective bromination of **66** using NBS, following the conditions indicated previously in this work, led to bromo derivative **65** with **81%** yield (Scheme 32).



Scheme 32

The absence of the proton signal at position 5 as well as the slight deshielding of H-2 allowed to confirm the proposed structure for **65** (Figure 86).

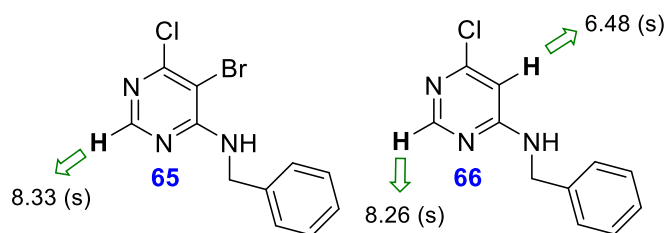
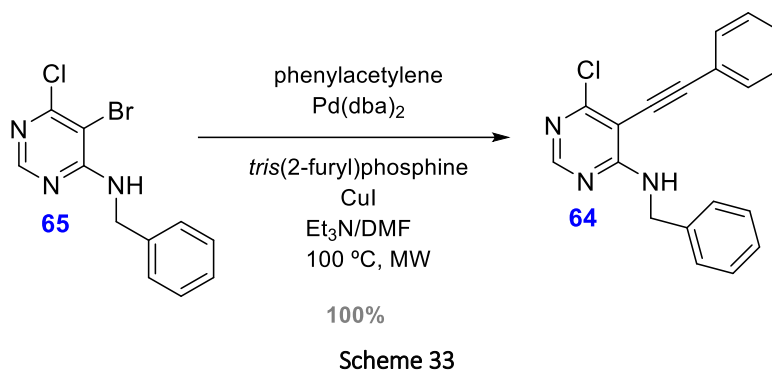
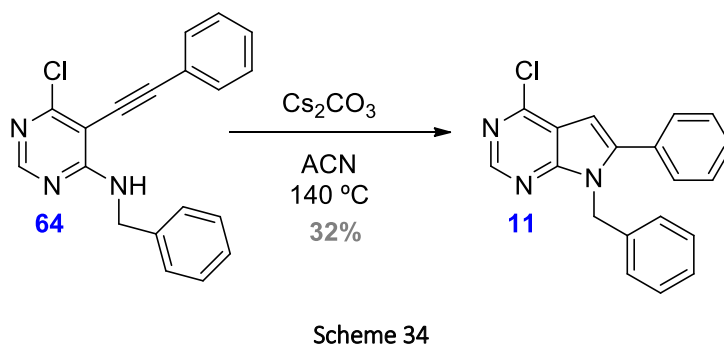


Figure 86. ^1H NMR spectroscopic data of **65** and **66**

The cross-coupling reaction between bromopyrimidine **65** and phenylacetylene using the Sonogashira reaction conditions resulted in obtaining **64** with **100%** yield (Scheme 33).



Pyrimidine **64** was then treated with caesium carbonate, within acetonitrile, and heated to 140 °C to obtain pyrrolo[2,3-*d*]pyrimidine **11** with **32%** yield (Scheme 34).



The appearance of a singlet at 6.66 ppm in the ^1H NMR spectrum corresponding to the proton at position 5 indicated that the *endo-dig* cyclization were successfully carried out to give the expected pyrrolo[2,3-*d*]pyrimidine system (Figure 87).

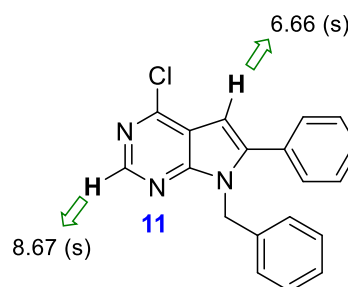
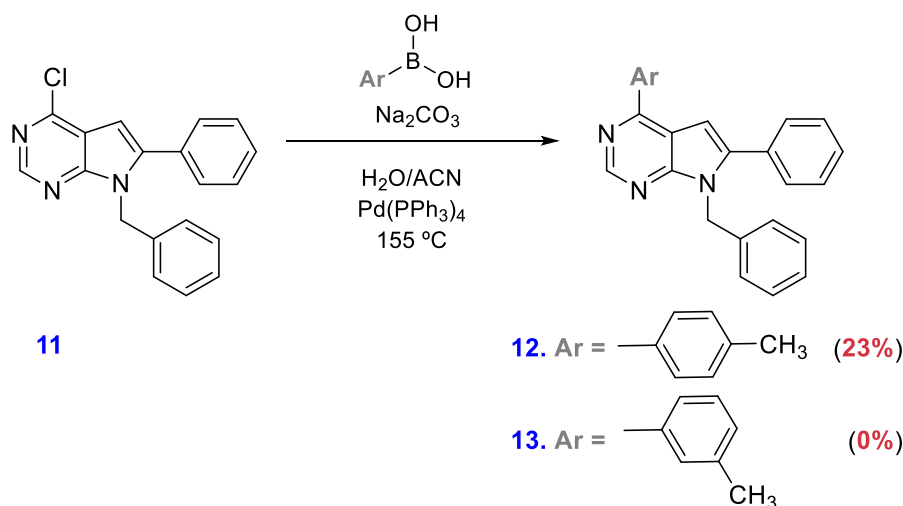


Figure 87. ^1H NMR spectroscopic data of **11**

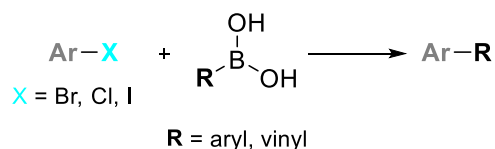
3.1.2.8. Preparation of 7-benzyl-6-phenyl-4-(*p*-tolyl)-7H-pyrrolo[2,3-*d*]pyrimidine (**12**) and 7-benzyl-6-phenyl-4-(*m*-tolyl)-7H-pyrrolo[2,3-*d*]pyrimidine (**13**)

Arylation of 4-chloropyrrolo[2,3-*d*]pyrimidine **11** led to diaryl derivatives **12** and **13** (Scheme 35).



Scheme 35

For the formation of the C-C bond leading to **12** and **13**, the Suzuki-Miyaura cross-coupling reaction was chosen, which involves the reaction of an aryl halide with boronic acid or derivatives, and palladium-phosphine complex as catalyst (Scheme 36).^{130,131}



Scheme 36

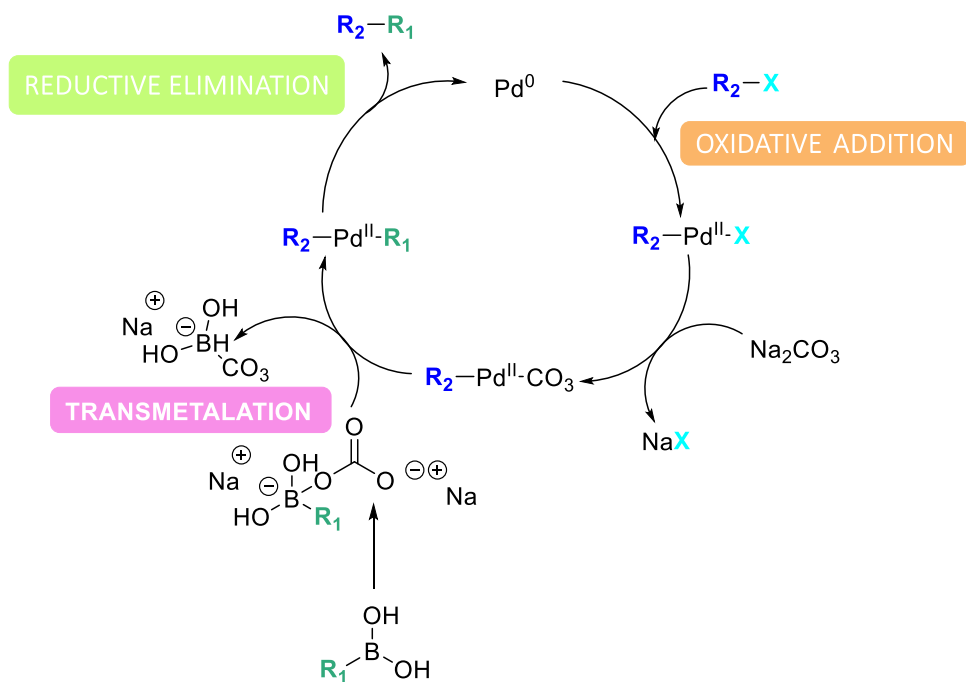
The Suzuki-Miyaura reaction is one of the most widely used coupling reactions in the synthesis of both organic academic and industrial compounds.¹³² This reaction has presented certain advantages over other organometallic reactions presented above. First, it is worth highlighting the mild conditions that make reactivity compatible with various substituents that may be present in the reagents. It should also be considered that the catalyst is a palladium complex used in very small quantities and can be recovered at the end of the reaction. Likewise, boronic acid derivatives with lower toxicity

¹³⁰ A. Suzuki. *Angew. Chem. Int. Ed.* **2011**, *50*, 6722-6737.

¹³¹ N. Miyaura, A. Suzuki. *Chem. Rev.* **1995**, *95*, 2457-2483.

¹³² J. P. Corbet, G. Mignani. *Chem. Rev.* **2006**, *106*, 2651-2710.

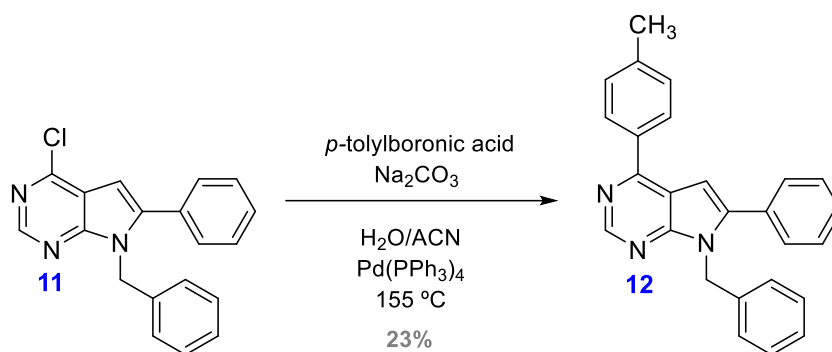
than tin or nickel derivatives of other reactions are also used. The mechanism for this reaction was proposed by A. Suzuki himself and is indicated below (Scheme 37).¹³⁰



Scheme 37

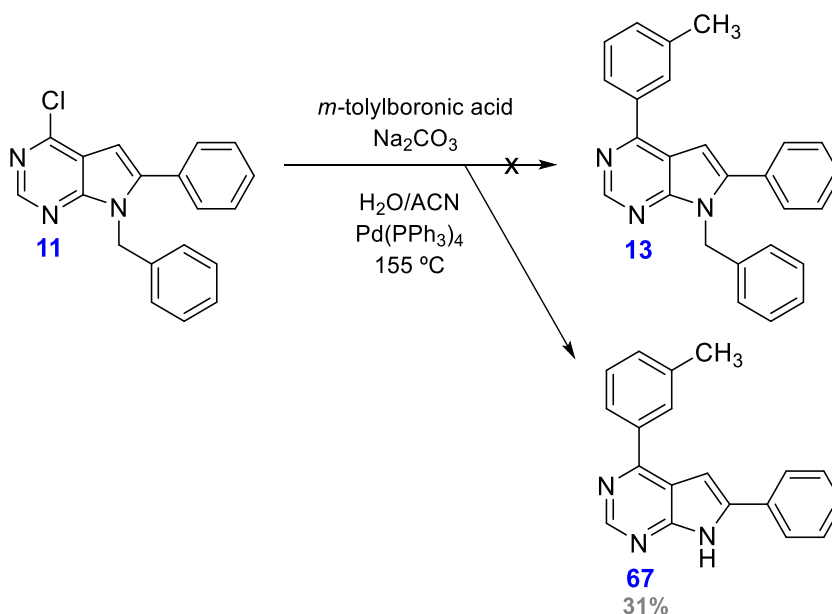
- Palladium (II) is reduced to palladium (0) by the loss of the two initial substituents.
- Palladium (0) is inserted into the Ar-X compound and again oxidized to palladium (II).
- With the help of the base, the boronic acid in the transmetalation stage is converted into salt, and so increases the nucleophilic power of the R group, thus facilitating its transfer to the palladium derivative.
- Finally, the C-C bond is formed by a reductive elimination step that allows the palladium (0) to be recovered, and so the cycle can begin again.

Applying the Suzuki-Miyaura conditions, the preparation of **12** was achieved with **23%** of yield (Scheme 38).



Scheme 38

On the other hand, the coupling between **11** and *m*-tolylboronic acid did not allow obtaining the expected compound **13**. Instead, the derivative debenzylated **67** was isolated with **31%** yield (Scheme 39).



Scheme 39

3.1.3. Preparation of Combretastatin A4 (CA4) derivatives

In this work, a series of analogues structural compounds of *Combretastatin A4* also containing two aromatic rings were prepared, and they maintained the *cis* configuration.

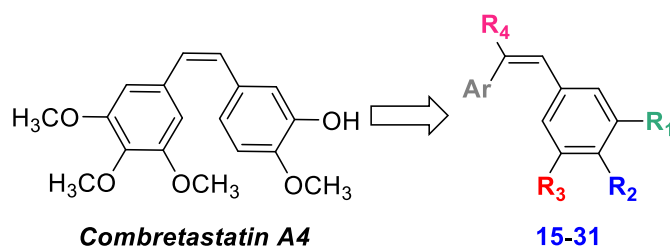


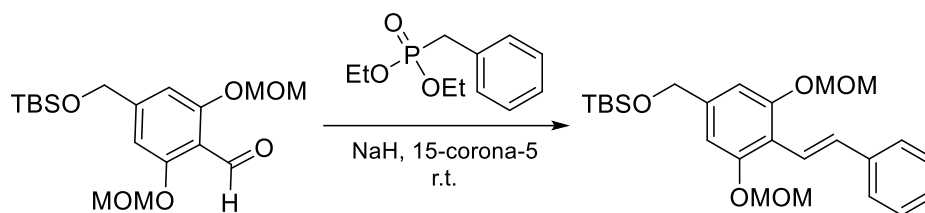
Figure 88. CA4 derivatives

3.1.3.1. Preparation of 15-31

Different methods for the preparation of stilbenes have been described in the literature and from these the following have been selected:

- Horner-Wadsworth-Emmans (HWE) method modified¹³³

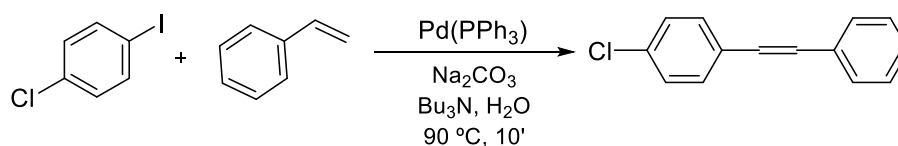
In this case, the appropriate aldehyde is condensed with the corresponding phosphonate to provide the expected stilbene in a single step (Scheme 40).



Scheme 40

- Applying Heck reaction conditions

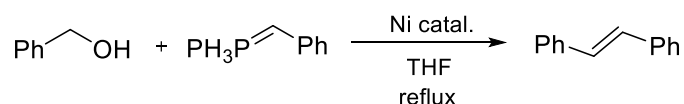
It is a cross-coupling reaction between the halogenated derivative and the vinylbenzene under Heck's reaction conditions,¹³⁴ using a Pd (0) derivative as a catalyst in the presence of sodium carbonate which acts as the base (Scheme 41).



Scheme 41

- Through the olefination of benzyl alcohol and phosphoranes using the modified Wittig method

The reaction between benzyl alcohols and phosphorus ylides using nickel nanoparticles as a catalyst, allows access to stilbene-type derivatives with high yields (Scheme 42).¹³⁵



Scheme 42

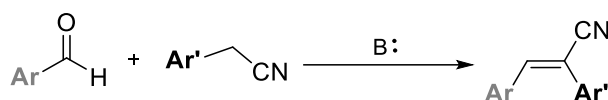
¹³³ C. H. Kuder, J. D. Neighbors, R. J. Hohl, D. F. Wiemer. *Bioorg. Med. Chem.* **2009**, *17*, 4718-4723.

¹³⁴ a) Y. Xu, S. Xu, L. Zhang, F. Zhang, Y. Ceui. *Lingxi Huagong* **2008**, *25*, 1083-1086. b) J. Tsuji. *Palladium Reagents and Catalysts: Innovations in Organic Synthesis*. Ed. Wiley: Chichester, **2004**. c) R. F. Heck. *J. Am. Chem. Soc.* **1968**, *90*, 5518-5526. d) R. F. Heck, J. P. Nolley. *J. Org. Chem.* **1972**, *37*, 2320-2322.

¹³⁵ F. Alonso, P. Riente, M. Yus. *Eur. J. Org. Chem.* **2009**, 6034-6042.

- By Knoevenagel condensation

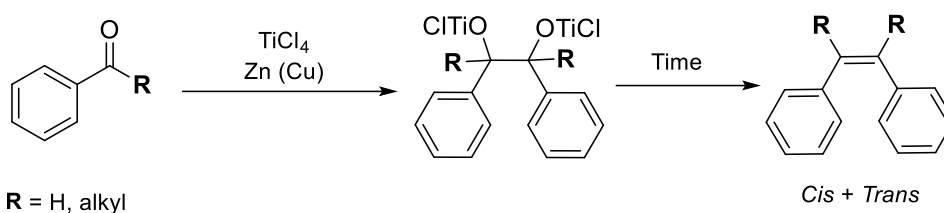
Condensation of an aldehyde with the corresponding aryl-acetonitrile, in basic medium (Et₃N, piperidine, NaH, KOH, *t*-BuOK), gives the expected olefin in low yields. This involves the addition of the nucleophile formatted “*in situ*” in the corresponding aldehyde, and a mixture of compounds of configuration (*E*) and (*Z*) is normally obtained (Scheme 43).¹³⁶



Scheme 43

- McMurry's method

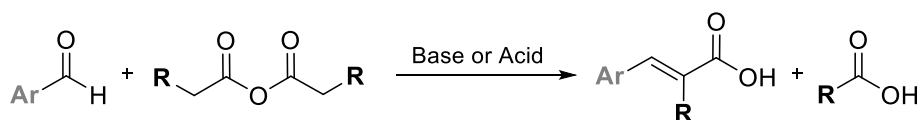
It is a coupling between carbonyl compounds, aldehydes, or ketones that leads to alkenes (Scheme 44).¹³⁷



Scheme 44

- Perkin type method

In this case it is the reaction between an aldehyde and an acid to provide the α,β -unsaturated carboxylic acid. This condensation can be carried out in an acid medium and also in a basic medium (Scheme 45).^{138,139}



Scheme 45

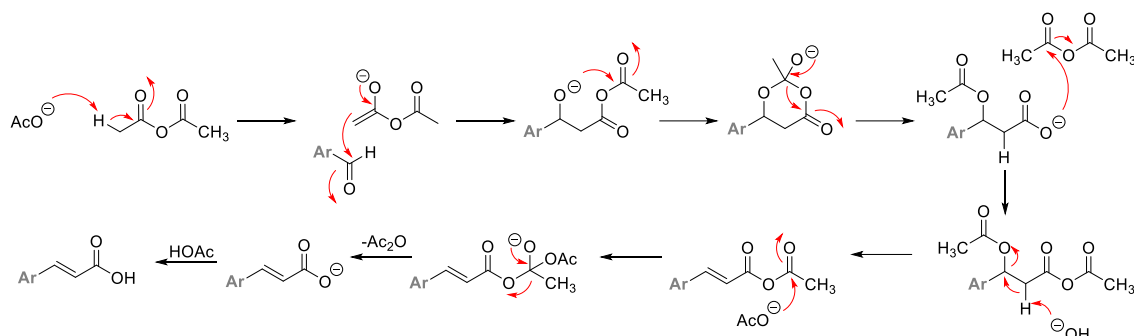
¹³⁶ N. Taha, Y. Sasson, M. Chidambaram. *Appl. Catal. A. General* **2008**, 217-224.

¹³⁷ J. E. McMurry, M. P. Fleming. *J. Am. Chem. Soc.* **1974**, 96, 4708-4709.

¹³⁸ a) W. H. Perkin. *J. Chem. Soc.* **1868**, 21, 53-61. b) G. Soladié, Y. Pooturel-Jocopé, J. Maignan. *Tetrahedron* **2003**, 59, 18, 3315-3321.

¹³⁹ W. H. Perkin. *J. Chem. Soc.* **1877**, 32, 660-674.

A probable mechanism for the indicated Perkin-type condensation reaction is outlined in detail below (Scheme 46).^{138,139}



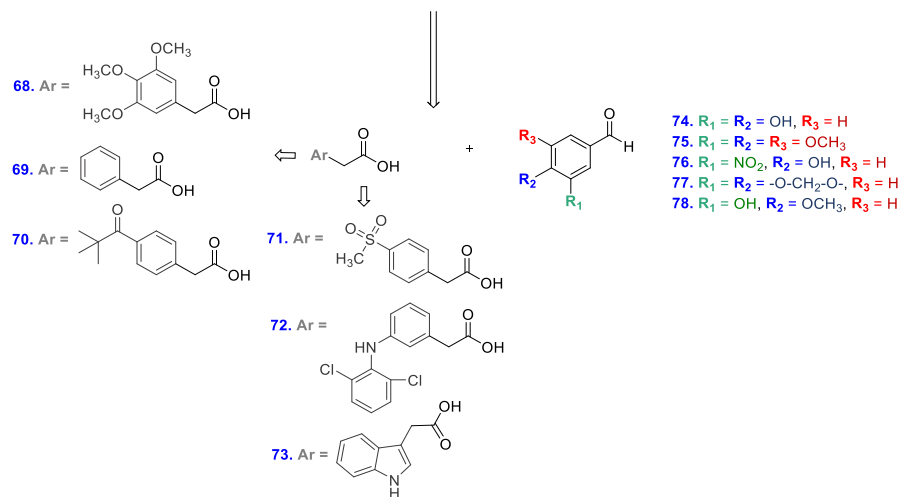
Scheme 46

In this thesis work, the preparation of the stilbene derivatives by applying the conditions of the Perkin type reaction was carried out.

The stilbene derivatives contain a carboxylic function and were prepared from the corresponding aldehydes and arylacetic acids as indicated in the following retrosynthetic analysis scheme (Scheme 47).



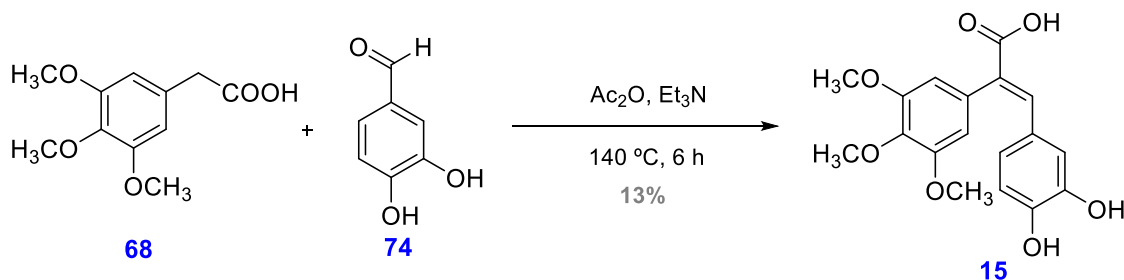
- | | |
|---|---|
| 16. Ar = 3,4,5-trimethoxyphenyl, $R_1 = R_2 = \text{OH}$, $R_3 = \text{H}$ | 15. Ar = 3,4,5-trimethoxyphenyl, $R_1 = R_2 = \text{OH}$, $R_3 = \text{H}$ |
| 18. Ar = 3,4,5-trimethoxyphenyl, $R_1 = R_2 = R_3 = \text{OCH}_3$ | 17. Ar = 3,4,5-trimethoxyphenyl, $R_1 = R_2 = R_3 = \text{OCH}_3$ |
| 20. Ar = 3,4,5-trimethoxyphenyl, $R_1 = \text{NO}_2$, $R_2 = \text{OH}$, $R_3 = \text{H}$ | 19. Ar = 3,4,5-trimethoxyphenyl, $R_1 = \text{NO}_2$, $R_2 = \text{OH}$, $R_3 = \text{H}$ |
| 22. Ar = 3,4,5-trimethoxyphenyl, $R_1 = R_2 = -\text{O}-\text{CH}_2-\text{O}-$, $R_3 = \text{H}$ | 21. Ar = 3,4,5-trimethoxyphenyl, $R_1 = R_2 = -\text{O}-\text{CH}_2-\text{O}-$, $R_3 = \text{H}$ |
| 24. Ar = phenyl, $R_1 = \text{OH}$, $R_2 = \text{OCH}_3$, $R_3 = \text{H}$ | 23. Ar = phenyl, $R_1 = \text{OH}$, $R_2 = \text{OCH}_3$, $R_3 = \text{H}$ |
| 27. Ar = methylsulfonylphenyl, $R_1 = \text{OH}$, $R_2 = \text{OCH}_3$, $R_3 = \text{H}$ | 25. Ar = 2-(2,6-dichlorophenylamino)phenyl, $R_1 = \text{OH}$, $R_2 = \text{OCH}_3$, $R_3 = \text{H}$ |
| 29. Ar = 2-(2,6-dichlorophenylamino)phenyl, $R_1 = \text{OH}$, $R_2 = \text{OCH}_3$, $R_3 = \text{H}$ | 26. Ar = methylsulfonylphenyl, $R_1 = \text{OH}$, $R_2 = \text{OCH}_3$, $R_3 = \text{H}$ |
| 31. Ar = 3-indolyl, $R_1 = \text{OH}$, $R_2 = \text{OCH}_3$, $R_3 = \text{H}$ | 28. Ar = 2-(2,6-dichlorophenylamino)phenyl, $R_1 = \text{OH}$, $R_2 = \text{OCH}_3$, $R_3 = \text{H}$ |
| | 30. Ar = 3-indolyl, $R_1 = \text{OH}$, $R_2 = \text{OCH}_3$, $R_3 = \text{H}$ |



Scheme 47

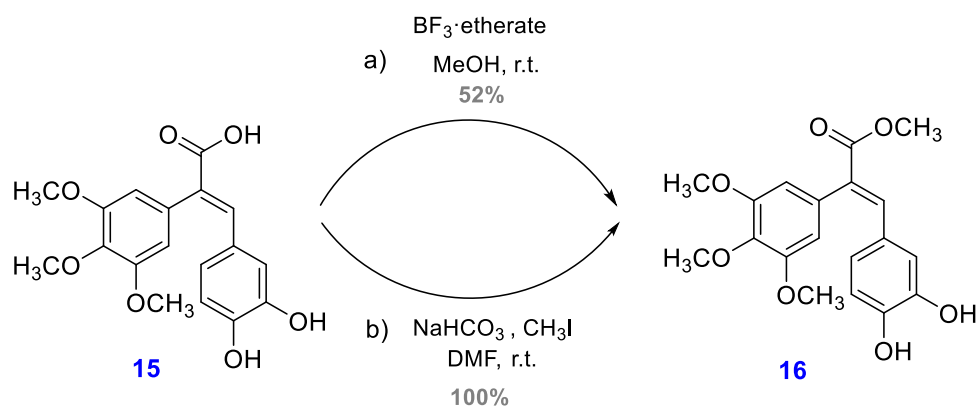
3.1.3.1.1. Preparation of (*E*)-3-(3,4-dihydroxyphenyl)-2-(3,4,5-trimethoxyphenyl)acrylic acid (**15**) and (*E*)-methyl 3-(3,4-dihydroxyphenyl)-2-(3,4,5-trimethoxyphenyl)acrylate (**16**)

The condensation of acid 3,4,5-trimethoxyphenylacetic (**68**) with 3,4-dihydroxybenzaldehyde (**74**) applying the conditions of the Perkin reaction mentioned above, allowed access to acid **15** with **13%** yield (Scheme 48).



Scheme 48

The conversion of carboxylic acid **15** to the corresponding methyl ester **16** was performed using two different methodologies. On the one hand, acid **15** was treated with methanol, which acts as a reagent and solvent, and a Lewis acid such as BF₃·etherate was added as catalyst in catalytic amount. Under these conditions, the expected methyl ester **16** was obtained with **52%** yield (Scheme 49). Acid **15** was alternatively treated with methyl iodide in DMF and in the presence of NaHCO₃ as a base. These latter conditions allow the isolation of ester **16** in quantitative yields (Scheme 49).

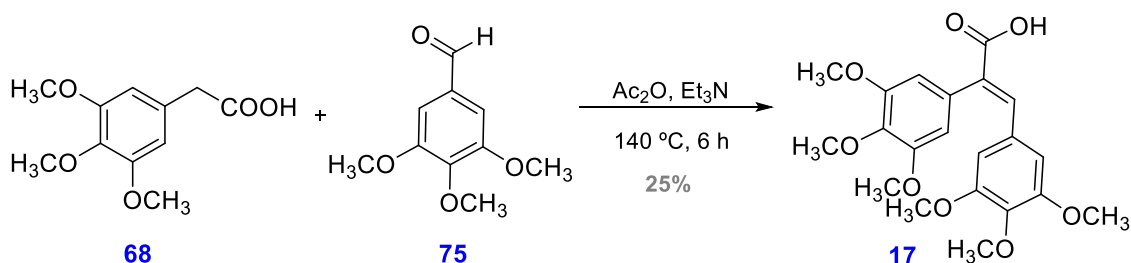


Scheme 49

It is noteworthy the high chemoselectivity of these transformations attributable to the mild conditions applied in each one of them, which make possible the alkylation exclusively of the carboxylic acid in the presence of phenol groups.

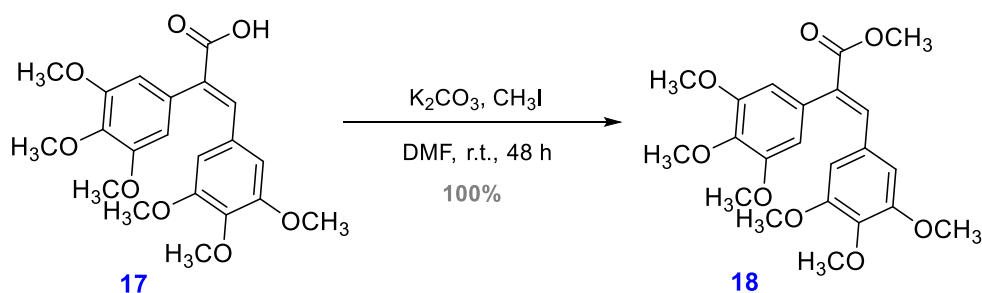
3.1.3.1.2. Preparation of (*E*)-2,3-bis(3,4,5-trimethoxyphenyl)acrylic acid (**17**) and (*E*)-methyl 2,3-bis(3,4,5-trimethoxyphenyl)acrylate (**18**)

According to the methodology of the Perkin reaction applied to the condensation of carboxylic acid **68** with aldehyde **75**, compound **17** was obtained with **25%** yield (Scheme 50).



Scheme 50

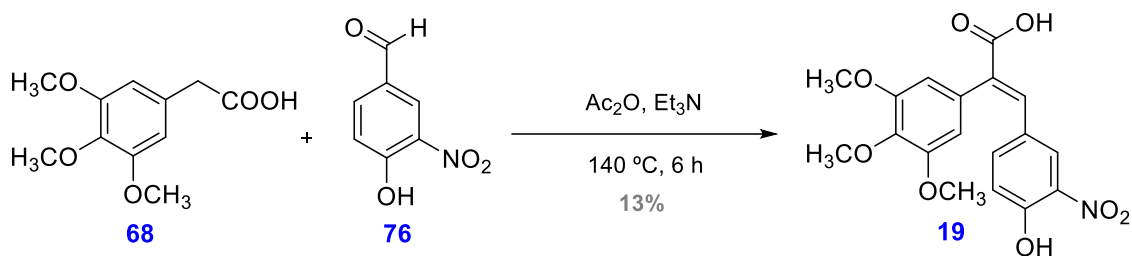
Subsequently, carboxylic acid **17** was subjected to the classical alkylation reaction by treatment with excess of methyl iodide in the presence of K₂CO₃. Under these conditions, methyl ester **18** was exclusively obtained with **100%** yield (Scheme 51).



Scheme 51

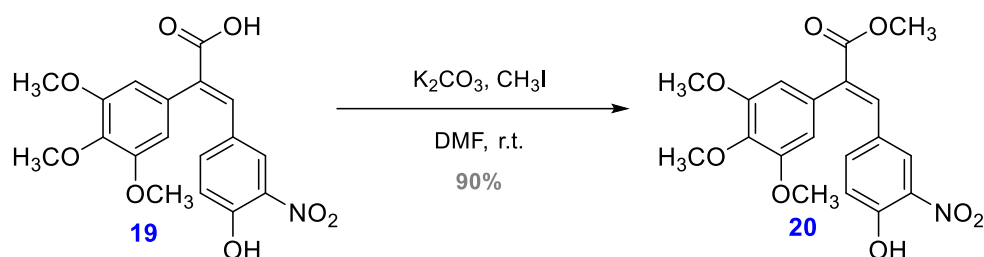
3.1.3.1.3. Preparation of (*E*)-3-(4-hydroxy-3-nitrophenyl)-2-(3,4,5-trimethoxyphenyl)acrylic acid (**19**) and (*E*)-3-methyl (4-hydroxy-3-nitrophenyl)-2-(3,4,5-trimethoxyphenyl)acrylate (**20**)

The preparation of the carboxylic acid **19** was achieved by condensation of 3,4,5-trimethoxyphenylacetic acid (**68**) with the 4-hydroxy-3-nitrobenzaldehyde (**76**) by applying the Perkin reaction conditions. The expected carboxylic acid was obtained with **13%** yield (Scheme 52).



Scheme 52

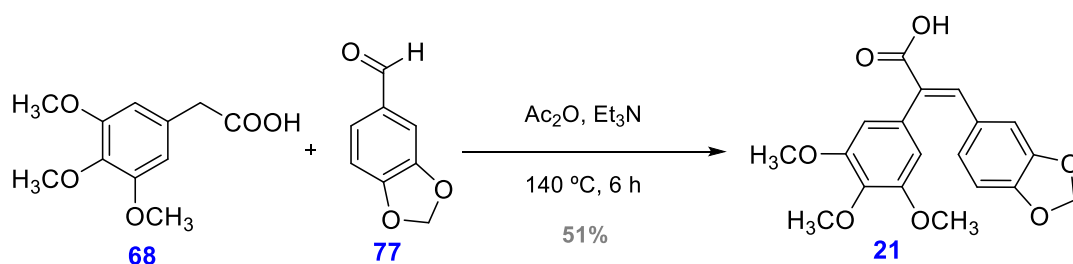
Then, by treating acid **19** with an excess of methyl iodide in the presence of potassium carbonate as a base within the DMF, the expected methyl ester **20** was obtained with **90%** yield (Scheme 53).



Scheme 53

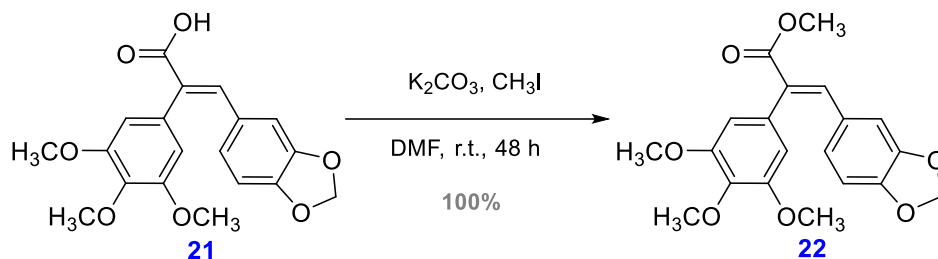
3.1.3.1.4. Preparation of (*E*)-3-(benzo[*d*][1,3]dioxol-5-yl)-2-(3,4,5-trimethoxyphenyl)acrylic acid (**21**) and (*E*)-methyl 3-(benzo[*d*][1,3]dioxol-5-yl)-2-(3,4,5-trimethoxyphenyl)acrylate (**22**)

The condensation of 3,4,5-trimethoxyphenylacetic acid (**68**) with piperonal (**77**) using the Perkin reaction conditions mentioned above in this work, allowed access to carboxylic acid **21** with **51%** yield (Scheme 54).



Scheme 54

Then, alkylation of carboxylic acid **21** was conducted by treatment with an excess of methyl iodide and in the presence of potassium carbonate in DMF. Thus the methyl ester **22** was obtained with **100%** yield (Scheme 55).



Scheme 55

The carboxylic acids (**15**, **17**, **19**, **21**) and methyl esters (**16**, **18**, **20**, **22**) mentioned above were identified by means of their ^1H NMR spectral data. Following are the most relevant data that highlight the proposed structure for each of these compounds.

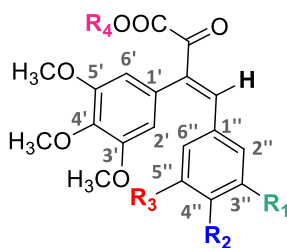
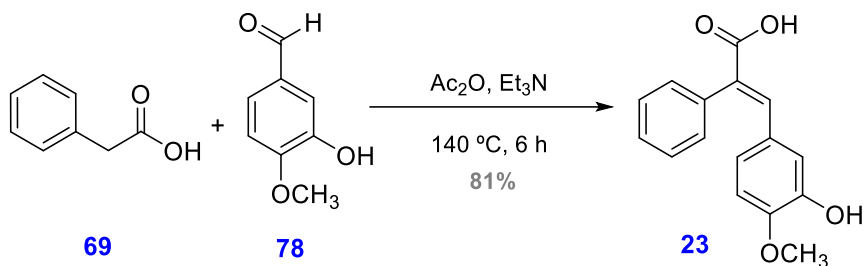


Table 17. Spectral data of different protons of acrylic acid derivatives

Compounds	H	H-2'/H-6'	H-2''	H-6''	R ₄ (H o CH ₃)
15	7.77	6.63	6.81	7.09	-
16	7.82	6.45	7.02	7.22	3.87 (CH ₃ -)
17	7.80	6.23	6.48	6.48	-
18	7.72	6.37	6.48	6.48	3.85 (CH ₃ -)
19	7.81	6.85	8.62	8.13	9.92 (H-)
20	7.70	6.50	8.35	8.09	3.79 (CH ₃ -)
21	7.85	6.46	6.68	6.81	-
22	6.84	6.41	6.62	6.73	3.89 (CH ₃ -)

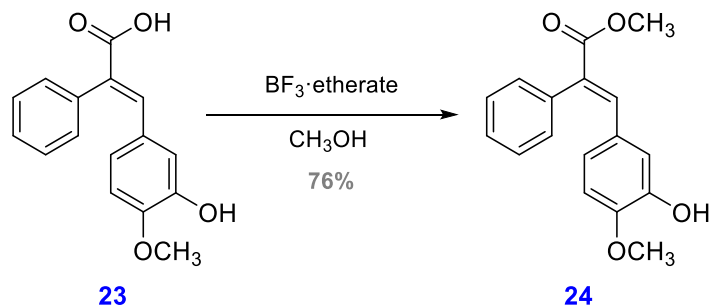
3.1.3.1.5. Preparation of (*E*)-3-(3-hydroxy-4-methoxyphenyl)-2-phenylacrylic acid (**23**) and (*E*)-methyl 3-(3-hydroxy-4-methoxyphenyl)-2-phenylacrylate (**24**)

The condensation of phenylacetic acid (**69**) with 3-hydroxy-4-methoxybenzaldehyde (**78**) under the Perkin reaction conditions mentioned in this work led to the obtaining of carboxylic acid **23** with **81%** yield (Scheme 56).



Scheme 56

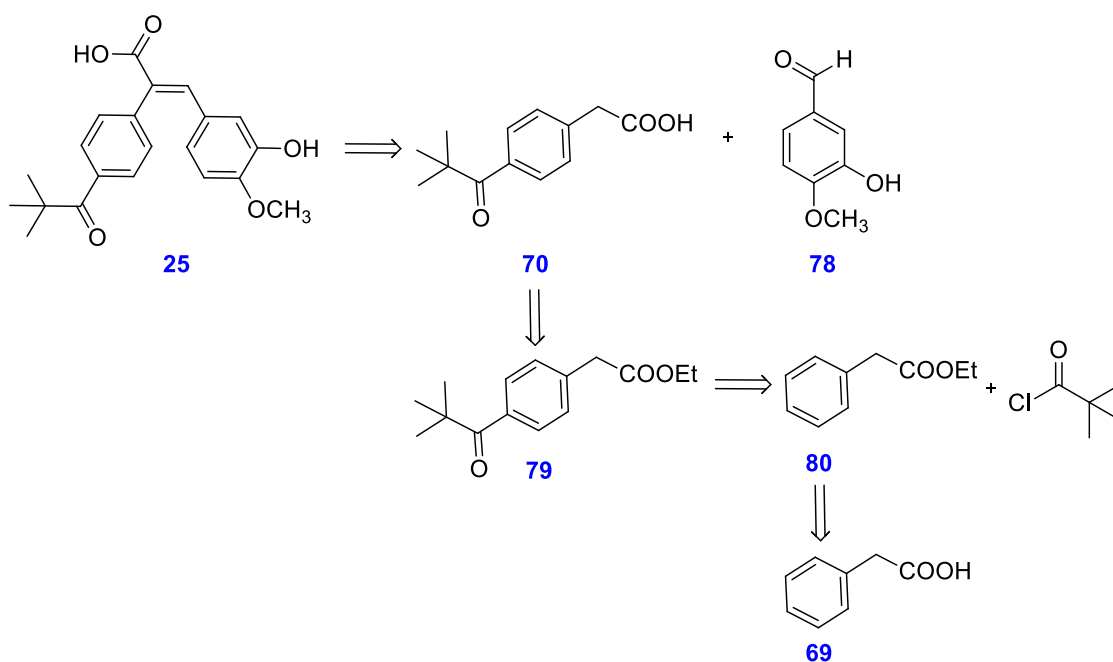
The esterification of the acrylic acid **23** was performed by methanol treatment in the presence of boron trifluoride etherate (BF₃·(CH₃-CH₂)₂O, Lewis acid) used as a catalyst. Under these conditions, the expected methyl ester was obtained with **76%** yield (Scheme 57).



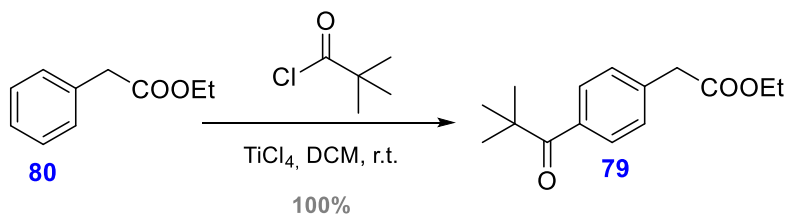
Scheme 57

3.1.3.1.6. Preparation of (*E*)-3-(3-hydroxy-4-methoxyphenyl)-2-(4-pivaloylphenyl)acrylic acid (**25**)

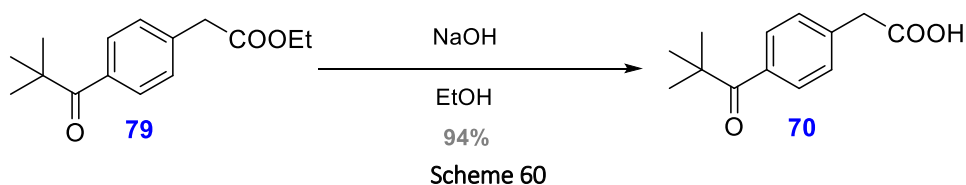
The following retrosynthesis analysis is proposed for the preparation of acid **25** (Scheme 58).



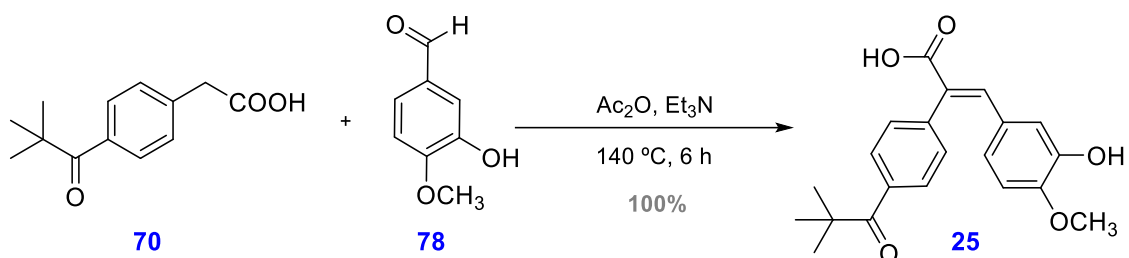
Alkylation of the arylacetic acid **69** with ethyl iodide in the presence of K_2CO_3 led to ethyl ester **80** in quantitative yields. The acylation of the ester **80** was then carried out with the pivaloyl chloride in the presence of $TiCl_4$ in dichloromethane with **100%** yield (Scheme 59).



Hydrolysis of ester **80** with sodium hydroxide-ethanol (4:1) led to acid **70** with **94%** yield (Scheme 60).



Then 2-(4-pivaloylphenyl) acetic acid (**70**) was condensed with 3-hydroxy-4-methoxybenzaldehyde (**78**) and the expected acid **25** was obtained with **100%** yield (Scheme 61).



Scheme 61

Spectroscopic NMR data indicated in Figure 89 confirmed the proposed structure for the carboxylic acid **25**.

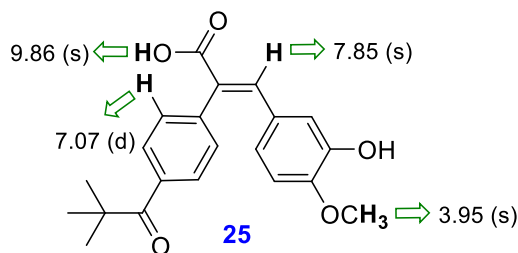
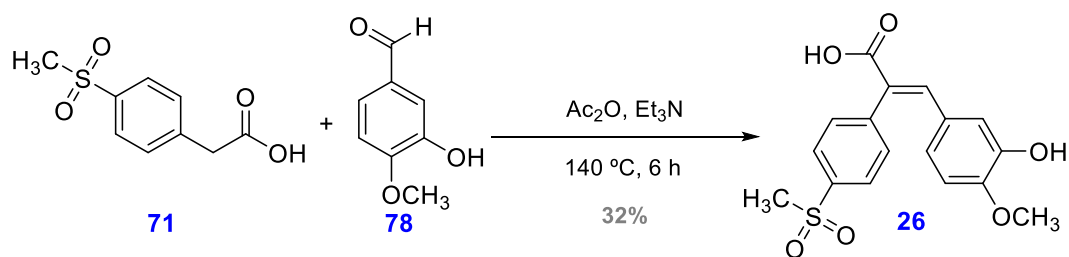


Figure 89. ^1H NMR spectroscopic data of **25**

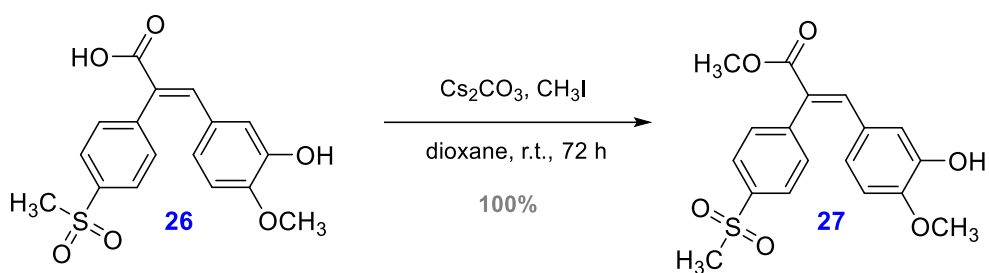
3.1.3.1.7. Preparation of (*E*)-3-(3-hydroxy-4-methoxyphenyl)-2-(4-(methylsulfonyl)phenyl) acrylic acid (**26**) and (*E*)-methyl 3-(3-hydroxy-4-methoxyphenyl)-2-(4-(methylsulfonyl)phenyl) acrylate (**27**)

Treatment of 2-(4-(methylsulfonyl)phenyl) acetic acid (**71**) with 3-hydroxy-4-methoxybenzaldehyde (**78**) under the conditions of the Perkin-type condensation reaction allowed to obtain the carboxylic acid **26** with **32%** yield (Scheme 62).



Scheme 62

The following was the alkylation of the acid **26** with methyl iodide and caesium carbonate used as a base. The low solubility of the starting acid **26** led to alkylation within 1,4-dioxane. This is how the expected methyl ester was obtained with **100%** yield (Scheme 63).



Scheme 63

The carboxylic acids **23**, **25** and **26** and the methyl esters **24** and **27** showed signals in the ^1H NMR spectrum that allowed each of them to be identified. The most prominent spectral data are shown below.

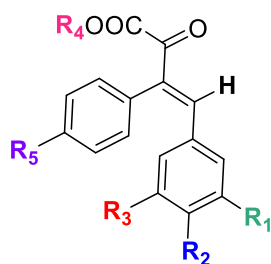


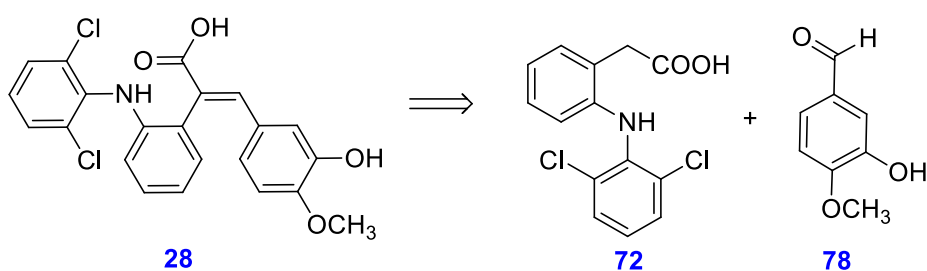
Table 18. Spectral data of different protons of acrylic acid derivatives

Compounds	H	H-2'/H-6'	H-2''	H-6''	R ₄ (H or CH ₃)
23	7.85	6.70 / 6.93	7.39 / 7.41	7.39 / 7.41	9.87 (H-)
24	7.75	6.54 / 6.61	7.21*	7.22*	3.84 (CH ₃ -)
25	7.85	7.07	6.70	6.94	9.86 (H-)
26	7.72	7.36	6.56	6.84	9.76 (H-)
27	7.84	7.51	6.67	6.99	3.81 (CH ₃ -)

* Interchangeables protons

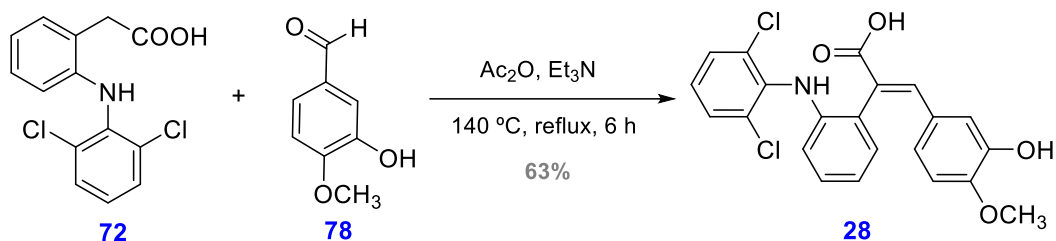
3.1.3.1.8. Preparation of (*E*)-2-(2-((2,6-dichlorophenyl)amino)phenyl)-3-(3-hydroxy-4-methoxyphenyl)acrylic acid (**28**) and (*E*)-methyl 2-(2-((2,6-dichlorophenyl)amino)phenyl)-3-(3-hydroxy-4-methoxyphenyl)acrylate (**29**)

The following retrosynthetic analysis is proposed for the synthesis of carboxylic acid **28** (Scheme 64).



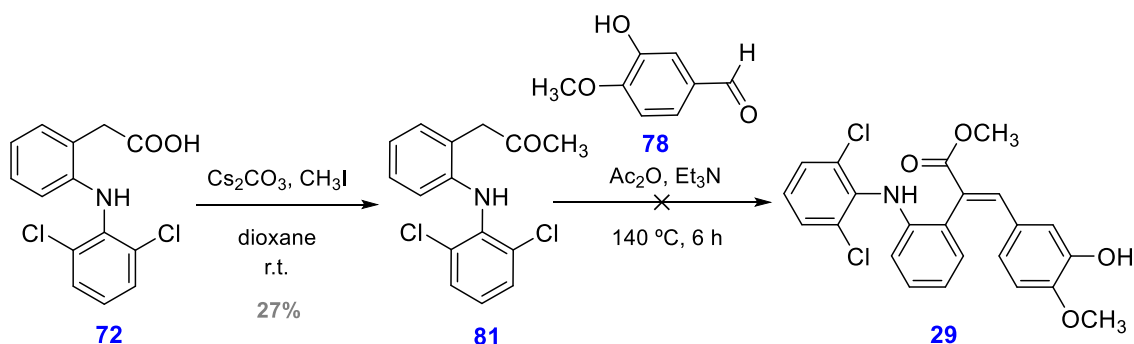
Scheme 64

The condensation of arylacetic acid **72** called *diclofenac* with the 3-hydroxy-4-methoxybenzaldehyde (**78**) using Perkin's conditions allowed the preparation of the expected acrylic acid **28** with **63%** yield (Scheme 65).



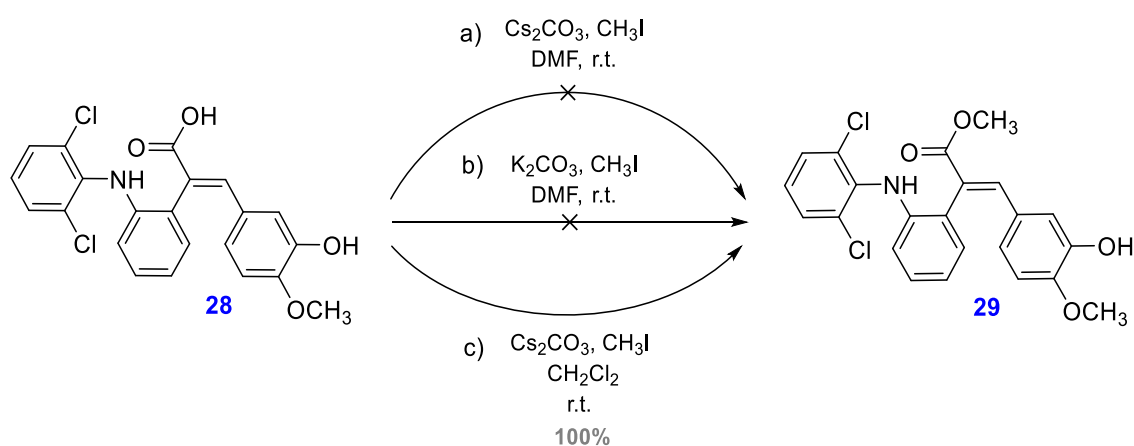
Scheme 65

The condensation of the aldehyde **78** with the ester **81** was also assayed, but under the tested conditions it did not allow the isolation of the expected ester **29** (Scheme 66).



Scheme 66

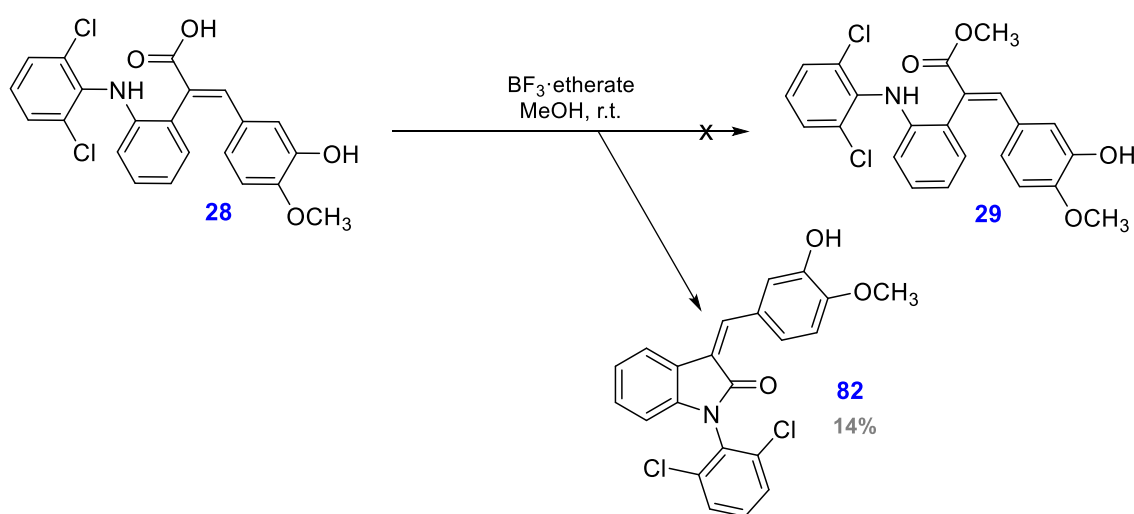
The formation of **29** by alkylation of carboxylic acid **28** was carried out with methyl iodide and in the presence of an inorganic base (Scheme 67).



Scheme 67

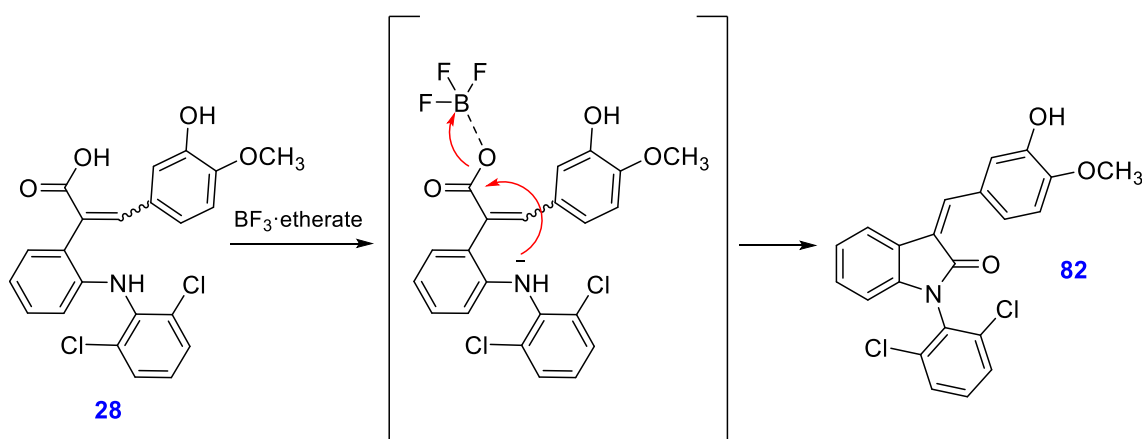
In the first case, caesium carbonate was used as the base and DMF was the solvent but the formation of the desired product **29** was not detected. It was neither possible to obtain it by changing potassium carbonate to caesium carbonate. Whereas using caesium carbonate, in dichloromethane, it successfully led to ester **29**, which yielded **100%**.

It should be noted that the esterification of acid **28** with methanol was also assayed in the presence of BF_3 ·etherate. In this case, the desired methyl ester was not obtained but the indole derivative **82** was isolated (Scheme 68).



Scheme 68

A possible mechanism for the formation of **82** is given below (Scheme 69).



Scheme 69

In this case, the boron trifluoride etherate was complexed with the carboxylic acid group and increases the electrophilic character thus facilitating the intramolecular attack by the amino group. A cyclization occurred, leading to the *N*-aryl indole derivative **82**.

The presence of two singlets in the ^1H NMR spectrum of **28** at 9.84 (COOH) and at 7.87 ppm (C=CH) confirmed the structure proposed for the acid **28**. However, in the ^1H NMR spectrum of **82** the 7.22 ppm signal that appeared as a double doublet, should be noted, and it was attributed to the proton of the benzylidene group (CH=C) (Figure 90). The structure proposed for **82** was confirmed by mass spectrometry (ESI +), theoretical mass 412.0429, and the experimental mass was 412.0511.

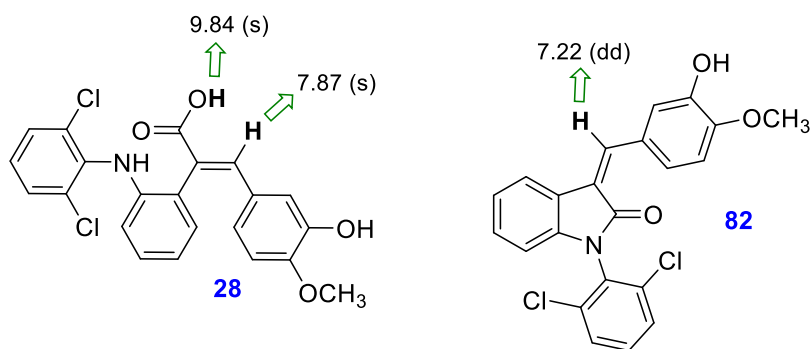
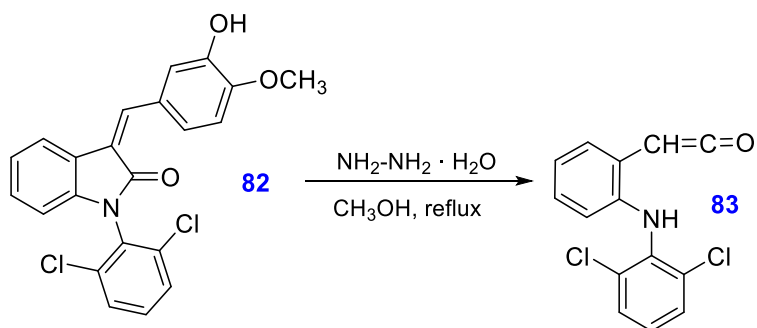


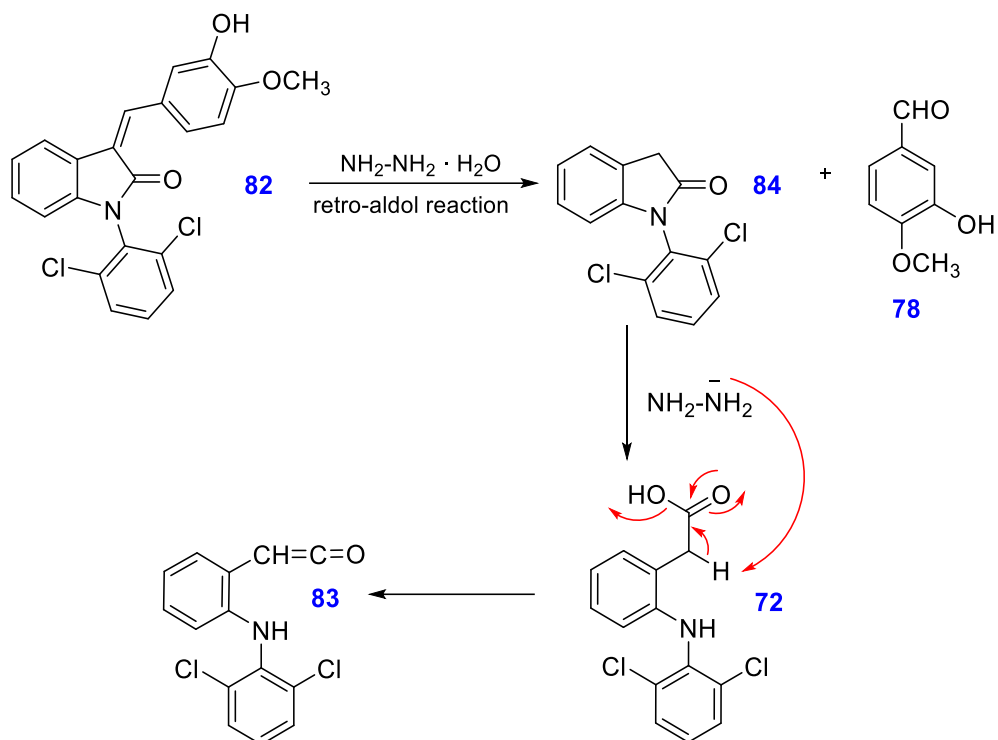
Figure 90. ^1H NMR spectral data of **28** and **82**

The treatment of indole **82** derivative with excess of hydrazine hydrate led to ketene **83** with **25%** yield (Scheme 70).



Scheme 70

The basic medium of hydrazine hydrate led to a retroaldolic reaction of **82** which resulted in indolone **84** and the corresponding aldehyde **78**. This indolone **84** was hydrolysed to form acetic acid **72** and again the hydrazine hydrate removed the proton of the benzylic position which, by removal of water, allowed the formation of ketene **83** (Scheme 71).



The signals at 71.0 (CH=) and at 173.6 (C=O) were perceived in the ^{13}C NMR spectrum, which allowed confirming the proposed structure for ketene **83** (Figure 91).

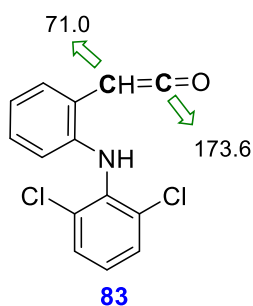
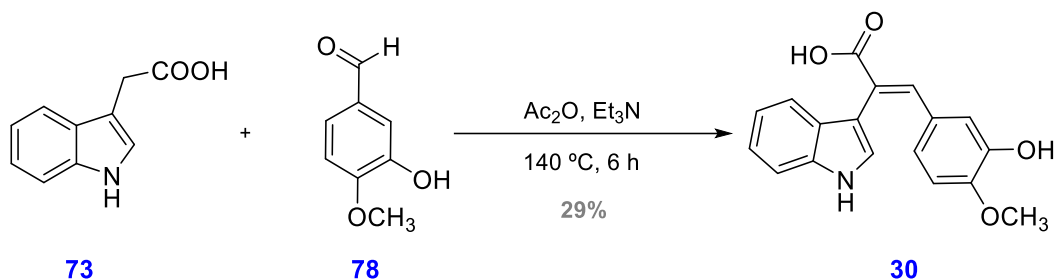


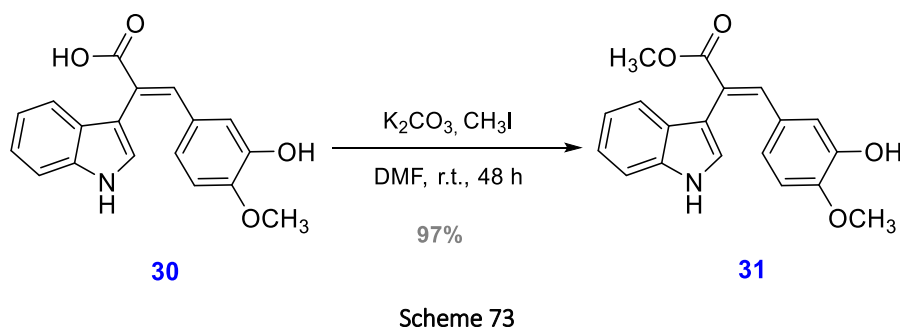
Figure 91. ^{13}C NMR spectral data of **83**

3.1.3.1.9. Preparation of (*E*)-3-(3-hydroxy-4-methoxyphenyl)-2-(1*H*-indol-3-yl)acrylic acid (**30**) and (*E*)-3-(3-hydroxy-4-methoxyphenyl)-2-(1*H*-indol-3-yl)methyl acrylate (**31**)

Applying the condensation procedure between carboxylic acid **73** and 3-hydroxy-4-methoxybenzaldehyde (**78**) applied to the derivatives mentioned above in this present thesis, acid **30** was obtained with **29%** yield (Scheme 72).



Next, the alkylation of acid **30** with methyl iodide, in DMF, and in the presence of potassium carbonate led to the carboxylic ester with **97%** yield (Scheme 73).



In the ¹H NMR spectrum, signals at 3.78 (s, CH₃-O-); 6.72 (d, H-5''); 7.38 (s, CH=); and 7.96 ppm (s, H-2') were perceived, which allowed to identify methyl ester **31** correctly (Figure 92).

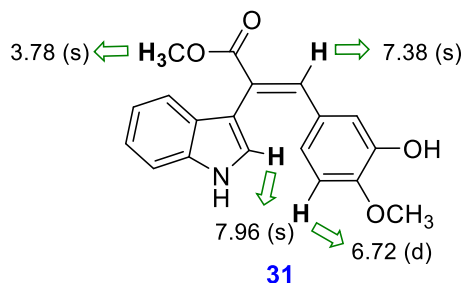
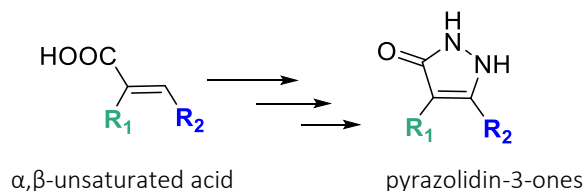


Figure 92. ¹H NMR spectroscopic data of **31**

3.1.3.2. Preparation of pyrazolidin-3-ones (32-38)

Pyrazolidin-3-ones considered internal hidrazides were obtained by treatment of α,β -unsaturated carboxylic acids with hydrazines under different conditions (Scheme 74).^{140,141,142,143,144,145,146,147,148,149,150,151,152,153,154,155,156,157}



Scheme 74

In this thesis work, the synthesis of pyrazolidin-3-ones was performed by ester treatment with hydrazine hydrate,^{158,159,160} as indicated below (Scheme 75).

¹⁴⁰ R. Kitamura. *Rept. Japan. Assoc. Adv. Sci.* **1942**, *16*, 541-544.

¹⁴¹ A. Spasov, B. Kurtev. *Ann. Univ. Sofia, Fac. Phys.-Math.* **1947**, *43* (Livre 2), 37-51.

¹⁴² H. T. Clarke, *Chemistry of Penicillin*. Ed. Oxford University Press: Oxford **1949**, 688-848.

¹⁴³ F. H. Stodola. *J. Org. Chem.* **1948**, *13*, 757-762.

¹⁴⁴ P. E. Gagnon, J. L. Boivin, R. N. Jones. *Can. J. Res.* **1949**, *27B*, 190-204.

¹⁴⁵ C. V. Otis. *J. Soc. Motion Picture Engrs.* **1949**, *52*, 534-539.

