



UNIVERSITAT<sup>DE</sup>  
BARCELONA

# Design and synthesis of engineered peptides to target undruggable PPIs: from *in silico* to *in vitro* studies

Pep Rivas Santos

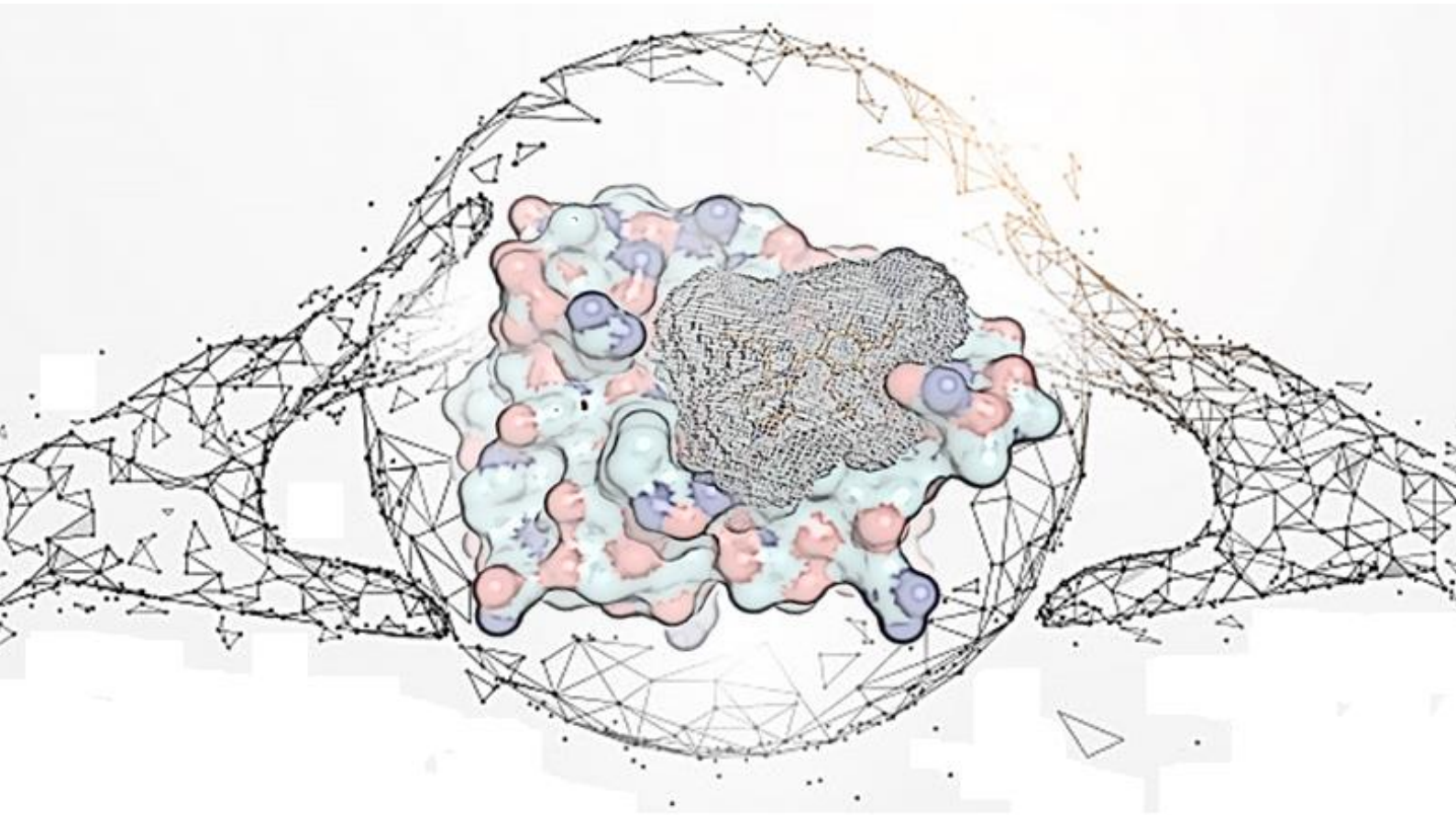


Aquesta tesi doctoral està subjecta a la llicència **Reconeixement 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution 4.0. Spain License.**

# Design and synthesis of engineered peptides to target undruggable PPIs: from *in silico* to *in vitro* studies



Pep Rivas Santos



UNIVERSITAT DE  
BARCELONA







*« (...) no proponemos más que soluciones que hayan sido demostradas por los hechos, teorías ratificadas por la razón y verdades confirmadas por la verificación de pruebas.*

*El objeto de nuestra enseñanza es que el cerebro del individuo llegue a ser el instrumento de su voluntad. Queremos que las verdades de la ciencia brillen por su propia luz e iluminen cada inteligencia de manera que llevadas a la práctica puedan dar a la humanidad, felicidad y bienestar sin exclusión de nadie y sin privilegios odiosos. »*

*Francesc Ferrer i Guardia*

*« (...) Weber ha aceptado mi propuesta para la tesis, y hasta se mostró muy satisfecho. Me alegro enormemente por las investigaciones que con este motivo tendré que desarrollar. »*

*Mileva Marić*





UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA I CIÈNCIES DE L'ALIMENTACIÓ

Dr. Roger Prades Cosano

Dr. Jesús Seco Moral

Design and synthesis of engineered peptides to target undruggable PPIs: from *in silico* to *in vitro* studies

Josep Rivas Santos, 2019





UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA I CIÈNCIES DE L'ALIMENTACIÓ

PROGRAMA DE DOCTORAT DE BIOTECNOLOGIA

Design and synthesis of engineered peptides to target undruggable PPIs: from *in silico* to *in vitro* studies

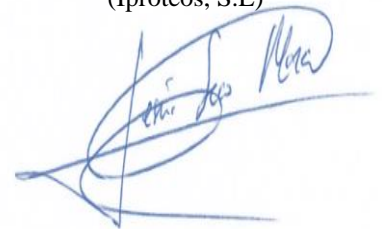
Memòria presentada per Josep Rivas Santos per optar al títol de doctor per la universitat de Barcelona

Tesi doctoral supervisada per:

Dr. Roger Prades Cosano  
(Iproteos, S.L)



Dr. Jesús Seco Moral  
(Iproteos, S.L)



Tutora:

Prof. Josefa Badia Palacín  
Universitat de Barcelona

Facultat de de Farmàcia i Ciències de l'Alimentació

Departament de Bioquímica i Fisiologia

Josep Rivas Santos, 2019



## Acknowledgments



Un altre cop torno a esborrar les línies que encapçalant aquest text, i és aleshores quan decideixo que em deixaré de formalismes, que finalitzar un projecte de vida, encara que sigui un punt i *seguit*, mereix uns agraïments sincers i que serveixen per fer saber tot allò que a vegades se'm fa tan difícil de dir de paraula. Viure és lluitar, per un mateix i per les persones que estimem. Doncs, aquesta tesi doctoral és el resultat de totes les lluites i esforços fets per molta gent que ha format o formen part de la meua vida i que sense elles hagués estat impossible arribar en aquest punt. Cada cop que hi penso apareixen nous noms, nous records, que em fan somriure i pensar “*buff... quina sort*”, però d'altra banda em dono per vençut en l'intent recopilar-los tots i opto per una via més pragmàtica i conformar-me, que no és poc, amb anomenar aquelles persones que d'alguna manera o altra han estat més directament relacionades amb aquesta tesi doctoral.

M'agradaria començar per les que van compartir amb mi el projecte d'Iproteos, són elles amb qui vaig estar dia a dia trencant-nos el cap per solucionar cada problema o per gaudir de cada èxit. Doncs, parlant d'Iproteos, gràcies **Teresa** per donar-me l'oportunitat en el seu moment de fer el TFM amb vosaltres i després el doctorat, però encara molt més agraït per la teua comprensió, **SEMPRE**, i ser tan pròxima amb tothom, al final es pot aprendre com a dirigir una empresa, però per tenir l'estima i el compromís de totes el les que la formen, com tu ho tenies, s'ha de ser bona persona, i això no s'aprèn, s'és. **Roger**, em considero molt afortunat per tot el que he après de tu, ets d'aquelles poques persones que tot el que diu és una nova lliçó, bàsicament ets un REFERENT, espero algun dia arribar al teu nivell, ets la persona més professional que mai he conegut. **Jesús**, sempre vas posar la meua formació i a mi per davant de tot, el respecte i el compromís no s'imposen, es guanyen, i jo ho tinc ben clar, tu tens tot el meu respecte i compromís. **Laura**, crec que inicialment ens va costar adaptar-nos l'un a l'altre, però ens em vam sortir!, i per mi vas acabar sent una persona molt important dins d'Iproteos i de qui sentia sempre el seu suport i confiança, gràcies. **Chiara**, sempre amb un somriure i una energia que transmet felicitat, els moments difícils són molt menys difícils amb tu! **Ana**, sense cap dubte una gran amiga a Iproteos, tant trist per no veure't cada matí quan vas marxar..., però tan content per tu, ets una crack! **Ariadna**, encara que sempre ens féssim la punyeta al lab i ocupessis sempre el meu espai, va ser una sort treballar amb tu,estic totalment convençut que arribaràs molt lluny, perquè no conec a ningú amb més capacitats i passió que tú. **Núria**, no sé que hagués fet sense tu, i el que tampoc sé, és com m'has aguantat i com m'aguantes, ompliria fulls i fulls i no seria suficient, així que on no entren paraules, de tot cor et dic GRÀCIES. **Sandra**, pensar que tu just començaves el doctorat quan ens vam conèixer... sempre t'admirat, tota la feina que vas arribar a fer, així i tot sempre que et vaig necessitar no vas dubtar ni un moment en ajudar-me, gràcies per ser una persona tan fàcil. **Ruben**, vaig aprendre molt de tu i la teua experiència, gràcies per tota la teua ajuda i consells. **Carlota**, ha sigut un plaer treballar amb tu, la teua tranquil·litat i eficiència a la feina no te preu, ets d'aquelles persones que saps que podria amb tot. Finalment, **Laura B**, demanar sempre una segona opinió, treballar conjuntament, preguntar, ajudar a les demes, són qualitats que tots podríem aprendre de tu.

No voldria pas oblidar-me d'agrair a tota la gent del group del Giralt; **Salva, Pep, Cris F., Edu, Martí, Cris G., Macarena, Cris D., Judith, Julia** per haver adoptat a tots els membres d'Iproteos com un més i mai tenir un no per resposta. Especial menció a l'**Adam**, qui a més d'un gran amic sempre troba el temps per ajudar a tothom sense

demanar absolutament res a canvi, es per això que si hi ha persones que són indispensables, Adam a tu aquesta paraula se't queda molt curta.

Agraïr també tota la gent de la planta del “subterrani”, **Dani, Omar, Gerardo, Juan, Ivan** i especialment a la **Sonia** i l'**Unai**, gràcies per ajudar-me cada vegada que ho he necessitat, que no han estat precisament poques. Però, sobretot gràcies per preocupar-vos sempre molt més enllà del que implicava oferir un servei.

Gràcies també a totes les investigadores amb les que he tingut la sort de col·laborar durant aquesta tesi, a la **Marina**, l'**Anabel-Lise** i la **Jenny** del grup del **Pere Roca**, a la **Dèbora** del grup de la **Neus Agell** i al **Josep Roma** i el **Gabriel** del VHIR.

Fora de l'àmbit del “Parc”, evidentment sé que tot el que ha passat durant els darrers anys no hagués estat possible sense les millors companyes de pis que es pot tenir, qui sempre m'han recolzat i han fet que la frustració que a vegades representa la ciència fos molt més fàcil de portar. Gràcies **Manu, Olga** i **Tània**. Llavors en aquest mateix sentit, a totes les amigues, **Andrea, Joan, Rafa, Imanol, Patri, Guille, Rebeca, Ira, Leti, Iñigo** i **Ramón**. També als que són els amics d'ençà que tinc memòria, l'**Alex**, el **Dani**, el **Borja** el **Franc**, el **Pablo**, el camí és llarg i cadascú tria el seu però l'estima d'uns pels altres ens manté units en la distància i el temps.

Finalment, a qui sempre he agraït tot el que he aconseguit, cada etapa que he finalitzat, a vosaltres, perquè sou les arrels, el tronc i les fulles de la meva vida. Gràcies **Júlia**, Gràcies **Mama**, Gràcies **Papa**, per donar-m'ho tot sense tenir res, per utilitzar l'estima per moneda i per ensenyar-me que la felicitat és no conformar-se.

# Content





## Table of Contents

<b>Abbreviations</b> .....	<b>2</b>
Abbreviations and acronyms.....	2
Natural and non-natural amino acids .....	5
Tail cappings.....	8
Coupling reagents and solid synthesis resins.....	9
<b>Chapter 1: Introduction</b> .....	<b>10</b>
1.1. Current status of the pharmaceutical field .....	12
1.2. Peptides.....	13
1.3. Peptidomimetics .....	17
1.4. Iproteos approach .....	20
<b>Objectives</b> .....	<b>20</b>
<b>Chapter 2: IPROTech</b> .....	<b>26</b>
2.1. Introduction .....	28
2.2. IPRO Library.....	30
2.3. IPRO Filter.....	31
2.4. IPRO Docking.....	32
2.5. IPRO permeability .....	35
2.6. Experimental section.....	36
2.7. Physicochemical and ADME evaluation .....	39
<b>Chapter 3: Disruption of the mechanical clutch Talin-Vinculin</b> .....	<b>44</b>
3.1 Introduction: Talin-Vinculin.....	46
3.2. IPROTech hit identification.....	47
3.3. Experimental evaluation .....	50
3.4. Conclusions .....	55
<b>Chapter 4: Targeting the RAD51-BRCA2 interaction in the DNA repair mechanism</b> .....	<b>58</b>
4.1. Introduction: RAD51-BRCA2.....	60
4.2. IPROTech hit identification.....	62
4.3. Experimental evaluation .....	67
4.4. Hit optimization.....	69
4.5. Conclusions .....	75
<b>Chapter 5: Targeting the protein downstream of K-RAS</b> .....	<b>76</b>
5.1. Introduction: KRAS-Effectors.....	78
5.2. IPROTech hit identification.....	80
5.3. Experimental evaluation .....	88
5.4. Hit optimization.....	90

5.5. Conclusions .....	95
<b>Chapter 6: Inhibition of the Retromer-L2 interaction to prevent Human Papillomavirus endosomal release .....</b>	<b>98</b>
6.1. Introduction: Retromer-L2 .....	100
6.2. IPROTech hit identification.....	101
6.3. Experimental evaluation .....	104
6.4. Conclusions .....	106
<b>Chapter 7: Discussion.....</b>	<b>108</b>
7.1. IPROTech .....	110
7.2. Talin-Vinculin.....	112
7.3. RAD51-BRCA2.....	113
7.4. RAS-Effectors.....	113
7.5. Retromer-L2 protein .....	114
<b>Conclusions .....</b>	<b>116</b>
<b>Materials and Methods.....</b>	<b>120</b>
1.1. General protocol for computational studies.....	122
1.2. General protocol for peptide synthesis.....	124
1.3. Liquid chromatography .....	127
1.4. Physicochemical properties experiments .....	128
1.5. Biophysical properties evaluation.....	130
1.6. Cell-based experiments.....	132
<b>References .....</b>	<b>134</b>
<b>Product characterization.....</b>	<b>148</b>
2.1. Talin-Vinculin inhibitors .....	150
2.3. RAD51-BRCA2 inhibitors .....	151
2.3. RAS-Effectors inhibitors .....	152
2.4. Retrome-L2 inhibitors .....	154

## *Abbreviations*



## **Abbreviations and acronyms**

ACN: Acetonitrile

ACH:  $\alpha$ -cyano-4-hydroxycinnamic acid

ADME: Absorption, Distribution, Metabolism, and Excretion

BBB: Blood-brain barrier

CNS: Central Nervous System

CRO: Contract Research Organization

DBU: 1,8-Diazabicyclo[5.4.0]undec-7-ene

DCM: Dichloromethane

DIAD: Diisopropyl azodicarboxylate

DIC: *N,N'*-Diisopropylcarbodiimide

DIEA: *N,N*-Diisopropylethylamine

DMEM: Dulbecco modified eagle medium

DMF: *N,N*-Dimethylformamide

DMSO: Dimethyl sulfoxide

DPPA: Diphenyl phosphoryl azide

EtOH: Ethanol

eq: Equivalent

ESI: Electro Spray Ionization

EtOAc: Ethyl acetate

FA: Formic acid

FDA: Food and Drug Administration

Fmoc: Fluorenylmethyloxycarbonyl chloride

HPLC: High Pressure Liquid Chromatography

HPLC-MS: High Pressure Liquid Chromatography-Mass Spectra

KCN: Potassium cyanide

GIT: Gastrointestinal tract

GLP: Good Laboratory Practice

ip: Intraperitoneal

iv: Intravenous

IC<sub>50</sub>: Half maximal inhibitory concentration

K<sub>i</sub>: Inhibitor constant

LC-MS/MS: Liquid chromatography coupled to tandem mass spectrometry

MALDI-TOF: Matrix-Assisted Laser Desorption/Ionization-Time-Of-Flight

MW: Molecular weight

NaHCO<sub>3</sub>: Sodium Bicarbonate

N/E: Not evaluated

NMR: Nuclear Magnetic Resonance

NR: Not Reported

*o*-NBS: 2-nitrobenzenesulfonyl chloride

PAMPA: Parallel Artificial Membrane Permeability Assay

Papp: Apparent permeability

PDA: Photodiode array

PDB: Protein Data Bank

Pe: Effective permeability

PEG: Polyethyleneglycol

P-gp: p-glycoprotein

PK: Pharmacokinetics

PVDF: Poly(vinylidene fluoride)

RMSD: Root mean-square deviation

rt: Room temperature

RT: Retention Time

t: Time

T: Transport

t<sub>1/2</sub>: Half-life time

TBTU: *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate

TEMED: Tetramethylethylenediamine

TFA: Trifluoroacetic acid

TIS: Triisopropylsilane

UHPLC: Ultra-High Performance Liquid Chromatography

UV: Ultraviolet

WB: Western blot

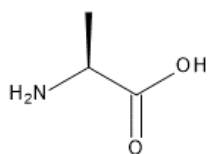
$\epsilon$ : Molar absorbance coefficient



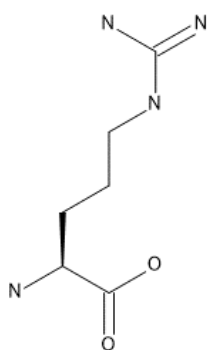
## Natural and non-natural amino acids

### Natural amino acids

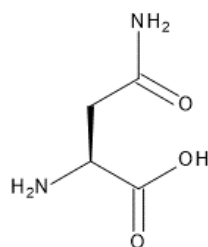
Ala  
L-Alanine



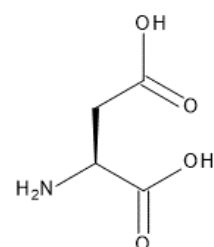
Arg  
L-Arginine



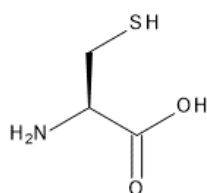
Asn  
L-Asparagine



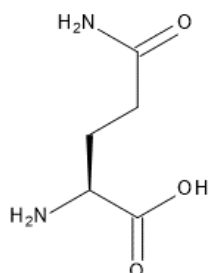
Asp  
L-Aspartic acid



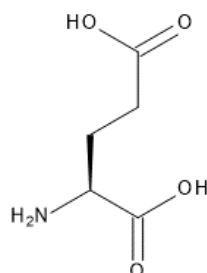
Cys  
L-Cys



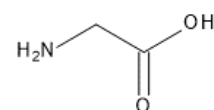
Gln  
L-Glutamine



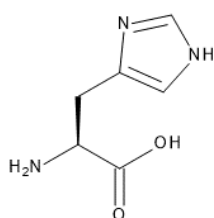
Glu  
L-Glutamic acid



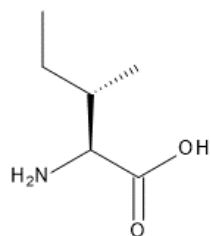
Gly  
L-Glycine



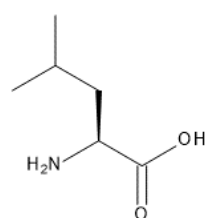
His  
L-Histidine



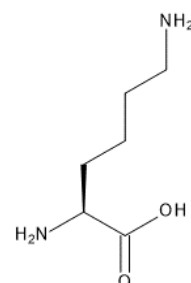
Ile  
L-Isoleucine



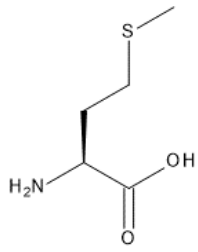
Leu  
L-Leucine



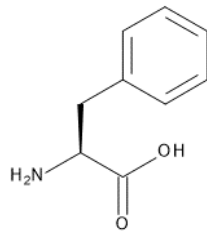
Lys  
L-Lysine



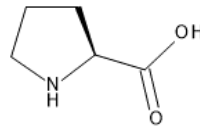
Met  
L-Methionine



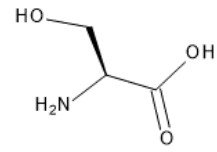
Phe  
L-Phenylalanine



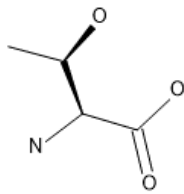
Pro  
L-Proline



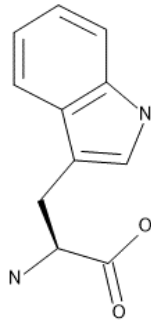
Ser  
L-Serine



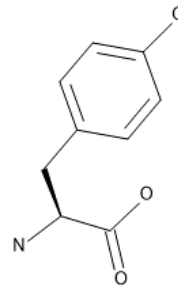
Thr  
L-Threonine



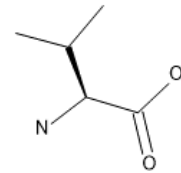
Trp  
L-Tryptophan



Tyr  
L-Tyrosine

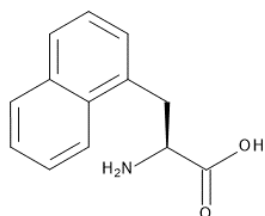


Val  
L-Valine

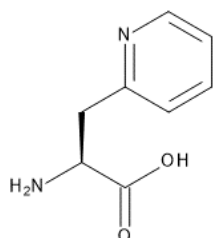


Non-natural amino acids

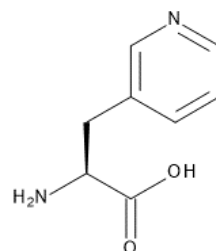
1-Nal  
L-1-Naphthylalanine



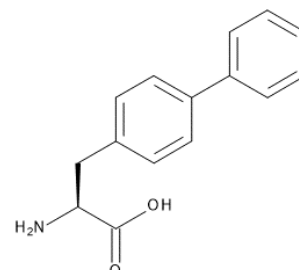
2-Pal  
L-2-Pyridylalanine



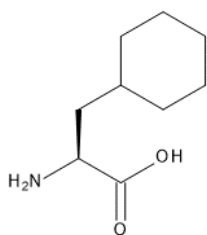
3-Pal  
L-3-Pyridylalanine



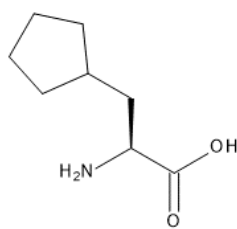
Bip  
L-Biphenylalanine



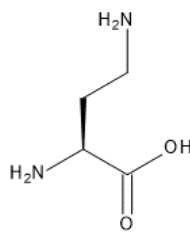
Cha  
L-Cyclohexylalanine



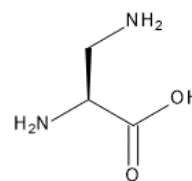
Cpa  
L-Cyclopentylalanine



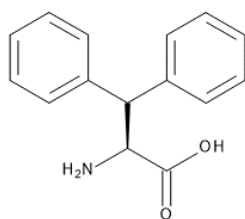
Dab  
L-Diaminobutyric acid



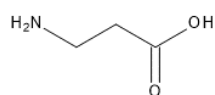
Dap  
L-Diaminopropionic acid



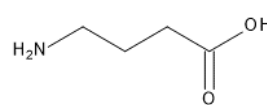
Dip  
L-Diphenylalanine



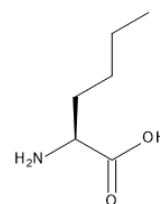
$\beta$ -Ala  
L- $\beta$ -Alanine



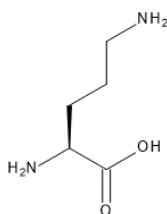
Gaba  
L-aminobutyric acid



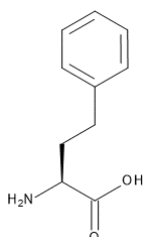
Nle  
L-Norleucine



Orn  
L-Ornithine



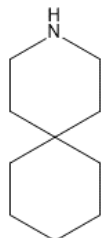
HPhe  
L-Homophenylalanine



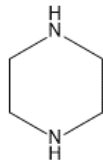
## Tail cappings

### C-terminal

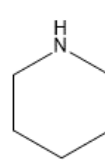
3-azaspiro[5.5]undecane



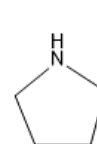
Piperazine



Piperidine

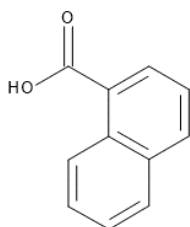


Pyrrolidine

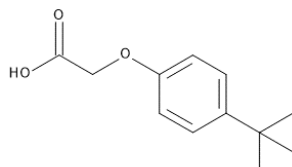


### N-terminal

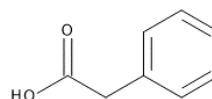
1-Napthoic acid



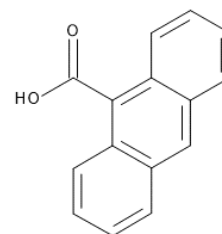
2-(-4-tert-butylphenoxy)acetic acid



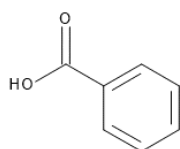
Phenylacetic acid



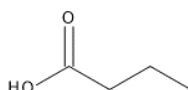
9-Anthracenecarboxylic acid



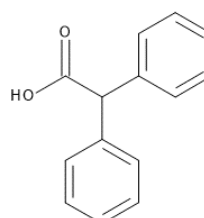
Benzoic acid



Butyric acid

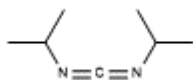


Diphenylacetic acid

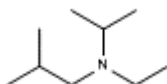


## Coupling reagents and solid synthesis resins

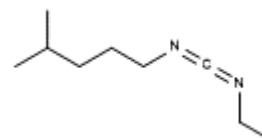
DIC  
N,N-Diisopropyl-Carbodiimide



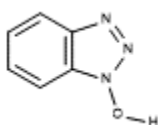
DIEA  
N,N-Diisopropyl Ethylamine



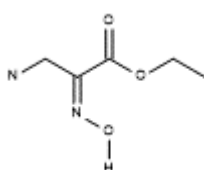
EDC  
N-(3-Dimethylaminopropyl)-N-Ethylcarbodiimide



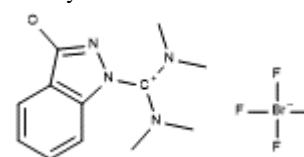
HOAt  
1-hydroxy-7-azabenzotriazole



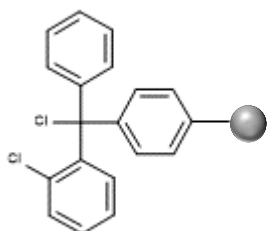
Oxyma  
Ethyl cyano-glyoxylate-2-oxime



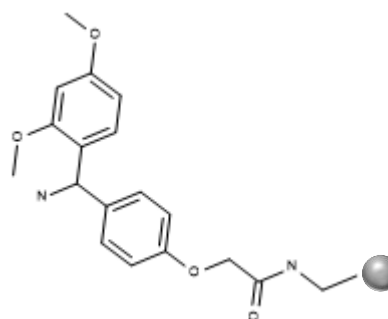
TBTU  
2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethylammonium tetrafluoroborate



Resin  
2-Chlorotrityl chloride resin



Resin  
Rink Amide-Chemmatrix



## Chapter 1: Introduction



## 1.1. Current status of the pharmaceutical field

During the 1950s and the following decades the modern pharmaceutical industry sector was developed and as a result, the number of approved drugs grew up systematically reaching a steady average rate of 20-30 New Molecular Entities Entries (NMEs) per year and with some fruitful exceptions as for example in 1996 which was the most successful year, ending up with the approval of 53 new entries by the Food and Drug Administration (FDA).<sup>1</sup>

Although the pharmaceutical industry had experienced a steady rate in the number of NMEs registered by the FDA, during the 2000s, due to limitations of the small molecules to interact with new challenging therapeutic targets, such protein-protein interactions, many diseases still remained untreated.<sup>2</sup> Fortunately, this tendency has suffered a turning point with the biological-based drugs boom, leading into an outstanding new record in registrations during 2018. In the last year, the FDA's Center for Drug Evaluation and Research (CDER) approved 59 novel drugs. Meaning a new record for the biopharmaceutical sector and a culmination of a new 5-year trend if compared with the previous two decades.<sup>3</sup> From this 59 NMEs, still 64 % were small molecules although a final overcome by biological-based drugs is expected to happen in the coming years.<sup>4</sup> Nowadays, these new biopharmaceutical drugs are mainly antibodies although siRNA and cell-based therapies are being finally awarded with a FDA approval and therefore opening the door to a new biotechnology industry era.

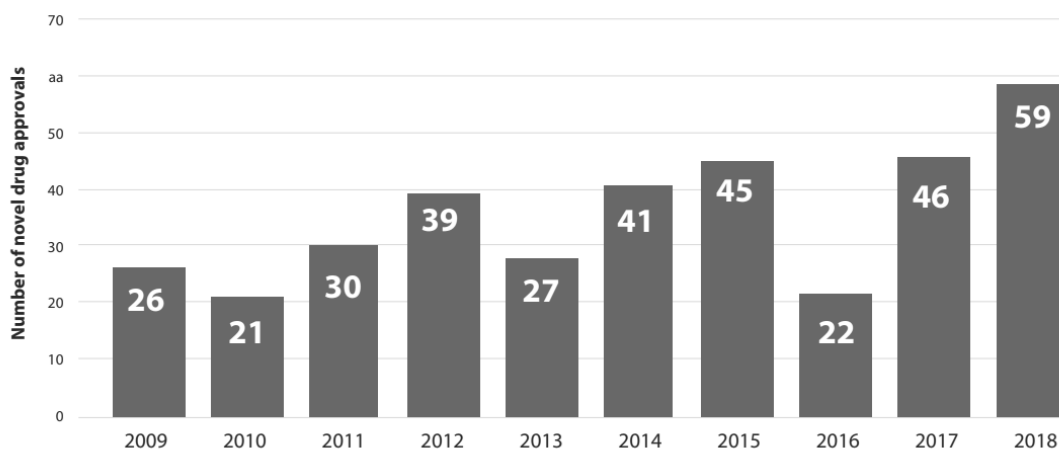


Figure 1. Number of new approved drugs by the FDA per year. There has been a considerable increase in the number of molecules awarded with the FDA's approval (except for 2016). Figure reproduced from the CDER report "2018 New Drug Therapy Approvals".

To explain the decay in the number of NCE along the last years, one plausible explanation is the problems to modulate difficult protein targets, often classified as undruggable by their topological features, overall represented by a flat and extended binding site protein surface region with no clefts where a small molecule can properly interact and be accommodated. This phenomena is more pronounced in protein-protein interactions (PPIs), that constitute a promising approach to treat unmet diseases but whose modulations is specially challenging.<sup>5</sup>



In this regard, antibodies (with averaged molecular weight of 150 kDa)<sup>6</sup> which are considerable larger structures than traditional small molecules, are able to cover larger protein surface areas and consequently disrupt more efficiently a PPI without the need of well-defined cavity in the protein surface.<sup>7</sup> However, the use of antibody based therapies is highly limited by their possible immunogenicity response (this is particularly relevant for chronic treatments), and the lack of permeability across biological barriers, such as the cell membrane, which challenge its delivery to the target location.

Fortunately, these drawbacks can be overcome for example by humanizing the antibody or by introducing some post-modification to increase their stability.<sup>8</sup> But, beyond the high production cost of biologics, the main handicap of antibodies is related with their poor or null permeability across biological barriers and therefore making practically impossible their applicability for intracellular targets. This also has an impact on their route of administration, which must to be parenteral.

One plausible classification of PPIs is done based on their cellular location. If intracellular (iPPIs), their modulation is considered more challenging and no effective solutions have been envisaged so far. Normally, iPPIs are involved in complex protein networks in which a malfunction in its biological activity can trigger an uncontrolled domino effect resulting in an unwanted health issue<sup>8</sup>.

Hence, design molecules able to cross biological barriers and disrupt intracellular PPI constitutes a field of intensive research within drug discovery paradigm, with future impact in relevant diseases such as cancer,<sup>8</sup> neuropathies<sup>9</sup> or cardiopathies.<sup>10</sup>

To understand the approach conducted in this PhD dissertation, it is necessary to introduce the use of peptides and peptidomimetics, whose main properties are defined below.

## 1.2. Peptides

In this scenario, the use of peptides have emerged as a promising tool to modulate the biological activity of PPIs.<sup>11</sup> Briefly, peptides are found in the middle of the chemical-space between traditional small molecules and more sophisticated agents as for example antibodies.<sup>12</sup> Peptides have between 2-50 residues length, and weight that may vary between 200 Da to 10.000 Da.<sup>13</sup> Indeed, although peptide average ( $\sim 500 \text{ \AA}^2$ ) solvent-accessible surface area (SASA) buried is less than half of a protein-protein interaction surface area, they bind in a more optimized planar shape and also more hydrogen bonds per interface are involved.<sup>14</sup> Usually SASA is used to compute the necessary transfer energy of a molecule to be transferred from polar solvent to a non-polar solvent or vice versa. However, SASA is also a good measure of the representative protein surface that is accessible for a possible ligand and also allows the differentiation between hot spots on the protein interface from those that are null spots.

The application of peptides for therapeutic purposes was established in the 1920s, when insulin was the first isolated therapeutic peptide from canine and bovine pancreas, and was used for the treatment of Diabetics.<sup>15</sup> Since those early days in drug discovery field, several peptide drugs have gained market access and the trend keeps growing up (Figure 2).<sup>16</sup>

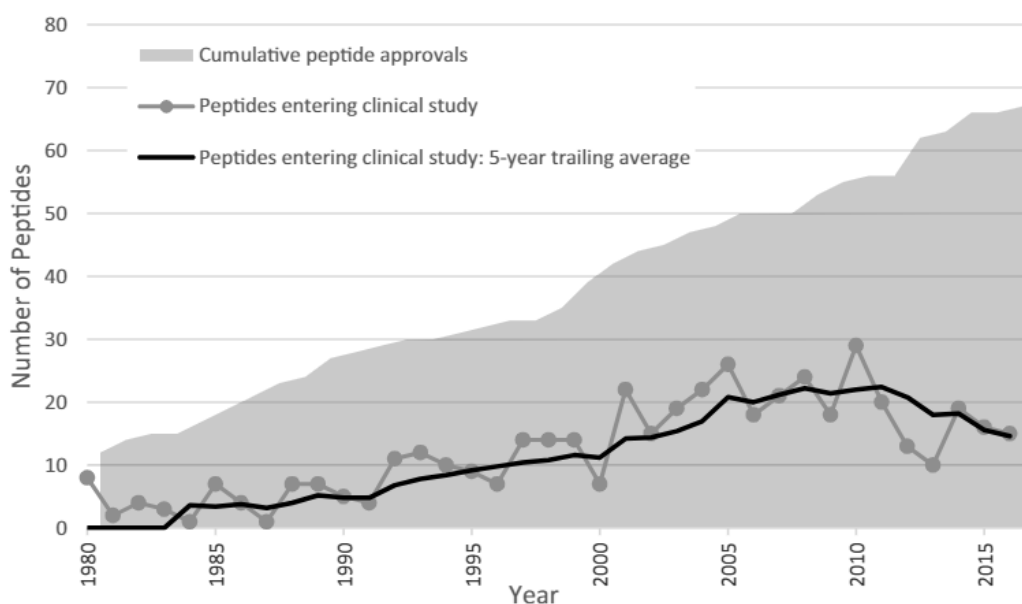


Figure 2. Cumulative plot of the development of therapeutic peptides in major pharmaceutical markets. It is considered entering in clinical studies when phase I or pilot human study starts. *Image taken from reference 15.*

The vast majority of developed peptide drugs are usually derived from natural products or bigger proteins binding motifs, then limiting the sources to develop *de novo* peptides and consequently their historical weight in the drug discovery field. Furthermore, although peptides generally are non-toxic, they show low stability in front proteolytic activity and poor permeability across biologic barriers.<sup>16</sup>

Then, stability and permeability of peptides are the major drawbacks for their use as therapeutic drugs and therefore limiting their application

### 1.2.1. Stability

Toxicity and bioavailability are properties straightly linked to the peptide stability. Compound cleavage could lead in smaller peptides with off-target effects, then having a possible toxic effect when administered *in vivo*<sup>16</sup>. On the other hand, a low resistance to enzymatic proteases will be translated into a lower concentration in blood, narrowing the compound bioavailability, and therefore dropping its potency as a consequence of a lower concentration in blood.

Hence, improving the drug potency by keeping the therapeutic dose longer time in the body, plus a reduction of toxic effects, as a consequence of less unwanted cleaved sub-products of the peptide, will enhance the metabolic profile to progress in the drug pipeline for a given peptide.<sup>17</sup>

In addition to this, oral drug administration remains as the preferred route of administration for the pharmaceutical industry because of the patient compliance and the feasibility to administer large amounts of the drug. In this scenario, compound's metabolic stability is a key property for an efficient distribution after oral dosage. Upon oral administration, the compound must remain stable to the strong acidic conditions of the stomach, which can compromise the peptide integrity. Later, it must be highly stable

to the first pass of metabolism, and once in blood stream must to be stable against plasma proteases.

However, peptide proteolytic resistance to the body enzymatic activity and strong acidic conditions of the stomach is very low, promoting fast degradation rates. Indeed, digestive track strong conditions prevents most of the therapeutic peptides form be administered orally, and then have to be administered intravenously.<sup>18</sup> Likewise in all situations, peptides are rapidly cleared from the body system and excreted, thus limiting their circulation time and as a consequence a possible therapeutic application.<sup>19</sup>

### 1.2.2. Permeability

Biological barriers such as, the skin barrier, the gastro intestinal track (GIT) or the blood brain barrier (BBB) prevent the absorption of xenobiotics by the body. The, drug absorption relies on a number of factors and drug properties, such compound's solubility, physical and metabolic stability, route of administration, status of the biological barrier in physiological and pathological conditions, and formulation. Permeation across biological barriers is mandatory for a drug to reach the target location.<sup>20</sup> In particular, when the target is extravascular located and the route of administration of the compound is non-parenteral.

Mechanisms of transport can be classified in active and passive transport.

### 1.2.3. Active transport

Molecules are internalized by membrane transporters in a saturable energy-dependent process. There are three main active transport mechanism, carrier-mediated transport (CTM), receptor-mediated transport (RTM) and absorptive-mediated transport (AMT).

In CTM is based in the specific recognition of a substrate by its carrier transporter protein, which assist the translocation of the substrate through the cell. The molecule to be transported is bound to endothelial membrane transport protein, which drive the biological barrier transport.<sup>21</sup> Since CTM is a vesicular-based mechanism of transport, its capacity for transportation is limited to small molecules.

RTM is a vesicle-based transport, rather than a stereoselective carrier, that occurs in three steps, receptor-mediated endocytosis of the compound present in blood stream, movement through the endothelial cytoplasm and finally the compound is released by exocytosis. Then, RTM receptors are characteristic for their function to recognize particular types of endogenous molecules.<sup>22</sup> Therefore, a drug design can be done to play the role as a RTM ligand in order to cross for example the cell membrane.

AMT, unlike CTM and RTM, is not dependent of any plasma membrane receptor or transported protein. This mechanism is driven by the cationic or anionic charge of the transported molecule. The charged molecule interaction with the counterpart lipid membrane charges and triggers endocytosis, which can lead to transcytosis afterwards.<sup>23</sup> Since AMT is a vesicular mechanism of action, larger cargos can be transported and also no competition with endogenous compounds present in the blood stream is found.

Although peptides are unlike to cross biological barriers, cell-penetrating peptides (CPPs), which are peptides type comprising between 5-30 residues have enhanced permeability. There are more than 100 known CPPs that are divided in three major

categories; amphipathic, catatonics and hydrophobics.<sup>24</sup> The use of CPPs as shuttle to deliver different *cargos* as for example DNAs, siRNAs, or proteins, has been successfully applied and translated from the basic research to clinics.<sup>25</sup>

However, CPPs cellular uptake is not fully understood, unspecificity and formation of membrane pores are factors that can produce toxic effects. Moreover, most of the currently known CPPs have a low stability and their applicability has to be studied almost case by case. Once internalized, endosomal release remains a challenge nowadays. Another drawback of the use of shuttles is that the physico-chemical properties of this one can be altered by the *cargo*, then losing its transport capacity.

#### 1.2.4. Passive diffusion

Passive diffusion is the most common mechanism for drug transport. This transport mechanism is mediated by concentration gradient of the transported molecule from both sides of the corresponding biologic barrier.<sup>21</sup>

Concretely, passive diffusion is an intrinsic property that relies in the compound physicochemical features and as opposed to active transport. In this regard, in 1997 it was reported by Lipinski that the majority of oral administered drugs that have a FDA approval follow the physicochemical features ( $MW \leq 500$  g/mol,  $\leq 5$  hydrogen bond donors,  $\leq 10$  hydrogen bond acceptors and  $\log P \leq 5$ ) described in the Lipinski's rule-of-5 (Ro5).<sup>26,27</sup> However, peptides-like compounds usually escape from the chemical-space comprehended by Ro5 and in general terms are non-permeable. Lipophilicity is the major determinant of peptide permeability by passive diffusion, although should be balanced with a moderate or good solubility in aqueous environment. In this regard, medicinal chemistry approaches aimed to circumvent a permeability limitation of a given molecule have been focused to increase the overall molecule lipid solubility by blocking hydrogen bond-forming groups on the parent molecule. However, an excess of lipidization may enhance binding to plasma protein which could translate into molecule retention on the lipid membrane.

Fortunately, some natural peptides, such as Cyclosporine A (CSA),<sup>28</sup> a 11-amino acid long cyclic peptide, is a clear example that there is life beyond the Ro5 (bRo5). CSA has a chameleonic behavior that allows to this cyclic peptide switch its tridimensional conformation by forming intramolecular hydrogen bonds and then, increasing or decreasing its Polar Surface Area (PSA)<sup>29</sup> depending of the medium polarity,<sup>30</sup> that the peptide is located in giving moment, *e.g.* plasma or intestinal fluids versus cell membrane.

Indeed, most of the currently approved drugs are transported across biological barriers through passive diffusion. Unfortunately, the knowledge about which rules drive the passive diffusion for those molecules that are bRo5 is limited, hampering the *in silico* prediction and its applicability for proper drug discovery development.

#### 1.2.5. Status of peptides as therapeutic agents

In conclusion, natural peptides are often associated to a low proteolytic resistance and negligible permeability across biological barriers, narrowing their applicability as future drugs. To overcome poor permeability, CPPs are implemented as *cargo* carriers, but also toxicity issues related with the shuttle release or the complex stability can be found.

However, the above drawbacks are leveraged by the capacity of peptides to target large and flat areas of a protein surface with high affinity and by their ability to naturally modulate the biological activity of proteins.

Hence, peptides are the most suitable molecules to unlock those therapeutic targets that were thought to be undruggable, such as PPIs,<sup>31</sup> thus explaining why they have called the attention in the drug discovery field during the last years. But an optimization of their intrinsic physico-chemical properties is needed for a successful application of peptides in the drug discovery field.

### **1.3. Peptidomimetics**

The above disclosed needs of the pharmaceutical industry to develop new drugs to target those proteins that were thought to be undruggable in the past, especially those that are intracellular PPIs, along with the progress made in the peptide chemistry, particularly with the development of the Solid-Phase Peptide Synthesis (SPPS),<sup>32</sup> has been translated in an increase of peptide based drug that have reached the market. Indeed between the years 2011 and 2017, from the 22 new peptide entities approved by FDA 20 had some chemical modification.<sup>33</sup>

Interestingly, new feasible applications of peptides are being published as for example the before mentioned use of cell penetrating peptides (CPPs) as drug shuttles to delivery non-permeable drugs to their site of action<sup>24,25</sup> or the use of peptides for labeling techniques.<sup>34,35</sup>

Furthermore, new strategies devoted to screen *de novo* peptide-like structures against a given target have been developed, as for example phage display,<sup>36</sup> ribosomal synthesis of peptides<sup>37</sup> and *in silico* techniques.<sup>38</sup> In these compound screening campaigns non-natural amino acids or chemically modified amino acids and modifications on the peptide backbone are included in many occasions.

Generally, peptide-like based drugs, peptidomimetics, are characterized by a peptide structure core, which usually has been tuned in order to overcome the major drawbacks of natural peptides; poor metabolic stability and low permeability across biological barriers. Moreover the current synthesis techniques allow the optimization of the peptide sequence to gain more affinity and specificity for the target protein along with the improvement of its biological profile.<sup>16</sup>

Peptidomimetics are based in a peptide sequence modified in order to improve its biophysical properties along with a potency improvement. Some of the most frequent modifications are substitution of L-amino acid by their chiral version (D-amino acid) or non-natural amino acids, selective backbone *N*-methylation and peptide cyclization are recurrently found in a peptidomimetic structures.<sup>39</sup>

#### 1.3.1. Peptide sequence modifications

Amino acid substitution by D-amino acids or non-natural generally improves the ADME properties of the compound. Meaning, an increase in front proteolytic degradation, as the points of the peptide sensitive to metabolic cleavage or modifications are replaced or deleted.<sup>40,41</sup> Besides, by introducing non-proteinogenic amino acids, as non- $\alpha$  amino

acids, where the amino group has been shifted further from the carboxylic group, or side-chain modifications, as Lysine derivatives (Citrulline, Ornithine and Homo-Arginine), the affinity of a peptide by its target protein can be optimized. Nowadays, more than 100 Fmoc protected non-natural amino acids are commercially available. Thus, exponentially increasing the number of possible combinations even for short peptides (3-5 amino acids) accessible from SPPS synthesis.

Other structure-changing modifications are staple peptides, in which the  $\alpha$ -helical secondary structure is stabilized by intramolecular bonds.<sup>42,43</sup> Also, disulfide bonds or derivate as lactams, olefins and thioethers are being used to stabilize the peptide tridimensional structure<sup>44</sup>.

In addition, some other post-modifications are included in the peptide synthesis such as bond-surrogates (despsipetides, pseudo-peptides or peptoids), *N*-terminus and *C*-terminus capping. Moreover designing a peptide pro-drug, is a well-known technique to increase the overall peptide ADME properties. A good example of this, is the attachment of polyethylene glycol (PEG) to a peptide sequence. PEG is an amphiphilic polymer linked to the active peptide sequence that increases the peptide stability and permeability. Then, once the coupled peptide is internalized and reaches the target location, PEG gets cleaved from the peptide sequence. Then, the active peptide can effectivity interact with the therapeutic target meanwhile PEG is fully excreted after proteolytic degradation, therefore being non-toxic.<sup>21</sup>

### 1.3.2. *N*-methylation

Selective backbone *N*-methylation of solvent exposed N-H bonds is a widely used approach to reduce the polarity and the number of hydrogen bonds donors for a given peptide in order to improve the peptide permeability.<sup>45,46</sup> A complete backbone *N*-methylation could be detrimental from a permeability point of view, as normally the consequent loss of solubility cannot be balanced enough with the lipophilicity gain.<sup>47</sup>

Furthermore, selective *N*-methylation of peptides bonds increases the peptide stability by reducing the number of sensible proteolytic points along the compound sequence.<sup>48</sup> Also, the steric hindrances related with the *N*-methylation blocks the number of possible conformations and stabilizes the amide bond cis-trans equilibrium.<sup>49,50</sup>

Currently, there are several *N*-methylated Fmoc protected commercial available amino acids, but also, selective *N*-methylation can easily be performed during SPPS process without the cleave of the growing peptide from the polymeric support, thus allowing site-selective *N*-Methylation<sup>51</sup>. In addition to this, new approaches include *N*-alkylation with longer and more variety of alky chains.<sup>52,53</sup>

### 1.3.3. Cyclization

Engineering cyclic peptides to target PPI is a hot topic within the drug discovery field.<sup>54</sup> These molecules have the capacity to bind larger protein interfaces and they outperform linear peptides in terms of proteolytic stability and permeability across biological barriers.<sup>55</sup>

Macrocycles rotation angles are more constrained than their lineal counterparts, meaning a confined conformational space. When the bioactive conformation is the most favored

by molecule structural restrains, the energetic entropic penalty cost to adopt it is lowered if compared with the lineal version and therefore its potency enhanced.<sup>56</sup>

Naturally occurring cyclic peptides are present in the drug market, as for example the before mentioned CSA for the treatment of psoriatic arthritis, another good example is Romidepsin, an anticancer drug.<sup>57</sup> Besides the discovery of cyclic natural products, new methods to engine *de novo* cyclic peptides to inhibit PPI have been developed. Aside from SPPS methods, in 1985 the 2018 Nobel prize in chemistry, George P. Smith develop the invention of phage-displayed to screen peptide libraries,<sup>58</sup> a widely used technique that has been optimized to code non-proteinogenic building blocks or scaffolds to obtain bicyclic peptides.<sup>59</sup> A similar approach are mRNA display libraries with a library size of  $10^{12}$  (in front of  $10^9$  of Phage display) and also, a technique optimized by Prof. Suga and coworkers to add non-natural amino acid or post-modification in the ribosomal synthesis code.<sup>37,60</sup>

Although the constantly increasing interest in macrocyclic structures and after several successful cyclic compounds have reached the market, their complexities related with structural prediction have limited *in silico* screening techniques<sup>61,62</sup>. Briefly, correct starting point conformation for a cyclic structure is determining for an accurate prediction, as different low conformation can be wrongly found as the most stable conformation biasing all subsequent results. Thus, the design of macrocyclic compounds remains in an early stage and then, a more spread use of cyclic peptides in drug development programs is narrowed.

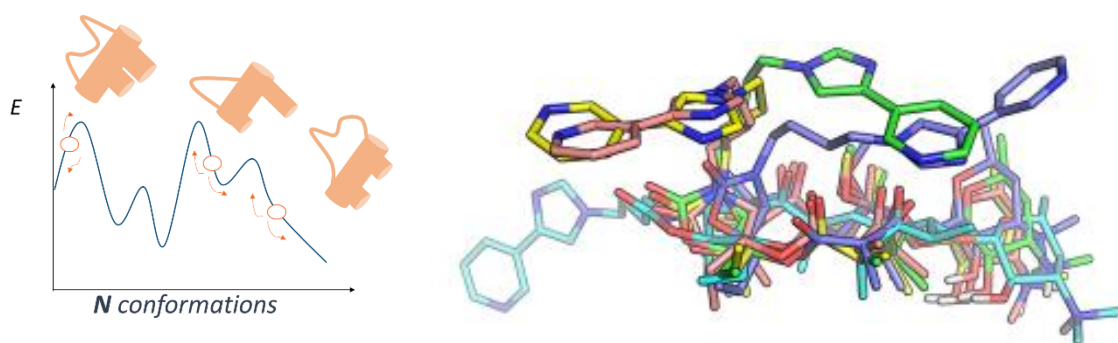


Figure 3. Graphic representation of conformational complexity of macrocyclic structures. *In silico* studies are biased by the starting point conformation. Although long-term Molecular Dynamics<sup>55</sup> or Distance Geometry<sup>56</sup> can be applied to overcome this issue, the associated computational cost narrow the screening scope. On left the different low energy conformations that can be accessible for a given cyclic peptide is represented. On the right, all conformations that from and energy point of view are accessible for the cyclic peptide *c*(HomoPhe-Nle-Dip-Gaba- $\beta$ -Ala).

As final remarks, peptidomimetic solutions are able to overcome the main handicaps of peptides as drugs. In fact, continuous evolution of the different strategies to obtain tuned peptides in order to optimize their structures or to improve both, physico-chemical properties and bioactive function, has lead into a golden-age era of peptides.

However, due to a lack of an extended background in the peptidomimetic field, the knowledge about how to predict the behavior of different molecules is limited. Meaning, an expensive trial-and-error process to find *novo* structures or to improve initial hit candidates.

## 1.4. Iproteos approach

Iproteos was founded in 2012 by Prof. Ernest Giralt and Dr. Teresa Tarragó with the idea to apply peptidomimetic novel structures to target intracellular PPIs of therapeutic interest. Since its inception, the company contribution in the drug discovery field has been recognized by public and private institutions with relevant awards or funding campaigns. Altogether, Iproteos track has allowed the company to expand its initial pipeline from the initial molecule for the treatment of Cognitive Impairment Associated with Schizophrenia (CIAS) to the development of new drugs for Epilepsy, Artherosclerosis and several projects related with cancer disease. More information about Iproteos can be found in the company's website, [www.iproteos.com](http://www.iproteos.com).

The design of new peptidomimetics molecules is based in an in-house built platform, IPROTech. This technology lies in a set of computational and experimental tools aimed to exploit a proprietary *in silico* library (IPRO Library) to identify, through state-of-the-art computational approaches, the most promising small peptidomimetic structures able to bind to the desired target with large affinity and selectivity, but also with high membrane permeability. By using structure-based approaches, such as peptide docking and protein-based molecular dynamic simulations, a throughout survey of the desired target is conducted in a sequential manner identifying and prioritizing the synthesis of the more promising candidate structures.

A through non-confidential description of IPROTech is done in Chapter 2 of this thesis. Graphically, IPROTech can be illustrated as depicted in Figure 4.

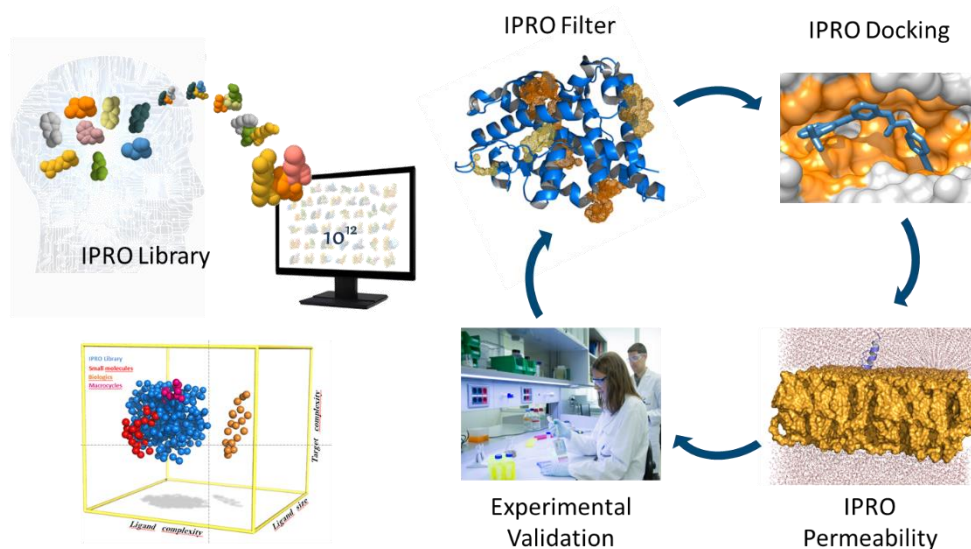


Figure 4. Illustrative representation of the main tools and steps of the Iprotech. A large in-house peptidomimetic virtual library (IPRO Library) is screened through the peptide based computational tools (IPRO Filter, IPRO Docking and IPRO Permeability). Afterwards, selected molecules are synthesized and evaluated. Based on the experimental results, the same process is repeated until obtain a promising drug candidate.





## Objectives



The main objective of this thesis was to evaluate the capacity of the proprietary technology platform IPROTech to generate peptidomimetics able to modulate intracellular protein-protein interactions. Hence, technology validation by its successful application on high scientific interest cases was a major milestone of this work. For this purpose, here, four cases of challenging PPI have been studied and evaluated, namely: Talin-Vinculin, RAD51-BRCA2, K-RAS-Effectors and Retromer-L2 protein. With the successful results obtained in this thesis, Iproteos had the plans to be positioned as a key partner for any drug discovery program which aims to target an intracellular protein-protein interaction.

The constant integration of peptide docking protocol steps, new amino acids and terminal capping building blocks or other new approaches to improve the compounds physicochemical properties was fundamental to achieve higher success rates in the different company's project.

Hence, the work done during this thesis has been structured around the following objectives:

- 1) Synthesis of high purity (>95 %) novel peptidomimetics that had been designed applying the IPROTech technology in order to disrupt the PPI of interest.
- 2) The company's focus is targeting intracellular PPIs, therefore permeability across biological barriers is a key asset. Meaning an evaluation of the permeability of the synthesized compounds across biological barriers and/or their capability to be uptake by cells.
- 3) In order to have permeable compounds, the increase of hydrophobicity detriment their solubility. A good solubility is essential for any drug candidate and as a consequence study new approaches to solve this issue are mandatory to overcome the solubility bottle neck.
- 4) Demonstrate the IPROTech efficacy by its application in relevant scientific projects where therapeutic mechanism of action is related with the disruption of a PPI.
- 5) Integration of all generated data into the company know-how in order to optimize the IPROTech efficacy.
- 6) Include new business cases in the company's pipeline to diversify the number of projects and molecules that can reach clinical phases.



## Chapter 2: IPROTech



## 2.1. Introduction

IProTech is the acronym of Iproteos proprietary technology, consisting in a set of algorithms devoted to design and engineer permeable peptidomimetic structures aimed to inhibit intracellular PPIs. The computational platform is combined with adjacent experimental studies to verify prediction results. Once the computational evaluation of a given PPI is completed, the compounds suggested as more active (from a *in silico* viewpoint) are then synthesized and characterized in terms not only of permeability, but also stability and potential activity versus the desired protein target.

Once the former experimental results are obtained in the first screening, the prediction is reinforced with wet-lab data. Based on sub-sequent iteration, the predictions are refined towards an optimal results, both from an *in silico* and experimental viewpoint. Ideally, after 3-4 iteration, initial hit compounds are fully optimized to a lead candidate structure that is immediately subjected to PK/PD *in vivo* studies.

IProTech exploits a proprietary *in silico* library of peptidomimetics, whose size is currently estimated around  $10^8$ . This virtual library, coined as IPRO Library differs from other commercial catalogs that are purchasable in terms of ligand size and physico-chemical properties. Unlike commercial libraries, whose chemical space can be easily screened with modern IT infrastructure ( $10^5$ - $10^6$ ) in days or weeks, the large size of IProTech forces to apply logical steps to reduce the large number of compounds to be initially screened to an amenable number. To do so, a set of proprietary algorithms were designed to efficiently navigate and to select from IPRO Library the more promising (*in silico*) compounds too be tested. Additional algorithms, named IPRO Filter, IPRO Docking and IPRO permeability, are approaches to leverage not only the computational resources, but also guide predictions towards a scenario with the less number of computational workload and to overpass the known limitations observed when applying modern Computer Aided Drug Design (CADD) software to peptides and peptidomimetics. The CADD software accuracy is especially limited in those cases where peptide size clearly exceeds the size of small molecules. To develop accurate and novel algorithms, with a keen interest in execution efficiency, three main branches were created, consisting in the software herein denoted as IPRO Filter, IPRO Docking and IPRO Permeability. These algorithms are sequentially applied to screen new peptidomimetics sequences to target a PPI of interest. The development of this technology has been carried out keeping the focus on the particularities of peptide based-molecules.

After IProTech identifies the best molecules to be synthesized, we proceed to synthesize them through solid-phase peptide synthesis (SPPS) following a Fmoc/*t*But strategy. Synthesis are performed manually because of the low length of the compounds, the high price of some of the non-natural amino acids used and because backbone specific *N*-alkylations are added frequently. Therefore, automated synthesis methodologies cannot be applied in such designs in straightforward and inexpensive way.

Finally, the high purity synthesized compounds (usually > 95 %, HPLC area/area) are evaluated *in vitro* to screen positive hits and to validate the computational model, which is a key point for the further optimization of the peptide-like sequences. Other experiments related with the compounds biophysicochemical properties, such as



permeability, solubility and stability are also included as evaluation tools to select the most promising hit.

To validate the IPROTech platform, the company has settled different collaboration projects with leading research groups from the academia. These collaborators have a long track record of the study of a particular iPPI with a possible or validated therapeutic application.

In the context of this thesis, these groups were responsible for the efficacy (*in vitro*) evaluation of each project after the identification of potential peptidomimetics based on IPROTech. The collaboration agreements were done under win/win scope, with the aim to first, validate the advantages of IPROTech where other computational approaches previously failed measured in terms of active molecules for the desired PPI; and secondly, to provide a compound tool to collaborators to continue their research and demonstrate the therapeutic utility of targeting the corresponding PPI.

Briefly, the established collaborations were;

1. Cellular and molecular mechanobiology group from the Institute for Bioengineering of Catalonia (IBEC). The group is led by Dr. Pere Roca-Cusachs and the collaboration was settled to discover inhibitors of the Talin-Vinculin protein-protein interaction as possible target to modulate the growth of solid tumors. The information about this PPI as well as the approach made by Iproteos is throughout discussed in the chapter 3 of this thesis.
2. The team of Dr. Alessandro Sartori from the Institute of Molecular Cancer Research (IMCR) from the University of Zurich. The project aimed to disrupt the interaction between RAD51-BRCA2 through a synergistic treatment with Olaparib, a poli ADP ribosa polimerasa inhibitor (PARPi), for the treatment of several types of cancers. The information about this PPI as well as the approach made by Iproteos is throughout discussed in the chapter 4 of this thesis.
3. Signaling and checkpoints of cell cycle group from the department of Biomedical Sciences from the University of Barcelona. The group is led by Dr. Neus Agell and its research interest are related with the role of K-RAS in cancer. Therefore, the collaboration was carried out to develop inhibitors of the interactions of K-RAS and its effector proteins. The information about this PPI as well as the approach made by Iproteos is throughout discussed in the chapter 5 of this thesis.
4. The Membrane and Trafficking Lab from the CICbioGune led by Dr. Aitor Hierro had published the crystal structure of the Retromer protein, a transmembrane protein which is used by the Human Papillomavirus to infect the cells. The idea behind this project was develop peptides that can compete with the virus for the recognition of the Retromer protein. The information about this PPI as well as the approach made by Iproteos is throughout discussed in the chapter 6 of this thesis.