¹⁴⁶ H. Dorn. *Chem. Heterocycl. Compd. USSR* **1981**, 3-31.

¹⁴⁷ R. M. Claramunt, Elgero. *J. Org. Proc. Prep. Int.* **1991**, *23*, 273-320.

¹⁴⁸ J. Steve, M. A. Horvat, J. H. Eds. Golob, *Nova Science Publishers, Inc.* **2008**, 129-182.

¹⁴⁹ G. Varvounis, Y. Iamegos, G. Pilidis. *Adv. Heterocycl. Chem.* **2001**, *80*, 75-165.

¹⁵⁰ T. Eicher, S. Hauptmann. *The Chemistry of Heterocycles*, 2nd Ed. **2003**.

¹⁵¹ J. Marchand-Brynaert, L. Ghosez, G. Lukacs, M. Ohno. *Springer* **1990**. DOI: [10.1007/978-3-642-75617-7_20](https://doi.org/10.1007/978-3-642-75617-7_20)

¹⁵² S. Hanessian, G. McNaughton-Smith, H.-G. Lombart, W. D. Lubell. *Tetrahedron* **1997**, *53*, 12789-12854.

¹⁵³ M. I. Konaklieva, B. J. Plotkin. *Curr. Med. Chem. Anti-Infect. Agents* **2003**, *2*, 287-302.

¹⁵⁴ H. L. White, J. L. Howard, B. R. Cooper, F. E. Soroko, J. D. McDermed, K. J. Ingold, R. A. Maxwell. *Neurochem.* **1982**, *39*, 271-271.

¹⁵⁵ K. Kawai, J. Hasegawa, H. Shiojiri, Y. Nozawa, K. Tsurumi, H. Fujimura. *Experientia* **1985**, *41*, 490-492.

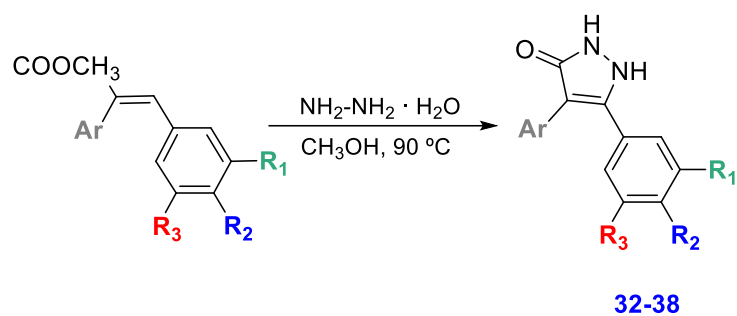
¹⁵⁶ C. Cucurou, J. P. Battioni, D. C. Thang, N. H. Nam, D. Mansuy. *Biochemistry* **1991**, *30*, 8964-8970.

¹⁵⁷ E. M. Kosower, E. Hershkowitz. *Chem. Abstr.* **1995**, *122*, 214077.

¹⁵⁸ A. Novak, J. Bezensek, U. Groselj, A. Golobic, B. Stanovnik, J. Svete. *Arkivoc* **2011**, 18-28.

¹⁵⁹ A. Novak, M. Stefanic, U. Groselj, M. Hrast, M. Kasunic, S. Gobec, B. Stanovnik, J. Svete. *Tetrahedron* **2013**, *69*, 6648-6665.

¹⁶⁰ A. Novak, M. Stefanic, U. Groselj, M. Hrast, M. Kasunic, S. Gobec, B. Stanovnik, J. Svete. *Helv. Chim. Acta* **2014**, *97*, 245-267.

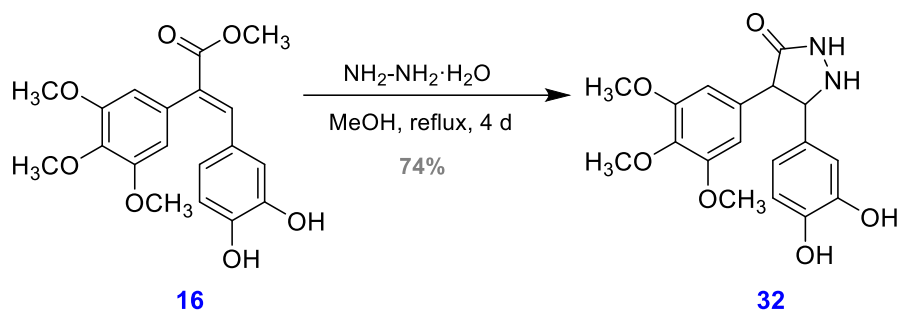


Scheme 75

Table 19. Structure of compounds **16-31**

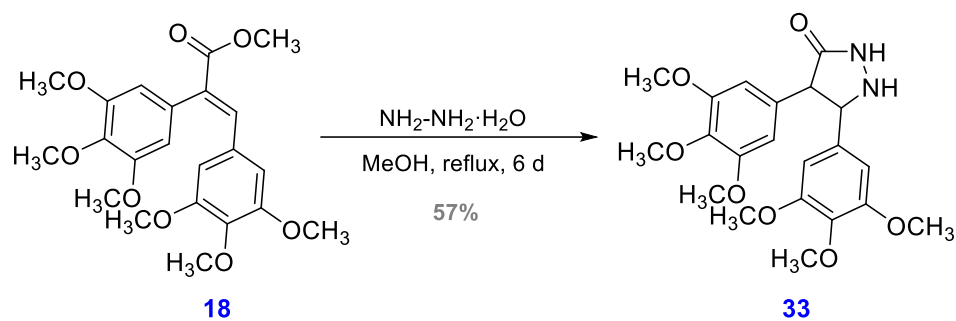
Compounds	Ar	R ₁	R ₂	R ₃
16	3,4,5-trimethoxyphenyl	-OH	-OH	-H
18	3,4,5-trimethoxyphenyl	-OCH ₃	-OCH ₃	-OCH ₃
20	3,4,5-trimethoxyphenyl	-NO ₂	-OH	-H
22	3,4,5-trimethoxyphenyl	-O-CH ₂ -O-		-H
24	phenyl	-OH	-OCH ₃	-H
27	methylsulfonylphenyl	-OH	-OCH ₃	-H
31	3-indolyl	-OH	-OCH ₃	-H

The pyrazolidin-3-one **32** was obtained with **74%** yield by adding hydrazine hydrate to the α,β -unsaturated ester **16** by heating under reflux of methanol for 4 days (Scheme 76).



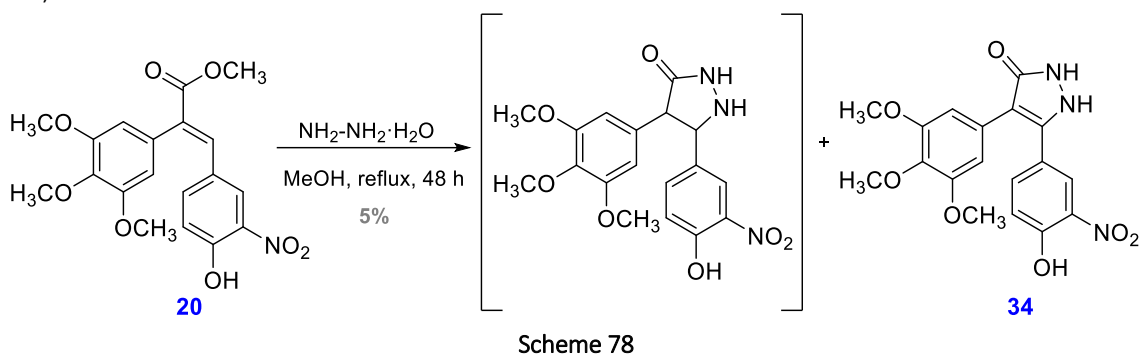
Scheme 76

Following the same conditions in the treatment of ester **18** with excess of hydrazine hydrate, pyrazolidin-3-one **33** was obtained with **57%** yield (Scheme 77).

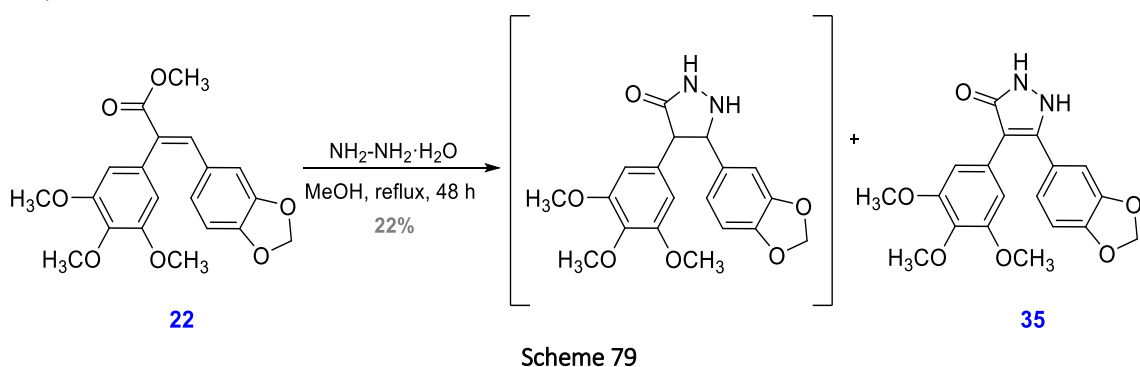


Scheme 77

The treatment of the methyl ester **20** with excess of hydrazine hydrate allowed the isolation of pyrazol-3-one **34** in only **5%** yield. It should be pointed out that the presence of the nitro group in the structure of the starting ester hinders the reaction with the hydrazine and also led to the oxidation of pyrazolidin-3-one to pyrazol-3-one (Scheme 78).



Pyrazol-3-one **35** was obtained by treatment of ester **22** with hydrazine hydrate following the conditions mentioned above in this same work, providing a yielding of **22%** (Scheme 79).



When the treatment was extended up to 7 days, hydrazide **85** was obtained as a secondary product with **22%** yield.

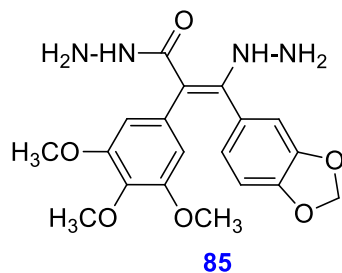
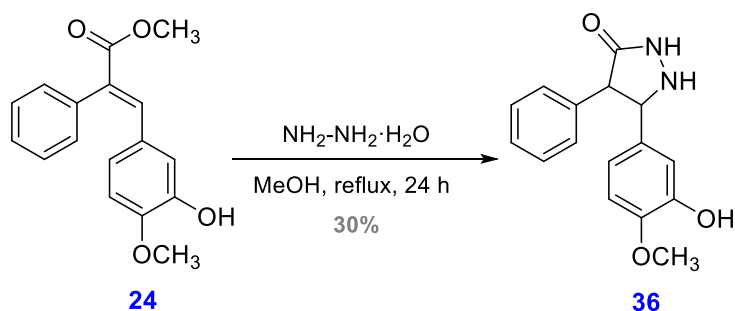


Figure 93. Structure of secondary product **85**

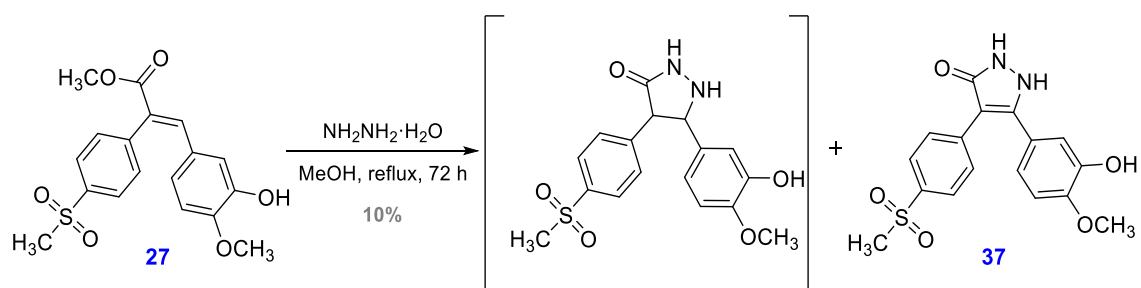
This derivative was obtained due to excess of hydrazine hydrate that allowed the addition to the α,β -unsaturated ester and at the same time the hydrazide was formed.

Treatment of ester **24** with hydrazine hydrate led to pyrazolidin-3-one **26** with **30%** yield (Scheme 80).



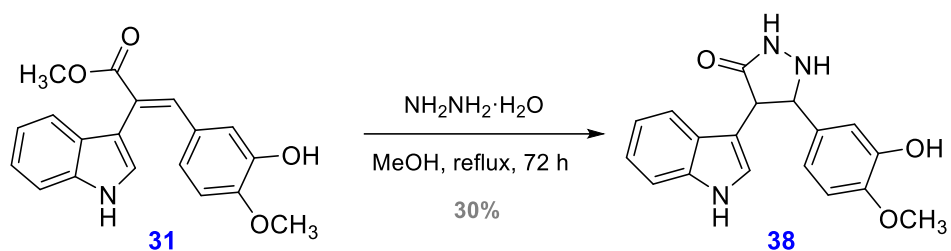
Scheme 80

Following the general conditions indicated above, ester **27** gave access to the pyrazol-3-one **37** with only **10%** yield (Scheme 81).



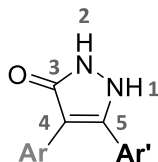
Scheme 81

Likewise, the addition of hydrazine hydrate to the methyl ester **31** led to the pyrazolidin-3-one **38** with **30%** yield (Scheme 82).



Scheme 82

^{13}C NMR data allowed the assignment of the indicated compounds **32-38**, so the quaternary carbons C-4 and C-5 of pyrazol-3-ones **35** and **37** showed the aromatization of this nucleus, whereas for pyrazolidin-3-ones **32**, **33**, **36** and **38** the C-4 and C-5 carbons appeared in the form of methynes (CH-) in more shielding fields. As for this compounds group, they all had a clear signal corresponding to the carbonyl group that resonated between 153.2 and 176.5 (Table 20).



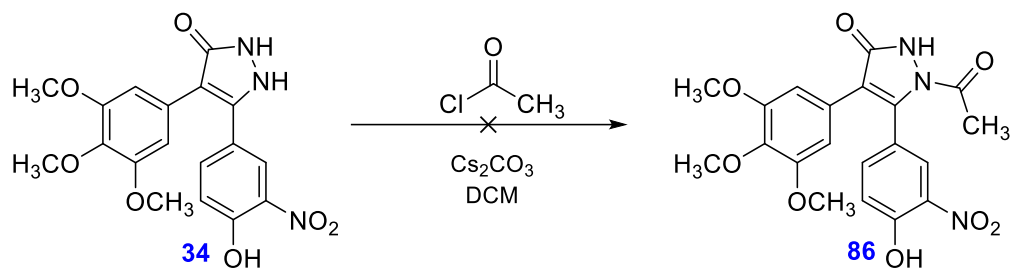
In addition, it should be noted that pyrazolidin-3-ones usually show signals corresponding to two specific rotamers, whereas these rotamers are not seen with products that have a pyrazol-3-one nucleus.

Table 20. Spectral data of compounds **32-38**

Compounds	Ar	Ar'	C-3	C-4	C-5
32	3,4,5-trimethoxyphenyl	3,4-dihydroxyphenyl	173.2 (B) 174.5 (A)	39.1	49.0
33	3,4,5-trimethoxyphenyl	3,4,5-trimethoxyphenyl	173.4	40.0	29.6
34	3,4,5-trimethoxyphenyl	3-nitro-4-methoxyphenyl	-	-	-
35	3,4,5-trimethoxyphenyl	piperonal	170.6	102.4	153.0
36	phenyl	3-hydroxy-4-methoxyphenyl	173.7 (A) 176.5 (B)	54.7	53.4
37	4-methylsulfonylphenyl	3-hydroxy-4-methoxyphenyl	153.2	151.8	96.5
38	3-indolyl	3-hydroxy-4-methoxyphenyl	169.5 (A) 175.1 (B)	40.4 (A) 40.3 (B)	44.7 (A) 44.6 (B)

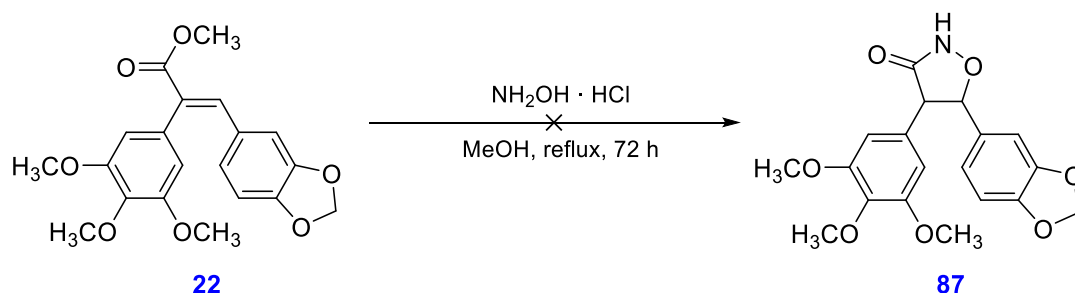
A : majority isomer / B: minority isomer

Selective *N*-acylation of pirazol-3-one **34** did not lead to the expected product **86** (Scheme 83).



Scheme 83

The treatment of the ester **22** with hydroxylamine was also assayed to obtain the corresponding isoxazolidin-3-one, but despite trying different conditions, in no case could the desired product be isolated (Scheme 84).



Scheme 84

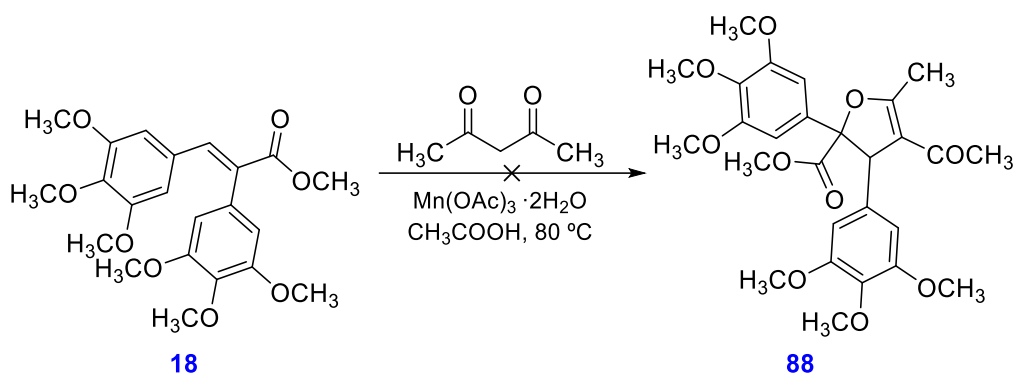
3.1.3.2.1. Preparation of substituted furans 4-acetyl-3-(3,4-dimethoxyphenyl)-5-methyl-2-(3,4,5-trimethoxyphenyl)-2,3-dihydrofuran (**39**) and 4-acetyl-2-(3,4-dimethoxyphenyl)-5-methyl-3-(3,4,5-trimethoxyphenyl)-2,3-dihydrofuran (**40**)

Different protocols for the preparation of substituted furans are known.^{161,162}

In this work, it was first considered the treatment of ester **18** with acetylacetone, using manganese (III) acetate in acidic medium, but in these conditions the isolation of the expected dihydrofuran **88** was not possible (Scheme 85).

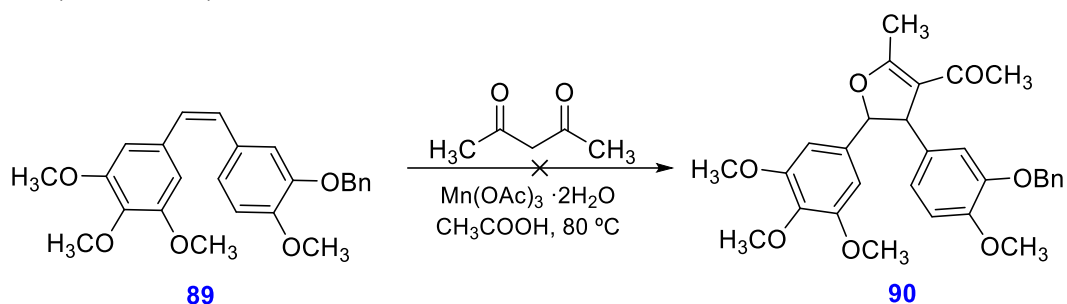
¹⁶¹ Y. Wu, Z. Huang, Y. Luo, D. Liu, Y. Deng, H. Yi, J.-F. Lee, C. W. Pao, J.-L. Chen, A. Lei. *Org. Lett.* **2017**, *19*, 2330-2333.

¹⁶² J. T. Yu, B. Shi, H. Peng, S. Sun, H. Chu, Y. Jiang, J. Cheng. *Org. Lett.* **2015**, *17*, 3643-3645.



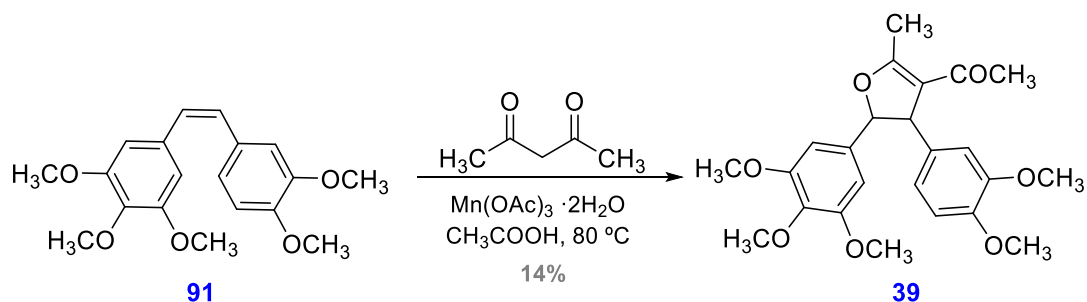
Scheme 85

Treatment of stilbene **89** (*cis*) with acetylacetone was also tested under the same conditions as in the previous section, but the expected product could not be isolated either (Scheme 86).



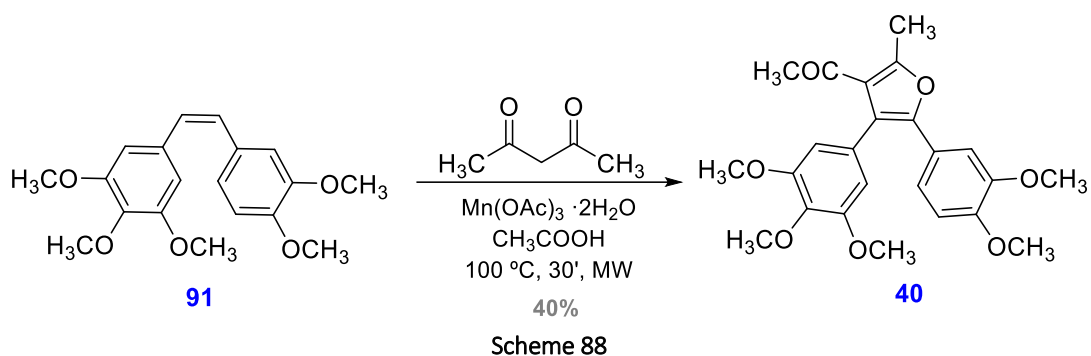
Scheme 86

However, when applying these conditions to stilbene **91** (*cis*), the isolation and identification of dihydrofuran **39** was possible (Scheme 87).

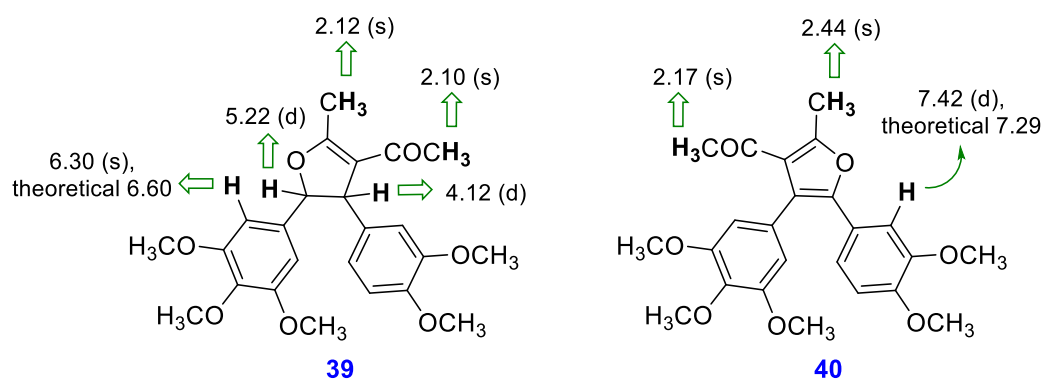


Scheme 87

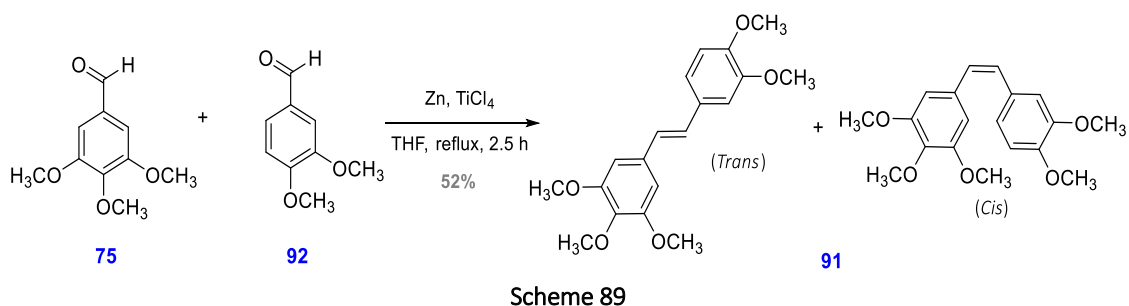
Within the same work, the treatment of stilbene **91** (*cis*) with acetylacetone in the presence of manganese (III) acetate and, in this case, with microwave irradiation led to furan **40** with **40%** yield (Scheme 88).



In the ^1H NMR spectrum of **39**, signals assignable to the protons at positions 2 (5.22 ppm (d)) and 3 (4.12 ppm (d)) appeared, besides the singlets at 2.10 ppm ($\text{CH}_3\text{-O-}$) and 2.12 ($\text{CH}_3\text{-Ar}$), which confirmed the proposed structure, whereas the spectrum of **40** did not show the signals of the groups -CH in positions 2 and 3 (Figure 94).



The McMurry reaction conditions were applied to the condensation of aldehydes **75** and **92** to obtain stilbene **91**. A mixture of *cis* and *trans* stilbenes were obtained with **52%** yield. Under these conditions the *cis* isomer was predominant (Scheme 89).



The ^1H NMR spectra showed a 7.04 ppm doublet with a $J = 16$ Hz attributable to the protons of the double bond in the *trans* configuration, while for the *cis* isomer, the double bond protons appeared at 6.92 ppm in the form of a doublet with a constant coupling of 12 Hz (Figure 95).

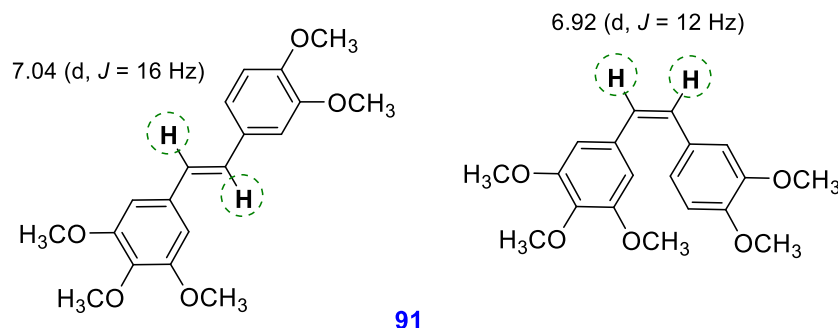


Figure 95. ^1H NMR spectral data of **91** *cis/trans*

3.2. Preparation of new KRAS protein inhibitory agents

As mentioned in the introduction, KRAS protein inhibitors are compounds of a very diverse structure that inhibit the protein, most of them directly on the ATP binding site itself.

3.2.1. Optimization of potential KRAS inhibitors or activators from lead **P14**

3.2.1.1. Search of the binding site of the lead **P14** by performing gaussian accelerated molecular dynamics simulations (GaMD)

Given that it is a great challenge to directly inhibit the KRAS protein as well as to find effectors that can act directly on said protein, the purpose of this work is to find the binding site of a compound which is known to act on this target. In previous work, it was possible to determine the activity of lead **P14**. It was concluded that **P14** was able to inhibit KRAS and act on the MAPK signalling pathway.¹⁰¹ For this reason, the aim is to determine the binding site, as well as, to identify what interactions it establishes with the protein. Once the binding mode is determined, a search will be made for other small organic molecules that could act in the same place.

The **P14** compound is a small organic molecule with two different groups, one is the hydrophobic part and the other represents the polar part, which consists [REDACTED] where π -stacking and VdW interactions could be established within the protein, while the polar groups ([REDACTED]) could establish hydrogen bonds (Figure 96).

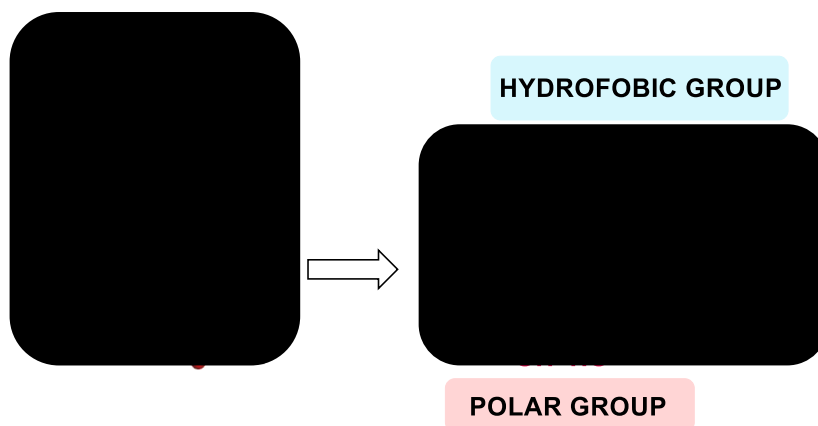


Figure 96. Main structural moieties of **P14**

These particular conditions are very interesting since the hydrophobic group can retain the ligand in one position while the polar groups give the direction of the ligand. To find out where the **P14** binds, it was first placed in 8 different positions of the protein box, giving 8 separately PDB's complex, where each complex differed only by the position of **P14**. The HiConMD program was used to carry out this process. The purpose of this kind of **P14** positioning was to place the **P14** ligand at different initial locations in order to obtain different starting points. With this in mind, the aim was to find out the last position of the ligand in the last point of the trajectory, being in theoretically (for very long MD) the same for all ligands. However, later it will be verified that the final position of the ligand was never the same.

After ligand positioning, each complex was solvated and energy minimized. Then, 6 ns of conventional molecular dynamics (cMD) was performed in order to adapt the system. This was followed by a stepped GaMD, first of 300 ns, those structures found in the water being discarded. A second 300 ns were made and those structures located in the water were discarded again. At this point, 200 ns more were completed, and those structures interacting with the protein which also showed good energy of binding were prepared for 200 ns more. Four of the eight initial structures performed 1200 ns of GaMD and after that, only the best complexes were selected to carry out cMD calculations. In particular

300 ns of cMD were performed for complexes **3**, **5** and **6**.

During the first 300 ns, complex **2** and **4** went to the solvent and were discarded for further calculations. As shown in Figure 97a, the free energy of binding (MM-GBSA) of complex **4** calculated for the last point of the trajectory was 0 kcal/mol, which indicated that the ligand was in the solvent; the same happened for complex **2**, where the ligand was surrounded by water molecules (Figure 97b).

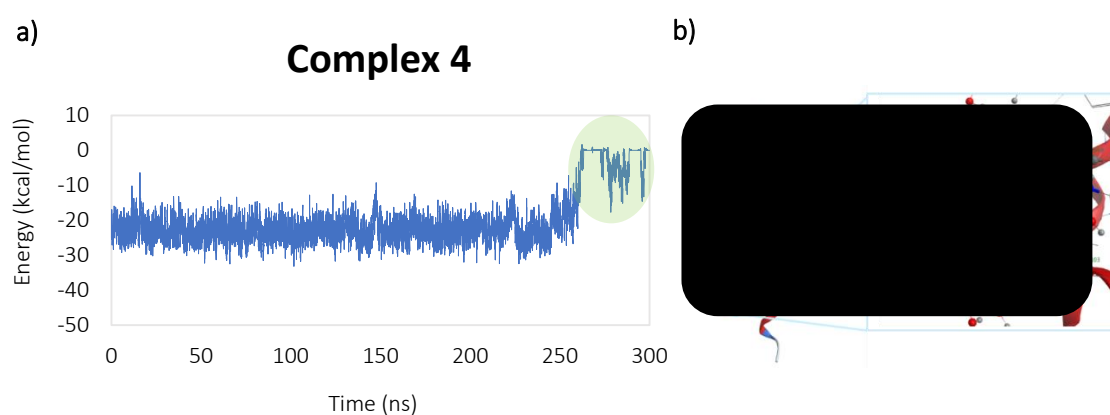
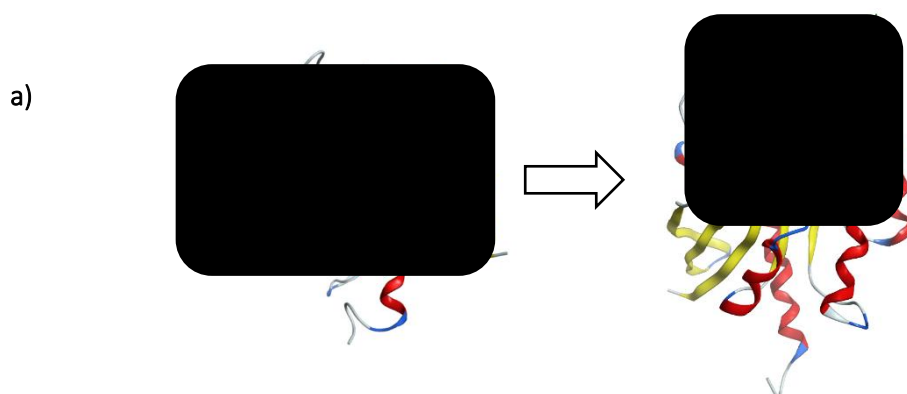


Figure 97. a) MM-GBSA calculated for 300 ns of GaMD for complex **4**; b) The last point of the trajectory for complex **2**

Complex **1**, from the initial position to the 800 ns was placed in pocket $\alpha 4$, but from the 800 to 1000 ns went to the solvent and was discarded for additional calculations (Figure 98a). The complex achieved an average energy of -11.0 kcal/mol, according to the MM-GBSA performed (800 to 900 ns). The decrease of energy was significant since the ligand was placed in the solvent (Figure 98b).



b)

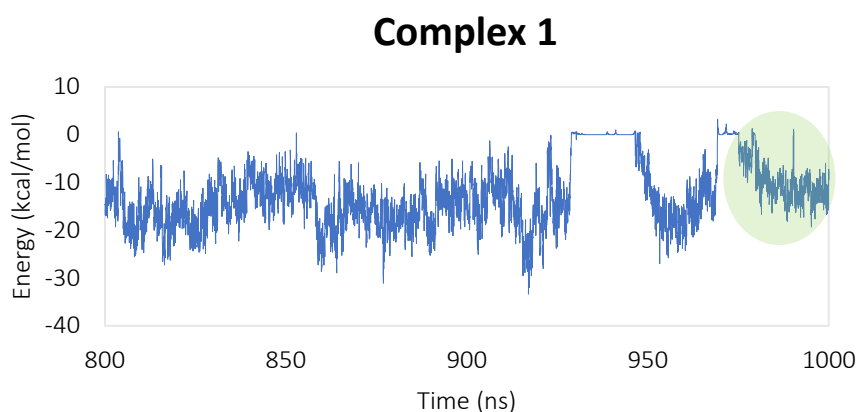


Figure 98. a) Position of **P14** ligand from 800 to 1000 ns of GaMD; **b)** Energy values from 800 to 1000 ns of trajectory

Complex **3** achieved a binding free energy of -15.0 kcal/mol for the first 850 ns of trajectory. However, from 850 to 10000 ns the energy of binding increased considerably to -28.0 kcal/mol being stable the last 100 ns. In this case, the ligand was placed in pocket α_4 during the entire trajectory (Figure 99a, b).

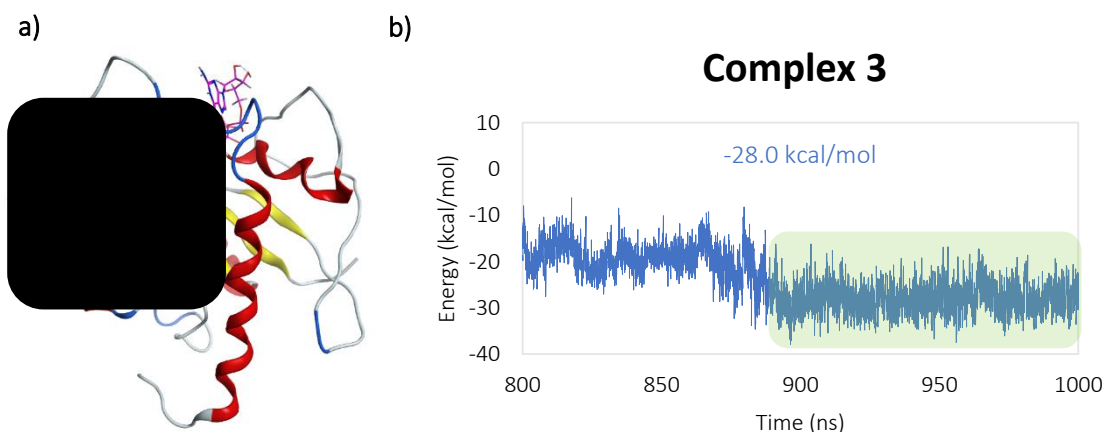


Figure 99. a) Ligand position; **b)** Energy values (MM-GBSA) of **P14** during the 800-1000 ns of GaMD

As regards complex **5**, after carrying out 1000 ns of GaMD calculations, the ligand was not stabilized. The energy values of the MM-GBSA showed an average of energy of -40.0 kcal/mol (Figure 100), setting the ligand into an allosteric pocket, something that had never been described before in the literature.

Complex 5

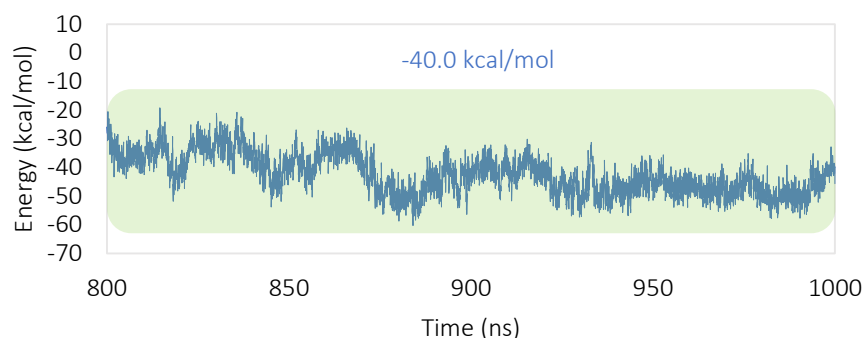


Figure 100. MM-GBSA from 800 to 1000 ns of GaMD for **P14** in complex **5**

In this period of time (800 to 1000 ns) the ligand got inside the protein through the ████████ of the hydrophobic group and the pyrazole ring.

The lack of stability was due to the fact that the ligand first goes into the protein and needs a position readjustment and consequently needs to open the α -helices. Anyway, the energy values were better in comparison to the other complexes studied.

On the other hand, complex **6** and **8** were placed at P4 pocket (Figure 101) after carrying out 1000 ns of GaMD trajectory.

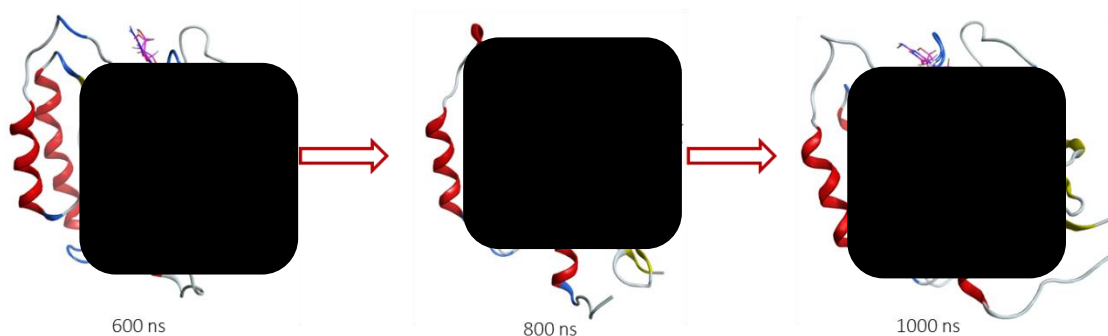


Figure 101. Position of the ligand during the 600 to 1000 ns aMD for complex **6**

The MM-GBSA was performed for both complexes. The binding free energy was ranged between -20.0 to -22.0 kcal/mol being both stable over time (Figure 102).

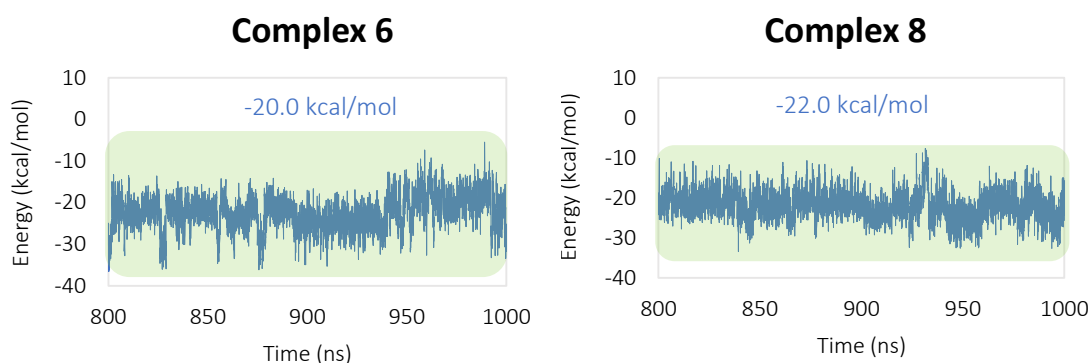


Figure 102. MM-GBSA values for complex **6** and **8**

Complex **7** was stable during the 600 to 800 ns of GaMD, whereas during the next 100 ns the ligand went to the solvent (800 to 900 ns) and consequently the energy values decreased. Finally the last 100 ns of the GaMD the ligand was re-stabilized, but it was practically in contact with the solvent all the time (Figure 103).

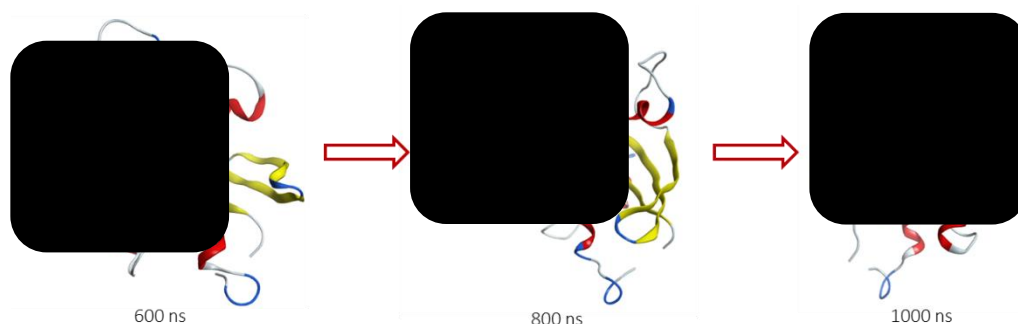


Figure 103. Trajectory of **P14** during the 600 to 1000 ns of GaMD trajectory for complex **7**

The best energetically complexes were **3**, **5**, **6** and **8**. Next step consisted in performing 200 ns more of GaMD in order to confirm the stability of each complex.

Complex **3** was energetically stabilized with an average free energy of binding around -30.0 kcal/mol, from 1050 to 1200 ns of GaMD trajectory (Figure 104a). The position in pocket α_4 did not change during this time. However, the ligand was almost in the solvent and the interactions with the protein were poor (Figure 104b).

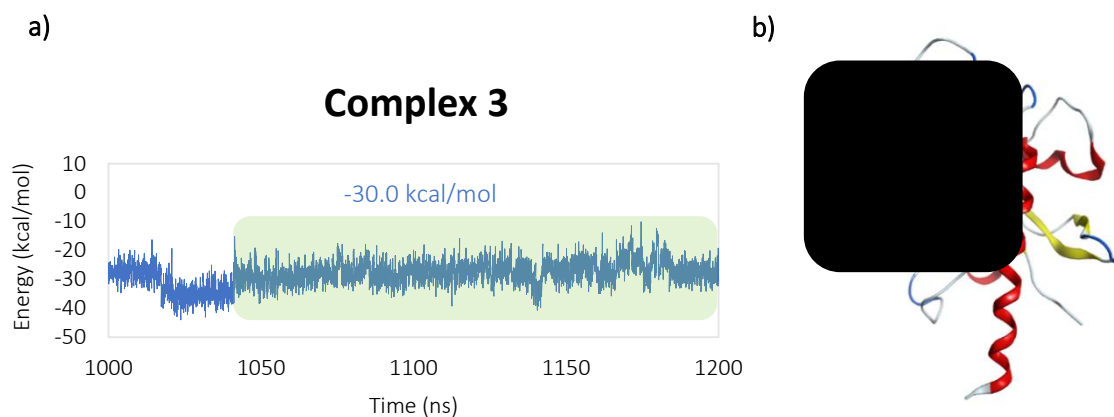


Figure 104. a) MM-GBSA calculation (1000 to 1200 ns) for complex 3; **b)** P14 position ($\alpha 4$)

Complex 5 was not stabilized over time from 1000 to 1200 ns of GaMD trajectory because it was still looking the best energetically position in order to achieve the maximum number of protein-ligand interactions.

The MM-GBSA calculated an average energy of binding of -45.0 kcal/mol, turning out to be the complex with the highest free energy of binding compared to all the others (Figure 105a). The ligand was kept in the same position as it was previously (from 600 to 1000 ns of GaMD trajectory) in this calculation and a new allosteric binding site was found, being of great interest for the research (Figure 105b).

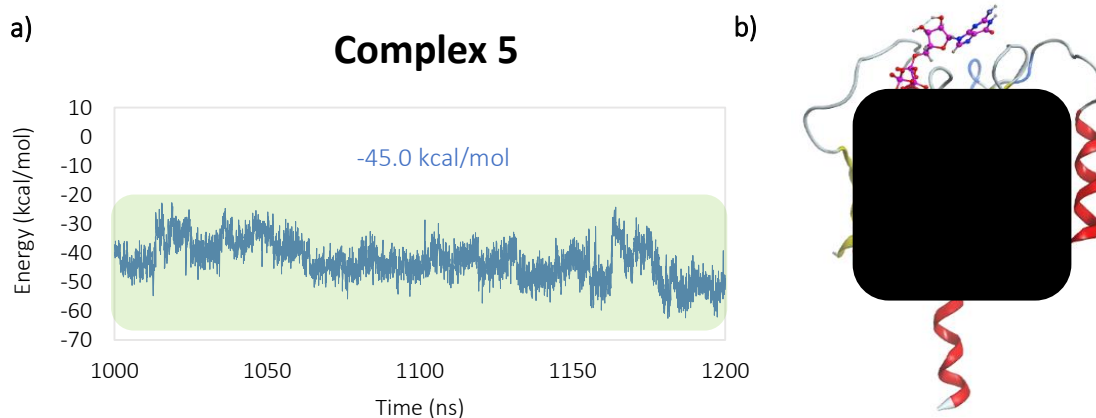


Figure 105. a) MM-GBSA values (1000 to 1200 ns) for complex 5; **b)** P14 position (allosteric pocket)

Complex **6** was energetically stabilized from 1032 ns of GaMD trajectory. The ligand did not change the position (P4 pocket) since the last trajectory calculated (600-1000 ns). The energy values improved, being -28.0 kcal/mol the average energy (Figure 106a, b). However, complex **8**, which was placed in the same pocket, was in contact with water molecules at the last point of the trajectory. Consequently, an evident increase in energy was observed until reaching a binding free energy of -14.0 kcal/mol.

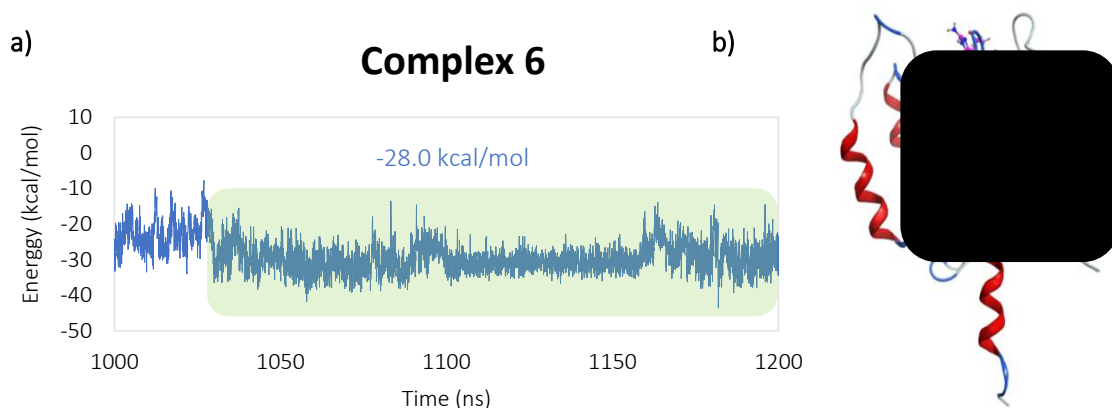


Figure 106. a) MM-GBSA calculation (1000 to 1200 ns) for complex **6**; b) **P14** position (P4)

To sum up, complex **3** was placed in pocket α_4 being most of the time in contact with the solvent. Complex **5** was the most promising positioned ligand, in which a new allosteric binding site was found. Complexes **6** and **8** were positioned in the same place, corresponding to P4 pocket. The energy and stability were different for each of the complexes which permitted to discard complex **8** and use complex **6** for further calculations.

After analyzing all the systems, complexes **3**, **5** and **6** were used to carry out more molecular dynamic calculations. As the binding free energy obtained with GaMD cannot be extrapolated to the real energy, further cMD calculations were performed. Next step consisted of 300 ns of conventional molecular dynamics using the last point of GaMD as the starting point for cMD.

Starting with complex **3**, it is noteworthy that the position of the ligand observed in GaMD was the same (in α_4 pocket) for the 300 ns of cMD trajectory. Comparing the energy values of GaMD and cMD, there was an energy improvement and also in stability (Figure 107).

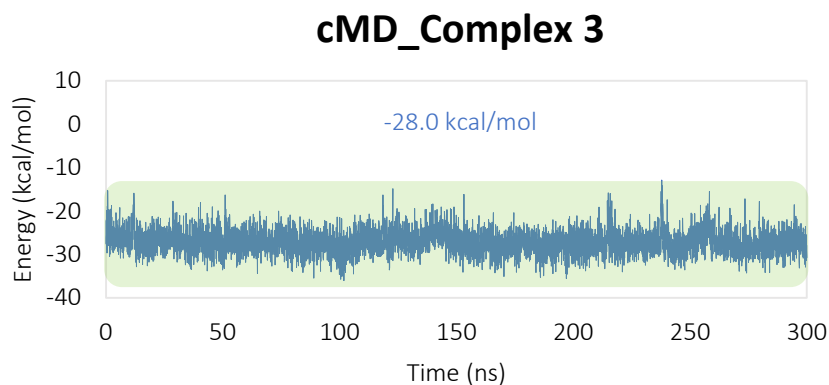


Figure 107. MM-GBSA obtained over 300 ns of cMD

Major interactions between the ligand and the protein were 3 hydrogen bonds. The first hydrogen bond took place with the [REDACTED]; the second hydrogen bond established was with [REDACTED], and the third, between the [REDACTED]. The [REDACTED] was surrounded for solvent and non-important hydrophobic interactions were observed (Figure 108).

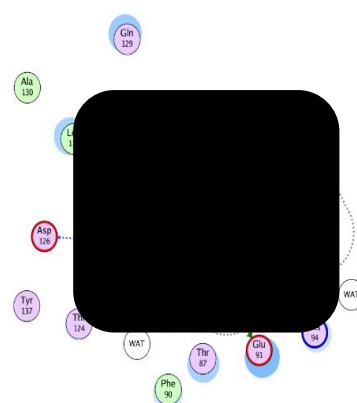


Figure 108. Main interactions established by **P14** and the protein

For complex **5**, the ligand was placed in a new allosteric binding site which has never been described before in the literature. The entry of the ligand to the protein was curious since SwII and $\alpha 3$ needed to be opened by **P14**. A slight deformation of the protein was observed when the hydrophobic group and one pyrazole ring got inside (Figure 109).

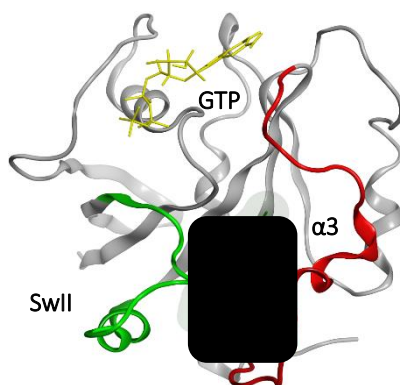


Figure 109. Entrance of **P14** between SwII and $\alpha 3$

Most of the ligand-protein interactions were hydrophobic although polar interactions were found. As shown in Figure 110, 3 hydrogen bonds were established, giving in this way the correct direction of the ligand inside the protein. The interactions were between the [REDACTED] (Figure 110a), the [REDACTED] (Figure 110b) and the last one, [REDACTED] (Figure 110c). The hydrophobic interactions (Figure 110 in blue) were established thanks to the methyls and the [REDACTED] of the ligand. In this case, π -stacking and VdW interactions were identified, the most remarkable residues to carry out such interactions were [REDACTED].

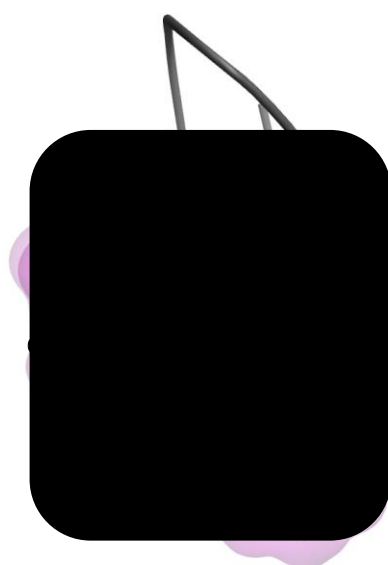


Figure 110. Hydrogen bond interactions represented in pink and in blue the hydrophobic interactions. After identifying the chemical interactions, the MM-GBSA was also calculated. In this case, the complex was not stabilized due to the subsequent changes made by the ligand during its entrance inside the protein. The energy values were better in conventional molecular dynamics than in accelerated molecular dynamics. The average of free binding energy was approximately -50.0 kcal/mol (Figure 111).

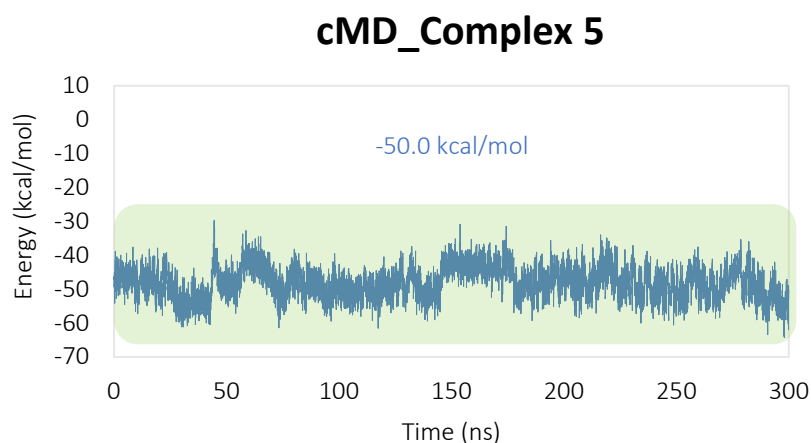


Figure 111. MM-GBSA for complex 5

Finally for complex 6, it should be noted that the ligand was found in the same position in which was placed in GaMD trajectory. The ligand did not move from P4 pocket.

The free energy of binding for this complex was about an average energy of -28.0 kcal/mol.

Nevertheless the interactions performed were poor. The [REDACTED] was the only that established an interaction with the protein and the [REDACTED] observed were with the solvent (Figure 112).

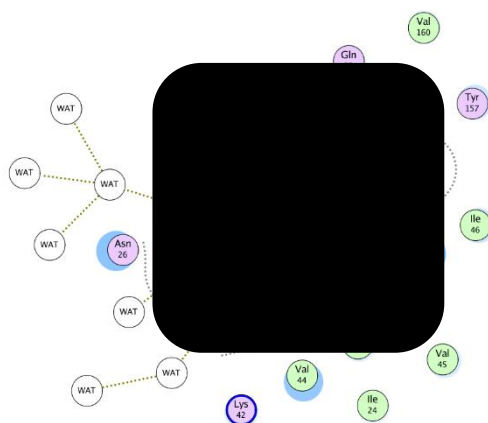


Figure 112. Interactions established between P14 and the protein

3.2.2. Preparation of KRAS inhibitors

The multicomponent reactions contain three or more reagents in the same recipient to form a final product. These reactions have advantages over time, easy manipulation, experimental simplicity and atomic economy (Figure 113).¹⁶³

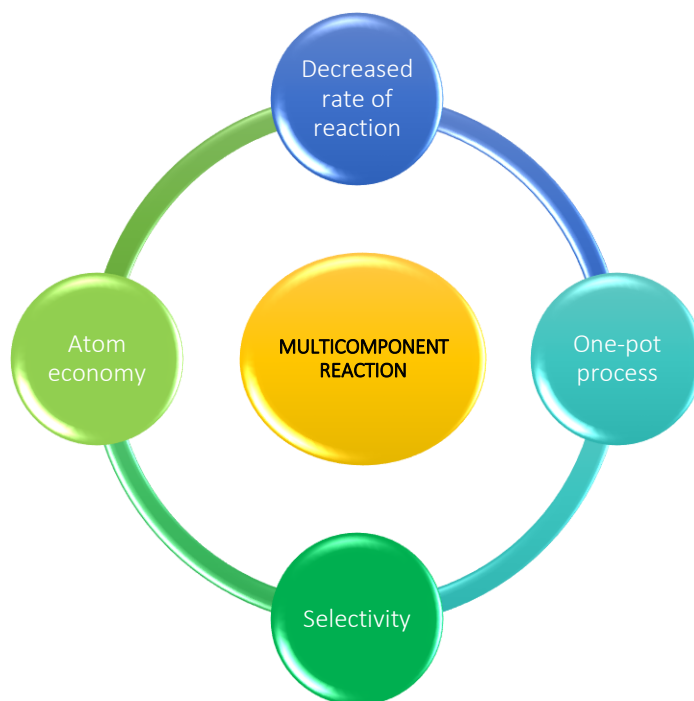
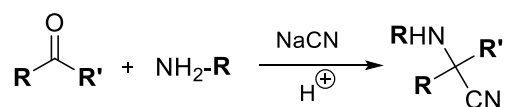


Figure 113. Advantages of multicomponent reactions¹⁶³

It is generally applied to classical reactions that develop in more than one stage and in this case are combined in a single reaction. Multicomponent reactions are common in Combinatorial Chemistry and in the synthesis of heterocyclic derivatives very common for the preparation of new drugs.¹⁶⁴

The first multicomponent reaction is the Strecker reaction (1850). This reaction involves the treatment of imines with aldehydes in the presence of sodium or potassium cyanide and in acidic medium to form amino-nitriles (Scheme 90).¹⁶⁵



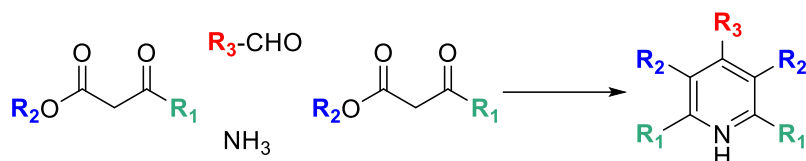
Scheme 90

¹⁶³ a) L. El Kaim, L. Grimaud. *Tetrahedron* **2009**, *65*, 2153-2171. b) R. W. Armstrong, A. P. Combs, P. A. Tempest, T. A. Keating. *Chem. Res.* **1996**, *29*, 123-131. c) T. J. J. Müller. *J. Org. Chem.* **2011**, *7*, 960-961.

¹⁶⁴ A. Dömling, W. Wang, K. Wang. *Chem. Rev.* **2012**, *112*, 3083-3135.

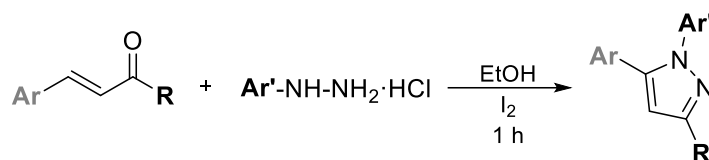
¹⁶⁵ S. Masumoto, H. Usuda, M. Suzuki, M. Kanai, M. Shibasaki. *J. Am. Chem. Soc.* **2003**, *125*, 5634-5635.

In 1882, Hantzsch first applied multicomponent reactions for the synthesis of heterocyclic systems. In this case, 1,4-dihydropyridines are obtained by condensation of β -ketoesters with aldehydes in the presence of ammonia (Scheme 91).¹⁶⁶



Scheme 91

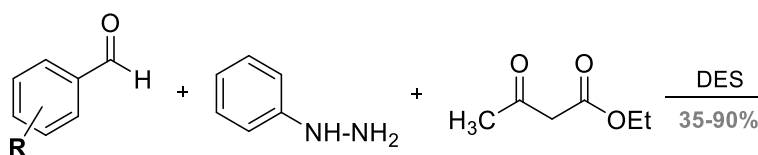
The classic preparation of pyrazolic nuclei is the condensation of aldehydes or α,β -unsaturated ketones with hydrazines in the presence of iodine, within ethanol (Scheme 92).¹⁶⁷



Scheme 92

In this case, for the preparation of [REDACTED] the application of a multicomponent reaction containing aldehyde or ketone, hydrazine and ethyl acetylacetate was considered. Different reaction conditions are described in the bibliography, and are summarized below:

- Shankarling et al.¹⁶⁸ describe the preparation of [REDACTED] using a Deep Eutectic Solvent (DES) and ultrasounds in water, thus obtaining the expected compounds with yields ranging from **35-90%** (Scheme 93).



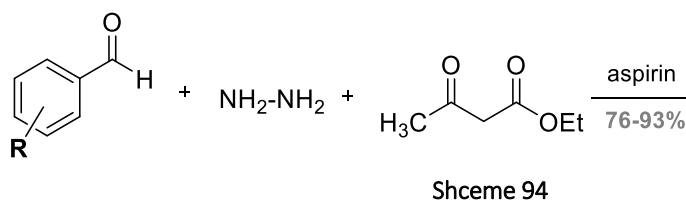
Scheme 93

¹⁶⁶ A. Kumar, R. A. Maurya. *Synlett* **2008**, 883-885.

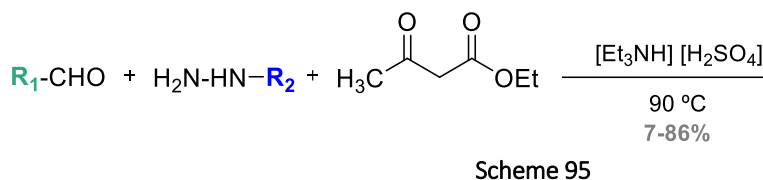
¹⁶⁷ X. Zhang, J. Kang, P. Niu, J. Wu, W. Yu. *J. Org. Chem.* **2014**, 79, 10170-10178.

¹⁶⁸ S. S. Kamble, G. S. Shankarling. *Chem. Select.* **2018**, 3, 2032-2036.

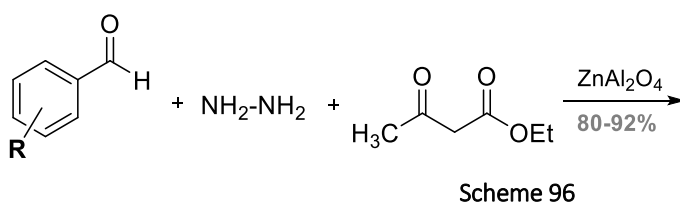
- Lashkari et al.¹⁶⁹ present the preparation [redacted] with the same type of condensation indicated above but using the acetylsalicylic acid (aspirin) as a catalyst (Scheme 94).



- Zhou et al.¹⁷⁰ describe the use of trimethylamine bisulfate as a catalyst and operate without solvent at 90 °C (Scheme 95).



- Safaei-Ghoni¹⁷¹ describes condensation of aldehyde, hydrazine and ketoester in the presence of ZnAl₂O₄ in the form of nanoparticles to obtain [redacted] with yields ranging from 80 to 92% (Scheme 96).



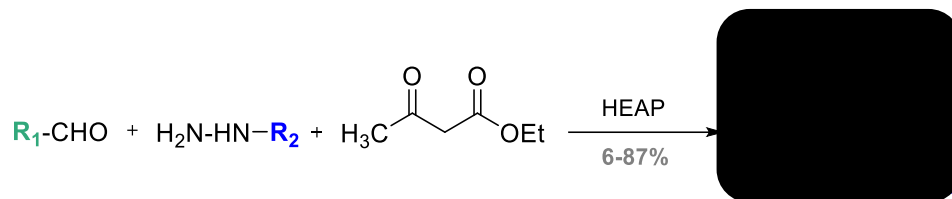
- Zhou et al.¹⁷² describe the condensation with the same compounds of the previous reaction but using 2-hydroxy-ethylammonium propionate (HEAP) and no solvent. The yields obtained range from 6-87% (Scheme 97).

¹⁶⁹ M. Fatahpour, F. N. Sadeh, N. Hazeri, M. T. Maghsoodlou, M. Lashkari. *J. Iran Chem. Soc.* **2017**, *14*, 1945-1956.

¹⁷⁰ Z. Zhou, Y. Zhang. *J. Chil. Chem. Soc.* **2015**, *60*, 2992-2996.

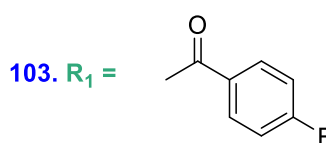
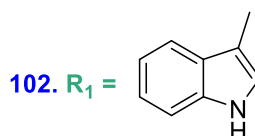
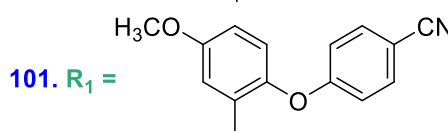
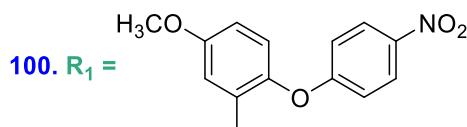
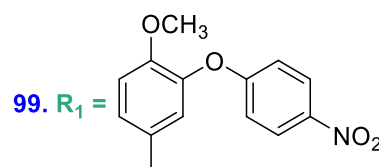
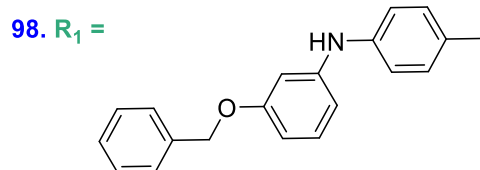
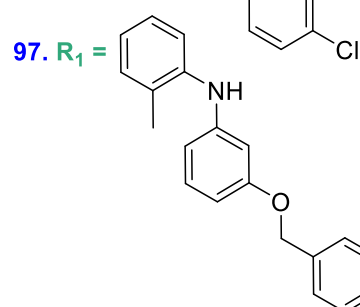
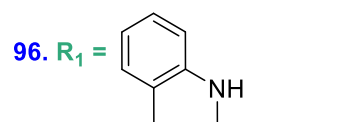
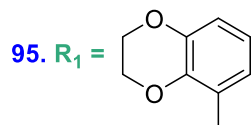
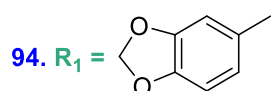
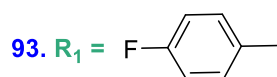
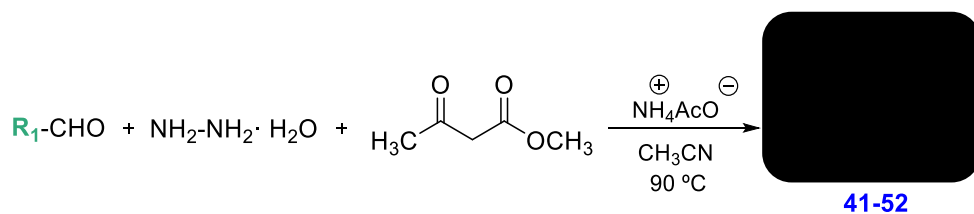
¹⁷¹ J. Safaei-Ghoni, B. Khojastehbakht-Koopaei, H. Shahbazi-Alavi. *RSC. Adv.* **2014**, *4*, 46106-46113.

¹⁷² Z. Zhou, Y. Zhang. *Green Chem. Lett. Rev.* **2014**, *7*, 18-23.



Scheme 97

In this work, the condensation of different aryl aldehydes with hydrazine hydrate and methyl acetoacetate was carried out under multicomponent reaction conditions, and using ammonium acetate as a catalyst (Scheme 98).