## 2.2. IPRO Library

IPRO Library constitutes a database in which allocated compounds are a product of virtual combination of different amino acid (natural and non-natural). The selection of building block that compose the peptidomimetic sequence were selected on basis of the know-how of Iproteo's team on cell permeability and SPPS synthesis. By combinatory methods, the sequences are encoded into suitable IT format to yield at the end a number of unique compounds around  $10^8$  structures.

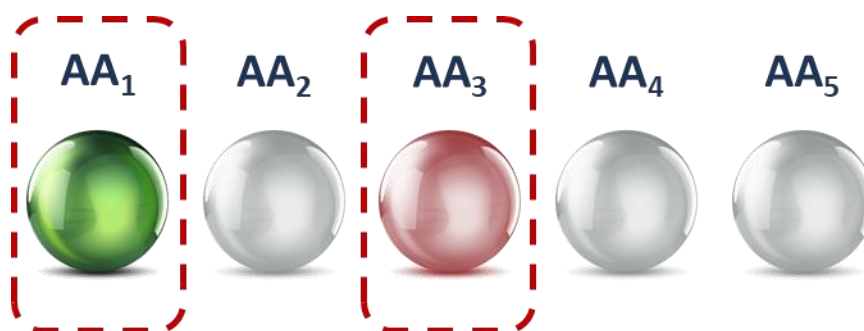


Figure 5. Commercially available Fmoc-Aa-OH natural, non-natural, N-alkylated and D-amino acids are found in the virtual library.

The library building blocks are continuously subjected to evaluation. Mainly those structures that show SPPS limitations or undesired properties (*e.g.* short half-life) are replaced by a backup building block with similar physicochemical properties (Figure 6). Furthermore, the library is continuously updated with the most recent commercially available building-blocks.

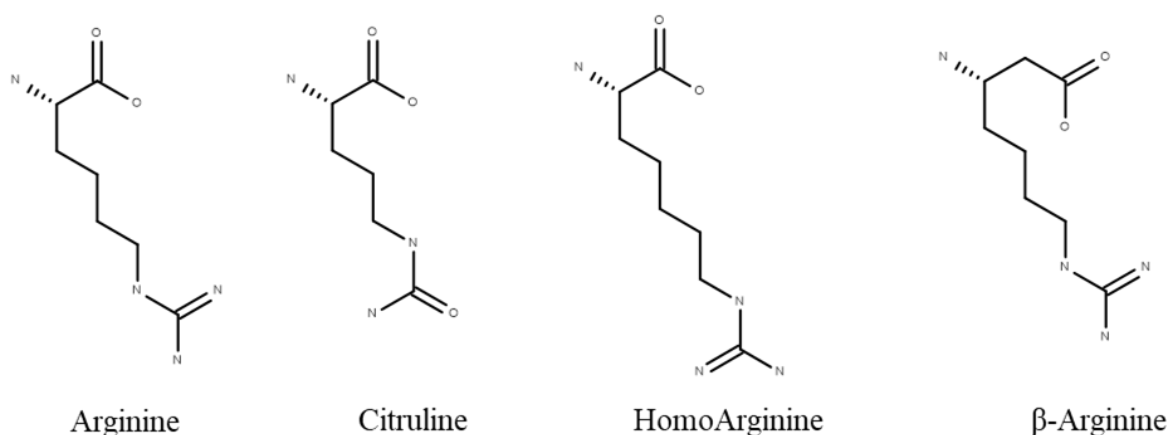


Figure 6. L-Arginine is a natural amino acid, whereas L-Citruline, L-HomoArginine and L- $\beta$ -Arginine are analogous derivatives from the original one. Switching from one to the others allows the optimization of a peptidomimetic sequence.

Because the number of virtual structures that define the boundaries of IPRO Library chemical exceeds computationally amenable resources of the most advanced IT infrastructure, a clustering distribution approach was conducted to retain chemical space representation while reducing the number of linear sequence structures, usually by

applying a virtual library not exceeding 250000 molecules. This reduction was done taking into account that the peptide sequence diversity is preserved in terms of building block occurrence and peptide size (N=2, 3, 4 and 5). Then, forming building blocks of top ranked IPRO Docking solutions can be permuted by other amino acids analogs from the same cluster family and screened again for refinement.

All comprehended building blocks of the IPRO Library have been parametrized by applying AMBERTools package<sup>63</sup> and converted to a suitable format (PBQT) by using Open Babel,<sup>64</sup> which is the standard format required for AMBER to carry out molecular dynamics analysis.

### 2.3. IPRO Filter

This tool allows the navigation along IPRO Library; it explores the target protein surface in order to identify potential binding sites and describes the desired physicochemical properties that contribute to the binding process, also termed druggability analysis. Therefore, an approximated 100-fold library reduction can be performed efficiently, which is regularly translated in a pull of 100000 compounds that will be evaluated *in silico*.

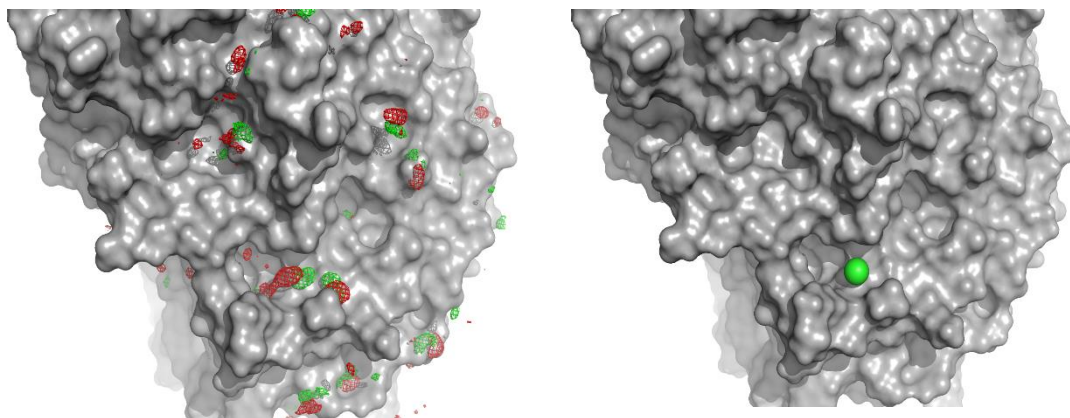


Figure 7. A solvent box over the protein structure with volume/volume proportion of 20 % (organic/water molecules) is constructed. Then, the number of interactions and the stability of these interactions between the organic probes and the protein surface area reveals the potential hotspots. Left, density representation of probe analyzed along the simulation (PDB: 1H2W). Right, coordinate representation of main region where probe is concentrated represented as green sphere. The physico-chemical properties of probe demand a comparable amino acid in the peptide sequence (in this case a Serine or Threonine) in order to achieve desired potency.

Normally, this type of target evaluation is carried out by analyzing simple geometric and energetic metrics, such as the curvature of the protein surface and degree of hydrophobicity.<sup>65,66</sup>

However, PPI are typically flat and can undergo through conformational changes when a small molecule ligand binds to them. Therefore, Molecular Dynamics (MD) based approaches are applied to understand the existence of hotspots at PPI interfaces. Particularly, Mixed-Solvent Molecular Dynamics (MSMD) simulations allow the protein under study to reorganize its tridimensional structure, and therefore new binding sites that are not present in the *apo* or protein-bound conformation can be found.<sup>67,68</sup>

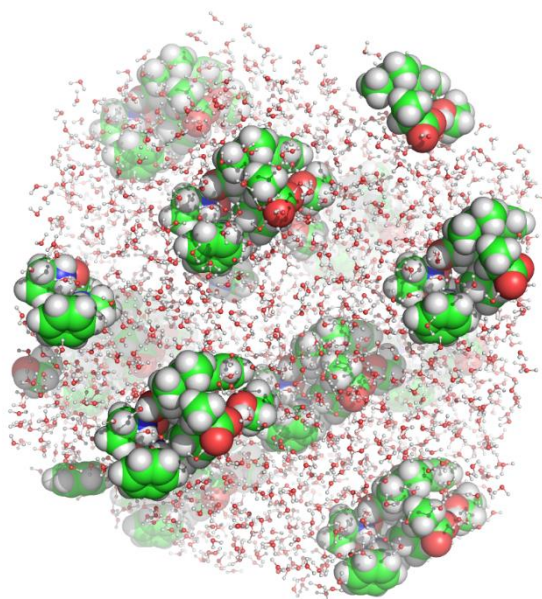


Figure 8. An octahedral solvent box is replicated around the protein surface according to an expanding factor of 7-10 Å. The solvent box represents aromatic, hydrophobic, small-polar, negatively-charged and positively-charged moieties. To avoid aggregation of organic probes, heating and cooling effect steps are performed to sparse organic molecules homogenously around the protein.

In all the cases, MSMD were limited to one or two organic probes immersed in the solvent box, running alternative MD replicas. For example, isopropanol was a probe used to quantify the maximal affinity achievable at different patches on the protein surface. In the case of IPRO Filter, multiple organic probes are tested simultaneously, yielding not only preferences in terms of probe affinity, but also probe competition, that is particularly useful when optimizing hit structures to lead structure.

In conclusion, IPRO Filter is in-house developed MSMD protocol which has been focused to evaluate the suitability of peptide sequences by studying the tendency of organic probes to be allocated in a protein surface region that normally is from biological relevance.

## 2.4. IPRO Docking

In the last decades, large amounts of structural data of small molecules have been the fundamental benchmark to develop and train docking algorithms. These algorithms are based in a Monte Carlo approach (MC)<sup>69</sup> to estimates the potential binding energy (kcal/mol) between a ligand and a protein. Theoretically, the obtained score value correlates with the experimentally biophysical data, allowing a low cost and straightforward screening of large virtual libraries. As a result, a large plethora of docking software has emerged in terms of commercial package<sup>70-73</sup> (Glide, MOE, Gold and ICM) or open/source initiatives<sup>74,75</sup> (rDock, AutoDock or AutoDock Vina), with similar performance and accuracy.

The current state-of-the-art algorithms are subtle modifications of the above methods, where in some cases more computationally demanding but sophisticated post-filtering approaches are coupled at the end of the docking workflow with the aim to further refine results and reject solutions that are known to be responsible of false positives

predictions.<sup>76</sup> Another critical point is the exploration of the molecule conformational landscape, which is given by its rotatable bonds 10-15  $N_{rot}$ . A larger number of rotatable bonds will imply a larger number of conformations that the molecule can adopt. Then, this number of possible rotations can be translated in a search of 10000-20000 conformations per molecule. Larger and more flexible small molecules require longer execution times to obtain the desired accuracy, meaning a high extra computational cost.

When comparing a peptide to a small molecule, the former is primarily characterized by the larger flexibility not only of the backbone, but also of the side-chains of the amino acids. In average, every amino acid of the peptide sequences contributes with a minimum of 3 and a maximum of 7  $N_{rot}$ . Thus, a small natural peptide sequence of 3 amino acids, overpass the flexibility of any small molecule. While increasing the peptide size to larger number of amino acids (4-7), the averaged  $N_{rot}$  is exceeding the capabilities of modern docking software. To mitigate this effect, a plausible approach consisting in expand the MC search to millions of conformations would imply longer time computing time, a non-amenable number even for modern IT resources.

Thus, the application of current docking techniques to peptides delimits the high throughput scale to evaluate libraries of only some hundreds or thousands of peptides due to the exponential increase of possible tridimensional conformations. However, most of these conformations are artefactual, as they consider angles bonds that are not accessible for peptides backbone bonds according to the Ramachandran plot.<sup>77</sup> Ramachandran plot determines which are the dihedral angles  $\psi$  against  $\phi$  of that each amino can have and the frequency of each combination.

By using LEAD-PEPS benchmark<sup>78</sup> as a recent published training set composed of crystal structures of peptide bound to its target protein, it was found that more than 80 % of the explored conformations along the MC simulations were not feasible from a geometric point of view.

In this scenario, a proprietary docking platform, IPRO Docking has been designed and set based on the core of AutoDock Vina (AD Vina) and Smina (a fork of AD Vina)<sup>79</sup> taking into consideration Ramachandran information. In the IPRO Docking a biasing function has been successfully incorporated in the code to predict bioactive conformations of linear peptides, with correct pose prediction accuracy larger than 80 % (for 4-residues sequences). The algorithm discriminates between true and false conformations by restricting their flexibility of all backbone amide bonds to each amino acid Ramachandran plot.

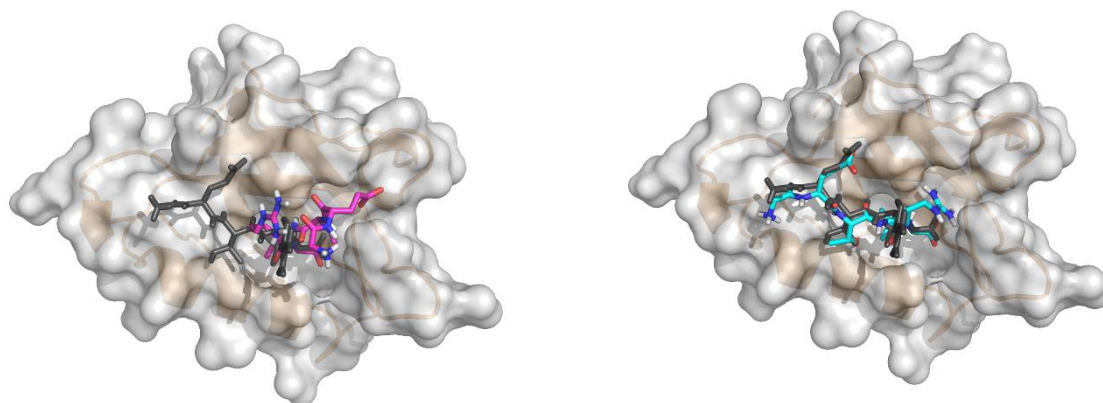
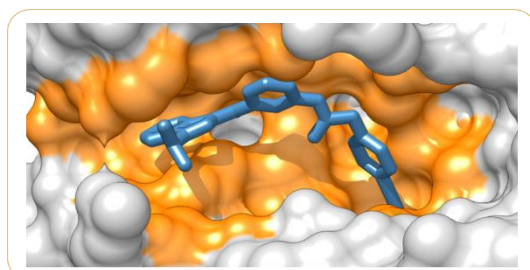


Figure 9. Left, in pink color the docking prediction of penta-peptide (PDB: 3NFK) based on AutoDock Vina (RMSD: 7.7 Å). Right, in blue color the prediction of same peptide using IPRO Docking (RMSD: 2.4 Å). Minor differences are found in the IPRO Docking prediction affecting to the arginine sidechain, whose orientation differ when comparing X-ray and docking prediction (but in both cases, solvent-exposed oriented)

The technology described here was trained using a benchmark of 22 peptides ranging from 3 to 7 amino acids (LEADS-PEP benchmark). Those peptides of the benchmark that performed a covalent bond or where the crystallographic resolution was lower than 3Å were not taken into account for IPRO Docking evaluation. IPRO Docking was compared to AD Vina (gold-standard docking software), and same parameters were used (identical box parameters, exhaustiveness set to 32 and a predefined random number for reproducibility issues (docking softwares use a random number to start calculations, therefore predefining this number, allows the proper reproduction under the same conditions of the docking results). In both cases, the root-mean square deviation (RMSD) values when comparing the first docked conformation (AD Vina / IPRO Docking) with the bioactive conformation found in the X-ray structure, was extracted and compared.

The benchmarking results demonstrated that IPRO Docking consistently yields crystallographic-like predictions when compared to original AD Vina regardless of the peptide size.



	IPRO Docking %	AutoDockVina %
≤ 3-residue	100	67
4-residue	80	30
5-residue	63	36
≥ 6-residue	77	44

Figure 10. IPRO Docking is a peptide focused docking which also includes non- natural and D-aminoacids screening. The obtained results are post-processed through an in-house software which relies in Ramachandran dihedrals angles plot to discard false positives. Comparative data have been obtained from re-docking of active peptides (RMSD<3.5 Å). The results of the percentage of reproduced structures applying each docking technology to the LEADs Database.

Summarizing, for a screen set of peptidomimetic sequences, IPRO Docking returns numerical score (*i.e.* the higher in absolute value, the better) which is associated to each peptidomimetic conformation proposed by the software. Peptides with a score larger than predefined threshold (often set to -7.0 kcal/mol) are sub-selected for further computational evaluation. The selection of this threshold is supported by the empirical rule-of-thumb where compounds around -7.5 kcal/mol should exhibit (if perfectly translated the score to the experimental evaluation) to a potency of 10 to 100  $\mu\text{M}$ .<sup>80</sup> The selected threshold constitutes an acceptable value of potency for an initial hit.

## 2.5. IPRO permeability

Passive diffusion through biological barriers can be translated into a physics-based process where desolvation, diffusion and re-solvation mechanisms take place sequentially. The solvation/de-solvation mechanisms requires that the molecule must have a balanced equilibrium between fundamental molecule size and lipophilicity but also, polarity and conformational dynamics.<sup>81</sup> Chameleonic molecules, which usually are have a large cyclic structures, with the ability to hide and expose polar groups by forming intramolecular interactions.<sup>82</sup> Due this flexibility, these type of molecules can transit between aqueous medium to a more organic/hydrophobic one.<sup>29</sup> However, the mechanisms that drive this amphiphilic behavior is not well understood and therefore its application in the design of new molecules limited.

For polar molecules, the cost of desolvation and later entering a hydrophobic environment (*e.g.* lipid bilayer of cell membranes) is too high. Hence, this case the permeability of the polar molecule through a biological barrier by means of passive diffusion is null or very low. On the other hand, the opposite can be assumed for highly lipophilic compounds that could be retained into biological membrane after desolvation, and for which the re-solvation process is forbidden from an energetically point of view.

For that reason, descriptors accounting for structure flexibility and polarity were introduced into permeability and oral absorption estimations. These structural descriptors are described by Veber rule-of-thumb,<sup>29</sup> which considers the number of rotatable bonds ( $N_{\text{rot}} < 10-20$ ) and polarity derived from topological polar surface area (TPSA)<sup>81</sup> as the main descriptors for positive membrane permeability and not only molecular size and lipophilicity.

Static descriptors, such as MW, HBD, HBA,  $N_{\text{rot}}$  and LogP are intrinsic properties of a particular molecule, are not influenced or modified along the transport mechanism. However, the TPSA is constituted by dynamic descriptor whose value is dependent of the environment and the conformation adopted by the molecule in such environment. Then, TPSA calculation is made on tabulated atomic contributions in seconds of computation, but transferability between tabulated values and molecule of interest is only applicable if chemical spaces are similar.<sup>83</sup> Under this scenario, complex molecules, such as the ones needed to modulate iPPI's, clearly exceed the limitations of TPSA descriptor. A clear TPSA limitation is described by Lokey *et al.* by comparing different regio-isomers with identical polarity estimation but different experimental permeability.<sup>84</sup> Consequently,



description of TPSA for complex molecules may lead to incorrect biological membrane permeability prediction with a negative impact along the drug development process.

A more appropriate polar surface area (PSA) representation can be explained by exploring the conformational landscape of a given molecule to identify low-energy conformation and their associated PSA in an aqueous environment.<sup>85</sup> However, correlation between ensemble of low-energy conformations and PSA representation can be a misleading criterion. Assuming conformational selection is commonly followed by clustering techniques, reducing the permeability classification to few conformations is a potential source of errors. For example, ligand (1-(3,4-dimethoxyphenyl)-3-(3-imidazol-1-ylpropyl)thiourea, with 9 rotatable bonds and whose crystallographic structure has been reported in two different PDB entries (4f9v and 3pb7) features a difference in RMSD lower than 0.25 units, but their polar surface representation are 129 and 135 Å<sup>2</sup> respectively.

To circumvent these aforementioned limitations, IPRO Permeability is a ligand-based algorithm that renders a binary permeability profile based on the input of the structure that was virtually created by the software. The predictions done with IPRO Permeability were compared with manually curated and in-house permeability data obtained through PAMPA assays to validate the algorithm approach. More than 200 peptidomimetic structures, with different topological composition (*e.g.* linear and cyclic peptides) and composition (*e.g.* natural and non-natural amino acids) were used to validate the narrow window where predictions agree with experimentally determined permeability.

Ideally, once the whole *in silico* process is completed, and all post-docking filtering tools are applied, about 10-20 *de novo* peptidomimetics structures are obtained. From that final pull, the more interesting sequences are selected by visual inspection and synthesized by means of SPPS.

## 2.6. Experimental section

SPPS is a straightforward and well established method for the synthesis of short peptides.<sup>86</sup> However, the yield of synthesis drops dramatically when non-natural amino acids *N*-alkylations or post-modifications after peptide-resin cleavage are included in the peptide synthesis process.<sup>87</sup>

On average, peptidomimetics developed by Iproteos contain a large percentage of non-natural amino acids. The high diversity in non-natural amino acids allows the design of more potent peptides. During virtual screening of a pharmacophoric point of the target protein surface, a largest number of amino acid-like family members can be accessed. Moreover, a superior metabolic stability for peptides that incorporated non-natural amino acids is expected, as these challenge enzymatic recognition.<sup>86</sup>

Frequently, at least one N-H from the peptide backbone, which in the docking model is predicted to not be interacting with the target protein, is *N*-alkylated in order to decrease the PSA and increase permeability.



Amide terminus motif at the *C*-terminal or hydrophobic cappings for both *N*-terminal and *C*-terminal parts of the compound are common modifications in Iproteos's peptidomimetics that also differ from lineal natural peptides. These ending modifications aim to reduce the molecule polarity by hiding exposed polar hydrogens, but also are devoted to increase the Van der Waals contacts with the protein surface.



First, the resin was selected based on the final product. For all those lineal peptides that contain an amide motif in the C-terminal, the Rink-Amide ChemMatrix resin was selected for the synthesis. On the other hand, 2-Chlorotrytil chloride resin was selected for those compounds containing a carboxylic acid at the C-terminal peptides, lineal peptides with a secondary amine capping and head-to-tail cyclic peptides.

Once, the first Fmoc-protected amino acid (step 1) is added, the Fmoc-protecting group of this one is removed with a mixture of 20 % of piperidine in DMF. This step is followed by a colorimetric test to assess the removal of the protecting group. Next amino acids were coupled consecutively over the peptide sequence. TBTU and DIEA were used for couplings onto a primary amine, while Oxyma pure and DIC were used for coupling over secondary amines or re-couplings. Both, the amino acid couplings and the removal of the Fmoc group was monitored by means of Kaiser test<sup>88</sup> (primary amines) or the chloranil test<sup>88</sup> (secondary amines) depending on the type of amine which should be detected (step 2) to assess the extent of the coupling reactions. *N*-alkylation was performed on-resin allowing selective peptide backbone *N*-alkylation. Once the lineal peptide sequence was completed, the molecule was cleaved from the resin using an acidic treatment. When the desired product did not contain post-modifications a treatment with a mixture of TFA 95 %; TIS 2,5 %; H<sub>2</sub>O 2,5 % for 90 min, was performed in order to remove amino acid later side-chain protecting groups and obtain the crude of the compound (step 3 and 4). When a head-to-tail cyclization or C-terminal capping were required a mild acidic mixture (DCM 95 %; TFA 5 %) was used to preserve the side-chain protecting groups (step 5).

For cyclization, lineal sequences containing a carboxylic acid group at the C-terminal were dissolved in DMF with NaHCO<sub>3</sub> and DPPA to obtain its cyclization (step 6). Contrarily, to obtain a lineal peptide with a C-terminal capping, the desired secondary amine (typically a piperidine) to be coupled is dissolved with the peptide in DCM, HOAt and EDC (step 7). Afterwards, for both scenarios, lateral side-chain groups were removed by adding acid mixture of TFA 95 %; TIS 2,5 %; H<sub>2</sub>O 2,5 % for 90 min in order to obtain the desired final product.

The crude of the desired product was analyzed by reverse phase high-performance liquid chromatography (RP-HPLC), coupled either to a diode-array or mass spectrometry (MS) detector, or by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF) characterization. Peptidomimetics were purified by RP-HPLC at semi-preparative scale and the pure product was fully characterized by RP-HPLC, RP-HPLC-MS, MALDI-TOF and amino acids analysis.

## 2.7. Physicochemical and ADME evaluation

Biophysical based experiments, typically, permeability across biological barrier, stability in human serum or rat plasma, solubility and biological activity evaluation against the protein of interest, are endorsed experiments for pure compounds. After evaluation, all generated data is collected and introduced in the IPROTech software, where deeper and long-term analyses are implemented to refine the *in silico* predictions. After analysis, a new subset of predicted candidates is synthesized and evaluated following the protocol mentioned before.

### 2.7.1. Solubility experiments

High solubility in aqueous medium is desired for any new drug candidate.<sup>89</sup> However, in order to achieve a permeable bRo5 molecule, the lipophilicity and hydrophobicity of these compounds is exponentially increased in to the detriment of aqueous solubility.<sup>90</sup>

Particularly, poor soluble compounds can drastically affect the preliminary studies results. For example, the peptidomimetic real compound concentration in solution during the assay could diverge from the nominal concentration, therefore underestimating its potency as the real concentration would be lower than the applied. Another scenario would be the formation of aggregates that can precipitate in aqueous buffer (Figure 11) and could produce cell death in *in vitro* assays.

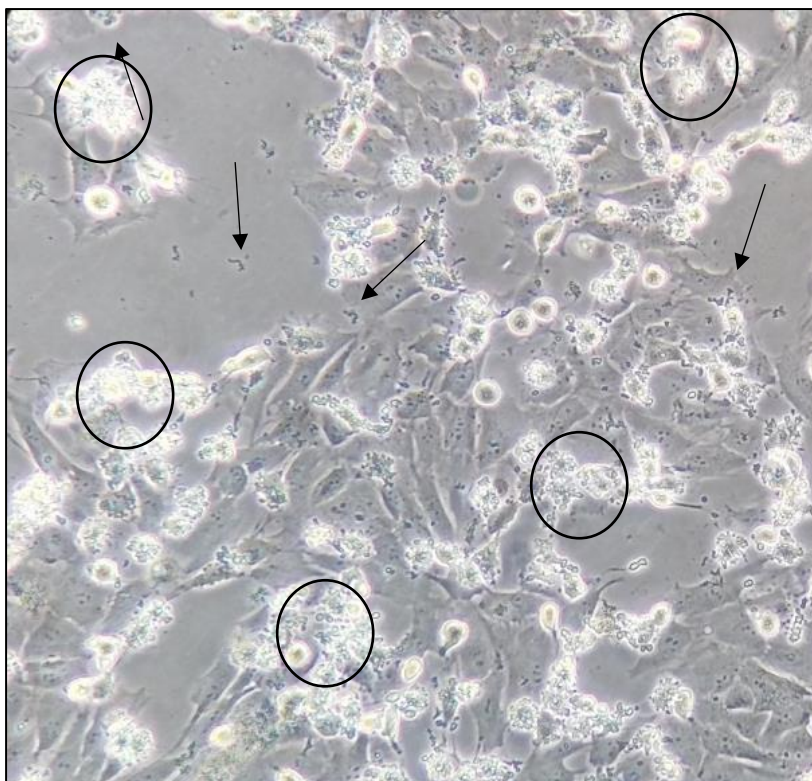


Figure 11. IPR-474 is a peptidomimetic designed to inhibit the interaction with K-Ras and its effectors (chapter 3). Although the molecule was fully soluble at 10 mM in DMSO, when was incubated at 50  $\mu$ M for 2 h in cell medium precipitated. This precipitation could explain the effect of cell death observed under the microscopy. Arrows are pointing out IPR-474 aggregates whereas death cells are enclosed in circles.

On this basis, the solubility of the synthesized compounds was evaluated in the same conditions as the planned experiments. These studies allowed to calculate the real maximum compound concentration when the experiment is performed and therefore preventing from a biased result.

DMSO is the most common and well-established solvent agent for *in vitro* experiments. DMSO is a colorless polar aprotic solvent that can solve both polar and nonpolar molecules. However, its use can be detrimental for NMR or Surface Plasmon Resonance experiments as could mask the N-H signal, not to talk about the cytotoxicity of the solvent.<sup>88</sup> Furthermore, DMSO solvating capacity for very hydrophobic and none soluble

molecules is limited, which represents a major concern when these molecules are need to be solved at high concentrations (>1 mg/mL) to be administered *in vivo*.

In this regard, exploring other excipients that can increase the peptidomimetics solubility and are suitable for both *in vitro* and *in vivo* was a key point for the proper drug development process.<sup>91</sup>

Cyclodextrins (CDs) are cylindrical oligosaccharides complexes that can cage hydrophobic molecules in their inner cavity. CDs hydroxyl groups are exposed to the aqueous medium forming hydrogen bonds with water molecules, then playing as shell around the CD-molecule complex.<sup>90</sup>

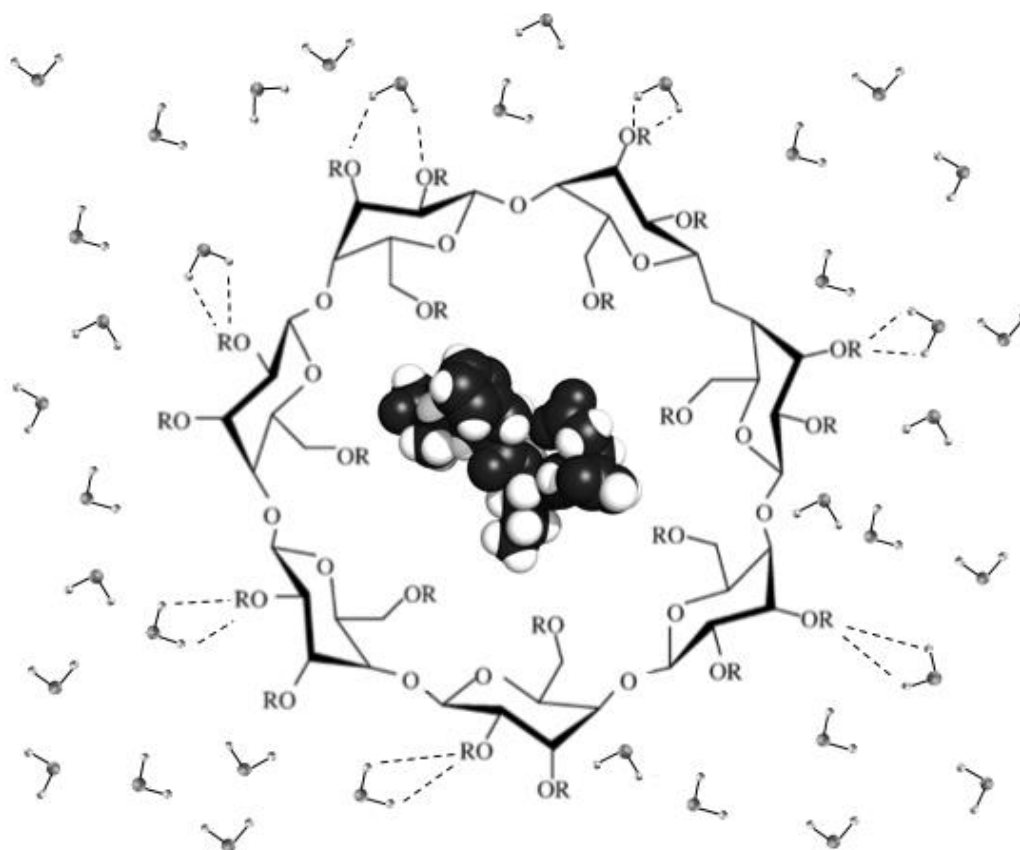


Figure 12. Graphic representation of common cyclodextrins family structure, shielding a peptide molecule in its inner cavity from water molecules.