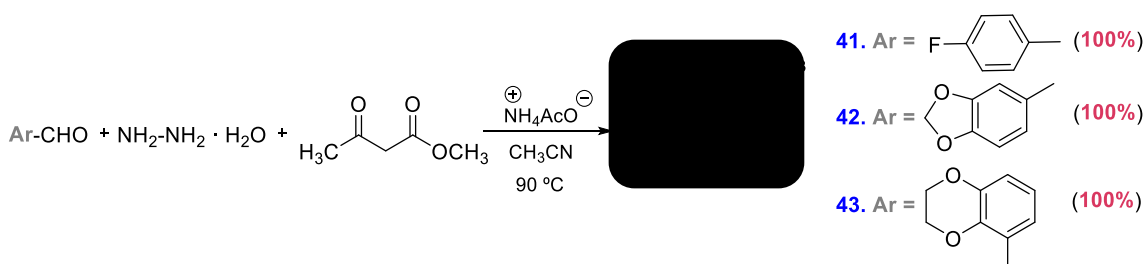
Scheme 98

3.2.2.1. Preparation of [REDACTED]

(41), [REDACTED] (42) and

(43)

Applying the indicated methodology (Scheme 99), [REDACTED] 41, 42, 43 were obtained



with 100% yield.

Scheme 99

It should be noted that the obtained compounds had a high polarity that hinders solubility in organic solvents. The signals corresponding to the [REDACTED] nucleus 6.99 (dt, H-2', H-6') and 7.12 ppm (dt, H-3', H-5') as well as the singlet of the [REDACTED] position that resonated at 4.72 ppm allowed to confirm the structure proposed for [REDACTED] 41.

On the other hand, the singlets at 5.98 corresponding to the methylene group and at 4.64 ppm attributable to the benzylic proton confirmed the structure of 42.

In the case of [REDACTED] 43, signals were perceived in the ^1H NMR spectrum at 4.14 assignable to the [REDACTED] group and a singlet at 4.99 corresponding to the proton of the [REDACTED] position (Figure 114).

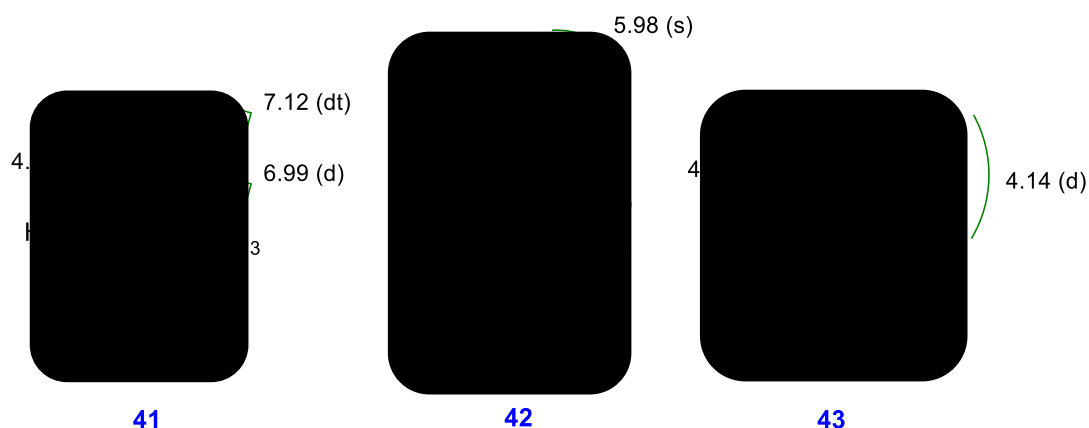


Figure 114. ^1H NMR spectroscopic data of 41, 42 and 43

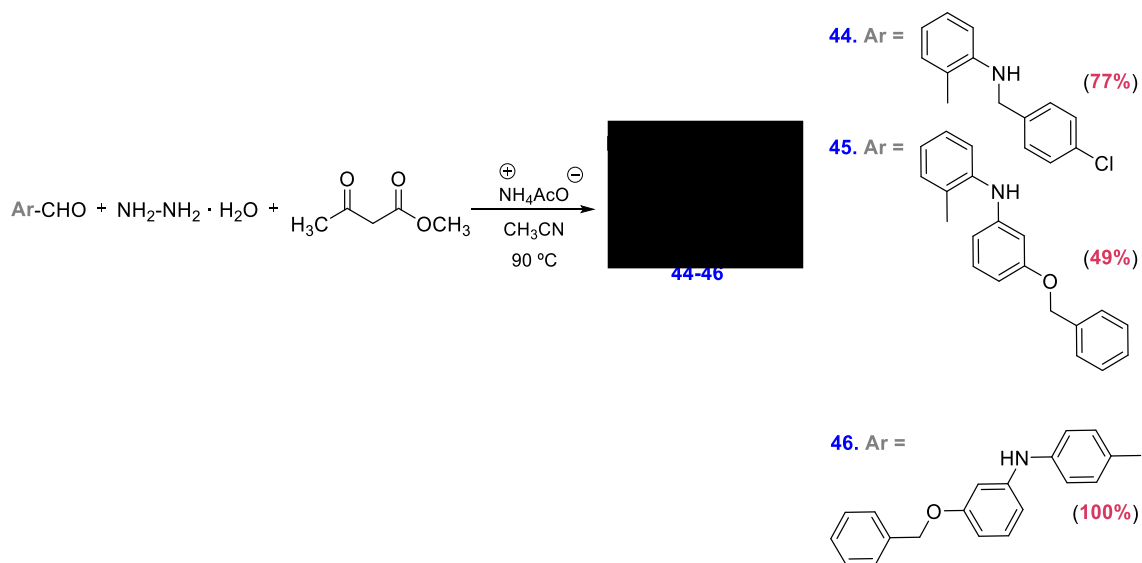
3.2.2.2. Preparation of

(44),

(45) and

(46)

44, 45 and 46 were obtained by the condensation of the corresponding arylaldehydes 96-98 with hydrazine hydrate and with methyl acetylacetae (Scheme 100) following the methodology developed in this thesis.



Scheme 100

These were completely identified by the ^1H NMR data (Figure 115). It is worth highlighting the singlets at 4.32 and 4.42 ppm that integrated one and two protons for 44. Likewise, 45 presented two singlets at 4.80 and 5.72 ppm corresponding to the protons (CH- and $-\text{O-CH}_2-$). With respect to 46, the singlet stood out at 5.01 (Ar-CH-) and 5.19 ($-\text{O-CH}_2-$).

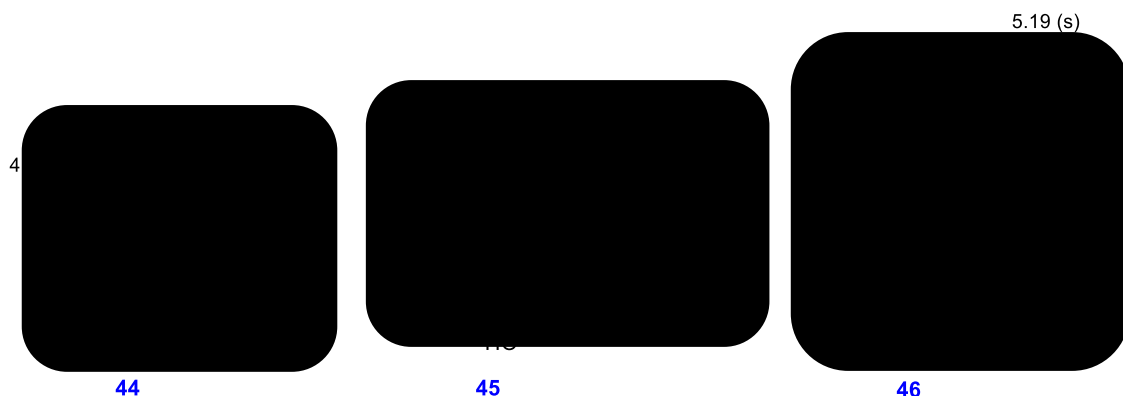
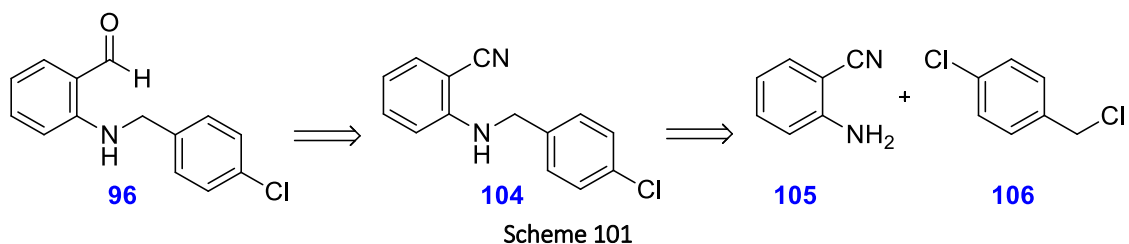
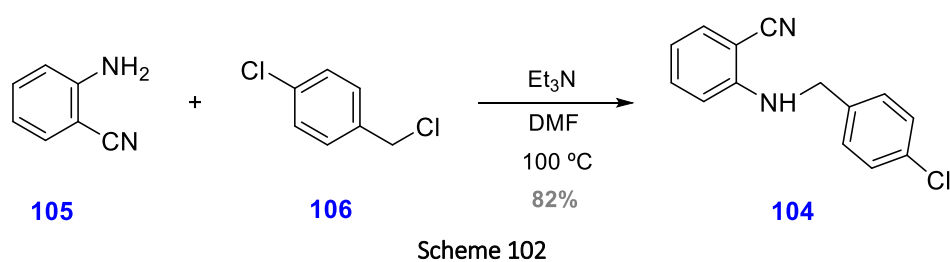


Figure 115. ^1H NMR spectral data of 44, 45 and 46

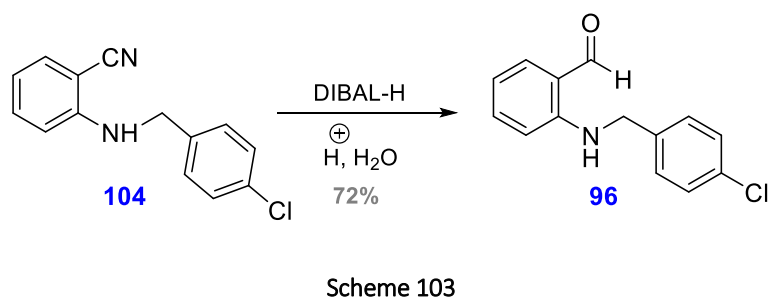
Aldehyde **96** was obtained according to the following retrosynthesis (Scheme 101).



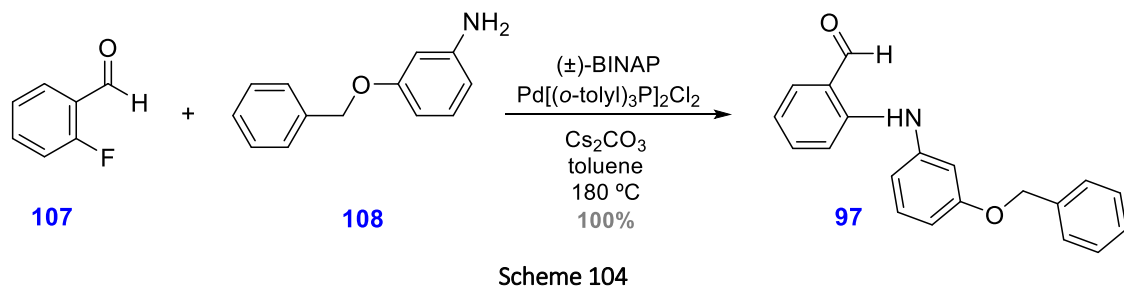
For the preparation of **104**, the alkylation of aniline **105** was performed with *p*-chlorobenzyl chloride, heated to 100 °C, in DMF (Scheme 102).



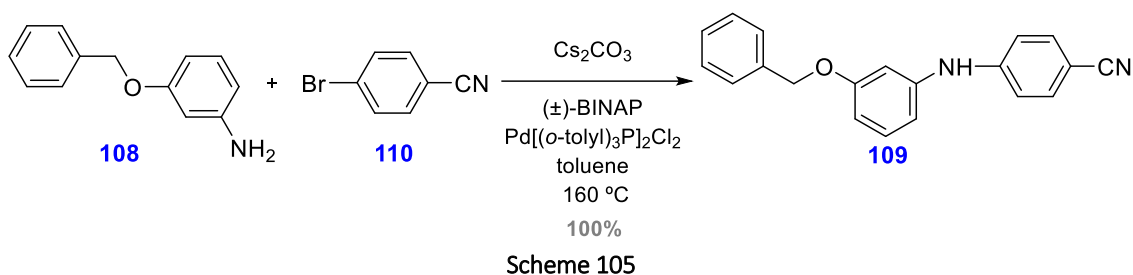
Then, selective reduction of benzonitrile **104** with DIBAL-H led to the expected aldehyde **96** with 72% yield (Scheme 103).



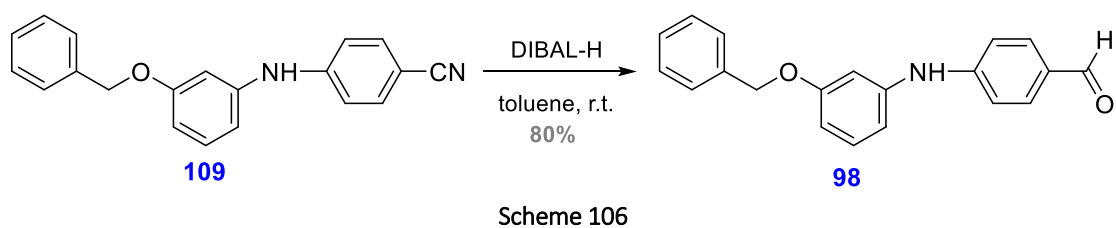
Aldehyde **97** was prepared from 2-fluorobenzaldehyde (**107**) and 3-benzyloxyaniline (**108**) through cross-coupling reaction conditions by palladium derivatives as catalyst. This is how the expected product was obtained with 100% yield (Scheme 104).



Aldehyde **98** was also prepared from 4-bromobenzonitrile (**110**) and 3-benzyloxyaniline (**108**) by palladium catalysed cross-coupling reaction and in the presence of (±)-BINAP as a ligand and caesium carbonate playing the base role (Scheme 105).



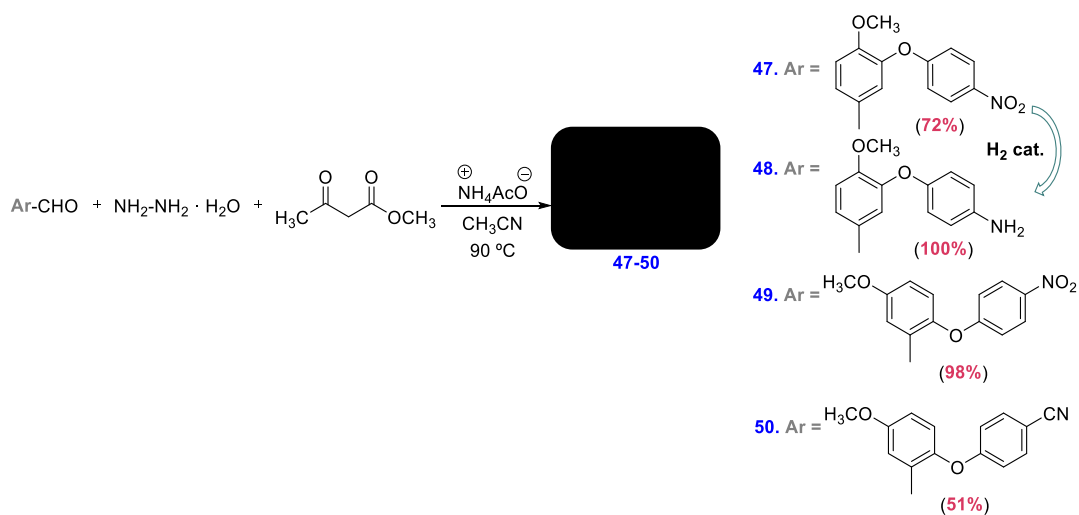
Benzonitrile **109** was reduced with DIBAL-H, within the toluene, and provided aldehyde **98** with 80% yield (Scheme 106).




3.2.2.3. Preparation of

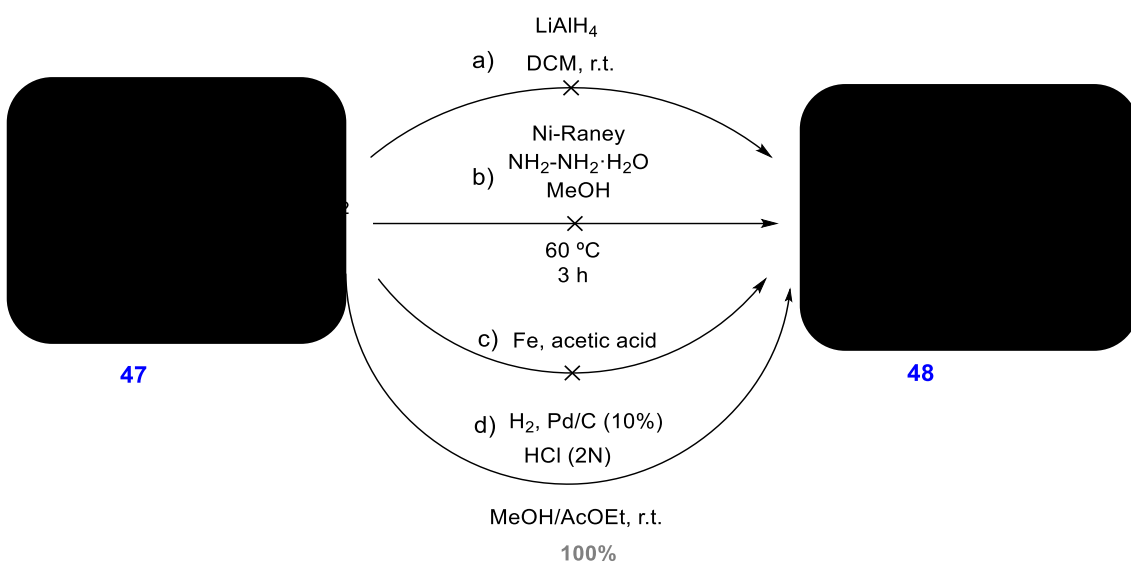
[REDACTED] (**47**), [REDACTED]
[REDACTED] (**48**), [REDACTED]
[REDACTED] (**49**) and [REDACTED]
[REDACTED] (**50**)

The condensation of synthesized aldehydes **99**, **100** and **101** with hydrazine hydrate and methyl acetylacetate under “one-pot” conditions as previously indicated allowed access to [REDACTED] **47**, **49** and **50** respectively with yields of **72**, **98** and **51%** (Scheme 107).




Scheme 107

The  **47** was attempted to be reduced to the corresponding aniline by different methods (LiAlH_4 , Ni-Raney and also with Fe in acidic medium) but none of these procedures provided the corresponding anilinic derivative. However, the catalytic hydrogenation of **47** led to the reduced derivative **48** with 100% yield (Scheme 108).



Scheme 108

Singlets at 4.48, 4.90, 4.80 and 5.20 ppm appeared in the ^1H NMR spectra of compounds **47-50**, attributable to the proton at the  position of each one of them, which allowed confirming that the condensation reaction was successful (Figure 116).

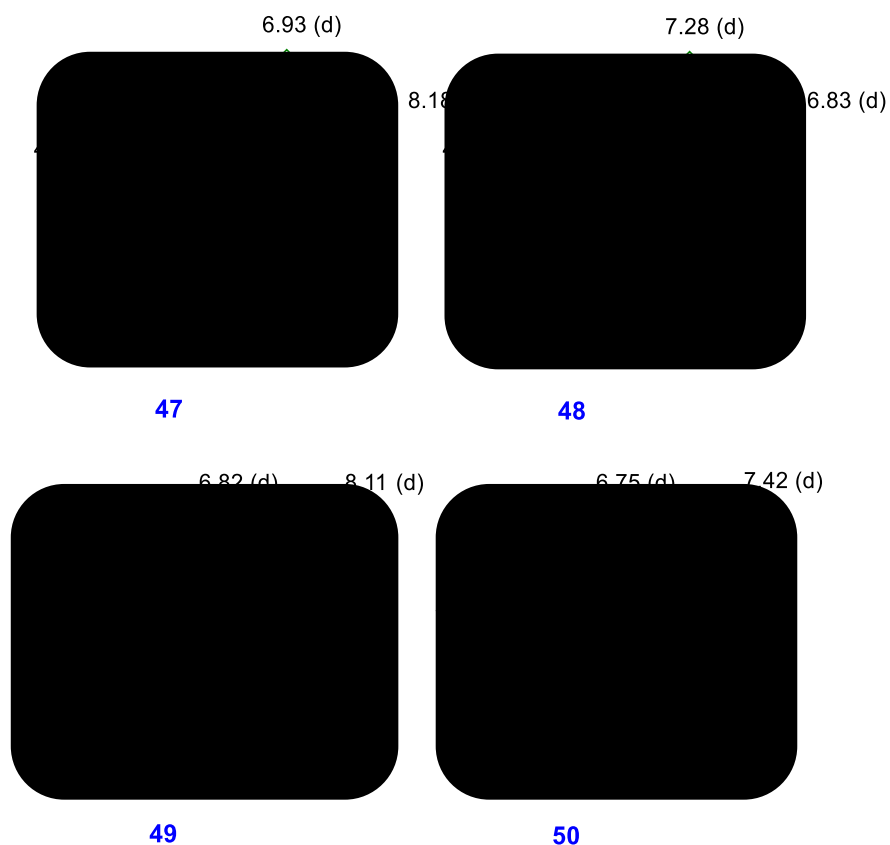
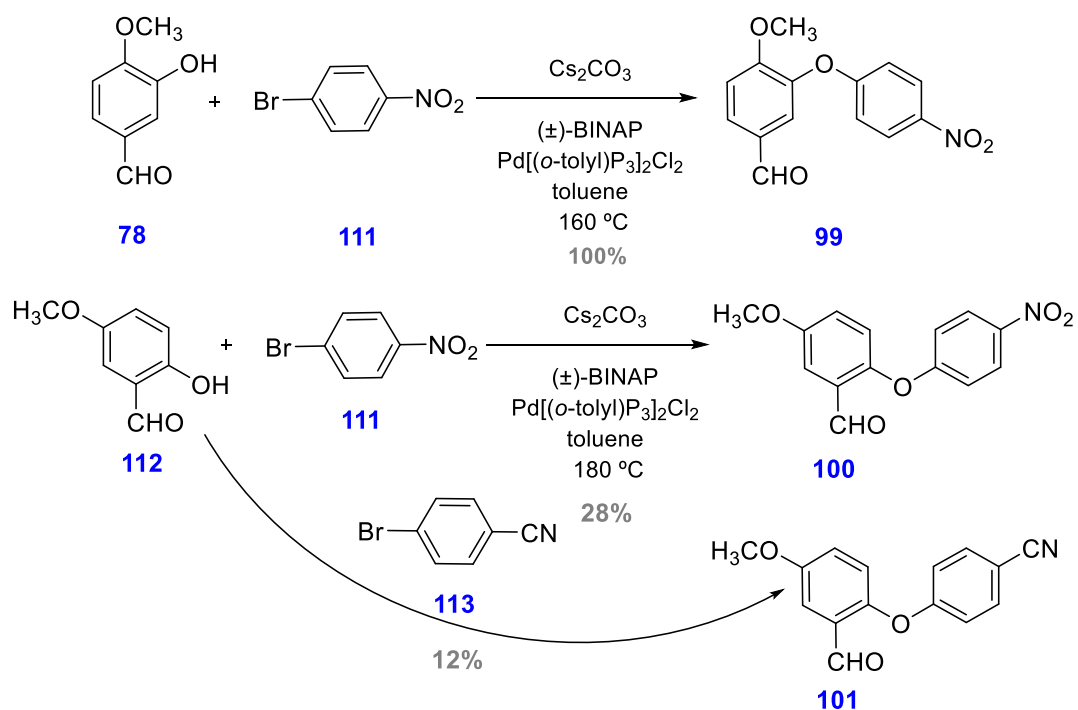


Figure 116. ^1H NMR spectroscopic data of **47-50**

Aldehydes **99**, **100** and **101** were obtained from the properly substituted phenol by a cross-coupling reaction with the corresponding aryl bromide (Scheme 109).



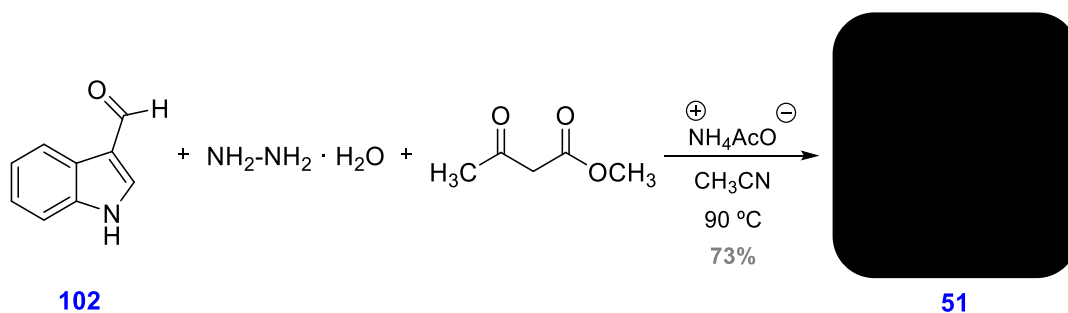
Scheme 109

The formation of diaryl ethers is not as described as diaryl amines and it is usually more difficult than these. A methodology for obtaining these types of compounds was recently developed in our research group.¹⁷³

3.2.2.4. Preparation of [REDACTED]

(51)

The application of the condensation conditions applied to benzaldehydes were also applied to heterocyclic aldehydes such as 3-indolecarbaldehyde (**102**) and thus the formation of [REDACTED] **51** was achieved with 73% yield (Scheme 110).



Scheme 110

In the ^1H NMR spectrum of [REDACTED] **51**, a singlet was perceived at 5.01 (CH-Ar) and another at 7.28 ppm (proton of position 2 of the indole), signals that corroborated the proposed structure for [REDACTED] **51** (Figure 117).

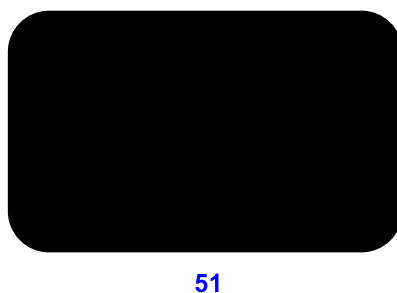


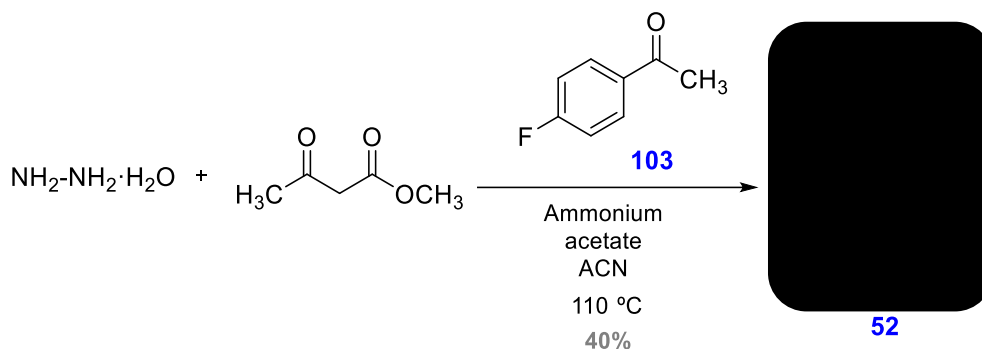
Figure 117. ^1H NMR spectral data of **51**

¹⁷³ L. Navarro, M. D. Pujol. *Tetrahedron Lett.* **2015**, *56*, 1812-1816.

3.2.2.5. Preparation of the [REDACTED]

[REDACTED] (52)

In order to study the viability of the multicomponent condensation reaction carried out in the preparation of the previous compounds, acetyl-*p*-fluoro-benzene (**103**) was applied to the ketone condensation, thus achieving compound **52** with **40%**, lower yields than those obtained with aldehyds (Scheme 111).



Scheme 111

In addition, to indicate that so far no reference was found in the literature in regard to these types of condensations applied to ketones.

The presence of a singlet at 2.20 ppm assignable to the methyl group bounded to the benzylic carbon in the ^1H NMR spectrum of **52** should be noted (Figure 118).

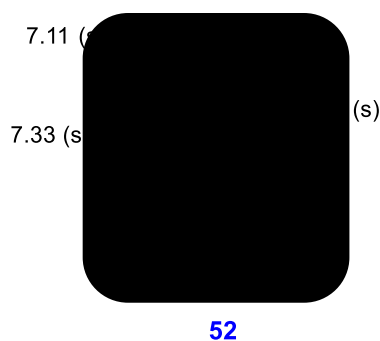
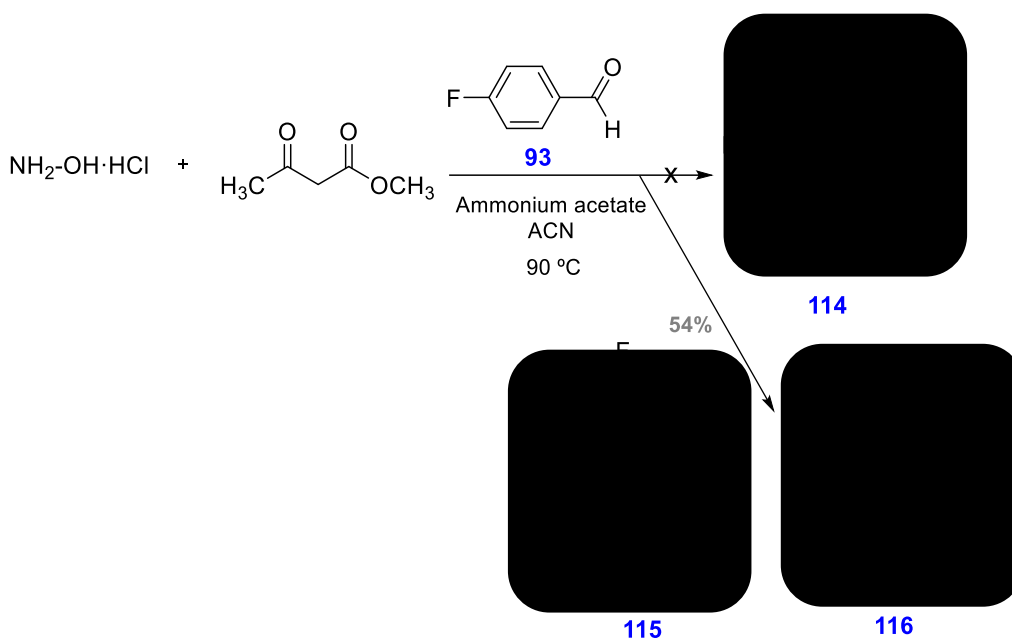


Figure 118. ^1H NMR spectral data of **52**

At the same time, this multicomponent reaction was extended to the preparation of isoxazoles using hydroxylamine instead of hydrazine hydrate and, in this case the desired isoxazolone **114** was not achieved (Scheme 112).



Scheme 112

Under these conditions, a mixture of isoxazoles **115** and **116**, were identified by NMR and mass spectrometry (Figure 119).

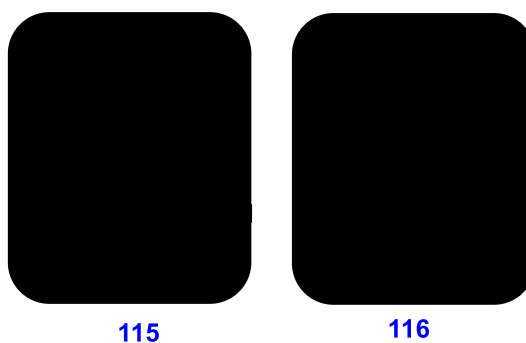


Figure 119. Isoxazoles **115** and **116**

3.3. Biological results

3.3.1. Biological results related to cardiovascular diseases

3.3.1.1. Biological activity evaluated at the Istituto Superiore di Sanità (Rome)

As explained above cardiovascular diseases are the leading cause of death in Spain.

For this reason it was decided to evaluate the antioxidant activity of two synthesized compounds, compound **21** and compound **38**. These two compounds are potential PCSK9 protein inhibitors. Due to its therapeutic indication, it is intended to evaluate their biological activity as antioxidants against pro-inflammatory mechanisms in *in-vitro* cultures of human umbilical vein endothelial cells (HUVEC).

These types of cells were chosen because they allow conducting studies related to endothelial dysfunction, such as oxidative stress, angiogenesis, hypoxia, inflammatory processes, cardiovascular complications, method of action and protection of different compounds in CV diseases, and the signalling of different pathways. The ultimate goal is to use the synthesized compounds to counteract the oxidative stress or inflammation produced by TNF- α (Tumor Necrosis Factor Alpha), as well as to see the ICAM-1 (Intracellular Adhesion Molecule 1) levels in response to the activation or damage to the endothelial cells.

To determine the cytotoxicity of compounds **21** and **38**, the test WST-8/CCK8 (2-(2-methoxy-4-nitrophenyl)-3-(4-nitrophenyl)-5-(2,4-disulfophenyl)-2H-tetrazolium), was performed, which is a colorimetric method that allows to measure the metabolic activity of cells, and therefore determines cell viability (number of cells). The yellowish tetrazole WST-8 is reduced by metabolically active cells to another product called *formazan* (orange), which can be quantified by spectrophotometric methods. Therefore, if the cells are not viable, they will not be able to carry out this reaction.

3.3.1.1.1. Cell viability

The compounds were incubated at increasing concentrations for 24 hours on two different cell lines: HUVEC/EGM-2 (Endothelial Cell Growth Medium 2) and HUVEC/EBM-2 (Endothelial Cell Growth Basal Medium 2). The first cell line was in growth conditions where the growth factors were included, whereas in the second cell line there were no growth factors and only endothelial cells were present. Therefore, it is evident that the growth of the cells in the second cell line was more difficult and for this reason the cell viability should be reduced.

In these tests, cell viability is compared in percentage to the control with the different tested concentrations of the compound. Below are two bar graphs corresponding to each cell line.

- Compound **21**

As shown in Figure 121, cell viability is higher under growth conditions, since under EBM-2 conditions cell viability is decreased at almost all concentrations. A significant decrease in cell viability is perceived at 50 μ M, which is the highest concentration.

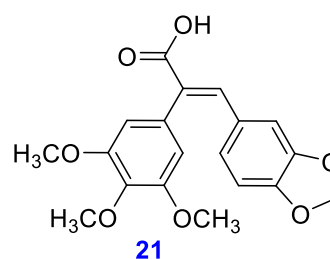


Figure 120. Structure of compound **21**

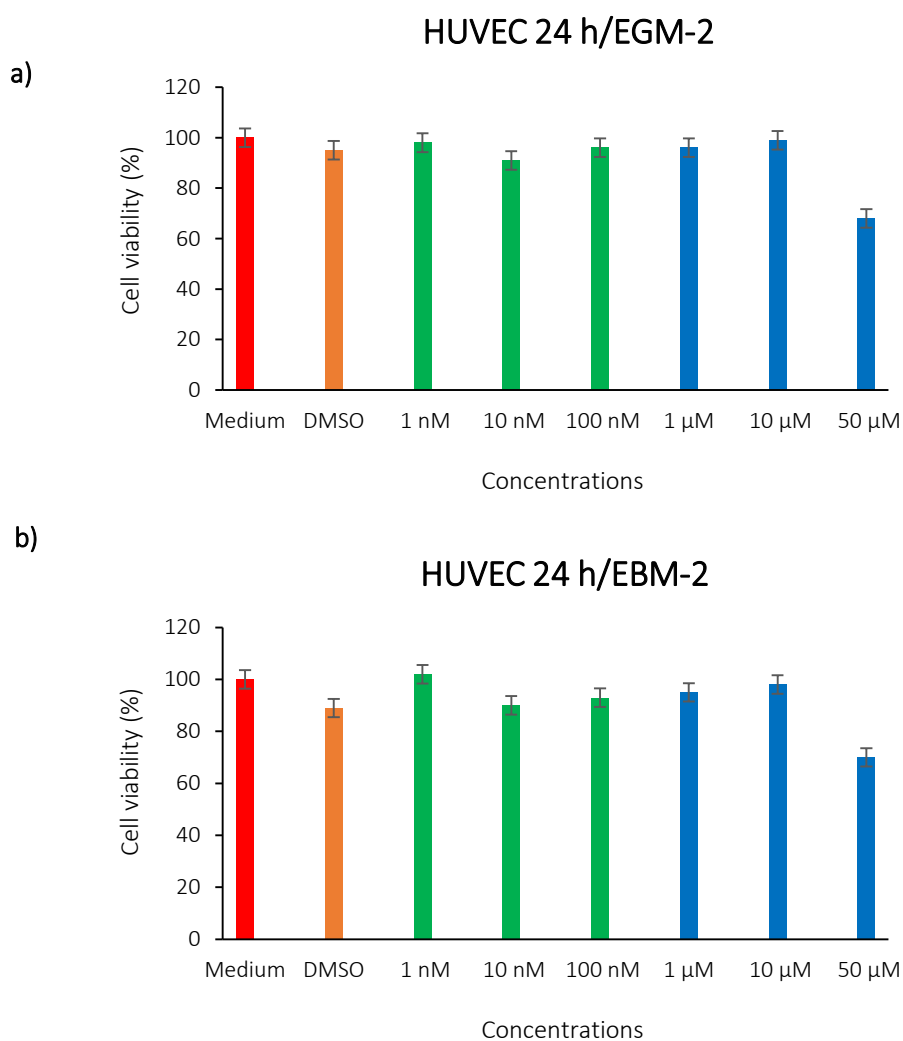


Figure 121. Relation between cell viability and concentration of compound **21**: **a)** Under growing conditions; **b)** In EBM-2 conditions

- Compound **38**

Cell viability remained above 80% at all tested concentrations under growth conditions, whereas under EBM-2 conditions there is a significant decrease at the highest tested concentration (Figure 123).

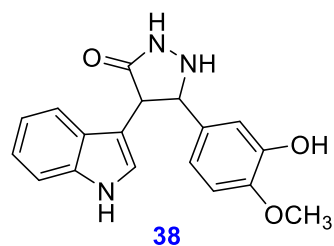


Figure 122. Structure of compound **38**

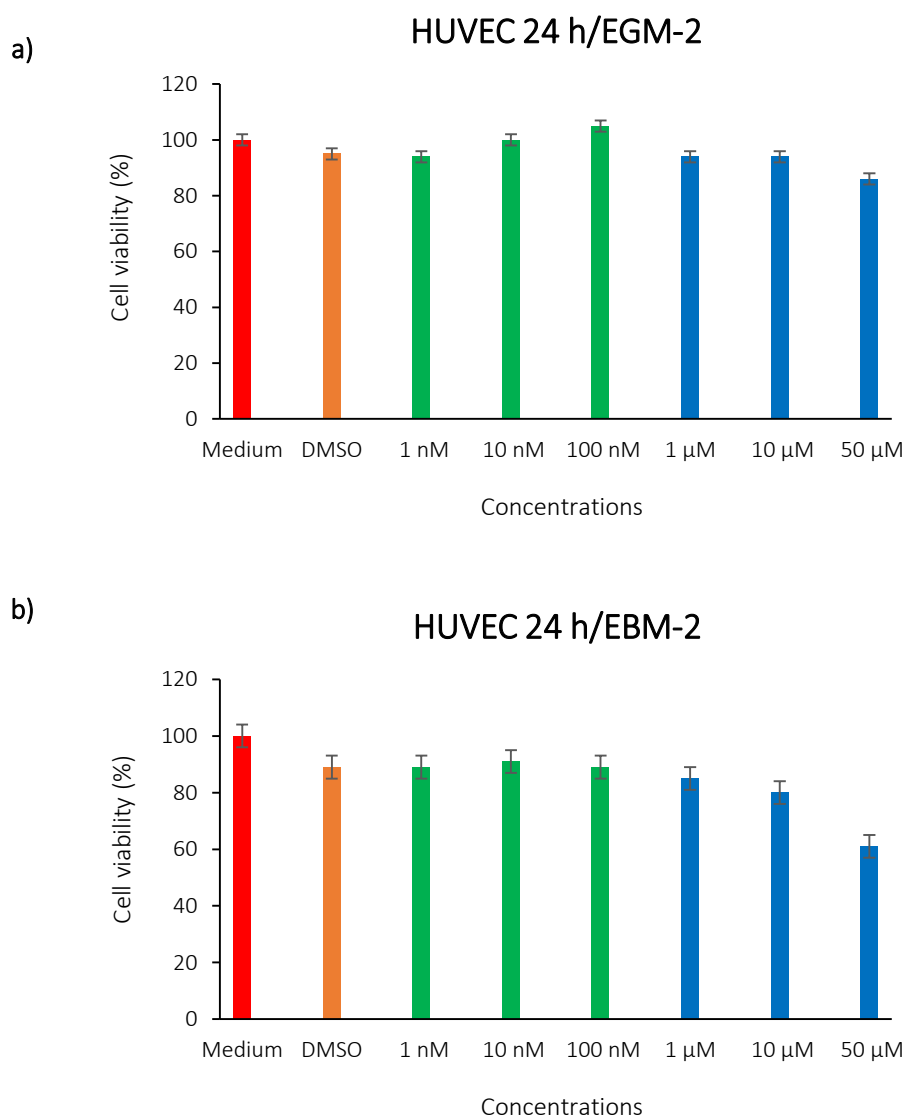


Figure 123. Relation between cell viability and concentration of compound **38**: **a)** Under growth conditions; **b)** In EBM-2 conditions

As it can be seen, both compounds could be used in all tested concentrations (except at 50 μM) as they do not significantly decrease the number of endothelial cells.

3.3.2. Biological results related to KRAS

3.3.2.1. Biological activity evaluated at the Hospital Clinic

As explained earlier in Section 1 of this present thesis, a mutation in KRAS protein leads to the development of different types of cancer, specifically, the colorectal cancer. It is for this reason that different compounds were synthesized to be tested in DLD-1 (KRAS^{WT/G13D}) human colorectal adenocarcinoma cell line in order to obtain different potential inhibitors of the protein and, consequently, colorectal cancer inhibitors. In this case the compounds tested were **41**, **42**, **46** and **47**.

These types of cells were chosen because they allow conducting studies related to MAPK-signalling pathway, proliferation and angiogenesis; as can be seen all associated in carcinogenic processes. And, the most important, the mutation G13D is found in the 20% of colorectal cancers.

The main goal of this study is to determine which compounds have the ability to inhibit the signalling pathways of the MAPKs and which of these could be a potential inhibitor or activator of KRAS protein.

At first, to determine the cytotoxicity of compounds **41**, **42**, **46** and **47**, the test MTS 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, was performed. This test is used to measure cellular metabolic activity as an indicator of cell viability, proliferation and cytotoxicity. It is a colorimetric assay in which the active cells are available to convert the MTS (yellow pale tetrazolium salt) to a purple *formazan*, which can be quantified by spectrophotometric methods. If the cells are not viable they can not perform this reduction.

After performing the cell viability assay the activation of ERK and AKT were assessed by Western Blot (WB) using specific antibodies against the phosphorylated (active) forms of these proteins.

As it can be seen, the results include the fact that compound **47** kills colorectal cancer cells containing KRAS oncogenic mutation without affecting normal cells. More assays were carried out to extend the studies with compounds **P14** and **47**, such as cell signalling and cell survival tests using pancreatic ductal adenocarcinoma cancer (PDAC) cell lines.

3.3.2.1.1. Cell viability in colorectal adenocarcinoma cell line

The compounds were incubated at increasing concentrations for 24 hours on DLD-1 (KRAS^{WT/G13D}) human colorectal adenocarcinoma cell line at different concentrations (25 μ M, 50 μ M, 75 μ M, 100 μ M and 125 μ M). The results are also compared with the Retinal Pigment Epithelial (RPE) cell line which is used as control cell line.

In this test, cell viability is compared in percentage to the control with the different tested concentrations of the compound. Below are two line graphs corresponding to the cell viability of compounds **42** and **47** (Figure 124) which show cell death. However, compounds **41** and **46** show no effect in cells of the colorectal adenocarcinoma cell line.

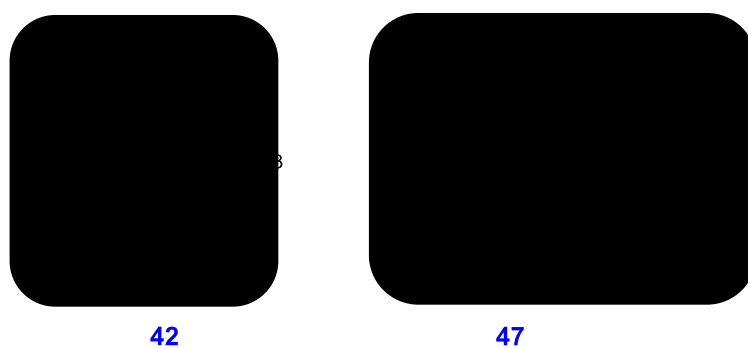


Figure 124. Structure of compounds **42** and **47**

- Compound **42**

A significant decrease in cell viability was perceived at 100 μ M (Figure 125). There are no significant differences between DLD-1 and DLD-1 KO. However, the differences in cell viability between DLD-1 and RPE are significant at 100 μ M.

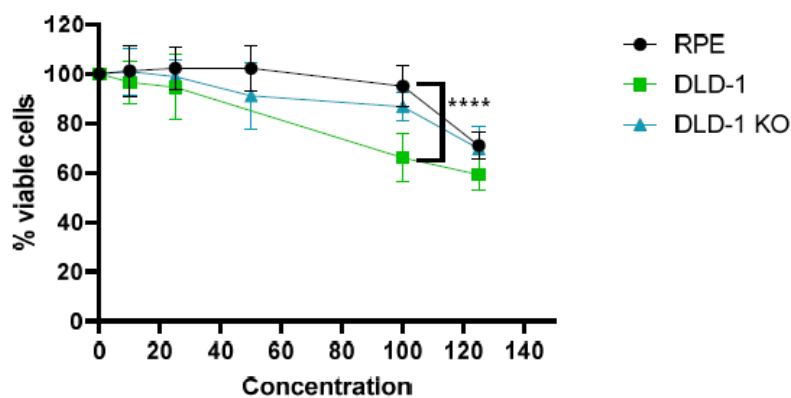


Figure 125. Relation between cell viability and concentration of compound **42**

- Compound 47

As can be seen in Figure 126 the differences in cell viability between RPE and DLD-1 KO; and, RPE and DLD-1 are significant at 50 μ M.

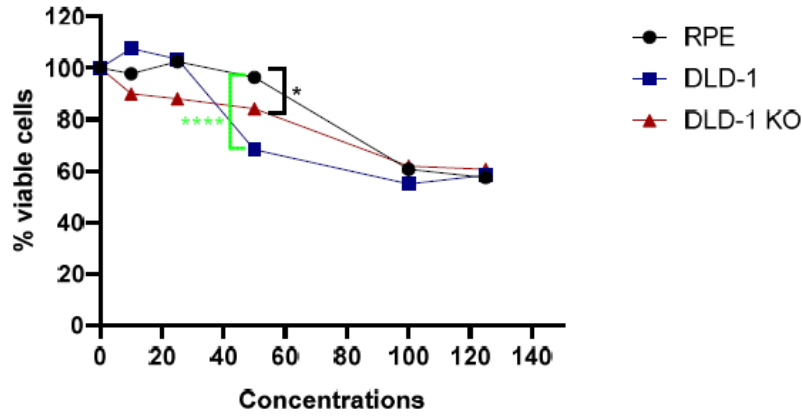


Figure 126. Relation between cell viability and concentration of compound 47

3.3.2.1.2. Cell signalling analysis in colorectal adenocarcinoma cell line

The compounds were incubated at serum limiting conditions for 24 hours. After that time, the cells were treated with 100 μ M for all compounds for 3 hours. The total amount and phosphorylation levels of RAF, ERK, and AKT proteins were detected by Western Blot, with specific antibodies against said proteins and against their phosphorylated forms (Figure 127).

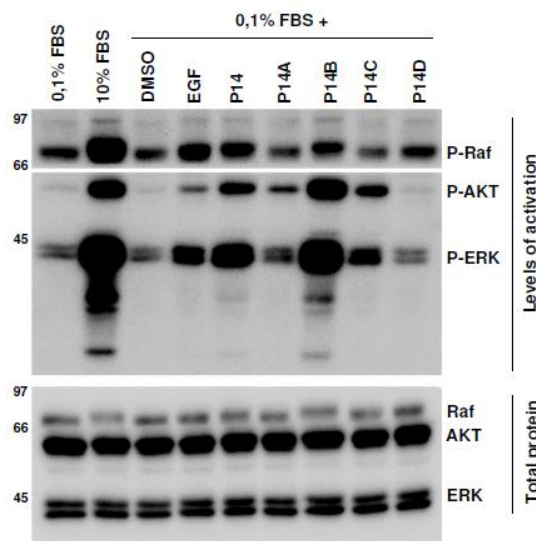


Figure 127. Western blot using antibodies against proteins of the MAPK pathway in the presence of compounds P14, 41 (P14A), 47 (P14B), 42 (P14C) and 46 (P14D)

As shown in Figure 123 compound **P14** and **47** activates RAF, ERK and AKT. However, compound **47** activates at higher levels than **P14** itself.

Compound **41** as shown in the results, would surely have no effect on the signalling pathway of MPAKs.

Compound **42** only activates ERK and AKT but have no effect on RAF. Otherwise, compound **46** inhibits ERK and AKT.

As mentioned above, the goal is to search new inhibitors of the KRAS protein or its pathway. That is why the results obtained are at least different from those expected. In this case there are three pathway activators (**P14**, **47** and **42**) and one inhibitor (**46**). However, the activation process also leads to cell death. This suggests that the mechanism of action of these compounds are different from any of those already described in the literature.

In this particular case, the activating compounds are structurally similar to **P14**, where two aromatic rings with small substituents are found. In contrast, the inhibitory compound has a much larger structure, which probably will be bound to another binding site of KRAS protein (Figure 128).

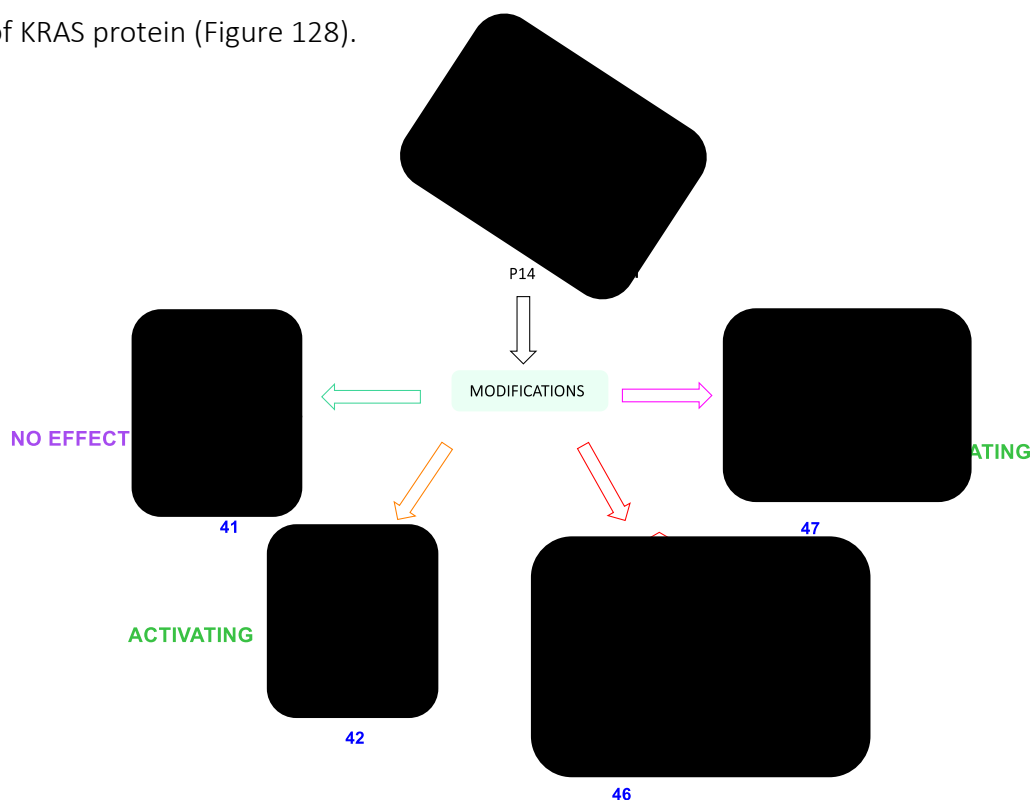


Figure 128. Result produced by the compounds on the MAPK pathway

3.3.2.1.3. Cell viability in PDAC cell line

The compounds **P14** and **47** were incubated at increasing concentrations for 24-48 hours on six different PDAC cell lines and the non-transformed immortalized cell line RPE (control cell line). The PDAC cell lines used were the following: 8988-T, PANC-1, SW1990, MPANC-96, HPAF and PA-TU. The MTS assay was performed, and the results obtained show that no of these compounds has any effect on the PDAC cells.

Table 21. Cell survival after 24-28 hours of treatment (MTS)

DRUG	RPE	8988-T	PANC-1	SW1990	MPANC-96	HPAF	PA-TU
P14	No data	∅	∅	∅	∅	∅	∅
47	∅	∅	∅	∅	∅	∅	∅

∅ No effect

3.3.2.1.4. Cell signalling in PDAC cell line

The compounds **P14** and **47** were incubated at serum limiting conditions for 3 hours by using the same PDAC cell lines explained above. The total amount and phosphorylation levels of RAF, ERK, and AKT proteins were detected by Western Blot, with specific antibodies against said proteins and against their phosphorylated forms. The results obtained are summarized in Table 22.

Table 22. Cell signalling after 3 hours at serum limiting

DRUG	RPE	8988-T	PANC-1	SW1990	MPANC-96	HPAF	PA-TU
P14	+ ERK	+AKT/ERK	+AKT/ERK	+AKT/ERK	+AKT/ERK	+AKT/ERK	+AKT/ERK
47	∅	∅	∅	∅	∅	∅	∅

+ Activation; ∅ No effect

As shown in Table 22 **P14** activates ERK and AKT proteins in the six cell lines, whereas compound **47** has no effect in any of the six PDAC cell lines.

As the study of compounds **P14** and **47** did not show significant results in PDAC cells another type of study will be proposed. With this in mind, analysis by co-immunoprecipitation will be performed in order to determine if compound **47** is breaking oncogenic KRAS interaction with their effectors P110/PI3K and RAF.

4. EXPERIMENTAL SECTION

4.1. Physical Chemistry laboratory materials and methods

4.1.1. Preparation of new inhibitory agents of PCSK9 protein involved in cholesterol homeostasis

4.1.1.1. Design of new enzyme inhibitors through molecular modeling studies

4.1.1.1.1. Search for new allosteric binding sites (Pockets)

System preparation

The initial crystallographic structure of PCSK9 was taken from the PDB with the access code 4NE9.¹⁰² This structure corresponds to a dimer, each one in complex with EFG-(A5) peptide (GTNECLDNNGGCASHVCNDLKIGYECL*),¹⁰² encoding a fragment of the LDL-R. Assuming that both monomers are very similar, it was decided to work with chain B, which corresponds to the catalytic and C-terminal domain (residues from 153 to 682), where the peptide and chains A, C and D were removed. The Swiss modeling program^{174,175} was then used to add the missing residues of the protein, the template was taken from the 4NE9 fasta sequence obtained from UNIPROT.¹⁷⁶ The structure obtained was complete with all missing residues.

Finally, hydrogens were added to all protein residues to their corresponding protonation states at pH 7 and the side chains orientations were established using the Protein Preparation Wizard module of Maestro Version 10.6.014.¹⁷⁷

Next, the protein was placed in a cubic periodic box filled with TIP3P¹⁷⁸ explicit water molecules, setting a minimum distance of 15 Å between the solute and the box walls. Water molecules closer than 2.0 Å to any complex atom were removed. Then, four Na⁺ ions were added to neutralize the system at the positions of lowest electrostatic potential using the LEaP module of AMBER 14.¹⁷⁹

¹⁷⁴ A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede. *Nucleic Acids Res.* **2018**, *46*, W296-W303.

¹⁷⁵ The UniProt Consortium. UniProt: the universal protein knowledge base. *Nucleic Acids Res.* **2017**, *45*, D158-D169.

¹⁷⁶ <https://www.uniprot.org/uniprot/Q8NBP7> (03/04/2020)

¹⁷⁷ Schrödinger Release 2020-1: Maestro, Schrödinger, LLC, New York, NY, **2020**.

¹⁷⁸ W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein. *J. Chem. Phys.* **1983**, *79*, 926-935.

¹⁷⁹ T. Case, D. T. E. Cheatham III, C. Simmerling, A. Roitberg, J. Wang, W. A. Göetz, I. Kolossváry. *Amber 14 Reference Manual*, University of California, San Francisco.

Energetic minimizations

Before starting the molecular dynamics calculations the structure was first relaxed. Thus, energetic minimizations were carried out in order to eliminate possible steric stress using the AMBER ffSB14¹⁸⁰. A cutoff of 10 Å was employed for the nonbonded interactions and, electrostatic interactions were treated using the particle-mesh Ewald (PME) method.¹⁸¹ Minimizations were performed following multi-step procedure: (1) minimization of water molecules and ions through 10,000 steps using the steepest descent method¹⁸² by keeping fixed the C α atoms of the protein by applying harmonic positional restrictions of 1 kcal/mol·Å⁻²; (2) all the atoms were allowed to move during 10,000 steps of the conjugate gradient method.¹⁸³

Molecular dynamics simulation

After minimization, the structure was heated to 300 K stepwisely at a rate of 30 K every 20 ps, fixing the C α atoms of the protein with harmonic positional restrictions of 0.5 kcal/mol·Å⁻², by coupling the system in a thermal bath, using Langevin thermostat algorithm¹⁸⁴ with a collision frequency of 2 ps⁻¹ under the NVT ensemble. Subsequently, 100 ps simulation was performed at constant pressure, (NPT ensemble (Isothermal-isobaric ensemble)), without keeping fixed any atoms of the protein for density equilibration. The SHAKE algorithm¹⁸⁵ was used in all the MD simulations to constrain all bonds involving hydrogen atoms. Mass repartitioning was also used to allow the use of an integration step of 4 fs.

Then, four different systems were prepared, differentiating each one by their initial velocity distribution that was generated at 300 K, 295 K, 305 K and 310 K. Once established the initial temperature, 300 ns length of GaMD was carried out by using a cutoff of 9 Å. The upper limit of the standard deviation of the total potential boost was set to 3 and the upper limit of the standard deviation of the dihedral potential boost was

¹⁸⁰ J. A. Maier, C. Martínez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling. *J. Chem. Theory Comput.* **2015**, *11*, 3696-3713.

¹⁸¹ T. Darden, D. York, L. Pedersen. *J. Chem. Phys.* **1993**, *98*, 10089-10092.

¹⁸² M. Chaichian, A. Demichev. *Path Integrals in Physics Stochastic Process and Quantum Mechanics*, Taylor and Francis, Vol. 1, **2001**, 174.

¹⁸³ M. R. Hestens, E. Stiefel. *J. Res. Natl. Bur. Stand.* **1952**, *49*, 409.

¹⁸⁴ M. P. Allen, D. J. Tildesley. *Computer Simulation of Liquids* (Second Ed.), Oxford University Press **2017**.

¹⁸⁵ J. P. Ryckaert, G. Ciuccotti, H. J. C. Berendsen. *J. Comput. Phys.* **1977**, *23*, 327-341.

set to 6. The results of the GaMD calculations were consequently used to study the flexibility of the protein.

Root-Mean Square Deviation (RMSD) and Root-Mean Square Fluctuation (RMSF)

The RMSD and RMSF were computed using the structures of the MD simulation and the starting structure as a reference, in order to measure the similarity among the two superimposed atomic coordinates. Cpptraj program from AMBER 14¹⁷⁹ was used to calculate the RMSD and RMSF. An iterative process was applied to perform the calculation, which was carried out in two steps: (1) superposition of the coordinates of all protein residues with the reference structure; (2) the residues with RMSF below 2 Å were selected to recalculate the RMSD and RMSF by superimposing the C α of the selected residues with the reference structure. That process provided information of the local conformational flexibility of the residues and allowed to identify the invariant nucleus of the protein, defined by a set of residues that exhibited the lowest fluctuations along the MD trajectory.

Radius of Gyration (Rg)

The radius of gyration was calculated for the four corresponding systems in order to measure the root-mean square distance of the selected atoms (C α residues of the protein) from their center. Cpptraj program from AMBER 14¹⁷⁹ was used. The calculation of the Rg brought information about the compactness of the protein.

Clustering

In order to select a group of structures representing the greatest structural diversity of the PCSK9 protein similar structures in GaMD simulations were grouped into 10 different clusters using the average linkage algorithm¹⁸⁶, as implemented in the cpptraj module of AMBER 14.¹⁷⁹ For this process, the RMSD of the C α located in the secondary structure of the protein with a larger RMSF was used as distance. A total of 220 amino acids were selected.

¹⁸⁶ L. Rokach, O. Maimon. *Clustering Methods, Data Mining and Knowledge*, Eds: Springer, Boston, M.A., 2005, 321-352.

Principal components analysis (PCA)

The PCA permits to reduce the number of dimensions needed to describe protein dynamics through a decomposition process that describes motions from the largest to the smallest spatial scales. A covariance matrix (3N x 3N symmetric matrix) was built up using the atomic coordinates of the C α of every residue of the secondary structure and then was diagonalized to produce a set of eigenvectors, as well as their corresponding eigenvalues. To perform these calculations all the frames resulting from the four GaMD, which corresponded to the 1200 ns length, were grouped using the cpptraj module of AMBER 14.¹⁷⁹

Binding site and affinity prediction

As the allosteric binding sites for this protein have not been discovered yet, the aim of this work is to identify the pockets of the protein. For this purpose, different programs were used, such as FTMap,^{103,104} DeepSite¹⁰⁵ and PockDrug.^{106,107,108,109,110} In addition, Fpocket^{111,112} was used to determine a score drug of the discovered pockets. To carry out this job the representatives of the first five clusters (0, 1, 2, 3, 4) obtained from clustering were used.

FTMap^{103,104} is a computational mapping server which works in the following way: the structures were uploaded and the same program used probes (small organic molecules) in order to score the probe poses using a detailed energy expression in order to find the hot spot of an average-size protein in less than one hour. The site that contained the largest number of probes was considered the main hot spot. The hot spots were classified according to their Boltzmann averaged energies and the probes with the lowest average energies were retained. The program used 16 different organic molecules as probes (ethanol, isopropanol, isobutanol, acetone, acetaldehyde, dimethyl ether, cyclohexane, ethane, acetonitrile, urea, methylamine, phenol, benzaldehyde, benzene, acetamide, and *N,N*-dimethylformamide).

This program also provides the druggability. A site to be druggable must comply the following requirements: (1) the cluster strength (*S*), described as the number of probes within the main hot spot, which should be $S > 16$ to be druggable by some ligand.

If the cluster strength, $S < 13$, it is considered an undruggable site; (2) the existence of at least one possible hot spot within 8 Å (measured distance from the center of mass between two hot spots) from the main hot spot.

Deepsite program¹⁰⁵ defines a pocket considering: (1) the distance of a predicted point for a pocket have to be closer than a known distance to the geometric center of the real binding site. It is considered a pocket if the values are ranged between 4 and 20 Å; and, (2) the discretized volumetric overlap, which considers the pocket shape and size of the predicted pocket.

The PDB structures were uploaded and with 1-2 minutes the program identified the main pockets, then a maximum of 3 predefined pockets were obtained. A whole predicted volumetric map of the pocket was provided where the distance of each pocket center was 2 Å and 3 predicted centers were attained.

PockDrug program^{106,107,108,109,110} have two different ways to define a pocket of a protein. One based on ligand proximity information and the second method is based on decomposing a protein structure into Voronoi polyhedrals. The second method was carried out, which is basically a geometry-based method. It extracted all the pockets from the protein surface using spheres with different diameters. The druggability was given as well, if it was higher than 0.5 it was considered a druggable pocket.

Fpocket program^{111,112} is a geometric based pocket detection algorithm based on the concept of alpha spheres. The program use the QHULL library to perform Voronoi tessellation to gather alpha spheres. An alpha sphere is a sphere that contacts four atoms on its boundary and contains no internal atom. Alpha sphere radii reflect the local curvature defined by the four atoms. At first the calculation of the alpha spheres and the followed clustering was performed. Then, merged nearby pockets and removed the small ones and finally, the score and rank pockets were calculated.

4.1.1.1.2. Screening of libraries (Docking)

Protein preparation

The structure of the protein was obtained from the Protein Data Bank (PDB) with the entry code 4NE9.¹⁰² This structure corresponds to a dimer, each one in complex with the truncated peptide EFG-(A5) (residues from 293 to 318)¹⁰². This peptide interacts with the PCSK9 similarly to the wild-type EFG-(A). Assuming that both monomers are very similar, it was decided to work with one monomer composed by chain B, C and D. Where chain C is the prodomain (residues from 61 to 152), chain B corresponds to the catalytic domain (residues from 153 to 454) and C-terminal domain (residues from 455 to 682), and chain D match with the EFG-(A5) peptide (GTNECLDNNGGC SHVCNDL KIGYECL*).¹⁰² The Swiss modeling program,^{174,175} was then used to add the missing residues in the protein, the template used was taken from the fasta 4NE9 sequence obtained from UNIPROT.¹⁷⁶ The structure obtained was complete with all missing residues (169-175, 450-453, 573-581, 660-662).¹⁷⁶

Finally, hydrogens were added to all protein residues to their corresponding protonation states at pH 7 and the side chains orientations were established using the Protein Preparation Wizard module of Maestro Version 10.6.014.¹⁷⁷

Next, the protein was placed in a cubic periodic box filled with TIP3P¹⁷⁸ explicit water molecules, imposing a minimum distance of 15 Å between the solute and the box walls. Water molecules closer than 2 Å to any complex atom were removed. Then, 5 Na⁺ ions were required to neutralize the system at the positions of lowest electrostatic potential to neutralize the complex using the Leap module from AMBER 14.¹⁷⁹

Energetic minimizations

Before starting the docking process, the structure was first relaxed. Thus, energetic minimizations were carried out in order to eliminate possible steric stress using AMBER ff14SB¹⁸⁰ force field. A cutoff of 10 Å was employed for the nonbonded interactions and, electrostatic interactions were treated using the PME method.¹⁸¹ Minimizations were performed following multi-step procedure: (1) minimization of the protein through 10,000 cycles by using the steepest descent method¹⁸² was performed by keeping fixed the C α atoms of the peptide, through applying harmonic positional restrictions of

1 kcal/mol·Å⁻²; (2) the protein and the peptide were allowed to move without any restriction during 10,000 cycles.

Molecular dynamics simulation

The minimized structure was heated to 300 K, the temperature was raised at a constant rate of 30 K every 20 ps, fixing the C α atoms of the peptide with a force constant of 0.5 kcal/mol·Å⁻², by coupling the system in a thermal bath, using Langevin thermostat algorithm¹⁸⁴ with a collision frequency of 2 ps⁻¹ under the NVT ensemble.

The density of the system was equilibrated around 1 g/mL by carrying out 1 ns of simulation. The C α atoms of the peptide were not allowed to move by applying a constant strength of 0.5 kcal/mol·Å⁻² and a cutoff of 10 Å. The SHAKE algorithm¹⁸⁵ was used in all the MD simulations to constrain all bonds involving hydrogen atoms allowing to use an integration step of 2 fs.

Finally, 100 ns of cMD was performed within the NVT ensemble. No restrictions were imposed to any atom of the system. The coordinates of the trajectory were written every 5,000 steps (10 ps).

Two different systems were settled up. [System 1](#), where the receptor was the prodomain (residues from 61 to 152) and the catalytic and C-terminal domain (residues from 153 to 682), while the peptide (residues from 293 to 318) was the ligand. And, [System 2](#), where the receptor was the peptide (293-318) and the catalytic domain (153-682) and, the prodomain (residues from 61 to 152) was the ligand.

The selection of the receptor and the ligand was based on the idea of carrying out a docking process where the willing was to find protein inhibitors that could bind to the peptide site, and also inhibitors that could bind in the catalytic domain. For that reason, the knowledge of the stability and energy convergence of the peptide and the prodomain over time was so important for the following steps.

Clustering

The method used for clusterization was the hierarchical agglomerative or bottom-up approach.¹⁸⁶ This method consisted in grouping the last 4,000 frames corresponding to

the last 40 ns of cMD. The program used was cpptraj module of AMBER 14¹⁷⁹. The residues near 4.5 Å from the peptide or the prodomain were used to be superposed with the reference structure. For **System 1**, the C α root-mean square deviation of residues 153-155, 158, 195, 222, 237-239, 243, 367, 369, 372, 374-381 were used, while the residues taken for **System 2** were 186-187, 209, 212, 226, 253, 257-263, 288-292, 297-314, 316-318, 326-327, 384-387; as a measurement of the distance between two conformations. The rationale of using these residues was that they were close to the ligand and therefore more likely to interact with the peptide or the prodomain depending on the system to study. As a result, for **System 1** and **2**, two and one clusters were obtained respectively. Obtaining these clusters allowed to identify the main structural conformations of the protein over the last 40 ns length of cMD.

Binding free energy calculations

Free energy of binding was calculated using the Molecular Mechanic Generalized Born (MM-GBSA)¹⁸⁷ algorithm enable in the AMBER package.