From cyclodextrin family, hydrophilic cyclodextrins, particularly hydroxypropyl  $\beta$ -cyclodextrin (H $\beta$ -CD) is considered safe and well-tolerated in humans when administrated orally or intravenous,<sup>92</sup> being eliminated through the kidneys via glomerular filtration. Certainly, H $\beta$ -CD safety application is translated into a frequently use in marketed drug formulations.<sup>93</sup>

In conclusion, H $\beta$ -CD is a suitable solving excipient by its capacity to potentially increase the solubility of highly lipophilic and hydrophobic molecules in aqueous mediums. Moreover, broadening possible *in vivo* studies for not soluble compounds. Here, H $\beta$ -CD was successfully applied as excipient in different projects allowing higher solubilization

of peptidomimetics under study, if compared with the use of DMSO. As expected, no toxic effect related with the use of H $\beta$ -CD was found neither on cells nor in mice studies.

### 2.7.2. Parallel Artificial Membrane Permeability Assay (PAMPA)

PAMPA is a lipid-based transport assay method that measures the effective constant permeability ( $P_e$ ) of compounds across an artificial lipid membrane. It is used to determine whether a compound can potentially cross a biological barrier by passive diffusion. Usually, the assay is done in a sandwich-like 96-well plate comprising an acceptor and donor well-plate that are separated by a porous PDVF membrane coated with a mixture of phospholipids that mimics cell membrane lipid layer composition. Compounds under study are loaded into the donor compartments (basal plate), whereas acceptor compartments (apical plate) are loaded with the assay buffer (Figure 13). After incubation time, the compound amount in each compartment is quantified by HPLC-MS. This information is use for the calculation of  $P_e$ , percentage of transport and percentage of compound retained in the lipid membrane after the transport assay.

Several factors can influence PAMPA permeability, such as membrane composition, incubation time, unstirred water layer or DMSO concentration. Another critical point is that PAMPA was intentionally designed to screen small molecule libraries, and underestimation due to membrane retention can be found for highly hydrophobic molecules.<sup>88</sup> However, PAMPA still represents a robust and valid tool that allows rapid high-throughput screening of large number of compounds if it is compared to other tissue-based transport systems or cell-based assays (e.g. Caco-2, and MDCK transport assays). PAMPA is a non-cell-based assay which does not provide information regarding active transport or efflux transporters.

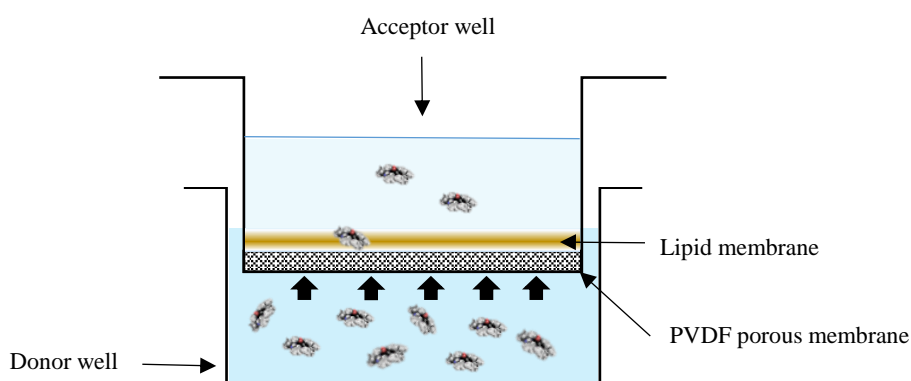


Figure 13. PAMPA, an artificial phospholipid membrane immobilized on a filter is placed between a donor and acceptor compartments. At the start of the test, a drug is introduced in the donor compartment. After the permeation period has finalized (16 h), the concentration of drug is evaluated in both compartments by HPLC-MS.

### 2.7.3. Rat plasma stability

Drug metabolic stability in plasma is critical factor for a proper drug development planning. With the exception of prodrugs, plasma labile structural compounds are removed by pharmaceutical companies from preclinical studies<sup>94,95</sup> due to the rapid degradation in the blood stream which limits they potency and also can lead to a toxic byproducts.

In order to evaluate the stability of drugs in front metabolic activity, fast and inexpensive experimental techniques have been developed. In this regard, a commonly used *in vitro* technique is the determination of drug stability. The proteolytic activity of rat plasma enzymes is more aggressive than human plasma enzymes which compromise the integrity of the molecule under study.<sup>96</sup> Moreover, rodents are a preferred animal to start *in vivo* studies. Thus, representing a well-established go-no-go decision for those compounds that are rapidly degraded in very few minutes.<sup>97</sup> Moreover, in terms of preclinical studies, mice/rat studies are frequently conducted, as are the most popular rodent specie to study the pharmacokinetic, efficacy and toxicological profile of drug candidates. On this basis, a compound which exhibits a rapid clearance and short half-lives will be removed from further experiments, even if it could be more stable in human based experiments.

Therefore, rat plasma stability studies can be used to understand the half-life of a compound and the clearance process prior the drug is administered *in vivo* for preclinical studies.

Rat plasma stability experiment is a technique in which the molecules of interest are diluted in a mixture of PBS pH 7.4 and rat plasma 1:1. The stability of the compound is analyzed at several time points, usually between  $t_0$  min and  $t_{120}$  min and samples are analyzed by a HPLC-MS. The incubation time of the experiment is 2 h as is considered that after this time the rat plasma is no longer active.<sup>96</sup>

Chapter 3: Disruption of the mechanical clutch Talin-  
Vinculin





### 3.1 Introduction: Talin-Vinculin

Cells are permanently exposed to mechanical forces, which directly affects their biological functions such as migration, differentiation, proliferation, and others<sup>98</sup>. These forces are transmitted to cell neighbors or towards the cytoplasm and nucleus compartments by linkage of different proteins that can pull or unfold themselves, playing as a mechanical engine, known as mechanical clutch.

One of the key players in the mechanically sensitive process, so-called mechanosensing, is Talin. Talin is a 2541 residue cytoskeletal  $\alpha$ -helix bundle protein which is made of a large C-terminal rod domain and a N-terminal FERM.<sup>99</sup> Then, applied forces to cell surface are transmitted from the cell membrane to the inner compartments by acting protein filaments. Acting filaments produce Talin  $\alpha$ -helix bundles unfolding, this process stretches Talin which also led exposed Vinculin binding sites. Talin interacts with Vinculin to form a stable protein-protein interaction complex. The PPI formation requires structural unfolding of Talin into separated helical elements. Some of these helices, particularly helix 3 (H3) and helix 4 (H4) of the R3 domain (res. 787-911), when unfolded, can form stable interactions with binding site of Vinculin. Talin unfolding and sub-sequent Talin-Vinculin PPI formation promotes a signaling cascade that ultimately impacts in nucleus flattening and pore formation.

On the C-terminal rod is found a mechanosensitive switch domain R3, which is force-dependent and comprehends 4- $\alpha$ -helices. Under mechanical stress, over 5 pN force gradient,<sup>100</sup> R3  $\alpha$ -helices unfold and consequently exposing the Vinculin binding sites that are buried when Talin is in a folded state.<sup>101</sup> This process of R3 unfolding and the following Vinculin binding initiates the whole Talin's unfolding process and furthermore inhibits Talin's refolding.<sup>102</sup>

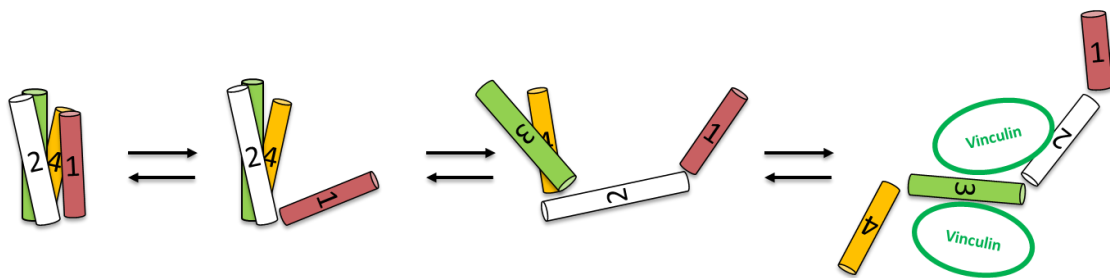


Figure 14. Under mechanical stress, the 4  $\alpha$ -helix of the R3 domain starts the unfolding process gradually, as the mechanical forces over the protein are incremented. Finally, the complete unfolding of R3 domain exposes the vinculin binding sites, allowing Vinculin to interact with the R3 domain which triggers the full unfolding of Talin.

Summarizing, force transmission is mediated by Talin, a mechanosensitive cytoskeletal protein, that unfolds above a threshold force in tissue stiffness and binds to Vinculin, resulting in nuclear flattening and the opening of nuclear pores, thereby increasing Yes-associated protein (YAP) nuclear import.<sup>103</sup>

Then, Talin unfolding and the subsequently Vinculin binding leads to YAP's nuclear translocation. YAP's nuclear translocation promotes the activation of TEAD transcription factors, which induces cell proliferation.<sup>104</sup> YAP plays a major role in the progression of

cancer diseases<sup>105</sup> when it is located in the cell nucleus, and the Talin-Vinculin are necessary enhancers.

In this regard, cancer cells present higher cell membrane rigidity, meaning increased tissue stiffness and as a consequence a displacement of Talin equilibrium towards unfolded conformations leading to an increase of Vinculin binding. Then, this equilibrium displacement ends up with higher levels of YAP inside the nucleus and as a final effect an increase in cancer cell proliferation.

Consequently, this intracellular PPI, Talin-Vinculin, represents a new challenging therapeutic target in the cancer field, so far unexplored. In a similar approach, Goult *et al.*, designed a R3 mutant, R3-IVVI, which has a more hydrophobic buried core and consequently is more resistant to unfold and less prone to bind to Vinculin.<sup>106,107</sup> Thus, validating R3 domain as a potential target to disrupt the Vinculin binding to Talin.

On this basis, targeting the R3 domain that binds to Vinculin should inhibit the PPI and therefore block the unfolding cascade,<sup>108</sup> which hypothetically would provoke solid tumors cell death by preventing YAP translocation. For this purpose, the use of peptidomimetics, which can bind more efficiently to flat and undruggable areas related with PPIs than small molecules, could disrupt the Talin recognition by Vinculin. This approach emerged as a feasible challenge for the Iproteos technology platform.

To demonstrate the advantages of using IPROTech when targeting a challenging intracellular PPI, such as Talin-Vinculin, a research collaboration between Iproteos and the Institute for Bioengineering of Catalonia (IBEC, group of Cellular and molecular mechanobiology, led ICREA researcher Pere Roca-Cusachs, a Key Opinion Leader in the field of mechanosensing) was set.

Dr. Roca-Cusachs has track-record on cell mechanosensing and the complex Talin-Vinculin, publishing during last years in high-peer reviewed journals<sup>109</sup> and having the infrastructure and partners to evaluate the activity of the molecules designed, synthesized, and characterized by Iproteos.

At the present writing moment of this thesis, this was an ongoing project which aimed to obtain a lead candidate, and a family of derivative that could be protected for a patent and be studied for the treatment of solid tumors.

### **3.2. IPROTech hit identification**

After a preliminary evaluation of Talin structure and the mechanism of action that initiates the mechanosensing process, it was decided to target the R3 domain of Talin in its folded state, as targeting an isolated  $\alpha$ -helix motif would require a longer peptide sequence to increase the number of interactions with the receptor protein.<sup>110</sup> The rationale behind this approach was to block the folded conformation of Talin preventing its unfolding and ultimately the binding of Vinculin to R3. Then, if this PPI is inhibited YAP will not be uptake by the nucleus.

### 3.2.1. Proof-of-concept of peptidomimetic approach

First, the IPRO Filter was applied to Talin, identifying the critical hotspots of R3  $\alpha$ -helix folded bundle for drug design, special attention was paid to the presence of hydrophobic and aromatic groups. The protein target (PDB: 1u89) was downloaded from Protein Data Bank. From the NMR conformation ensemble, the first tridimensional structure was select.

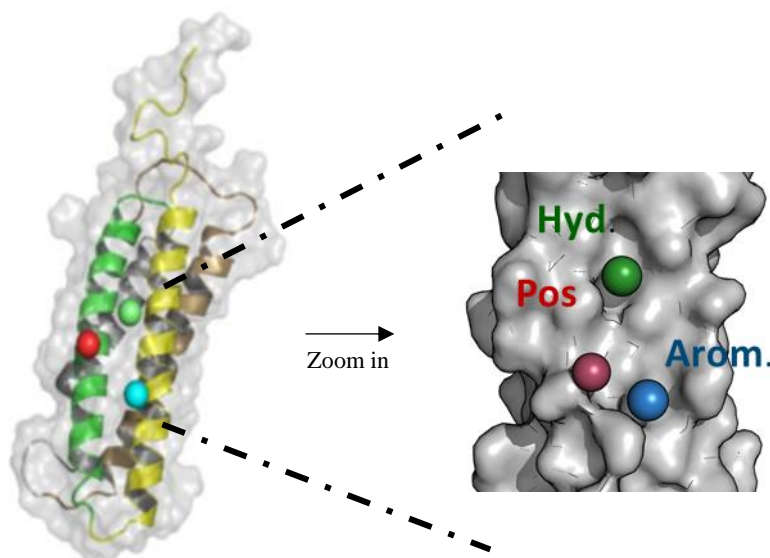


Figure 15. Pharmacophore protein surface representation of Talin after IPRO Filter evaluation. The presence of aromatic, polar and hydrophobic/aromatic features are highlighted as cyan, red and green spheres respectively. Talin helices are highlighted in colors (H3 in green, H4 in yellow).

Initial design identified 4  $\alpha$ -helical peptides (2016/IP-12-01, 2016/IP-12-02, 2016/IP-12-03, 2016/IP-12-04) able to interact with different faces of the Talin R3 domain (this work was carried out previously to the present thesis). Design was inspired in the structural evidences by which  $\alpha$ -bundles favorably interact with complimentary  $\alpha$ -helices. This approach was complimented by the design of a smaller fifth non-helical peptidomimetic (2016/IP-12-05), engineered to interact and be accommodated within a hydrophobic cleft allocated between helices 3 and 4 (H3 and H4).

After a first efficacy assay in which it was evaluated the peptidomimetics capacity to prevent YAPs nuclear translocation when incubated with Mouse Embryonic Friboblast cells (MEFs). 2016/IP-12-05, the only linear peptidomimetic of the 5 compounds (Figure 16, drug 5), was selected by its favorable activity and because is more feasible to base a screening campaign in a linear short peptidomimetic (5 amino acids sequence length) rather than in a  $\alpha$ -helix peptidomimetic with >10 amino acid sequence length (drugs 1-4). Importantly, 2016/IP-12-05 represented the first proof-of-concept of a peptidomimetic approach that inhibits the interaction between Talin and Vinculin and consequently reducing YAP's nucleus uptake.

## YAP Ratio in MEFs Treated with Drugs 1-5

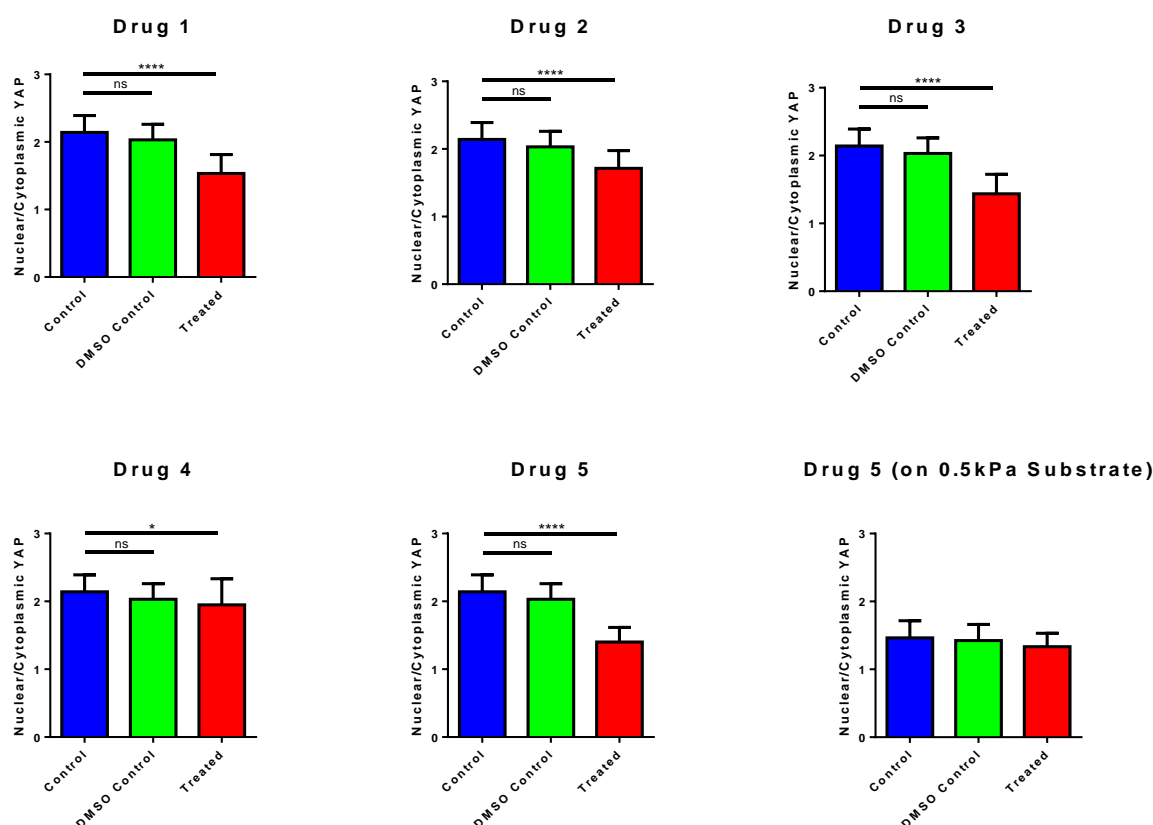


Figure 16. Drugs were incubated with MEFs (Mouse Embryonic Fibroblasts Talin 1<sup>-/-</sup>), for 1 hour at 60  $\mu$ M. Drug 5 (2016/IP-12-05) showed the most promising results. Drug 5 was also tested on 0.5 kPa gels and it was found that there is no effect on a soft substrate. No significant differences between the control and DMSO control or treated cells on the soft substrate were found.

### 3.2.2. First generation of Talin/Vinculin inhibitors

The identification of 2016/IP-12-5 was a proof-of-concept approach to determine the suitability of IPROTech to identify active compounds. Based on experimental evidences, a second iteration was conducted, screening from IPRO Library other potential candidates with similar binding mode and compatibility with previously reported pharmacophoric representation of the binding site (obtained through IPRO Filter). For this reason, the new screened structures were enriched accordingly to IPRO Filter results, yielding a library of more than 100000 structures to be evaluated by IPRO Docking and IPRO Permeability as detailed in Materials and Methods chapter of this thesis.

To escape from linear peptide proposals derived from the IPROTech, and because impossibility to conduct directly docking of cyclic structures, in cases where a *in silico* compound, yielding a good potency and suitable orientation between *N*- and *C*- terminus regions, we proceed to manually generated cyclic analogs.

Once docking and permeability analysis were completed, top-ranked structures were visually inspected to identify the more promising structures, measured in terms of balanced properties (docking score, *in silico* permeability, non-redundant amino acid

composition). As a result, a total of 25 peptidomimetics were proposed for synthesis. Synthesis and purification of peptidomimetics is detailed in Material and Methods chapter of this thesis and required a total time of 8 weeks. Compounds were synthesized with high purity and underwent to experimental evaluation. Initial hits were considered if they displayed a cytotoxic effect larger than 50 % at 50  $\mu$ M in 2 h incubation time, in at least in one of the cell lines relevant to the mechanism of action.

### 3.3. Experimental evaluation

*In vitro* efficacy evaluation was done by cell proliferation assay by the colorimetric assay MTT. From the first generation, 25 evaluated compounds, a total of 3 hits were identified, 2017/IP18-06, 2017/IP-18-22 and 2017/IP-18-32.

The hit-rate of first IPROTech evaluation was of 12 % (3 out of 25 compounds were active) with the presence of non-redundant (unique) scaffold structures (linear and cyclic, different capping groups), thus, allowing us to expand the family variability in eventual hit-to-lead optimization process. Moreover, the optimal peptide length was found to be 3 amino acids (cyclic version required additional amino acids to cyclize the active motif).

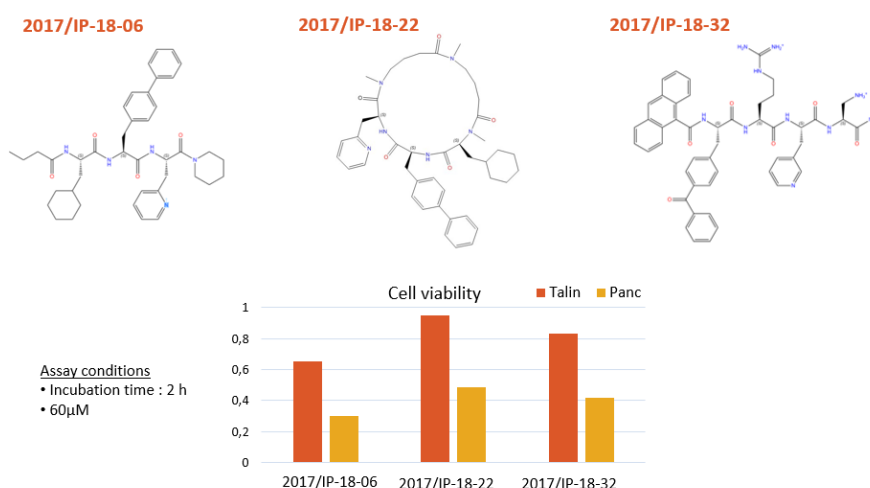


Figure 17. Drugs were incubated with MEFs (Mouse Embryonic Fibroblasts Talin 1<sup>-/-</sup>), and pancreatic cancer cell line, Panc-1 for 1 hour at 60  $\mu$ M. Next, cell death was measured by MTT colorimetric assay. DMSO at 5 % in cell medium was used as control for both cell lines (same proportion was applied when incubated with the peptides).

#### 3.3.1. Hit optimization

2017/IP-18-06 and 2017/IP18-32 were the most active hits in the cell viability measurement. Then, by analyzing the peptide conformation obtained by IPRO Docking (representative of the bioactive conformation), a common structural re-arrangement in terms of protein-ligand interactions and structure-activity relationship (SAR) analysis was derived from both hits.

The biphenyl motif of 2017/IP-18-06 was promoting a more stable hydrophobic and aromatic interaction than single phenyl of 2017/IP-18-32, the presence of Arginine and three aromatic rings capping group of 2017/IP-18-32 were also identified as positive contributors to the binding energy.

### 3.3.2. Second generation of Talin/Vinculin inhibitors

By structurally deconvolution and human-based analysis of initial hits, the optimization was led by the introduction of the amino acids (or family derivatives of each amino acid) identified in SAR analysis. Combination of building blocks of desired peptide length (3 amino acids) ended up in the generation of multiple structural mimetics of initial hits that were evaluated by docking methods.

Additionally, the hit-optimization design included structural modifications supported by the replacement of carboxamide by piperidine at the C-terminal part of the compounds. This replacement had the advantage of introducing a new hydrophobic sub-unit into the peptide design with the aim to enhance through hydrophobic interactions (protein-peptidomimetic) while reducing polarity by removal of highly polar carboxamide moiety.

Hence, constitutive elements (9-anthracenic N-terminus group; one positively charged amino acid; biaryl sidechain, and piperidine C-terminus capping) were combined into a 3-amino acid length peptide sequence, which showed highest docking score and positive *in silico* predicted permeability. Moreover, N-methylated derivatives were also included in the design. These elements were used to create *ad hoc* a custom-made library. The structures generated in the combinatorial process (200 structures) were subjected to IPRO Docking protocol. Top docking ranked structures were visually inspected. Structures with additional number of protein-peptidomimetic contacts and their ability to match pharmacophores revealed by IPRO Filter were submitted to permeability profiling through IPRO Permeability. Finally, a total of 12 molecules were selected for synthesis and purified as described in the Materials and methods chapter of this thesis.

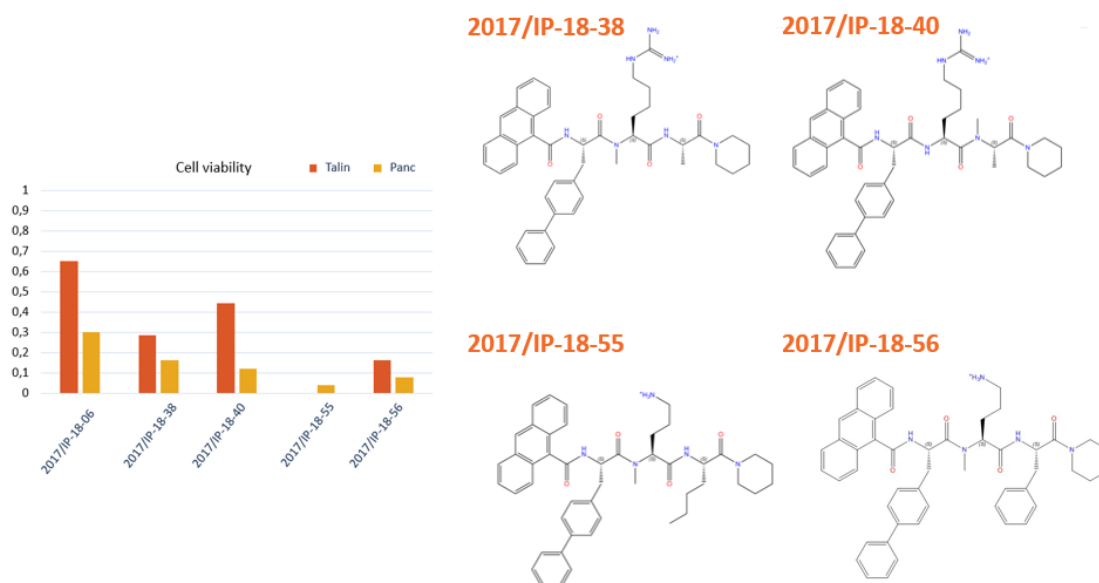


Figure 18. Cell proliferation assay evaluation of optimized-hit molecules. Increase in inhibitory activity was observed in a range of cell survival from 1-to near 40 %. The results of the initial hit (2017/IP-18-06) are showed in the figure as reference.

Experimental evaluation confirmed 4 optimized hit structures (2017/IP-18-38, 2017/IP-18-40, 2017/IP-18-55 and 2017/IP-18-56), with a hit-rate of 33 % (4 out of 12 structures) of active compounds. From these 4 molecules, a lead candidate (2017/IP-18-55) was selected for further experimental evaluation (solubility, rat plasma stability and mechanism of action validation).

In this regard, the activity of the 2 more potent candidates (2017/IP-18-55 and 2017/IP-18-56) was evaluated in a different cancer cell type, SH-SY5Y cells (Figure 19). This cell line is derived from a neuroblastoma cancer and are widely used for the evaluation of drug candidates in neurodegenerative diseases.<sup>111</sup> This cell line has a less stiff tissue and therefore is expected a loss of activity for a given compound that is producing cell death by direct inhibition of the mechanosensing mechanism.



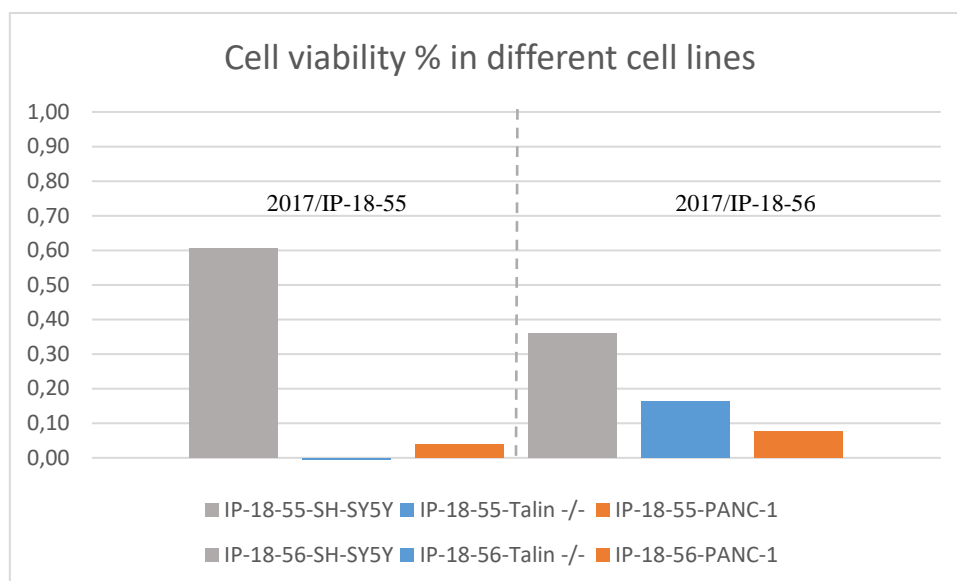


Figure 19. 2017/IP-18-55 and 2017/IP-18-56 were incubated for 1 h with SH-SY5Y cells in the same conditions of evaluation in Talin -/- and PANC-1 cells. In this figure the results obtained for both peptides when incubated in SH-SY5Y cells (grey bars) was compared with the previous results obtain when the peptides were incubated with Talin -/- (blue bars) and PANC-1 (orange bars).

Thus, showing a specificity of the compounds for stiff cancer cells and therefore corroborating the mechanism of action. Actually, both peptidomimetics were less effective in SH-SY5Y cell line than in the other two cell lines, for which Talin is speculated to have a larger impact. Hence, pointing that its killing effect over cancer cells is related with the mechanosensing cell process.

### 3.3.3. Lead Characterization

In order to properly evaluate of the effect of the compounds in cells cultures, 2017/IP-18-55 solubility in cell culture medium (DMEM) supplemented with 5 % DMSO at 37° C was studied. The, the solubility concentration found for this compound was over the half maximal inhibitory concentration (IC<sub>50</sub>), ~5 μM in cells, being 28 μM. Then, not having a limiting impact in the MTT assay results, because the peptidomimetic would always be soluble in higher concentration than its IC<sub>50</sub>. However, such poor solubility limited the applicability of the compound in other *in vitro* experiments or even more importantly for its *in vivo* evaluation.

Moreover, to carry out Surface Plasmon Resonance (SPR) experiments (experiments in progress) and Nuclear Magnetic Resonance (NMR) experiments (also in progress), an alternative excipient to DMSO was a crucial step in order to avoid possible signal interferences. Tween20 at low concentration is well tolerated in NMR experiments and can increase the solubility of a given compound. On this basis, 2018/IP-18-55 was found to be soluble at final concentration of 45 μM in NMR buffer at 0.025 % Tween20 at RT.

Solvent	Conc. (μM)
DMEM at 5 % DMSO	28 ± 2
NMR buffer at 0.025 % Tween20	45 ± 1

Finally, H $\beta$ -CD was the excipient used by safety track record in *in vivo* studies. Then, solubility of 2017/IP-18-55 was evaluated at 15 % and 5 % of H $\beta$ -CD in phosphate buffer pH= 7.4. All data were expressed as the mean  $\pm$  SD.

Solvent	Conc. (mM)
PBS at 15 % H $\beta$ -CD	5,055 $\pm$ 0.047
PBS at 5 % H $\beta$ -CD	2,446 $\pm$ 0.034

On the other hand, high drug stability in front proteolytic activity of plasma enzymes is a key feature for its successful therapeutic application. For this reason, 2017/IP-18-55 was incubated with rat plasma, which presents higher proteolytic activity than human plasma. The peptide was fully stable after being incubated 2 h in rat plasma. Rat plasma proteolytic activity is considered inactive after 2 h, Benfluorex is a small molecule used as positive control to corroborate rat plasma proteolytic activity.

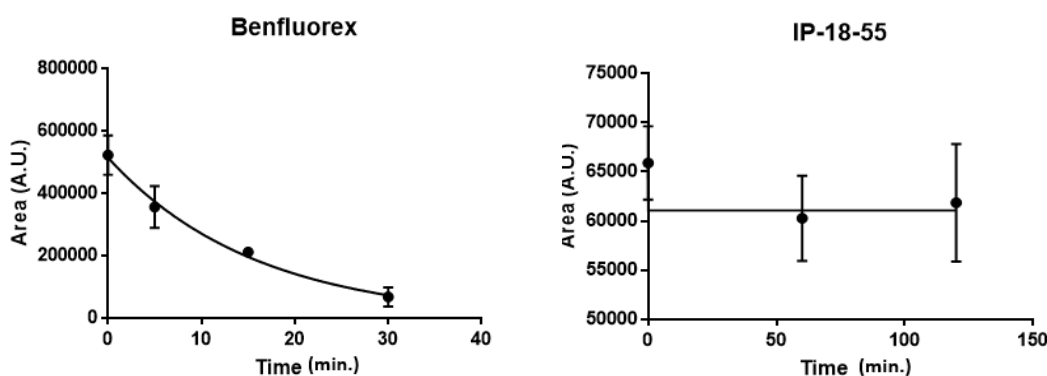


Figure 20. Benfluorex, a small molecule was used as a positive control, and 2017/IP-18-55 were incubated at 10  $\mu$ M at 2,5 % DMSO with Rat Plasma/ PBS 1:1 for 2 h. For Benfluorex samples were harvested at 0, 5, 15 and 30 min. Meanwhile, For 2017/IP-18-55 time points were 0, 60 and 120 min. The final concentration of 2017/IP-18-55 did not decrease after 2h incubation time, meaning an extraordinary stability in front rat plasma proteolytic activity. Error bars represent  $\pm$  SD (n=2).

Next, to validate if the compound was being internalized by cells, 2018/IP-18-55 was incubated at 60  $\mu$ M with SH-SY5Y cells for standard incubation time, 2 h. After incubation and following the described protocol in the Materials and methods chapter of this thesis, the quantity of the peptide internalized by seeded cells was calculated. 3.5 % of it was internalized by cells.

Finally, by applying Traction Force Microscopy (TFM) technique, the rational design of the compound mechanism of action to inhibit Talin's unfolding was studied (Figure 21). Under force stress, R3 domain of Talin is the first  $\alpha$ -helix bundle to unfold over a rigidity threshold ( $\sim$  5 kPa, blue dots). When Talin is incubated with 2018/IP-18-55 this unfolding process is inhibited (orange dots), even at high gel stiffness. Therefore, validating peptide mechanism of action, which prevents Talin unfolding under force stress. However, the therapeutic effect of the stabilization of the folded conformation of R3 is not proved in

this biophysical experiment as neither the straight interaction with Vinculin nor the YAP's nuclear translocation were evaluated.

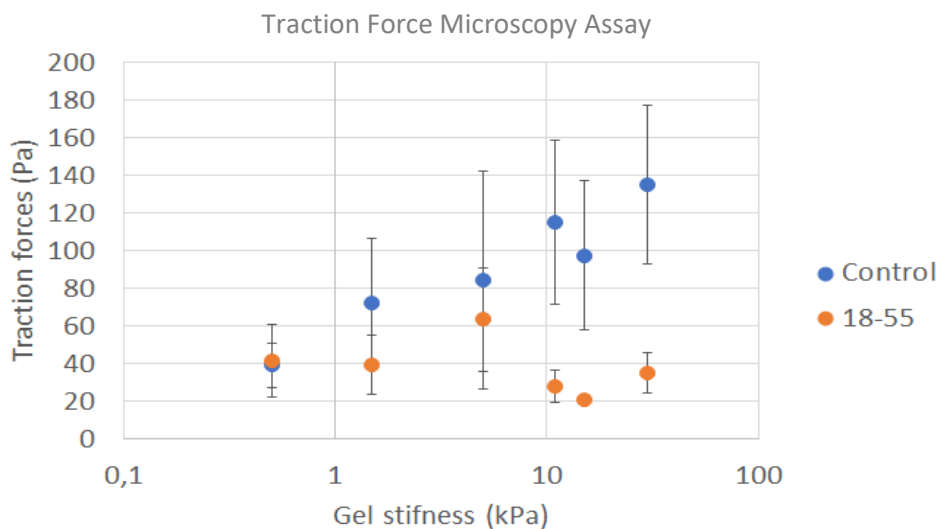


Figure 21. Talin sets a rigidity threshold that triggers increased force transmission. Average forces exerted by cells plated on fibronectin-coated polyacrylamide gels of increasing rigidity. Difference between control cells (blue) and cells incubated with compound 2017/IP-18-55 (orange) show that the peptide prevents Talin's unfolding

### 3.4. Conclusions

Talin-Vinculin PPI was subjected to IPROTech evaluation. Disruption of the complex interface was conducted via stabilization of one of the protein partners (Talin), whose unfolding was tagged as key role in the activation of the cell proliferation cascade.

By studying Talin folded conformation (PDB code: 1U89), IPRO Filter revealed the importance of hydrophobic and aromatic sub-units to be inserted into the peptidomimetic design. Accordingly, a sub-set library from proprietary IPRO Library (100000) was used in an exploratory study to identify by IPRO Docking, the most active molecules. From this analysis, and sub-sequent IPRO permeability study, 25 peptidomimetic sequences were synthesized and experimentally evaluated. In the first screening, a total of 3 hits were identified, with different composition and topology, paving the way for hit optimization.

After careful analysis between activities and structures of active hits (SAR analysis), and later deconvolution of its primary active residues or building blocks, a new battery of structural mimetics was designed by applying IPROTech. A total of 12 potentially active inhibitors were synthesized, identifying 4 new structures with larger *in vitro* affinity.

From hit optimization procedure, a lead candidate, 2017/IP-18-55, was selected and profiled in terms of solubility and rat plasma stability. The lead candidate, with an  $IC_{50}$  lower than 3  $\mu$ M when evaluated in cells plus a good solubility profile in different excipients (H $\beta$ -CD and Tween20) and remarkable rat plasma stability profile (> 2 h). Moreover, it was proven to be internalized in cells by passive diffusion mechanism and to effectively induce the desired pharmacological effect (e. g. prevent the unfolding and inactivate the YAP translocation as speculated by initial design hypothesis).

These results encouraged Dr. Pere Roca's team to carry out exhaustive *in vitro* experiments as NMR and SPR. Additionally, *in vivo* experiments in pancreatic cancer-induced mice were carried out with 2017/IP-18-55 (data not shown). Unfortunately, both biophysical experiments only showed some binding effect signal over the R3 domain at concentrations  $>100 \mu\text{M}$ . Furthermore, experiments in pancreatic cancer mice models did not show any therapeutic activity when administered with the peptidomimetic.

Hence, although initially TFMs experiments validated the stabilization of R3 by 2017/IP-18-55, these results were not consistent with NMR and SPR data. The lack of any determining evidence that can validate the mechanism of action of the peptidomimetic along with no significant efficacy in tumor size reduction frustrated the possible applications of 2017/IP-18-55 for the treatment of solid tumors. Thus, emphasizing how important is to have a valid technique beyond simple proliferation assay to discern between hits and false positives. A false positive peptidomimetic could provoke cell death by a different mechanism and then compromise a second-round screening campaign.

In this scenario, Dr. Pere Roca's Lab has planned to introduce a new screening technique by incubating candidate compounds with Talin knockout cell line and pancreatic cancer cell line PANC-1. Thus, allowing straightforward screening of those compounds that directly target the desired mechanism. Therefore, a positive compound should selectively kill cells where Talin is expressed (PANC-1) and not the Talin knockout cell line. Meanwhile a false positive theoretically will kill cells from both cell lines.

In this regard, it has been planned to screen back all positive compounds in order to find some that could be able to produce this effect, and also to evaluate any future drug. This process will determine the perspectives of this Talin/Vinculin project.



Chapter 4: Targeting the RAD51-BRCA2 interaction in  
the DNA repair mechanism



#### 4.1. Introduction: RAD51-BRCA2

Environmental factors<sup>112</sup> or DNA replication<sup>113</sup> during cell division can promote damage in the DNA integrity. The result of DNA damage can produce single strand-breaks (SSBs) or double strand-breaks (DSBs), which can induce cell death or even worse, deletions, translocation or other mutations that can evolve in a health issue.

To overcome sequence errors and facilitate proper cell-cycle machinery, DNA presents several partly overlapping repair mechanisms.<sup>114</sup> Nucleotide-Excision Repair (NER) and Base-Excision Repair (BER) are devoted to solve SSBs meanwhile Homologous Recombination (HR) and Non-Homologous End Joining (NHEJ) are focused in DSBs.<sup>115</sup>

In this regard, some types of cancer are featured by a quick proliferation compared to normal cells, and therefore their dependency of DNA repair mechanisms is more critical. Poly(ADP-ribose) polymerases (PARPs) play a key role in DNA repair mechanisms, particularly in BER. Then, for those types of cancer with a pronounced cell proliferation, inhibition of PARPs could drastically stop cancer growth and eventually reduce tumor size. Exploiting this idea appeared PARPs inhibitors.

The idea behind PARP inhibitors is synthetic lethality, which consist in the use of PARP inhibitors (PARPi) in cancer cells that are already deficient in another DNA-Repair mechanism, as for example BRCA-deficient tumor cells.<sup>116</sup> BRCA1 and BRCA2 genes are known to have a great impact in DNA repair mechanism, particularly in HR. Therefore if only one gene is working properly, PARP1, and it gets inhibited, this action will induce cell death, meanwhile normal cells will revert the situation by activating another compensatory pathway.<sup>117</sup>

In 2014 the FDA approved the first PARP inhibitor from AstraZenca, a small molecule called Olaparib. Other compounds like Niraparib, Veliparib, Rucaparib have been approved more recently.<sup>118</sup> This, shows the high pharmaceutical interest behind this novel approach.<sup>119</sup> However, the application of PARP inhibitors is limited by the capacity of DNA repair mechanism recovery.

Indeed, fast mutation of cancer cells can lead spontaneously to DNA repair activities reactivation or also tumor cells can survive to PARPi by activating other DNA repair mechanism paths. Examples of acquired resistance is the case of BRCA1- and BRCA2-deficient carcinoma cells, or myeloid chronic leukemia cancer and some types of Breast Cancer.<sup>120</sup>

Homologous Recombination is one of the main causes of cancer resistance, and the interaction between BRCA2 and RAD51 (Bacterial RecA homolog DNA recombinase) is a central step in the HR activation cascade. Briefly, BRCA2 sequesters RAD51, facilitating the formation of RAD51-single stranded DNA filaments complex which starts HR.<sup>121,122</sup> Moreover, exploration of different types of PARPi-resistant tumors pointed out an inverse relationship between RAD51 presence in the nucleus and clinical efficacy of Olaparib.<sup>123</sup> More interestingly from a therapeutic point view, it is the synergistic effect of Olaparib and RAD51/BRCA2 inhibitors.<sup>123,124</sup>



BRCA2, a 3418-amino acids protein, binds RAD51 by an evolutionary conserved BRC motif, which is present in the protein as series of 8 short-peptides with a total of 35 involved residues and a common binding motif “FxxA”, it can also bind by the C-terminal domain of BRCA2, although this last is only able to bind RAD51 nucleoprotein filaments.<sup>125</sup>

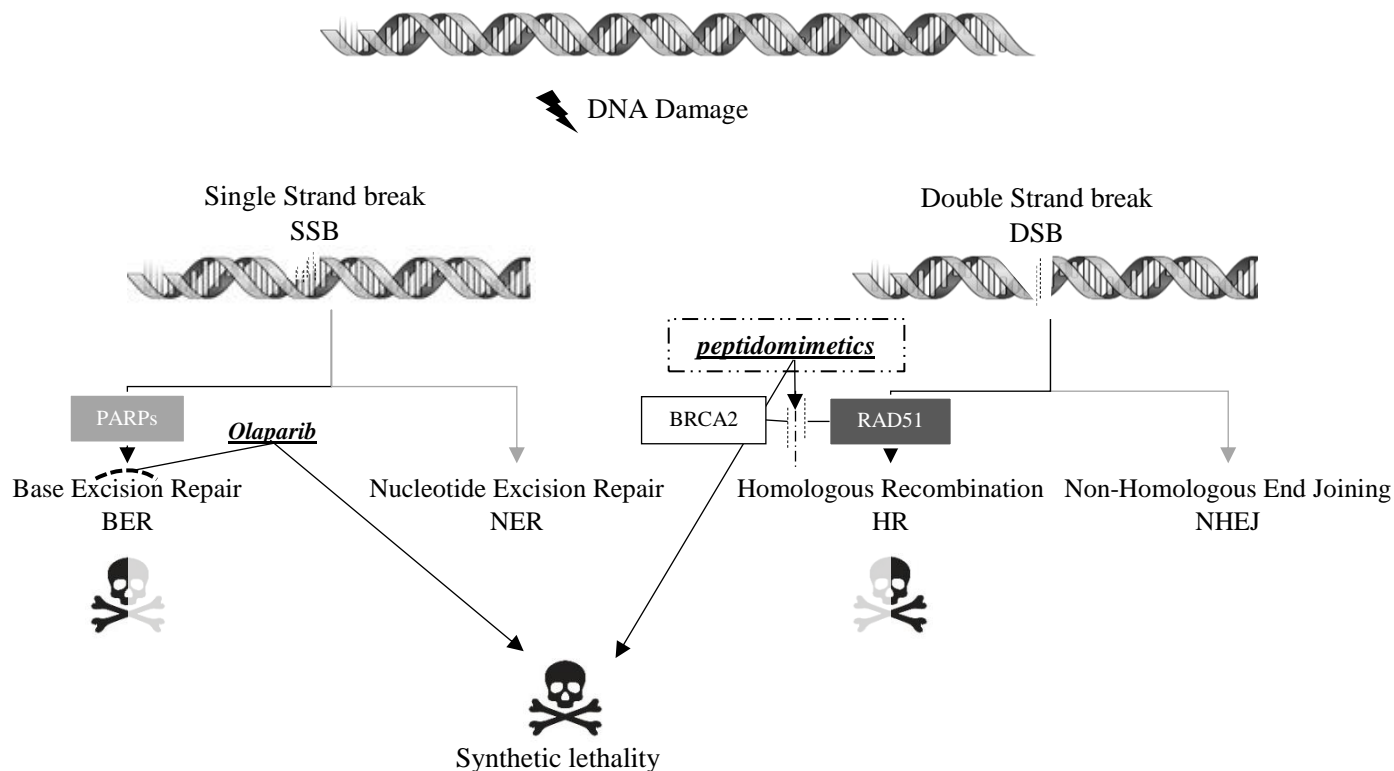


Figure 22. After DNA damage single strand break or double strand break mechanism repairs are activated. BER is the main mechanism of SSB and PARPs protein play a key role in this process. On the other hand, HR is the most relevant repair mechanism of DSB and the RAD51-BRCA2 PPI switches on the downstream protein cascade. If one mechanism is inhibited, for example by using Olaparib (a PARPi small molecule) the other can overcome the situation in order to repair the DNA. However, if both mechanisms are inhibited simultaneously, cells will die automatically. This is the rationality behind the Synthetic lethality approach.

More recently, it was published a set of small molecules disruptors of RAD51-BRCA2 that in combination with Olaparib lead to cell death in a Pancreatic Cancer cell line.<sup>126</sup> However, after a virtually screen 1.5 million compounds and develop a chemical library (41 analog compounds) from the initial hit ( $EC_{50} = 53 \pm 3 \mu\text{M}$ ), the best candidate showed an  $EC_{50} = 8 \pm 2 \mu\text{M}$ . Another example of small molecules targeting BRCA2-RAD51 was published by Scott *et al.* They published a fragments hits library, evaluated with isothermal titration assay calorimetry assay (ITC)<sup>127</sup> with a milimolar potency ( $570 \mu\text{M}$ ) to low-micromolar potency ( $18 \mu\text{M}$ ). This work was guided by the binding site of a tetrapeptide derived from a BRC repeat, the Ac-FHTA-NH<sub>2</sub> peptide.

Altogether, it is worth to highlight again how challenging is disrupt a PPI with small molecules, even applying if an extensive effort is done by using fragment derived libraries or high demanding computational resources.

On the other hand, A Trenner *et al.*,<sup>128</sup> as a proof-of-concept, applied the use of peptides to inhibit the BRCA2-RAD51 PPI. To do that, the BRCA2 BRC repeat able to inhibit BRCA2-RAD51 interaction was fused to the CPP nona-arginine (R9). Peptide

coincubation with Olaparib at 1  $\mu$ M concentration lead to reduction in cellular viability in HeLa cells with an IC<sub>50</sub> value around 10  $\mu$ M. However, the low peptide metabolic stability when incubated with cells ( $t_{1/2} < 1$  h) along with a low potency challenges the future therapeutic application of this strategy. Nevertheless, this work opened the door to optimize the peptide sequence in order to improve both stability and activity.

Guided by the interesting work done by Prof. Alexander Sartori's group, from the Institute of Molecular Cancer Research (University of Zurich), about the DNA repair mechanism, and particularly about the BRCA-RAD51 inhibition with a BCR-CPP, Iproteos's team decided to study this PPI and the applicability of its technology to obtain peptidomimetics able to modulate this interaction.

After an extensive bibliographic search and a visual inspection of the RAD51 surface area, it was decided that IPROTech could be successfully applied in this PPI. For the success of the project, a collaboration with Sartori's group was established. In this collaboration Sartori's team will carried out the efficacy evaluation of the compounds designed and synthesized by Iproteos.

#### **4.2. IPROTech hit identification**

The reported PDB structure (PDB; 4B3B) of a short peptide Ac-FHTA-NH<sub>2</sub> derived from the BCR4 repeat that binds to RAD51 was used as a template for the *in silico* studies. Here, IPRO Filter application step over the RAD51 protein surface was circumvented as the docking studies were directly focused on the same Ac-FHTA-NH<sub>2</sub> binding site.

Likewise, IPRO Docking was applied on the acetylate tetrapeptide in order to establish a relationship between the Docking results and the possible experimental binding, as the peptide  $K_d = 250 \pm 50$   $\mu$ M, measured by ITC.

Next, a 80000 tri- and tetrapeptidomimetics library was virtually screened against the binding site of the short peptide mentioned before. Top 10000 ranked structures were docked again but increasing the exhaustivity of the computational process, were the time of calculation for each molecule was duplicated. Then, 100 best docked peptides were studied by applying the IPRO Permeability tool in order to discard those molecules that were predicted to be not permeable across biological barriers.

Finally, visual inspection was conducted to select a total of 5 peptidomimetics with different binding mode, stable binding and good protein-ligand contacts. These structures were proposed for synthesis (see Table 1).

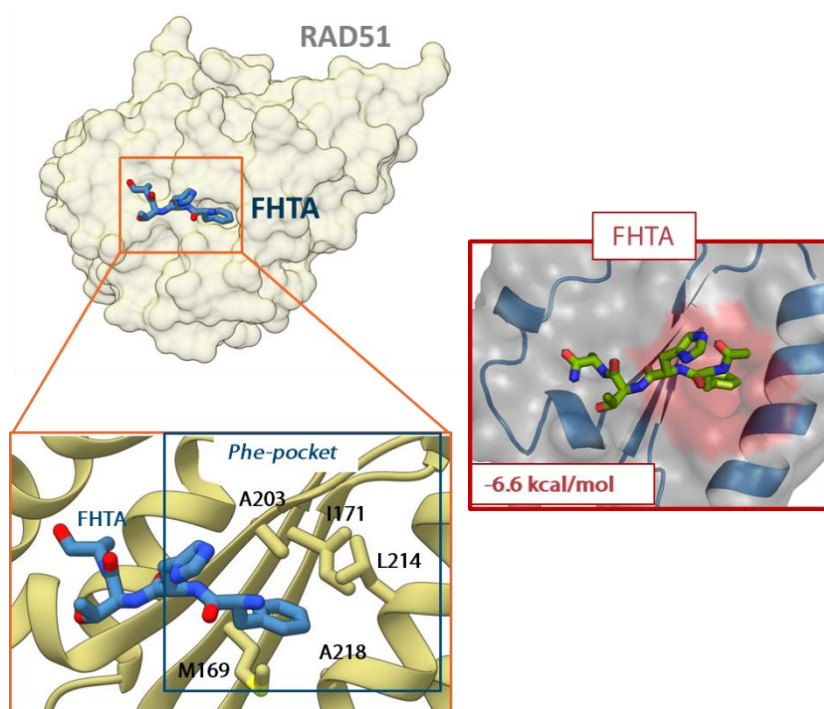
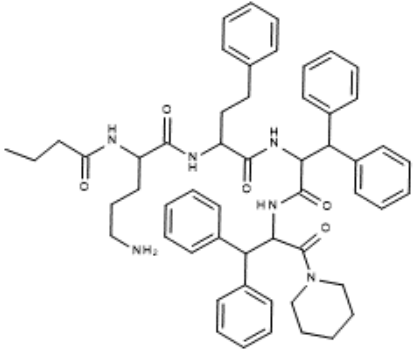
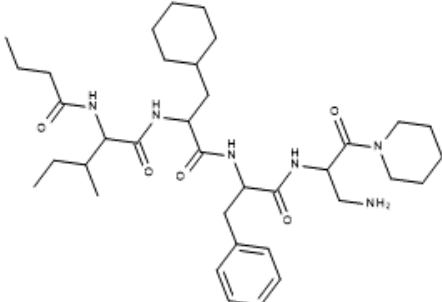
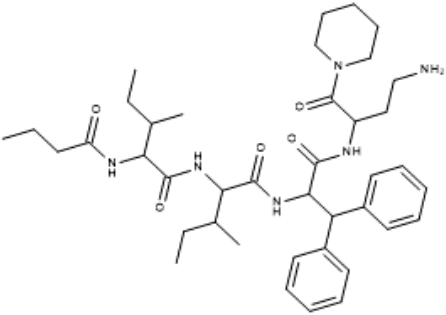
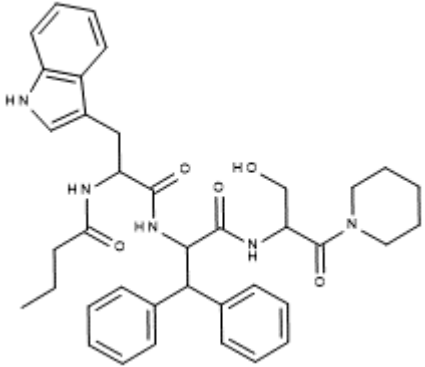


Figure 23. Image obtained from the PDB structure 4B3B. IPRO Docking was applied to see if the *in silico* tool was able to reproduce the Ac-FHTA-NH<sub>2</sub> peptide pose when re-docked. Moreover, it was expected that all those peptidomimetics with a better docking score than -6.6 kcal will improve binding potency when evaluated experimentally.

Iproteos Code	Sequence	Formula	Structure	Docking score (Kcal/mol)	IPRO Permeability (Å <sup>2</sup> )
IP-15-01	Butyryl-Ornithine-Homophenylalanine-Diphenylalanine-Piperidide	C <sub>54</sub> H <sub>64</sub> N <sub>6</sub> O <sub>5</sub>	 <p>The structure shows a linear peptide chain with four amino acid residues: Butyryl, Ornithine, Homophenylalanine, and Diphenylalanine, terminated with a piperidide group. The side chains include a butyl group, a 3-aminopropyl group, a 2-phenylethyl group, and a 1,1-diphenylethyl group.</p>	-9,6	145
IP-15-02	Butyryl-Isoleucine-Cyclohexylalanine-Phenylalanine-Diaminopropionic acid-Piperidide	C <sub>36</sub> H <sub>58</sub> N <sub>6</sub> O <sub>5</sub>	 <p>The structure shows a linear peptide chain with five amino acid residues: Butyryl, Isoleucine, Cyclohexylalanine, Phenylalanine, and Diaminopropionic acid, terminated with a piperidide group. The side chains include a butyl group, an isobutyl group, a cyclohexylmethyl group, a benzyl group, and a 2-aminoethyl group.</p>	-9,4	134

<p>IP-15-03</p>	<p>Butyryl-Isoleucine-Isoleucine- Diphenylalaline- Diaminobutyric acid-Piperidide</p>	<p><math>C_{40}H_{60}N_6O_5</math></p>	 <p>The structure shows a linear peptide chain starting with a butyryl group, followed by two isoleucine residues, a diphenylalaline residue, and a diaminobutyric acid residue. The diaminobutyric acid is linked to a piperidine ring via its carboxyl group.</p>	<p>-9,2</p>	<p>138</p>
<p>IP-15-04</p>	<p>Butyryl-Tryptophan- Diphenylalanine-Serine- Piperidide</p>	<p><math>C_{38}H_{45}N_5O_5</math></p>	 <p>The structure shows a linear peptide chain starting with a butyryl group, followed by a tryptophan residue, a diphenylalanine residue, and a serine residue. The serine is linked to a piperidine ring via its carboxyl group.</p>	<p>-9,2</p>	<p>141</p>

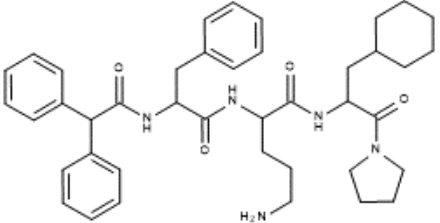
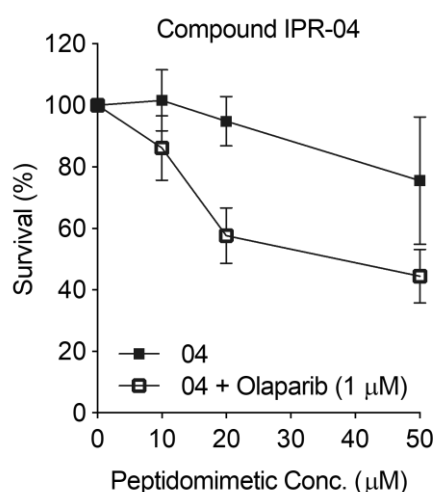
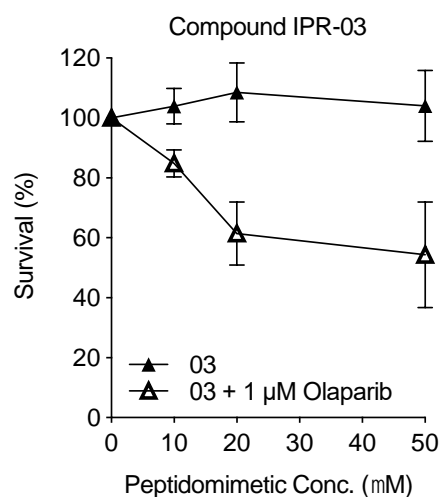
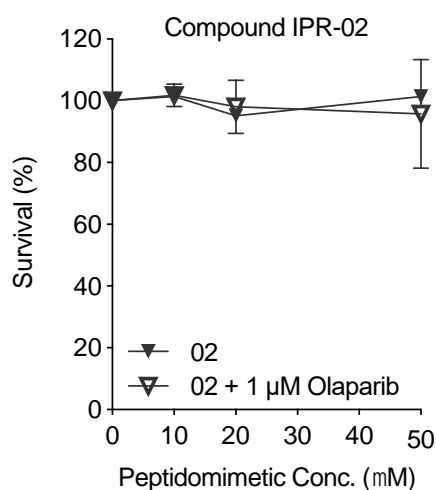
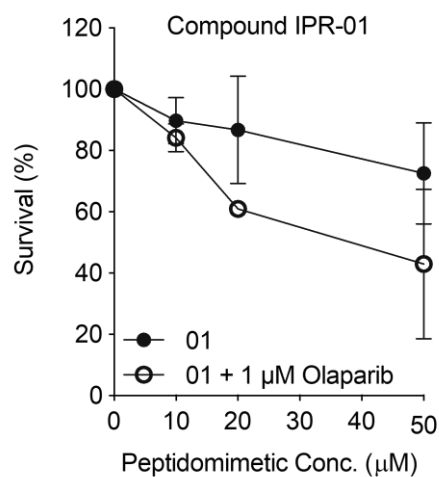
IP-15-05	Diphenyl-Phenylalanine- Ornithine-Cyclohexylalanine- Pyrrolidide	$C_{41}H_{53}N_5O_4$		-7,1	145
----------	--	----------------------	---	------	-----

Table 1. After a final visual inspection, 5 peptidomimetic sequences were selected for synthesized. The batch includes tri- and tetra-peptidomimetics. All sequences have a secondary amine capping for the C-terminal and a hydrophobic group as *N*-terminal capping. Based in the benchmarked data, IPRO Permeability threshold was set at 170 and therefore for all peptides were predicted as permeable. It was decided to not extend to more than 5 molecules the first batch instead of 10 as the resources available at that moment were limited.

### 4.3. Experimental evaluation

A clonogenic survival assay was used<sup>129</sup> to perform a first *in vitro* screening of the compounds efficacy in cells, in which the cell death is measured at different compound concentrations with the presence or absence of Olaparib. Compounds were evaluated by co-incubating them at different concentrations (10, 20 and 50  $\mu\text{M}$ ) in 1  $\mu\text{M}$  of Olaparib with HeLa cells for 10 days. Furthermore, same peptide concentrations were incubated with HeLa cells without the presence of Olaparib.

A positive hit should be inactive or have a low potency when cells are not treated with Olaparib at the same time, as cancer cells could survive by activating the repair mechanism that is not inhibited. On the other hand, co-treatment with the peptidomimetic and Olaparib should be translated in a synergistic effect which eventually will cause cell death.