For each complex, the binding free energy could be calculated with the following equation:

$$\Delta G^{\circ}_{\text{bind}} = \Delta G^{\circ}_{\text{complex}} - \Delta G^{\circ}_{\text{ligand}} - \Delta G^{\circ}_{\text{receptor}}$$

$$\Delta G^{\circ} = \Delta H^{\circ} - T\Delta S^{\circ} + \Delta G_{\text{solv}}$$

Where enthalpy (ΔH°) is the internal energy (ΔU°), which is calculated by AMBER. ΔU is the sum of different terms, corresponding to the electrostatic ($\Delta G^{\circ}_{\text{elec}}$), $\Delta G^{\circ}_{\text{vdw}}$ and $\Delta G^{\circ}_{\text{intramolecular}}$ molecular mechanics energies. The intramolecular energy is decomposed by stretching, torsion angle and dihedral link. The entropic term depends on the vibrational state of two binding partners. While, the solvation free energy (ΔG_{solv}) is the sum of polar and non-polar terms. ΔG_{polar} is calculated by the Generalized Born (GB) method¹⁸⁸ and the $\Delta G_{\text{non-polar}}$ is calculated according to the equation $\Delta G_{\text{non-polar}} = \gamma\text{SASA} + \beta$. Where the linear combination of pairwise overlap (LCPO) method¹⁸⁹ was used to

¹⁸⁷ P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, T. E. Cheatham. *Acc. Chem. Res.* **2000**, *33*, 889-897.

¹⁸⁸ V. Tsui, D. A. Case. *Biopolymers* **2000**, *56*, 275-291.

¹⁸⁹ J. Weiser, P. S. Shenkin, W. C. Still. *J. Comp. Chem.* **1999**, *20*, 217-230.

calculate the solvent-accessible surface area (SASA), and the values for γ and β constants were 0.0072 kcal/mol·Å⁻² and 0.00 kcal/mol respectively.

In this present work, the Onufiev-Bashford-Case (OBC)¹⁹⁰ generalized born method (igb=2) was used. The MM-GBSA¹⁸⁷ was calculated for the 100 ns length of cMD for both systems ([System 1](#) and [System 2](#)) using as starting point the corresponding minimized structures. The energy values corresponded to the full trajectory.

Per-residue energy decomposition

In order to find a trustworthy pharmacophore for docking calculations, the per-residue energy decomposition was calculated using the MMGBSA.py module of AMBER 14.¹⁷⁹ The construction of the pharmacophore was based on target bond pharmacophore modeling approach, where the per-residue energy decomposition had to be calculated. This method consists of calculating the contribution of each residue to the total binding free energy (MM-GBSA).¹⁸⁷

The calculations were performed for the 100 ns of cMD. For each system the binding free energy was calculated in order to obtain the highly energetically contributing amino acids needed for the pharmacophore construction.

Hydrogen bound analysis

The hydrogen bonds established between each receptor-ligand was determined using the cpptraj module from AMBER 14¹⁷⁹ over the last 40 ns of cMD. The parameters used to define a hydrogen bond were the shown below:

- The distance between the acceptor hydrogen and the donor had to be less than 3.5 Å
- The angle (θ) between the acceptor and donor hydrogen must be equal or greater than 135°
- The percentage at which the hydrogen bond satisfies both conditions was quantified with an occupancy rate of >60%

¹⁹⁰ A. Onufriev, D. Bashford, D. A. Case. *Proteins Struct. Funct. Genet.* **2004**, 55, 383-394.

Pharmacophore construction

Once the clusters for each systems and the amino acids with the highest energy contribution were identified, the pharmacophore was built. In this case, the receptor structure (PCSK9) was known and the pharmacophore was constructed through the amino acids of the ligand with highest energy contribution.

Three pharmacophores were made. For **System 1** there were 2 clusters so 2 pharmacophores were constructed based on the MM-GBSA decomposition results.

While for **System 2**, just one cluster was found, so one pharmacophore of the prodomain was built.

The program used was MOE¹⁹¹ which permits the identification of different pharmacophore points, hydrophobic, donor and acceptor hydrogen bond. The query editor tool from MOE was used to design the pharmacophore points where the radius of each point was 1.5 Å.

Docking

The docking calculations were performed using the MOE program.¹⁹¹ Two different databases (Selleck Drug[®] and Fragments[®]) with possible candidates to bind favorably to the receptors were used. The Selleck Drug[®] database contained 915 approved drugs while the Fragments[®] database contained 1488 fragments of these drugs.

The ligand positioning was based on the pharmacophore positioning method.¹⁹² Once the ligand was anchored to the binding site (pharmacophore points) was possible to refine the poses with energy minimization and as well as near the ligand receptor side chains. The London dG function¹⁹³ was used to determine the binding free energy of each compound pose in the first step while a force field function was used in the second.

¹⁹¹ MOE (The Molecular Operating Environment) Version 2009.10, software available from Chemical Computing Group Inc., 1010 Sherbrooke Street West, Suite 910, Montreal, Canada H3A 2R7.

¹⁹² D. Oranit, D. Schneidman-Duhovny, Y. Inbar, R. Nussinov, H. J. Wolfson. *J. Chem. Inf. Model.* **2009**, *49*, 2333-2343.

¹⁹³ C. R. Corbeil, C. I. Williams, P. Labute. *J. Comput. Mol. Des.* **2012**, *26*, 775-786.

System preparation (clusters+compounds from databases)

Prior energetic minimization, the system was prepared. The parameters of the selected compounds required for the molecular dynamics simulations were obtained using Antechamber program from AMBER 14¹⁷⁹ software package. Next, each system obtained in the docking process was placed in a cubic periodic box filled with TIP3P¹⁷⁸ explicit water molecules, imposing a minimum distance of 15 Å between the solute and the box walls. Water molecules closer than 2 Å to any complex atom were removed. Then, counter-ions were added at the positions of lowest electrostatic potential in order to neutralize the system by using the LEaP module from AMBER 14.¹⁷⁹

Energy minimization

Before starting the molecular dynamics simulations, the 500 selected compounds with the best free energy of binding were first relaxed. Thus, energetic minimizations were carried out in order to eliminate possible steric stress using AMBER ff14SB¹⁸⁰ force field for the protein and gaff (General Amber Force Field)¹⁹⁴ for the selected compounds. The simulations were performed under PBC. A cutoff of 9 Å was employed for the nonbonded interactions and, electrostatic interactions were treated using the PME.¹⁸¹ Minimizations were performed following multi-step procedure: (1) minimization of the protein through 10,000 cycles using the steepest descent method¹⁸² by keeping fixed the atoms of the protein applying harmonic positional restrictions of 5 kcal/mol·Å⁻²; (2) the backbone atoms (C, C α , N) of the protein were fixed with the same constant strength, but the side chains, water molecules and counter-ions were allowed to move during 10,000 cycles; (3) finally, 10,000 cycles were performed without restrictions being the whole system allowed to move freely.

Binding free energy calculations

Once the structures were minimized, the MM-GBSA/PBSA methods¹⁸⁷ were used to calculate the binding free energy of the selected systems.

¹⁹⁴ J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, A. D. Case. *J. Comput. Chem.* **2004**, 25, 1157-1174.

ΔG_{polar} can be calculated by the Generalized Born (GB) method¹⁸⁸ or by solving the Poisson-Boltzman (PB) equation,¹⁹⁵ for MM-GBSA and MM-PBSA algorithms respectively. The $\Delta G_{\text{non-polar}}$ is calculated according to the equation $\Delta G_{\text{non-polar}} = \gamma \text{SASA} + \beta$. Where the linear combination of pairwise overlap (LCPO) method¹⁸⁹ was used to calculate the solvent-accessible surface area (SASA), and the values for γ and β constants were 0.0072 kcal/mol·Å⁻² and 0.00 kcal/mol for GB and, 0.00542 kcal/mol·Å⁻² and 0.92 kcal/mol for PB respectively.

In this present work, the Onufiev-Bashford-Case (OBC)¹⁹⁰ generalized born method (igb=2) was used. The MM-GBSA/PBSA¹⁸⁷ was calculated for the 10 ns length of cMD for the two studied systems ([System 1](#) and [System 2](#)).

Consensus

The 100 best energetically compounds were selected. For [System 1](#), as two clusters were selected and two similar processes were done, a self made program was used to identify and select which molecules and poses were the same in both representatives clusters. Then, the best compounds were picked to carry out molecular dynamic simulations. From Selleck Drug[®] database 74 compounds were selected and from Fragments[®] Database 63 compounds. On the other hand, for [System 2](#), the same program was used and, as a result, from Selleck Drug[®] and Fragments[®] databases, 64 and 77 compounds were respectively obtained.

Conventional molecular dynamics calculations

The energetically minimized structures were heated to 300 K. The temperature was raised at a constant rate of 30 K every 20 ps, fixing the atoms (C α , C, O, N) of the receptor with a force constant of 5 kcal/mol·Å⁻², by using Langevin thermostat algorithm¹⁸⁴ with a collision frequency of 2 ps⁻¹ under NVT ensemble.

The density of the system was equilibrated near 1 g/mL by carrying out 1 ns of simulation under NPT ensemble. The backbone atoms (C α , C, O, N) were fixed using a restrain of 5 kcal/mol·Å⁻², a cutoff of 9 Å was applied to the nonbonded interactions.

¹⁹⁵ R. Luo, L. David, M. K. Gilson. *J. Comput. Chem.* **2002**, *23*, 1244-1253.

The SHAKE algorithm was used in all the MD simulations to constrain all bonds involving hydrogen atoms allowing to use an integration step of 2 fs.

Finally, a production of 10 ns of cMD was carried out within NVT ensemble. No restrictions were imposed to any atom of the system. The coordinates of the trajectory were written every 5,000 steps (10 ps). For this step a cutoff of 9 Å was used to speed out the calculations.

Selection of the best compounds

The criterion of selection was carried out with the most likely to inhibit the PCSK9 protein based on the following requirements:

- The compound must be among the top ten energy positions (GB/PB)
- The compound must be positioned in a known pocket (defined by FTMap program)
- The calculated free energy of binding must tend to stabilize over the time, as a sign that the compound binds to the protein favorably

Afterward, the selected compounds were used to perform 40 ns of cMD within NVT ensemble. No restrictions were imposed to any atom of the system. The coordinates of the trajectory were written every 5,000 steps (10 ps).

4.1.1.1.3. Rational design of inhibitors based on the structure of the ligands

Protein preparation

Explained above at section 4.1.1.1.1. (*System preparation without solvation*).

BOX Ligands preparation

Six different ligands were built with MOE version 2009.10.¹⁹¹ Then, ligands **P1** and **P2** were energy minimized with a force field from MOE. Ligands **P4** and **P5** were energy minimized with AM1 hamiltonian from MOE. Ligand **P3** and **P8** were energy minimized with Gaussian, both using MP2/6-31G** level.

The parameters and charges of all ligands were calculated using Antechamber from AMBER16.¹⁹⁶ A repulsion center was introduced in all ligands in order to prevent the aggregation of the ligands in further calculations with many ligands in the system.

Each ligand was placed in a cubic box filled with TIP3P¹⁷⁸ explicit water molecules, setting a minimum distance of 15 Å between the solute and the box walls, and water molecules closer than 2 Å to any atom were removed. Then, counter-ions were added to neutralize the system at the positions of lowest electrostatic potential using the LEaP module of AMBER 16.¹⁹⁶ Minimization of water molecules and the ligand through 10,000 steps was carried out.

After minimization, the BOX was heated to 300 K stepwisly at a rate of 30 K every 20 ps, without fixing any atom of the protein, by coupling the system in a thermal bath, using Langevin thermostat algorithm¹⁸⁴ with a collision frequency of 2 ps⁻¹ under NVT ensemble. Subsequently, 100 ps simulation was performed at constant pressure (NPT ensemble) without keeping fixed any atom of the protein. The SHAKE algorithm¹⁸⁵ was used in all the MD simulations to constrain all bonds involving hydrogen atoms allowing to use an integration step of 2 fs.

Then, 100 ns of cMD was performed by using a cutoff of 9 Å and the coordinate's trajectory were written every 5000 steps (10 ps).

System preparation

The BOX's were used to solvate the protein, thus ligands were pieced together at different positions under PBC. A distance of 16 Å was established between the solute and the box walls, and water molecules closer than 1 Å to any atom were removed. Then, counter-ions were added to neutralize the system at the positions of lowest electrostatic potential using the LEaP module of AMBER 16.¹⁹⁶

¹⁹⁶ D. A. Case, R. M. Betz, D. S. Cerutti, T. E. Cheatham III, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao, P. A. Kollman. AMBER 2016, University of California, San Francisco.

The system was different for each in their ligands number. For system **1**, 34 molecules of ligand **P1** were placed. For systems **2**, **4** and **6**, the same number of molecules were placed of **P2**, **P4** and **P8** respectively.

In contrast, system **3** contained 22 molecules of **P3** and system **5** enclosed 25 molecules of **P5**.

Energy minimization

Before starting the molecular dynamics calculations the system was first relaxed. Thus, energetic minimizations were carried out in order to eliminate possible steric stress using AMBER ffSB14 force field.¹⁹⁶ A cutoff of 10 Å was employed for the nonbonded interactions and, electrostatic interactions were treated using the PME.¹⁸¹ Minimization of water molecules and ions through 10,000 steps using the steepest descent method¹⁸² without keeping fixed any atom of the protein.

Molecular dynamics simulation

After minimization, the system was heated to 300 K stepwisly at a rate of 30 K every 20 ps, without fixing any atom of the protein, by coupling the system in a thermal bath, using Langevin thermostat algorithm¹⁸⁴ with a collision frequency of 2 ps⁻¹ under NVT ensemble. Subsequently, 100 ps simulation was performed at constant pressure (NPT ensemble) without keeping fixed any atom of the protein. The SHAKE algorithm¹⁸⁵ was used in all the MD simulations to constrain all bonds involving hydrogen atoms allowing to use an integration step of 2 fs.

Accelerated molecular dynamics (GaMD)

Four different systems were prepared, differentiating each one by their initial velocity distribution that was generated at 300 K, 295 K, 305 K and 310 K. The upper limit of the standard deviation of the total potential boost was set to 3 and the upper limit of the standard deviation of the dihedral potential boost was set to 5. Finally, 300 ns length of GaMD was performed for all systems. Mass repartitioning was also used to allow the use of an integration step of 4 fs. The cutoff was 11 Å and a force switching region of 8 was settled up, where the force cutoff smoothly approaches 0 between the regions of the

fswitch value to the cut value. The coordinate's trajectory were written every 5,000 steps (20 ps).

fdMD program

In order to analyze properly each ligand trajectory the fdMD program was used. The ligands were separated independently from the system; the waters were removed and each ligand was detached from the protein in order to observe each molecular dynamics calculation. The interactions between the protein and the ligand were established and, as well as, the free energy of binding. For instance, ligand **P1**, which contained 34 ligands, after running the fdMD, as a result, the obtaining of 34 independent trajectories was achieved. This allowed the ligands to be analyzed independently of each other. With this in mind, the goal was to spot the location and the stability of the ligand in a concrete position.

Ligands selection

The process for selecting the best ligands was based on the following parameters:

- Stay more than 100 ns in the same position
- Be placed in a concrete pocket (defined by FTMap program)
- Energy convergence (MM-GBSA) from the last 100 ns of the GaMD calculation

After carrying out the selection process, each ligand was separated (fdMD program) from the GaMD simulation in order to obtain a system composed of a ligand and the protein, which were used for further calculations.

System solvation (protein+selected ligand)

Next, the system was placed in a cubic periodic box filled with TIP3P¹⁹⁷ explicit water molecules, setting a minimum distance of 15 Å between the solute and the box walls. Water molecules closer than 2.2 Å to any complex atom were removed. Then, counterions were added to neutralize the system at the positions of lowest electrostatic potential using the LEaP module of AMBER 16.¹⁹⁶

¹⁹⁷ W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein. *J. Chem. Phys.* **1983**, 79, 926-935.

Conventional molecular dynamics (cMD)

The system was heated to 300 K stepwisly at a rate of 30 K every 20 ps, fixing the backbone atoms (N, O, C, C α) of the system with harmonic positional restrictions of 1 kcal/mol·Å⁻², by coupling the system in a thermal bath, using Langevin thermostat algorithm¹⁸⁴ with a collision frequency of 2 ps⁻¹ under NVT ensemble. Subsequently, 400 ps of simulation was performed at constant pressure (NPT ensemble) by keeping fixed the backbone atoms (N, O, C, C α) of the protein and the ligand with harmonic positional restrictions of 1 kcal/mol·Å⁻². The SHAKE algorithm¹⁸⁵ was used in all the MD simulations to constrain all bonds involving hydrogen atoms to use an integration step of 2 fs. Afterward, 100 ns length of cMD calculations were performed using a cutoff of 9 Å, no atoms were fixed and the coordinates of the trajectory were written every 5,000 steps (20 ps). Mass repartitioning was also used to allow the use of an integration step of 4 fs. Then, the selection process was carried out again and 100 ns of cMD was performed under the same conditions.

Binding free energy calculations

The free energy of binding was calculated using the MM-GBSA algorithm¹⁸⁷ enable in the AMBER package¹⁹⁶ as explained above at section 4.1.1.1.2. (*Binding free energy calaculations*).

In this work, the OBC¹⁹⁰ generalized born method (igb=2) was used. Firstly, the MM-GBSA¹⁸⁷ was calculated for the 300 ns of GaMD simulation, and, after the selection process the free energy of binding was calculated for the 200 ns of cMD.

Ligands modification

The ligands selected were modified with MOE in order to increase their size. The modifications were based on the possible hydrogen bonds that they could create and according to the hydrophobic interactions that they could establish with the protein. Only ligands **P1**, **P2**, **P3** and **P4** were modified.

Ligands preparation

The modified ligands were energy minimized with Gaussian at the HF/6-31G** level. The parameters and charges of all ligands were calculated using the restrained electrostatic potential (RESP) module from Antechamber 16.¹⁹⁶

Next, the protein and the ligand were placed in a cubic box. The box was filled with TIP3P¹⁷⁸ explicit water molecules, setting a minimum distance of 16 Å between the solute and the box walls, and water molecules closer than 1 Å to any atom were removed. Then, counter-ions were added in order to neutralize the system at the positions of lowest electrostatic potential using the LEaP module of AMBER 16.¹⁹⁶

Energy minimization

Before starting the molecular dynamics calculation the structure was first relaxed. Thus, energetic minimizations were carried out in order to eliminate possible steric stress using AMBER ffSB14 force field.¹⁹⁶ A cutoff of 10 Å was employed for the nonbonded interactions and, electrostatic interactions were treated using the PME method.¹⁸¹ Minimization of water molecules and ions through 10,000 steps using the steepest descent method¹⁸² was performed without keeping fixed any atom of the protein.

Molecular dynamics simulation

After minimization, the system was heated to 300 K stepwisly at a rate of 30 K every 20 ps, fixing the backbone atoms (C α , C, O, N) of the protein and the ligand with harmonic positional restrictions of 1 kcal/mol·Å⁻²; by coupling the system in a thermal bath, using Langevin thermostat algorithm¹⁸⁴ with a collision frequency of 2 ps⁻¹ under NVT ensemble. Subsequently, 400 ps of simulation was performed at constant pressure (NPT ensemble) by keeping fixed the backbone atoms (C α , C, O, N) of the protein and the ligand for density equilibration after applying harmonic positional restrictions of 1 kcal/mol·Å⁻². The SHAKE algorithm was used in all the MD simulations to constrain all bonds involving hydrogen atoms. Mass repartitioning was also used to allow the use of an integration step of 4 fs.

Conventional molecular dynamics (cMD)

For system **1**, 100 or 300 ns length of cMD was carried out. On the other hand, for system **2**, **3** and **4** just 100 ns of cMD was performed. The restrictions were imposed to the the backbone atoms (C α , C, O, N) of the protein and the ligand applying harmonic positional restrictions of 1 kcal/mol·Å². Mass repartitioning was also used to allow the use of an integration step of 4 fs. A cutoff of 9 Å was used and the coordinate's trajectory were written every 5,000 steps (20 ps).

4.1.2. Preparation of new KRAS protein inhibitory agents

*4.1.2.1. Optimization of potential KRAS inhibitors or activators from lead **P14***

*4.1.2.1.1. Search of the binding site of the lead **P14** by performing gaussian accelerated molecular dynamics simulations (GaMD)*

Protein preparation¹⁰¹

The initial crystallographic structure of KRAS was taken from the PDB with the access code 4DSN.¹⁹⁸ This structure was loaded with the GTP, corresponding an active conformation. This protein carries an oncogenic mutation on codon 12 (G12D).

As there is no crystallography of the HVR yet, due to its high motility, the 11 lacking residues (from 178 to 188) were added and attached to the globular domain. The sequence added was as followed:

K-K-K-S-K-T-K-C-V-I-M

Then, the peptide chain was generated, protonated and then linked to the crystallography structure in the same coordinate system.

The MOE version 2009.10¹⁹¹ added the hydrogens to all the residues of the protein at pH 7.

¹⁹⁸ T. Maurer, L. S. Garrenton, A. Oh, K. Pitts, D. J. Anderson, N. J. Skelton, B. P. Fauber, B. Pan, S. Malek, D. Stokoe, M. J. Ludlam, K. K. Bowman, J. Wu, A. M. Giannetti, M. A. Starovasnik, I. Mellman, P. K. Jackson, J. Rudolph, W. Wang, G. Fang. *Proc. Natl. Acad. Sci.* **2012**, *109*, 5299-52304.

P14 Preparation

Gaussian calculations using a HF/6-31G* level were performed in order to optimize the ligand. Then, the parameters were calculated using RESP module from AMBER 16.¹⁹⁶

Complex preparation

The main idea in preparing the complex was to set the ligand in 8 random positions of the protein box with an intramolecular distance between the ligand and the protein of 15 Å. The HiConMD program was used and gave separately 8 different complexes, each differentiated by the position of the ligand.

Each complex was placed in a cubic periodic box filled with TIP3P¹⁷⁸ explicit water molecules, setting a minimum distance of 15 Å between the solute and the box walls. Water molecules closer than 2 Å to any complex atom were removed. The counter-ions were added in order to neutralize the system, where 7 Na⁺ ions were needed. LEaP module of AMBER 16¹⁹⁶ was used.

Energetic minimizations

Prior to GaMD, each complex was first relaxed. Each one was energy minimized using AMBER ff14SB.¹⁹⁶ A cutoff of 10 Å was employed for the nonbonded interactions and, electrostatic interactions were treated using the PME method.¹⁸¹ Only one step was needed to minimize the system, which consisted in 10,000 steps of the steepest descent method¹⁸² without fixing any atom of the protein.

Molecular dynamics simulation (MD)

After minimization, each complex was heated gradually to 300 K at a rate of 30 K every 20 ps, where no atoms of the protein were fixed, using Langevin thermostat algorithm¹⁸⁴ with a collision frequency of 2 ps⁻¹ under NVT ensemble. After heating, 100 ps of simulation was carried out at constant pressure (NPT ensemble) in order to equilibrate the density of the system (around $\rho = 1$ g/mL) without fixing any atom of the protein.

Accelerated molecular dynamics simulation (GaMD)

Once each complex was minimized, 1200 ns of GaMD was performed by using a cutoff of 12 Å. Mass repartitioning was also used to allow the use of an integration step of 4 fs. The upper limit of the standard deviation of the total potential boost was set to 3 and the upper limit of the standard deviation of the dihedral potential boost was set to 4. No restraints were applied.

Conventional molecular dynamics simulation (cMD)

cMD simulation was only performed with the complexes that showed the best energetic stability over the last 300 ns length of GaMD by using as starting point the last position of the complex.

A cMD of 300 ns length was performed by using a cutoff of 9 Å. Mass repartitioning was also used to allow the use of an integration step of 4 fs.

Binding free energy calculations

The free energy of binding was calculated using the MM-GBSA algorithm¹⁸⁷ enable in the AMBER package¹⁹⁶ as explained above at section 4.1.1.1.2. (*Binding free energy calaculations*).

In this present thesis, the OBC¹⁹⁰ generalized born method (igb=2) was used. The binding free energies were calculated for the 1200 ns of GaMD (for each complex). The best energetically complexes were used for the 300 ns lenght of cMD, where the last point of the GaMD was used as starting point for the cMD. Then, the MM-GBSA was also calculated for the 300 ns.

4.2. Organic Chemistry laboratory materials and methods

Proton and carbon nuclear magnetic spectroscopy (^1H and ^{13}C NMR) respectively spectrums were performed with a Mercury-400 spectrophotomer (400 and 100.6 MHz, respectively) using CDCl_3 , $(\text{CD}_3)_2\text{CO}$, DMSO-d_6 or other deuterated solvents with TMS as a reference. Chemical shifts are expressed as parts per million (ppm).

ESI Mass spectrums were carried out with Agilent LC/MSD-ToF mass spectrophotomer (Chemistry Faculty, University of Barcelona).

Melting points were determined with a Gallenkamp model MFB.595.010M device with an internal thermometer and were adjusted with an external thermometer.

Column chromatography was performed manually on silica gel Merck 60 (40-60 cm) eluting with mixtures of different solvents or through an automatic CombiFlash[®] R_f system equipped with UV-Vis (PN 68-5230-008) and silica gel RediSep R_f (4 and 12 grams).

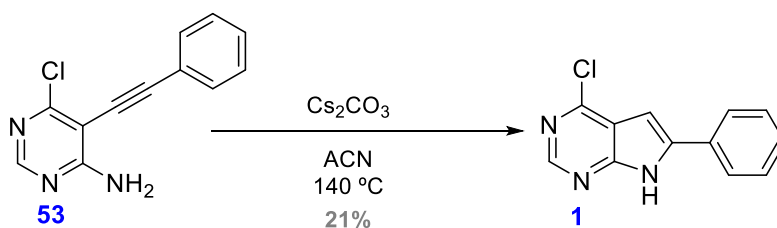
TLC plate 60 F₂₅₄ Merck was used as thin layer chromatography.

Microdistillations were performed in a Büchi GKR-50 Glass Tube Oven.

The microwaves reactions were carried out with CEM Discover LabMate and the temperature were adjusted by an IR sensor.

All reagents and organic used were recognized analytical grade or were purified before their use. Commercial products were obtained from Sigma-Aldrich-Merck.

Preparation of 4-chloro-6-phenyl-7H-pyrrolo[2,3-d]pyrimidine (**1**)



Procedure

A 50 mL round-bottomed flask specially equipped with a screw cap for coupling reactions, previously flame-dried under argon, was charged with the starting material **53** (0.295 g, 1.28 mmol, 1 eq) dissolved in 10 mL of ACN. Then, caesium carbonate (0.418 g, 1.28 mmol, 1 eq) was added to the reaction mixture. The resulting mixture was stirred and heated at $140 \pm 10\text{ }^\circ\text{C}$ for 24 hours.

Work-up

The reaction mixture was diluted with water (20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

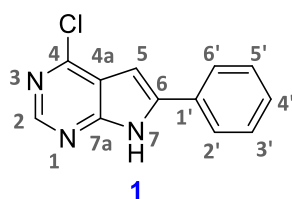
Purification

The crude reaction was purified by column chromatography using a CombiFlash[®] R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a polarity of hexane/ethyl acetate (30:70).

Yield: 21%

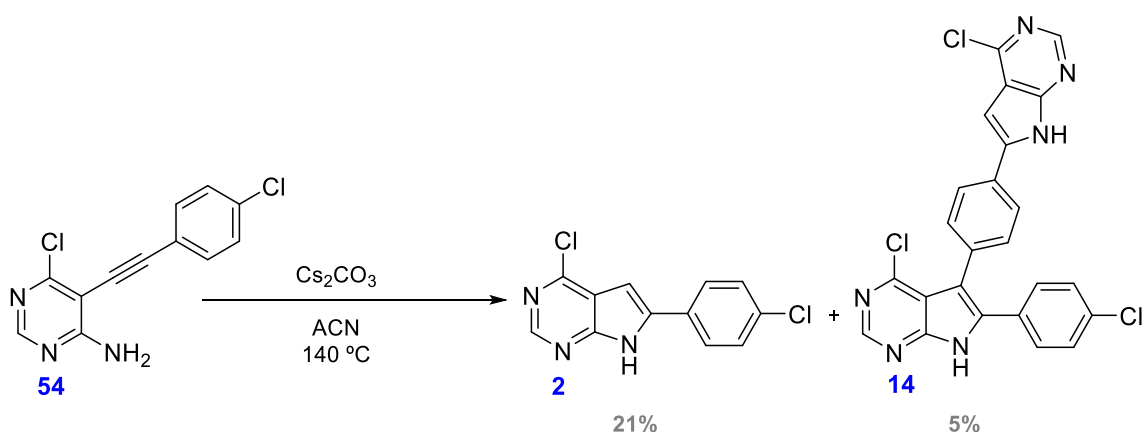
Analytical data

- R_f: 0.051 (Hexane/Ethyl Acetate (7:3))
- Aspect: brown semisolid



- ^1H NMR (DMSO- d_6 , 400 MHz) δ (ppm), 7.10 (s, 1H, H-5); 7.42 (t, $J = 7.3$ Hz, 1H, H-4'); 7.51 (t, $J = 7.6$ Hz, 2H, H'-3', H-5'); 8.01 (t, $J = 7.6$ Hz, 2H, H-2', H-6'); 8.15 (s, 1H, H-2); 12.90 (s, 1H, NH).
- ^{13}C NMR (CDCl $_3$, 100.6 MHz) δ (ppm), 95.5 (CH, C-5); 117.8 (C, C-4a); 125.9 (CH, C-2', C-6'); 129.0 (CH, C-3', C-5'); 129.1 (CH, C-4'); 130.1 (C, C-6); 140.3 (C, C-1'); 149.8 (C, C-4); 150.4 (CH, C-2); 153.0 (C, C-7a).

Preparation of 4-chloro-6-(4-chlorophenyl)-7H-pyrrolo[2,3-d]pyrimidine (**2**)



Procedure

A 50 mL round-bottomed flask specially equipped with a screw cap for coupling reactions, previously flame-dried under argon, was charged with the starting alkyne **54** (0.450 g, 1.70 mmol, 1 eq) dissolved in 10 mL of ACN. Then, caesium carbonate (0.555 g, 1.70 mmol, 1 eq) was added to the reaction mixture. The resulting mixture was stirred and heated at $140 \pm 10^\circ\text{C}$ for 24 hours.

Work-up

The reaction mixture was diluted with water (20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

Purification

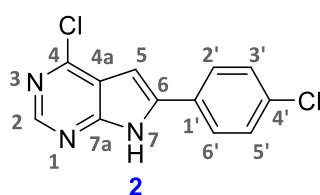
The crude reaction was purified by column chromatography using a CombiFlash[®] R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a polarity of hexane/ethyl acetate (20:80). A subproduct was obtained during purification. The

subproduct, 4-chloro-5-(4-chloro-7*H*-pyrrolo[2,3-*d*]pyrimidine-6-yl)phenyl)-6-(4-chloro phenyl)-7*H*-pyrrolo[2,3-*d*]pyrimidine (**14**) eluted with a polarity of hexane/ethyl acetate (50:50).

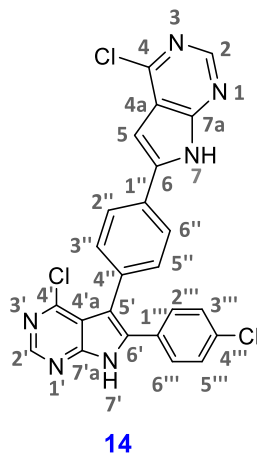
Yield: **21%** (Compound **2**), **5%** (Compound **14**)

Analytical data of compound **2** and **14**

- R_f: 0.051 (Hexane/Ethyl Acetate (7:3)), 0.1 (Hexane/Ethyl Acetate (7:3))
- Melting point: 168-171 °C (Ethyl Acetate) (x2)
- Aspect: yellow solid, brown solid

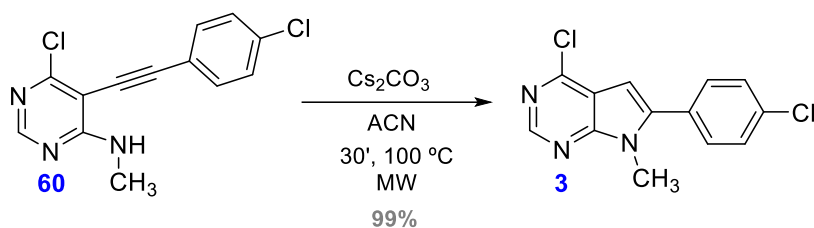


- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 6.28 (s, 1H, H-5); 7.13 (d, *J* = 8.6 Hz, 2H, H-3', H-5'); 7.61 (d, *J* = 8.6 Hz, 2H, H-2', H-6'); 8.47 (s, 1H, H-2).
- ¹³C NMR (CDCl₃, 100.6 MHz) δ(ppm), 102.0 (C, C-4a); 103.1 (CH, C-5); 128.5 (CH, C-2', C-6'); 128.8 (CH, C-3', C-5'); 129.0 (C, C-6); 133.5 (C, C-1'); 133.6 (C, C-4'); 154.3 (C, C-4); 155.9 (CH, C-2); 158.2 (C, C-7a).



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 6.25 (s, 1H, H-5); 7.05 (d, *J* = 8.7 Hz, 2H, H-3'', H-5''); 7.24 (d, *J* = 8.7 Hz, 2H, H-3''', H-5'''); 7.37 (d, *J* = 8.7 Hz, 2H, H-2'', H-6''); 7.53 (d, *J* = 8.7 Hz, 2H, H-2''', H-6'''); 8.31 (bs, 1H, NH); 8.40 (s, 1H, H-2); 8.58 (s, 1H, H-2').
- HRMS ESI (+) *m/z*: calculated mass for C₂₄H₁₄N₆Cl₃ 491.0267, found 491.0358.

Preparation of 4-chloro-6-(4-chlorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidine (**3**)



Procedure

A 10 mL round-bottomed flask specially equipped with a screw cap for microwaves under argon, was charged with the starting alkyne **60** (0.1 g, 0.36 mmol, 1 eq) dissolved in 4.5 mL of ACN. Then, caesium carbonate (0.362 g, 1 mmol, 3 eq) was added to the reaction. The reaction mixture was introduced to the microwaves oven (100 kPa) and heated at 100 ± 10 °C under stirring for 30 minutes.

Work up

After cooling, the solvent was removed under vacuum. Water (3 X 20 mL) was added to the reaction mixture and the crude of reaction was extracted with ethyl acetate (3 X 20 mL). The combined organic layers were washed with Na_2CO_3 and brine (15 mL). Then, the combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was evaporated under reduced pressure.

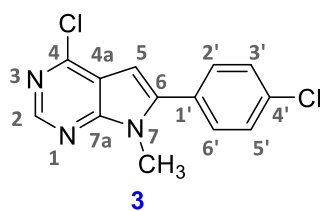
Purification

No purification by column chromatography was performed since the desired product was obtained. The product was directly used for the next step since the product was afford with enough purity as shown in the ^1H spectrum.

Yield: 99%

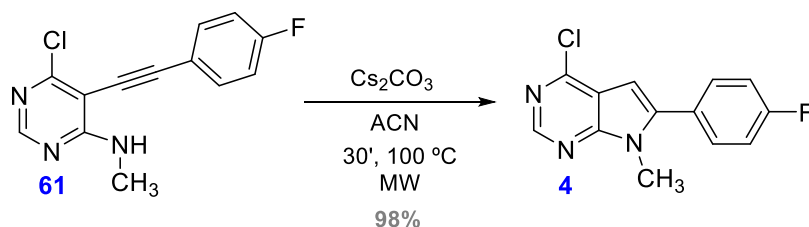
Analytical data

- R_f : 0.01 (Hexane/Ethyl acetate (7:3))
- Melting point: 168-170 °C (Ethyl Acetate)
- Aspect: yellow solid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 3.85 (s, 3H, -N- CH_3); 6.64 (s, 1H, H-5); 7.45-7.54 (m, 4H, H-2', H-3', H-5', H-6'); 8.67 (s, 1H, H-2).
- ^{13}C NMR (CDCl_3 , 100.6 MHz) δ (ppm), 30.5 (C, -NH- CH_3); 99.1 (CH, C-5); 117.6 (C, C-4a); 129.3 (CH, C-2', C-6'); 130.5 (CH, C-3', C-5'); 135.7 (C, C-1'); 142.0 (C, C-4'); 150.7 (CH, C-2); 151.6 (C, C-6); 152.7 (C, C-4); 153.1 (C, C-7a).
- HRMS ESI (+) m/z: calculated mass for $\text{C}_{13}\text{H}_{10}\text{N}_3\text{Cl}_2$ 278.0246, found 278.0251.

Preparation of 4-chloro-6-(4-fluorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidine (**4**)



Procedure

A 10 mL round-bottomed flask specially equipped with a screw cap for microwaves under argon, was charged with the starting material **61** (0.150 g, 0.57 mmol, 1 eq) dissolved in 4.5 mL of ACN. Then, caesium carbonate (0.390 g, 1 mmol, 2 eq) was added to the reaction. The reaction mixture was introduced to the microwaves oven (100 kPa) and heated at 100 ± 10 °C under stirring for 30 minutes.

Work up

After cooling, the solvent was removed under vacuum. Water (3 X 20 mL) was added to the reaction mixture and the crude of reaction was extracted with ethyl acetate (3 X 20 mL). The combined organic layers were washed with an aqueous solution of Na_2CO_3 and brine (15 mL). Then, the combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was evaporated under reduced pressure.

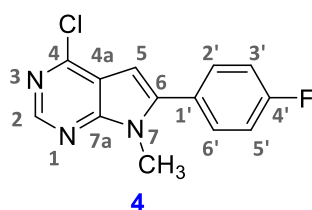
Purification

No purification by column chromatography was performed since the desired product was obtained. The product was directly used for the next step since the product was afforded with enough purity as shown in the ^1H spectrum.

Yield: **98%**

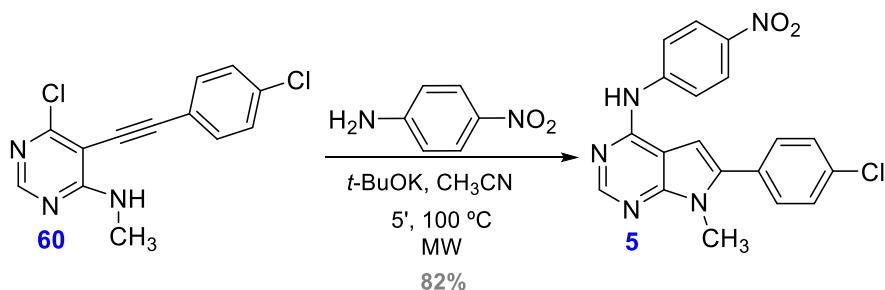
Analytical data

- R_f : 0.011 (Hexane/Ethyl acetate (7:3))
- Melting point: 120-124 °C (Ethyl Acetate)
- Aspect: orange solid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 3.84 (s, 3H, -N- CH_3); 6.61 (s, 1H, NH); 7.16-7.29 (m, 2H, H-3', H-5'); 7.46-7.58 (m, 2H, H-2', H-6'); 8.66 (s, 1H, H-2).
- ^{13}C NMR (CDCl_3 , 100.6 MHz) δ (ppm), 30.4 (C, -NH- CH_3); 98.8 (C, C-5); 116.2 (d, $J = 22$ Hz, C-3', C-5'); 117.6 (C, C-4a); 126.9 (d, $J = 4$ Hz, C-1'); 131.2 (d, $J = 9$ Hz, C-2', C-6'); 142.2 (C, C-7a); 150.6 (CH, C-2); 152.0 (d, $J = 89$ Hz, C-4'); 161.7 (C, C-4); 165.0 (C, C-6).
- HRMS ESI (+) m/z : calculated mass for $\text{C}_{13}\text{H}_{10}\text{N}_3\text{ClF}$ 262.0542, found 262.0546.

Preparation of 4-nitrophenylamino-6-(4-chlorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidine (**5**)



Procedure

A 10 mL round-bottomed flask specially equipped with a screw cap for microwaves under argon, was charged with the starting alkyne **60** (0.050 g, 0.18 mmol, 1 eq) dissolved in 5

mL of ACN. Then, 4-nitroaniline (0.049 g, 0.36 mmol, 2 eq) and potassium *tert*-butoxide (0.122 g, 1 mmol, 5 eq) were added to the reaction. The reaction mixture was introduced into the microwaves oven (100 kPa) and heated at 100 ± 10 °C under stirring for 5 minutes.

Work up

After cooling, the solvent was removed under vacuum. The crude reaction was extracted with diethyl ether/ethyl acetate 1/1 (3 X 20 mL). The organic layer was washed with water (2 X 20 mL) and brine. The combined organic phases were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was removed under reduced pressure.

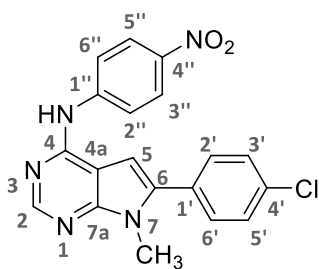
Purification

The crude reaction was purified by column chromatography using a CombiFlash[®] R_f provided with a UV-Vis detector. The eluent used was hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of dichloromethane/acetone (98:2).

Yield: **82%**

Analytical data

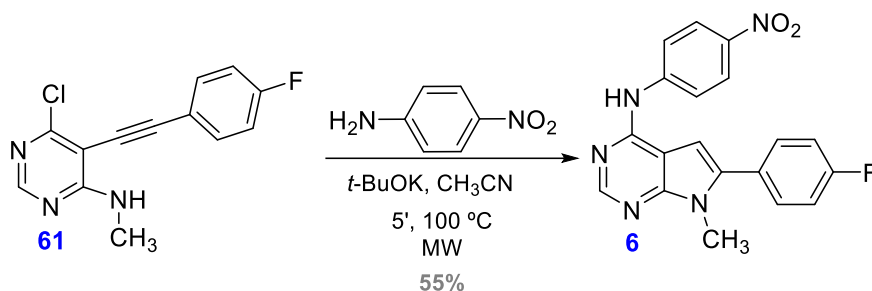
- R_f: 0.23 (Dichloromethane/Acetone (8:2))
- Melting point: 295-297 °C (Ethyl Acetate)
- Aspect: orange solid



5

- ¹H NMR (DMSO-d₆, 400 MHz) δ(ppm), 3.80 (s, 3H, CH₃-N-); 7.05 (s, 1H, H-5); 7.60-7.64 (m, 2H, H-3'', H-5''); 7.68-7.72 (m, 2H, H-2'', H-6''); 8.22-8.28 (m, 4H, H-2', H-3', H-5', H-6'); 8.52 (s, 1H, H-2); 10.10 (s, 1H, NH).

Preparation of 4-nitrophenyl-6-(4-fluorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidine (**6**)



Procedure

A 10 mL round-bottomed flask specially equipped with a screw cap for microwaves under argon, was charged with the starting alkyne **61** (0.050 g, 0.19 mmol, 1 eq) dissolved in 5 mL of ACN. Then, the 4-nitroaniline (0.052 g, 0.38 mmol, 2 eq) and potassium *tert*-butoxide (0.130 g, 1 mmol, 5 eq) were added to the reaction. The reaction mixture was introduced into the microwaves oven (100 kPa) and heated at $100 \pm 10\text{ }^\circ\text{C}$ under stirring for 5 minutes.

Work up

After cooling, the solvent was removed under vacuum. The crude reaction was extracted with diethyl ether/ethyl acetate 1/1 (3 X 20 mL). The organic layer was washed with water (2 X 20 mL) and brine. The combined organic phases were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was removed under reduced pressure.

Purification

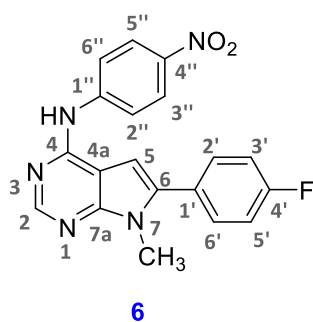
The crude reaction was purified by column chromatography using a CombiFlash[®] R_f provided with a UV-Vis detector. The eluent used was a mixture of dichloromethane and acetone with increasing polarity. The desired product eluted with a mixture of dichloromethane/acetone (98:2).

Yield: **55%**

Analytical data

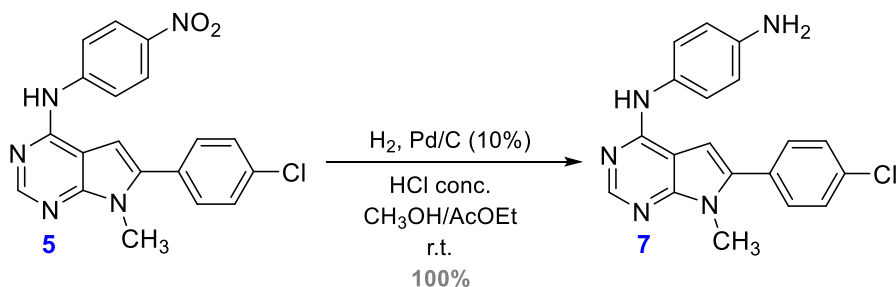
- R_f: 0.19 (Dichloromethane/Acetone (8:2))
- Melting point: 265-267 °C (Ethyl Acetate)

- Aspect: yellow solid



- ^1H NMR (DMSO- d_6 , 400 MHz) δ (ppm), 3.78 (s, 3H, $\text{CH}_3\text{-N}$); 7.01 (s, 1H, H-5); 7.37-7.47 (m, 2H, H-3'', H-5''); 7.69-7.74 (m, 2H, H-2'', H-6''); 8.25 (s, 4H, H-2', H-3', H-5', H-6'); 8.52 (s, 1H, H-2); 10.08 (s, 1H, NH).

Preparation of 4-(4-aminophenylamine)-6-(4-chlorophenyl)-pyrrolo[2,3-d]pyrimidine (**7**)



Procedure

A 50 mL round-bottomed flask for catalytic hydrogenation was charged with compound **5** (0.032 g, 0.084 mmol, 1 eq) dissolved in a mixture of 10 mL of methanol and 10 mL of ethyl acetate. The catalyst Pd-C 10% (10% p/p) was then added. Finally, five drops of hydrochloric acid (37%) was added to the solution in order to catalyze the reaction. The reaction mixture was stirred at room temperature for 72 hours, under hydrogen at atmospheric pressure.

Work-up

The theoretical volume required for the reaction was 4 mL. Given that the hydrogenation apparatus was not hermetical and may leak, the consumed volume was 18 mL, higher than expected.

The crude reaction was filtered by means of a pleated filter and washed with 20 mL of methanol and was then collected in a round-bottomed flask of 100 mL capacity.

Finally, ethyl acetate and methanol were removed under reduced pressure. The crude product was extracted with dichloromethane (3 X 20 mL) and washed with NaOH (2N). Then, the combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was evaporated under reduced pressure.

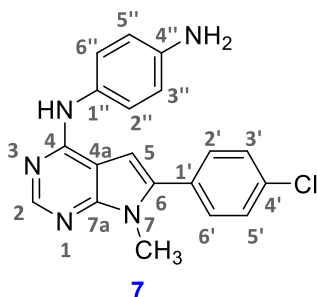
Purification

No purification by column chromatography was performed since the desired product was obtained with an amount of low quantity and also due to solubility problems. The product was obtained as shown in the ¹H spectrum.

Yield: **100%**

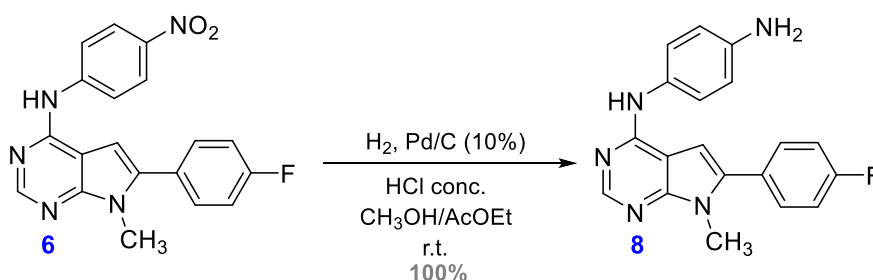
Analytical data

- R_f: 0.02 (Hexane/Ethyl Acetate (7:3))
- Melting point: 295-297 °C
- Aspect: white-yellowish solid



- ¹H NMR ((CD₃)₂CO, 400 MHz) δ(ppm), 3.77 (s, 3H, CH₃-N-); 4.50 (bs, 1H, NH); 6.47 (s, 1H, H-5); 6.68 (d, *J* = 8.7 Hz, 2H, H-3'', H-5''); 7.44 (d, *J* = 8.7 Hz, 2H, H-2'', H-6''); 7.53 (d, *J* = 8.6 Hz, 2H, H-3', H-5'); 7.59 (d, *J* = 8.6 Hz, 2H, H-2', H-6'); 8.28 (s, 1H, H-2).

Preparation of 4-(4-aminophenylamine)-6-(4-fluorophenyl)-pyrrolo[2,3-d]pyrimidine (**8**)



Procedure

A 50 mL round-bottomed flask for catalytic hydrogenations was charged with compound **6** (0.032 g, 0.088 mmol, 1 eq) dissolved in a mixture of 10 mL of methanol and 10 mL of ethyl acetate. The catalyst Pd-C 10% (10% p/p) was then added. Finally, one drop of hydrochloric acid (37%) was added to the solution in order to catalyze the reaction. The reaction mixture was stirred at room temperature for 24 hours, under hydrogen at atmospheric pressure.

Work-up

The theoretical volume required for the reaction was 4 mL. Given that the hydrogenation apparatus was not hermetical and may leak, the consumed volume was 15 mL, higher than expected.

The crude reaction was filtered by means of a pleated filter and washed with 20 mL of methanol and was then collected in a round-bottomed flask of 100 mL capacity. Finally, ethyl acetate and methanol were removed under reduced pressure. The crude product was extracted with dichloromethane (3 X 20 mL) and washed with NaOH (2N). Then, the combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was evaporated under reduced pressure.

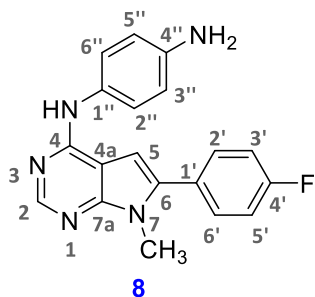
Purification

No purification by column chromatography was performed since the desired product was obtained with an amount of low quantity and also due to solubility problems. The product was obtained with enough purity as shown in the ¹H spectrum.

Yield: 100%

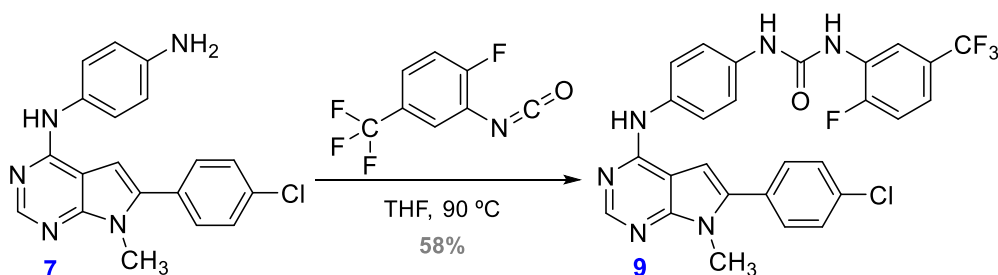
Analytical data

- R_f: 0.1 (Hexane/Ethyl Acetate (7:3))
- Melting point: 267-270 °C
- Aspect: white-yellowish solid



- ¹H NMR ((CD₃)₂CO, 400 MHz) δ(ppm), 3.78 (s, 3H, CH₃-N-); 6.45 (s, 1H, H-5); 6.67 (d, *J* = 8.5 Hz, 2H, H-3'', H-5''); 7.44 (d, *J* = 8.5 Hz, 2H, H-2'', H-6''); 7.49 (d, *J* = 7.5 Hz, 2H, H-3', H-5'); 7.57 (d, *J* = 7.5 Hz, 2H, H-2', H-6'); 8.25 (bs, 1H, NH); 8.27 (s, 1H, H-2).

Preparation of 1-(4-((6-(4-chlorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidin-4-yl)amino)phenyl)-3-(2-fluoro-5-(trifluoromethyl)phenyl)urea (9)



Procedure

A 50 mL round-bottomed flask equipped with a magnetic stirring bar and equipped with a reflux condenser, was charged with the amine **7** (0.030 g, 0.085 mmol, 1 eq) dissolved in 20 mL of THF. Then, the corresponding isocyanate (*d* = 1.418 g/mL, 0.0125 mL/g, 0.085 mmol, 1 eq) was added and the reaction mixture was stirred and heated at 90 ± 10 °C for 5 days.

Work-up

THF was removed under vacuum. A ¹H spectrum was carried out and the desired compound was obtained.

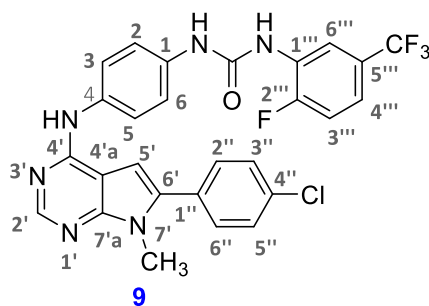
Purification

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (20:80).

Yield: **58%**

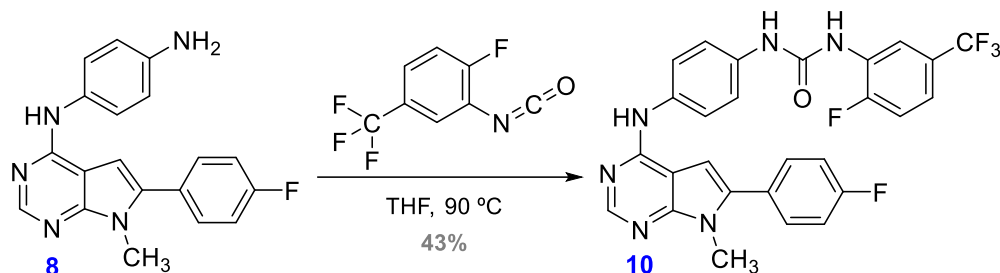
Analytical data

- R_f: 0.02 (Hexane/Ethyl Acetate (7:3))
- Melting point: >300 °C (Methanol)
- Aspect: white solid



- ¹H NMR ((CD₃)₂CO, 400 MHz) δ(ppm), 3.82 (s, 3H, CH₃-N-); 6.77 (s, 1H, H-5'); 7.36-7.38 (m, 2H, H-3''', H-4'''); 7.40 (s, 1H, H-6'''); 7.53 (d, *J* = 9 Hz, 2H, H-2, H-6); 7.56 (d, *J* = 8.7 Hz, 2H, H-3'', H-5''); 7.65 (d, *J* = 9 Hz, 2H, H-3, H-5); 7.87 (d, *J* = 8.7 Hz, 2H, H-2'', H-6''); 8.33 (bs, 1H, NH); 8.38 (s, 1H, H-2'); 8.62 (bs, 1H, NH); 8.83 (bs, 1H, NH).

Preparation of 1-(2-fluoro-5-(trifluoromethyl)phenyl)-3-(4-((6-(4-fluorophenyl)-7-methyl-7H-pyrrolo[2,3-d]pyrimidin-4-yl)amino)phenyl)urea (**10**)



Procedure

A 50 mL round-bottomed flask equipped with a magnetic stirring bar and equipped with a reflux condenser, was charged with the amine **8** (0.033 g, 0.098 mmol, 1 eq) dissolved in 20 mL of THF. Then, the corresponding isocyanate (d = 1.418 g/mL, 0.014 mL/g, 0.098 mmol, 1 eq) was added and the reaction mixture was stirred and heated at 90 ± 10 °C for 48 hours.

Work-up

THF was removed under vacuum. A ^1H spectrum of the final product was carried out confirming the proposed structure for **10**.

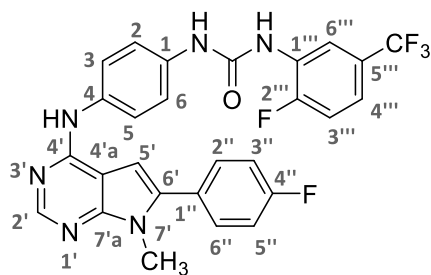
Purification

The crude reaction was purified by column chromatography using a CombiFlash[®] R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (20:80).

Yield: 43%

Analytical data

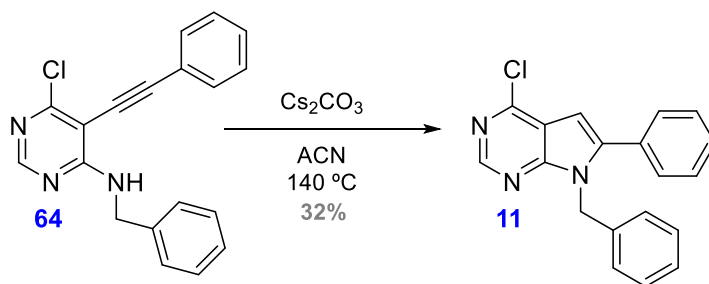
- R_f: 0.05 (Hexane/Ethyl Acetate (6:4))
- Melting point: >300 °C (Methanol)
- Aspect: white solid



10

- ^1H NMR ($(\text{CD}_3)_2\text{CO}$, 400 MHz) δ (ppm), 3.76 (s, 3H, $\text{CH}_3\text{-N}$); 6.60 (s, 1H, H-5'); 7.22-7.24 (m, 1H, H-3'''); 7.27 (s, 1H, H-6'''); 7.33 (d, $J = 7.4$ Hz, 1H, H-4'''); 7.38 (d, $J = 7.6$ Hz, 2H, H-2, H-6); 7.40 (d, $J = 8.9$ Hz, 2H, H-3'', H-5''); 7.49 (d, $J = 7.6$ Hz, 2H, H-3, H-5); 7.73 (d, $J = 8.9$ Hz, 2H, H-2'', H-6''); 8.15 (bs, 1H, NH); 8.25 (s, 1H, H-2'); 8.45 (bs, 1H, NH); 8.69 (bs, 1H, NH).

Preparation of 7-benzyl-4-chloro-6-phenyl-7H-pyrrolo[2,3-d]pyrimidine (**11**)



Procedure

A 50 mL round-bottomed flask specially equipped with a screw cap for coupling reactions, previously flame-dried under argon, was charged with 4-amino-*N*-benzyl-6-chloro-5-(phenylethynyl)pyrimidine (**64**) (0.252 g, 1.03 mmol, 1 eq) dissolved in 10 mL of ACN. Then, caesium carbonate (0.235 g, 0.72 mmol, 0.7 eq) was added to the reaction mixture. The resulting mixture was stirred and heated at 140 ± 10 °C for 24 hours.

Work-up

The reaction mixture was diluted with water (20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

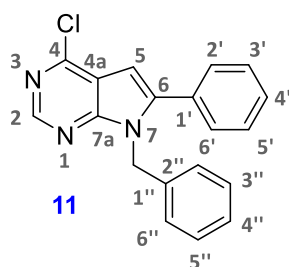
Purification

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (90:10).

Yield: **32%**

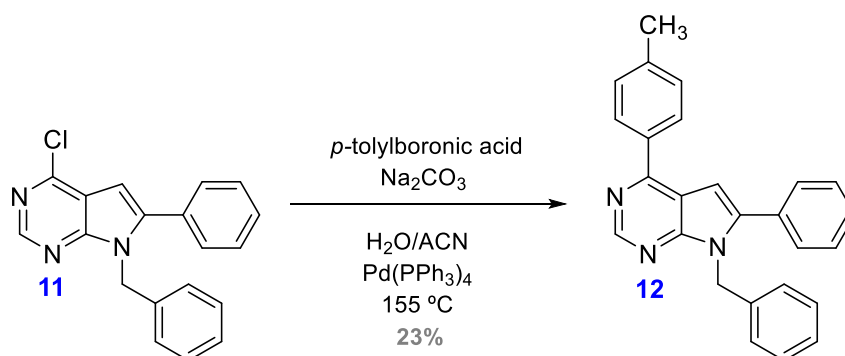
Analytical data

- R_f: 0.69 (Hexane/Ethyl Acetate (7:3))
- Aspect: yellow semisolid



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 5.52 (s, 2H, CH₂-N-); 6.66 (s, 1H, H-5); 7.19 (t, *J* = 6.5 Hz, 1H, H-4''); 7.21 (d, *J* = 6 Hz, 1H, H-4'); 7.32-7.34 (m, 2H, H-2'', H-6''); 7.35-7.37 (m, 2H, H-3'', H-5''); 7.38-7.39 (m, 2H, H-3', H-5'); 7.42 (d, *J* = 7 Hz, 2H, H-2', H-6'); 8.67 (s, 1H, H-2).

Preparation of 7-benzyl-6-phenyl-4-(*p*-tolyl)-7H-pyrrolo[2,3-*d*]pyrimidine (**12**)



Procedure

A 50 mL round-bottomed flask specially equipped with a screw cap for coupling reactions, previously flame-dried under argon, was charged with 7-benzyl-4-chloro-6-phenyl-7H-pyrrolo[2,3-*d*]pyrimidine (**11**) (0.053 g, 0.17 mmol, 1 eq) dissolved in 10 mL of ACN and

1 mL of water. Then, carbonate sodium (0.035 g, 0.33 mmol, 2 eq) and *p*-tolylboronic acid (0.024 g, 0.18 mmol, 1.1 eq) were added to the reaction mixture. Finally, phenyl phosphine in catalytic amount was added. The resulting mixture was stirred and heated at 155 ± 10 °C for 72 hours.

Work-up

The reaction mixture was diluted with water (20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

Purification

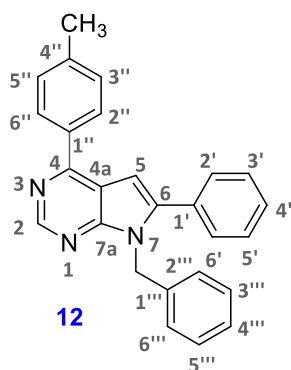
The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (70:30). An excess of *p*-tolylboronic acid was observed and another column chromatography was performed.

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (80:20).

Yield: 23%

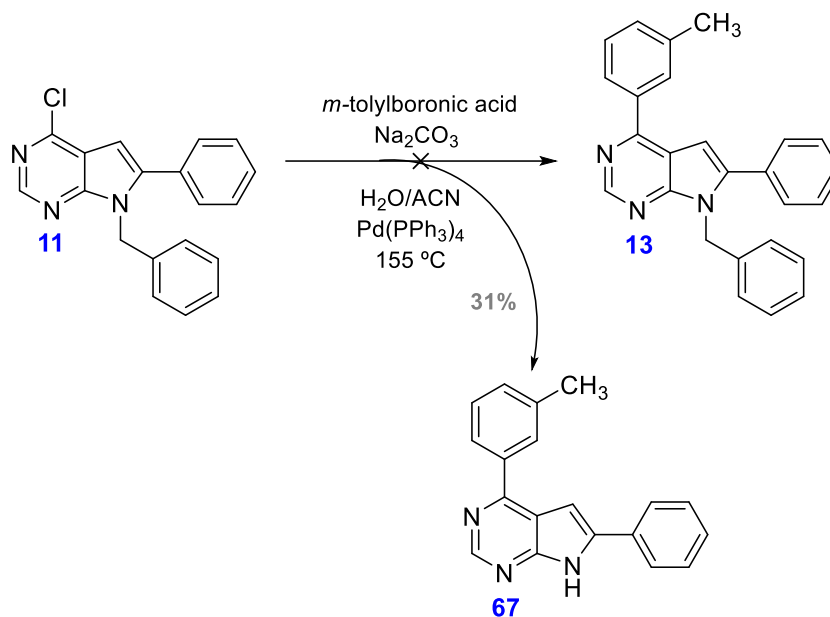
Analytical data

- R_f: 0.65 (Hexane/Ethyl Acetate (7:3))
- Aspect: orange semisolid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 1.97 (s, 3H, CH_3 -); 5.47 (s, 2H, $\text{CH}_2\text{-N}$ -); 6.28 (s, 1H, H-5); 7.23 (t, $J = 6.3$ Hz, 1H, H-4'''); 7.24 (t, $J = 6$ Hz, 1H, H-4'); 7.29 (d, $J = 6.3$ Hz, 4H, H-3'', H-5'', H-2''', H-6'''); 7.34 (d, $J = 6.9$ Hz, 2H, H-2'', H-6''); 7.36 (d, $J = 6.8$ Hz, 2H, H-3''', H-5'''); 7.41 (dt, $J_1 = 2, J_2 = 7.4$ Hz, 2H, H-3', H-5'); 7.72 (dd, $J_1 = 2, J_2 = 7.4$ Hz, H-2', H-6'); 8.30 (s, 1H, H-2).

Preparation of 7-benzyl-6-phenyl-4-(*m*-tolyl)-7H-pyrrolo[2,3-*d*]pyrimidine (**13**)



Procedure

A 50 mL round-bottomed flask specially equipped with a screw cap for coupling reactions, previously flame-dried under argon, was charged with 7-benzyl-4-chloro-6-phenyl-7H-pyrrolo[2,3-*d*]pyrimidine (**11**) (0.053 g, 0.17 mmol, 1 eq) dissolved in 10 mL of ACN and 1 mL of water. Then, carbonate sodium (0.035 g, 0.33 mmol, 2 eq) and *m*-tolylboronic acid (0.024 g, 0.18 mmol, 1.1 eq) were added to the reaction mixture. Finally, phenyl phosphine in catalytic amount was added. The resulting mixture was stirred and heated at $155 \pm 10^\circ\text{C}$ for 48 hours.

Work-up

The reaction mixture was diluted with water (20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

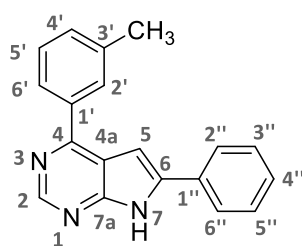
Purification

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The final product was surprisingly desbenzylated and the compound expected was not obtained. The subproduct achieved eluted with a mixture of hexane/ethyl acetate (80:20).

Yield: **31%**

Analytical data

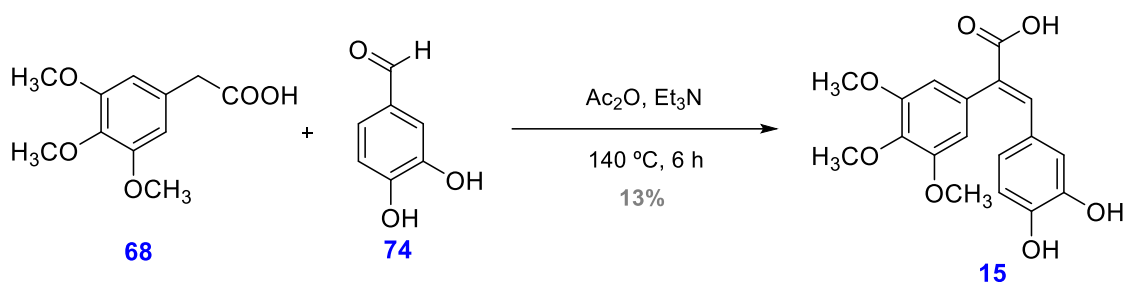
- R_f: 0.68 (Hexane/Ethyl Acetate (7:3))
- Aspect: brown semisolid



87

- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 2.25 (s, 3H, CH₃-Ar-); 6.28 (s, 1H, H-5); 7.22 (d, *J* = 4.1 Hz, 1H, H-4'); 7.23 (t, *J* = 6.3 Hz, 1H, H-4''); 7.26-7.27 (m, 1H, H-4''); 7.28-7.29 (m, 1H, H-5'); 7.29 (d, *J* = 7 Hz, 2H, H-2'', H-6''); 7.29-7.31 (m, 2H, H-3'', H-5''); 7.32-7.33 (m, 1H, H-6'); 7.48 (d, *J* = 2 Hz, 1H, H-2'); 8.30 (s, 1H, H-2).

Preparation of (*E*)-3-(3,4-dihydroxyphenyl)-2-(3,4,5-trimethoxyphenyl)acrylic acid (**15**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon and equipped with a reflux condenser was charged with the starting aldehyde **74** (1 g, 4.42 mmol, 1 eq) dissolved in anhydrous acetic acid (5 mL). Then, 2-(3,4,5-trimethoxyphenyl)acetic acid (**68**) (0.610 g, 4.42 mmol, 1 eq) and triethylamine (2.5 mL, $d = 0.726 \text{ g/mL}$, 17.94 mmol, 4 eq) were added to the mixture. The reaction mixture was heated at reflux of anhydrous acetic acid ($140 \pm 10 \text{ }^\circ\text{C}$) under stirring for 6 hours.

Work-up

The crude product was microdistilled (7.5 mm Hg) under vacuum at $150 \pm 10 \text{ }^\circ\text{C}$. Then, the crude was dissolved with dichloromethane and extracted with an aqueous solution of HCl (2N) (3 X 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was removed under reduced pressure.

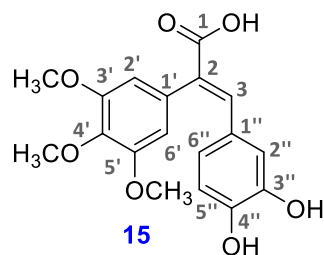
Purification

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was hexane and ethyl acetate with increasing polarity. The desired product eluted with a polarity of hexane/ethyl acetate (50:50). The product obtained was a mixture of rotamers or conformers.