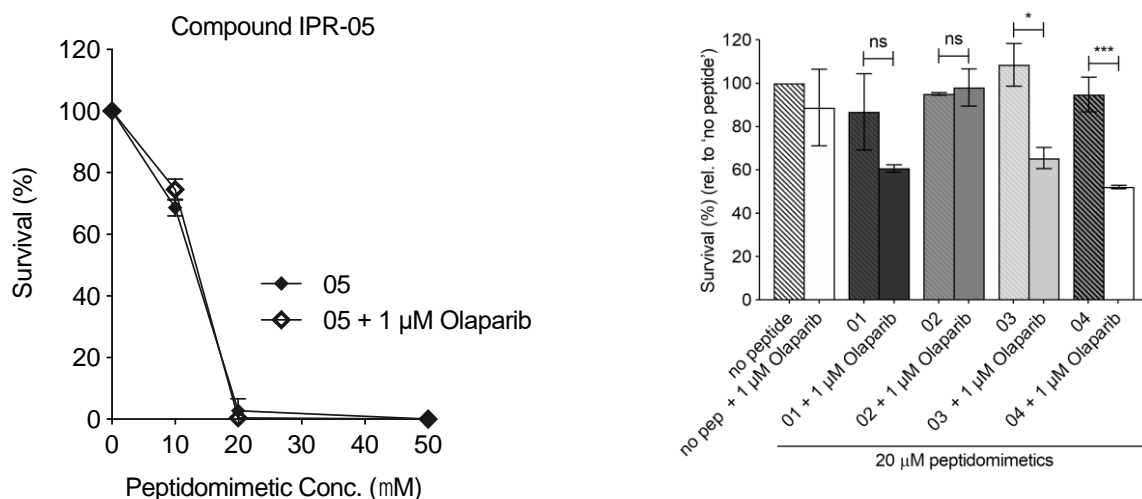


Figure 24. Peptidomimetics (IP-15-01; IPR-01, IP-15-02; IPR-02, IP-15-03; IPR-03, IP-15-04; IPR-04 and IP-15-05; IPR-05) were incubated with HeLa cells for 10 days at 0.25 % DMSO and 10 % Fetal Calf Serum (FCS) at 10, 20 and 50 μM. Cell medium was replaced by fresh cell medium, which contained the same compound concentration, on daily basis. In parallel, the same experiment was performed co-incubating the compound under study with Olaparib (1 μM). All compounds were compared at 20 μM by establishing a relationship with no drug treated cells. In these conditions, IP-15-04 was the compound with higher potency.

From the 5 synthesized peptidomimetics, IP-15-01, IP-15-03 and IP-15-04 showed the desired synergistic effect when co-incubated with Olaparib. More interestingly, IP-15-03 was inactive without the presence of Olaparib, but when the PARPi was applied in the cell culture there was a significant cell death. However, the most potent compound was IP-15-04, when the effect of all peptidomimetics at 20 μM with 1 μM of Olaparib is compared.

Iproteos Code	H <sub>2</sub> O at 5 % DMSO (μM)	DMEM at 0,25 % DMSO + 10 % FCS (μM)	Pe (cm/s)	Transport %	Retention %	Internalization in cells %
IP-15-01	193 ± 8	46 ± 3	1,0E-08 ± 8,0E-09	0,1 ± 0,01	25,7 ± 0,9	0,1 ± 0,07
IP-15-03	349 ± 2	50 ± 6	2,6E-08 ± 1,0E-09	0,2 ± 0,03	16,5 ± 4,1	0
IP-15-04	82 ± 16	48 ± 2	0	0	66,6 ± 2,5	0,4 ± 0,06

Table 2. All experiments were carried out as is described in Materials and Methods chapter. Solubility in H<sub>2</sub>O at 5 % DMSO was performed as a control to discard any possible issue related with a low solubility. However, due the low solubility of the three peptidomimetics, their solubility in the same assay conditions was evaluated. PAMPA assay (Pe, Transport % and Retention %) along with cell internalization % were done as evaluation of peptidomimetics permeability. Data are expressed as the mean ± SD.



For the 3 positive hits, their physicochemical properties were studied. The three peptidomimetics had a very low solubility in 5 % DMSO in water, and therefore their solubility in the *in vitro* assay conditions was also evaluated to discard any lack of efficacy due to low solubility. The maximum solubility of IP-15-01, IP-15-03 and IP-15-04 in cell culture medium was calculated to be 46  $\mu\text{M}$ , 50  $\mu\text{M}$  and 48  $\mu\text{M}$  respectively. As the maximum concentration studied in cells was 50  $\mu\text{M}$  no significant issues related with compounds solubility was expected. On the other hand, PAMPA evaluation showed a very low or null permeability across the artificial lipid membrane, of the assay, this permeability data is in good agreement with the internalization assay in SH-SY5Y cells results.

At this point we hypothesized that the hydrophobic features of the three identified compounds scape form the evaluation capacity of the PAMPA assay. Hence, this technique could underestimate the compounds permeability. Furthermore, we were not able to obtain a positive result neither by evaluating its permeability by studying their internalization when incubated in cells. This could be explained for a very low internalization of the compounds when incubated with cells, with an internalized quantity which would be out of the limit of the detection but that at same time could be potent enough to block the RAD51-BRCA2 interaction once located inside the cell.

Fortunately, the disappointing permeability results were balanced by the outstanding results obtained in cells and therefore, motivated us for a second iterative cycle of the technology in order to optimize the potency of the final candidates. For this purpose, IP-14-04 which showed the higher activity structure was used as structure template.

#### 4.4. Hit optimization

IP-15-04 structure has two natural amino acids, threonine and serine, and one non-natural amino acid, diphenylalanine. In addition, a piperidide group and a butryl moiety had been coupled in the C-terminal and N-terminal of the lineal peptide, respectively.

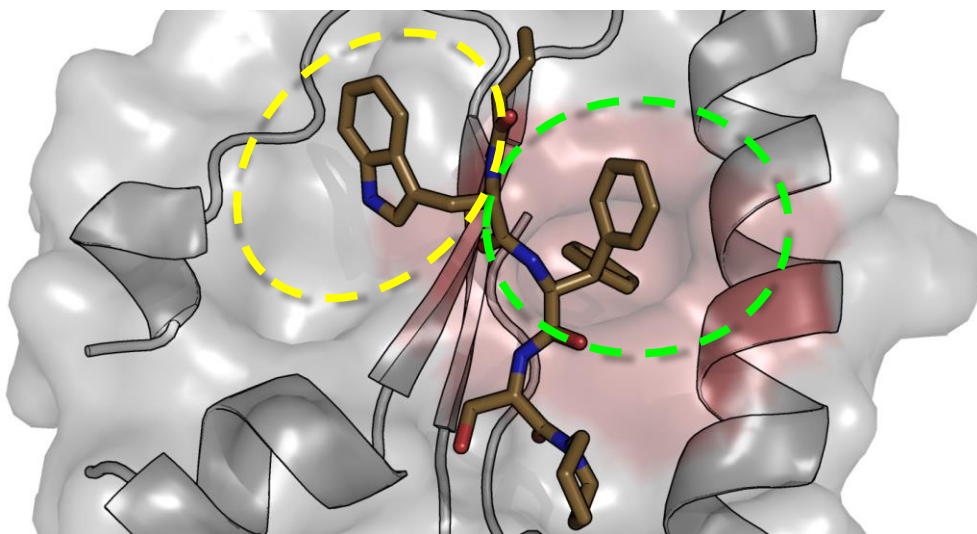


Figure 25. Docking result of IP-15-04 bind to RAD51 (PDB 4B3B). The lateral side-chain of the threonine and diphenylalanine were selected for their interaction in two surface pockets of RAD51 for further optimization.

The same IPROTech Protocol was applied over a library of peptidomimetics derived from the IP-15-04. This library was targeted against the same binding site of the parent compound, with a distance margin of 5 Å from the center of the protein surface pockets (Figure 27). From this second campaign, three lineal tripeptidomimetics with slight modifications of their lateral side-chain groups were selected. These optimized molecules were obtained after applying the same filters, which had been previously applied for the first batch. Moreover one peptidomimetic more was added to the pull in order to investigate and validate the docking model (IPR-465). IPR-465 contained the exact same structure as the parent compound. However, the design included *N*-alkylation of the nitrogen group of the tryptophan, as this nitrogen was predicted by the docking model as a key interaction for the recognition of RAD51.

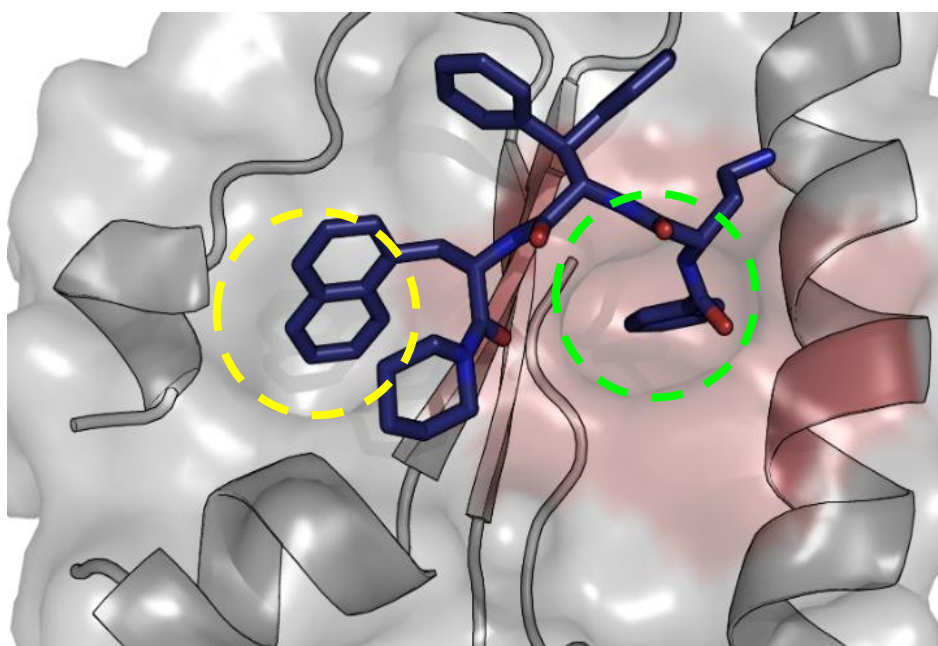
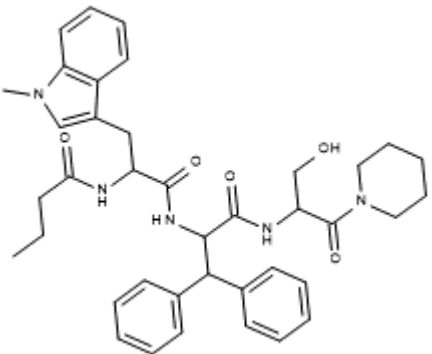
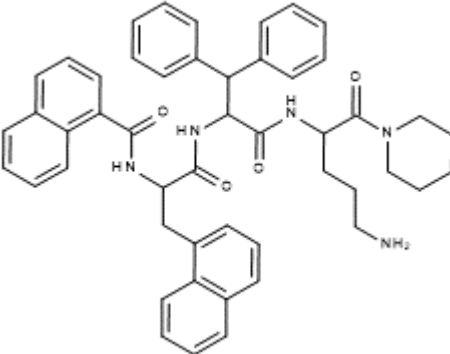


Figure 27. 1-Naphthylalanine lateral side-chain (yellow circle) and benzyl from the C-terminal capping (green circle) of IPR-469 perfectly fit in the two pockets in the RAD51 surface. All compounds of the second generation were designed following this binding mode, where two aromatic groups were pointing to this two pockets.

Iproteos Code	Sequence	Formula	Structure	Docking score (Kcal/mol)	IPRO Permeability (Å <sup>2</sup> )
IPR-465	Butyryl- <i>N</i> -Methyl-tryptophan-Diphenylalanine-Serine-Piperidide	C <sub>39</sub> H <sub>47</sub> N <sub>5</sub> O <sub>5</sub>	 <p>The structure shows a peptide backbone with five residues: butyryl, N-methyltryptophan, diphenylalanine, serine, and piperidide. The butyryl group is attached to the N-methyltryptophan residue. The diphenylalanine residue has two phenyl rings attached to its alpha-carbon. The serine residue has a hydroxyl group on its beta-carbon. The piperidide residue is a six-membered ring with a nitrogen atom and a carbonyl group.</p>	-8,6	123
IPR-467	1-Naphthyl-1-Napthylalanine-Diphenylalanine-Ornithine-Piperidide	C <sub>49</sub> H <sub>51</sub> N <sub>5</sub> O <sub>4</sub>	 <p>The structure shows a peptide backbone with five residues: 1-naphthyl-1-naphthylalanine, diphenylalanine, ornithine, and piperidide. The 1-naphthyl-1-naphthylalanine residue has two naphthyl groups attached to its alpha-carbon. The diphenylalanine residue has two phenyl rings attached to its alpha-carbon. The ornithine residue has a primary amine group on its side chain. The piperidide residue is a six-membered ring with a nitrogen atom and a carbonyl group.</p>	-10,1	148

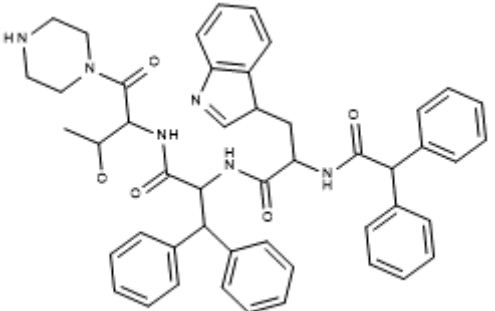
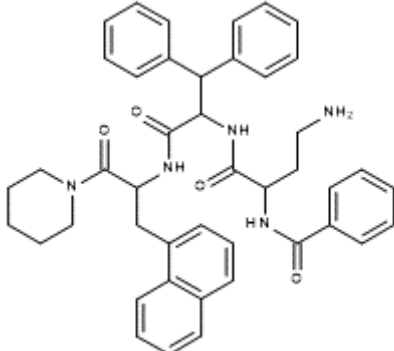
IPR-468	Diphenyl- Tryptophan- Diphenylalanine- Threonine-Piperazine	$C_{48}H_{56}N_6O_5$		-8,1	169
IPR-469	Benzyloxyl-2,4- diaminobutyric acid- Diphenylalanine-1- Naphthylalanine- Piperidine	$C_{44}H_{47}N_5O_4$		-9,6	157

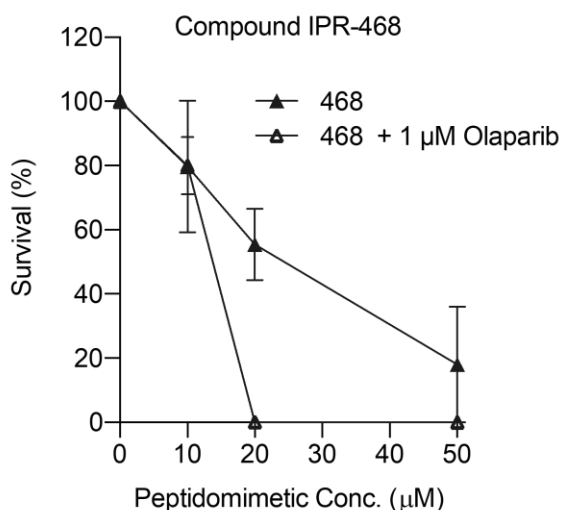
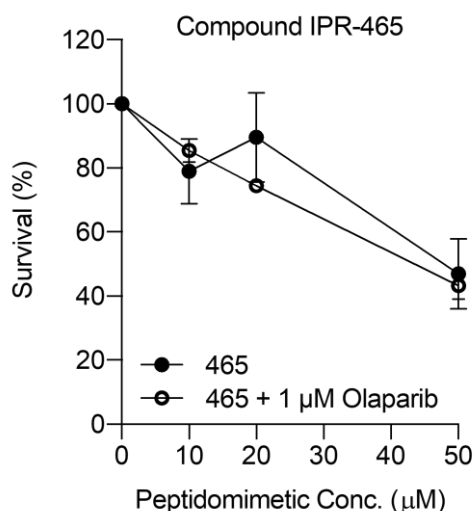
Table 3. Three new peptidomimetics that respected the binding mode of the parent compound, IP-15-04, were selected after visual inspection. The docking score were similar to parent peptide (-9.2 kcal/mol), but the high stability of the docking poses after run short-MD was determinant to their selection for *in vitro* validation. Furthermore, IPR-465 was added as a negative control in order to validate the docking model. The peptidomimetic only has modified a *N*-methylation in the nitrogen from the tryptophan later side-chain. It was expected a loss of potency for this compound.

All compounds contained the piperidide capping in the C-terminal except IPR-468 which had a piperazine group instead the piperidide motif. This means a second nitrogen which can play as a hydrogen bond donor. On the N-terminal position, only IPR-465 preserved the butyryl capping. For the other compounds butyryl moiety was substituted by one benzoic acid (IPR-469), naphthoic acid (IPR-467), or 2,2-diphenylacetic acid (IPR-468) to increase the hydrophobicity of the compound.

Interestingly, all docking models pointed the nitrogen of the tryptophan latera side-chain group straight to the surface pocket. To evaluate the contribution of this nitrogen in the binding, IPR-465 decorated with an N-methylation in this position, expecting a loss of activity when evaluated experimentally.

The non-natural amino acid, diphenylalanine was maintained in all sequence, whereas the serine was replaced for other amino acids containing a lateral-chain proton donor group; ornithine for IPR-467, threonine for IPR-468 and diaminobutyric acid for IPR-469.

After compound synthesis, purification and quantification, these were sent to Sartori's lab for the *in vitro* efficacy evaluation. The protocol used for the evaluation of the second set of compounds was the same one applied for the first batch.



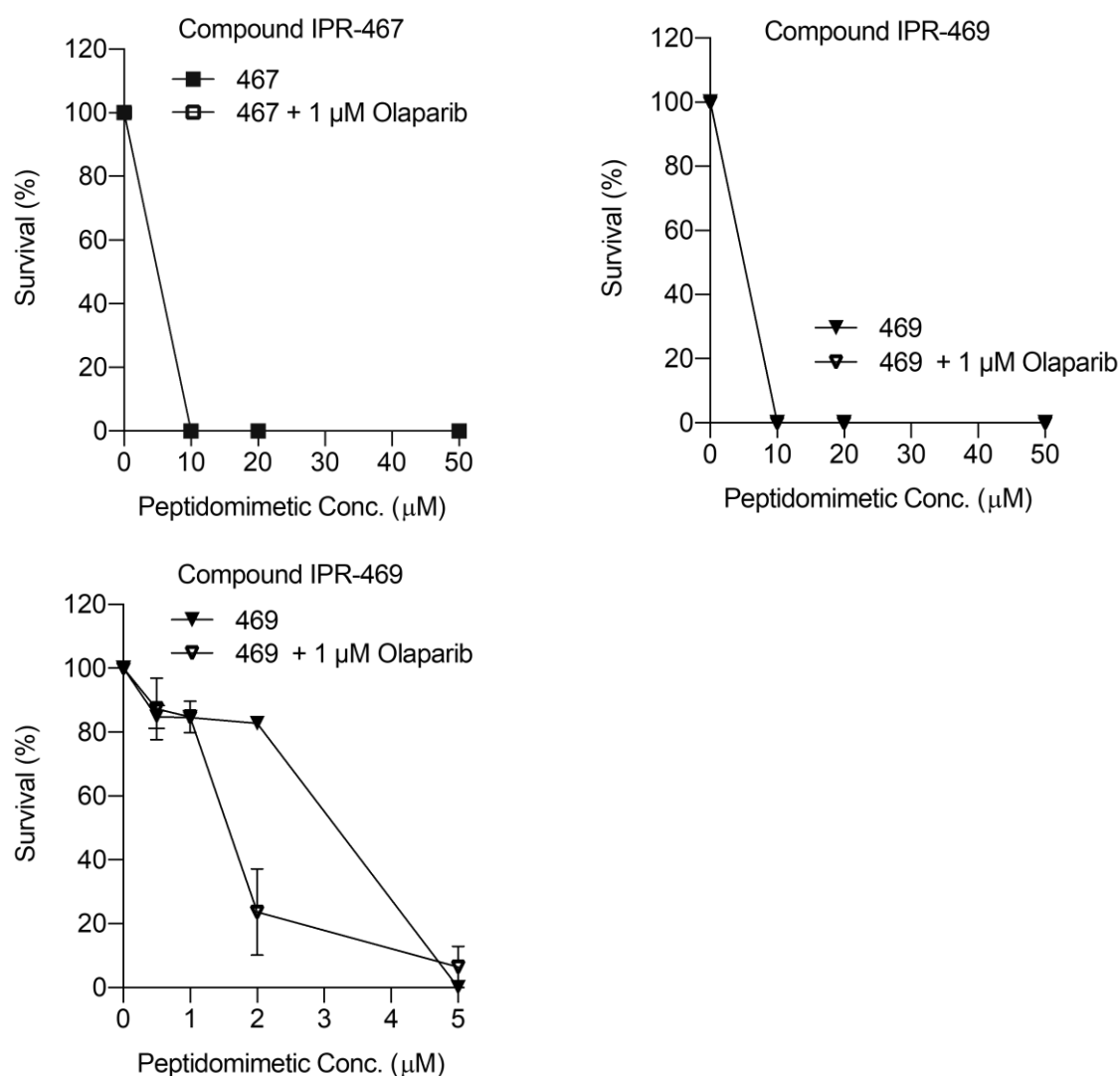


Figure 26. Second generation of compounds inspired in the IP-15-04 structure evaluated in the same conditions as the parent peptidomimetic in a clonogenic proliferation assay with HeLa cells for 10 days at 0.25 % DMSO and 10 % FCS at 10, 20 and 50 µM. As it was expected IPR-465 showed a loss of activity when compared with IP-15-04. IPR-468 slightly improved the performance of the original compound, and IPR-467 and IPR-469 killed the cells at concentration > 10 µM with or without Olaparib. For IPR-469 it was confirmed that at 1-5 µM concentration range there was a synergistic effect with Olaparib.

After the *in vitro* evaluation of the second set of molecules, results showed a highly potent peptidomimetic, IPR-469 that when co-incubated with Olaparib reduced the proliferation of cancer cells with an IC<sub>50</sub> around 1,5 µM. IPR-468 also showed some activity, with a IC<sub>50</sub> between 10-20 µM, IPR-467 and IPR-469 killed tumor cells at concentration range of 1-10 µM. For these two compounds, in order to optimize recourse, the efficacy range between 0-10 µM was evaluated in depth for IPR-469. It could be expected that IPR-467 could show the same synergistic effect at narrow concentration window as IPR-469. Both compounds have a very similar structure, since were designed based on the same parent compound.

IPR-465 showed a loss of potency if compared with its parent peptide, IP-15-04, therefore validating the docking model.

Due the outstanding synergic effect of IPR-469 with Olaparib additional experiments were decided to be conducted to evaluate the feasibility to use this compound in animal models. In this regard, the stability of selected compound in rat plasma was studied (Figure 28).

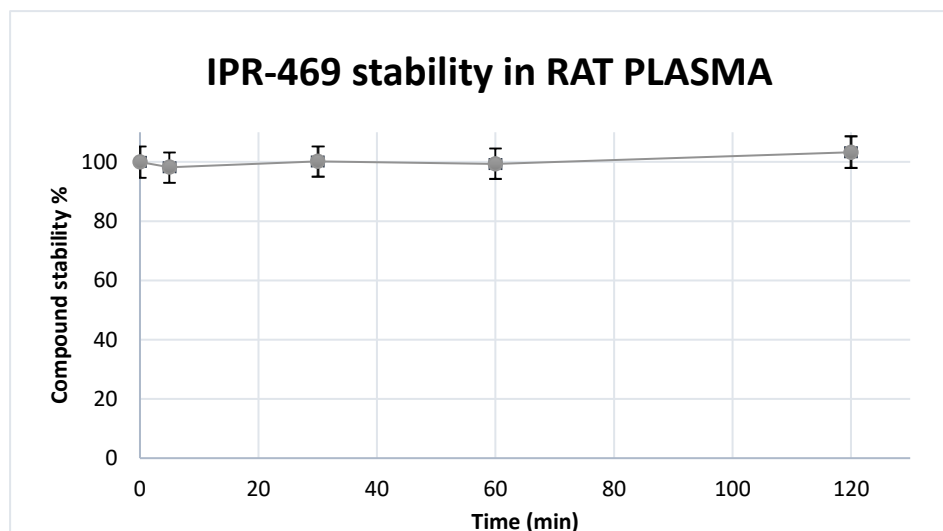


Figure 28. IPR-469 was incubated in rat plasma for 2 h, the estimated maximum time of enzymatic activity. Sample point were taken at different times (x axis in min), and after applied the protocol described in the Materials and Methods chapter, were injected in HPLC to calculate the percentage of degradation (y axis). It was concluded that after 2h the peptide was completely stable. Error bars represent  $\pm$  SD (n=2).

#### 4.5. Conclusions

The use of published the crystallographic structure of RAD51 (PDB: 4B3B) bind to a short-peptide (Ac-FHTA-NH<sub>2</sub>), has facilitated to build a solid docking model to run the IPROTech. This process allowed us to obtain five structures that were selected for synthesis and *in vitro* evaluation. After evaluation, three peptidomimetics were found to be active. For this evaluation compound 2015/IP-15-04 was selected for further optimization in a second round of the technology.

After the second cycle of the technology four compounds were selected for synthesis and evaluation. From these, IPR-469 showed a synergistic effect with Olaparib with an IC<sub>50</sub> < 2  $\mu$ M. This compound was also found to have a high stability in rat plasma, has been obtained. Moreover, the parent peptidomimetic (IP-15-04) also show a promising efficacy with an IC<sub>50</sub> < 30  $\mu$ M. Indeed, future experiments could let into another positive candidate, as a narrow window of concentrations will be studied for IPR-467.

These peptidomimetics outperform all previous published molecules and therefore can represent an interesting project for those pharmaceutical companies with PARPi drugs in their pipeline.

Future experiments will be conducted to validate the mechanism of action validation of the generated molecules.

*Chapter 5: Targeting the protein downstream of K-RAS*





## 5.1. Introduction: KRAS-Effectors

RAS genes are an oncogene family (HRAS, NRAS and KRAS) related with cell proliferation processes which have been found to be highly mutated in human cancer. Indeed, RAS genes were the first discovered oncogenes, in 1982.<sup>130</sup> From the three RAS isoforms, KRAS is the most frequently mutated in RAS-driven cancers, 86 % times, in front of N-RAS 11 % and HRAS 3 %. Interestingly, RAS mutated genes are present in the most deadly cancer types, 90 % of pancreas cancer, 45 % colon cancer and 25 % lung cancer.<sup>131</sup>

However despite the prevalence of RAS mutation in the most deadly types of cancer and after over 30 years of intensively research, still no RAS inhibitor has gained the approval of drug regulatory authorities.<sup>132</sup>

RAS plays a central role in the switch on/off of cell growth process. Catalyzed by guanine nucleotide exchange factors (GEFs), Guanosine 5'-Triphosphate (GTP) bounds to RAS at picomolar range, which activates the protein downstream signals by binding GTP-K-RAS to its effectors (RAS-binding proteins, mainly RAF, PI3K and RAL). When RAS is mutated, the inactivation process whereby GTPase-activating proteins (GAPs) promote GTP hydrolyzation to form the RAS-GDP complex, is inhibit.<sup>133</sup>

Strategies that attempt the direct inhibition of the GTP binding to RAS have been proved an unsuccessful approach, due the picomolar potency of GTP to recognize RAS. Only few low-micromolar compounds have been reported until date,<sup>134,135</sup> therefore discouraging further studies that following this strategy.

Another unsuccessful approach was to prevent the immobilization of RAS to the cell membrane which is mandatory for RAS signaling activation.<sup>136</sup> RAS has a C-terminal CAAX farnesyl moiety. Then, farnelysation of the C-terminal allows RAS to be recognized by farnesyltransferases and then traffic through the cell cytosol to cell membrane. Farnesyltransferase inhibitors (FTIs) have been published<sup>137,138</sup> to be able of block the localization of RAS in the lipid bilayer. Unfortunately, from the three isoforms of RAS only HRAS driven-cancers have shown anti-cancer activity when treated with FTIs in clinical trials<sup>139,140</sup>. This fact could be explained as K-RAS and N-RAS are also substrates of other enzymes that can play as compensatory pathway and then apply post-translational modifications in the protein C-terminals as well.<sup>141</sup> Hence, although targeting other trafficking enzymatic activities of RAS could be of therapeutic relevance, so far no successful advances have been obtained on this line.<sup>142</sup>

In this scenario, the most promising approach is target the RAS signaling protein cascades, which are initiated after the binding of GTP activated RAS with its effectors. Respectively, there are three major RAS-Effector proteins that trigger different signal downstream, RAF, PI3K and RAL.<sup>143</sup> RAF activates the RAF-MEK-ERK cascade, therefore targeting these proteins has led to the development of new approved drugs for specific cancers. For example, Vemurafenib<sup>144</sup> is a B-RAF enzyme inhibitor approved for the treatment of V600E melanoma. Cobimetinib<sup>145</sup> is a MEK1 inhibitor that stabilizes the RAF-MEK complex and has gained FDA approval for combinatory treatment with

Vemurafenib in B-RAF-mutated melanoma resistance cancer. Ulixertinib<sup>146</sup> is a first-in-class ATP-competitive ERK1/2 inhibitor for the treatment of solid tumors in in patients with NRAS-, BRAF V600-, and non-V600 BRAF-mutant cancers.

Although for different protein nodes, promising inhibitors are reaching from bench to patients, resistance to direct inhibition of downstream proteins of the K-RAS-driven cancers as a factor of upregulation or compensatory effects of other signaling pathways could limit its future therapeutic application.

In conclusion, unsuccessful approaches to target the GTP binding site of RAS or modulation of its bioactive function through inhibition of RAS traffic across the cell cytosol and the limits of targeting downstream proteins has move the focus in RAS research to design inhibitors that bind straightly to RAS-effectors binding site of RAS surface protein.