Yield: 13%

Analytical data

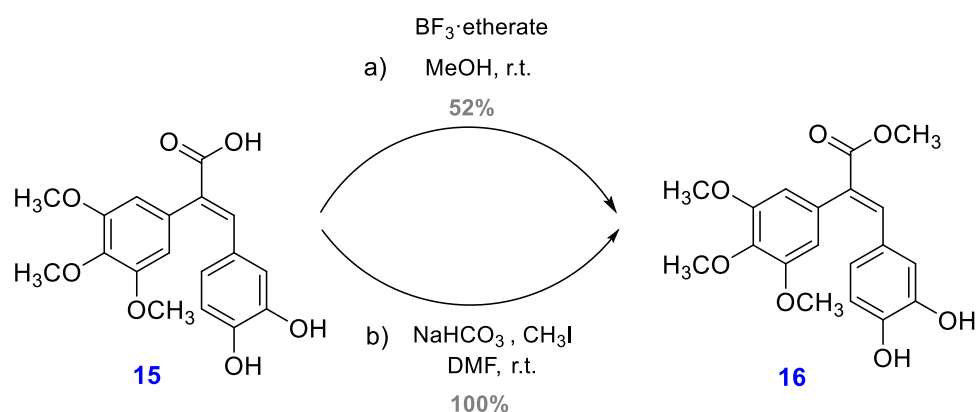
- R_f: 0.13 (Hexane/Ethyl Acetate (4:6))
- Aspect: white semisolid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 3.69 (s, 3H, $\text{CH}_3\text{-O-}$); 3.79 (s, 6H, $\text{CH}_3\text{-O-}$ (x2)); 6.63 (s, 2H, H-2', H-6'); 6.81 (d, $J = 2$ Hz, 1H, H-2''); 6.84 (s, 1H, OH); 6.94-6.96 (m, 2H, H-5'', OH); 7.09 (dd, $J_1 = 2$, $J_2 = 8$ Hz, 1H, H-6''); 7.77 (s, 1H, H-3).
- ^{13}C NMR (CDCl_3 , 100.6 MHz) δ (ppm), 56.1 ($\text{CH}_3\text{-O-}$ (x2)); 60.9 ($\text{CH}_3\text{-O-}$); 104.2 (CH, C-2, C-6') (A); 106.4 (CH, C-2', C-6') (B); 123.3 (CH, C-2'') (B); 123.5 (CH, C-2'') (A); 125.6 (CH, C-6'') (B); 126.6 (CH, C-6'') (A); 129.2 (CH, C-5'') (A); 129.8 (CH, C-5'') (B); 130.3 (C, C-2); 132.2 (C, C-1') (A); 132.7 (C, C-1') (B); 134.1 (C, C-1'') (A); 135.4 (C, C-1'') (B); 137.9 (C, C-3'') (B); 138.6 (C, C-3'') (A); 140.4 (CH, C-3); 141.8 (C, C-4') (A); 141.9 (C, C-4') (B); 142.9 (C, C-4''); 153.2 (C, C-3', C-5') (A); 153.7 (C, C-3', C-5') (B); 172.3 (C, C=O) (A); 172.4 (C, C=O) (B).

A: majority isomer / B: minority isomer

Preparation of (*E*)-methyl 3-(3,4-dihydroxyphenyl)-2-(3,4,5-trimethoxyphenyl)acrylate (16)



Procedure A

A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon, was charged with compound **15** (0.973 g, 2.81 mmol, 1 eq) dissolved in methanol (10 mL). Then, 2 mL of boron trifluoride etherate ($d = 1.15$ g/mL, 16.21 mmol,

6 eq) was slowly added. The resulting mixture was maintained under constant stirring at room temperature for 10 days.

Work-up

Methanol was removed under reduced pressure. Then, 15 mL of water was added to the reaction mixture and the crude product was extracted with ethyl acetate (3 X 20 mL). The organic phases were washed with an aqueous sodium bicarbonate solution (3 X 20 mL). Finally, the combined organics layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was removed under reduced pressure with toluene.

Purification

The crude product was used in the next step without further purification since the product was obtained with enough purity as shown by TLC and ¹H spectrum. The obtained product was a mixture of rotamers or conformers.

Yield: 52%

Procedure B

A 100 mL round-bottomed flask equipped with a reflux condenser and a magnetic stirring bar was charged with the initial acid **15** (0.186 g, 0.53 mmol, 1 eq) dissolved in 15 mL of DMF. Then, NaHCO₃ (0.047 g, 0.55 mmol, 1 eq) and 0.5 mL of methyl iodide (d = 2.28 g/mL, 24.09 mmol) were added to the reaction mixture. The resulting mixture was stirred at room temperature for 48 hours during which 1 mL more of methyl iodide was added in order to obtain the desired product.

Work-up

The reaction mixture was diluted with water (15 mL) and extracted with diethyl ether (3 X 15 mL). The combined organic layers were washed with water (3 X 15 mL), dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

Purification

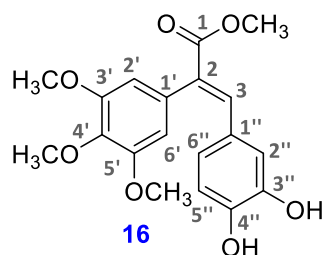
The crude product was used in the next step without further purification since the product was obtained with enough purity as shown by TLC and ¹H spectrum.

In this case a mixture of rotamers or conformers was obtained.

Yield: **100%**

Analytical data

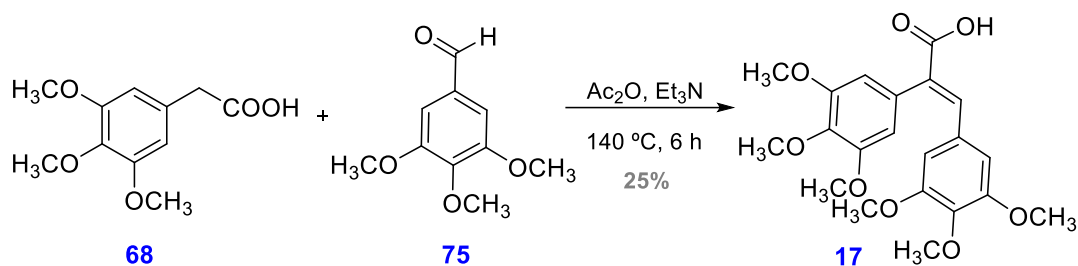
- R_f: 0.47 (Hexane/Ethyl Acetate (1:1))
- Aspect: yellow semisolid



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 3.86 (s, 3H, CH₃-O-); 3.87 (s, 3H, CH₃-O-C=O); 3.88 (s, 6H, CH₃-O- (x2)); 6.45 (s, 2H, H-2', H-6') (B); 6.66 (s, 2H, H-2', H-6') (A); 6.87 (d, *J* = 8 Hz, 1H, H-5''); 6.90 (bs, 1H, OH); 7.02 (d, *J* = 2 Hz, 1H, H-2''); 7.03 (bs, 1H, OH); 7.22 (dd, *J*₁ = 2, *J*₂ = 8 Hz, 1H, H-6''); 7.82 (s, 1H, H-3).
- ¹³C NMR (CDCl₃, 100.6 MHz) δ(ppm), 52.5 (CH₃-O-); 56.1 ((CH₃-O-) (B)); 56.2 ((CH₃-O-) (A)); 60.9 (CH₃-O-C=O); 103.9 (CH, C-2', C-6') (A); 106.4 (CH, C-2', C-6') (B); 123.0 (CH, C-2'') (A); 123.4 (CH, C-2'') (B); 125.5 (CH, C-6'') (B); 126.3 (CH, C-6'') (A); 129.2 (CH, C-5') (B); 129.4 (CH, C-5') (A); 129.9 (C, C-1') (B); 132.2 (C, C-1') (A); 132.4 (C, C-1'') (B); 132.9 (C, C-1'') (A); 134.3 (C, C-2) (A); 135.8 (C, C-2) (B); 137.9 (C, C-4') (B); 138.7 (C, C-4') (A); 139.9 (CH, C-3); 141.8 (C, C-3'') (B); 141.9 (C, C-3'') (A); 142.0 (C, C-4'') (A); 142.8 (C, C-4'') (B); 153.3 (C, C-3', C-5') (A); 153.7 (C, C-3', C-5') (B); 167.9 (C, C=O) (B); 168.0 (C, C=O) (A).

A: majority isomer / B: minority isomer

Preparation of (*E*)-2,3-bis(3,4,5-trimethoxyphenyl)acrylic acid (**17**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon and equipped with a reflux condenser was charged with the starting aldehyde **75** (0.867 g, 4.42 mmol, 1 eq) dissolved in anhydrous acetic acid (5 mL). Then, 2-(3,4,5-trimethoxyphenyl)acetic acid (**68**) (1 g, 4.42 mmol, 1 eq) and triethylamine (2.5 mL, $d = 0.726 \text{ g/mL}$, 17.94 mmol, 4 eq) were added to the mixture. The reaction mixture was heated at reflux of anhydrous acetic acid ($140 \pm 10 \text{ }^\circ\text{C}$) under stirring for 6 hours.

Work-up

The crude product was microdistilled (7.5 mm Hg) under vacuum at $150 \pm 10 \text{ }^\circ\text{C}$. Then, the crude product was dissolved with dichloromethane and extracted with an aqueous solution of HCl (2N) (3 X 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was removed under reduced pressure.

Purification

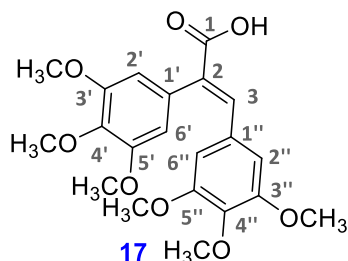
The amount obtained of crude product was 1.8 g, only 1 g of crude product was purified.

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (50:50). The product obtained was a yellowish solid. In one hand, 0.807 g were not purified. However, this amount of product was used to carry out a different assay in similar conditions since the crude product was obtained with enough purity.

Yield: 50%

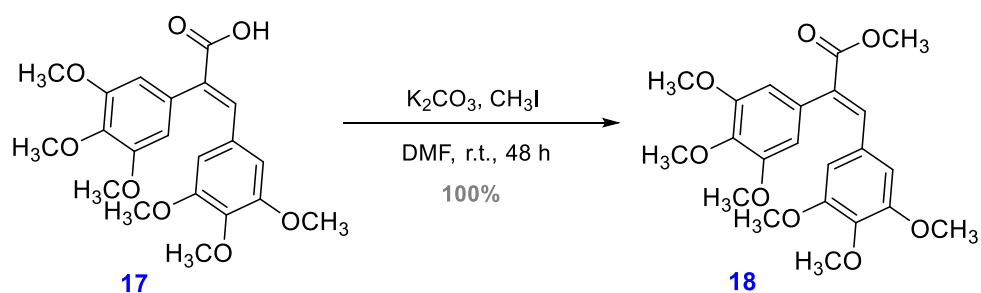
Analytical data

- R_f: 0.29 (Hexane/Ethyl Acetate (2:8))
- Melting point: 80-83 °C (Dichloromethane)
- Aspect: yellowish solid



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 3.57 (s, 6H, CH₃-O- (x2)); 3.77 (s, 6H, CH₃-O- (x2)); 3.78 (s, 3H, CH₃-O-C-4'); 3.82 (s, 3H, CH₃-O-C-4''); 6.23 (s, 2H, H-2', H-6'); 6.48 (s, 2H, H-2'', H-6''); 7.80 (s, 1H, H-3).
- ¹³C NMR (CDCl₃, 100.6 MHz) δ(ppm), 55.7 (CH₃-O- (x2)); 56.2 (CH₃-O- (x2)); 60.7 (CH₃-O-); 60.8 (CH₃-O-); 104.1 (CH, C-2'', C-6''); 108.4 (CH, C-2', C-6'); 129.1 (C, C-1''); 130.3 (C, C-1'); 131.1 (C, C-2); 137.7 (C, C-4''); 139.5 (C, C-4'); 142.3 (CH, C-3); 152.6 (CH, C-3'', C-5''); 153.8 (C, C-3', C-5'); 172.7 (C, C=O).

Preparation of (*E*)-methyl 2,3-bis(3,4,5-trimethoxyphenyl)acrylate (**18**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar was charged with (*E*)-2,3-bis(3,4,5-trimethoxyphenyl)acrylic acid (**17**) (0.414 g, 1.02 mmol, 1 eq) dissolved in 10 mL of DMF. Then, potassium carbonate (0.212 g, 1.53 mmol, 1.5 eq) and 1 mL of methyl iodide (d = 2.28 g/mL, 32.13 mmol) were added to the reaction mixture. All compounds were handled under argon conditions.

The resulting mixture was stirred at room temperature for 48 hours, during which 1 mL more of methyl iodide was added in order to obtain the desired compound **18**.

Work-up

The reaction mixture was diluted with water (15 mL) and extracted with diethyl ether (3 X 15 mL). The combined organic layers were washed with water (3 X 20 mL), dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

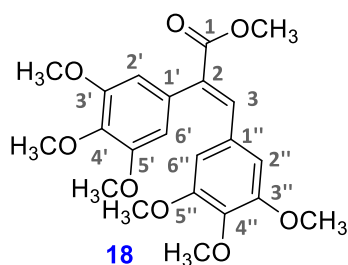
Purification

The crude product was used in the next step without further purification since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: **100%**

Analytical data

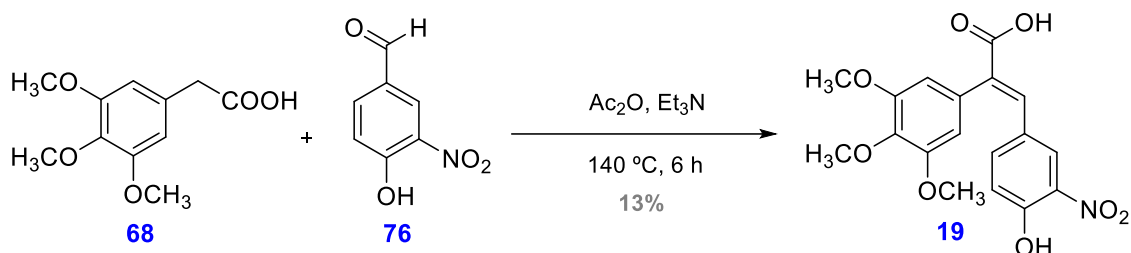
- R_f : 0.73 (Hexane/Ethyl Acetate (2:8))
- Aspect: yellow semisolid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 3.59 (s, 6H, $\text{CH}_3\text{-O-}$ (x2)); 3.80 (s, 6H, $\text{CH}_3\text{-O-}$ (x2)); 3.81 (s, 3H, $\text{CH}_3\text{-O-C-4}'$); 3.82 (s, 3H, $\text{CH}_3\text{-O-C-4}''$); 3.85 (s, 3H, $\text{CH}_3\text{-O-C=O}$); 6.37 (s, 2H, H-2', H-6'); 6.48 (s, 2H, H-2'', H-6''); 7.72 (s, 1H, H-3).
- ^{13}C NMR (CDCl_3 , 100.6 MHz) δ (ppm), 52.4 ($\text{CH}_3\text{-O-C=O}$); 55.7 ($\text{CH}_3\text{-O-}$ (x2)); 56.2 ($\text{CH}_3\text{-O-}$ (x2)); 60.7 ($\text{CH}_3\text{-O-}$); 60.8 ($\text{CH}_3\text{-O-}$); 106.6 (CH, C-2'', C-6''); 108.0 (CH, C-2', C-6'); 129.5 (C, C-1''); 131.1 (C, C-1'); 131.6 (C, C-2); 137.5 (C, C-4''); 139.1 (C, C-4'); 140.3 (CH, C-3); 152.6 (CH, C-3'', C-5''); 153.7 (C, C-3', C-5'); 168.1 (C, C=O).

Preparation of (*E*)-3-(4-hydroxy-3-nitrophenyl)-2-(3,4,5-trimethoxyphenyl)acrylic acid (**19**)

Procedure



A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon and equipped with a reflux condenser was charged with the starting aldehyde **76** (0.738 g, 4.41 mmol, 1 eq) dissolved in anhydrous acetic acid (5 mL). Then, 2-(3,4,5-trimethoxyphenyl)acetic acid (**68**) (1 g, 4.41 mmol, 1 eq) and triethylamine (2.5 mL, $d = 0.726 \text{ g/mL}$, 17.94 mmol, 4 eq) were added to the mixture. The reaction mixture was heated at reflux of anhydrous acetic acid ($140 \pm 10 \text{ }^\circ\text{C}$) under stirring for 6 hours.

Work-up

The crude product was microdistilled (7.5 mm Hg) under vacuum at $150 \pm 10 \text{ }^\circ\text{C}$ (was not 100% effective due to vacuum problems). Then, the crude product was dissolved with dichloromethane and extracted with an aqueous solution of HCl (2N) (3 X 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was removed under reduced pressure.

Since there was Ac₂O in the reaction mixture a saturated aqueous NaHCO₃ solution (3 X 15 mL) was added. Then, the crude reaction was extracted with diethyl ether (3 X 15 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was evaporated under reduced pressure. The solid obtained was the initial aldehyde. On the other hand, the aqueous phase was extracted with dichloromethane and diluted with an aqueous HCl solution (2N) (3 X 20 mL). Finally, the combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was evaporated under reduced pressure. The desired product was then obtained.

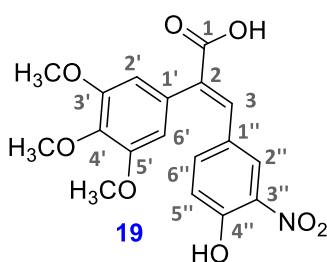
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: **13%**

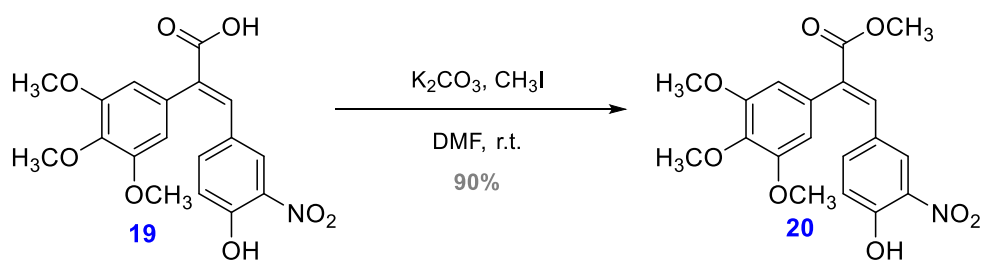
Analytical data

- R_f : 0.77 (Hexane/Ethyl Acetate (1:1))
- Melting point: 211-214 °C (Dichloromethane)
- Aspect: yellow solid



- ^1H NMR ((CDCl_3 , 400 MHz) δ (ppm), 3.83 (s, 3H, $\text{CH}_3\text{-O-}$); 3.85 (s, 6H, $\text{CH}_3\text{-O-}$ (x2)); 6.85 (s, 2H, H-2', H-6'); 7.32 (d, $J = 8.8$ Hz, 1H, H-5''); 7.81 (s, 1H, H-3); 8.13 (dd, $J_1 = 2, J_2 = 8.8$ Hz, 1H, H-6''); 8.62 (d, $J = 2$ Hz, 1H, H-2''); 9.92 (s, 1H, COOH).

Preparation of (E)-3-methyl (4-hydroxy-3-nitrophenyl)-2-(3,4,5-trimethoxyphenyl) acrylate (20)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon was charged with the initial acid **19** (0.314 g, 0.84 mmol, 1 eq), potassium carbonate (0.173 g, 1.25 mmol, 1.5 eq) and 1 mL of methyl iodide (d = 2.28 g/mL, 32.12 mmol).

The reaction mixture was stirred at room temperature for 4 days and the system was closed with a reflux condenser. During these days 1 mL more of methyl iodide was added to the resulting mixture in order to obtain the final compound.

Work-up

The reaction mixture was diluted with water (15 mL) and extracted with diethyl ether (3 X 15 mL). The combined organic layers were washed with water (3 X 20 mL), dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure. The final solid was the initial acid and the desired compound was not obtained as shown in the ^1H spectrum.

The reaction was carried out again following the same procedure and work-up. The final solid was the desired compound **20** with enough purity to carry out the next step.

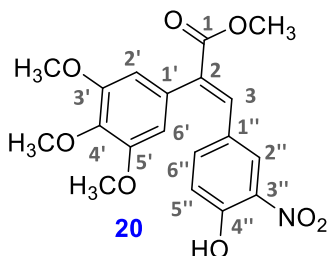
Purification

A ^1H spectrum was carried out in order to confirm the achievement of the final compound.

Yield: 90%

Analytical data

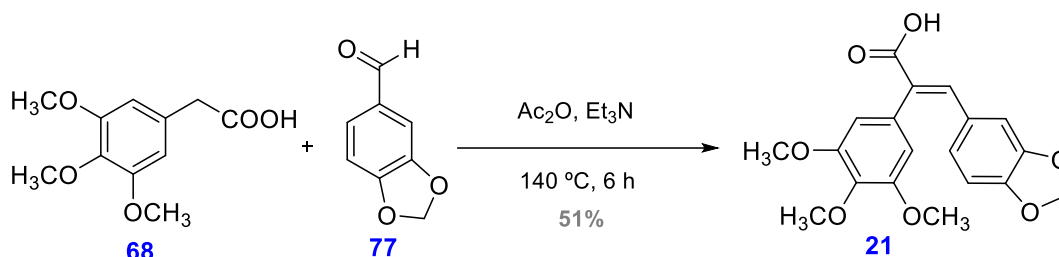
- R_f : 0.53 (Hexane/Ethyl Acetate (1:1))
- Melting point: 104-106 °C (Ethyl acetate)
- Aspect: yellow solid



- ^1H NMR ((CDCl_3 , 400 MHz) δ (ppm), 3.71 (s, 3H, $\text{CH}_3\text{-O-}$); 3.79 (s, 3H, $\text{CH}_3\text{-O-C=O}$); 3.85 (s, 6H, $\text{CH}_3\text{-O-}$ (x2)); 6.50 (s, 2H, H-2', H-6'); 6.88 (d, $J = 8.8$ Hz, 1H, H-5''); 7.70 (s, 1H, H-3); 8.09 (dd, $J_1 = 2$, $J_2 = 8.8$ Hz, 1H, H-6''); 8.35 (d, $J = 2$ Hz, 1H, H-2''); 10.0 (s, 1H, OH).

Preparation of (*E*)-3-(benzo[*d*][1,3]dioxol-5-yl)-2-(3,4,5-trimethoxyphenyl)acrylic acid (**21**)

Procedure



A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon and equipped with a reflux condenser was charged with the piperonal (**77**) (0.663 g, 4.41 mmol, 1 eq) dissolved in anhydrous acetic acid (5 mL). Then, 3,4,5-trimethoxyphenylacetic acid (**68**) (1 g, 4.41 mmol, 1 eq) and triethylamine (2.5 mL, $d = 0.726$ g/mL, 17.94 mmol, 4 eq) were added to the mixture. The reaction mixture was heated at reflux of anhydrous acetic acid (140 ± 10 °C) under stirring for 6 hours.

Work-up

The crude product was microdistilled (7.5 mm Hg) under vacuum at 150 ± 10 °C. Then, the crude product was dissolved with dichloromethane and extracted with an aqueous solution of HCl (2N) (3 X 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was removed under reduced pressure.

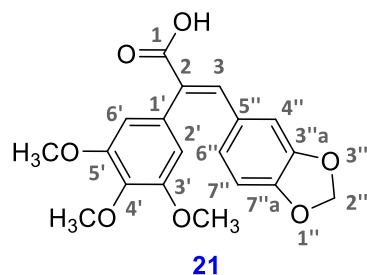
Purification

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (60:40).

Yield: **51%**

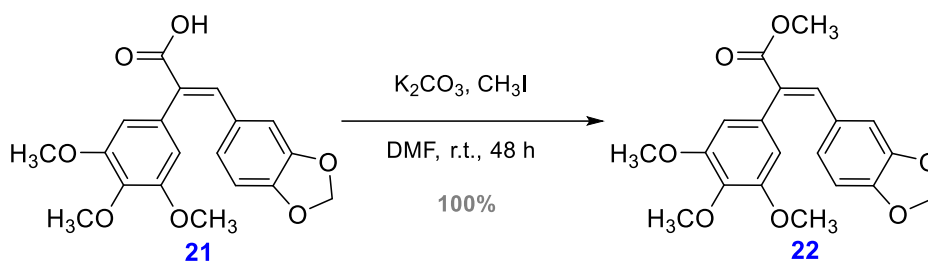
Analytical data

- R_f: 0.33 (Hexane/Ethyl Acetate (1:1))
- Aspect: yellow semisolid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 3.79 (s, 6H, $\text{CH}_3\text{-O-}$ (x2)); 3.90 (s, 3H, $\text{CH}_3\text{-O-}$); 5.92 (s, 2H, $-\text{O-CH}_2\text{-O-}$); 6.46 (s, 2H, H-2', H-6'); 6.68 (s, 1H, H-4''); 6.69 (d, $J = 8$ Hz, 1H, H-7''); 6.81 (d, $J = 8$ Hz, 1H, H-6''); 7.85 (s, 1H, H-3).

Preparation of (*E*)-methyl 3-(benzo[*d*][1,3]dioxol-5-yl)-2-(3,4,5-trimethoxyphenyl) acrylate (22)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon was charged with (*E*)-3-(benzo[*d*][1,3]dioxol-5-yl)-2-(3,4,5-trimethoxyphenyl)acrylate acid (**21**) (0.781 g, 2.18 mmol, 1 eq), potassium carbonate (0.451 g, 3.27 mmol, 1.5 eq) and 1 mL of methyl iodide ($d = 2.28$ g/mL, 16.03 mmol, 7 eq). The reaction mixture was stirred at room temperature for 48 hours.

Work-up

The reaction mixture was diluted with water (15 mL) and extracted with diethyl ether (3 X 15 mL). The combined organic layers were washed with water (3 X 20 mL), dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

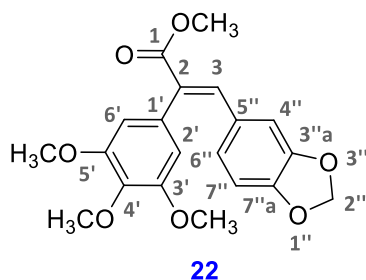
Purification

The crude product was directly used for the next synthetic step since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: **100%**

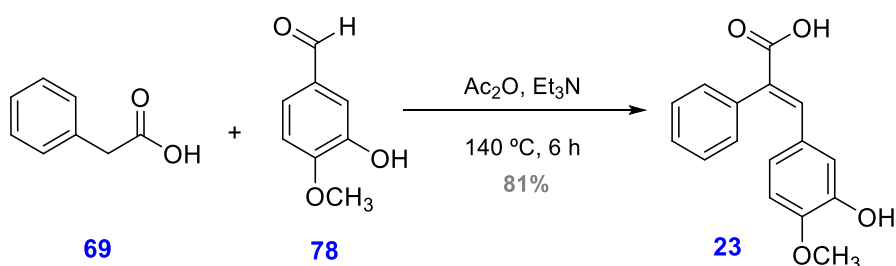
Analytical data

- R_f: 0.66 (Hexane/Ethyl Acetate (4:6))
- Aspect: brown semisolid



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 3.77 (s, 6H, CH₃-O- (x2)); 3.87 (s, 3H, CH₃-O-); 3.89 (s, 3H, CH₃-O-C=O); 5.89 (s, 2H, -O-CH₂-O-); 6.41 (s, 2H, H-2', H-6'); 6.62 (d, *J* = 2 Hz, 1H, H-4''); 6.66 (d, *J* = 8 Hz, 1H, H-7''); 6.73 (d, *J*₁ = 2, *J*₂ = 8 Hz, 1H, H-6''); 6.84 (s, 1H, H-3).
- ¹³C NMR (CDCl₃, 100.6 MHz) δ(ppm), 52.3 (CH₃-O-); 56.1 (CH₃-O- (x2)); 60.9 (CH₃-O-C=O); 101.3 (-O-CH₂-O-); 106.3 (CH, C-2', C-6'); 108.1 (CH, C-6''); 109.6 (CH, C-7''); 126.8 (CH, C-4''); 130.1 (C, C-1'); 131.2 (C, C-5''); 133.3 (C, C-2); 137.8 (C, C-4'); 140.2 (CH, C-3); 147.5 (C, C-7''a); 148.6 (C, C-3''a); 153.3 (C, C-3', C-5'); 168.3 (C, C=O).

Preparation of (*E*)-3-(3-hydroxy-4-methoxyphenyl)-2-phenylacrylic acid (23**)**



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon and equipped with a reflux condenser was charged with the starting aldehyde **78** (0.619 g, 4.07 mmol, 1 eq) dissolved in anhydrous acetic acid (5 mL).

Then, 2-phenylacetic acid (**69**) (0.556 g, 4.07 mmol, 1 eq) and triethylamine (2.5 mL, $d = 0.726$ g/mL, 17.94 mmol, 4 eq) were added to the mixture. The reaction mixture was heated at reflux of anhydrous acetic acid (140 ± 10 °C) under stirring for 6 hours.

Work-up

The crude product was microdistilled (7.5 mm Hg) under vacuum at 150 ± 10 °C. Then, the crude was dissolved with dichloromethane and extracted with an aqueous solution of HCl (2N) (3 X 20 mL). The combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was removed under reduced pressure.

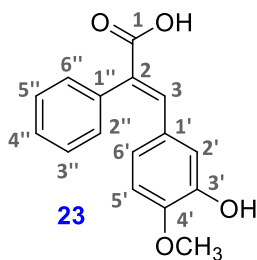
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: **81%**

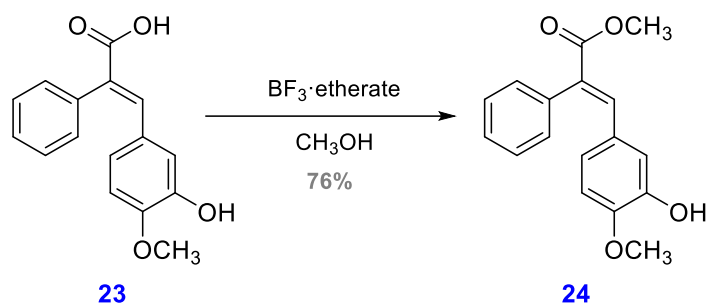
Analytical data

- R_f : 0.33 (Hexane/Ethyl Acetate (1:1))
- Aspect: yellow semisolid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 3.78 (s, 3H, $\text{CH}_3\text{-O-}$); 6.70 (d, $J = 2$ Hz, 1H, H-2'); 6.74 (d, $J = 8.7$ Hz, 1H, H-5'); 6.93 (dd, $J_1 = 2$, $J_2 = 8.7$ Hz, 1H, H-6'); 7.36-7.37 (m, 1H, H-4''); 7.38-7.39 (m, 2H, H-3'', H-5''); 7.39-7.41 (m, 2H, H-2'', H-6''); 7.85 (s, 1H, H-3); 9.87 (s, 1H, COOH).

Preparation of (*E*)-methyl 3-(3-hydroxy-4-methoxyphenyl)-2-phenylacrylate (**24**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon, was charged with compound **23** (0.900 g, 3.31 mmol, 1 eq) dissolved in methanol (10 mL). Then, 2 mL of boron trifluoride etherate ($d = 1.15 \text{ g/mL}$, 40.50 mmol) was slowly added. The resulting mixture was maintained under constant stirring at room temperature for 6 days. In three times 1 mL more of boron trifluoride etherate was added in order to obtain the final product.

Work-up

Methanol was removed under reduced pressure. Then, 15 mL of water was added to the reaction mixture and the crude product was extracted with ethyl acetate (3 X 20 mL). The combined organic layers were washed with an aqueous sodium bicarbonate solution (3 X 20 mL). Finally, the organic phases were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was removed under reduced pressure with toluene.

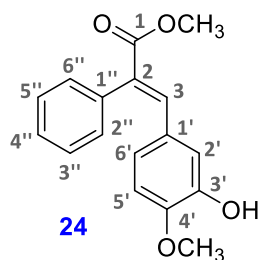
Purification

The crude product was used in the next step without further purification since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: 76%

Analytical data

- R_f : 0.93 (Hexane/Ethyl Acetate (1:1))
- Aspect: brown semisolid

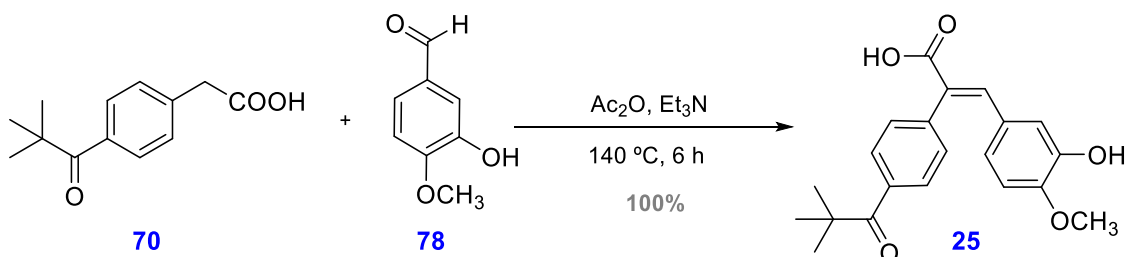


- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 3.76 (s, 3H, $\text{CH}_3\text{-O-}$); 3.84 (s, 3H, $\text{CH}_3\text{-O-C=O}$); 6.54 (d, $J = 2$ Hz, 1H, H-2'); 6.61 (dd, $J_1 = 2$, $J_2 = 8.4$ Hz, 1H, H-6'); 6.64 (d, $J = 8.4$ Hz, 1H, H-5'); 7.21 (d, $J = 2$ Hz, 1H, H-2''); 7.22 (d, $J = 2$ Hz, 1H, H-6''); 7.36-7.40 (m, 3H, H-3'', H-4'', H-5''); 7.75 (s, 1H, H-3).

*Interchangeable protons

- ^{13}C NMR (CDCl_3 , 100.6 MHz) δ (ppm), 52.3 ($\text{CH}_3\text{-O-}$); 55.8 ($\text{CH}_3\text{-O-C=O}$); 110.1 (CH, C-5'); 116.4 (CH, C-2'); 124.0 (CH, C-6'); 127.8 (CH, C-4''); 127.9 (C, C-1''); 128.7 (CH, C-3'', C-5''); 129.7 (CH, C-2'', C-6''); 130.4 (C, C-1'); 136.1 (C, C-2); 140.4 (C, C-3); 145.0 (C, C-3'); 147.5 (C, C-4'); 168.5 (C, C=O).

Preparation of (*E*)-3-(3-hydroxy-4-methoxyphenyl)-2-(4-pivaloylphenyl)acrylic acid (**25**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon and equipped with a reflux condenser was charged with the starting aldehyde **78** (0.375 g, 2.45 mmol, 1 eq) dissolved in anhydrous acetic acid (5 mL). Then, 2-(4-pivaloylphenyl)acetic acid (**70**) (0.541 g, 2.45 mmol, 1 eq) and triethylamine (2.5 mL, $d = 0.726$ g/mL, 17.94 mmol, 7 eq) were added to the mixture. The reaction mixture was heated at reflux of anhydrous acetic acid (140 ± 10 °C) under stirring for 6 hours.

Work-up

The crude product was microdistilled (7.5 mm Hg) under vacuum at 150 ± 10 °C. Then, the crude product was dissolved with dichloromethane and extracted with an aqueous solution of HCl (2N) (3 X 20 mL). The combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was removed under reduced pressure.

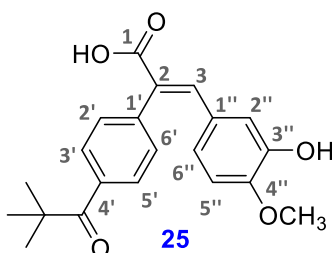
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: **100%**

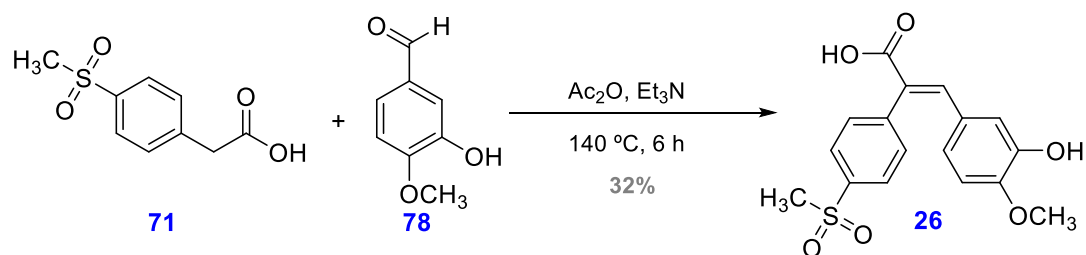
Analytical data

- R_f : 0.13 (Hexane/Ethyl Acetate (1:1))
- Aspect: brown semisolid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 1.21 (s, 9H, CH_3 - (x3)); 3.95 (s, 3H, CH_3 -O-); 6.70 (d, $J = 2$ Hz, 1H, H-2''); 6.74 (d, $J = 8.6$ Hz, 1H, H-5''); 6.94 (dd, $J_1 = 2$, $J_2 = 8.6$ Hz, 1H, H-6''); 6.95 (bs, 1H, OH); 7.07 (d, $J = 8.4$ Hz, 2H, H-2', H-6'); 7.76 (dd, $J_1 = 2$, $J_2 = 8.4$ Hz, 2H, H-3', H-5'); 7.85 (s, 1H, H-3); 9.86 (s, 1H, COOH).

Preparation of (E)-3-(3-hydroxy-4-methoxyphenyl)-2-(4-(methylsulfonyl)phenyl)acrylic acid (**26**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon and provided with a reflux condenser, was charged with the initial aldehyde **78** (0.710 g, 4.66 mmol, 1 eq) dissolved in anhydrous acetic acid (5 mL). Then, 2-(4-(methylsulfonyl)phenyl)acetic acid (**71**) (1 g, 4.66 mmol, 1 eq) and trimethylamine (2.5 mL, $d = 0.726\text{ g/mL}$, 36.26 mmol, 8 eq) were added and the reaction mixture was stirred and heated at $140 \pm 10\text{ }^\circ\text{C}$ for 6 hours.

Work-up

The crude product was microdistilled (7.5 mm Hg) under vacuum at $150 \pm 10\text{ }^\circ\text{C}$. Then, the crude was dissolved with dichloromethane and extracted with an aqueous solution of HCl (2N) (3 X 20 mL). The combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was removed under reduced pressure.

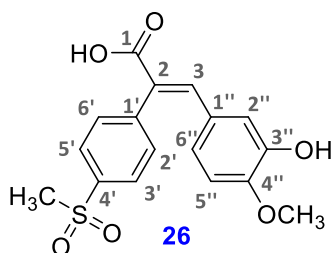
Purification

The crude reaction was purified by column chromatography using a CombiFlash[®] R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (30:70).

Yield: 32%

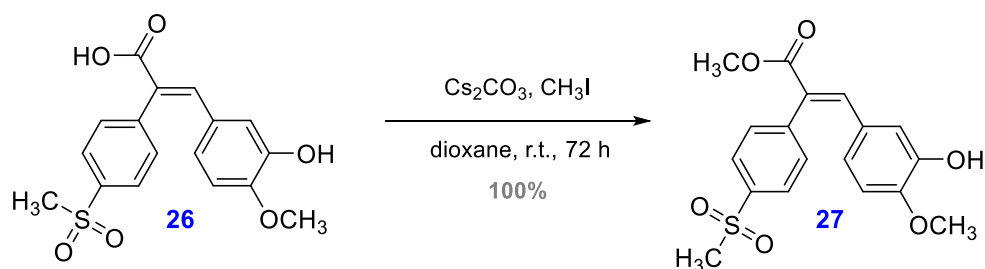
Analytical data

- R_f: 0.05 (Hexane/Ethyl Acetate (1:1))
- Aspect: yellow semisolid



- ^1H NMR ($(\text{CD}_3)_2\text{CO}$, 400 MHz) δ (ppm), 3.03 (s, 3H, $\text{CH}_3\text{-S-}$); 3.64 (s, 3H, $\text{CH}_3\text{-O-}$); 6.56 (d, $J = 2$ Hz, 1H, H-2''); 6.81 (d, $J = 8.6$ Hz, 1H, H-5''); 6.84 (d, $J_1 = 2, J_2 = 8.6$ Hz, 1H, H-6''); 7.36 (d, $J = 8.2$ Hz, 2H, H-2', H-6'); 7.72 (s, 1H, H-3); 7.82 (d, $J = 8.2$ Hz, 2H, H-3', H-5'); 9.76 (bs, 1H, COOH).
- ^{13}C NMR ($(\text{CD}_3)_2\text{CO}$, 100.6 MHz) δ (ppm), 43.5 ($\text{CH}_3\text{-S-}$); 55.5 ($\text{CH}_3\text{-O-}$); 112.3 (CH, C-5''); 124.9 (CH, C-2''); 126.8 (C, C-1''); 127.5 (CH, C-2', C-6'); 129.6 (CH, C-6''); 130.4 (C, C-2); 130.9 (CH, C-3', C-5'); 139.5 (C, C-1'); 139.6 (CH, C-3); 140.5 (C, C-4'); 142.4 (C, C-3''); 152.4 (C, C-4''); 167.9 (C, COOH).

Preparation of (*E*)-methyl 3-(3-hydroxy-4-methoxyphenyl)-2-(4-(methylsulfonyl)phenyl)acrylate (**27**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar and a reflux condenser was charged with the initial acid **26** (0.262 g, 0.75 mmol, 1 eq) dissolved in 10 mL of dioxane. Then, caesium carbonate (0.294 g, 0.90 mmol, 1.2 eq) and methyl iodide (1 mL, $d = 2.28$ g/mL, 32.13 mmol) were added to the reaction mixture. The mixture was stirred at room temperature for 72 hours in which 1 mL more of methyl iodide was added to the reaction mixture.

Work-up

The reaction mixture was diluted with 20 mL of water and extracted with dichloromethane/ethyl acetate 1/1 (3 X 20 mL).

The organic layer was dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was removed under reduced pressure with the aid of toluene.

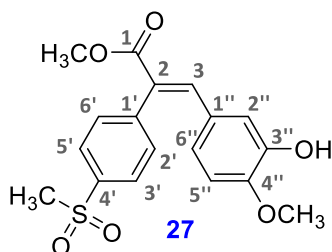
Purification

The crude product was directly used for the next synthetic step since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: **100%**

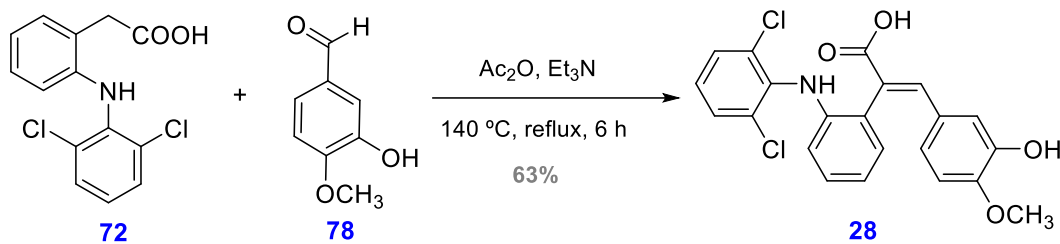
Analytical data

- R_f : 0.36 (Hexane/Ethyl acetate (3:7))
- Aspect: yellow semisolid



- ^1H NMR ($(\text{CD}_3)_2\text{CO}$, 400 MHz) δ (ppm), 3.18 (s, 3H, $\text{CH}_3\text{-S-}$); 3.75 (s, 3H, $\text{CH}_3\text{-O-}$); 3.81 (s, 3H, $\text{CH}_3\text{-O-C=O}$); 6.67 (d, $J = 2\text{ Hz}$, 1H, H-2''); 6.98 (d, $J = 8\text{ Hz}$, 1H, H-5''); 6.99 (dd, $J_1 = 2, J_2 = 8\text{ Hz}$, 1H, H-6''); 7.51 (d, $J = 8.6\text{ Hz}$, 2H, H-2', H-6'); 7.84 (s, 1H, H-3); 7.98 (d, $J = 8.6\text{ Hz}$, 2H, H-3', H-5'); 8.16 (bs, 1H, OH).

Preparation of (*E*)-2-(2-((2,6-dichlorophenyl)amino)phenyl)-3-(3-hydroxy-4-methoxyphenyl)acrylic acid (**28**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flamed-dried under argon and with a reflux condenser, was charged with the starting aldehyde **78** (0.256 g, 1.68 mmol, 1 eq) dissolved in anhydrous acetic acid (5 mL). Then, 2-(2-((2,6-dichlorophenyl)amino)phenyl)acetic acid (**72**) (0.5 g, 1.68 mmol, 1 eq) and triethylamine (2.5 mL, $d = 0.726$ g/mL, 17.94 mmol, 10 eq) were added to the mixture. The reaction mixture was heated at reflux of anhydrous acetic acid (140 ± 10 °C) under stirring for 6 hours.

Work-up

The crude product was microdistilled (7.5 mm Hg) under vacuum at 150 ± 10 °C in order to eliminate the anhydrous acetic acid. Then, the crude was dissolved with dichloromethane and extracted with an aqueous solution of HCl (2N) (3 X 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was removed under reduced pressure.

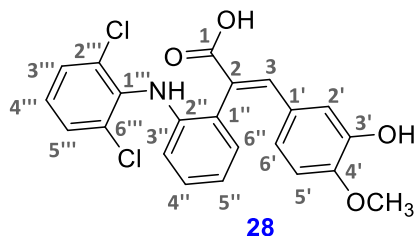
Purification

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (15:85).

Yield: 63%

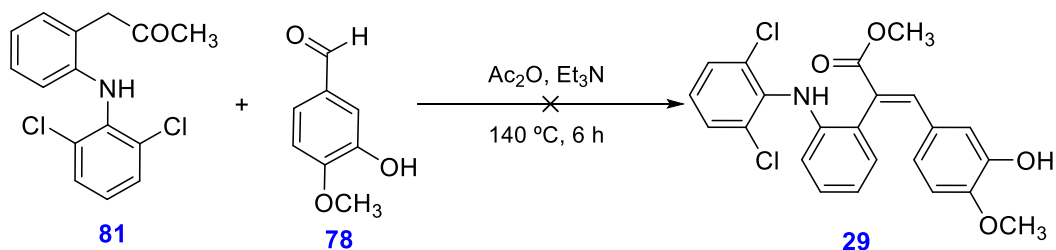
Analytical data

- R_f: 0.53 (Hexane/Ethyl Acetate (1:1))
- Aspect: brown semisolid



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 3.91 (s, 3H, CH₃-O-); 6.40 (d, *J* = 8 Hz, 1H, H-5'); 6.98 (t, *J* = 7 Hz, 1H, H-5''); 7.07 (d, *J* = 8 Hz, 1H, H-6'); 7.20 (t, *J* = 7 Hz, 1H, H-4''); 7.35-7.44 (m, 2H, H-4''', H-3''); 7.51 (t, *J* = 4 Hz, 2H, H-3''', H-5'''); 7.59 (s, 1H, H-2'); 7.76 (d, *J* = 8.4 Hz, 1H, H-6''); 7.87 (s, 1H, H-3); 8.08 (bs, 1H, OH); 8.52 (bs, 1H, NH); 9.84 (bs, 1H, COOH).
- ¹³C NMR (CDCl₃, 100.6 MHz) δ(ppm), 56.0 (CH₃-O-); 109.2 (CH, C-5'); 112.0 (CH, C-2'); 122.6 (CH, C-5''); 123.0 (CH, C-3''); 123.4 (CH, C-4'''); 124.3 (CH, C-6'); 128.2 (CH, C-6''); 128.5 (C, C-1'); 128.90-129.0 (CH, C-3''', C-5'''); 130.1 (CH, C-1''); 130.2 (C, C-4''); 135.8 (C, C-2); 137.6 (C, C-1'''); 139.3 (C, C-2''); 140.2 (C, C-2''', C-6'''); 142.1 (C, C-3'); 152.5 (C, C-4'); 168.8 (C, C=O); 190.0 (CH, C-3).

Preparation of (*E*)-methyl 2-(2-((2,6-dichlorophenyl)amino)phenyl)-3-(3-hydroxy-4-methoxyphenyl)acrylate (**29**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon with a reflux condenser was charged with the starting ester **81** (0.238 g, 0.76 mmol, 1 eq) dissolved in anhydrous acetic acid (5 mL). Then, triethylamine (2.5 mL, *d* = 0.726 g/mL, 17.94 mmol) was added to the mixture. After 30 minutes the substituted phenyl aldehyde **78** (0.117 g, 0.76 mmol, 1 eq) was added.

The reaction mixture was heated at reflux of anhydrous acetic acid (140 ± 10 °C) under stirring for 6 hours.

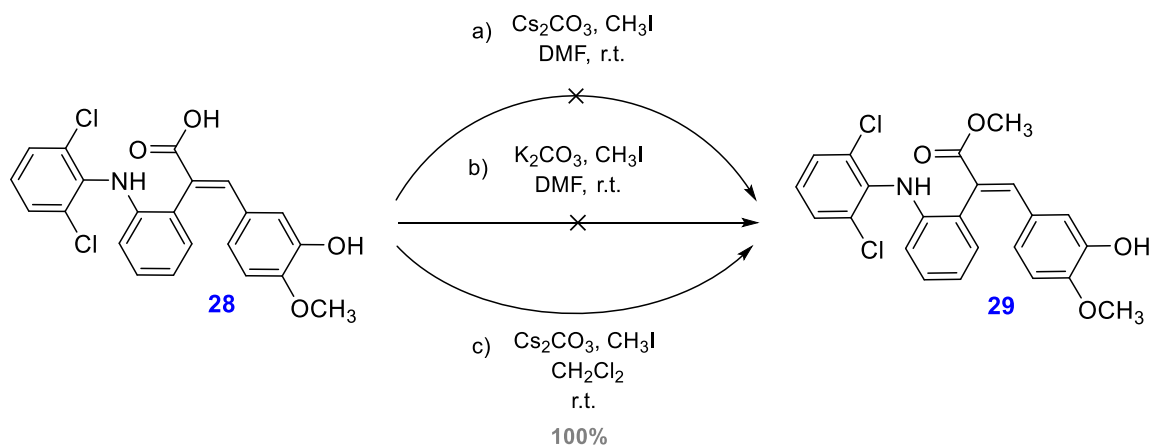
Work-up

The crude product was microdistilled (7.5 mm Hg) under vacuum at 150 ± 10 °C. Then, the crude was dissolved with dichloromethane and extracted with an aqueous solution of HCl (2N) (3 X 20 mL). The combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was removed under reduced pressure.

Purification

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The final product was unwanted because the initial unreacted products were obtained.

Preparation of (*E*)-methyl 2-(2-((2,6-dichlorophenyl)amino)phenyl)-3-(3-hydroxy-4-methoxyphenyl)acrylate (**29**)



Procedure A

A 100 mL round-bottomed flask equipped with a magnetic stirring bar and a reflux condenser was charged with the initial acid **28** (0.650 g, 1.51 mmol, 1 eq) dissolved in 10 mL of dimethylformamide. Then, caesium carbonate (0.590 g, 1.81 mmol, 1.2 eq) and methyl iodide (1 mL, $d = 2.28$ g/mL, 32.13 mmol) were added. The mixture was stirred at room temperature for 4 days in which 1 mL more of methyl iodide was added to the reaction mixture.

Work-up

The reaction mixture was diluted with water (15 mL) and extracted with diethyl ether (3 X 15 mL). The combined organic layers were washed with water (3 X 15 mL), and the organic phase was dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure, only the initial compound was recovered in this case.

Procedure B

A 100 mL round-bottomed flask equipped with a magnetic stirring bar and a reflux condenser was charged with the initial acid **28** (0.405 g, 0.94 mmol, 1 eq) dissolved in 10 mL of DMF. Then, potassium carbonate (0.195 g, 1.41 mmol, 1.5 eq) and methyl iodide (1 mL, $d = 2.28 \text{ g/mL}$, 64.25 mmol) were added. The mixture was stirred at room temperature for 48 hours in which in three times 1 mL more of methyl iodide was added to the reaction mixture.

Work-up

The reaction mixture was diluted with water (15 mL) and extracted with diethyl ether (3 X 15 mL). The combined organic layers were washed with water (3 X 15 mL), the organic phases were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

The product obtained was the initial product, as shown by ^1H NMR spectrum.

On the other hand, the aqueous phases obtained from the extraction with diethyl ether were acidified with an aqueous solution of HCl (1N). Then, they were extracted with dichloromethane (3 X 15 mL). The combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure. As a result, the isolated product was the unreacted aldehyde. So, an unwanted retrosynthesis was produced.

Purification

No purification was performed since no new product was observed by TLC or ^1H NMR.

Procedure C

A 100 mL round-bottomed flask equipped with a magnetic stirring bar and a reflux condenser was charged with the acid **28** (0.168 g, 0.39 mmol, 1 eq) dissolved in 10 mL of dichloromethane. Then, under 0 ± 10 °C conditions (ice-bath) caesium carbonate (0.153 g, 0.47 mmol, 1.2 eq) and 1 mL of methyl iodide ($d = 2.28$ g/mL, 48.20 mmol) were added to the reaction mixture. The resulting mixture remained for 72 hours at room temperature under constant stirring, where during this time in two occasions 1 mL of methyl iodide was added.

Work-up

The reaction mixture was diluted with water (20 mL) and extracted with dichloromethane (3 X 20 mL). Then, the combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure. As shown in the ^1H NMR spectrum the product was obtained plus the initial acid.

The crude mixture was basified with an aqueous sodium bicarbonate solution until a basic pH was reached. Then, the aqueous layer was basified and the extraction was performed with diethyl ether (3 X 20 mL). The combined organic layers were separated and were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was evaporated under reduced pressure. The product achieved was the desired.

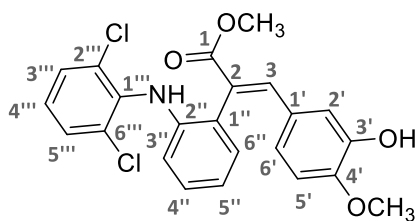
Purification

The crude product was directly used for the next step since the product was obtained with enough purity as shown by TLC and ^1H NMR.

Yield: **100%**

Analytical data

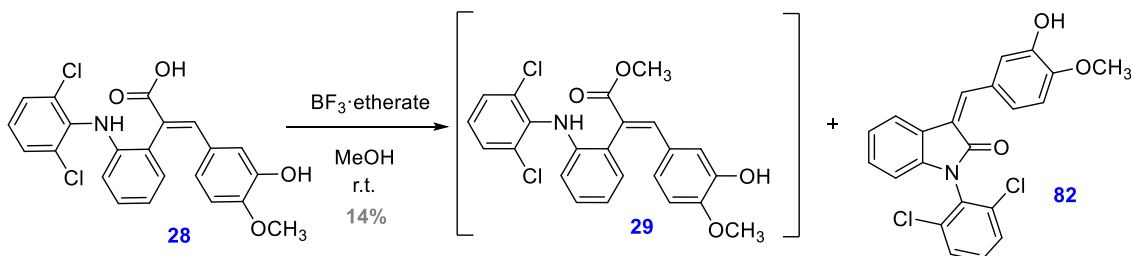
- R_f : 0.29 (Hexane/Ethyl Acetate (1:1))
- Aspect: brown semisolid



29

- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 3.92 (s, 3H, $\text{CH}_3\text{-O-}$); 3.98 (s, 3H, $\text{CH}_3\text{-O-C=O}$); 6.40 (d, $J = 8$ Hz, 1H, H-5'); 6.98 (d, $J = 8$ Hz, 1H, H-6'); 7.20 (t, $J = 8.8$ Hz, 1H, H-5''); 7.29 (d, $J = 2$ Hz, 1H, H-2'); 7.36-7.40 (m, 2H, H-4''', H-6'''); 7.51 (d, $J = 8$ Hz, 1H, H-3''); 7.52 (d, $J = 8$ Hz, 1H, H-6''); 7.60 (dt, $J_1 = 2, J_2 = 8$ Hz, 1H, H-4''); 7.91 (s, 1H, H-3); 8.68 (bs, 1H, NH).
- ^{13}C NMR (CDCl_3 , 100.6 MHz) δ (ppm), 56.00 ($\text{CH}_3\text{-O-}$); 56.2 ($\text{CH}_3\text{-O-C=O}$); 109.3 (CH, C-5'); 111.1 (CH, C-2'); 112.4 (CH, C-6'); 118.8 (C, C-4'''); 122.4 (C, C-1'); 122.5 (CH, C-3''); 127.3 (CH, C-5''); 127.4 (C, C-1''); 128.2 (CH, C-3''', C-5'''); 129.0 (CH, C-6''); 130.6 (C, C-2); 130.7 (CH, C-4''); 135.8 (C, C-1'''); 135.9 (C, C-2''); 138.8 (C, C-2''', C-6'''); 148.8 (C, C-3'); 150.6 (C, C-4'); 167.3 (C, C=O); 191.0 (CH, C-3).

Preparation of (*E*)-methyl 2-(2-((2,6-dichlorophenyl)amino)phenyl)-3-(3-hydroxy-4-methoxyphenyl)acrylate (**29**)



Procedure

A 100 mL round-bottomed flask with a magnetic stirring bar, previously flame-dried under argon, was charged with product **28** (0.476 g, 1.1 mmol, 1 eq) dissolved in methanol (10 mL). Then, 2 mL of boron trifluoride etherate ($d = 1.15$ g/mL, 32.41 mmol) was slowly added. The resulting mixture was maintained under constant stirring at room temperature for 72 hours. During this time 2 mL more of boron trifluoride etherate was added to the reaction mixture since TLC did not show any new product formatted.

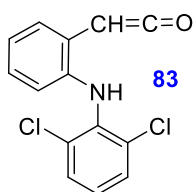
Work-up

Methanol was removed under reduced pressure. Then, 15 mL of water was added to the reaction mixture and the crude product was extracted with ethyl acetate (3 X 20 mL). The combined organic layers were washed with an aqueous sodium bicarbonate solution (3 X 20 mL). Finally, the combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was removed under reduced pressure with toluene.

Purification

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. While the desired product was not shown, the (*Z*)-1-(2,6-dichlorophenyl)-3-(3-hydroxy-4-methoxybenzylidene)indolin-2-one (**82**) was eluted with a mixture of hexane/ethyl acetate (60:40).

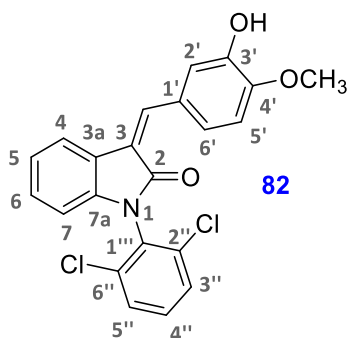
The treatment of this product with hydrazine monohydrate gave the following product, 2-(2-((2,6-dichlorophenyl)amino)phenyl)ethenone (**83**):



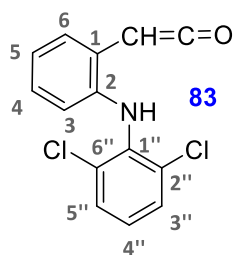
Yield: **14%, 25%**

Analytical data for compounds **82** and **83**

- R_f: 0.32 (Hexane/Ethyl Acetate (6:4)), 0.93 (Hexane/Ethyl Acetate (6:4))
- Aspect: yellow semisolid (x2)

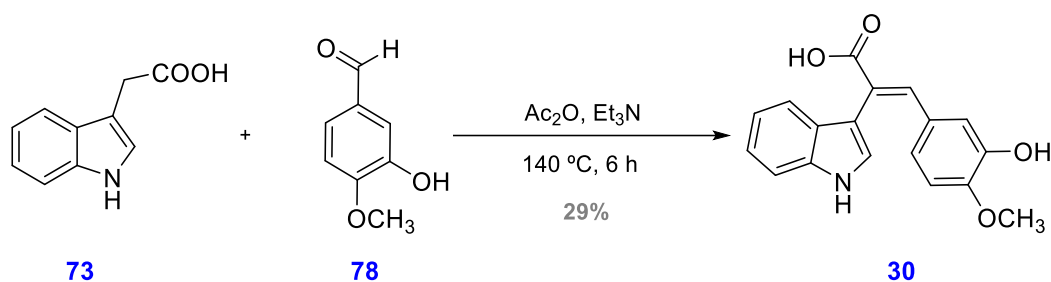


- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 3.89 (s, 3H, $\text{CH}_3\text{-O-}$); 6.33 (d, $J = 8$ Hz, 1H, H-5'); 6.87 (d, $J = 8$ Hz, 1H, H-6'); 6.90 (dd, $J_1 = 1$, $J_2 = 7.6$ Hz, 1H, H-2'); 7.09 (dt, $J_1 = 1$, $J_2 = 7.7$ Hz, 1H, H-4''); 7.10 (dd, $J_1 = 1$, $J_2 = 7.7$ Hz, 1H, H-6); 7.22 (dd, $J_1 = 2$, $J_2 = 8$ Hz, 1H, CH=); 7.27 (d, $J = 7$ Hz, 1H, H-7); 7.28 (t, $J = 7$ Hz, 1H, H-5); 7.43 (d, $J_1 = 1$, $J_2 = 7.7$ Hz, 2H, H-3'', H-5''); 7.80 (bs, 1H, OH); 7.83 (d, $J = 7.7$ Hz, 1H, H-4).
- ^{13}C NMR (CDCl_3 , 100.6 MHz) δ (ppm), 55.9 ($\text{CH}_3\text{-O-}$); 108.8 (CH, C-2'); 110.1 (CH, C-5'); 115.6 (CH, C-5); 122.8 (C, C-3); 123.0 (CH, C-6'); 124.7 (C, C-3a); 125.0 (CH, C-4''); 128.8 (C, C-1'); 128.9 (CH, C-3'', C-5''); 130.5 (CH, C-4); 130.6 (CH, C-7); 138.7 (C, C-7a); 135.8 (C, C-2'', C-6''); 142.0 (C, C-1''); 145.1 (C, C-3'); 148.1 (C, C-4'); 167.3 (C=O).
- HRMS ESI (+) m/z: calculated mass for $\text{C}_{22}\text{H}_{16}\text{NO}_3\text{Cl}_2$ 412.0429, found 412.0511.
- HRMS ESI (-) m/z: calculated mass for $\text{C}_{22}\text{H}_{14}\text{NO}_3\text{Cl}_2$ 410.0429, found 410.0358.



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 6.40 (dd, $J_1 = 0.5$, $J_2 = 7.8$ Hz, 1H, H-6); 7.09 (dt, $J_1 = 1$, $J_2 = 7.8$ Hz, 1H, H-5); 7.20 (dt, $J_1 = 1$, $J_2 = 7.8$ Hz, 1H, H-4); 7.32 (d, $J = 7.8$ Hz, 1H, H-3); 7.37 (dt, $J_1 = 0.7$, $J_2 = 7.9$ Hz, 1H, H-4'); 7.50 (dd, $J_1 = 0.7$, $J_2 = 7.9$ Hz, 2H, H-3', H-5').
- ^{13}C NMR (CDCl_3 , 100.6 MHz) δ (ppm), 71.0 (CH, CH=C=O); 109.1 (CH, C-3); 123.0 (CH, C-4'); 124.0 (C, C-1); 124.2 (CH, C-4); 127.9 (CH, C-5); 129.0 (CH, C-3', C-5'); 130.9 (CH, C-6); 135.5 (C, C-2', C-6'); 143.3 (C, C-1'); 144.7 (C, C-2); 173.6 (C, C=O).
- HRMS ESI (+) m/z: calculated mass for $\text{C}_{14}\text{H}_{10}\text{NOCl}_2$ 278.0139, found 278.0154.

Preparation of (*E*)-3-(3-hydroxy-4-methoxyphenyl)-2-(1*H*-indol-3-yl)acrylic acid (**30**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon and with a reflux condenser was charged with the starting aldehyde **78** (0.434 g, 2.85 mmol, 1 eq) dissolved in anhydrous acetic acid (5 mL). Then, 2-(1*H*-indol-3-yl)acetic acid (**73**) (0.5 g, 2.85 mmol, 1 eq) and triethylamine (2.5 mL, *d* = 0.726 g/mL, 17.93 mmol, 6.3 eq) were added to the mixture. The reaction mixture was heated at reflux of anhydrous acetic acid (140 ± 10 °C) under stirring for 6 hours.

Work-up

The crude product was microdistilled (7.5 mm Hg) under vacuum at 150 ± 10 °C. Then, the crude was dissolved with dichloromethane and extracted with an aqueous solution of HCl (2N) (3 X 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was removed under reduced pressure.

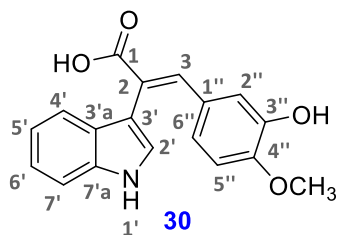
Purification

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (60:40).

Yield: 29%

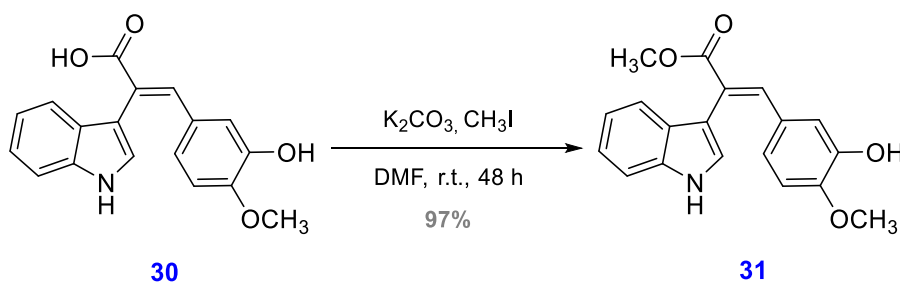
Analytical data

- R_f: 0.28 (Hexane/Ethyl Acetate (1:1))
- Aspect: brown semisolid



- ^1H NMR ($(\text{CD}_3)_2\text{CO}$, 400 MHz) δ (ppm), 3.01 (s, 3H, $\text{CH}_3\text{-O-}$); 6.90 (d, $J = 8.7$ Hz, 1H, H-5''); 7.00 (d, $J = 2$ Hz, 1H, H-2''); 7.15 (d, $J = 6.7$ Hz, 1H, H-7'); 7.17 (t, $J = 6.7$ Hz, 1H, H-6'); 7.20 (d, $J = 6.7$ Hz, 1H, H-4'); 7.32 (dt, $J_1 = 2, J_2 = 6.7$ Hz, 1H, H-5'); 7.74 (s, 1H, H-3); 7.97 (s, 1H, H-2'); 8.43 (d, $J = 8.7$ Hz, 1H, H-6'').

Preparation of (*E*)-methyl 3-(3-hydroxy-4-methoxyphenyl)-2-(1*H*-indol-3-yl)acrylate (**31**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon was charged with (*E*)-3-(3-hydroxy-4-methoxyphenyl)-2-(1*H*-indol-3-yl)acrylic acid (**30**) (0.258 g, 0.834 mmol, 1 eq), potassium carbonate (0.17 g, 1.25 mmol, 1.5 eq) and 1 mL of methyl iodide ($d = 2.28$ g/mL, 64.59 mmol). The mixture was stirred at room temperature for 48 hours. In order to promote the formation of the final compound, in three times 1 mL more of methyl iodide was added during this time.

Work-up

The reaction mixture was diluted with water (15 mL) and extracted with diethyl ether (3 X 15 mL). The combined organic layers were washed with water (3 X 20 mL), dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

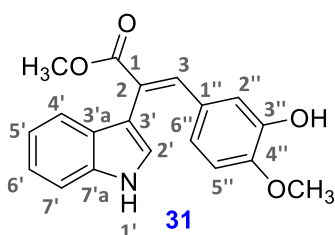
Purification

The crude product was directly used for the next synthetic step since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: **97%**

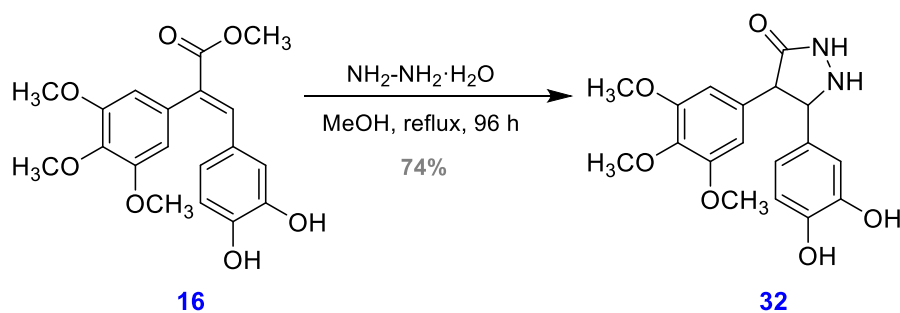
Analytical data

- R_f: 0.42 (Hexane/Ethyl Acetate (1:1))
- Aspect: yellow semisolid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 3.75 (s, 3H, $\text{CH}_3\text{-O-Ar}$); 3.78 (s, 3H, $\text{CH}_3\text{-O-C=O}$); 6.72 (d, $J = 8.5$ Hz, 1H, H-5''); 6.88 (d, $J = 2$ Hz, 1H, H-7'); 7.07 (dd, $J_1 = 2$, $J_2 = 8.5$ Hz, 1H, H-6'); 7.18-7.20 (m, 2H, H-4', H-2''); 7.34 (t, $J = 4$ Hz, 1H, H-5'); 7.38 (s, 1H, H-3); 7.96 (s, 1H, H-2'); 8.48 (d, $J = 8.5$ Hz, 1H, H-6'').

Preparation of 5-(3,4-dihydroxyphenyl)-4-(3,4,5-trimethoxyphenyl)pyrazolidin-3-one (**32**)



Procedure

A 100 mL round bottomed flask equipped with a reflux condenser and a magnetic stirring bar was charged with the starting ester **16** (0.200 g, 0.55 mmol, 1 eq) dissolved in 30 mL of methanol. Then, an excess of an aqueous solution of hydrazine (1 mL, $d = 1.011$ g/mL, 35% water, 40.39 mmol) was added. The reaction mixture was stirred at reflux

temperature of methanol (80 ± 10 °C) for 96 hours, during this time 1 mL of hydrazine hydrate was added in order to obtain the final product.

Work-up

Methanol was removed under reduced pressure. Then, the reaction mixture was diluted with water (15 mL) and the organic phase was extracted with ethyl acetate/diethyl ether 1/1 (3 X 20 mL). The combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

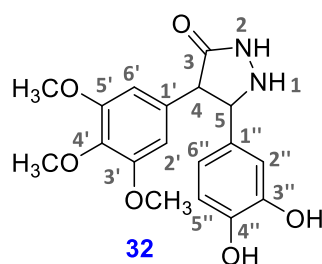
Purification

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a polarity of hexane/ethyl acetate (5:95). The product obtained was a mixture of rotamers or conformers as shown in the ¹H spectrum.