However, RAS surface lacks well-defined pockets and few binders have been described so far.<sup>147</sup>

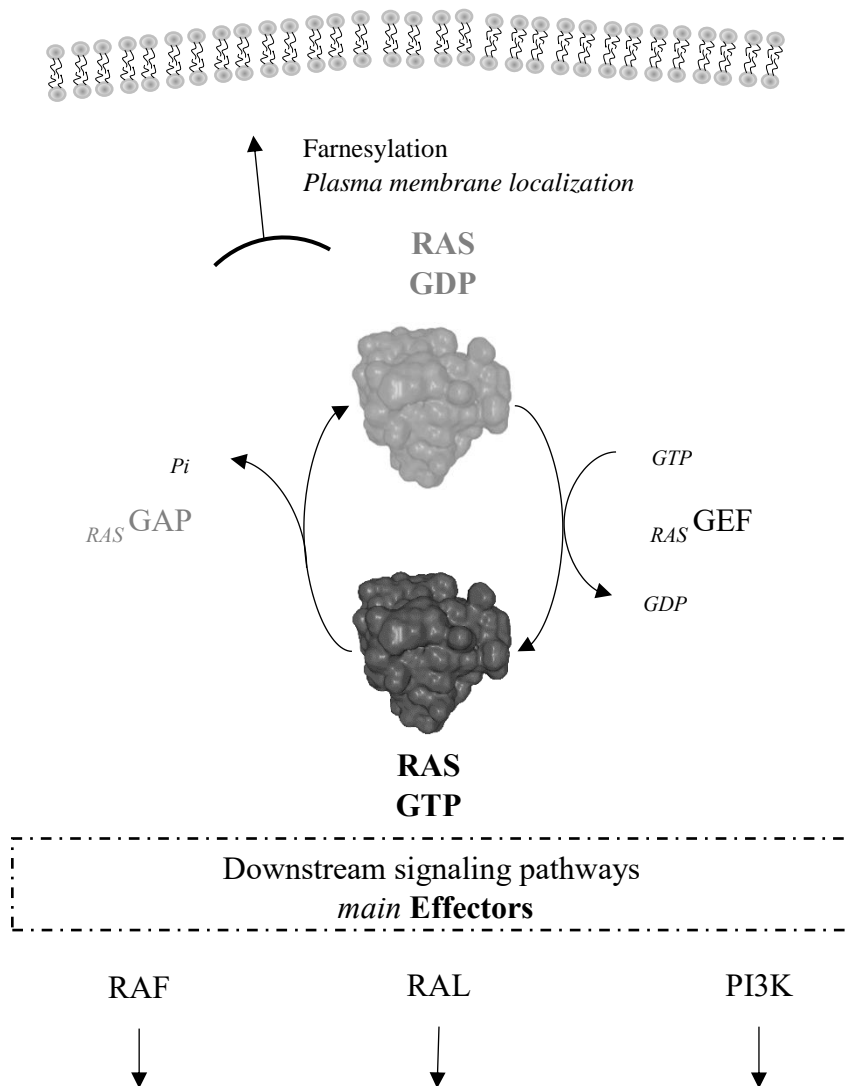


Figure 29. Activation of RAS into a RAS GTP-bound conformation by guanine nucleotide exchange factors (GEFs) proteins state, allows RAS to associate with its protein effectors and start the downstream protein cascade. Moreover, activated RAS can be recruited by the Farnesyl transferase enzymes (FTs) and traffic through the cytosol until reach the cell membrane.

Iproteos's team selected RAS-effector PPI as very challenging and relevant therapeutic target. A preliminary evaluation of the effectors binding site in the surface of RAS shed the possibility to modulate the interaction with the use of peptidomimetics. As a partner of the project, the group Signaling and checkpoints of cell cycle group from the department of Biomedical Sciences from the University of Barcelona led by the Dr. Neus Agell with a high expertise in the RAS field, was in charge of the evaluation of the biological activity of the designed and synthesized compounds.

## 5.2. IPROTech hit identification

IPROTech was applied to afford potent and permeable peptidomimetics as drug candidates. Our approach was to design peptidomimetics that will bind to Ras-effector

binding site, blocking the possibility to interact with effectors proteins, thus, reducing its oncological activity.

The identification of RAS pharmacophoric sites was performed by the structural comparison of the GTPase-RAS in complex with several effector proteins (such as phosphoinositide 3-kinase (PI3K), Bry2RBD, RalGDS, Phospholipase C, NORE1A and RAF) and applying the IPRO Filter standard protocol.

A total of one aromatic and three negatively charged residues (*i.e.* Asp33, Glu37, Asp38 and Tyr64) were identified on the RAS interface. Interestingly, these residues are conserved in the interactions with virtually all the effector proteins. Interestingly, it can be found a high positively charged propensity between Asp33 and Asp38 (Figure 30, Table 4) which favors the interaction with a possible binder.

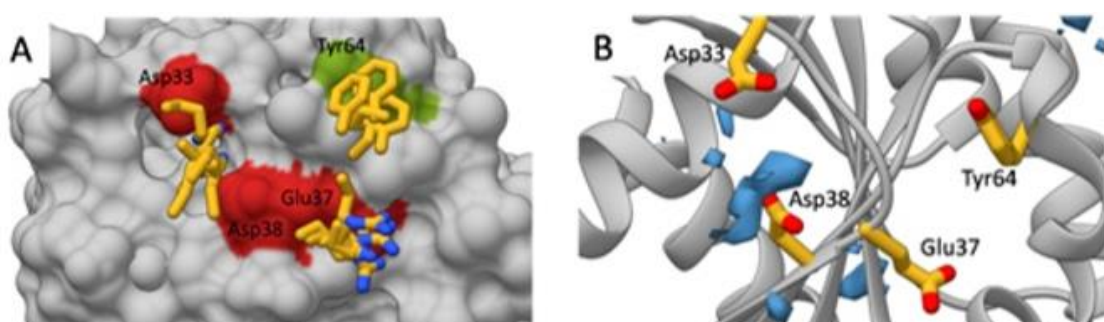


Figure 30. GTPase-RAS protein pharmacophoric sites (PDB: 5P21): (A) GTPase-RAS in complex with IPROfilter probes. Highly conserved residues among RAS effectors proteins are underlined in yellow sticks. RAS protein surface colored in gray, residues involved in intermolecular contacts with highly conserved residues among effector proteins are underlined in red (Asp33, Glu37, Asp38) and green (Tyr64). (B) IPROTech IPROfilter prediction of more relevant residues for the design of a new set of peptidomimetics.

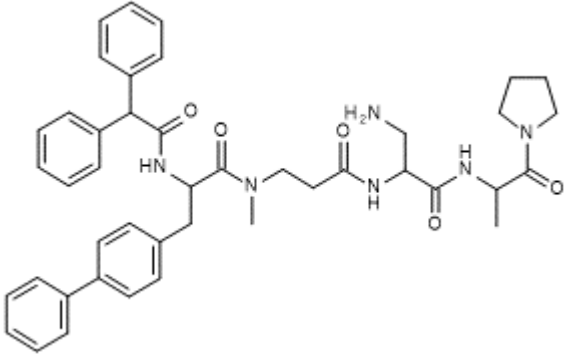
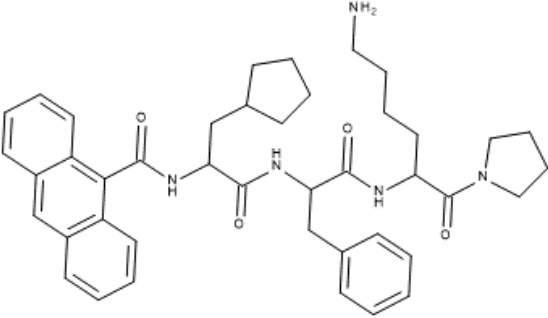
Afterwards, an initial library of more than 80,000 tri- and tetra-peptides, formed by both natural and non-natural amino acids, and containing at least one positive charged residue per sequence, was generated. The peptidomimetic screening was performed by SMINA docking program, using the K-RAS GTPase crystal structure as receptor (PDB 5P21) and setting the binding site around the negatively charged hotspots residues.

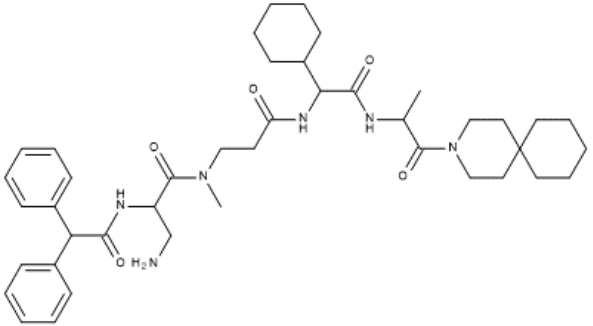
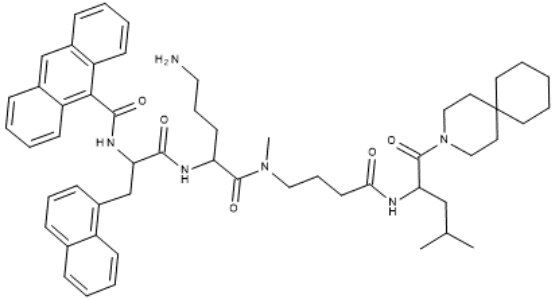
An extended computational survey of the RAS interface resulted in a set of peptidomimetics able to bind to the target binding site with a theoretical potency larger than  $-8.0$  kcal/mol ( $\mu$ M-nM expected experimental inhibitory potency range scale) while exhibiting good *in silico* permeability profile, whose averaged Polar Accessible Surface Area (PASA) is lower than  $150 \text{ \AA}^2$ . Those parameters are representative of the potential activity and permeability of the designed peptidomimetics.

To reject potential docking false negatives, short implicit peptide-binding site MD simulations was conducted to identify peptidomimetics with stable binding mode (RMSD along simulation smaller than  $3 \text{ \AA}$ ). The idea was to enrich the selection with compounds able to bind, from a thermodynamic point of view, but preserving the binding target area, *in silico* properties and peptides structure.

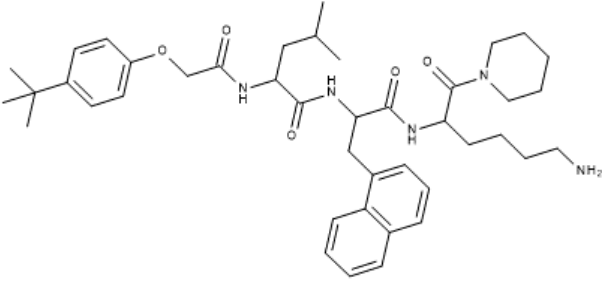
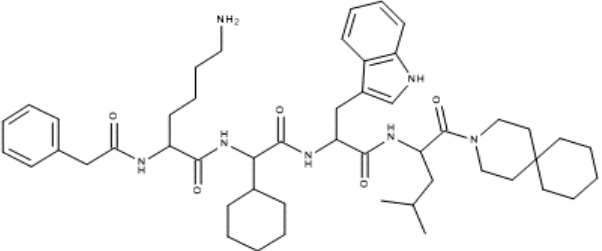
After a final visual inspection, 9 sequences were decided to be synthesized, quantified and by following the protocol described in the Materials and Methods chapter of this thesis. To evaluate the synthesized compounds, cell based experiments were carried out by Debora Cabot a PhD student at the Dr. Neus Agell group at that time.

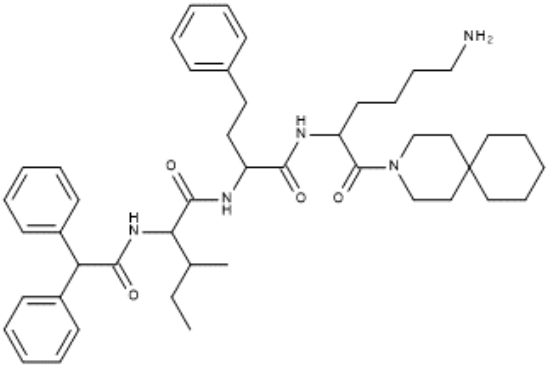
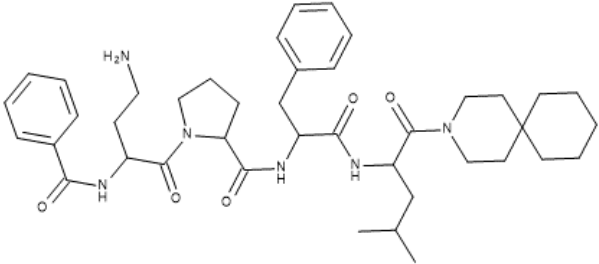
For the most potent candidates, if they are finally proved to be specifically targeting those cells that dependent of K-RAS, further experiments will carried out in order to investigate the possible compounds application for the treatment of K-RAS dependent cancers, and therefore intellectual property of the compound will be seek.

Iproteos Code	Formula	Structure	Docking score (Kcal/mol)	IPRO Permeability (Å <sup>2</sup> )
IP-14-01	C <sub>43</sub> H <sub>51</sub> N <sub>6</sub> O <sub>5</sub>	 <p>The structure of IP-14-01 is a complex molecule featuring a central chain of amide bonds. From left to right, it includes: a benzamide group with a phenyl substituent on the nitrogen; a methylene group; a secondary amide with a methyl group on the nitrogen; a propyl chain; a primary amide with a methyl group on the nitrogen; a secondary amide with a methyl group on the nitrogen; and finally, a pyrrolidine ring attached to the chain via its nitrogen atom.</p>	-10,2	128
IP-14-02	C <sub>42</sub> H <sub>52</sub> N <sub>5</sub> O <sub>4</sub>	 <p>The structure of IP-14-02 is a complex molecule featuring a central chain of amide bonds. From left to right, it includes: a benzamide group with a naphthalene substituent on the nitrogen; a methylene group; a secondary amide with a pyrrolidine ring on the nitrogen; a methylene group; a secondary amide with a methyl group on the nitrogen; a propyl chain with a primary amine group (NH<sub>2</sub>) at the end; a secondary amide with a methyl group on the nitrogen; and finally, a pyrrolidine ring attached to the chain via its nitrogen atom.</p>	-9,7	93

IP-14-03	$C_{42}H_{61}N_6O_5$		-9,6	105
IP-14-04	$C_{54}H_{69}N_6O_5$		-9,1	123



IP-14-05	C <sub>42</sub> H <sub>60</sub> N <sub>5</sub> O <sub>5</sub>		-8,9	111
IP-14-06	C <sub>49</sub> H <sub>72</sub> N <sub>7</sub> O <sub>5</sub>		-8,4	116

IP-14-07	$C_{46}H_{64}N_5O_4$	 <p>The chemical structure of IP-14-07 is a complex molecule featuring a central chain of amide bonds. From left to right, it includes: a biphenyl group attached to a carbon atom; a secondary amide group; a carbon atom bonded to a methyl group and an ethyl group; another secondary amide group; a carbon atom bonded to a benzyl group and a hydrogen atom; a third secondary amide group; a carbon atom bonded to a hydrogen atom and a decyl chain ending in a primary amine group; and finally, a decyl chain attached to a nitrogen atom within a bicyclic piperidine-decane system.</p>	-8,4	86
IP-14-08	$C_{41}H_{59}N_6O_5$	 <p>The chemical structure of IP-14-08 is a complex molecule with a central chain of amide bonds. From left to right, it includes: a benzamide group; a carbon atom bonded to a primary amine group and a hydrogen atom; a secondary amide group; a carbon atom bonded to a hydrogen atom and a pyrrolidine ring; another secondary amide group; a carbon atom bonded to a hydrogen atom and a benzyl group; a third secondary amide group; a carbon atom bonded to a hydrogen atom and a decyl chain ending in a primary amine group; and finally, a decyl chain attached to a nitrogen atom within a bicyclic piperidine-decane system.</p>	-8,3	133

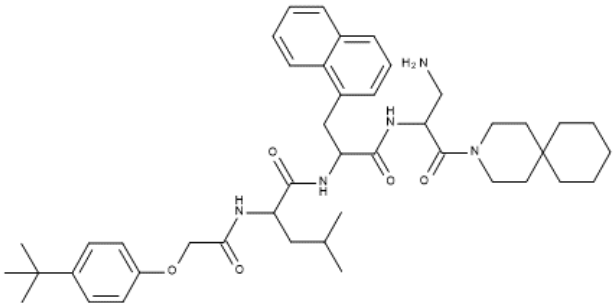
IP-14-09	C <sub>44</sub> H <sub>62</sub> N <sub>5</sub> O <sub>5</sub>		-8,0	104
----------	---	--	------	-----

Table 4. After a final visual inspection, 9 peptidomimetic sequences were selected for synthesis. The batch includes tri- and tetra-peptidomimetics. All sequences have a secondary amine capping at the C-terminal and a hydrophobic group as N-terminal capping. Based in the benchmarked data, IPRO Permeability threshold was set at 170 Å and therefore for all peptidomimetics below this threshold were predicted as permeable.

### 5.3. Experimental evaluation

RAS-GTP signaling protein was activated and the inhibitory effect of the synthesized peptidomimetics was evaluated in retinal pigment epithelium (RPE) cells. RPE cells were seed in culture plate for 48 h. After that period, a first treatment of 10 min with Epidermal Growth Factor (EGF), 50 ng/mL, was performed to active the RAS signaling pathways.<sup>148</sup> Next, GTPase activating protein (GAP120), a positive control,<sup>149</sup> and peptidomimetics (50  $\mu$ M) were incubated for 2 h with cells. Finally, Western Blotting (WB) detection of effector proteins was obtained through immunoprecipitation (IP).

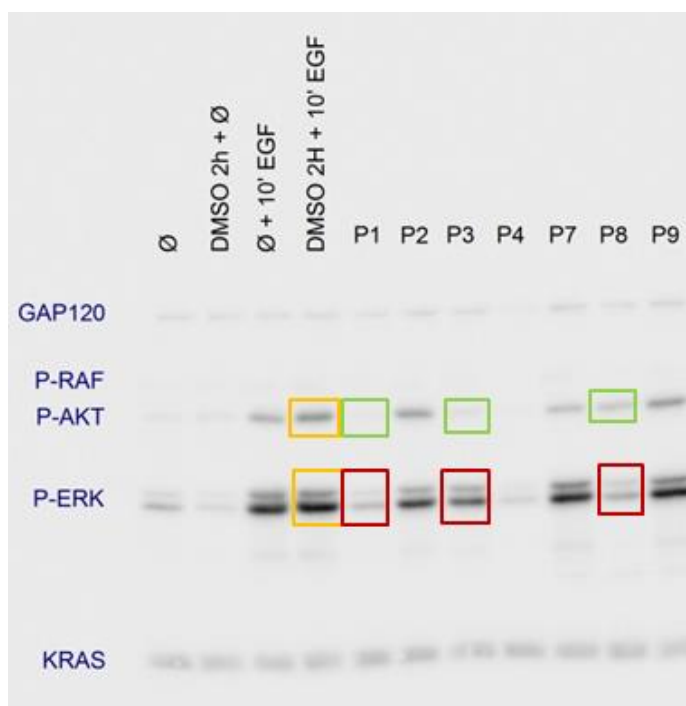


Figure 31. RAS Western Blot of the evaluation of the synthesized compounds IP-14-01 (P1), IP-14-02 (P2) IP-14-03 (P3), IP-14-04 (P4), IP-14-07 (P7), IP-14-08 (P8) and IP-14-09 (P9). GTPase activating protein (GAP120) is an inhibitor of RAS activity. DMSO was the solvating agent to dilute the peptides and then after dilution with cell medium the final percentage was 0.5 %.

Positive compounds should inhibit signaling RAS pathways that are activated when RAS binds with its effectors. As has been mentioned there are three major pathways, and their inhibition was directly, RAF phosphorylation, or indirectly, AKT and ERK phosphorylation, evaluated by Western Blot. RAF protein binds to RAS, AKT is in the PI3K pathway and ERK in the RAL pathway. Therefore, to consider a peptidomimetic as a positive hit, it should inhibit the activity of the three protein cascades, and as a consequence the phosphorylation of the RAF (P-RAF), AKT (P-AKT) and ERK (P-ERK) not be observed.

Unfortunately, compounds IP-14-05 and IP-14-06 were not soluble when diluted in cell medium and it was found that peptide IP-14-04 formed aggregates and precipitated (Figure 31) over the cells causing cell death, therefore discarding its read out as well.

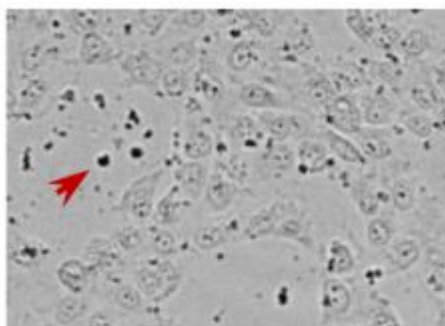


Figure 32. Once IP-14-04 was diluted in cell culture medium with RPE cells, formed aggregates (black dots pointed by the red arrow), which end up causing cell death.

Hence, from the initial batch, only 6 compounds were properly evaluated in cells and 4 peptidomimetics were discarded as a consequence of their low solubility. From these 6 evaluated molecules, IP-14-01, IP-14-03 and IP-14-08 were able to inhibit three RAS-effector protein cascades. Being IP-14-01 the most potent of them when the WB was analyzed (figure 31).

Once the most promising candidates were selected, it was evaluated their biophysical properties by studying solubility in 5 % DMSO in water, permeability through biological barriers (PAMPA assay) and internalization in cells.

Iproteos Code	Solubility (mM)	Pe (cm/s)	Transport %	Retention %	Cell internalization %
IP-14-01	1,333 ± 0,013	5,6E-10 ± 3,1E-10	0	55,7 ± 7,1	13,5 ±
IP-14-03	0,549 ± 0,090	1,6E-09 ± 3,6E-10	0	77,9 ± 2,4	0
IP-14-08	1,317 ± 0,022	0	0	97,6 ± 5,0	0

Table 5. PAMPA assay (Pe, Transport % and Retention %) along with cell internalization % were to evaluate peptidomimetic permeability. All experiments were carried out as is described in Materials and Methods chapter. Solubility in H<sub>2</sub>O at 5 % DMSO was performed as a control to discard any possible issue related with a low solubility. Data are expressed as the mean ± SD.

Solubility was measured in H<sub>2</sub>O at 5 % DMSO and the three compounds showed a good solubility that assured that experiments at > 100 µM concentrations could be done without the risk of compound precipitation. The PAMPA assay showed a high retention %, negligible transport and null permeability *Pe* for the three peptidomimetics. Oppositely, when IP-14-01 was incubated with SH-SY5Y cells for 2 h at 60 µM, 13,5 % of the peptideomimetic quantity was uptake by cells. However, this result cannot be reproduced when IP-14-03 and IP-14-08 were incubated with cells.

Based on *in vitro* results, we assumed that peptidomimetics were retained in the lipid membrane of the PAMPA assay as a consequence of high and bulky hydrophobic groups. Therefore, being their permeability capacity underestimated by PAMPA assay.

Due to the inhibitory effect on RAS-effector signaling cascades of IP-14-01 along its high cell internalization value, compound IP-14-01 was further optimized through a second round of the IPROTech.

#### 5.4. Hit optimization

A second generation of peptidomimetics was designed following the same *in silico* protocol applied previously but basing the new compounds design in IP-14-01 structure. In this regard, 4 new peptidomimetics were eluted once the computational studies were complete.

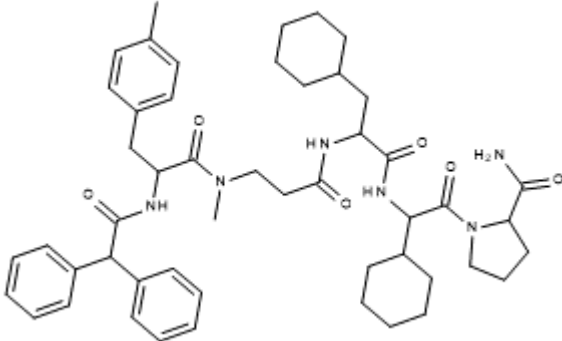
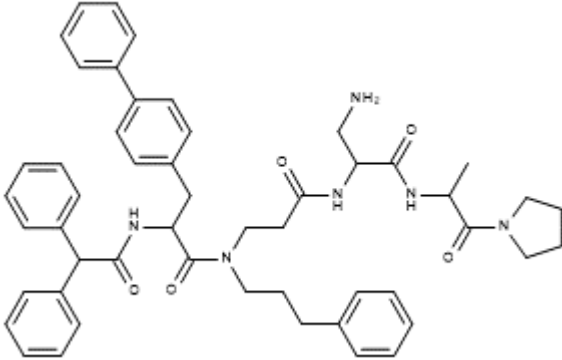
IPR-471 was intentionally designed to increase the number of amino acids on the peptidomimetic backbone. On this basis, the C-terminal was extended by removing the secondary amine that had been used as capping for a proline plus a carboxamide.

IPR-472 had almost the same structure than IP-14-01 but the *N*-methyl alkylation of the  $\beta$ -Alanine was substituted by longer carbon chain attached to an aromatic group, propylbenzene. The idea behind this was to increase the overall compound hydrophobicity along with an increase of the n-alkyl shielding capacity. A 4-carbon chain, instead a methyl, would present a certain degree of flexibility allowing the 6-carbon aromatic ring wrap the molecule, thus reducing its polarity in an aqueous environment. Moreover, it could be expected that the number of contacts with the protein surface would be favored as well.

IPR-473 was the most conservative proposal as only an alanine was substituted by an isoleucine and a *N*-methylation was added in one amide bond of the sequence backbone. This new molecule was expected to completely preserve the binding mode of the parent compound but adding a few more contacts in order to slightly optimize the potency.

IPR-474 had two substitutions, an alanine by a cyclohexylglycine and the substitution of the acid with a polar group for another amino acid with a polar group, a threonine.

Finally, all peptidomimetics preserved the same sequence ending by keeping the three same amino acids following the same order and the diphenyl *N*-terminal (table 6).

Iproteos Code	Formula	Structure	Docking score (Kcal/mol)
IPR-471	$C_{50}H_{66}N_6O_6$	 <p>The chemical structure of IPR-471 is a complex molecule featuring a central chain of amide bonds. On the left, there is a diphenylmethyl group attached to an amide nitrogen. The chain continues through several amide linkages, including a piperidine ring and a cyclohexane ring. On the right, it terminates in a proline ring system with an amino group.</p>	-10,5
IPR-472	$C_{51}H_{59}N_6O_5$	 <p>The chemical structure of IPR-472 is a complex molecule with a central amide chain. It features a diphenylmethyl group on the left, a piperidine ring, and a cyclohexane ring. The chain ends on the right with a proline ring system and an amino group.</p>	-10,2

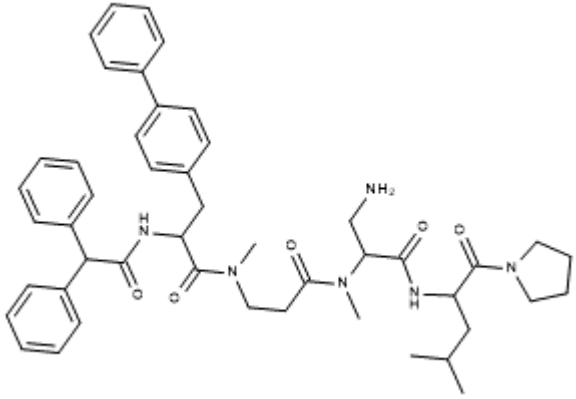
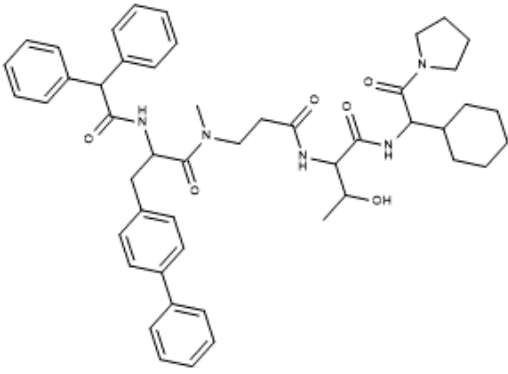
IPR-473	C <sub>47</sub> H <sub>59</sub> N <sub>6</sub> O <sub>5</sub>	 <p>The chemical structure of IPR-473 is a complex molecule featuring a central chain of amide bonds. It includes a diphenylmethyl group, a benzyl group, a dimethylamino group, a primary amine group, a secondary amine group, and a pyrrolidine ring.</p>	-9,6
IPR-474	C <sub>49</sub> H <sub>59</sub> N <sub>5</sub> O <sub>6</sub>	 <p>The chemical structure of IPR-474 is a complex molecule featuring a central chain of amide bonds. It includes a diphenylmethyl group, a benzyl group, a dimethylamino group, a secondary amine group, a hydroxyl group, and a pyrrolidine ring.</p>	-9,6

Table 6. Applying the same docking methodology previously used for the screening of the 1<sup>st</sup> round of peptidomimetics, a second generation was generated, four new molecules, by basing the screening of new molecules in the structure of IP-14-01.



Furthermore, an additional effort was done to reevaluate those peptidomimetics from the first generation that were not soluble on the *in vitro* assay conditions or formed aggregates when incubated with cells. For those compounds, their solubility was studied in PBS with 15 %  $\beta$ -cyclodextrin (table 7).

Iproteos Code	Conc. in PBS $\beta$ -cyclodextrin 15 % ( $\mu$ M)
IP-14-04	780 $\pm$ 30
IP-14-05	849 $\pm$ 43
IP-14-06	844 $\pm$ 4
IP-14-07	766 $\pm$ 19
IP-14-09	949 $\pm$ 8

Table 7. All peptidomimetics were dilute at 1 mM in phosphatase with at 15 %  $\beta$ -cyclodextrin. Then were let under constant agitation for 24 h, centrifuged and finally, the supernatant of each one was injected in the HPLC and compared with a 1mM solution of each compound in ACN/H<sub>2</sub>O. This control that was used to determine the real solubility in PBS at 15 %  $\beta$ -cyclodextrin. Data are expressed as the mean  $\pm$  SD.

Hence, the four IP-14-01 (P1) derivatives (IPR-471; P1.1, IPR-472; P2.2, IPR-473; P3.3 and IPR-474; P4.4) along with peptidomimetics dissolved in PBS with 15 % of  $\beta$ -cyclodextrin (IP-04; P4, IP-14-05; P5, IP-14-06; P6, IP-14-07; P7 and IP-14-09; P9) were evaluated *in vitro*. In order to have a clear read out, the experiments were ran in HA-KRAS-G12V (mutant K-RAS) transfected HeLa cells<sup>148</sup> applying the same conditions used for RPE cells.

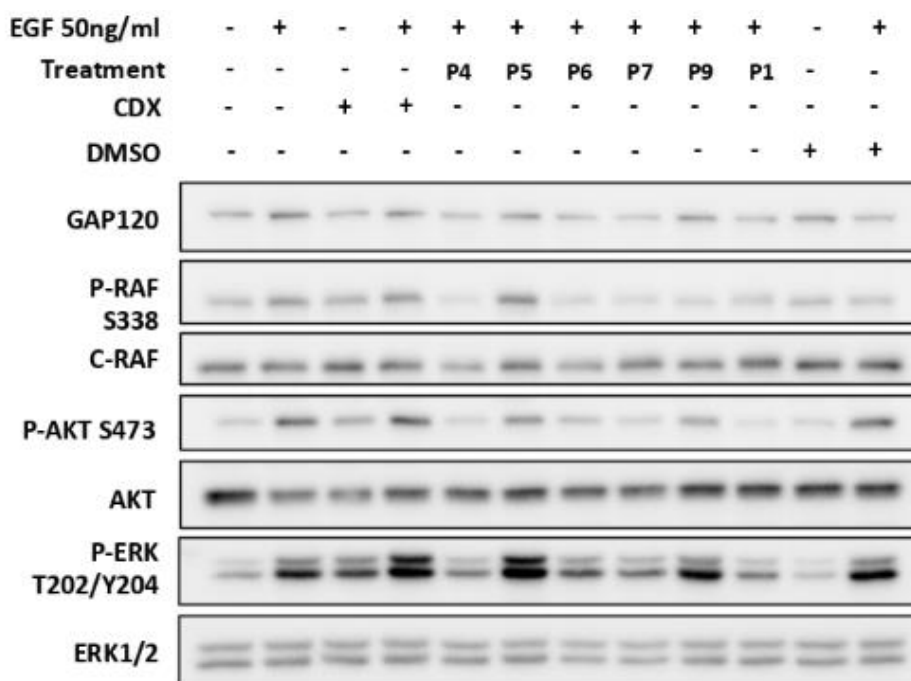


Figure 33. RAS western blot was repeated in the same conditions except for the use of  $\beta$ -cyclodextrin at 0.5 % instead of DMSO used in the evaluation of compounds IP-14-01 (P1), IP-14-02 (P2) IP-14-03 (P3), IP-14-04 (P4), IP-14-07 (P7), IP-14-08 (P8) and IP-14-09 (P9). GTPase activating protein (GAP120) is an inhibitor of RAS activity. DMSO was the solvating agent to dilute the peptides and then after dilution with cell medium the final percentage was 0.5 %.