Yield: **74%**

Analytical data

- R_f: 0.07 (Hexane/Ethyl Acetate (4:6))
- Aspect: yellow semisolid

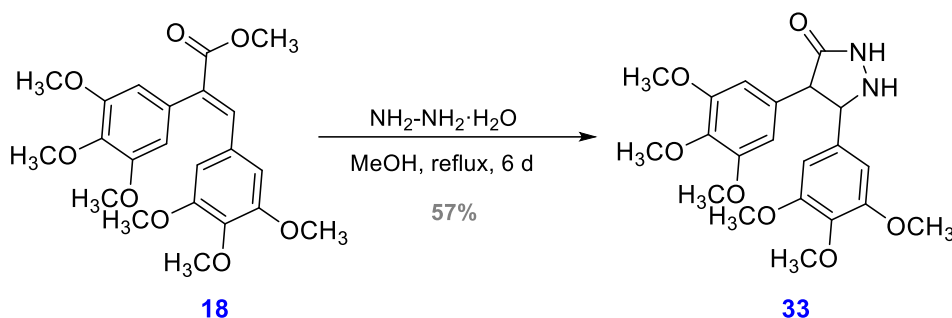


- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 2.76-2.82 (m, 1H, H-4); 3.72 (s, 3H, CH₃-O-); 3.73 (s, 6H, CH₃-O- (x2)); 4.61-4.63 (m, 1H, H-5); 6.51 (bs, 2H, OH (x2)); 6.54 (s, 2H, H-2', H-6'); 6.59 (d, *J* = 3 Hz, 1H, H-2''); 6.62 (d, *J* = 8 Hz, 1H, H-5''); 6.63 (d, *J* = 8 Hz, 1H, H-6''); 8.2 (bs, 1H, NH); 8.3 (bs, 1H, NH).
- ¹³C NMR (CDCl₃, 100.6 MHz) δ(ppm), 39.1 (CH, C-4); 49.0 (CH, C-5); 55.9 (CH₃-O- (x2)); 60.6 (CH₃-O-); 104.9 (CH, C-2', C-6') (B); 105.4 (CH, C-2', C-6') (A); 114.8 (CH, C-5'') (A); 114.9 (CH, C-5'') (B); 115.8 (CH, C-2'') (B); 115.9 (CH, C-2'') (A); 120.7

(CH, C-6'); 131.3 (C, C-1') (A); 131.8 (C, C-1') (B); 135.2 (C, C-1'') (B); 135.3 (C, C-1'') (A); 136.6 (C, C-4') (A); 136.9 (C, C-4') (B); 142.4 (C, C-4'') (A); 142.7 (C, C-3'') (A); 143.7 (C, C-4'') (B); 143.9 (C, C-3'') (B); 152.8 (C, C-3', C-5') (A); 153.0 (C, C-3', C-5') (B); 173.2 (C, C=O) (B); 174.5 (C, C=O) (A).

A: majority isomer / B: minority isomer

Preparation of 4,5-bis(3,4,5-trimethoxyphenyl)pyrazolidin-3-one (**33**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar and a reflux condenser was charged with the initial ester **18** (0.266 g, 0.63 mmol, 1 eq) in 30 mL of methanol. Then, an excess of an aqueous solution of hydrazine was added (1 mL, $d = 1.011 \text{ g/mL}$, 35% water, 60.58 mmol). The reaction was heated at reflux temperature of $80 \pm 10 \text{ }^\circ\text{C}$ under stirring for 6 days, in two times 1 mL more of hydrazine hydrate was added to the reaction mixture during these days in order to reach the desired product.

Work-up

Methanol was firstly evaporated under reduced pressure and then 20 mL of water was added to the mixture and the crude reaction was extracted with diethyl ether/ethyl acetate 1/1 (3 X 20 mL). The organic layer was dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was removed under reduced pressure.

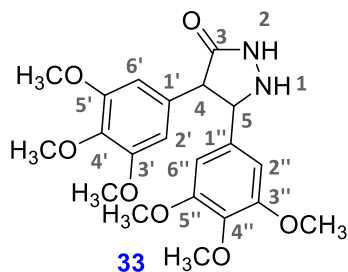
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: **57%**

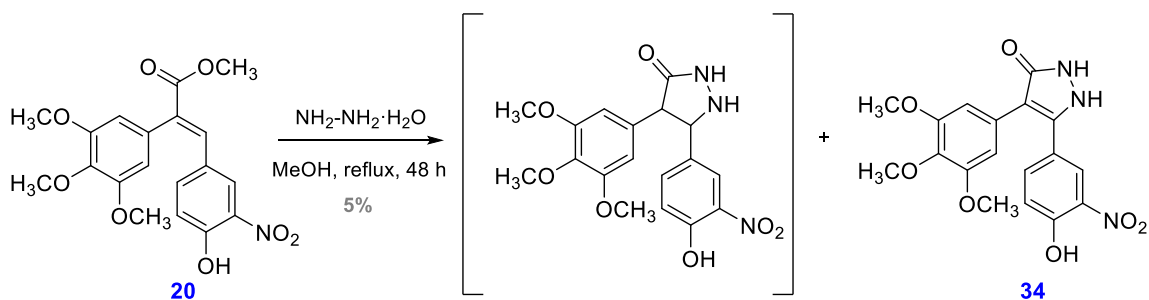
Analytical data

- R_f: 0.10 (Hexane/Ethyl Acetate (2:8))
- Aspect: orange-brown semisolid



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 2.83-2.88 (m, 1H, H-5); 3.29-3.38 (m, 1H, H-4); 3.67 (s, 6H, CH₃-O- (x2)); 3.69 (s, 3H, CH₃-O-); 3.73 (s, 9H, CH₃-O- (x3)); 6.22 (s, 2H, H-2', H-6'); 6.45 (s, 2H, H-2'', H-6'').
- ¹³C NMR (CDCl₃, 100.6 MHz) δ(ppm), 29.6 (CH, C-5); 40.0 (CH, C-4); 53.7 (CH₃-O-); 56.0 (CH₃-O- (x2)); 56.1 (CH₃-O- (x2)); 60.7 (CH₃-O-); 105.1 (CH, C-2', C-6'); 105.9 (CH, C-2'', C-6''); 128.1 (C, C-4'); 128.9 (C, C-4''); 136.4 (CH, C-1'); 137.3 (CH, C-1''); 152.9 (C, C-3', C-5'); 153.2 (C, C-3'', C-5''); 173.4 (C, C=O).

Preparation of 5-(4-hydroxy-3-nitrophenyl)-4-(3,4,5-trimethoxyphenyl)-1H-pyrazol-3(2H)-one (**34**)



Procedure

A 100 mL round bottomed flask equipped with a magnetic stirring bar and a reflux condenser was charged with the starting ester **20** (0.292 g, 0.75 mmol, 1 eq) dissolved in 30 mL of methanol. Then, an aqueous solution of hydrazine (1 mL, d = 1.011 g/mL, 35% water, 40.39 mmol) was added. The reaction mixture was stirred at reflux temperature of methanol (80 ± 10 °C) for 48 hours, during this time 1 mL more of hydrazine hydrate was added to the reaction mixture.

Work-up

Methanol was removed under reduced pressure. Then, the reaction mixture was diluted with water (15 mL) and the organic phase was extracted with diethyl ether/ethyl acetate 1/1 (3 X 20 mL). The combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

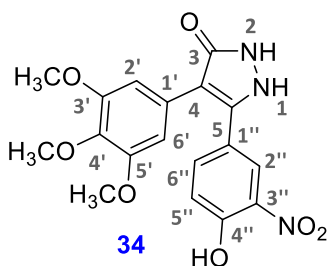
Purification

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (40:60). In this case, the oxidized compound was obtained.

Yield: 5%

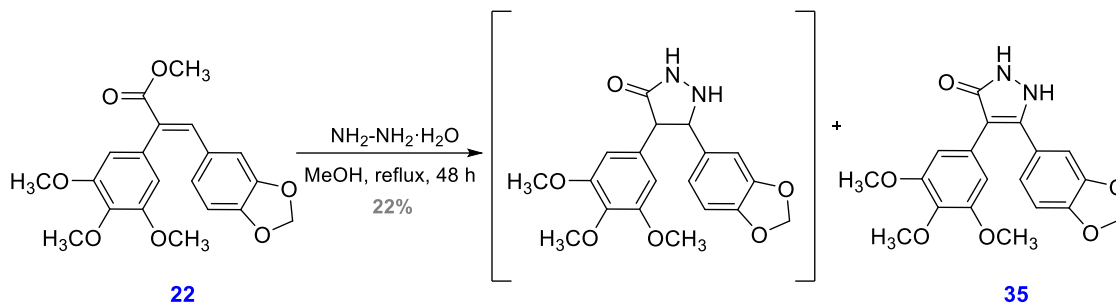
Analytical data

- R_f: 0.22 (Hexane/Ethyl Acetate (1:1))
- Aspect: yellow solid



- ¹H NMR ((CDCl₃, 400 MHz) δ(ppm), 3.96 (s, 6H, CH₃-O- (x2)); 4.02 (s, 3H, CH₃-O-); 6.23 (s, 2H, H-2', H-6'); 7.16 (d, *J* = 8.7 Hz, 1H, H-5''); 7.67 (s, 1H, OH); 7.74 (dd, *J*₁ = 2, *J*₂ = 8.7 Hz, 1H, H-6''); 8.35 (d, *J* = 2 Hz, 1H, H-2''); 8.60 (s, 2H, NH).

Preparation of 5-(benzo[d][1,3]dioxol-5-yl)-4-(3,4,5-trimethoxyphenyl)-1H-pyrazol-3(2H)-one (35)



Procedure

A 100 mL round bottomed flask equipped with a magnetic stirring bar and a reflux condenser was charged with the starting ester **22** (0.236 g, 0.63 mmol, 1 eq) in 30 mL of methanol. Then, an excess of an aqueous solution of hydrazine (1 mL, $d = 1.011 \text{ g/mL}$, 35% water, 41.23 mmol) was added. The reaction mixture was stirred at reflux temperature of methanol ($80 \pm 10 \text{ }^\circ\text{C}$) for 48 hours, during which 1 mL of hydrazine hydrate was added in order to promote the formation of the final product.

Work-up

Methanol was removed under reduced pressure. Then, the reaction mixture was diluted with water (15 mL) and the organic phase was extracted with diethyl ether/ethyl acetate 1/1 (3 X 20 mL). The combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

Purification

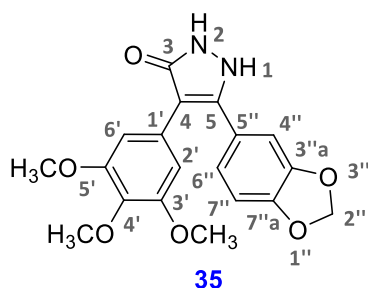
The crude reaction was purified by column chromatography using a CombiFlash[®] R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product was not obtained since the oxidized derivative was achieved and eluted with a mixture of hexane/ethyl acetate (40:60).

Yield: **22%**

Analytical data

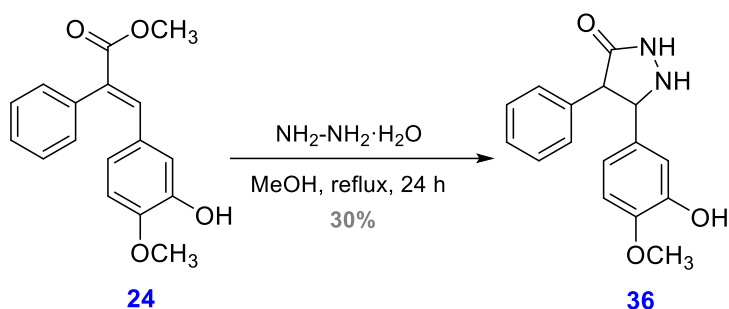
- R_f: 0.1 (Hexane/Ethyl Acetate (1:1))

- Aspect: yellow semisolid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 3.83 (s, 3H, $\text{CH}_3\text{-O-}$); 3.84 (s, 6H, $\text{CH}_3\text{-O-}$ (x2)); 5.90 (s, 2H, $-\text{O-CH}_2\text{-O-}$); 6.50 (s, 2H, H-2', H-6'); 6.57 (d, $J = 2$ Hz, 1H, H-4''); 6.62 (d, $J = 7.8$ Hz, 1H, H-7''); 6.68 (dd, $J_1 = 2$, $J_2 = 7.8$ Hz, 1H, H-6'').
- ^{13}C NMR (CDCl_3 , 100.6 MHz) δ (ppm), 55.4 ($\text{CH}_3\text{-O-}$ (x2)); 59.8 ($\text{CH}_3\text{-O-}$); 101.2 (CH_2 , $-\text{O-CH}_2\text{-O-}$); 102.4 (C, C-4); 105.3 (CH, C-2', C-6'); 107.9 (CH, C-4''); 108.6 (CH, C-7''); 123.9 (CH, C-6''); 129.7 (C, C-1''); 131.2 (C, C-5''); 136.3 (C, C-4'); 147.2 (C, C-7''a); 147.4 (C, C-3''a); 153.0 (C, C-5); 153.6 (C, C-3', C-5'); 170.6 (C, C=O).

Preparation of 5-(3-hydroxy-4-methoxyphenyl)-4-phenylpyrazolidin-3-one (**36**)



Procedure

A 100 mL round bottomed flask equipped with a magnetic stirring bar and a reflux condenser was charged with the starting ester **24** (0.693 g, 2.42 mmol, 1 eq) dissolved in 30 mL of methanol. Then, an excess of an aqueous solution of hydrazine (1 mL, $d = 1.011$ g/mL, 35% water, 40.39 mmol) was added. The reaction mixture was stirred at reflux temperature of methanol (80 ± 10 °C) for 24 hours, during which 1 mL of hydrazine hydrate was added to the reaction mixture.

Work-up

Methanol was removed under reduced pressure. Then, the reaction mixture was diluted with water (15 mL) and the organic phase was extracted with diethyl ether/ethyl acetate 1/1 (3 X 20 mL). The combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

Purification

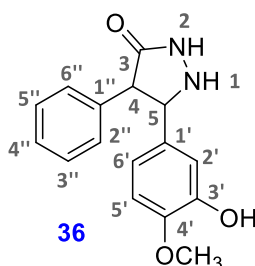
The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (15:85).

The product obtained was a brown semisolid. In this case, the product obtained was not 100% pure. Another extraction was carried out with diethyl ether (3 X 20 mL) and diluted with an aqueous HCl solution (1N) (3 X 10 mL) in order to reach a neutral pH. Solid carbonate potassium was added in catalytic quantities to the aqueous phase in order to obtain a neutral pH, then was extracted with dichloromethane (3 X 20 mL). The combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was evaporated under reduced pressure. Rotamers or conformers of the final product was achieved.

Yield: 30%

Analytical data

- R_f: 0.05 (Hexane/Ethyl Acetate (1:1))
- Aspect: brown semisolid



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 2.88 (d, *J* = 6.2 Hz, 1H, H-4) (*Cis*, A); 2.92 (d, *J* = 6.2 Hz, 1H, H-4) (*Cis*, B); 3.38-3.43 (m, 1H, H-5) (*Trans*, A/B); 3.50 (t, *J* = 6.2 Hz,

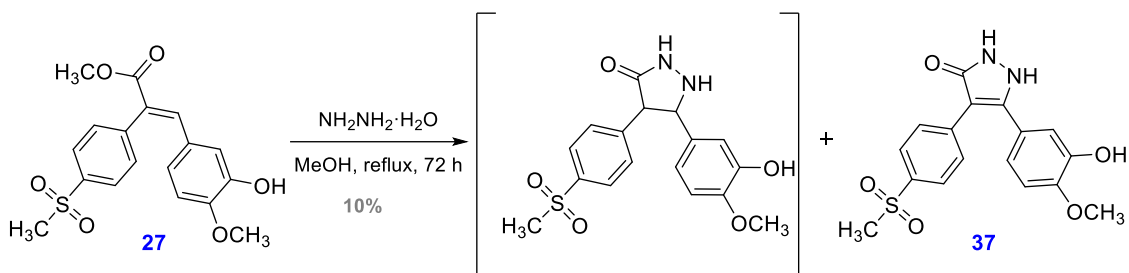
1H, H-5) (*Cis*, A/B); 3.80 (s, 3H, CH₃-O-) (A); 3.84 (s, 3H, CH₃-O-) (B); 3.95 (d, *J* = 11 Hz, 1H, H-4) (*Trans*, A); 4.64 (d, *J* = 11 Hz, 1H, H-4) (*Trans*, B); 6.55 (dd, *J*₁ = 2, *J*₂ = 8 Hz, 1H, H-6') (A); 6.59 (dd, *J*₁ = 2, *J*₂ = 8 Hz, 1H, H-6') (B); 6.63 (d, *J* = 8 Hz, 1H, H-5') (B); 6.67 (d, *J* = 8 Hz, 1H, H-5') (A); 6.71 (d, *J* = 2 Hz, 1H, H-2') (A); 6.78 (bs, 1H, NH); 6.79 (d, *J* = 2 Hz, 1H, H-2') (B); 6.96 (bs, 1H, NH); 7.19-7.25 (m, 3H, H-3'', H-4'', H-5'') (A/B); 7.29 (d, *J* = 5 Hz, 2H, H-2'', H-6'') (A/B); 7.44 (bs, 1H, OH).

A: majority isomer / B: minority isomer

- ¹³C NMR (CDCl₃, 100.6 MHz) δ(ppm), 53.4 (CH, C-5); 54.7 (CH, C-4); 55.7 (CH₃-O-) (A); 55.8 (CH₃-O-) (B); 110.2 (CH, C-2') (A); 110.9 (CH, C-2') (B); 113.3 (CH, C-5') (B); 115.1 (CH, C-5') (A); 119.2 (CH, C-6') (B); 120.4 (CH, C-6') (A); 126.9 (CH, C-4'') (A); 127.3 (CH, C-4'') (B); 127.5 (CH, C-3'', C-5'') (A); 127.9 (CH, C-3'', C-5'') (B); 128.6 (CH, C-2'', C-6'') (A); 128.7 (CH, C-2'', C-6'') (B); 132.4 (C, C-1'') (A); 135.3 (C, C-1'') (B); 137.3 (C, C-1') (B); 139.1 (C, C-1') (A); 145.0 (C, C-4') (A); 145.3 (C, C-3') (A); 145.9 (C, C-4') (B); 147.0 (C, C-3') (B); 173.7 (C, C=O) (A); 176.5 (C, C=O) (B).

A: majority isomer / B: minority isomer

Preparation of 5-(3-hydroxy-4-methoxyphenyl)-4-(4-(methylsulfonyl)phenyl)-1*H*-pyrazol-3(2*H*)-one (37)



Procedure

A 100 mL round bottomed flask equipped with a magnetic stirring bar and a reflux condenser was charged with the starting ester **27** (0.272 g, 0.27 mmol, 1 eq) dissolved in 30 mL of methanol. Then, an aqueous solution of hydrazine (1 mL, *d* = 1.011 g/mL, 35% water, 40.39 mmol) was added. The reaction mixture was stirred at reflux temperature of methanol (80 ± 10 °C) for 72 hours, during this time 1 mL of hydrazine hydrate was added to the reaction mixture.

Work-up

Methanol was removed under reduced pressure. Then, the crude mixture was diluted with water (15 mL) and the organic phase was extracted with diethyl ether/ethyl acetate 1/1 (3 X 20 mL). The combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

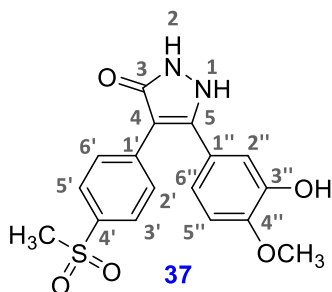
Purification

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. In this case, the product obtained was the oxidized compound, which eluted with a mixture of hexane/ethyl acetate (20:80).

Yield: **10%**

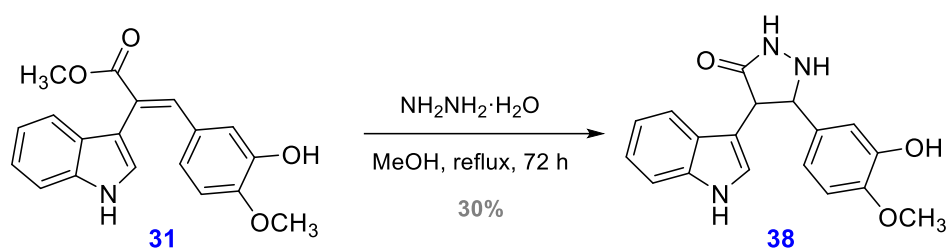
Analytical data

- R_f: 0.08 (Hexane/Ethyl Acetate (1:1))
- Aspect: brown semisolid



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 3.02 (s, 3H, CH₃-S-); 3.86 (s, 3H, CH₃-O-); 6.67 (dd, J₁ = 2, J₂ = 8.4 Hz, 1H, H-6''); 6.76 (d, J = 8.4 Hz, 1H, H-5''); 6.84 (d, J = 2 Hz, 1H, H-2''); 7.36 (d, J = 8.4 Hz, 2H, H-2', H-6'); 7.91 (d, J = 8.4 Hz, 2H, H-3', H-5').
- ¹³C NMR (CDCl₃, 100.6 MHz) δ(ppm), 64.3 (CH₃-S-); 70.7 (CH₃-O-); 96.5 (C, C-5); 110.7 (CH, C-5''); 112.3 (CH, C-2''); 117.9 (CH, C-6''); 127.6 (CH, C-3', C-5'); 127.0 (C, C-1''); 129.2 (CH, C-2', C-6'); 134.3 (C, C-1'); 139.5 (C, C-4'); 145.7 (C, C-3''); 145.9 (C, C-4''); 151.8 (C, C-4); 153.2 (C, C=O).

Preparation of 5-(3-hydroxy-4-methoxyphenyl)-4-(1*H*-indol-3-yl)pyrazolidin-3-one (**38**)



Procedure

A 100 mL round bottomed flask equipped with a magnetic stirring bar and a reflux condenser was charged with the starting ester **31** (0.264 g, 0.85 mmol, 1 eq) dissolved in 30 mL of methanol. Then, an excess of an aqueous solution of hydrazine (1 mL, $d = 1.011$ g/mL, 35% water, 61 mmol) was added. The reaction mixture was stirred at reflux temperature of methanol (80 ± 10 °C) for 72 hours, in two times 1 mL more of hydrazine hydrate was added.

Work-up

Methanol was removed under reduced pressure. Then, the reaction mixture was diluted with water (15 mL) and the organic phase was extracted with ethyl acetate (3 X 20 mL). The combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

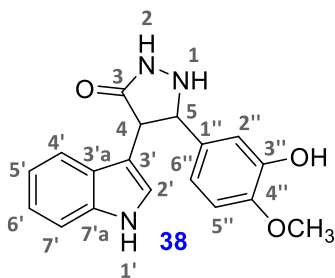
Purification

The crude reaction was purified by column chromatography using a CombiFlash[®] R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (90:10).

Yield: 30%

Analytical data

- R_f: 0.05 (Hexane/Ethyl Acetate (1:1))
- Melting point: 55-57 °C (Ethyl Acetate/Methanol)
- Aspect: brown solid



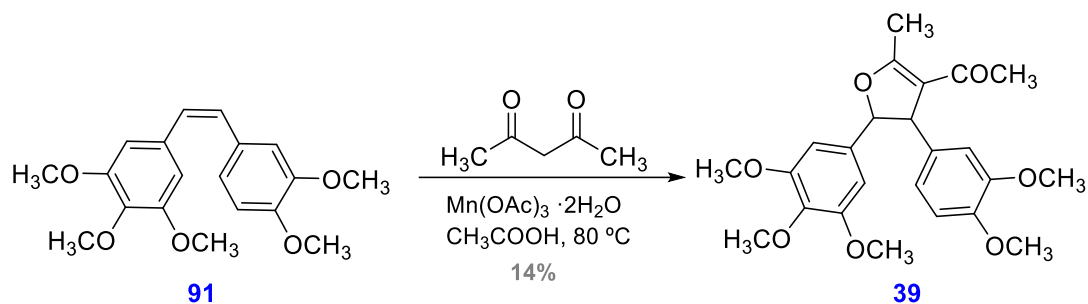
- $^1\text{H NMR}$ ($(\text{CD}_3)_2\text{CO}$, 400 MHz) δ (ppm), 2.96 (dt, $J_1 = 3.5$, $J_2 = 9$ Hz, 1H, H-5) (B); 3.42 (dt, $J_1 = 3.5$, $J_2 = 9$ Hz, 1H, H-5) (A); 3.77 (s, 3H, $\text{CH}_3\text{-O-}$) (A); 3.77 (s, 3H, $\text{CH}_3\text{-O-}$) (B); 4.15 (dd, $J_1 = 3.5$, $J_2 = 5.5$ Hz, 1H, H-4) (B); 5.17 (dd, $J_1 = 3.5$, $J_2 = 5.5$ Hz, 1H, H-4) (A); 6.67 (d, $J_1 = 1.5$, $J_2 = 8.2$ Hz, 1H, H-5''); 6.75 (d, $J = 8.20$ Hz, 1H, H-6''); 6.79 (cs, 1H, H-2''); 6.97-7.11 (m, 2H, H-5', H-6') (A)/(B); 7.24 (d, $J = 2$ Hz, 1H, H-2') (A); 7.29 (d, $J = 2$ Hz, 1H, H-2') (B); 7.35 (d, $J = 8$ Hz, 1H, H-7') (A); 7.38 (d, $J = 8$ Hz, 1H, H-7') (B); 7.77 (d, $J = 8$ Hz, 1H, H-4') (B); 7.81 (d, $J = 8$ Hz, 1H, H-4') (A); 8.87 (bs, 1H, NH); 9.35 (bs, 1H, NH); 10.05 (bs, 1H, OH); 10.15 (bs, 1H, NH-indole).

A: majority isomer / B: minority isomer

- $^{13}\text{C NMR}$ ($(\text{CD}_3)_2\text{CO}$, 100.6 MHz) δ (ppm), 40.3 (CH, C-4) (B); 40.4 (CH, C-4) (A); 44.6 (CH, C-5) (B); 44.7 (CH, C-5) (A); 55.4 ($\text{CH}_3\text{-O-}$) (B); 55.4 ($\text{CH}_3\text{-O-}$) (A); 111.2 (CH, C-7') (B); 111.5 (CH, C-7') (A); 114.2 (C, C-3') (A); 114.3 (C, C-3') (B); 115.9 (CH, C-5''); 118.4 (CH, C-4') (B); 118.6 (CH, C-4') (A); 119.0 (CH, C-2''); 119.5 (CH, C-6''); 119.9 (CH, C-5') (B); 120.0 (CH, C-5) (A); 121.0 (CH, C-2') (B); 121.2 (CH, C-2') (A); 122.8 (CH, C-6') (A); 123.2 (CH, C-6') (B); 126.8 (C, C-3'a) (A); 127.0 (C, C-3'a) (B); 133.5 (C, C-7'a) (A); 133.8 (C, C-7'a) (B); 136.6 (C, C-1'') (B); 136.7 (C, C-1'') (A); 145.7 (C, C-3'') (B); 145.8 (C, C-3'') (A); 148.7 (C, C-4'') (A); 149.0 (C, C-4'') (B); 169.5 (C, C=O) (A); 175.1 (C, C=O) (B).

A: majority isomer / B: minority isomer

Preparation of 4-acetyl-3-(3,4-dimethoxyphenyl)-5-methyl-2-(3,4,5-trimethoxyphenyl)-2,3-dihydrofuran (39)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar and a reflux condenser, previously flame-dried under argon was charged with manganese (III) acetate (0.270 g, 1.008 mmol) in 3 mL of glacial acetic acid ($d = 1.05 \text{ g/mL}$, 52.46 mmol). The reaction mixture was heated at $80 \pm 10 \text{ }^\circ\text{C}$ for 15 minutes. Then, was cooled down to $0 \pm 10 \text{ }^\circ\text{C}$ with an ice-bath and the starting stilbene **91** (0.164 g, 0.49 mmol, 1 eq) was added dropwise dissolved in a mixture of 0.74 mL of acetyl acetone ($d = 0.97 \text{ g/mL}$, 7.21 mmol) and 5 mL of glacial acetic acid. After that, the resulting mixture was heated at $80 \pm 10 \text{ }^\circ\text{C}$ under constant stirring for 24 hours.

Work-up

The crude reaction was diluted with saturated NaHCO_3 (20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the organic phases were washed with water (3 X 20 mL). Finally, the combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was evaporated under reduced pressure.

Purification

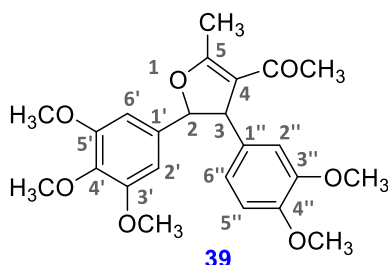
The crude reaction was purified by column chromatography using a CombiFlash[®] R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (50:50). The product obtained was a mixture of the two final products. As two products were obtained a second purification was needed in order to separate both from the mixture.

The resulting mixture was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (60:40). The desired product **39** was separated through purification.

Yield: 14%

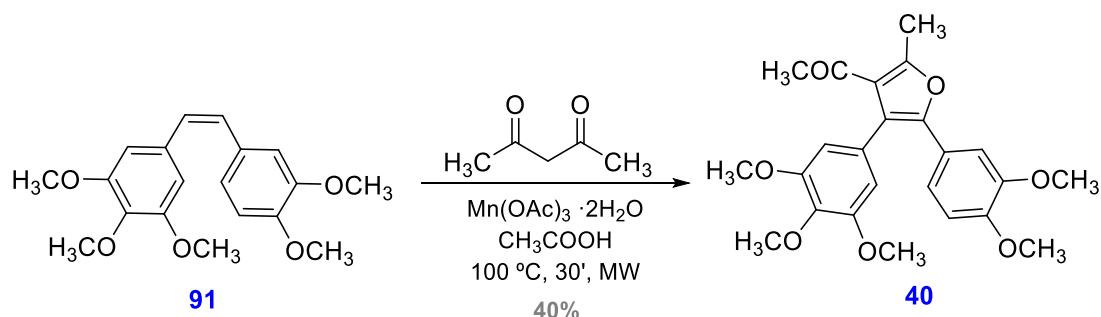
Analytical data

- R_f: 0.3 (Hexane/Ethyl Acetate (3:5))
- Aspect: orange semisolid



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 2.10 (s, 3H, CH₃-C=O); 2.12 (s, 3H, CH₃-Ar); 3.72 (s, 6H, CH₃ (x2)); 3.77 (s, 6H, CH₃-O- (x2)); 3.81 (s, 3H, CH₃-O-); 4.12 (d, *J* = 7 Hz, 1H, H-3); 5.23 (d, *J* = 7 Hz, 1H, H-2); 6.30 (s, 2H, H-2', H-6'); 6.40 (d, *J* = 7 Hz, 1H, H-5''); 6.50 (d, *J* = 2 Hz, 1H, H-2''); 6.78 (dd, *J*₁ = 2, *J*₂ = 7 Hz, 1H, H-6'').
- HRMS ESI (+) *m/z*: calculated mass for C₂₄H₂₈O₇, 429.1835, found 429.1884.

Preparation of 4-acetyl-2-(3,4-dimethoxyphenyl)-5-methyl-3-(3,4,5-trimethoxyphenyl)-2,3-dihydrofuran (**40**)



Procedure

A 10 mL round-bottomed flask specially equipped with a screw cap for microwaves and a magnetic bar under argon was charged with manganese (III) acetate (0.270 g, 1.008

mmol, 1 eq) dissolved in 5 mL of glacial acetic acid ($d = 1,05 \text{ g/mL}$, 87.42 mmol). Then, the starting stilbene **91** (0.073 g, 0.22 mmol, 1 eq) was added dropwise dissolved in 0.74 mL of acetyl acetone ($d = 0.97 \text{ g/mL}$, 7.21 mmol). The resulting mixture was introduced into the microwaves oven (100 kPa) and heated at $100 \pm 10 \text{ }^\circ\text{C}$ for 30 minutes under constant stirring.

Work-up

The crude reaction was diluted with an aqueous saturated solution of NaHCO_3 (20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the organic phases were washed with water (3 X 20 mL). Finally, the combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was evaporated under reduced pressure.

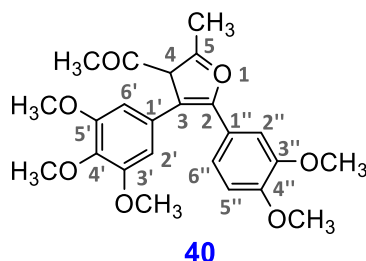
Purification

The crude reaction was purified by column chromatography using a CombiFlash[®] R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (35:65). The desired product was achieved as a mixture of rotamers or conformers.

Yield: 40%

Analytical data

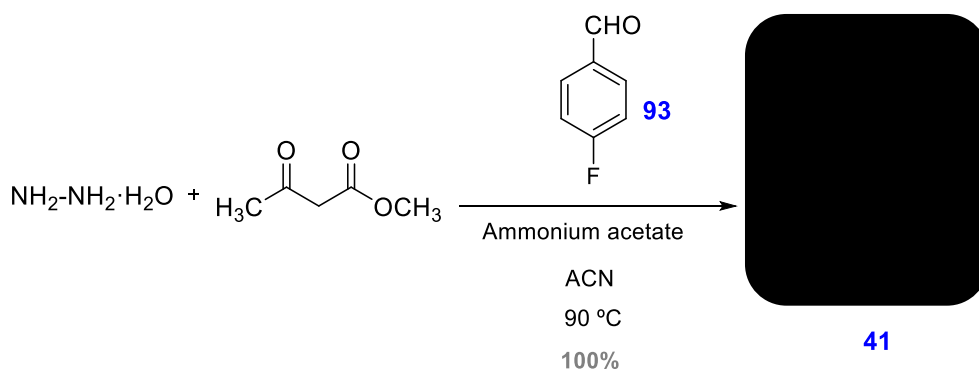
- R_f: 0.46 (Hexane/Ethyl Acetate (4:6))
- Aspect: orange semisolid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 2.17 (s, 3H, $\text{CH}_3\text{-C=O}$); 2.44 (s, 3H, $\text{CH}_3\text{-Ar}$); 3.87 (s, 3H, $\text{CH}_3\text{-O-}$); 3.93 (s, 6H, $\text{CH}_3\text{-O-}$ ($\times 2$)); 3.94 (s, 6H, $\text{CH}_3\text{-O-}$ ($\times 2$)); 6.89 (d, $J = 8$ Hz, 1H, H-5''); 7.02 (s, 2H, H-2', H-6') (A); 7.06 (s, 2H, H-2', H-6') (B); 7.37 (dd, $J_1 = 2$, $J_2 = 8.2$ Hz, 1H, H-6''); 7.42 (d, $J_1 = 2$ Hz, 1H, H-2'').

A: majority isomer / B: minority isomer

Preparation of [REDACTED] **(41)**



Procedure

A 50 mL round-bottomed flask equipped with a magnetic stirring bar and cooled down to 0 ± 10 °C with an ice-bath, was charged with hydrazine hydrate ($d = 1.03$ g/mL, 0.5 mL, 10 mmol, 2 eq), methyl acetoacetate (1.07 mL, 10 mmol, 2 eq), 4-fluorobenzaldehyde (**93**) (0.53 mL, 5 mmol, 1 eq) and finally with ammonium acetate (0.77 g, 10 mmol, 2 eq). All compounds were dissolved in 5 mL of ACN. The resulting mixture was closed with a reflux condenser and the system was heated at 90 ± 10 °C under constant stirring for 24 hours.

Work-up

During the reaction time a spontaneous solid was formed, which was filtered and dried under reduced pressure.

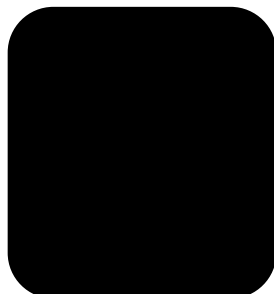
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: **100%**

Analytical data

- R_f: 0.125 (Hexane/Ethyl Acetate (4:6))
- Melting point: 244-246 °C (Methanol)
- Aspect: white solid

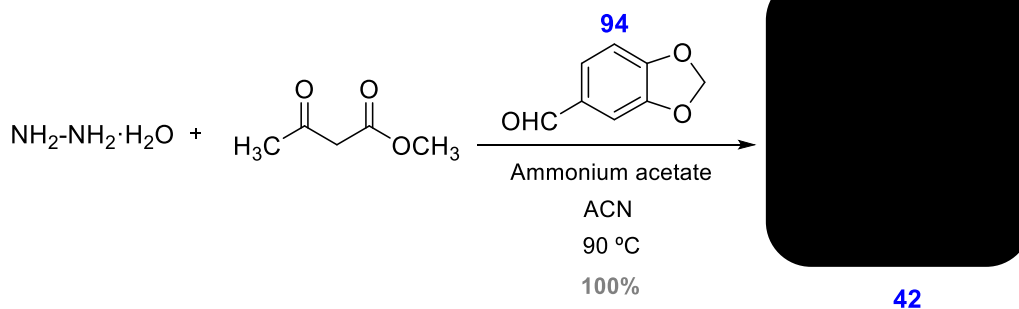


41

- ¹H NMR (DMSO-d₆, 400 MHz) δ(ppm), 2.01 (s, 6H, CH₃- (x2)); 4.72 (s, 1H, Ar-CH-); 6.99 (dt, J₁ = 2, J₂ = 8.8 Hz, 2H, H-2', H-6'); 7.12 (dt, J₁ = 2, J₂ = 8.8 Hz, 2H, H-3', H-5').
- ¹³C NMR (DMSO-d₆, 100.6 MHz) δ(ppm), 10.8 (CH₃-Ar- (x2)); 32.5 (CH, Ar-CH-); 104.6 (C, C-4 (x2)); 114.6 (CH, d, J = 20 Hz, C-3', C-5'); 129.6 (CH, d, J = 7.5 Hz, C-2', C-6'); 139.8 (C, C-3); 140.1 (C, C-3); 159.6 (C, d, J = 242 Hz, C-4'); 161.4 (C, C-5 (x2)).

Preparation of

(42)



Procedure

A 50 mL round-bottomed flask equipped with a magnetic stirring bar and cooled down to 0 ± 10 °C with an ice-bath, was charged with hydrazine hydrate (d = 1.03 g/mL, 0.5 mL, 10 mmol, 2 eq), methyl acetoacetate (1.07 mL, 10 mmol, 2 eq), benzo[d][1,3]dioxole-5-carbaldehyde (**94**) (0.751 g, 5 mmol, 1 eq) and finally with ammonium acetate (0.771 g, 10 mmol, 2 eq). All compounds were dissolved in 5 mL of ACN.

The resulting mixture was closed with a reflux condenser and the system was heated at 90 ± 10 °C under constant stirring for 24 hours.

Work-up

The solvent was evaporated under vacuum. Then, ^1H spectrum of the residue was carried out in order to identify the desired product.

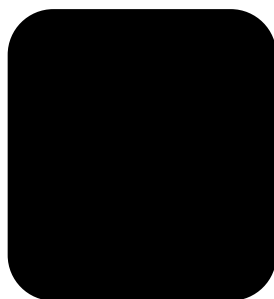
Purification

No purification by column chromatography was performed since the product was obtained with enough purity (**100%**) as shown by TLC and ^1H spectrum.

Yield: **100%**

Analytical data

- R_f : 0.128 (Hexane/Ethyl Acetate (1:1))
- Melting point: 251-253 °C (Methanol)
- Aspect: orange solid

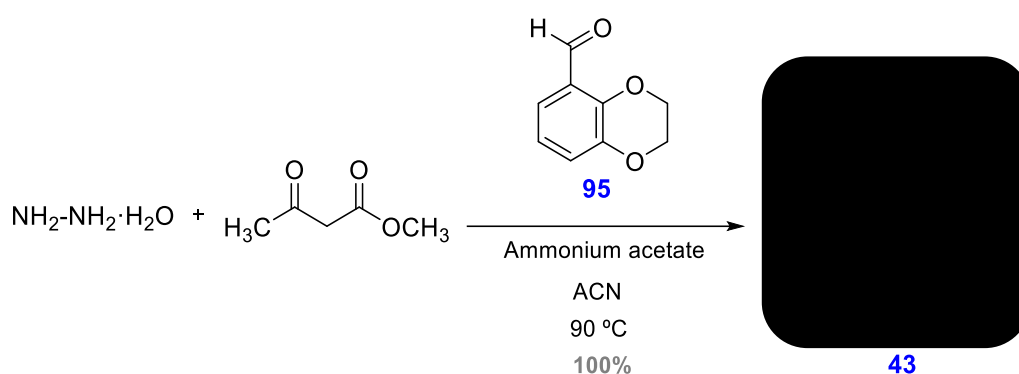


42

- ^1H NMR (DMSO- d_6 , 400 MHz) δ (ppm), 2.04 (s, 6.4, CH_3 (x2)); 4.64 (s, 1H, Ar-CH-); 5.98 (s, 2H, O- CH_2 -O-); 6.60 (d, $J = 7$ Hz, 1H, H-6'); 6.67 (s, 1H, H-4'); 6.70 (d, $J = 7$ Hz, 1H, H-7').
- ^{13}C NMR (DMSO- d_6 , 100.6 MHz) δ (ppm), 10.9 (CH_3 (x2)); 39.2 (CH, Ar-CH-); 100.9 (CH_2 , O- CH_2 -O-); 104.5 (C, C-4 (x2)); 107.9 (CH, C-7'); 108.6 (CH, C-4'); 120.6 (CH, C-6'); 138.5 (C, C-5'); 140.3 (C, C-3 (x2)); 145.3 (C, C-7'a); 147.3 (C, C-3'a); 161.2 (C, C-5); 173.2 (C, C-5 bis).

Preparation of

(43)



Procedure

A 50 mL round-bottomed flask equipped with a magnetic stirring bar and cooled down to $0 \pm 10\text{ }^\circ\text{C}$ with an ice-bath, was charged with hydrazine hydrate ($d = 1.03\text{ g/mL}$, 0.06 mL, 2.43 mmol, 2 eq), methyl acetoacetate (0.14 mL, 2.43 mmol, 2 eq), 2,3-dihydrobenzo[*b*][1,4]dioxin-5-carbaldehyde (**95**) (0.2 g mL, 1.22 mmol, 1 eq) and finally with ammonium acetate (0.096 g, 2.43 mmol, 2 eq). All compounds were dissolved in 5 mL of ACN. The resulting mixture was closed with a reflux condenser and the system was heated at $90 \pm 10\text{ }^\circ\text{C}$ under constant stirring for 24 hours.

Work-up

During the reaction time a spontaneous solid was formed, which was filtered and dried under reduced pressure.

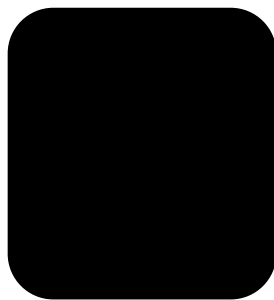
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: 100%

Analytical data

- R_f : 0.04 (Hexane/Ethyl Acetate (5:5))
- Melting point: 205-210 $^\circ\text{C}$ (Methanol)
- Aspect: yellowish solid

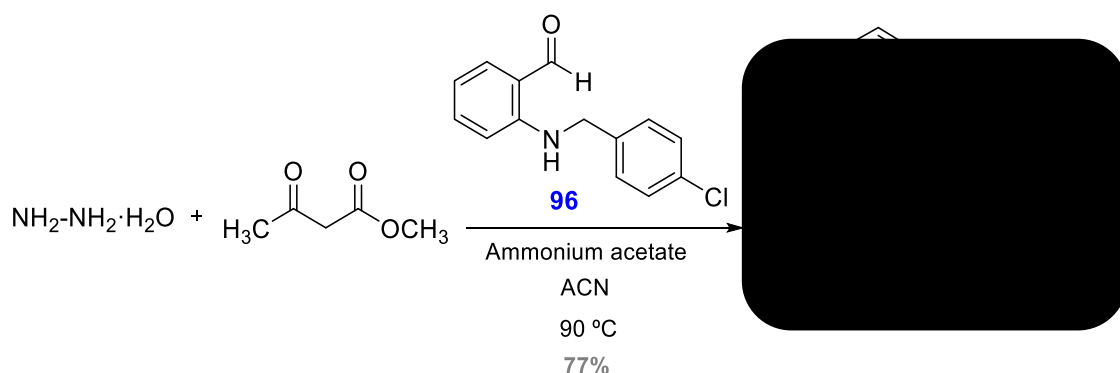


43

- ^1H NMR (DMSO- d_6 , 400 MHz) δ (ppm), 2.00 (s, 6H, CH_3 (x2)); 4.14 (d, $J = 8$ Hz, 4H, $\text{CH}_2\text{-O-}$ x2); 4.99 (s, 1H, CH-); 6.60 - 6.62 (m, 2H, H-7', H-8'); 7.12 (d, $J = 8$ Hz, 1H, H-6').
- ^{13}C NMR (DMSO- d_6 , 100.6 MHz) δ (ppm), 9.4 (CH_3 (x2)); 25.15 (CH, Ar-CH); 62.6-62.9 (CH_2 , $\text{CH}_2\text{-O-}$); 87.8 (CH, C-4' (x2)); 102.6 (C, C-4 (x2)); 113.4 (CH, C-8'); 118.6 (CH, C-7'); 120.2 (CH, C-6'); 131.5 (C, C-3 (x2)); 138.7 (C, C-3 (x2)); 139.2 (C, C-8'a); 141.7 (C, C-4'a); 160.1 (C, C-5 (x2)); 175.5 (C, C=O).

Preparation of

(44)



Procedure

A 50 mL round-bottomed flask equipped with a magnetic stirring bar and cooled down to 0 ± 10 °C with an ice-bath, was charged with hydrazine hydrate ($d = 1.03$ g/mL, 0.06 mL, 1.25 mmol, 2 eq), methyl acetoacetate (0.14 mL, 1.25 mmol, 2 eq), the aldehyde **96** (0.154 g, 0.63 mmol, 1 eq) and finally with ammonium acetate (0.096 g, 1.25 mmol, 2 eq). All compounds were dissolved in 5 mL of ACN. The resulting mixture was closed with reflux condenser and the system was heated at 90 ± 10 °C under constant stirring for 24 hours.

Work-up

During the reaction time a spontaneous solid was formed, which was filtered and dried under reduced pressure.

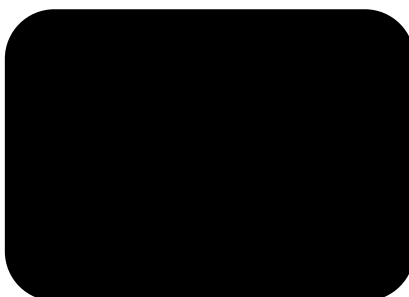
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: **77%**

Analytical data

- R_f : 0.05 (Hexane/Ethyl Acetate (5:5))
- Melting point: 185-188 °C (Methanol)
- Aspect: yellowish solid

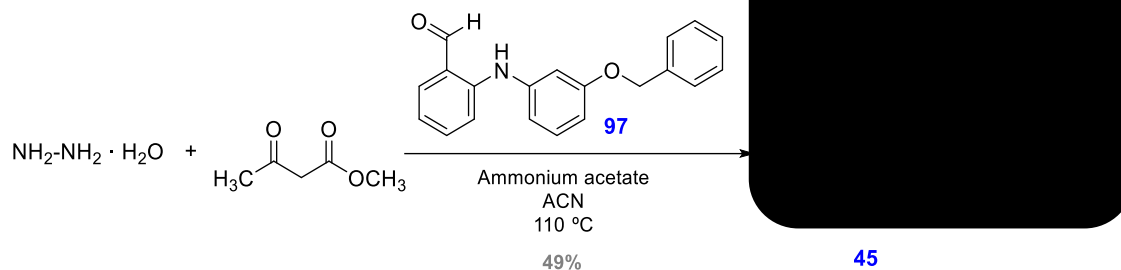


44

- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 1.99 (s, 6H, CH_3 (x2)); 4.32 (s, 1H, Ar-CH-); 4.42 (s, 2H, CH_2 -Ar); 6.55 (d, $J = 8$ Hz, 2H, H-2'', H-6''); 6.60 (d, $J = 8$ Hz, 2H, H-3'', H-5''); 7.33-7.34 (m, 4H, H-3', H-4', H-5', H-6'); 8.51 (s, 1H, NH).

Preparation of

(45)



Procedure

A 50 mL round-bottomed flask equipped with a magnetic stirring bar and cooled down to 0 ± 10 °C with an ice-bath, was charged with hydrazine hydrate ($d = 1.03$ g/mL, 0.048 mL, 0.98 mmol, 2 eq), methyl acetoacetate (0.11 mL, 0.98 mmol, 2 eq), the initial aldehyde **97** (0.150 g, 0.49 mmol, 1 eq) and finally with ammonium acetate (0.076 g, 0.98 mmol, 2 eq). All compounds were dissolved in 5 mL of ACN. The resulting mixture was closed with a reflux condenser and the system was heated at 110 ± 10 °C under constant stirring for 24 hours.

Work-up

The solvent was removed under pressure. Then, the crude product was crystallized with ethyl acetate. The solid obtained was filtered and dried under pressure.

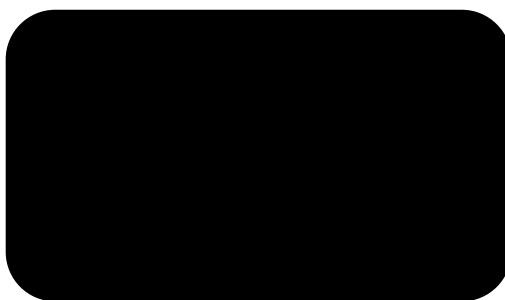
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: 49%

Analytical data

- R_f : 0.08 (Hexane/Ethyl Acetate (3:7))
- Melting point: 150-155 °C (Methanol)
- Aspect: brown solid



45

- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 2.10 (s, 6H, CH_3 (x2)); 4.8 (s, 2H, $\text{CH}_2\text{-O-}$); 5.72 (s, 1H, Ar-CH-); 6.85 (d, $J = 9$ Hz, 2H, H-3', H-6'); 6.85-6.87 (m, 1H, H-4''); 6.92-6.95 (m, 2H, H-4', H-5'); 7.02 (t, $J = 7.6$ Hz, 2H, H-3''', H-5'''); 7.08-7.10 (m, 2H, H-4'', H-5''); 7.16 (d, $J = 7$ Hz, 1H, H-6''); 7.19 (d, $J = 2$ Hz, 1H, H-2''); 7.54 (d, $J = 7.7$ Hz, 1H, H-2''')*; 7.60 (d, $J = 7.7$ Hz, 1H, H-6''')*; 8.50 (bs, 2H, OH).

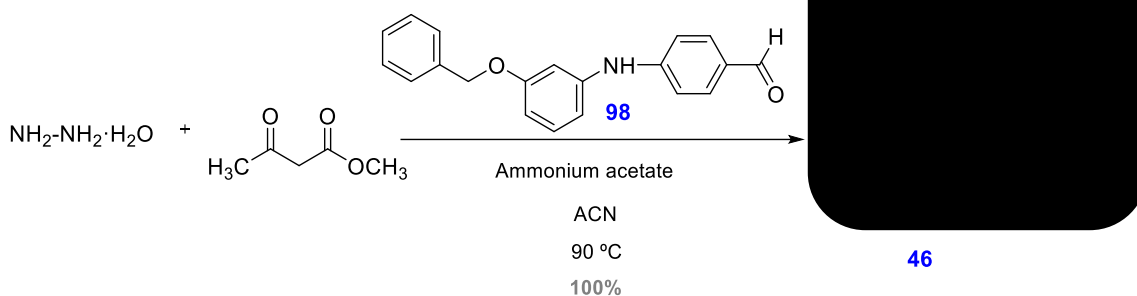
*Interchangeable protons

- ^{13}C NMR (CDCl_3 , 100.6 MHz) δ (ppm), 10.2 (CH_3 (x2)); 30.0 (CH, Ar-CH-); 52.4 (CH_2 , $\text{CH}_2\text{-O-}$); 101.3 (C, C-4 (x2)); 108.4 (CH, C-2''); 115.0 (CH, C-6''); 115.2 (CH, C-6'); 115.4 (CH, C-5'); 124.1 (CH, C-3'); 124.3 (CH, C-4'); 125.3 (CH, C-4''); 127.5 (C, C-1'); 128.0 (CH, C-2''', C-6'''); 128.1 (CH, C-4'''); 128.2 (CH, C-3''', C-5'''); 129.0 (CH, C-5''); 136.8 (C, C-1'''); 140.1 (C, C-3 (x2)); 151.4 (C, C-1''); 158.4 (C, C-2'); 162.0 (C, C-3''); 168.4 (C, C-5); 168.7 (C, C-5).
- HRMS ESI (+) m/z : calculated mass for $\text{C}_{28}\text{H}_{28}\text{N}_5\text{O}_3$ 482.2114, found: 482.2130.

Preparation of

(46)

Procedure



A 50 mL round-bottomed flask equipped with a magnetic stirring bar and cooled down to 0 ± 10 °C with an ice-bath, was charged with hydrazine hydrate ($d = 1.03$ g/mL, 0.058 mL, 1.19 mmol, 2 eq), methyl acetoacetate (0.128 mL, 1.19 mmol, 2 eq), 4-((3-(benzyloxy)phenyl)amino)benzaldehyde (**98**) (0.181 g, 0.597 mmol, 1 eq) and finally with ammonium acetate (0.092 g, 1.19 mmol, 2 eq). All components were dissolved in 5 mL of ACN. The resulting mixture was closed with a reflux condenser and the system was heated at 90 ± 10 °C under constant stirring for 24 hours.

Work-up

The solvent was evaporated under vacuum. Then, ^1H spectrum of the obtained residue was carried out in order to identify the desired product.

Purification

No purification by column chromatography was performed since the product was obtained with enough purity (**100%**) as shown by TLC and ^1H spectrum.

Yield: **100%**

Analytical data

- R_f : 0.11 (Hexane/Ethyl Acetate (1:1))
- Melting point: 180-183 °C (Methanol)
- Aspect: orange-red solid

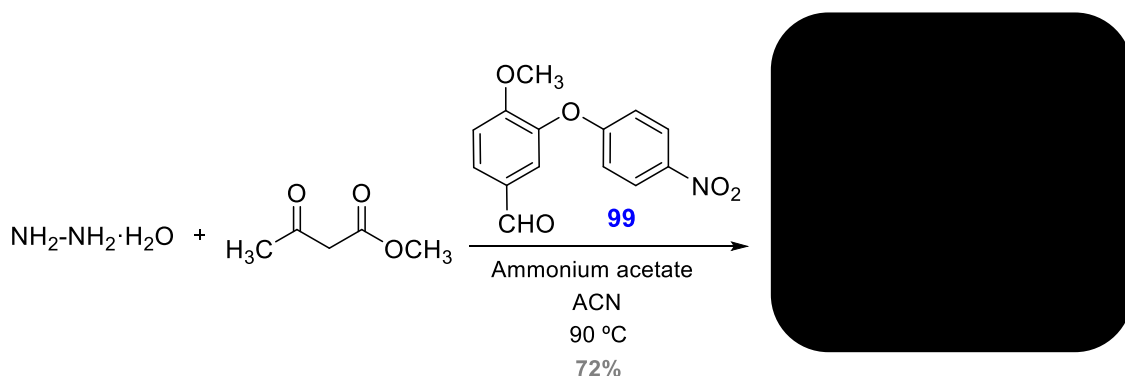


46

- ^1H NMR (DMSO- d_6 , 400 MHz) δ (ppm), 2.13 (s, 6H, CH_3 (x2)); 5.01 (s, 1H, CH-); 5.19 (s, 2H, $\text{CH}_2\text{-O-}$); 6.35 (dd, $J_1 = 2$, $J_2 = 8$ Hz, 1H, H-4''); 6.61 (d, $J = 8$, 1H, H-6''); 6.65 (d, $J = 2$ Hz; H-2''); 6.94 (d, $J = 8.5$ Hz, 2H, H-2', H-6'); 7.04 (d, $J = 8.5$ Hz, 2H, H-3', H-5'); 7.02 (t, $J = 8$ Hz, 1H, H-5''); 7.26 (t, $J = 7.12$ Hz, 1H, H-4'''); 7.35 (t, $J = 7.12$ Hz, 2H, H-3''', H-5'''); 7.43 (d, $J = 7.12$ Hz, 2H, H-2''', H-6'''); 7.98 (bs, 1H, NH); 8.3 (bs, 2H, OH).
- ^{13}C NMR (DMSO- d_6 , 100.6 MHz) δ (ppm), 11.6 (CH_3 (x2)); 32.6 (CH, Ar-CH-); 69.4 ($\text{CH}_2\text{-O-}$); 89.3 (CH, C-3', C-5'); 102.3 (CH, C-2''); 104.8 (CH, C-4''); 105.7 (C, C-4 (x2)); 109.0 (CH, C-6''); 117.9 (CH, C-4'''); 127.9 (CH, C-2''', C-6'''); 128.1 (CH, C-3''', C-5'''); 128.8 (CH, C-2', C-6'); 130.2 (CH, C-5''); 135.7 (C, C-1'); 137.7 (C, C-1'''); 139.9 (C, C-3 (x2)); 140.1 (C, C-4'); 159.6 (C, C-1''); 161.5 (C, C-3''); 172.7 (C, C-5 (x2)).

Preparation of

(47)



Procedure

A 50 mL round-bottomed flask equipped with a magnetic stirring bar and cooled down to 0 ± 10 °C with an ice-bath, was charged with hydrazine hydrate ($d = 1.03$ g/mL, 0.23 mL, 4.81 mmol, 2 eq), methyl acetoacetate (0.52 mL, 4.81 mmol, 2 eq), 4-methoxy-3-(4-nitrophenoxy)benzaldehyde (**99**) (0.785 mL, 2.40 mmol, 1 eq) and finally with ammonium acetate (0.37 g, 4.81 mmol, 2 eq). All components were dissolved in 5 mL of ACN. The resulting mixture was closed with a reflux condenser and the system was heated at 90 ± 10 °C under stirring for 24 hours.

Work-up

The solvent was evaporated under vacuum. Then, ^1H spectrum was carried out in order to identify the desired product.

Purification

The crude reaction was purified by column chromatography on silica gel eluting with mixtures of hexane/ethyl acetate and methanol with increasing polarity. The desired product eluted with a polarity of ethyl acetate/methanol (80:20).

Yield: 72%

Analytical data

- R_f : 0.78 (Methanol)
- Melting point: 201-203 °C (methanol)

- Aspect: brown-orange solid

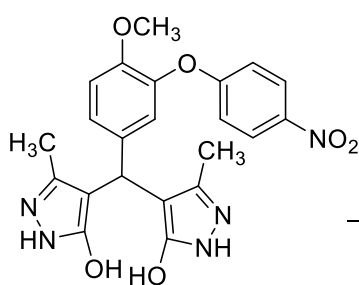


47

- ^1H NMR (DMSO- d_6 , 400 MHz) δ (ppm), 2.1 (s, 6H, CH_3 (x2)); 3.67 (s, 3H, $\text{CH}_3\text{-O-}$); 4.48 (s, 1H, CH-Ar); 6.89 (d, $J = 2$ Hz, 1H, H-2'); 6.93 (d, $J = 9$ Hz, 2H, H-2'', H-6''); 7.01 (d, $J = 8.6$ Hz, 1H, H-5'); 7.03 (dd, $J = 2$, $J = 8.6$ Hz, 1H, H-6'); 8.18 (d, $J = 9$ Hz, 2H, H-3'', H-5'').
- ^{13}C NMR (DMSO- d_6 , 100.6 MHz) δ (ppm), 10.7 (CH_3 (x2)); 32.4 (CH, Ar-CH-); 56.2 ($\text{CH}_3\text{-O-}$); 104.5 (C, C-4 (x2)); 113.4 (CH, C-5'); 115.9 (CH, C-2'', C-6''); 122.0 (CH, C-6'); 126.2 (CH, C-2'); 126.4 (CH, C-3'', C-5''); 137.2 (C, C-3 (x2)); 141.2 (C, C-4''); 142.1 (C, C-3'); 149.4 (C, C-4'); 162.9 (C, C-5); 163.6 (C, C-1'').
- HMRS ESI (+) m/z : calculated mass for $\text{C}_{22}\text{H}_{22}\text{N}_5\text{O}_6$ 452.1496, found 452.1500.

Preparation of

(48)



47

H_2 , Pd/C (10%)

HCl (2N)
MeOH/AcOEt, r.t.

100%



Procedure

A 50 mL round-bottomed flask for catalytic hydrogenations was charged with compound **47** (0.091 g, 0.20 mmol, 1 eq) dissolved in a mixture of 10 mL of methanol and 5 mL of ethyl acetate. The catalyst was then added Pd-C 10% (10% p/p). Then, three drops of hydrochloric acid (2N) were added to the solution in order to catalyze the reaction. The reaction mixture was stirred at room temperature for 8 days.

Work-up

The theoretical volume required for the reaction was 9 mL. Given that the hydrogenation apparatus was not hermetical and may leak, the consumed volume was 145 mL, higher than expected.

The crude reaction was filtered by means of a pleated filter and washed with 20 mL of methanol and was then collected in a round-bottomed flask of 100 mL capacity. Finally, methanol was removed under reduced pressure.

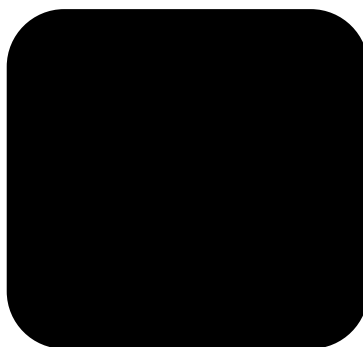
Purification

In this case no purification by column chromatography was performed since as shown in the ^1H spectrum the product was pure.

Yield: **100%**

Analytical data

- R_f : 0,05 (Ethyl Acetate/Methanol (8:2))
- Melting point: 160-162 °C (Methanol)
- Aspect: brown solid



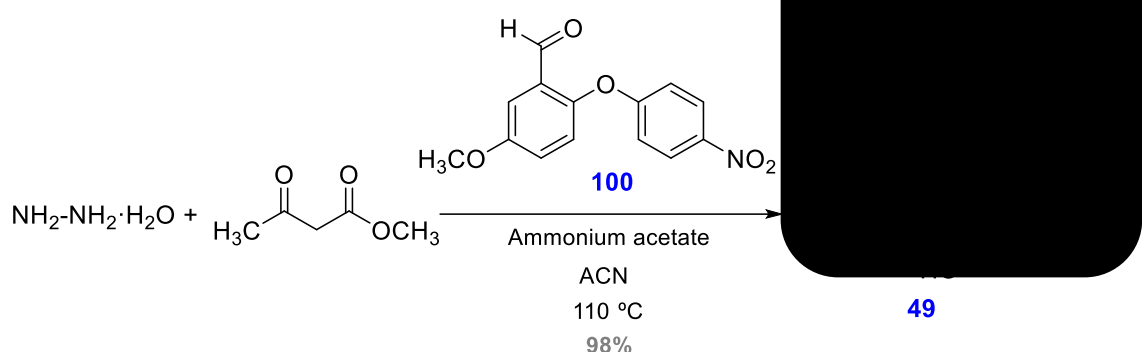
48

- ^1H NMR (DMSO- d_6 , 400 MHz) δ (ppm), 1.99 (s, 6H, CH_3 (x2)); 4.90 (s, 1H, Ar-CH-); 6.30 (dd, $J_1 = 2$, $J_2 = 8.8$ Hz, 1H, H-6'); 6.83 (d, $J = 8.9$ Hz, 2H, H-3'', H-5''); 6.97 (d, $J = 2$ Hz, 1H, H-2'); 7.06 (d, $J = 8.8$ Hz, 1H, H-5'); 7.28 (d, $J = 8.9$ Hz, H-2'', H-6'').

Preparation of [REDACTED]

(49)

Procedure



A 50 mL round-bottomed flask equipped with a magnetic stirring bar and cooled down to 0 ± 10 °C with an ice-bath, was charged with hydrazine hydrate ($d = 1.03$ g/mL, 0.036 mL, 0.72 mmol, 2 eq), methyl acetoacetate (0.079 mL, 0.72 mmol, 2 eq), the starting aldehyde **100** (0.100 g, 0.36 mmol, 1 eq) and finally with ammonium acetate (0.056 g, 0.72 mmol, 2 eq). All components were dissolved in 5 mL of ACN. The resulting mixture was closed with a reflux condenser and the system was heated at 110 ± 10 °C under stirring for 24 hours.

Work-up

The solvent was evaporated under vacuum. Then, ^1H spectrum of the obtained residue was carried out in order to identify the expected product.

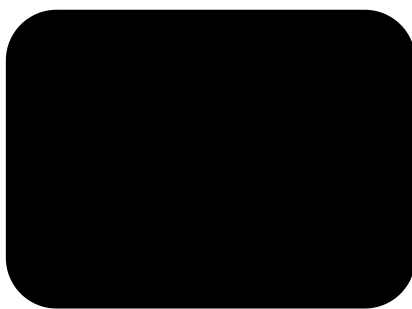
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: 98%

Analytical data

- R_f : 0.04 (Hexane/Ethyl Acetate (3:7))
- Melting point: 200-205 °C (Methanol)
- Aspect: yellow solid



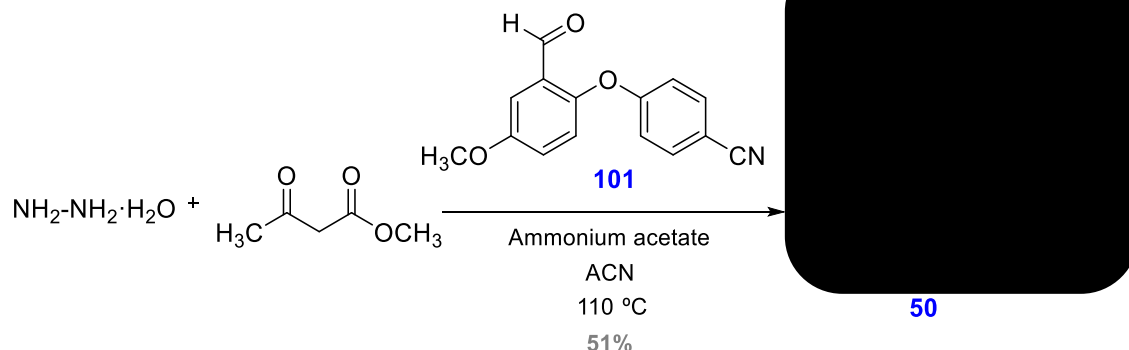
49

- ^1H NMR (DMSO- d_6 , 400 MHz) δ (ppm), 1.92 (s, 6H, CH_3 (x2)); 3.70 (s, 3H, $\text{CH}_3\text{-O-}$); 4.80 (s, 1H, Ar-CH-); 6.79-6.81 (m, 1H, H-4'); 6.82 (d, $J = 9.2$ Hz, 2H, H-2'', H-6''); 6.89 (d, $J = 8.7$ Hz, 1H, H-6'); 7.21 (d, $J = 3$ Hz, 1H, H-3'); 8.11 (d, $J = 9.2$ Hz, 2H, H-3'', H-5'').
- HRMS ESI (+) m/z : calculated mass for $\text{C}_{22}\text{H}_{22}\text{N}_5\text{O}_6$ 452.1492, found 452.1623.

Preparation of

(50)

Procedure



A 50 mL round-bottomed flask equipped with a magnetic stirring bar and cooled down to 0 ± 10 °C with an ice-bath, was charged with hydrazine hydrate ($d = 1.03$ g/mL, 0.014 mL, 0.292 mmol, 2 eq), methyl acetoacetate (0.03 mL, 0.292 mmol, 2 eq), the initial aldehyde **101** (0.063 g, 0.146 mmol, 1 eq) and finally with ammonium acetate (0.022 g, 1.25 mmol, 2 eq). All components were dissolved in 5 mL of ACN. The resulting mixture was closed with a reflux condenser and the system was heated at 90 ± 10 °C under constant stirring for 24 hours.

Work-up

During the reaction a spontaneous solid was obtained, which was filtered and dried under reduced pressure.

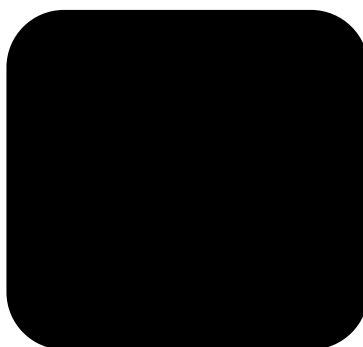
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: **51%**

Analytical data

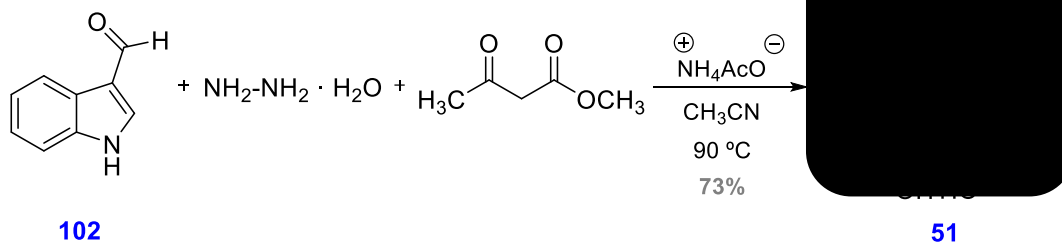
- R_f : 0.42 (Hexane/Ethyl Acetate (8:2))
- Melting point: 203-206 °C (Ethyl Acetate)
- Aspect: brown solid



50

- ^1H NMR (DMSO- d_6 , 400 MHz) δ (ppm), 2.01 (s, 6H, CH_3 - (x2)); 5.20 (s, 1H, CH-); 5.97 (bs, 1H, H-6'); 6.57 (t, $J = 8$ Hz, 1H, H-4'); 6.75 (d, $J = 8.5$ Hz, 2H, H-2'', H-6''); 7.27 (d, $J = 8$ Hz, 1H, H-3'); 7.42 (d, $J = 8.5$ Hz, 2H, H-3'', H-5'').
- HRMS ESI (+) m/z : calculated mass for $\text{C}_{23}\text{H}_{22}\text{N}_5\text{O}_4$ 432.1594, found 432.1599.

Preparation of XXXXXXXXXX (51)



Procedure

A 50 mL round-bottomed flask with a magnetic stirring bar and cooled down to 0 ± 10 °C with an ice-bath, was charged with hydrazine hydrate (d = 1.03 g/mL, 0.43 mL, 5.52 mmol, 2 eq), methyl acetoacetate (0.60 mL, 5.52 mmol, 2 eq), the initial aldehyde **102** (0.400 g, 2.76 mmol, 1 eq) and finally with ammonium acetate (0.400 g, 5.52 mmol, 2 eq). All compounds were dissolved in 5 mL of ACN. The resulting mixture was closed with a reflux condenser and the system was heated at 90 ± 10 °C under constant stirring for 24 hours.

Work-up

During the reaction a spontaneous solid was formed, which was filtered and dried under reduced pressure.

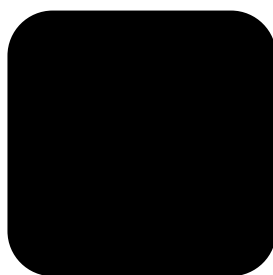
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: 73%

Analytical data

- R_f : 0.02 (Hexane/Ethyl Acetate (1:1))
- Melting point: 225-230 °C (Methanol)
- Aspect: orange solid



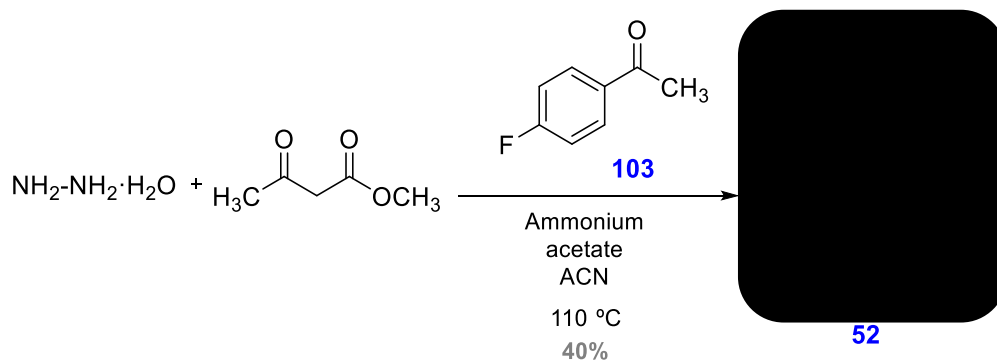
51

- ^1H NMR (DMSO- d_6 , 400 MHz) δ (ppm), 2.12 (s, 3H, CH_3); 2.23 (s, 3H, CH_3 -); 5.01 (s, 1H, CH-); 6.95 (t, $J = 7.5$ Hz, 1H, H-5'); 7.25 (t, $J = 7.5$ Hz, 1H, H-6'); 7.28 (s, 1H, H-2'); 7.53 (d, $J = 7.5$ Hz, 1H, H-7'); 7.89 (bs, 1H, NH); 8.07 (d, $J = 7.5$ Hz, 1H, H-4').

Preparation of

(52)

Procedure



A 50 mL round-bottomed flask equipped with a magnetic stirring bar and cooled down to 0 ± 10 °C with an ice-bath, was charged with hydrazine hydrate ($d = 1.03$ g/mL, 0.16 mL, 3.34 mmol, 2 eq), methyl acetoacetate (0.36 mL, 3.34 mmol, 2 eq), the initial ketone **103** (0.200 g, 1.67 mmol, 1 eq) and finally with ammonium acetate (0.250 g, 3.34 mmol, 2 eq). All components were dissolved in 5 mL of ACN. The resulting mixture was closed with a reflux condenser and the system was heated at 90 ± 10 °C under constant stirring for 24 hours.

Work-up

During the reaction time a spontaneous solid was formed, which was filtered and dried under reduced pressure.

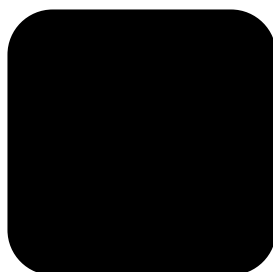
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: **40%**

Analytical data

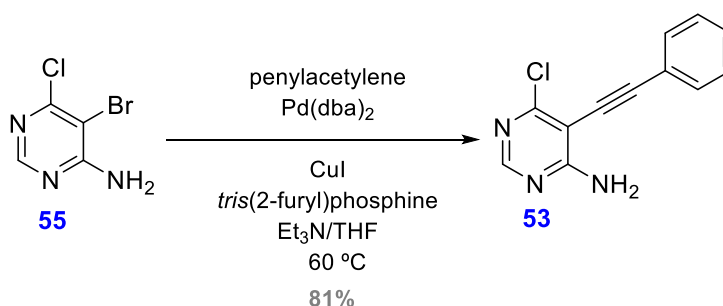
- R_f : 0.01 (Hexane/Ethyl Acetate (8:2))
- Melting point: 192-195 °C (Methanol)
- Aspect: brown-yellowish solid



52

- ^1H NMR (DMSO- d_6 , 400 MHz) δ (ppm), 2.20 (s, 3H, CH_3); 2.52 (s, 6H, CH_3 (x2)); 7.11 (t, $J = 8.5$ Hz, 1H, H-3'); 7.31 (t, $J = 8.5$ Hz, 1H, H-5'); 7.33 (t, $J = 7.7$ Hz, 1H, H-2'); 8.02 (t, $J = 7.7$ Hz, 1H, H-6').

Preparation of 4-amine-6-chloro-5-(phenylethynyl)pyrimidine (**53**)



Procedure

A 50 mL round-bottomed flask specially equipped with a screw cap for coupling reactions, previously flame-dried under argon, was charged with the starting material **55** (0.331 g, 1.58 mmol, 1 eq) dissolved in 8.47 mL of THF (0.15M).

Then, phenylacetylene (0.177 mL, $d = 0.930 \text{ g/mL}$, 1.62 mmol, 1 eq), 1.58 mL of triethylamine ($d = 0.726 \text{ g/mL}$, 11.33 mmol, 10 eq), CuI, Pd(dba)₂ and tris(2-furyl)phosphine were added in catalytic quantities. The reaction mixture was heated at $60 \pm 10 \text{ }^\circ\text{C}$ under stirring for 48 hours, during this time 0.6 mL more of triethylamine was added in order to promote the formation of the final compound.

Work-up

Ammonium chloride was added to the reaction mixture. Then, the crude product was diluted with water (3 X 20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was evaporated under reduced pressure.

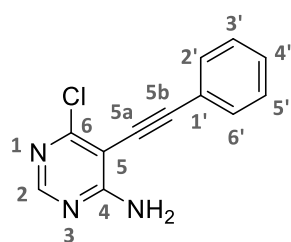
Purification

No purification by column chromatography was performed since the desired product were obtained with enough purity as shown in the ¹H spectrum.

Yield: **81%**

Analytical data

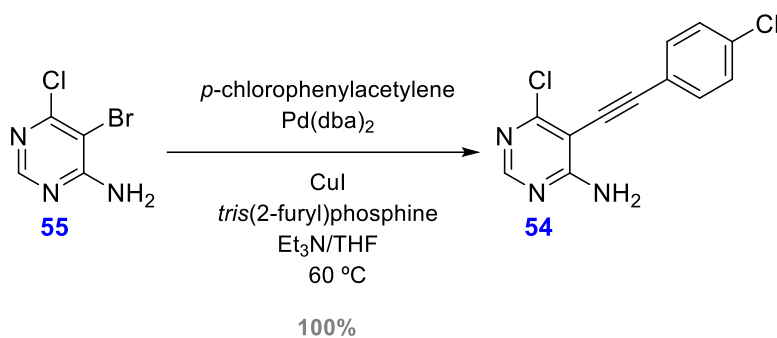
- R_f: 0.21 (Hexane/Ethyl Acetate (7:3))
- Aspect: brown-yellow semisolid



53

- ¹H NMR (CDCl₃, 400 MHz) δ (ppm), 7.08 (bs, 2H, NH₂); 7.17-7.29 (m, 3H, H-3', H-4', H-5'); 7.43 (d, $J = 7.2 \text{ Hz}$, 2H, H-2', H-6'); 8.14 (s, 1H, H-2).

Preparation of 4-amino-6-chloro-5-(4-chlorophenyl)ethynylpyrimidine (**54**)



Procedure

A 50 mL round-bottomed flask specially equipped with a screw cap for coupling reactions, previously flame-dried under argon, was charged with the starting bromopyrimidine **55** (0.260 g, 1.25 mmol, 1 eq) dissolved in 8.47 mL of THF (0.15M). Then, *p*-chlorophenylacetylene (0.177 mL, *d* = 0.930 g/mL, 1.20 mmol, 1 eq), 1.58 mL of triethylamine (*d* = 0.726 g/mL, 11.33 mmol, 10 eq), CuI, Pd(dba)₂ and tris(2-furyl)phosphine were added in catalytic quantities. The reaction mixture was heated at 60 ± 10 °C under stirring for 48 hours, during which 0.6 mL more of triethylamine was added in order to promote the formation of the final compound.

Work-up

Ammonium chloride was added to the reaction mixture. Then, the crude reaction was diluted with water (3 X 20 mL) and extracted with ethyl acetate (3 X 20 mL). Finally, the combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was evaporated under reduced pressure.

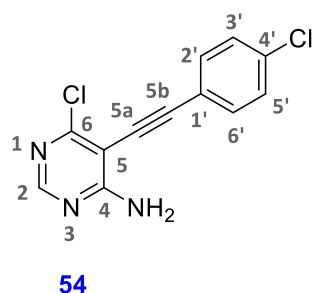
Purification

No purification by column chromatography was performed since the desired product was obtained with enough purity as shown in the ¹H spectrum.

Yield: 100%

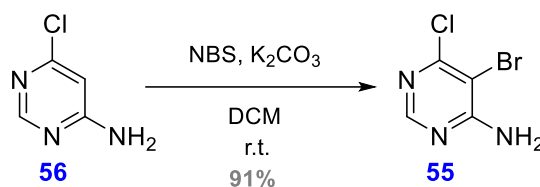
Analytical data

- R_f: 0.21 (Hexane/Ethyl Acetate (7:3))
- Aspect: brown semisolid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 5.58 (bs, 2H, NH_2); 7.30 (d, $J = 8.8$ Hz, 2H, H-3', H-5'); 7.45 (d, $J = 8.8$ Hz, 2H, H-2', H-6'); 8.45 (s, 1H, H-2).

Preparation of 4-amino-5-bromo-6-chloropyrimidine (**55**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon, was charged with the starting amine **56** (0.231 g, 1.78 mmol, 1 eq) dissolved in 15 mL of dichloromethane. Then, potassium carbonate (0.246 g, 1.78 mmol, 1 eq) and NBS (0.545 g, 1.96 mmol, 1.1 eq) were added to the reaction mixture. The resulting mixture was stirred constantly at room temperature for 24 hours.

Work-up

The crude reaction was diluted with a solution of sodium hydroxide (2N) (3 X 20 mL) and extracted with dichloromethane (3 X 20 mL). Then, the combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

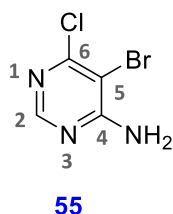
Purification

The crude product was used in the next step without further purification since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: 91%

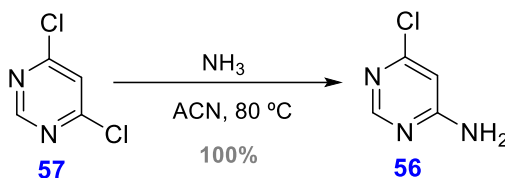
Analytical data

- R_f: 0.46 (Hexane/Ethyl Acetate (7:3))
- Melting point: 140-142 °C (Ethyl Acetate)
- Aspect: yellowish solid



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 6.93 (bs, 2H, NH₂); 8.17 (s, 1H, H-2).

Preparation of 4-amino-6-chloropyrimidine (56)



Procedure

A 50 mL round-bottomed flask specially equipped with a screw cap previously flamed-dried under argon, was charged with the starting material **57** (0.4 g, 2.68 mmol, 1 eq) dissolved in 10 mL of ACN. Then, 2.5 mL of ammonia (d = 0.73 g/mL, 107.15 mmol) was added to the reaction mixture. The resulting mixture was constantly stirred and heated at 80 ± 10 °C for 24 hours.

Work-up

The solvent and the ammonia were evaporated under reduced pressure. Then, ¹H spectrum was carried out.

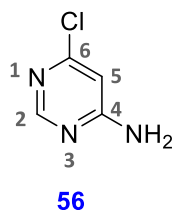
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown in the ¹H spectrum.

Yield: 100%

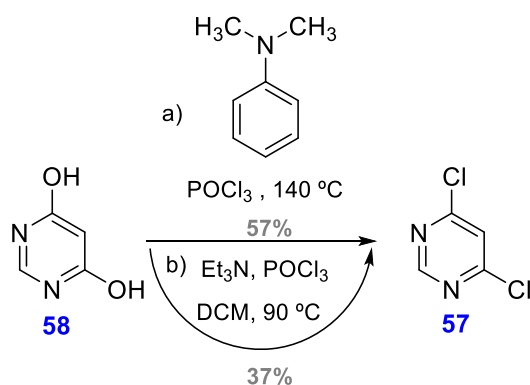
Analytical data

- R_f: 0.05 (Hexane/Ethyl Acetate (7:3))
- Melting point: 212-214 °C (Ethyl Acetate)
- Aspect: white solid



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 5.21 (bs, 2H, NH₂); 7.44 (s, 1H, H-5); 8.81 (s, 1H, H-2).

Preparation of 4,6-dichloropyrimidine (**57**)



Procedure A

A 100 mL round-bottomed flask with a stirring bar, equipped with a reflux condenser and cooled down with an ice-bath, was charged dropwise with 23.3 mL phosphoryl chloride (d = 1.64 g/mL, 92.65 mmol, 5 eq). Then, 4,6-dihydroxypyrimidine (**58**) (2 g, 17.84 mmol, 1 eq) was added dissolved in 3.65 mL of DMA (d = 0.956 g/mL, 31.50 mmol, 2 eq). The mixture was stirred under these conditions for 10 minutes. After that time, the resulting mixture was stirred at reflux temperature, 140 ± 10 °C for 2.5 hours.

Work-up

Phosphoryl chloride was evaporated under reduced pressure. The crude reaction was diluted with water (3 X 20 mL) and extracted with ethyl acetate (3 X 20 mL). The combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the

solvent was removed under reduced pressure. Finally, the crude obtained was microdistilled at 190 ± 10 °C with the aid of vacuum.

Purification

The final compound was not purified. In this case the yield was low due to the sublimation of product **57**.

Yield: 57%

Procedure B

A 50 mL round-bottomed flask with a stirring bar, equipped with a reflux condenser and cooled down with an ice-bath, was charged with the starting material **58** (2 g, 17.84 mmol, 1 eq) dissolved in 10 mL of DCM. Then trimethylamine (2.5 mL, $d = 0.726$ g/mL, 17.93 mmol, 1 eq) and 1.65 mL of phosphoryl chloride ($d = 1.64$ g/mL, 17.75 mmol, 1 eq) were added dropwise to the reaction mixture. The mixture was stirred at room temperature for 2 hours. After that time, the resulting mixture was stirred and heated at 90 ± 10 °C for 1 hour.

Work-up

Triethylamine, phosphoryl chloride and DCM were evaporated under reduced pressure with the aid of toluene and ^1H spectrum was carried out.

Purification

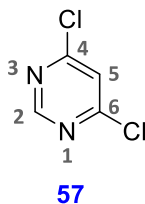
The final compound was not purified since the product showed enough purity. In this case the yield was low due to the sublimation of product **57**.

Yield: 37%

Analytical data

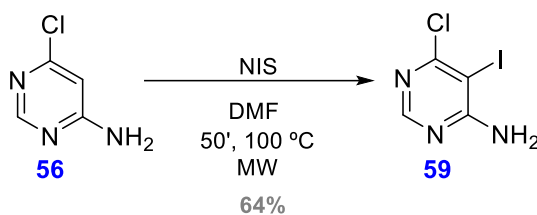
- R_f : 0.77 (Hexane/Ethyl Acetate (7:3))
- Melting point: 53-55 °C (Dichloromethane)

- Aspect: white needles



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 7.44 (s, 1H, H-5); 8.81 (s, 1H, H-2).

Preparation of 4-amine-6-chloro-5-iodopyrimidine (**59**)



Procedure

A 10 mL round-bottomed flask specially equipped with a screw cap was charged with the starting amine **56** (0.533 g, 4.12 mmol, 1 eq) dissolved in 3 mL of DMF. Then, NIS (1.85 g, 8.24 mmol) was added. The reaction mixture was introduced into the microwaves oven (100 kPa) and heated at 100 ± 10 °C under constant stirring for 50 minutes.

Work-up

After cooling, the solvent was removed. The crude reaction was diluted with water (3 X 20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the organic phase was washed with a solution of saturated tiosulfate (2 X 20 mL) and NaOH 10% (2 X 20 mL). Finally, the combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

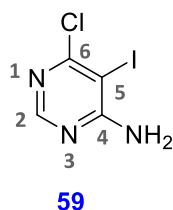
Purification

The desired compound was obtained with enough purity as shown in the ^1H spectrum.

Yield: **64%**

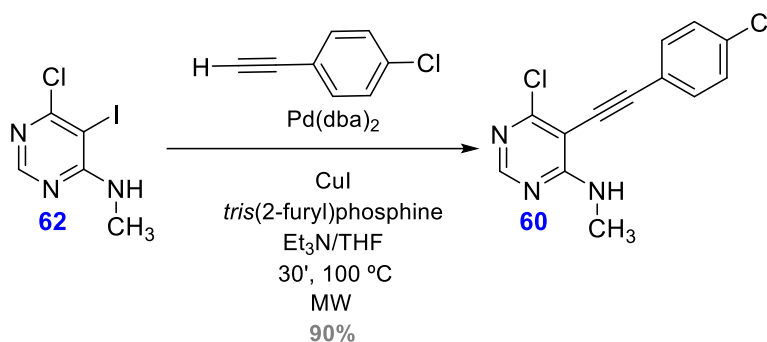
Analytical data

- R_f: 0.5 (Hexane/Ethyl acetate (7:3))
- Melting point: 189-191 °C (Ethyl Acetate)
- Aspect: white solid



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 5.73 (s, 2H, NH₂); 8.22 (s, 1H, H-2).
- ¹³C NMR (CDCl₃, 100.6 MHz) δ(ppm), 77.7 (C, C-5); 157.4 (CH, C-2); 163.8 (C, C-6); 164.3 (C, C-4).
- HRMS ESI (+) m/z: calculated mass for C₄H₄N₃Cl 255.9132, found 255.9129.

Preparation of 4-amino-6-chloro-5-((4-chlorophenyl)ethynyl)-*N*-methylpyrimidine (**60**)



Procedure

A 10 mL round-bottomed flask specially equipped with a screw cap for microwaves under argon, was charged with the starting material **62** (0.300 g, 1.11 mmol, 1 eq) dissolved in 1 mL of THF. Then, *p*-chlorophenylacetylene (0.29 mL, d = 0.930 g/mL, 2.22 mmol, 2 eq), 3.5 mL of triethylamine (d = 0.726 g/mL, 25.11 mmol, 6 eq), CuI, Pd(dba)₂ and tris(2-furyl)phosphine were added in catalytic quantities. The reaction mixture was introduced into the microwaves oven (100 kPa) and heated at 100 ± 10 °C under constant stirring for 30 minutes.

Work-up

Ammonium chloride was added to the reaction mixture. Then, was diluted with water (3 X 20 mL) and extracted with ethyl acetate (3 X 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was evaporated under reduced pressure.

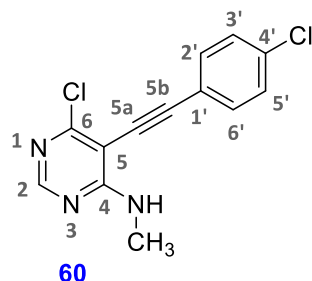
Purification

No purification by column chromatography was performed since the desired was obtained with enough purity as shown in the ¹H spectrum.

Yield: **90%**

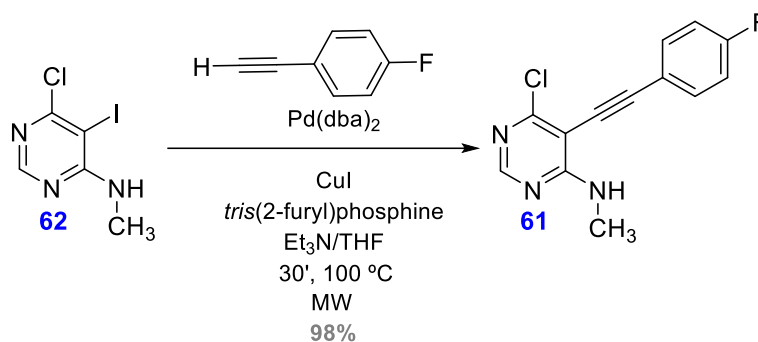
Analytical data

- R_f: 0.25 (Hexane/Ethyl Acetate (7:3))
- Melting point: 167-169 °C (Hexane)
- Aspect: beige solid



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 3.13 (d, 3H, -N-CH₃); 5.70 (s, 1H, NH); 7.37 (d, *J* = 8.8 Hz, 2H, H-3', H-5'); 7.48 (d, *J* = 8.8 Hz, 2H, H-2', H-6'); 8.36 (s, 1H, H-2).
- ¹³C NMR (CDCl₃, 100.6 MHz) δ(ppm), 28.5 (C, -NH-CH₃); 80.6 (C, C-5a); 100.6 (C, C-5b); 101.0 (C, C-5); 120.0 (C, C-1'); 129.0 (CH, C-3', C-5'); 132.9 (CH, C-2', C-6'); 135.6 (C, C-4'); 156.4 (CH, C-2); 159.4 (C, C-4); 162.9 (C, C-6).
- HRMS ESI (+) *m/z*: calculated mass for C₁₃H₁₀N₃Cl₂ 278.0246, found 278.0251.

Preparation of 4-amino-6-chloro-5-((4-fluorophenyl)ethynyl)-*N*-methylpyrimidine (**61**)



Procedure

A 10 mL round-bottomed flask specially equipped with a screw cap for microwaves under argon, was charged with the starting material **62** (0.3 g, 1.11 mmol, 1 eq) dissolved in 1 mL of THF. Then, *p*-fluorophenylacetylene (0.26 mL, $d = 0.930$ g/mL, 2.22 mmol, 2 eq), 3.5 mL of triethylamine ($d = 0.726$ g/mL, 25.11 mmol, 6 eq), CuI, Pd(dba)₂ and tris(2-furyl)phosphine were added in catalytic quantities. The reaction mixture was introduced into the microwaves oven (100 kPa) and heated at 100 ± 10 °C under constant stirring for 30 minutes.

Work-up

Ammonium chloride was added to the reaction mixture. Then, was diluted with water (3 X 20 mL) and extracted with ethyl acetate (3 X 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was evaporated under reduced pressure.

Purification

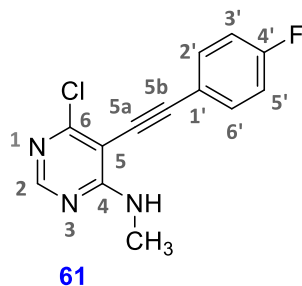
No purification by column chromatography was performed since the desired was obtained with enough purity as shown in the ¹H spectrum.

Yield: 98%

Analytical data

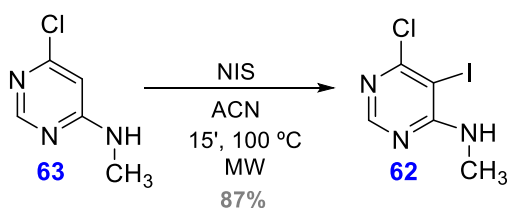
- R_f: 0.27 (Hexane/Ethyl Acetate (7:3))
- Melting point: 168-170 °C (Ethyl Acetate)

- Aspect: white solid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 3.13 (d, 3H, -N- CH_3); 5.70 (s, 1H, NH); 7.09 (d, $J = 8.8$ Hz, 2H, H-3', H-5'); 7.54 (dd, $J_1 = 8.8$, $J_2 = 5.3$ Hz, 2H, H-2', H-6'); 8.36 (s, 1H, H-2).
- ^{13}C NMR (CDCl_3 , 100.6 MHz) δ (ppm), 28.5 (C, -NH- CH_3); 79.3 (C, C-5a); 100.8 (C, C-5b); 101.1 (C, C-5); 116.1 (d, $J = 22$ Hz, C-3', C-5'); 118.1 (d, $J = 3$ Hz, C-1'); 133.8 (d, $J = 9$ Hz, C-2', C-6'); 156.3 (CH, C-2); 160.4 (d, $J = 170$ Hz, C-4'); 162.9 (C, C-4); 164.8 (C, C-6).
- HRMS ESI (+) m/z : calculated mass for $\text{C}_{13}\text{H}_{10}\text{N}_3\text{Cl}_1\text{F}$ 262.0542, found 262.0545.

Preparation of 4-amine-6-chloro-5-iodo-*N*-methylpyrimidine (**62**)



Procedure

A 10 mL round-bottomed flask specially equipped with a screw cap, was charged with the starting material **63** (2 g, 13.90 mmol, 1 eq) dissolved in 2 mL of ACN. Then, NIS (0.675 g, 3 mmol, 4 eq) was added. The reaction mixture was introduced into the microwave oven (100 kPa) and heated at 100 ± 10 °C under constant stirring for 15 minutes.

Work-up

After cooling, the solvent was evaporated under reduced pressure. The crude reaction was diluted with water (3 X 20 mL) and extracted with dichloromethane (3 X 20 mL). Then, the combined organic layers were washed with $\text{Na}_2\text{S}_2\text{O}_3$ and NaOH 10% (2 X 20 mL). The resulting organic phases were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

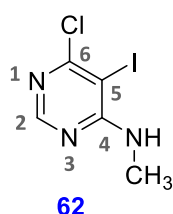
Purification

No purification by column chromatography was performed since the product was recrystallized with ethanol (20 mL).

Yield: **87%**

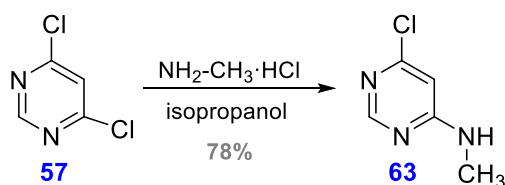
Analytical data

- R_f: 0.48 (Hexane/Ethyl Acetate (7:3))
- Melting point: 146-148 °C (Ethanol)
- Aspect: yellow crystals



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 3.06 (d, *J* = 4.9 Hz, 3H, CH₃); 5.64 (s, 1H, NH); 8.24 (s, 1H, H-2).
- ¹³C NMR (CDCl₃, 100.6 MHz) δ(ppm), 29.45 (C, CH₃); 79.6 (CH, C-5); 157.3 (CH, C-2); 162.2 (C, C-6); 163.0 (C, C-4).
- HRMS ESI (+) *m/z*: calculated mass for C₅H₆N₃ClI 269.9289, found 269.9291.

Preparation of 4-amino-6-chloro-*N*-methylpyrimidine (**63**)



Procedure

A 50 mL round-bottomed flask and equipped with a reflux condenser was charged with the dichloropyrimidine **57** (0.2 g, 1.34 mmol, 1 eq) dissolved in 10 mL of isopropanol. Then, methylamine (0.108 g, 1.61 mmol) was added to the reaction mixture. The resulting mixture was stirred and heated at 80 ± 10 °C for 24 hours.

Work-up

The solvent was evaporated under reduced pressure. Then, ¹H spectrum was carried out.

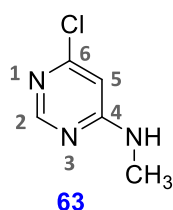
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown in the ^1H spectrum.

Yield: **78%**

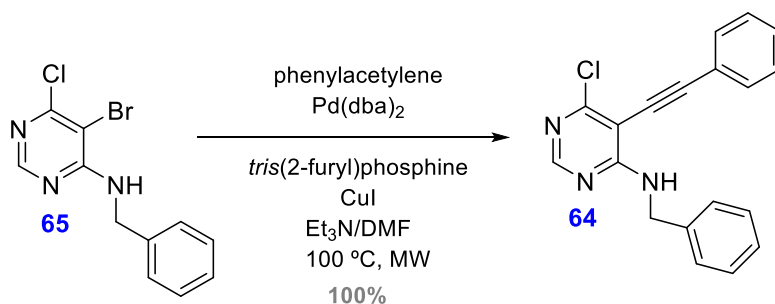
Analytical data

- R_f : 0.1 (Hexane/Ethyl Acetate (7:3))
- Melting point: 136-138 °C (Ethyl Acetate)
- Aspect: white solid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 2.92 (d, $J = 4.3$ Hz, 3H, CH_3); 5.87 (s, 1H, NH); 6.33 (s, 1H, H-5); 8.31 (s, 1H, H-2).
- ^{13}C NMR (CDCl_3 , 100.6 MHz) δ (ppm), 28.3 (C, CH_3); 99.9 (CH, C-5); 158.2 (CH, C-2); 159.8 (C, C-6); 164.0 (C, C-4).
- HRMS ESI (+) m/z : calculated mass for $\text{C}_5\text{H}_7\text{N}_3\text{Cl}$ 144.0323, found 144.0322.

Preparation of 4-amine-*N*-benzyl-6-chloro-5-(phenylethynyl)pyrimidine (**64**)



Procedure

A 10 mL round-bottomed flask specially equipped with a screw cap for microwaves under argon, was charged with the starting bromopyrimidine **65** (0.309 g, 1.03 mmol, 1 eq) dissolved in 3 mL of DMF.

Then, phenylacetylene (0.13 mL, $d = 1.109 \text{ g/mL}$, 1 eq), 0.5 mL of triethylamine ($d = 0.726 \text{ g/mL}$, 3.58 mmol, 3 eq), CuI, Pd(dba)₂ and *tris*(2-furyl)phosphine were added in catalytic quantities. The reaction mixture was introduced to the microwaves oven (100 kPa) and heated at $100 \pm 10 \text{ }^\circ\text{C}$ under constant stirring for 90 minutes.

Work-up

The reaction mixture was diluted with water (15 mL) and extracted with diethyl ether (3 X 15 mL). The combined organic phases were washed with water (3 X 20 mL). The combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

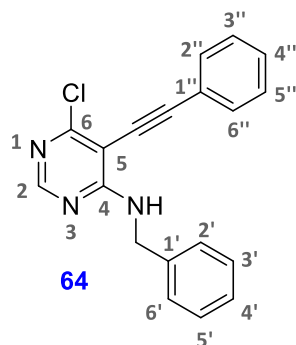
Purification

No purification by column chromatography was performed since the desired product was obtained with enough purity as shown in the ¹H spectrum.

Yield: **100%**

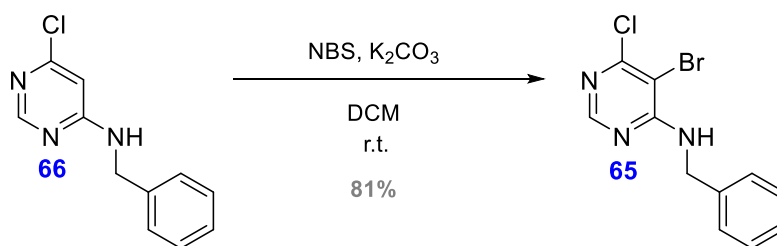
Analytical data

- R_f: 0.71 (Hexane/Ethyl Acetate (7:3))
- Aspect: brown-yellow semisolid



- ¹H NMR (CDCl₃, 400 MHz) δ (ppm), 4.56 (d, $J = 6 \text{ Hz}$, 2H, Ar-CH₂-); 5.98 (t, $J = 6 \text{ Hz}$, 1H, NH); 7.29-7.33 (m, 1H, H-4''); 7.34-7.37 (m, 3H, H-3', H-4', H-5'); 7.37 (d, $J = 4.3 \text{ Hz}$, 2H, H-2', H-6'); 7.51 (dd, $J_1 = 2, J_2 = 8 \text{ Hz}$, 2H, H-2'', H-6''); 7.61 (t, $J_1 = 2, J_2 = 8 \text{ Hz}$, 2H, H-3'', H-5''); 8.55 (s, 1H, H-2).

Preparation of 4-amine-*N*-benzyl-5-bromo-6-chloropyrimidine (**65**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar, previously flame-dried under argon, was charged with 4-amine-*N*-benzyl-6-chloropyrimidine (**66**) (0.277 g, 0.97 mmol, 1 eq) dissolved in 15 mL of dichloromethane. Then, potassium carbonate (0.185 g, 1.34 mmol, 1 eq) and NBS (0.262 g, 1.47 mmol, 1.1 eq) were added to the reaction mixture. The resulting mixture was stirred at room temperature for 24 hours.

Work-up

The crude reaction was diluted with a solution of sodium hydroxide (2N) (3 X 20 mL) and extracted with dichloromethane (3 X 20 mL). Then, the combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

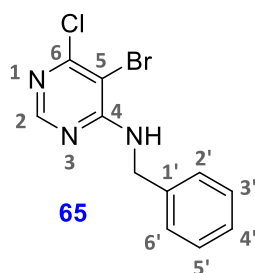
Purification

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (70:30).

Yield: 81%

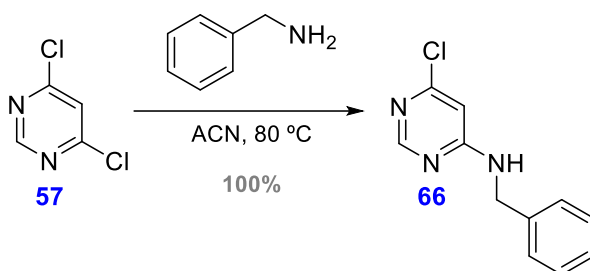
Analytical data

- R_f: 0.8 (Hexane/Ethyl acetate (7:3))
- Melting point: 150-153 °C (ethyl acetate)
- Aspect: white-yellowish solid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 4.72 (s, 2H, Ar- CH_2 -); 5.92 (bs, 1H, NH); 7.29-7.32 (m, 3H, H-2', H-4', H-6'); 7.36 (t, J = 7 Hz, 2H, H-3', H-5'); 8.33 (s, 1H, H-2).

Preparation of 4-amine-*N*-benzyl-6-chloropyrimidine (66)



Procedure

A 50 mL round-bottomed flask specially equipped with a screw cap for coupling reactions, previously flame-dried under argon, was charged with the starting 4,6-dichloropyrimidine (**57**) (0.2 g, 1.34 mmol, 1 eq) dissolved in 10 mL of ACN. Then, 0.176 mL of benzylamine ($d = 0.98$ g/mL, 1.6 mmol, 1.2 eq) was added to the reaction mixture. The resulting mixture was stirred and heated at 80 ± 10 °C for 24 hours.

Work-up

The solvent was evaporated under reduced pressure. Then, ^1H spectrum was carried out.

Purification

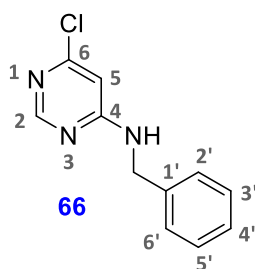
No purification by column chromatography was performed since the product was obtained with enough purity as shown in the ^1H spectrum.

Yield: 100%

Analytical data

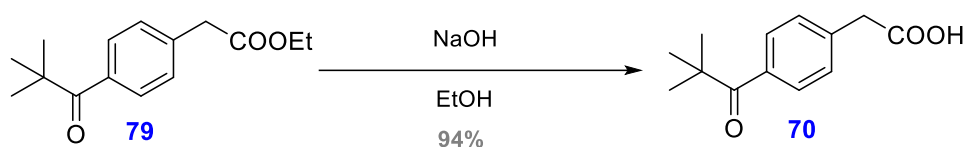
- R_f : 0.487 (Hexane/Ethyl Acetate (7:3))
- Melting point: 93-95 °C (Ethyl Acetate)

- Aspect: white solid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 4.50 (s, 2H, Ar- CH_2 -); 6.48 (s, 1H, H-5); 7.30-7.32 (m, 3H, H-2', H-3', H-6'); 7.34 (t, J = 6 Hz, 2H, H-3', H-5'); 8.26 (s, 1H, H-2).

Preparation of 2-(4-pivaloylphenyl)acetic acid (**70**)



Procedure

A 100 mL round-bottomed flask with a magnetic stirring bar was charged with the initial compound **79** (1.773 g, 7.14 mmol, 1 eq) dissolved in a mixture of 16 mL of an aqueous NaOH (2N) solution and 4 mL of ethanol (4:1). The reaction mixture was stirred at room temperature for 4.5 hours.

Work-up

Ethanol was removed under vacuum. HCl 2N was then added to the mixture in order to obtain a slightly acid pH (5-6). Dichloromethane (3 X 20 mL) was used to extract the crude reaction. The combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was removed under reduced pressure.

Purification

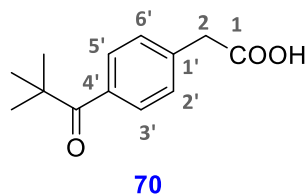
No purification by column chromatography was performed since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: 94%

Analytical data

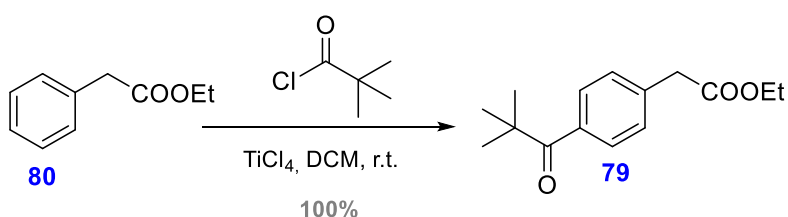
- R_f : 0.01 (Hexane/Ethyl Acetate (1:1))

- Melting point: 63-65 °C (Dichloromethane)
- Aspect: white-yellow solid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 1.22 (s, 9H, CH_3 - (x3)); 3.66 (s, 2H, H-2); 7.29 (d, $J = 7.2$ Hz, 2H, H-2', H-6'); 7.33 (dd, $J_1 = 1$, $J_2 = 7.2$ Hz, 2H, H-3', H-5').

Preparation of ethyl 2-(4-pivalophenyl)acetate (79)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar was charged with ethyl 2-phenylacetate (**80**) (1 g, 6.09 mmol, 1 eq) dissolved in 15 mL of dichloromethane and cooled down to 0 ± 10 °C with an ice-bath. Then, pivaloyl chloride (1.5 mL, $d = 0.98$ g/mL, 12.19 mmol, 2 eq) and titanium tetrachloride (1.4 mL, $d = 1.726$ g/mL, 12.76 mmol, 2 eq) were added dropwise to the reaction mixture. The mixture was stirred at room temperature for 24 hours. The system was closed with a CaCl_2 tube.

Work-up

The reaction mixture was cooled down by the addition of crushed ice. Then, the crude mixture was diluted with 10 mL of an aqueous HCl solution (2N). The crude reaction was cooled down at room temperature for 10 minutes. After that time, the reaction residue was diluted with water (3 X 20 mL) and extracted with dichloromethane (3 X 20 mL). Finally, the combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was removed under reduced pressure.

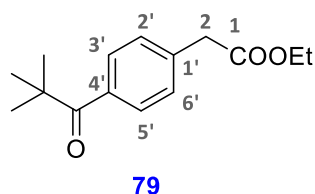
Purification

The crude product was used in the next step without further purification since the product was obtained with enough purity as shown by TLC and ^1H spectrum. A mixture of rotamers or conformers of the final product was obtained.

Yield: **100%**

Analytical data

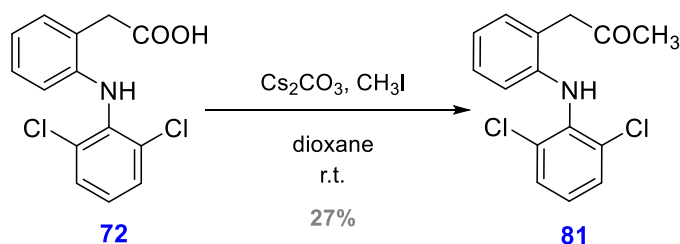
- R_f : 0.82 (Hexane/Ethyl Acetate (1:1))
- Aspect: green oil



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 1.23 (s, 9H, CH_3 - (x3)); 1.25 (t, $J = 7.1$ Hz, 3H, CH_3 - CH_2 -); 3.61 (s, 2H, H-2); 4.15 (q, $J = 7.1$ Hz, 2H, CH_2 -O-); 7.27-7.28 (m, 2H, H-2', H-6'); 7.32 (t, $J = 6$ Hz, 2H, H-3', H-5').
- ^{13}C NMR (CDCl_3 , 100.6 MHz) δ (ppm), 14.1 (CH_3 -); 27.0 (CH_3 -); 38.5 (C, $(\text{CH}_3)_2$); 41.4 (CH_2 , C-2); 60.9 (CH_2 -O-); 127.0 (CH, C-3', C-5') (B); 128.5 (CH, C-3', C-5') (A); 129.2 (CH, C-2', C-6') (A); 134.1 (CH, C-2', C-6') (B); 171.7 (C, COOH); 185.1 (C, Ar-C=O).

A: majority isomer / B: minority isomer

Preparation of 1-(2-((2,6-dichlorophenyl)amino)phenyl)propan-2-one (**81**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar and a reflux condenser was charged with the aryl acetic acid **72** (1 g, 3.38 mmol, 1 eq) dissolved in 10 mL of dioxane.

Then, caesium carbonate (1.1 g, 3.38 mmol, 1.2 eq) and methyl iodide (1 mL, d = 2.28 g/mL, 16.10 mmol, 5 eq) were added to the reaction mixture. The mixture was stirred at room temperature for 72 hours.

Work-up

The reaction mixture was diluted with water (20 mL) and extracted with dichloromethane (3 X 20 mL). The combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was evaporated under reduced pressure.

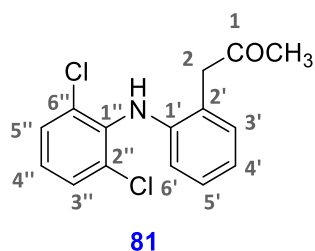
Purification

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (70:30).

Yield: **27%**

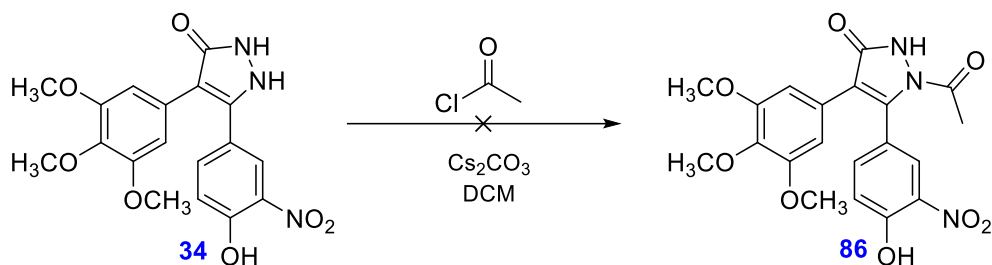
Analytical data

- R_f: 0.82 (Hexane/Ethyl Acetate (1:1))
- Melting point: 118-120 °C (Ethyl Acetate)
- Aspect: brownish solid



- ¹H NMR ((CD₃)₂CO, 400 MHz) δ(ppm), 2.09 (s, 2H, H-2); 3.90 (s, 3H, CH₃-O-); 6.55 (dd, J₁ = 3, J₂ = 8 Hz, 1H, H-6'); 6.79 (bs, 1H, NH); 6.95-6.99 (m, 2H, H-3', H-4'); 7.14 (t, J = 8 Hz, 1H, H-4''); 7.25 (dd, J₁ = 1.5, J₂ = 8 Hz, 1H, H-5'); 7.33 (dd, J₁ = 1.5, J₂ = 8. Hz, 2H, H-3'', H-5'').

Preparation of 1-acetyl-5-(4-hydroxy-3-nitrophenyl)-4-(3,4,5-trimethoxyphenyl)-1H-pyrazol-3(2H)-one (**86**)



Procedure

A 50 mL round-bottomed flask equipped with a magnetic stirring bar was charged with compound **34** (0.015 g, 0.037 mmol, 1 eq) dissolved in 15 mL of dichloromethane. Then, 0.0032 mL of acetyl chloride (d = 1.1 g/mL, 0.044 mmol, 1.2 eq) and caesium carbonate (0.014 g, 0.044 mmol, 1.2 eq) were added dropwise to the reaction mixture. The additions were always in ice-bath conditions (0 ± 10 °C). Finally, the resulting mixture was stirred at room temperature for 24 hours.

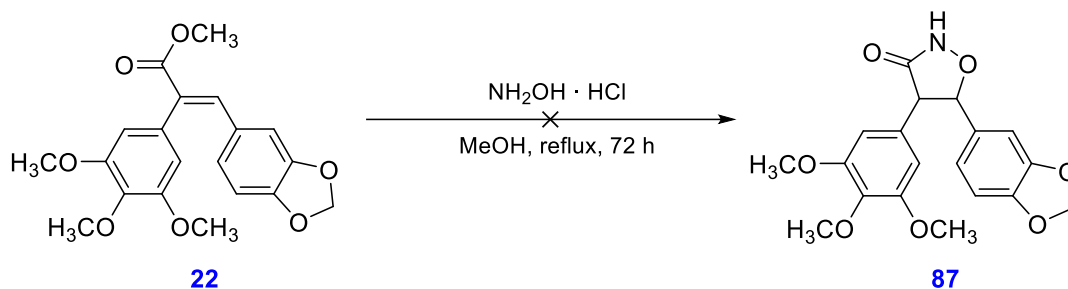
Work-up

The reaction mixture was diluted with water (15 mL) and extracted with dichloromethane (3 X 15 mL). Then, the combined organic layers were dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

Purification

The final solid was the initial aldehyde as shown in the ¹H spectrum. A retrosynthesis was carried out in this case.

Preparation of 5-(benzo[d][1,3]dioxol-5-yl)-4-(3,4,5-trimethoxyphenyl)isoxazolidin-3-one
(87)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar and a reflux condenser was charged with compound **22** (0.208 g, 0.56 mmol, 1 eq) dissolved in 30 mL of methanol. Then, hydroxylamine in catalytic amount was added to the mixture. The reaction was heated at reflux temperature of 80 ± 10 °C under stirring for 72 hours, during which hydroxylamine in excess and 4 drops of trimethylamine were added to the reaction mixture in order to obtain the desired product.

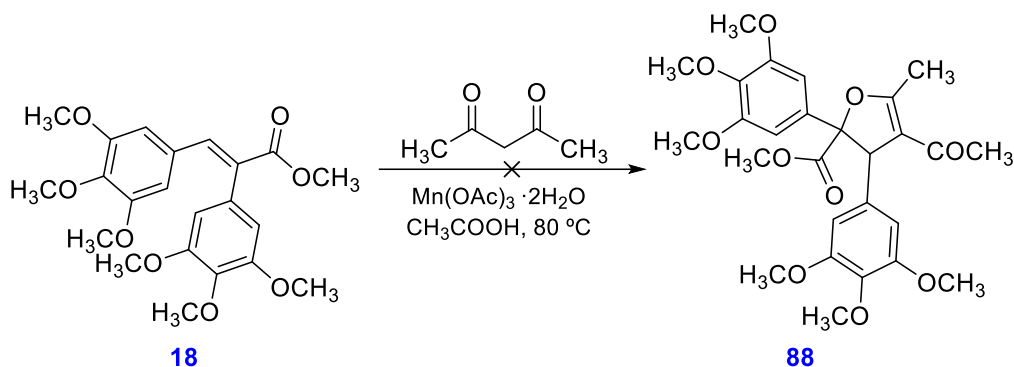
Work-up

Methanol was firstly evaporated under reduced pressure and then 20 mL of water was added and the crude reaction was extracted with diethyl ether/ethyl acetate 1/1 (3 X 20 mL). The organic layer was dried with anhydrous Na_2SO_4 , filtered under vacuum and the solvent was removed under reduced pressure.

Purification

No purification was performed since no new product was observed by TLC or ^1H spectrum.

Preparation of methyl 4-acetyl-5-methyl-2,3-bis(3,4,5-trimethoxyphenyl)-2,3-dihydrofuran-2-carboxylate (**88**)



Procedure

A 50 mL round-bottomed flask equipped with a magnetic stirring bar and a reflux condenser, previously flame-dried under argon was charged with manganese (III) acetate (0.270 g, 1.008 mmol) in 3 mL of glacial acetic acid ($d = 1.05 \text{ g/mL}$, 52.46 mmol). The reaction mixture was heated at $80 \pm 10^\circ\text{C}$ for 15 minutes. Then, was cooled down to $0 \pm 10^\circ\text{C}$ with an ice-bath and the carboxylic ester **18** (0.175 g, 0.41 mmol, 1 eq) was added dropwise dissolved in a mixture of 0.74 mL of acetyl acetone ($d = 0.97 \text{ g/mL}$, 7.21 mmol) and 5 mL of glacial acetic acid. After that, the resulting mixture was heated at $80 \pm 10^\circ\text{C}$ under stirring for 24 hours.

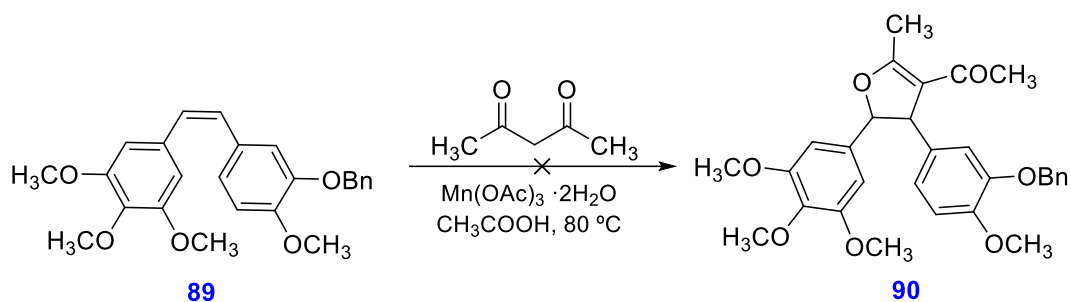
Work-up

The crude reaction was diluted with saturated NaHCO_3 (20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the organic phases were washed with water (3 X 20 mL). Finally, the combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was evaporated under reduced pressure.

Purification

After ^1H NMR was carried out the product obtained was not the desired.

Preparation of 1-(4-(3-(benzyloxy)-4-methoxyphenyl)-2-methyl-5-(3,4,5-trimethoxyphenyl)-4,5-dihydrofuran-3-yl)ethanone (**90**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar and a reflux condenser, previously flame-dried under argon was charged with manganese (III) acetate (0.270 g, 1.008 mmol) in 3 mL of glacial acetic acid (d = 1.05 g/mL, 52.46 mmol). The reaction mixture was heated at 80 ± 10 °C for 15 minutes. Then, was cooled down to 0 ± 10 °C with an ice-bath and the stilbene **89** (0.115 g, 0.29 mmol, 1 eq) was added dropwise dissolved in a mixture of 0.74 mL of acetyl acetone (d = 0.97 g/mL, 7.21 mmol) and 5 mL of glacial acetic acid. After that, the resulting mixture was heated at 80 ± 10 °C under stirring for 72 hours.

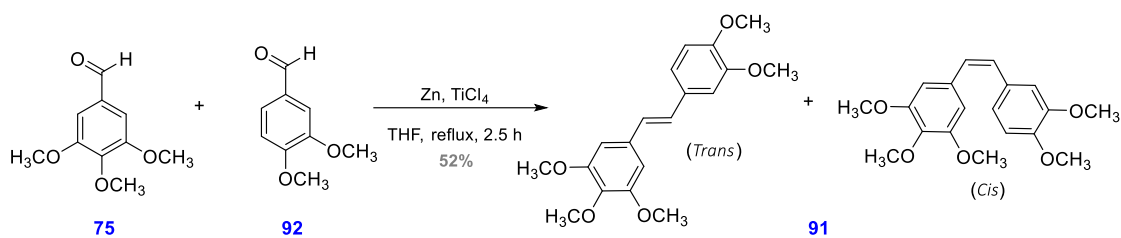
Work-up

The crude reaction was diluted with saturated NaHCO_3 (20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the organic phases were washed with water (3 X 20 mL). Finally, the combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was evaporated under reduced pressure.

Purification

The crude reaction was purified by column chromatography using a CombiFlash[®] R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired compound was not obtained in this case.

Preparation of (*E*)/(*Z*)-5-(3,4-dimethoxystyryl)-1,2,3-trimethoxybenzene (**91**)



Procedure

A 50 mL round-bottomed 3 necks flask with a magnetic stirring bar, previously flame-dried under argon was charged with zinc (0.8 g, 12 mmol, 5 eq) dissolved in 30 mL of dry THF. Then, the mixture was cooled down to 0 ± 10 °C with an ice-bath and titanium tetrachloride (0.65 mL, 6 mmol, 2.5 eq) was added dropwise and stirred for 30 minutes at room temperature. After that time, the reaction mixture was equipped with a reflux condenser and was stirred at 70 ± 10 °C for 2.5 hours. The mixture was cooled down to 0 ± 10 °C with an ice-bath and the initial aldehyde **92** (0.398 g, 2.39 mmol, 1 eq) and compound **75** (0.471 g, 2.39 mmol, 1 eq) were added to the reaction mixture in 30 mL of dry THF. Finally, the resulting mixture was heated at reflux temperature under constant stirring for 24 hours.

Work-up

First, the crude reaction was diluted with an aqueous sodium bicarbonate 10% (25 mL) and extracted with dichloromethane (3 X 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was removed under reduced pressure.

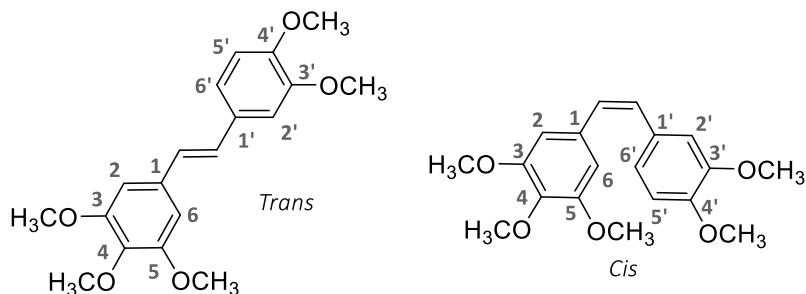
Purification

The crude reaction obtained was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (70:30). The mixture of *cis/trans* compounds was shown in ¹³C NMR spectrum.

Yield: 52%

Analytical data

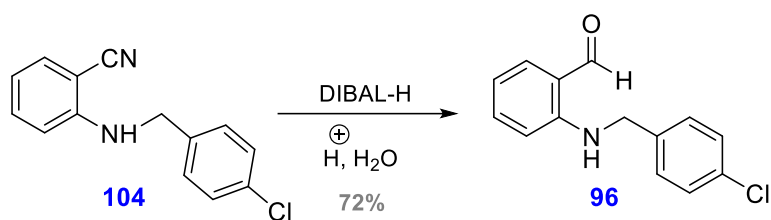
- R_f: 0.6 (Hexane/Ethyl Acetate (1:1))
- Aspect: yellowish oil



91

- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 3.62 (s, 3H, CH₃-O-) (*Trans*); 3.63 (s, 3H, CH₃-O-) (*Cis*); 3.82 (s, 3H, CH₃-O-) (*Trans*); 3.83 (s, 3H, CH₃-O-) (*Cis*); 3.87 (s, 3H, CH₃-O-) (*Trans*); 3.88 (s, 3H, CH₃-O-) (*Cis*); 3.91 (s, 3H, CH₃-O-) (*Cis*); 3.93 (s, 3H, CH₃-O-) (*Trans*); 3.94 (s, 3H, CH₃-O-) (*Cis*); 3.96 (s, 3H, CH₃-O-) (*Trans*); 6.59 (s, 2H, H-2, H-6); 6.72 (d, *J* = 8 Hz, 1H, H-5'); 6.84 (dd, *J*₁ = 2, *J*₂ = 8 Hz, 1H, H-6'); 6.90 (d, *J* = 2 Hz, 1H, H-2') (*Cis*); 6.91 (d, *J* = 2 Hz, 1H, H-2') (*Trans*); 6.92 (d, *J* = 12 Hz, 2H, CH=CH) (*Cis*); 7.04 (d, *J* = 16 Hz, 2H, CH=CH) (*Trans*).
- ¹³C NMR (CDCl₃, 100.6 MHz) δ(ppm), 55.6 (CH₃-O-); 55.7 (CH₃-O-); 55.8 (CH₃-O-); 55.81 (CH₃-O-); 55.85 (CH₃-O-); 55.9 (CH₃-O-); 56.0 (CH₃-O-); 56.1 (CH₃-O-); 56.2 (CH₃-O-); 56.07 (CH₃-O-); 103.3 (CH, C-2, C-6) (*Trans*); 103.5 (CH, C-2, C-6) (*Cis*); 107.0 (CH, C-2') (*Trans*); 107.1 (CH, C-2') (*Cis*); 110.9 (CH, C-5') (*Cis*); 111.0 (CH, C-5') (*Trans*); 122.2 (CH, C-6') (*Trans*); 122.3 (CH, C-6') (*Cis*); 126.5 (CH, CH=) (*Cis*); 126.7 (CH, CH=) (*Trans*); 127.9 (CH, CH=) (*Trans*); 128.0 (CH, CH=) (*Cis*); 130.1 (C, C-1') (*Cis*); 130.2 (C, C-1') (*Trans*); 133.0 (C, C-1) (*Cis*); 133.2 (C, C-1) (*Trans*); 135.9 (C, C-4) (*Cis*); 136.2 (C, C-4) (*Trans*); 143.7 (C, C-4') (*Trans*); 144.0 (C, C-4') (*Cis*); 147.3 (C, C-3') (*Cis*); 147.9 (C, C-3') (*Trans*); 152.9 (C, C-3, C-5) (*Cis*); 153.3 (C, C-3, C-5) (*Trans*).

Preparation of 2-((4-chlorobenzyl)amino)benzaldehyde (**96**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar under argon and cooled down 0 ± 10 °C with an ice bath, was charged with 2-((4-chlorobenzyl)amino)benzonitrile (**104**) (0.2 g, 0.824 mmol, 1 eq) dissolved in 10 mL of toluene. Then, DIBAL-H (d = 1.23 g/mL, 0.90 mL, 1.1 eq) was finally added dropwise. The resulting mixture was stirred at room temperature for 45 minutes.

Work-up

HCl (5N) solution was added dropwise to the crude reaction in order to hydrolyze the imine formatted. Then, ethyl acetate (3 X 20 mL) was used to extract. The organic phase was dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was evaporated under reduced pressure.

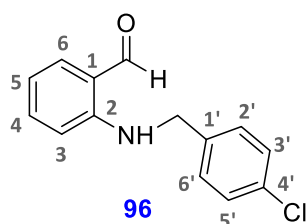
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown by TLC and ¹H spectrum. The product was used for the next synthetic step.

Yield: 72%

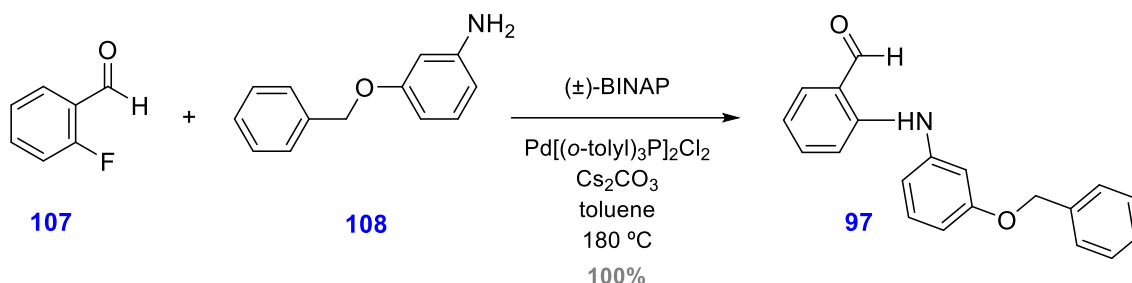
Analytical data

- R_f: 0.90 (Hexane/Ethyl Acetate (1:1))
- Aspect: yellow oil



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 4.45 (s, 2H, $\text{CH}_2\text{-N}$); 5.20 (bs, 1H, NH); 6.56 (d, $J = 8$ Hz, 2H, H-2', H-6'); 7.32-7.33 (m, 4H, H-3, H-4, H-5, H-6); 7.32 (d, $J = 8$ Hz, 2H, H-3', H-5'); 9.98 (s, 1H, CHO).

Preparation of 2-((3-(benzyloxy)phenyl)amino)benzaldehyde (**97**)



Procedure

A 50 mL round-bottomed flask specially equipped with a screw cap for coupling reactions, previously flame-dried under argon, was charged with the starting aldehyde **107** (0.42 mL, 4.02 mmol, 1 eq) dissolved in 10 mL of toluene. Then, the corresponding amine **108** (0.803 g, 4.02 mmol, 1 eq), caesium carbonate (1.60 g, 4.83 mmol, 1.2 eq), (\pm) -BINAP and palladium complex (II) in catalytic quantities were added. The reaction mixture was heated at $180 \pm 10\text{ }^\circ\text{C}$ under stirring for 48 hours.

Work-up

The reaction mixture was cooled down to room temperature, 20 mL of water was then added and ethyl acetate (3 X 20 mL) was used to extract the crude reaction. The combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was evaporated under reduced pressure.

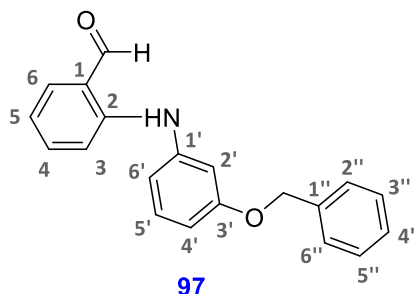
Purification

The crude product was used in the next step without further purification since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: 100%

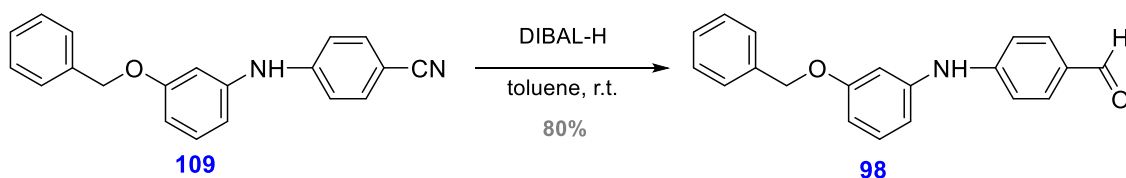
Analytical data

- R_f: 0.47 (Hexane/Ethyl Acetate (8:2))
- Aspect: brown semisolid



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 5.15 (s, 2H, CH₂-O-); 6.84 (d, *J* = 7.5 Hz, 1H, H-4'); 6.86-6.89 (m, 2H, H-2', H-6'); 7.13 (dt, *J*₁ = 2, *J*₂ = 7.6 Hz, 1H, H-5); 7.25 (t, *J* = 7.6 Hz, H-5'); 7.28 (d, *J* = 7.6 Hz, 1H, H-6); 7.30 (d, *J* = 7.6 Hz, 1H, H-3); 7.40 (t, *J* = 7 Hz, 2H, H-3'', H-5''); 7.45 (d, *J* = 7 Hz, 2H, H-2'', H-6''); 7.45-7.46 (m, 1H, H-4''); 8.17 (dd, *J*₁ = 2, *J*₂ = 7.6 Hz, 1H, H-4); 8.75 (s, 1H, CHO).

Preparation of 4-((3-(benzyloxy)phenyl)amino)benzaldehyde (**98**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar under argon and cooled down 0 ± 10 °C with an ice bath, was charged with 4-((3-(benzyloxy)phenyl)amino)benzonitrile (**109**) (0.200 g, 0.66 mmol, 1 eq) dissolved in 10 mL of toluene. Then, DIBAL-H (d = 1.23 g/mL, 0.73 mL, 1.1 eq) was finally added dropwise. The resulting mixture was stirred at room temperature for 45 minutes.

Table 23. Yield obtained depending on the time reaction

Starting material (mg)	Time (minutes)	Yield (%)
100	15	40
100	30	68
200	45	80

Work-up

HCl (5N) solution was added dropwise to the crude reaction in order to hydrolyze the formed imine. Then, ethyl acetate (3 X 20 mL) was used to extract. The organic phase was dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was evaporated under reduced pressure.

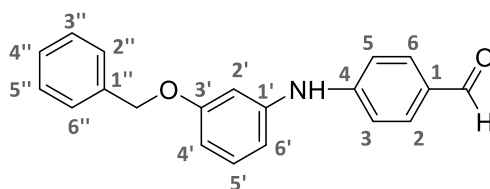
Purification

No purification by column chromatography was performed since the product was obtained with enough purity as shown by TLC and ¹H spectrum. The product was used for the next synthetic step.

Yield: **80%**

Analytical data

- R_f: 0.83 (Hexane/Ethyl acetate (1:1))
- Aspect: yellow oil

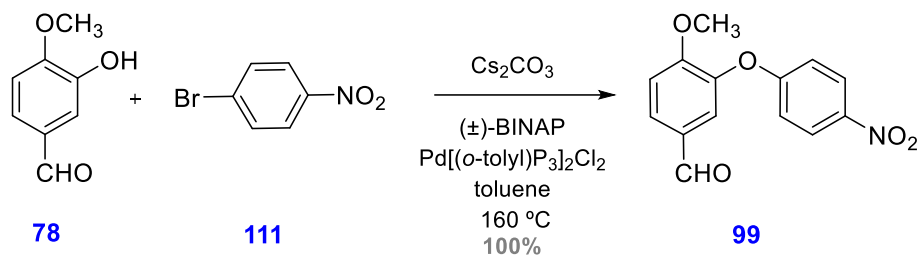


98

- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 4.95 (s, 2H, Ar-CH₂); 6.22 (bs, 1H, NH); 6.63 (dd, J₁ = 2, J₂ = 9 Hz, 1H, H-4'); 6.65 (d, J = 9 Hz, 1H, H-6'); 6.72 (t, J = 2 Hz, 1H, H-2'); 6.90 (d, J = 7 Hz, 2H, H-3, H-5); 7.15 (t, J = 7.7 Hz, 1H, H-4''); 7.24 (t, J = 7 Hz, 1H, H-5'); 7.28 (d, J = 7,7 Hz, 2H, H-2'', H-6''); 7.31 (t, J = 7.7 Hz, 2H, H-3'', H-5''); 7.72 (d, J = 7 Hz, H-2, H-6).

- ^{13}C NMR (CDCl_3 , 100.6 MHz) δ (ppm), 70.0 (CH_2 , Ar- CH_2); 107.6 (CH , C-2'); 110.1 (CH , C-4'); 113.6 (CH , C-6'); 114.9 (CH , C-3, C-5); 127.4 (CH , C-2'', C-6''); 128.0 (CH , C-2'', C-6''); 128.6 (CH , C-3'', C-5''); 129.1 (CH , C-4''); 130.3 (CH , C-5'); 136.9 (C , C-1''); 141.5 (C , C-1'); 149.6 (C , C-4); 159.8 (C , C-3'); 190.5 (CH , CHO).

Preparation of 4-methoxy-3-(4-nitrophenoxy)benzaldehyde (**99**)



Procedure

A 50 mL round-bottomed flask specially equipped with a screw cap for coupling reactions, previously flame-dried under argon, was charged with 3-hydroxy-4-methoxybenzaldehyde (**78**), (0.4 g, 2.63 mmol, 1 eq), 4-bromo-nitrobenzene (**111**), (0.584 g, 2.89 mmol, 1.1 eq), caesium carbonate (1.028 g, 3.15 mmol, 1.2 eq), (\pm)-BINAP and palladium (II) complex in catalytic quantities in 10 mL of toluene. The reaction mixture was heated at 160 ± 10 °C under stirring for 24 hours.

Work-up

The reaction mixture was diluted with water (20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was removed under reduced pressure.

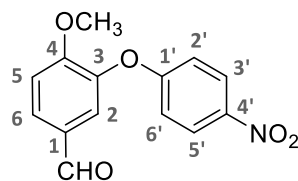
Purification

The crude product was directly used for the next synthetic step since the product was obtained with enough purity as shown by TLC and ^1H spectrum

Yield: **100%**

Analytical data

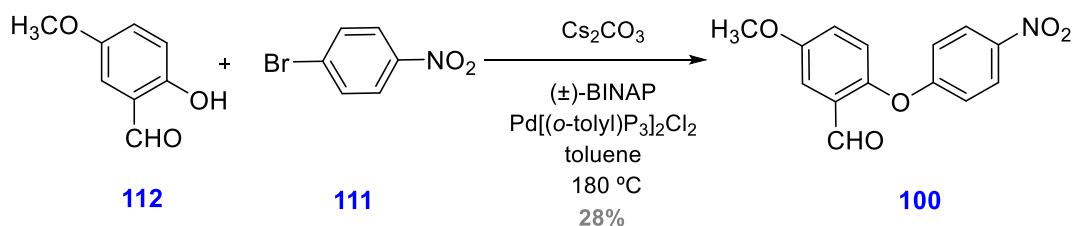
- R_f : 0.61 (Hexane/Ethyl Acetate (1:1))
- Aspect: brown-yellowish oil



99

- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 3.88 (s, 3H, $\text{CH}_3\text{-O}$); 6.93 (d, $J = 9$ Hz, 2H, H-2', H-6'); 7.16 (d, $J = 8.5$ Hz, 1H, H-5); 7.63 (s, 1H, H-2); 7.78 (dd, $J_1 = 2$, $J_2 = 8.5$ Hz, 1H, H-6); 8.17 (d, $J = 9$ Hz, 2H, H-3', H-5'); 9.88 (s, 1H, CHO).
- ^{13}C NMR (CDCl_3 , 100.6 MHz) δ (ppm): 56.3 ($\text{CH}_3\text{-O}$); 112.7 (CH, C-5); 115.9 (CH, C-2'); 116.1 (CH, C-6'); 122.4 (CH, C-2); 125.9 (CH, C-3', C-5'); 130.0 (CH, C-6); 130.4 (C, C-1); 142.8 (C, C-4'); 143.2 (C, C-3); 156.7 (C, C-4); 162.8 (C, C-1'); 189.9 (CH, CHO).

Preparation of 5-methoxy-2-(4-nitrophenoxy)benzaldehyde (**100**)



Procedure

A 50 mL round-bottomed flask specially equipped with a screw cap for coupling reactions, previously flame-dried under argon, was charged with the starting aldehyde **112** (0.4 mL, 3.29 mmol, 1 eq) dissolved in 10 mL of toluene. Then, 1-bromo-4-nitrobenzene (**111**) (0.730 g, 3.61 mmol, 1.1 eq), caesium carbonate (1.40 g, 4.27 mmol, 1.2 eq), (\pm)-BINAP and palladium complex (II) in catalytic quantities were added. The reaction mixture was heated at 180 ± 10 °C under constant stirring for 72 hours.

Work-up

The reaction mixture was cooled down to room temperature, 20 mL of water was then added and ethyl acetate (3 X 20 mL) was used to extract the crude reaction. The combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was evaporated under reduced pressure.

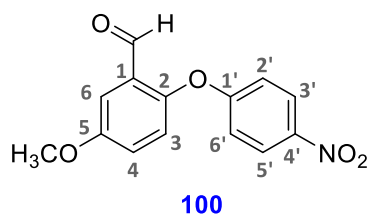
Purification

The crude reaction was purified by column chromatography using a CombiFlash® R_f provided with a UV-Vis detector. The eluent used was a mixture of hexane and ethyl acetate with increasing polarity. The desired product eluted with a mixture of hexane/ethyl acetate (80:20).

Yield: **28%**

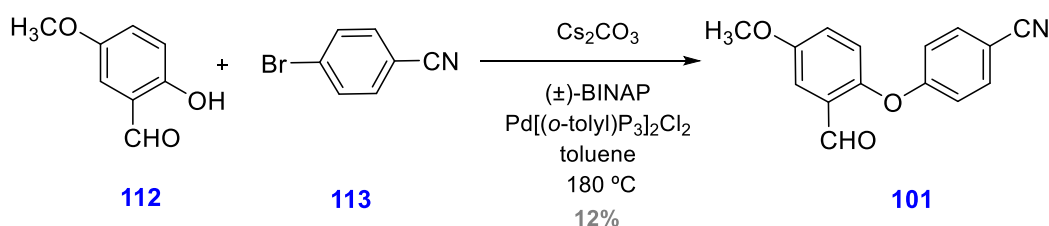
Analytical data

- R_f: 0.43 (Hexane/Ethyl Acetate (8:2))
- Melting point: 102-104 °C (Dichloromethane)
- Aspect: yellow solid



- ¹H NMR (CDCl₃, 400 MHz) δ(ppm), 3.98 (s, 3H, CH₃-O-); 7.04 (d, *J* = 9.2 Hz, 2H, H-2', H-6'); 7.05 (d, *J* = 8.9 Hz, 1H, H-3); 7.22 (dd, *J*₁ = 3, *J*₂ = 8.9 Hz, 1H, H-4); 7.43 (d, *J* = 3 Hz, 1H, H-6); 8.23 (d, *J* = 9.2 Hz, 2H, H-3', H-5'); 10.25 (s, 1H, CHO).

Preparation of 5-methoxy-2-(4-nitrophenoxy)benzaldehyde (**101**)



Procedure

A 50 mL round-bottomed flask specially equipped with a screw cap for coupling reactions, previously flame-dried under argon, was charged with 2-hydroxy-5-methoxybenzaldehyde (**112**) (0.25 mL, 1.25 mmol, 1 eq) dissolved in 10 mL of toluene. Then, 1-bromo-4-benzonitrile (**113**) (0.400 g, 2.16 mmol, 1.1 eq), caesium carbonate (0.500 g, 2.36 mmol, 1.2 eq), (±)-BINAP and palladium complex (II) in catalytic quantities were added. The reaction mixture was heated at 180 ± 10 °C under stirring for 72 hours.

Work-up

The reaction mixture was cooled down to room temperature, then 20 mL of water was added and ethyl acetate (3 X 20 mL) was used to extract the crude reaction. The combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was evaporated under reduced pressure.

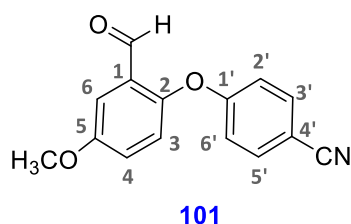
Purification

The crude product was used in the next step without further purification since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: **12%**

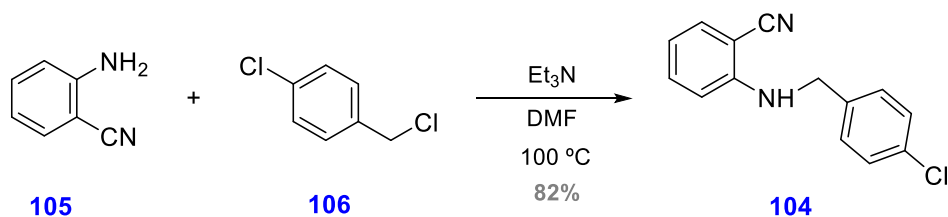
Analytical data

- R_f : 0.55 (Hexane/Ethyl Acetate (8:2))
- Aspect: brown solid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 3.82 (s, 3H, $\text{CH}_3\text{-O-}$); 6.93 (d, $J = 8$ Hz, 1H, H-3); 7.01 (d, $J = 2$ Hz, 1H, H-6); 7.15 (dd, $J = 2$ Hz, 1H, H-4); 7.51 (d, $J = 8.5$ Hz, 2H, H-2', H-6'); 7.63 (d, $J = 8.5$ Hz, 2H, H-3', H-5').

Preparation of 2-ciano-*N*-(4-chlorobenzyl)-aniline (**104**)



Procedure

A 100 mL round-bottomed flask equipped with a magnetic stirring bar and equipped with a reflux condenser was charged the initial 2-aminobenzonitrile (**105**) (0.500 g, 0.042 mmol, 1 eq) dissolved in 10 mL of DMF. Then, triethylamine (0.70 mL, 0.05 mmol, 1.2 eq) and 1-chloro-4-(chloromethyl)benzene (**106**) (0.681 g, 0.042 mmols, 1 eq) were added to the reaction mixture. The resulting mixture was stirred and heated at $100 \pm 10\text{ }^\circ\text{C}$ for 72 hours.

Work-up

The reaction mixture was diluted with water (15 mL) and extracted with diethyl ether (3 X 15 mL). The combined organic layers were washed with water (3 X 20 mL), dried over anhydrous sodium sulfate, filtered under vacuum and the solvent was removed under reduced pressure.

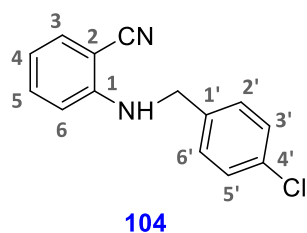
Purification

The crude product was used in the next step without further purification since the product was obtained with enough purity as shown by TLC and ^1H spectrum.

Yield: 82%

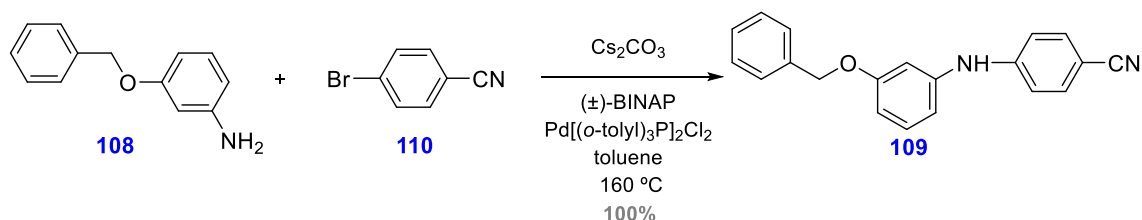
Analytical data

- R_f : 0.88 (Hexane/Ethyl Acetate (7:3))
- Melting point: 83-85 $^\circ\text{C}$ (Dichloromethane)
- Aspect: brown solid



- ^1H NMR (CDCl_3 , 400 MHz) δ (ppm), 4.41 (s, 2H, $\text{CH}_2\text{-N}$); 5.00 (bs, 1H, NH); 6.57 (d, $J = 8$ Hz, 1H, H-6); 6.66-6.68 (m, 1H, H-4); 7.27 (d, $J = 8$ Hz, 2H, H-2', H-6'); 7.31-7.35 (m, 3H, H-5, H-3', H-5'); 7.42 (m, 1H, H-3).

Preparation of 4-((3-(benzyloxy)phenyl)amino)benzonitrile (**109**)



Procedure

A 25 mL round-bottomed flask specially equipped with a screw cap for coupling reactions, previously flame-dried under argon, was charged with 3-(benzyloxy)aniline (**108**), (0.4 g, 2.00 mmol, 1 eq), 4-bromobenzonitrile (**110**), (0.365 g, 2.00 mmol, 1 eq), caesium carbonate (0.785 g, 2.40 mmol, 1.2 eq), (\pm)-BINAP and palladium (II) complex in catalytic quantities in 10 mL of toluene. The reaction mixture was heated at $160 \pm 10^\circ\text{C}$ under constant stirring for 24 hours.

Work-up

The reaction mixture was diluted with water (20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was evaporated under reduced pressure.

Purification

No purification by column chromatography was performed since the product was obtained with enough purity (100%) as shown by TLC and ^1H spectrum.

Yield: 100%

Analytical data

- R_f : 0.94 (Hexane/Ethyl Acetate (1:1))
- Melting point: $131\text{-}133^\circ\text{C}$ (Dichloromethane)

Work-up

At first, the ACN was evaporated under reduced pressure, crystallization with methanol was carried out and the solvent was finally removed.

The reaction mixture was diluted with water (3 X 20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was evaporated under reduced pressure.

Purification

The crude reaction was purified by column chromatography on silica gel eluting with mixtures of hexane/ethyl acetate and methanol with increasing polarity. The desired compound was not obtained under these conditions.

Procedure B

A 100 mL round-bottomed flask with a stirring bar, equipped with a reflux condenser was charged with chlorhydrate hydroxylamine (0.70 g, 10 mmol, 2 eq), methyl acetoacetate (1.07 mL, 10 mmol, d = 1.076 g/mL, 2 eq) dissolved in 2 mL of ACN, all added in room temperature. After 2 hours, the starting aldehyde **93** (0.53 mL, 5 mmol, d = 1.157 g/mL, 1 eq) and ammonium acetate (1.5 g, 20 mmol, 2 eq) were dissolved in 5 mL of ACN and added to the reaction mixture. Then, the reaction mixture was stirred at room temperature for 48 hours.

Work-up

At first, the ACN was evaporated under reduced pressure, crystallization with methanol was carried out and the solvent was finally removed.

The reaction mixture was diluted with water (3 X 20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the combined organic layers were dried over anhydrous Na₂SO₄, filtered under vacuum and the solvent was evaporated under reduced pressure.

Purification

A ¹H spectrum was carried out. The desired product was not obtained.

Procedure C

A 100 mL round-bottomed flask with a stirring bar, equipped with a reflux condenser and cooled down 0 ± 10 °C with an ice-bath, was charged with chlorhydrate hydroxylamine (0.70 g, 10 mmol, 2 eq), methyl acetoacetate (1.07 mL, 10 mmol, $d = 1.076$ g/mL, 2 eq), the starting aldehyde **93** (0.53 mL, 5 mmol, $d = 1.157$ g/mL, 1 eq) and finally ammonium acetate (1.5416 g, 20 mmol, 2 eq) dissolved in 5 mL of ACN. Then, the reaction mixture was stirred and heated at 90 ± 10 °C during 5 days.

Work-up

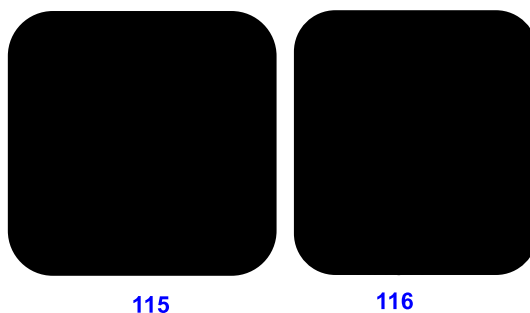
At first, the ACN was evaporated under reduced pressure, crystallization with methanol was carried out and the solvent was finally removed.

The reaction mixture was diluted with water (3 X 20 mL) and extracted with ethyl acetate (3 X 20 mL). Then, the combined organic layers were dried over anhydrous Na_2SO_4 , filtered under vacuum and the solvent was evaporated under reduced pressure.

Purification

A ^1H spectrum was realized. The desired compound was not obtained since a subproduct was obtained (**115** and **116**).

The analytical data for the subproducts obtained are the following:



- Yield: **54%** (mixture)
- Aspect: red semisolid

- ^1H NMR (DMSO- d_6 , 400 MHz) δ (ppm), 1.86 (s, 3H, CH_3 -); 1.91 (s, 3H, CH_3 -); 2.79 (t, $J = 12$ Hz, 1H, Ar-CH-); 3.61 (dd, $J_1 = 3$, $J_2 = 12$ Hz, 2H, -CH-C=O (x2)); 7.13 (d, $J = 8.8$ Hz, 2H, H-3', H-5'); 7.22 (d, $J = 8.8$ Hz, 2H, H-2', H-6') (A); 7.27-7.30 (m, 2H, H-3', H-5') (B); 7.62 (dd, $J_1 = 5.6$, $J_2 = 8.8$ Hz, 2H, H-2', H-6') (B); 8.20 (s, 2H, NH); 10.55 (s, 1H, N-OH).

A: majority isomer / B: minority isomer

- HRMS ESI (-) m/z : calculated mass for $\text{C}_{15}\text{H}_{16}\text{N}_4\text{O}_7\text{F}$ 3823.1003, found 383.1000.

4.3. Biochemical laboratory materials and methods

4.3.1. Biological activity evaluated at the Istituto Superiore di Sanità (Rome)

4.3.1.1. Cell cultures

Human endothelial cells obtained from the umbilical vein (HUVEC, Clonetics / Lonza, Basel, Switzerland), were amplified in 175 cm^2 flasks (Corning, Amsterdam, The Netherlands), in EBM-2 medium (Lonza) with the addition of the following growth factors (Lonza): 2% fetal bovine serum, fibroblast growth factor (FGF-2), heparin, vascular endothelial growth factor (VEGF), insulin-like growth factor (IGF), ascorbic acid, epidermal growth factor (EGF-2), gentamicin, amphotericin B, hydrocortisone (endothelial cell growth medium, EGM-2). Cell cultures were maintained at 37 °C in a humidified atmosphere with 5% CO_2 /95% air. For the experiments, cells between the 4th and 5th passage were used.

4.3.1.2. Cell viability

HUVEC were cultured in 96 well plates in EGM-2 (growing conditions) or EBM-2 (starvation conditions), at a concentration of 5,000 cells/well, (50,000 cells/ml, semi-confluent condition), and were treated with compounds **21** and **38** at different concentrations (1 nM, 10 nM, 100 nM, 1 μM , 10 μM , 50 μM). Compounds **21** and **38** were previously dissolved in dimethylsulphoxide (Farmitalia Carlo Erba, MI, Italy) to obtain a concentration of 100 mM; this solution was further diluted with EGM-2 or EBM-2 to obtain the final concentrations used in cell viability assay. Treatment with vehicle alone (dimethylsulphoxide) or medium alone were used as controls. Each sample was tested in triplicate.

After 24 hours of treatment the number of viable cells was assessed by the use of the Cell Counting Kit 8 (Abcam, Cambridge, UK): the WST-8/CCK8 solution, 10 μ l/well, was added and cells were incubated for further 3 hours. At the end of the culture period the plate was read at 450nm using a plate reader (BIO-RAD, Hercules, CA, USA). The assay is based on the formation of a water-soluble tetrazolium salt; WST-8 / CCK8 tetrazolium salt is reduced by cellular dehydrogenases combined with PMS (Phenazine Methyl Sulfate) electron mediator to an orange *formazan* product that is soluble in tissue culture medium. The amount of *formazan* produced is directly proportional to the number of living cells (Figure 129).¹⁹⁹

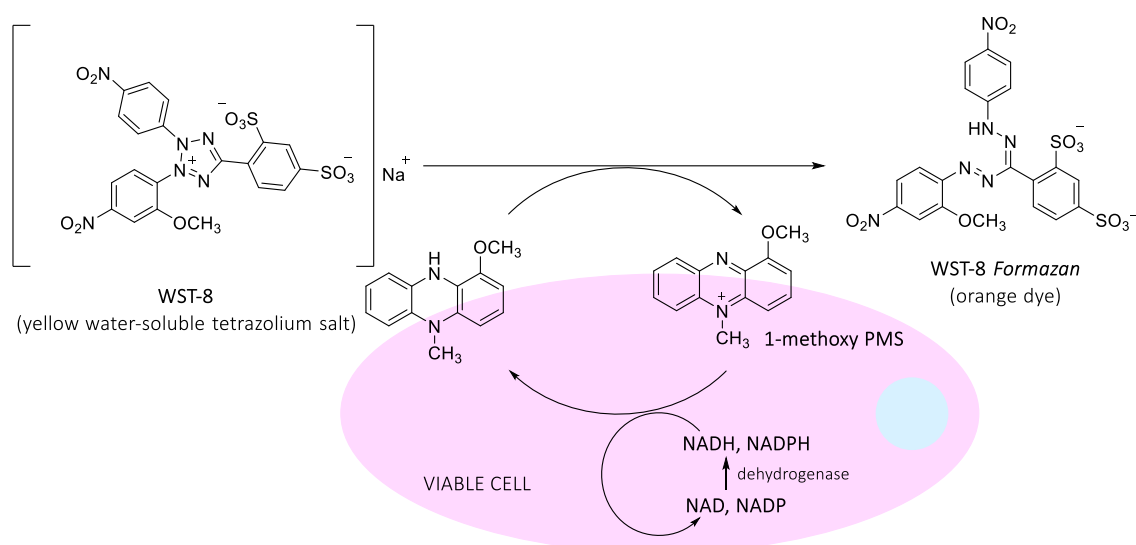


Figure 129. Cell viability detection mechanism of WST-8 / CCK8

4.3.2. Biological activity evaluated at the Hospital Clinic

4.3.2.1. Cell cultures

The methods used are part of a project developed at the Hospital Clinic of Barcelona under the direction of Neus Agell and the methodology is confidential.

4.3.2.2. Cell viability

The methods used are part of a project developed at the Hospital Clinic of Barcelona under the direction of Neus Agell and the methodology is confidential.

¹⁹⁹ <https://www.dojindo.eu.com/> (27/03/2020)

4.3.2.3. Cell signalling pathway

The methods used are part of a project developed at the Hospital Clinic of Barcelona under the direction of Neus Agell and it is a confidential methodology.

5. CONCLUSIONS

With the realization of this research work, relevant conclusions have been highlighted, as consequence of the molecular modeling studies, preparation of different synthetic routes, and not less important the evaluation of the biological activity of the compounds obtained.

5.1. Preparation of new inhibitory agents of PCSK9 protein involved in cholesterol homeostasis

5.1.1. Design of new enzyme inhibitors through molecular modeling studies

5.1.1.1. Search for new allosteric binding sites (Pockets)

- The joint use of multiple cMD and GaMD trajectories appears as a promising approach to study protein flexibility.
- A clustering approach applied in the PC space allows obtaining representatives that reflect the conformational diversity of the PCSK9 protein.
- Nearly 15 new binding sites were located, which opened the possibility to design PCSK9 inhibitors with an allosteric inhibition mechanism.

5.1.1.2. Screening of libraries (Docking)

- The docking performed on two allosteric binding sites of the PCSK9 protein was a good strategy in order to identify new inhibitors.
- From the proposed molecular docking it was possible to select from Selleck Drug[®] and Fragments[®] databases different compounds that could be good PCSK9 inhibitors.
- The main structural characteristics of the ligands for allosteric inhibition and the main interactions established with the protein were defined.

5.1.1.3. Rational design of inhibitors based on the structure of the ligands

- The use of the fdMD approach is a good method to determine the binding site of small organic molecules (fragments).
- The selection of initial fragments that allow growth from multiple points facilitates obtaining bigger compounds with good binding interactions.

- Different ligands with possible activity as PCSK9 inhibitors were obtained after modifying the initial fragments structures. Such ligands should be synthesized in order to evaluate their biological activity.

5.1.2. Preparation of pyrrolo[2,3-*d*]pyrimidines

- The preparation of 4-chloro-6-phenylpyrrolopyrimidine (**1**) from 4,6-dihydroxypyrimidine has been successful using a 5-step synthetic route in acceptable yields: a) dichlorination with POCl₃; b) amination; c) bromination by treatment with NBS; d) phenylacetylene addition under Sonogashira conditions; and, e) intramolecular cyclization. To highlight that step e) is limiting given that the product expected was obtained with **21%** yield. Analogue **2** (4-chloro-6-(4-chlorophenyl)pyrrolopyrimidine) was obtained with similar yields. The series of compounds with the *NH*-free group, the step c) changing the iodination for a bromination provided worse yields attributable to the lower stability of the iodinated derivative.
- In the synthesis of 4-chloro-7-methylpyrrolo[2,3-*d*]pyrimidines, the cyclization of intermediates **60** and **61** leads to the corresponding nucleus of pyrrolopyrimidines **3** (**99%**) and **4** (**98%**) with excellent yields. This fact suggests that the *NH*-free group makes it difficult to prepare this heterocyclic nucleus. However, intramolecular cyclization of the *N*-benzylated **64** derivative to the corresponding *N*-benzyl-4-chloro-6-phenylpyrrolo[2,3-*d*]pyrimidine **11** was also performed with **32%**.
- Treatment of pyrrolopyrimidines **3** and **4** with 4-nitroaniline, under conditions of the cross-coupling reaction catalyzed with Pd complexes, leads to diarylamines **5** and **6** respectively, with yields less than **10%**. Whereas when nitroaniline was added to intermediate acetylenes **60** and **61** pyrrolopyrimidines **5** and **6** were obtained directly with yields of **82** and **55%** respectively. This final strategy permits to reduce the total number of steps from 6 to 5.
- The formation of ureas **9** and **10** were successfully obtained by prior hydrogenation of nitro-derivatives **5** and **6**, and subsequent addition of the resulting anilines obtained **7** and **8** to 2-fluoro-5-trifluorophenylisocyanate.

The best conditions for the preparation of this ureas was to heat the mixture with THF at 90 °C.

- The arylation of *N*-benzyl-4-chloro-6-phenylpyrrolopyrimidine **11** with the *p*-methylboronic and *m*-methylboronic acids, following the Suzuki-Miyaura arylation conditions, leads to the desired product **12**, in the first case, with **23%** yield, while in the second case, compound **67** (**31%** yield) was obtained from the arylation at position 4 but *N*-disbenzylated.

5.1.3. Preparation of Combretastatin A4 (CA4) derivatives

- The 2,3-diarylacrylic acids **15**, **17**, **19**, **21**, **23**, **35**, **26**, **28**, and **30** have been obtained by condensation of the aryl aldehydes with the corresponding aryl acetic acids, through applying the Perkin reaction conditions, using anhydrid acetic acid, triethylamine and heating at 140 °C. The difficulty for the extraction and purification of the resulting acids leads to yields ranging from **13** to **100%**, (**15** (**13%**), **17** (**25%**), **19** (**13%**), **21** (**51%**), **23** (**81%**), **25** (**100%**), **26** (**32%**), **28** (**63%**) and **30** (**29%**)).
- Treatment of the carboxylic acid **28** with methanol and boron trifluoride etherate led to the derivative (*Z*)-1-(2,6-dichlorophenyl)-3-(hydroxy-4-methoxybenzylidene)indolin-2-one **82** (**14%**) instead of the methyl ester **29**. The boron trifluoride etherate activates the carboxylic acid and facilitates the attack of the aniline to form the indole derivative.

5.2. Preparation of new KRAS protein inhibitory agents

5.2.1. Optimization of potential KRAS inhibitors or activators from lead **P14**

5.2.1.1. Search of the binding site of the lead **P14** by performing gaussian accelerated molecular dynamics simulations (GaMD)

- Multiple GaMD calculations of the KRAS-**P14** system allows to the determination of a new binding mode for the **P14** ligand.
- The analysis of the interactions between **P14** and the KRAS protein in the new binding site is of special interest in order to synthesize new compounds that could fit in the same binding site but with better inhibition properties.

5.2.2. Preparation of KRAS inhibitors

- The [REDACTED] were successfully prepared with practically quantitative yields by using methyl acetoacetate, hydrated hydrazine and the corresponding aldehydes, in a multicomponent reaction, carried out in acetonitrile and in the presence of ammonium acetate, heating the resulting mixture to 90 °C .

5.3. Biological conclusions

5.3.1. Biological activity evaluated at the Istituto Superiore di Sanità (Rome)

- Compound **21** and **38** are useful in all tested concentrations as they do not significantly decrease the number of endothelial cells.
- The possible antioxidant activity of such compounds could not be studied due to COVID-19.

5.3.2. Biological activity evaluated at Hospital Clinic

- The results of the biological activity of the evaluated compounds derived from **P14; 41, 42, 46** and **47** demonstrate that compounds **42** and **47** showed activation of ERK and AKT, while derivative **46** showed inhibition of this same protein and compound **41** was inactive. The results do not allow to establish strong structure-activity relationships, but it is clear that the inactive derivative **41** has a 4-fluorophenyl group between the two pyrazoles, while **46** has a large group of benzyloxyphenylaminophenyl.
- The results obtained from the biological studies of the PDAC cell lines showed no effect of compounds **P14** and **47** in any of the six different cells tested in the cell survival analysis. However, **P14** showed ERK and AKT activation in the all PDAC cell lines, whereas compound **47** had no effect in any PDAC cell line tested.

5. CONCLUSIONS

Amb la realització d'aquest treball de recerca s'han obtingut conclusions rellevants, com a conseqüència de la modelització molecular, de la preparació de diferents rutes sintètiques, i també de l'avaluació biològica dels compostos obtinguts.

5.1. Preparació de nous agents inhibidors de la PCSK9 involucrats en l'homeòstasi del colesterol

5.1.1. Disseny de nous inhibidors enzimàtics mitjançant estudis de modelització molecular

5.1.1.1. Recerca de nous llocs d'unió al·lostèrics (Pockets)

- L'ús conjunt de múltiples trajectòries de cMD i GaMD resulta com un enfocament prometedor per estudiar la flexibilitat de la proteïna.
- El procés de clusterització aplicat a l'espai de PC permet obtenir les estructures representatives que reflecteixen la diversitat conformacional de la proteïna PCSK9.
- Al voltant de 15 nous llocs de unió van ser localitzats, obrint la possibilitat de dissenyar nous inhibidors de la PCSK9 amb un mecanisme de inhibició al·lostèric.

5.1.1.2. Screening of libraries (Docking)

- El procés de docking en dos llocs de unió al·lostèrics de la PCSK9 va ser una bona estratègia per tal de identificar nous inhibidors.
- A partir del procés de docking va ser possible seleccionar de la base de dades de Selleck Drug® i de Fragments® diferents compostos que podien ser bons inhibidors de la PCSK9.
- S'han definit les principals característiques estructurals dels lligands per inhibició al·lostèrica i les principals interaccions establertes amb la proteïna.

5.1.1.3. Disseny racional de inhibidors basat en l'estructura del lligand

- Un bon mètode per determinar el lloc de unió de molècules orgàniques petites (fragments) és mitjançant la utilització de fdMD.
- La selecció inicial dels fragments permet fer-los créixer des de múltiples punts facilitant la obtenció de compostos més grans amb millors interaccions de unió.

- Es va portar terme la obtenció de diferents lligands amb possible activitat inhibitoria enfront la PCSK9 després de modificar els fragments inicials. Aquests lligands haurien de ser sintetitzats per tal d'evaluar la seva activitat biològica.

5.1.2. Preparació de pyrrolo[2,3-*d*]pirimidines

- La preparació de la 4-cloro-6-fenilpirrolopirimidina (**1**) a partir de la 4,6-dihidroxipirimidina, ha estat possible mitjançant una ruta sintètica de 5 etapes amb rendiments acceptables: a) dicloració amb POCl₃, b) aminació, c) bromació per tractament amb NBS, d) addició de fenilacetilè sota condicions de Sonogashira i e) ciclació intramolecular. Cal destacar que l'etapa e) és limitant donat que es s'obté el producte esperat però amb un **21%** de rendiment. L'anàleg **2** (4-cloro-6-(4-clorofenil)pirrolopirimidina es va obtenir amb rendiments semblants. En aquesta sèrie amb el grup *NH*-lliure, el canvi de l'etapa c) per una iodació enlloc d'una bromació va proporcionar pitjors rendiments atribuïbles a la menor estabilitat del derivat iodat.
- En la síntesi de les 4-cloro-7-metilpirrolo[2,3-*d*]pirimidines, la ciclació dels intermediaris **60** i **61** van conduir als corresponents nuclis de pirrolopirimidines **3** (**99%**) i **4** (**98%**) amb rendiments excel·lents. Aquest fet posa de manifest que el grup *NH*-lliure dificulta la preparació d'aquest nucli heterocíclic. Així mateix, es va realitzar la ciclació intramolecular del derivat *N*-benzil **64** a la corresponent *N*-benzil-4-cloro-6-fenilpirrolo[2,3-*d*]pirimidina **11** amb un **32%**.
- El tractament de les pirrolopirimidines **3** i **4** amb la 4-nitroanilina, en les condicions de la reacció de cross-coupling catalitzada amb complexos de Pd, condueix a les diarilamines **5** i **6** respectivament, amb rendiment inferiors al **10%**. Mentre que quan la nitroanilina es va afegir als acetilens intermediaris **60** i **61** es van obtenir directament les pirrolopirimidines **5** i **6** amb rendiments del **82** i **55%** respectivament. Aquesta última estratègia sintètica suposa un estalvi al reduir-se el nombre total d'etapes de 6 a 5.
- Es van aconseguir amb èxit les urees **9** i **10** mitjançant hidrogenació prèvia dels nitro-derivats **5** i **6**, i posterior addició de les anilines resultants obtingudes **7** i **8** al 2-fluoro-5-trifluorofenilisocianat. Les millors condicions, per a la preparació de les urees, han estat escalfar la barreja amb THF a 90 °C.

- La arilació de l'*N*-benzil-4-cloro-6-fenilpirrolopirimidina **11** amb els àcids *p*-metilboronic i el *m*-metilboronic, seguint les condicions d'arilació de Suzuki-Miyaura, condueix al producte desitjat **12**, en el primer cas, amb un **23%** de rendiment, mentre que, en el segon cas, s'obté el compost **67** (**31%** de rendiment) procedent de l'arilació a la posició 4 però *N*-desbenzilat.

5.1.3. Preparació de derivats de la Combretastatin A4 (CA4)

- Els àcids 2,3-diaril acrílics **15**, **17**, **19**, **21**, **23**, **25**, **26**, **28** i **30** s'han obtingut per condensació dels aril aldehids amb els àcids aril acètics corresponents, aplicant les condicions de la reacció de Perkin, emprant anhídrid acètic, trimetilamina i escalfant a 140 °C. La dificultat per a l'extracció i purificació dels àcids resultants fa que s'obtinguin en rendiments que van del **13** al **100%** (**15** (**13%**), **17** (**25%**), **19** (**13%**) i **21** (**51%**), **23** (**81%**), **25** (**100%**), **26** (**32%**), **28** (**63%**) i **30** (**29%**)).
- El tractament de l'àcid carboxílic **28** amb metanol i trifluoro de boro eterat va conduir al derivat (*Z*)-1-(2,6-diclorofenil)-3-(3-hidroxi-4-metoxibenziliden)indolin-2-ona **82** (**14%**) enlloc de l'ester metílic **29**. El trifluoro de boro eterat activa l'àcid carboxílic i facilita l'atac de l'anilina formant així el derivat indòlic.

5.2. Preparació de nous agents inhibidors de KRAS

5.2.1. Optimització de potencials inhibidors de KRAS a partir del lead **P14**

5.2.1.1. Recerca del lloc d'unió del lead **P14** mitjançant simulacions de gaussian accelerated molecular dynamics (GaMD)

- Múltiples càlculs de GaMD del sistema KRAS-**P14** permet portar a terme la determinació de un nou lloc de unió del lligand **P14**.
- Conèixer les interaccions entre el **P14** i la proteïna és beneficiós per tal de sintetitzar nous compostos capaços de unir-se al mateix lloc de unió però amb millors propietats d'inhibició.

5.2.2. Preparació d'inhibidors de KRAS

- Els XXXXXXXXXX (**41-52**) es van preparar amb èxit i amb rendiments pràcticament quantitatius a partir d'acetoacetat de metil, hidrazina hidratada i els corresponents aldehids, en una reacció multicomponent, duta a terme en el si d'acetonitril i en presència d'acetat amònic, escalfant la mescla resultant a 90 °C.

5.3. Conclusions biològiques

5.3.1. Activitat biològica avaluada al Istituto Superiore di Sanità (Roma)

- El compost **21** i **38** són útils en totes les concentracions provades ja que no disminueixen significativament el nombre de cèl·lules endotelials.
- La possible activitat antioxidant d'aquests compostos no s'ha pogut estudiar a causa de la COVID-19

5.3.2. Activitat biològica avaluada a l'Hospital Clínic

- Els resultats de l'activitat biològica dels compostos avaluats derivats de **P14**, **41**, **42**, **46** i **47** mostren que els compostos **42** i **47** presenten activació d'ERK i AKT, mentre que el derivat **46** presenta inhibició d'aquesta mateixa proteïna i **41** fou inactiu. Els resultats no permeten establir unes relacions estructura-activitat fermes, però sí que és evident que el derivat inactiu **41** presenta un grup 4-fluorofenil entre els dos pirazoles, mentre el **46** presenta en la mateixa posició un grup voluminós de tipus benziloxifenilaminofenil.
- Els resultats obtinguts dels estudis biològics realitzats en línies cel·lulars de PDAC mostren que els compostos **P14** i **47** no tenen cap efecte a nivell de mort cel·lular en cap de les sis línies provades. No obstant, el compost **P14** va mostrar activació de ERK i AKT en línies cel·lulars de PDAC, mentre que el compost **47** no va mostrar efecte en cap de les sis línies cel·lulars.

6. BIBLIOGRAPHY

- ¹ L. Ose. *Tidsskr Nor Laegeforen* **2002**, *122*, 924-925.
- ² M. Pocovi, S. Castillo. *Cardiovascular Risk Factors* **2002**, *11*, 144156. DOI: [10.3305/nh.2015.32.6.9885](https://doi.org/10.3305/nh.2015.32.6.9885).
- ³ <https://ghr.nlm.gov/chromosome/19#resources> (03/04/2019)
- ⁴ P. Hopkins, P. Toth. *J. Clin. Lipidol.* **2011**, *5*, 9-17.
- ⁵ B. Nordestgaard, M. Chapman, S. Humphries. *Eur. Heart J.* **2013**, *34*, 3478-3490a.
- ⁶ M. Cuchel, E. Bruckert, H. Ginsberg. *Eur. Heart J.* **2014**, *35*, 2146-2157.
- ⁷ M. Benn, G. F. Watts, A. Tybjaerg-Hansen. *J. Clin. Endocrinol. Metab.* **2012**, *97*, 3956-3964.
- ⁸ B. Sjouke, D. Kusters, I. Kindt. *Eur. Heart J.* **2015**, *35*, 560-565.
- ⁹ T. L. Innerarity, K. H. Weisgraber, K. S. Arnold, R. W. Mahley, R. M. Krauss, G. L. Vega, S. M. Grundy. *Proceed. Nat. Aca. Sci. USA.* **1987**, *84*, 6919-6923.
- ¹⁰ L. F. Soria, E. H. Ludwig, H. R. Clarke, G. L. Vega, S. M. Grundy, B. J. McCarthy. *Proceed. Nat. Aca. Sci. USA.* **1989**, *86*, 587-591.
- ¹¹ N. B. Myant. *Atherosclerosis* **1993**, *104*, 1-18.
- ¹² J. R. Schaefer, H. Scharnagl, M. W. Baumstark, H. Schweer, L. A. Zech, H. Seyberth, K. Winkler, A. Steinmetz, W. März. *Arterioscler. Thromb. Vasc. Biol.* **1997**, *17*, 348-353.
- ¹³ X. Wang, X. Chen, X. Zhang, C. Su, M. Yang, W. He, Y. Du, S. Si, L. Wang, B. Hong. *EBioMedicine* **2020**, *52*, 102650. DOI: [10.1016/j.ebiom.2020.102650](https://doi.org/10.1016/j.ebiom.2020.102650).
- ¹⁴ A. N. Tsouka, C. C. Tellis, A. D. Tselepis. *Curr. Pharm. Des.* **2018**, *24*, 3622-3633.
- ¹⁵ T. F. Whayne. *Arch. Cardiol. Mexico* **2017**, *87*, 43-48.
- ¹⁶ Y.-W. Qian, R. J. Schmidt, Y. Zhang, S. Chu, A. Lin, H. Wang, X. Wang, T. P. Beyer, W. R. Bensch, W. Li. *J. Lipid Res.* **2007**, *48*, 1488-1498.
- ¹⁷ J. D. Horton, J. C. Choen, H. H. Hobbs. *J. Lipid Res.* **2009**, *50*, S172-S177.

- ¹⁸ N. G. Seidah. *Curr. Opin. Lipidol.* **2016**, *27*, 274-281.
- ¹⁹ E. R. Eden, R. P. Naoumova, J. J. Burden, M. I. McCarthy, A. K. Soutar. *Am. J. Hum. Genet.* **2001**, *68*, 653-660.
- ²⁰ C. K. Garcia, K. Wilund, M. Arca, G. Zuliani, R. Fellin, M. Maioli, S. Calandra, S. Bertolini, F. Cossu, N. Grishin, R. Barnes, J. C. Choen, H. H. Hobbs. *Science* **2001**, *292*, 1394-1398.
- ²¹ A. K. Soutar, R. P. Naoumova, L. M. Traub. *Arterioscler. Thromb. Vasc. Biol.* **2003**, *23*, 1963-1970.
- ²² M. H. Moghadasian, G. Salen, J. J. Frohlich, C. H. Scudamore. *Arch. Neurol.* **2002**, *59*, 527-529.
- ²³ D. Matías-Pérez, E. Pérez-Campos, I. A. García-Montalvo. *Nutr. Hosp.* **2015**, *32*, 2421-2426.
- ²⁴ A. C. Goldberg, P. N. Hopkins, P. P. Toth, C. M. Balantyne, D. J. Rader, J. G. Robinson, S. R. Daniels, S. S. Gidding, S. D. de Ferranti, M. K. Ito, M. P. McGowan, P. M. Moriarty, W. C. Cromwell, J. L. Ross, P. E. Ziajka. *J. Clin. Lipidol.* **2011**, *5*, S1-S8.
- ²⁵ P. J. Talmud, S. Shah, R. Whittall, M. Futema, P. Howard, J. A. Cooper, S. C. Harrison, K. Li, F. Drenos, F. Karpe, H. A. W. Neil, O. S. Descamps, C. Langenberg, N. Lench, M. Kivimaki, J. Whittaker, A. D. Hingorani, M. Kumari, S. E. Humphries. *Lancet.* **2013**, *381*, 1293-1301.
- ²⁶ G. F. Watts, S. Gidding, A. S. Wierzbicki, R. Alonso, W. V. Brown, E. Bruckert, K. K. Lin, P. Mata, K. G. Parhofer, F. J. Raal, R. D. Santos, W. G. Simpsons. *J. Clin. Lipidol.* Vol. 8, **2014**, 148-172.
- ²⁷ E. Youngblom, M. Pariani, J. W. Knowles. *Gene Rev.* **2014**, 4-7.
- ²⁸ M. A. Austin, C. M. Hutter, R. L. Zimmern, S. E. Humphries. *Am. J. Epidemiol.* **2004**, *160*, 407-420.
- ²⁹ A. Haase, A. C. Goldberg. *Curr. Opin. Lipidol.* **2012**, *23*, 282-289.
- ³⁰ D. From Marks, M. Thorogood, H. A. Neil, S. E. Humphries. *Atherosclerosis* **2003**, *168*, 1-14.
- ³¹ F. J. Raal, R. D. Santos. *Atherosclerosis* **2012**, *223*, 262-268.

- ³² A. Neil, J. Cooper, J. Betteridge, N. Capps, I. McDowell, P. Durrington, M. Seed, S. E. Humphries. *Eur. Heart J.* **2008**, *29*, 2625-2633.
- ³³ N. J. Stone, J. Robinson, A. H. Lichtenstein, C. N. B. Merz, C. B. Blum, R. H. Eckel, A. C. Goldberg, D. Gordon, D. Levy, D. M. Lloyd-Jones, P. McBride, J. S. Schwartz, S. T. Shero, S. C. Smith, K. Watson, P. W. F. Wilson. *J. Am. Coll. Cardiol.* **2014**, *63*, 2889-2934.
- ³⁴ M. Ito, M. McGowan, P. Moriarty. *J. Clin. Lipidol.* **2011**, *5*, S38-S45.
- ³⁵ F. J. Raal, G. J. Pilcher, V. R. Panz, H. E. van Deventer, B. C. Brice, D. J. Blom, A. D. Marais. *Circulation* **2011**, *124*, 2202-2207.
- ³⁶ V. E. Boucharie, A. C. Goldberg. *Cardiol. Clin.* **2015**, *33*, 169-179.
- ³⁷ <https://pubchem.ncbi.nlm.nih.gov/compound/Ezetimibe> (13/04/2019)
- ³⁸ C. Gagne, D. Gauder, E. Bruckert. *Circulation* **2002**, *105*, 2469-2475.
- ³⁹ R. Huijgen, E. Abbink, E. Bruckert, A. F. H. Stalenhoef, B. P. M. Imholz, P. N. Durrington, M. D. Trip, M. Eriksson, F. L. J. Visseren, J. R. Schaefer, J. J. P. Kastelein. *Clin. Ther.* **2010**, *32*, 615-625.
- ⁴⁰ F. Zieve, M. Kalin, S. Schwartz. *Clin. Ther.* **2007**, *29*, 74-83.
- ⁴¹ J. Malo, A. Parajuli, S. W. Walker. *Ann. Clin. Biochem.* **2020**, *1*, 7-25.
- ⁴² Y. J. Shimada, C. P. Cannon. *Eur. Heart J.* **2015**, *36*, 2415-2424.
- ⁴³ M. H. Davidson. *J. Clin. Lipidol.* **2013**, *3*, S11-S15.
- ⁴⁴ S. Xu, S. Luo, Z. Zhu, J. Xu. *Eur. J. Med. Chem.* **2019**, *162*, 212-233.
- ⁴⁵ M. Cuchel, L. Bloedon, P. O. Szapary, D. M. Kolansky, M. L. Wolfe, A. Sarkis, J. S. Millar, K. Ikewaki, E. S. Siegelman, R. E. Gregg, D. J. Rader. *N. Engl. J. Med.* **2007**, *2*, 148-156.
- ⁴⁶ E. Stein, R. Dufour, C. Gagne. *Circulation* **2012**, *19*, 2283-2292.
- ⁴⁷ F. J. Raal, R. D. Santos, D. J. Blom. *Lancet.* **2010**, *375*, 998-1006.
- ⁴⁸ G. R. Thompson, A. Catapano, S. Saheb. *Curr. Opin. Lipidol.* **2010**, *21*, 492-498.
- ⁴⁹ J. C. Choen, E. Boerwinkle, T. H. Mosley, H. H. Hobbs. *N. Eng. J. Med.* **2006**, *354*, 1264-1272.

- ⁵⁰ E. Stein, N. Honarpour, S. Wasserman. *Circulation* **2013**, *128*, 2113-2120.
- ⁵¹ E. Stein, S. Mellis, G. Yancopoulos. *N. Eng. J. Med.* **2012**, *366*, 1108-1118.
- ⁵² R. M. DeVay, D. L. Shelton, H. Liang. *J. Biol. Chem.* **2013**, 10805-10818.
- ⁵³ S. Poirier, G. Mayer. *Drug Des. Devel. Ther.* **2013**, *7*, 1135-1148.
- ⁵⁴ L. Wierod, J. Cameron, T. B. Strom, T. P. Leren. *Mol. Gen. Metab. Report* **2016**, 86-93.
- ⁵⁵ A. S. Peterson, L. G. Fong, S. G. Young. *J. Lipid Res.* **2008**, *6*, 1152-1156.
- ⁵⁶ D. E. Piper, S. Jackson, Q. Liu, S. T. Thibault, B. Shan, N. P. C. Walker. *Structure* **2007**, 545-552.
- ⁵⁷ <http://www.who.int> (07/03/2020)
- ⁵⁸ <http://www.cancer.gov> (07/03/2020)
- ⁵⁹ <http://www.cancerresearchuk.org> (08/03/2020)
- ⁶⁰ <http://www.gco.iarc.fr> (08/03/2020)
- ⁶¹ <http://www.ine.es> (08/03/2020)
- ⁶² <http://www.seom.org> (08/03/2020)
- ⁶³ M. Porru, L. Pompili, C. Caruso, A. Biroccio, C. Leonetti. *J. Exp. Clin. Cancer Res.* **2018**, 37-57.
- ⁶⁴ Y. Pylayeva-Gupta, E. Grabocka, D. Bar-Segi. *Nat. Rev. Cancer* **2011**, *11*, 761-774.
- ⁶⁵ A. D. Cox, C. J. Der. *Small GTPases* **2010**, *1*, 2-27.
- ⁶⁶ B. J. Grant, S. Lukman, M. J. Hocker, J. Sayyah, J. H. Brown, J. A. McCammon, A. A. Gorfe. *PLoS One* **2011**, *6*, 10.
- ⁶⁷ G. Buhrman, C. O'Connor, B. Zerbe, B. M. Kearney, R. Napoleon, E. A. Kovrigina, S. Vajda, D. Kozakov, E. L. Kovrigin, C. Mattos. *J. Mol. Biol.* **2011**, *4*, 773-789.
- ⁶⁸ P. Prakash, J. F. Hancock, A. A. Gorfe. *Proteins* **2015**, *83*, 898-909.
- ⁶⁹ A. A. Samatar, P. I. Poulikakos. *Nat. Rev Drug Discov.* **2014**, *13*, 928-942.
- ⁷⁰ A. Mullard. *Nat. Rev. Drug Discov.* **2019**, *18*, 887-891.

- ⁷¹ P. A. Konstantinopoulos, M. V. Karamouzis, A. G. Papavassiliou. *Nat. Rev. Drug Discov.* **2007**, *6*, 541-555.
- ⁷² M. Burotto, V. L. Chiou, J.-M. Lee, E. C. Kohn. *Cancer* **2014**, *120*, 3446-3456.
- ⁷³ E. Puxeddu, C. Durante, N. Avenia, S. Filetti, D. Russo. *Trends Endocrinol. Metab.* **2008**, *19*, 138-145.
- ⁷⁴ P. Liu, H. Cheng, T. M. Roberts, J. J. Zhao. *Nat. Rev. Drug Discov.* **2009**, *8*, 627-644.
- ⁷⁵ M. Cully, H. You, A. Levine, T. W. Mak. *Nat. Rev. Cancer* **2006**, *6*, 184-192.
- ⁷⁶ A. Takashima, D. V. Faller. *Expert Opin. Ther. Targets* **2013**, *17*, 507-531.
- ⁷⁷ S. Gysin, M. Salt, A. Young, F. McCormick. *Genes Cancer* **2011**, *2*, 359-372.
- ⁷⁸ I. R. Vetter, A. Wittinghofer. *Science* **2001**, *294*, 1299-1304.
- ⁷⁹ A. Takashima, D. V. Faller. *Expert Opin. Ther. Targets* **2013**, *17*, 507-531.
- ⁸⁰ R. Nazarian, H. Shi, Q. Wang, X. Kong, R. C. Koya, H. Lee, Z. Chen, M. K. Lee, N. Attar, H. Sazegar, T. Chodon, S. F. Nelson, G. McArthur, J. A. Sosman, A. Ribas, R. S. Lo. *Nature* **2010**, *468*, 973-977.
- ⁸¹ <http://www.chemdrug.com> (05/05/2016)
- ⁸² R. Dienstmann, J. Rodon, V. Serra, J. Tabernero. *Mol. Cancer Ther.* **2014**, *13*, 1021-1031.
- ⁸³ A. D. Cox, S. W. Fesik, A. C. Kimmelman, J. Luo, C. J. Der. *Nat. Rev. Drug Discov.* **2018**, *11*, 828-851.
- ⁸⁴ J. Malo, A. Parajuli, S. W. Walker. *Ann. Clin. Biochem.* **2020**, *1*, 7-25.
- ⁸⁵ T. Yamamoto, C. Lu, R. O. Ryan. *J. Biol. Chem.* **2011**, 5464-5470.
- ⁸⁶ S. Xu, S. Luo, Z. Zhu, J. Xu. *Eur. J. Med. Chem.* **2019**, 212-233.
- ⁸⁷ H.-M. Gu, A. Adijiang, M. Mah, D.-W. Zhang. *J. Lipid Res.* **2013**, *12*, 3345-3357.
- ⁸⁸ D.-K. Min, H.-S. Lee, N. Lee, C. J. Lee, H. J. Song, G. E. Yang, D. Yoon, S. W. Park. *Yonsei Med. J.* **2015**, *5*, 1251-1257. DOI: [10.3349/ymj.2015.56.5.1251](https://doi.org/10.3349/ymj.2015.56.5.1251).
- ⁸⁹ B. J. Evison, J. T. Palmer, G. Lambert, H. Treutlein, J. Zeng, B. Nativel, K. Chemello, Q. Zhu, J. Wang, Y. Teng, W. Tang, Y. Xu, A. K. Rathis, S. Kumar, A. K. Suchowerska, J. Parmer, I. Dixon, G. E. Kelly, J. Bonnar. *Bioorg. Med. Chem.* **2020**, *28*, 115344. DOI: [10.1016/j.bmc.2020.115344](https://doi.org/10.1016/j.bmc.2020.115344).
- ⁹⁰ C. Lammi, J. Sgrignani, A. Arnoldi, G. Lesma, C. Spatti, A. Silvani, G. Grazioso. *J. Med. Chem.* **2019**, *62*, 6163-6174.

- ⁹¹ J. Taechalestpaisarn, B. Zhao, X. Liang, K. Burgess. *J. Am. Chem. Soc.* **2018**, *140*, 3242-3249.
- ⁹² V. Prieur. Pyrrolo[2,3-*d*]pyrimidines: Conception, Synthèse, Fonctionnalisation. *Doctoral Thesis*. **December 2015**. University of Barcelona (Spain) and University of Orleans (France).
- ⁹³ R. Vardanyan, V. Hruby. *Synthesis of Best-Seller Drugs. Academic Press. 1st Edition*. **2016**, 495-547. ISBN: 9780124115248.
- ⁹⁴ M. Asif. *Org. Med. Chem.* **2017**, *2*, 555598. DOI: [10.19080/OMCIJ.2017.02.555598](https://doi.org/10.19080/OMCIJ.2017.02.555598).
- ⁹⁵ L. M. Greene, M. J. Meegan, D. M. Zisterer. *J. Pharmacol. Exp. Ther.* **2015**, *355*, 212-222.
- ⁹⁶ L. Vincent, P. Kermani, L. M. Young, J. Cheng, F. Zhang, K. Shido, G. Lam, H. Bompais-Vincent, Z. Zhu, D. J. Hicklin, P. Bohlen, D. J. Chaplin, C. May, S. Rafii. *J. Clin. Invest.* **2005**, *115*, 2992-3006.
- ⁹⁷ R. Singh, H. Kaur. *Synthesis* **2009**, *15*, 2471-2491.
- ⁹⁸ B. R. Hearn, S. J. Shaw, D. C. Myles. *Comprehensive Medicinal Chemistry* Vol. 6, Ed. Elsevier **2007**, 81-110.
- ⁹⁹ R. Mikstacka, T. Stefanski, J. Rózanski. *Cel. Mol. Biol. Lett.* **2013**, *18*, 368-397.
- ¹⁰⁰ L. Grau Valls. Disseny, síntesi i evaluació biològica de nous compostos potencialment antitumorals per inhibició enzimàtica. *Doctoral Thesis*. **June 2017**. University of Barcelona (Spain).
- ¹⁰¹ E. Garrido. New strategies of inhibition of K-Ras and study of the K-Ras/CaM interaction. *Doctoral Thesis*. **September 2017**. University of Barcelona (Spain).
- ¹⁰² C. I. Schroeder, J. E. Swedberg, J. M. Withka, J. Rosengren, M. Akcan, D. J. Clayton, N. L. Daly, O. Cheneval, K. A. Borzilleri, M. Griffor, I. Stock, B. Colless, P. Walsh, P. Sunderland, A. Reyes, R. Dullea, M. Ammirati, S. Liu, K. F. McClure, M. Tu, S. K. Bhattacharya, S. Liras, D. A. Price, D. J. Craik. *Chem. Biol.* **2014**, *21*, 284-294.
- ¹⁰³ D. Kozakov, E. L. Grove, D. R. Hall, T. Bohnuud, S. Mottarella, L. Luo, B. Xia, D. Beglov, S. Vajda. *Nat. Protoc.* **2015**, *10*, 733-755.
- ¹⁰⁴ C. H. Ngan, T. Bohnuud, S. E. Mottarella, D. Beglov, E. A. Villar, D. R. Hall, D. Kozakov, S. Vajda. *Nucleic Acids Res.* **2012**, 271-275.
- ¹⁰⁵ J. Jimenez, S. Doerr, G. Martínez-Rosell, A. S. Rose, G. de Fabritiis. *Bioinformatics* **2017**, *19*, 3036-3042.

- ¹⁰⁶ H. A. Mussein, A. Borrel, C. Geneix, M. Petitjean, L. Regad, A.-C. Camproux. *Nucleic Acids Res.* **2015**, *43*, 436-442.
- ¹⁰⁷ A. Borrel, L. Regad, H. G. Xhaard, M. Petitjean, A.-C. Camproux. *J. Chem. Inf. Model.* **2015**, *55*, 882-895.
- ¹⁰⁸ T. Krotzky, T. T. Rickmeyer, T. Fober, G. Klebe. *J. Chem. Inf. Model.* **2014**, *54*, 3229-3237.
- ¹⁰⁹ A. Krasowski, D. Muthas, A. Sarkar, S. Schmitt, R. Brenk. *J. Chem. Inf. Model.* **2011**, *51*, 2829-2842.
- ¹¹⁰ P. Schmidtke, X. Barril. *J. Med. Chem.* **2010**, *53*, 5858-5867.
- ¹¹¹ V. L. Guilloux, P. Schmidtke, P. Tuffery. *Bioinformatics* **2009**, *10*, 168.
- ¹¹² P. Schmidtke, V. L. Guilloux, J. Maupetit, P. Tuffery. *Nucleic Acids Res.* Vol. 38, **2010**, *2*, 582-589.
- ¹¹³ K. Sonogashira, Y. Tohda, N. Hagihara. *Tetrahedron Lett.* **1975**, *16*, 4464-4470.
- ¹¹⁴ D. H. Lee, H. Qiu, M. H. Cho, I. M. Lee, M. J. Jin. *Synlett* **2008**, 1657-1660.
- ¹¹⁵ C. W. D. Gallop, M. T. Chen, O. Navarro. *Org. Lett.* **2014**, *16*, 3724-3727.
- ¹¹⁶ S. Handa, J. D. Smith, Y. Zhang, B. S. Takala, F. Gallau, B. H. Lipshutz. *Org. Lett.* **2018**, *20*, 542-545.
- ¹¹⁷ J. E. Baldwin. *J. Chem. Soc. Chem. Commun.* **1976**, 734-736.
- ¹¹⁸ C. González Montero, F. J. Manso Platero, A. J. López Alba. *Avances en Odontología* ISSN, **2014**, 2340-3152.
- ¹¹⁹ J. L. Vialuc, F. Calmel, D. Bigg, E. Carilla, A. Stenger, P. Chopin, M. Briley. *J. Med. Chem.* **1994**, *37*, 689-695.
- ¹²⁰ P. S. Cohen, S. M. Cohen. *J. Chem. Educ.* **1996**, *73*, 883.
- ¹²¹ M. Leung, J.-L. Lai, K.-H. Leu, H.-H. Yu, H.-J. Hsiao. *J. Org. Chem.* **1996**, *61*, 4175-4179.
- ¹²² S. R. Jagtap, Y. P. Patil, A. Pauda, B. M. Bhanage. *Synth. Commun.* **2009**, *39*, 2093-2100.
- ¹²³ S. Kang, H.-K. Kim. *Tetrahedron* **2018**, *74*, 4036-4046.
- ¹²⁴ S. Hudson, M. Kiankarimi, M. W. Rowbottom, T. D. Vickers, D. Wu, J. Pontillo, B. Ching, W. Dwight, V. S. Goodfellow, D. Schwarz, C. E. Heise, A. Madan, J. Wen, W. Ban, H. Wang, W. S. Wade. *Biorg. Med. Chem.* **2006**, *16*, 4922-4930.
- ¹²⁵ a) Z. Hodnik, L. P. Masic, T. Tomasic. *J. Med. Chem.* **2014**, *57*, 4819-4833. b) B. Romano, D. Plano, I. Encio, J. A. Palop. *Biorg. Med. Chem.* **2015**, *23*, 1716-1727.

- ¹²⁶ a) D. Ke, C. Zhan, X. Li, A. D. Q. Li, J. Yao. *Tetrahedron* **2009**, *65*, 8269-8276. b) M. D. Mc Reynolds, K. T. Sprott, P. R. Hanson. *Org. Lett.* **2002**, *4*, 4673-4676.
- ¹²⁷ E. L. Luzina, A. V. Popov. *Eur. J. Med. Chem.* **2009**, *44*, 4944-4953.
- ¹²⁸ J. Clayden, L. Lemiègre, M. Helliiwell. *J. Org. Chem.* **2007**, *72*, 2302-2308.
- ¹²⁹ S. Porwanski, S. Menuel, X. Marsura, A. Marsura. *Tetrahedron Lett.* **2004**, *45*, 5027-5029 (and references cited therein).
- ¹³⁰ A. Suzuki. *Angew. Chem. Int. Ed.* **2011**, *50*, 6722-6737.
- ¹³¹ N. Miyaura, A. Suzuki. *Chem. Rev.* **1995**, *95*, 2457-2483.
- ¹³² J. P. Corbet, G. Mignani. *Chem. Rev.* **2006**, *106*, 2651-2710.
- ¹³³ C. H. Kuder, J. D. Neighbors, R. J. Hohl, D. F. Wiemer. *Bioorg. Med. Chem.* **2009**, *17*, 4718-4723.
- ¹³⁴ Y. Xu, S. Xu, L. Zhang, F. Zhang, Y. Ceui. *Lingxi Huagong* **2008**, *25*, 1083-1086. b) J. Tsuji. *Palladium Reagents and Catalysts: Innovations in Organic Synthesis*. Ed. Wiley: Chichester, **2004**. c) R. F. Heck. *J. Am. Chem. Soc.* **1968**, *90*, 5518-5526. d) R. F. Heck, J. P. Nolley. *J. Org. Chem.* **1972**, *37*, 2320-2322.
- ¹³⁵ F. Alonso, P. Riente, M. Yus. *Eur. J. Org. Chem.* **2009**, 6034-6042.
- ¹³⁶ N. Taha, Y. Sasson, M. Chidambaram. *Appl. Catal. A. General* **2008**, 217-224.
- ¹³⁷ J. E. McMurry, M. P. Fleming. *J. Am. Chem. Soc.* **1974**, *96*, 4708-4709.
- ¹³⁸ a) W. H. Perkin. *J. Chem. Soc.* **1868**, *21*, 53-61. b) G. Soladié, Y. Pooturel-Jocopé, J. Maignan. *Tetrahedron* **2003**, *59*, 18, 3315-3321.
- ¹³⁹ W. H. Perkin. *J. Chem. Soc.* **1877**, *32*, 660-674.
- ¹⁴⁰ R. Kitamura. *Rept. Japan. Assoc. Adv. Sci.* **1942**, *16*, 541-544.
- ¹⁴¹ A. Spasov, B. Kurtev. *Ann. Univ. Sofia, Fac. Phys.-Math.* **1947**, *43* (Livre 2), 37-51.
- ¹⁴² H. T. Clarke, *Chemistry of Penicillin*. Ed. Oxford University Press: Oxford **1949**, 688-848.
- ¹⁴³ F. H. Stodola. *J. Org. Chem.* **1948**, *13*, 757-762.
- ¹⁴⁴ P. E. Gagnon, J. L. Boivin, R. N. Jones. *Can. J. Res.* **1949**, *27B*, 190-204.
- ¹⁴⁵ C. V. Otis. *J. Soc. Motion Picture Engrs.* **1949**, *52*, 534-539.
- ¹⁴⁶ H. Dorn. *Chem. Heterocycl. Compd. USSR* **1981**, 3-31.
- ¹⁴⁷ R. M. Claramunt, Elgero. *J. Org. Proc. Prep. Int.* **1991**, *23*, 273-320.
- ¹⁴⁸ J. Steve, M. A. Horvat, J. H. Eds. Golob, *Nova Science Publishers, Inc.* **2008**, 129-182.
- ¹⁴⁹ G. Varvounis, Y. Iamegos, G. Pilidis. *Adv. Heterocycl. Chem.* **2001**, *80*, 75-165.

- ¹⁵⁰ T. Eicher, S. Hauptmann. *The Chemistry of Heterocycles*, 2nd Ed. **2003**.
- ¹⁵¹ J. Marchand-Brynaert, L. Ghosez, G. Lukacs, M. Ohno. *Springer* **1990**. DOI: [10.1007/978-3-642-75617-7_20](https://doi.org/10.1007/978-3-642-75617-7_20)
- ¹⁵² S. Hanessian, G. McNaughton-Smith, H.-G. Lombart, W. D. Lubell. *Tetrahedron* **1997**, *53*, 12789-12854.
- ¹⁵³ M. I. Konaklieva, B. J. Plotkin. *Curr. Med. Chem. Anti-Infect. Agents* **2003**, *2*, 287-302.
- ¹⁵⁴ H. L. White, J. L. Howard, B. R. Cooper, F. E. Soroko, J. D. McDermed, K. J. Ingold, R. A. Maxwell. *Neurochem.* **1982**, *39*, 271-271.
- ¹⁵⁵ K. Kawai, J. Hasegawa, H. Shiojiri, Y. Nozawa, K. Tsurumi, H. Fujimura. *Experientia* **1985**, *41*, 490-492.
- ¹⁵⁶ C. Cucurou, J. P. Battioni, D. C. Thang, N. H. Nam, D. Mansuy. *Biochemistry* **1991**, *30*, 8964-8970.
- ¹⁵⁷ E. M. Kosower, E. Hershkowitz. *Chem. Abstr.* **1995**, *122*, 214077.
- ¹⁵⁸ A. Novak, J. Bezensek, U. Groselj, A. Golobic, B. Stanovnik, J. Svete. *Arkivoc* **2011**, 18-28.
- ¹⁵⁹ A. Novak, M. Stefanic, U. Groselj, M. Hrast, M. Kasunic, S. Gobec, B. Stanovnik, J. Svete. *Tetrahedron* **2013**, *69*, 6648-6665.
- ¹⁶⁰ A. Novak, M. Stefanic, U. Groselj, M. Hrast, M. Kasunic, S. Gobec, B. Stanovnik, J. Svete. *Helv. Chim. Acta* **2014**, *97*, 245-267.
- ¹⁶¹ Y. Wu, Z. Huang, Y. Luo, D. Liu, Y. Deng, H. Yi, J.-F. Lee, C. W. Pao, J.-L. Chen, A. Lei. *Org. Lett.* **2017**, *19*, 2330-2333.
- ¹⁶² J. T. Yu, B. Shi, H. Peng, S. Sun, H. Chu, Y. Jiang, J. Cheng. *Org. Lett.* **2015**, *17*, 3643-3645.
- ¹⁶³ a) L. El Kaim, L. Grimaud. *Tetrahedron* **2009**, *65*, 2153-2171. b) R. W. Amstrong, A. P. Combs, P. A. Tempest, T. A. Keating. *Chem. Res.* **1996**, *29*, 123-131. c) T. J. J. Müller. *J. Org. Chem.* **2011**, *7*, 960-961.
- ¹⁶⁴ A. Dömling, W. Wang, K. Wang. *Chem. Rev.* **2012**, *112*, 3083-3135.
- ¹⁶⁵ S. Masumoto, H. Usuda, M. Suzuki, M. Kanai, M. Shibasaki. *J. Am. Chem. Soc.* **2003**, *125*, 5634-5635.
- ¹⁶⁶ A. Kumar, R. A. Maurya. *Synlett* **2008**, 883-885.
- ¹⁶⁷ X. Zhang, J. Kang, P. Niu, J. Wu, W. Yu. *J. Org. Chem.* **2014**, *79*, 10170-10178.

- ¹⁶⁸ S. S. Kamble, G. S. Shankarling. *Chem. Select.* **2018**, *3*, 2032-2036.
- ¹⁶⁹ M. Fatahpour, F. N. Sadeh, N. Hazeri, M. T. Maghsoodlou, M. Lashkari. *J. Iran Chem. Soc.* **2017**, *14*, 1945-1956.
- ¹⁷⁰ Z. Zhou, Y. Zhang. *J. Chil. Chem. Soc.* **2015**, *60*, 2992-2996.
- ¹⁷¹ J. Safaei-Ghoni, B. Khojastehbakht-Koopaei, H. Shahbazi-Alavi. *RSC. Adv.* **2014**, *4*, 46106-46113.
- ¹⁷² Z. Zhou, Y. Zhang. *Green Chem. Lett. Rev.* **2014**, *7*, 18-23.
- ¹⁷³ L. Navarro, M. D. Pujol. *Tetrahedron Lett.* **2015**, *56*, 1812-1816.
- ¹⁷⁴ A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede. *Nucleic Acids Res.* **2018**, *46*, W296-W303.
- ¹⁷⁵ The UniProt Consortium. UniProt: the universal protein knowledge base. *Nucleic Acids Res.* **2017**, *45*, D158-D169.
- ¹⁷⁶ <https://www.uniprot.org/uniprot/Q8NBP7> (03/04/2020)
- ¹⁷⁷ Schrödinger Release 2020-1: Maestro, Schrödinger, LLC, New York, NY, **2020**.
- ¹⁷⁸ W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein. *J. Chem. Phys.* **1983**, *79*, 926-935.
- ¹⁷⁹ T. Case, D. T. E. Cheatham III, C. Simmerling, A. Roitberg, J. Wang, W. A. Göetz, I. Kolossváry. *Amber 14 Reference Manual*, University of California, San Francisco.
- ¹⁸⁰ J. A. Maier, C. Martínez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling. *J. Chem. Theory Comput.* **2015**, *11*, 3696-3713.
- ¹⁸¹ T. Darden, D. York, L. Pedersen. *J. Chem. Phys.* **1993**, *98*, 10089-10092.
- ¹⁸² M. Chaichian, A. Demichev. *Path Integrals in Physics Stochastic Process and Quantum Mechanics*, Taylor and Francis, Vol. 1, **2001**, 174.
- ¹⁸³ M. R. Hestens, E. Stiefel. *J. Res. Natl. Bur. Stand.* **1952**, *49*, 409.
- ¹⁸⁴ M. P. Allen, D. J. Tildesley. *Computer Simulation of Liquids* (Second Ed.), Oxford University Press **2017**.
- ¹⁸⁵ J. P. Ryckaert, G. Ciuccotti, H. J. C. Berendsen. *J. Comput. Phys.* **1977**, *23*, 327-341.
- ¹⁸⁶ L. Rokach, O. Maimon. *Clustering Methods, Data Mining and Knowledge*, Eds: Springer, Boston, M.A., **2005**, 321-352.

- ¹⁸⁷ P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, T. E. Cheatham. *Acc. Chem. Res.* **2000**, *33*, 889-897.
- ¹⁸⁸ V. Tsui, D. A. Case. *Biopolymers* **2000**, *56*, 275-291.
- ¹⁸⁹ J. Weiser, P. S. Shenkin, W. C. Still. *J. Comp. Chem.* **1999**, *20*, 217-230.
- ¹⁹⁰ A. Onufriev, D. Bashford, D. A. Case. *Proteins Struct. Funct. Genet.* **2004**, *55*, 383-394.
- ¹⁹¹ MOE (The Molecular Operating Environment) Version 2009.10, software available from Chemical Computing Group Inc., 1010 Sherbrooke Street West, Suite 910, Montreal, Canada H3A 2R7.
- ¹⁹² D. Oranit, D. Schneidman-Duhovny, Y. Inbar, R. Nussinov, H. J. Wolfson. *J. Chem. Inf. Model.* **2009**, *49*, 2333-2343.
- ¹⁹³ C. R. Corbeil, C. I. Williams, P. Labute. *J. Comput. Mol. Des.* **2012**, *26*, 775-786.
- ¹⁹⁴ J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, A. D. Case. *J. Comput. Chem.* **2004**, *25*, 1157-1174.
- ¹⁹⁵ R. Luo, L. David, M. K. Gilson. *J. Comput. Chem.* **2002**, *23*, 1244-1253.
- ¹⁹⁶ D. A. Case, R. M. Betz, D. S. Cerutti, T. E. Cheatham III, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao and P. A. Kollman. AMBER **2016**, University of California, San Francisco.
- ¹⁹⁷ W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein. *J. Chem. Phys.* **1983**, *79*, 926-935.
- ¹⁹⁸ T. Maurer, L. S. Garrenton, A. Oh, K. Pitts, D. J. Anderson, N. J. Skelton, B. P. Fauber, B. Pan, S. Malek, D. Stokoe, M. J. Ludlam, K. K. Bowman, J. Wu, A. M. Giannetti, M. A. Starovasnik, I. Mellman, P. K. Jackson, J. Rudolph, W. Wang, G. Fang. *Proc. Natl. Acad. Sci.* **2012**, *109*, 5299-52304.
- ¹⁹⁹ <https://www.dojindo.eu.com/> (27/03/2020)