After the incubation time (2 h), from the first generation of peptidomimetics diluted in  $\beta$ -cyclodextrin, only IP-14-07 (P7) almost resembled the potency of IP-14-01 (P1) (Figure 33). On the other hand, IP-14-04 (P4) and IP-14-06 started killing cells. Therefore, although the use of  $\beta$ -cyclodextrin allowed the proper dissolution of the poorly soluble compounds in the in aqueous media, *i.e.* cell media, any of them outperformed IP-14-01.

The new generation of peptidomimetics based in IP-14-01 structure, was as well evaluated. It was found that IPR-471 (P1.1) and IPR-474 (P1.4) were considered inactive, whereas IP-14-02 (P1.2) was able to inhibit the RAF and AKT but not ERK. Interestingly, IPR-473 (P1.3) was found to be able to inhibit the three different protein cascades more efficiently even than the parent peptidomimetic.

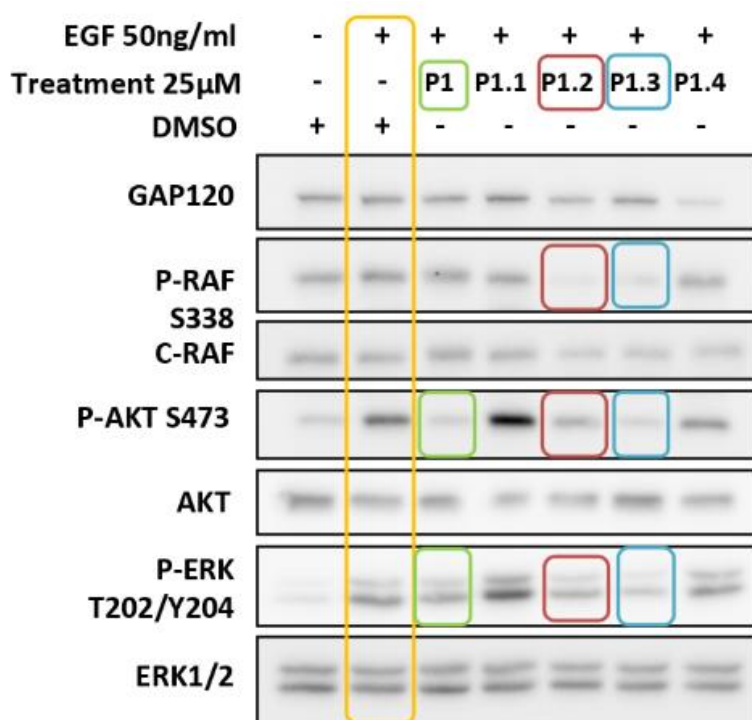


Table 34. RAS Western Blot of the evaluation of 2018/IP-14-01 (P1) and its derived peptidomimetics, IPR-471 (P1.1), IPR-472(P1.2), IPR-473 (P1.3) and IPR-474 (P1.4). For this experiment same protocol conditions of previous WB assay for the first generation of peptidomimetics was applied. Compounds were incubated in RPE cells at 50  $\mu$ M at 0.5 % DMSO for 2h. Not solubility issues were observed.

In this scenario, a potential therapeutic application of IPR-473 would be determined by its capacity to discriminate between normal cells and cancer cell lines, in which its inhibitory activity should produce cell death. The obtained information would be determinant prior to any effort to reach preclinical phases.

Hence, IPR-473 was incubated at several concentrations with cancer cell lines and normal cell line (RPE). Cell viability was measured by a MTS cell proliferation assay<sup>150</sup>. The same experiment was performed for the parent peptidomimetic, but not difference between normal cells and cancer cell lines was observed at any incubated concentrations, *i. e.* the parent compound did not show cell-line specificity (data not shown).

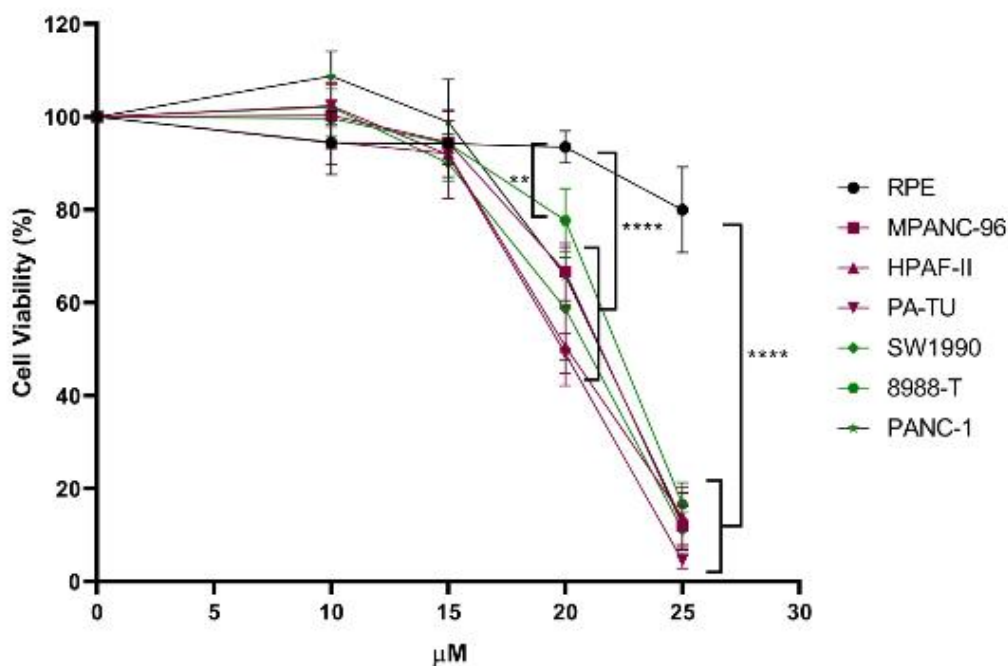


Figure 35. MTS viability assay was the technique used to measure the cell viability of 7 different cell lines. Cells were placed in an adequate volume of 10 % FBS-containing medium to get 200 cells / mL. Next, cells were cultured for 24 hours and then treated with the drugs (10  $\mu$ M, 15  $\mu$ M, 20  $\mu$ M and 25  $\mu$ M) for a further 24 hours incubation.

Human pancreatic cancer cells MPANC-96, Human pancreatic adenocarcinoma cells HPAF-II, Human pancreatic grade II adenocarcinoma PA-TU, Human pancreatic ductal adenocarcinoma SW1990, Human pancreas adenocarcinoma 8988-T and Human pancreatic ductal carcinoma PANC-1 are pancreatic tumor cell lines, meanwhile RPE are retinal pigment epithelium cells and therefore non cancerigenous.

The results obtained in the cell viability assay confirmed the potential therapeutic activity of IPR-473, which has a cytotoxic behavior for cancerigenous cells at concentrations higher than 15  $\mu$ M whereas there is no effect for normal cells.

## 5.5. Conclusions

After a first inspection of the binding site area of RAS with its effectors and the application of the IPRoTech computational tools, a set of 9 peptidomimetic sequences were selected for their predicted capacity to inhibit the oncogenic activity of RAS in its activate form.

From the initial pull, 3 compounds show an inhibitory activity in cells in the three more relevant effector signaling pathways. Not soluble compounds were incubated with cells with the presence of  $\beta$ -cyclodextrin instead of DMSO, which allowed a proper compound solubilization. These conditions allowed also to evaluate IP-14-07 which was able to inhibit the phosphorylation of RAF, AKT and ERK.

IP-14-01 showed the higher potency according the western blot analysis from the first 4 hits found, and therefore was used for hit optimization. Its structures and binding site were used to design 4 new peptidomimetics.

When the second peptidomimetics generation was evaluated in cells only IPR-473 outperformed the parent compound. Interestingly, IPR-473 differs from IP-14-01 in alanine for a leucine and in an extra *N*-methylation in the peptidomimetic backbone. This small difference in the molecule structure could help to improve its permeability through the cell membrane, but the improvement in the potency could be possible effect of a most favorable bioactive adoption due the *N*-methylation of the amide bond and increase of the number of contacts with the receptor protein.

Particularly, this increase of the IPR-473 potency respect from IP-14-01 was translated into a high specificity for pancreatic cancer cell lines in front of normal cells, when the peptidomimetic was evaluated in cell proliferation assay. Thus, opening the opportunity to carry out further experiments to evaluate the possible application of IPR-473 for the treatment of pancreatic cancer. Indeed, at the moment of writing of this thesis, the protection of the IP of IPR-473 is seek, as well a publication in a high impact journal.



*Chapter 6: Inhibition of the Retromer-L2 interaction to prevent Human Papillomavirus endosomal release*



## 6.1. Introduction: Retromer-L2

Human Papillomaviruses (HPVs) is the most common sexually transmitted virus. Although HPVs are associated with a wide range of pathologies, the most frequently health issues related with HPVs are genital warts and cervical cancer. Indeed, cervical cancer is the 4<sup>th</sup> most common type of cancer in women.<sup>151</sup>

Viruses have two main systems of virus content delivery into the host cytoplasm. Enveloped virus cross the lipid bilayer by a glycoprotein cover that fusion with the cell membrane.<sup>152</sup> On the other hand, non-enveloped virus disrupt the lipid bilayer to gain access, normally by means of a peptide or capsid protein that generates a pore into the cell membrane.<sup>153</sup>

However, HPVs have developed their own system to cross the cell membrane by the association of the viral material with the retrograde transport vesicles.<sup>154</sup> Retrograde transport is the mechanism used by cells to secrete proteins from the endoplasmic reticulum to the trans-Golgi network (TGN) for reuse. Within this mechanism, Retromer is an intracellular multi-protein complex that binds to the cytoplasmic domain of cellular transmembrane proteins and whose function is sequester those macromolecules that can be recycled for the cell body<sup>155,156</sup>.

Thereby, HPVs are internalized by cells membranes by endocytosis<sup>157</sup> and once uptakes, these are released from the endosomal envelop by binding to Retromer. Then, HPVs works as Retromer *cargo*. In fact, it was in 2015 when Popa *et al.*<sup>158</sup> demonstrated that the virus was not being released from the endosomal vesicle by itself, if not due the interaction with the Retromer protein. In the endosomal release mechanism, the C-terminal part of L2 protein, which is composed for 32 amino acids, HPV16 L2<sub>441-473</sub>, from the virus capsid bulges and acts as a cell penetrating peptide to escape from the lipid capsule and bind to the Retromer (Figure 36). This C-terminal motif contains a cationic sequence (RKRRKR) that resembles the TAT peptide motif. DiMaio *et al.*<sup>159</sup> demonstrated that permutation of the HPV16L2 cationic motif by TAT do not reduce the HPV transfection capacity whereas elimination of this amino acids from the HPV16 L2 protein is translated in a total deletion of the virus transfection. In the same direction, knockdown of the Retromer protein drastically decreases the HPVs infection rate, then validating the interaction within the virus and the cytosolic protein. This finding opened the door to a novel therapeutic approach for the treatment of the HPV based in the disruption of the interaction between Retromer and L2 protein, which should block the transfection of the HPVs through the cells.



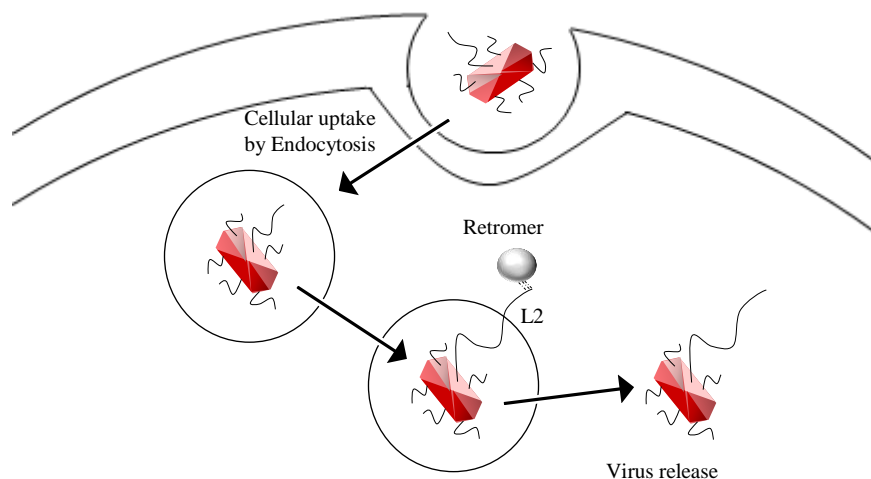


Figure 36. HPV capsid internalization is carried out by endocytosis-mediated mechanism. To escape from the endosomal envelope, the L2 protein which contains a C-terminal motif that plays the role of a cell-penetrating peptide, exits from the lipid capsule and binds to the Retromer protein, which is present in the cytosol. The interaction with the Retromer allows the virus to be released from the endosomal capsule and infect the cell.

In 2016 Dr. Aitor Hierro's group solved the crystallographic structure of the retromer, which allowed the understanding of how the protein is recruited by transmembrane proteins along with the recognition mechanism of the protein to sequester proteins or other macromolecules to be recycled.<sup>160</sup> As a continuation of this work, Hierro's lab crystallized HPV16 L2 motif bound to the retromer protein (unpublished structure).

Hence, guided by the results obtained by DiMaio *et al.* which suggest a possible therapeutic application for the infection of HPV and the crystallized structure of Hierro's group, it was hypothesized a possible design of peptidomimetics that can compete against L2 protein and therefore disrupt the interaction between L2-Retromer.

In this regard, a collaboration agreement with Dr. Aitor Hierro's group, Membrane Trafficking Lab of the CIC bioGUNE institute, was established. Iproteos was in charge of applying its technology platform to design peptidomimetics able to bind to the L2-retromer binding site and disrupt the L2-Retromer interaction in order to prevent HPV infection. The compounds obtained in this project are under an intellectual protection process.

## 6.2. IPROTech hit identification

The crystallographic structure of the retromer protein bound to Ac-DFY<sub>2</sub>LH-NH<sub>2</sub> peptide was used in this project to carry out the computational studies (crystal structure unpublished). This short-peptide was derived from the L2 motif that binds to the retromer protein.

441ADAGDFY<sub>2</sub>LHPSYYMLRKRRKRLPYFFSDVSLAA<sub>473</sub>.

First, and following the same methodology applied for the PPI RAD51-BRCA2 (Chapter 4), the bound peptide was re-docked to assess the feasibility of the IPROTech technology to be applied over the Retromer protein (Figure 37).

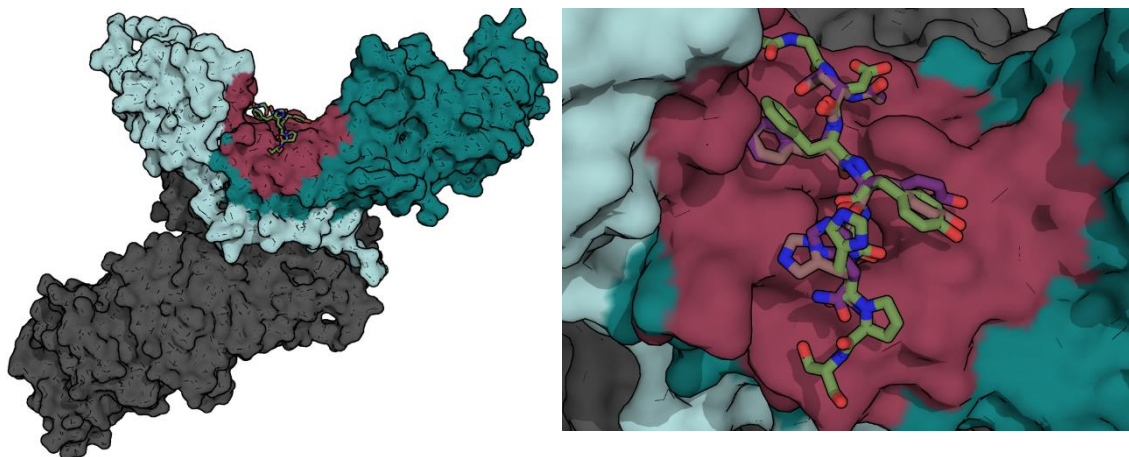


Figure 37. The bound Ac-DFYLH-NH<sub>2</sub> peptide to the Retromer protein (X-ray crystallographic structure unpublished) was re-docked by applying the IPRO Docking technology. Docking was performed at low (brown) and high (purple) exhaustively. The pose of the original peptide (green) was reproduced with a RMSD of 2.0 and 1.9 Å.

IPRO docking reproduced with good accuracy (RMSD = 2.0 Å at low exhaustiveness and docking score = -11.4 kcal/mol) the pose of the natural derived bound peptide (Figure 37). On basis to this preliminary result, a 13000 tetra- and penta-peptidomimetics library was screened against the same binding site of Ac-DFYLH-NH<sub>2</sub>. All compounds contained a butyryl moiety at the *N*-terminal part and piperidide group at the *C*-terminal part. In an initial step, docking step was performed at low exhaustiveness reducing the size of the library to a final pull of 600 molecules. For this set of molecules, a second docking analysis was performed increasing the exhaustiveness of the protocol. From this second docking analysis, top scored 200 structures were selected and evaluated using the IPRO Permeability filter, obtaining only 20 structures predicted to be permeable (poASA < 170). Finally, for this reduced number of molecules the stability of the Retromer-peptidomimetic was evaluated by short-MD simulations. After this analysis, 5 hit candidates were obtained.

From the 5 hit candidates, and taking into account the high success-rates obtained in the previous projects along with the resources limitations at that time, tipped the scale to select only the best top-ranked 2 compounds following the docking score, to be synthesized (Figure 38).

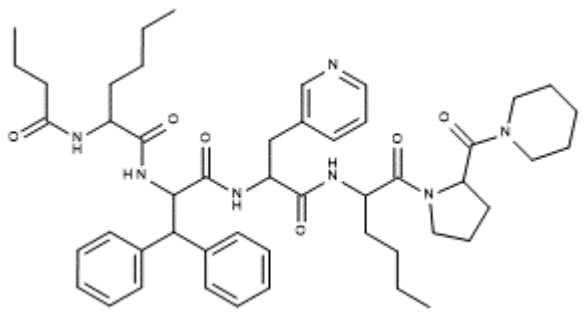
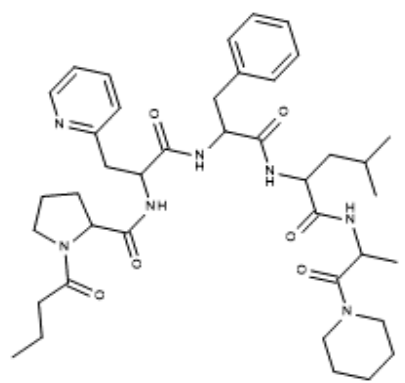
Iproteos Code	Formula	Structure	Docking score (Kcal/mol)	IPro Permeability (Å <sup>2</sup> )
IPR-463	C <sub>49</sub> H <sub>67</sub> N <sub>7</sub> O <sub>6</sub>		-11.4	141
IPR-464	C <sub>40</sub> H <sub>57</sub> N <sub>7</sub> O <sub>6</sub>		-9.7	135

Table 8. Once applied the IPro Docking over a penta- and tetra-peptidomimetic virtual library, and after post-docking filtering, two penta-peptidomimetics were selected for synthesis and experimental evaluation. Although any of both structures had a better score (-11.4 kcal/mol) than the binding motif of L2 (Ac-DFYLH-NH<sub>2</sub>), good docking scores (< -8.0 kcal/mol) and the same binding mode of the natural occurring peptide were determining factors for their selection.

Interestingly, both peptidomimetics reproduced the same binding mode of the original peptide. Then, even not being superior in terms of docking scores, positive hits could validate the docking model, and the chances to optimize the potency of the structure in a second round.

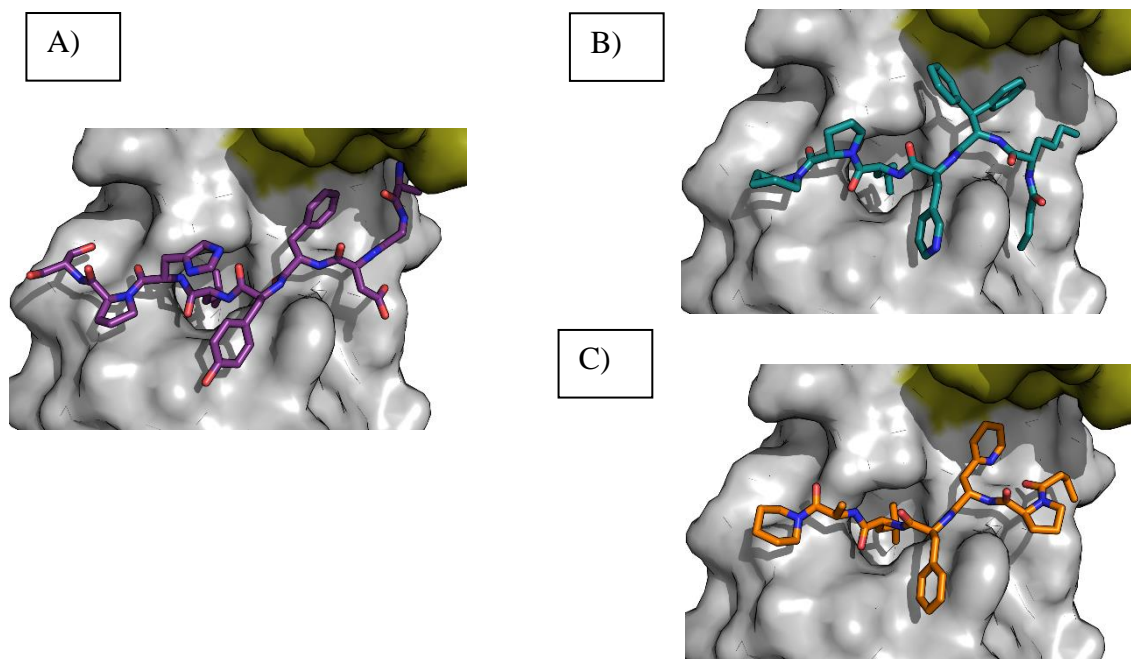


Figure 38. a) Zoom in the binding site of Ac-DFYLH-NH<sub>2</sub> to retromer B) Docking model of IPR-463 C) Docking model of IPR-464. The two *de novo* designed peptides mimic perfectly the binding mode of the crystalized short-peptide.

### 6.3. Experimental evaluation

As a first screening approach, Isothermal Titration Calorimetry (ITC) was the experimental technique applied to evaluate the binding of the compounds. ITC is a biophysical technique in which fluctuations of the temperature due the binding of a protein in solution to a possible binder is measured. These variations are translated into the binding potency based on reference cell.<sup>161</sup> ITC is binding evaluation technique that is independent from permeability or other parameters that are present in cell environment. Meaning a straight measure of the binding potency.

Unfortunately, IPR-463 was not soluble when it was dissolved in the ITC buffer (25 mM Hepes, NaCl 300 mM and TCEP 0,5 mM) even at micromolar concentration range. For this reason, and guided by other previous experiences, the solubility of IPR-463 was studied in the buffer assay at 15 % CDex.

Iproteos Code	Maximum Solubility in ITC buffer with 15 % H $\beta$ -cyclodextrin ( $\mu$ M)
IPR-463	345 $\pm$ 0,3

	Retromer + HPV	IPR-463	IPR-464
$K_d$ ( $\mu\text{M}$ )	49,2	32,3	7,7
$\Delta H$ (Kcal/mol)	-22,5	-3,6	-80
$T\Delta S$ (Kcal/mol)	16,6	-2,5	73
$\Delta G$ (Kcal/mol)	-5,9	-6,1	-8,7

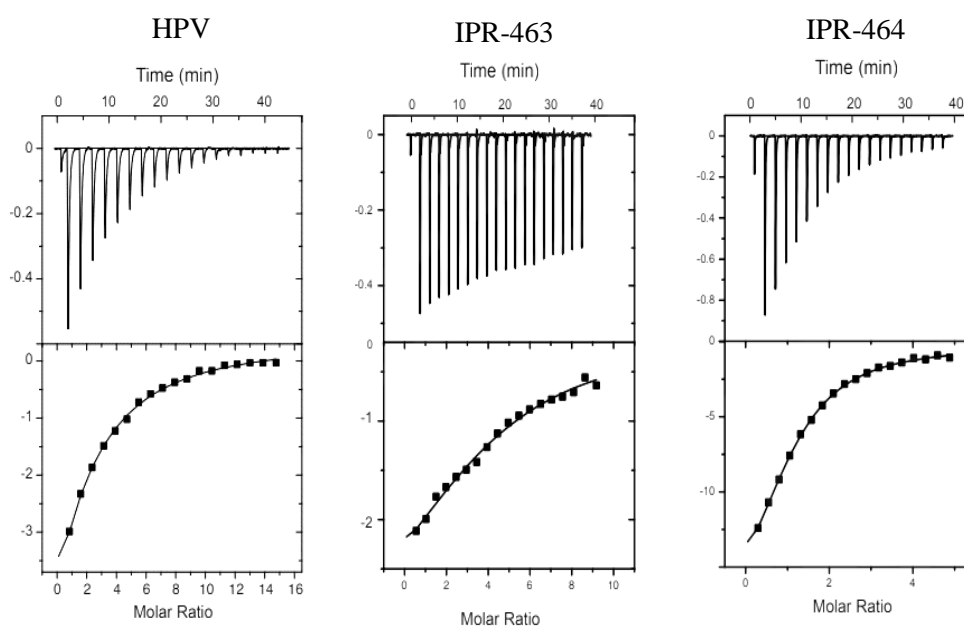


Figure 39. ITC experiment results; HPV is the short peptide Ac-ADAGDFYLHPSYYMLRKRK-NH<sub>2</sub> that corresponds to the C-terminal of the L2 protein of HPV, which binds to retromer protein. IPR-463,  $K_d = 32.3 \mu\text{M}$ , and IPR-464,  $K_d = 7.7 \mu\text{M}$  respectively, present a higher potency than the natural peptide. IPR-463 was evaluated in ITC buffer at 15 %  $\beta$ -cyclodextrin, and in order to discard any issue related with  $\beta$ -cyclodextrin, a control in these conditions without the presences of any molecule under study was carry out (data not shown). In this regard, no issue with  $\beta$ -cyclodextrin was found.

For ITC experiments, the molar extinction coefficient was measured in ACN/H<sub>2</sub>O to allow Dr. Aitor Hierro's team to calculate the real concentration when the biophysics assay was performed. Calibration curves were determined at 100, 50, 20 and 10  $\mu\text{M}$  points. The molar coefficient was calculated according the protocol described in Materials and Methods Chapter.

Iproteos Code	Molar extinction coefficient; $\epsilon$ ( $\text{L} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$ )
IPR-463	8601,7
IPR-464	3397,3

C-terminal derived peptide from the HPV's L2 protein that binds to the Retromer was used as a positive control during the ITC evaluation. As the main approach of this project was design a peptidomimetic that could compete against L2 for the binding to Retromer, then the peptides should have a higher potency than the natural peptide. The 20 amino acids peptide showed a  $K_d = 49.2 \mu\text{M}$  whereas both designed compounds bounded to the retromer at a lower concentration, IPR-463,  $K_d = 32.3 \mu\text{M}$ , and IPR-464,  $K_d = 7.7 \mu\text{M}$ . These positive results encouraged Dr. Aitor Hierro's team to attempt the crystallization IPR-464, the most potent of the two tested compounds with retromer protein and pursued future experiments focused to validate the docking model and the binding to the retromer.

#### **6.4. Conclusions**

ITC experiment results demonstrated the capacity of IPRoTech to generate peptidomimetics able to bind to the Retromer protein with a higher potency than the naturally occurring peptide, which is the responsible of HPV endosomal release into cytosol. This result validates, validating the docking model used, which has efficiently reproduced the binding motif of the Ac-DFYLH-NH<sub>2</sub> peptide and delivered two peptidomimetics with a superior potency than this peptide. For this project, the permeability of both molecules remains to be studied, a property that is considered to be a key factor for any possible therapeutic treatment as the target protein has an intracellular localization.

This ongoing project will further be investigated if IPR-464, the most potent of the two peptidomimetics evaluated, is able to compete for Retromer with the HPV L2 protein in cell based experiments. Also, the crystallization of IPR-464 and the Retromer protein will be carry out in order to validate the binding site of the compound and its mechanism of action.

Moreover, if all experiments succeed, it was agreed along with Dr. DiMaio's group, which has a strong background in therapeutic research in the field of HPV,<sup>162</sup> to investigate the use of IPR-464 in HPV relevant mouse model.



## Chapter 7: Discussion





## 7.1. IPROTech

Many major diseases, neurodegenerative diseases and rare diseases remain untreatable, as a consequence of a high complex mechanism that usually involve the interaction between proteins. Moreover, some of these therapeutic targets are located inside the cell which accounts for an increased difficulty for an increased difficulty to obtain successful drug candidate. Indeed, most challenging therapeutic targets are intracellular protein-protein interactions. PPIs are known for the lack of surface cavities, on the protein interaction areas, not being amenable for a small molecule approach and also their intracellular location that limits the application of antibodies or other macromolecular approaches that could bind the interface area with high affinity.

These complexities explain the pharmaceutical needs for new molecules than can cover different biochemical properties compared to traditional small molecules. Peptides-based molecules represent a solution to disrupt PPIs that cannot otherwise be targeted for small molecules as are able cover larger surface on the receptor protein and therefore establish more interactions.

At Iproteos an in-house technology platform (IPROTech) was developed to generate permeable peptidomimetics able tackle PPIs. This proprietary technology has been built to design new privilege peptidomimetic structures that are able to bind the protein of interest and are predicted to be permeable across biological barriers by passive diffusion. Despite there is a wide variety spectrum of computational tools to be applied for small molecules, few benchmark and computational work related with the application of peptides in the field of drug discovery has been published.

Furthermore, inclusion of non-natural amino acids and post modifications on the designed peptide sequences represents a key step to obtain optimized and potent candidates. Finally, for those compounds with a promising activity when evaluated *in vitro*, good solubility in aqueous environments is a physicochemical property require, also for the most successful molecules, prior to their evaluated *in vivo*.

On this basis, IPROTech was experimentally validated by it successfully application in the 4 projects that are detailed in this thesis. The technology was implemented to design potent PPI inhibitors in a high success-rate. More interestingly, the generated data has allowed to optimize the docking protocol and the post-docking results filtering to improve the accuracy of the *in silico* prediction.

<b>Project</b>	<b>Synthesized compounds</b>	<b>Hits</b>	<b>Succes-rate %</b>
Talin/vinculin	37	7	19
Rad51/BRCA2	10	5	50
Ras/Effectors	13	5	38
Retromer/L2	2	2	100
<b>All</b>	<b>62</b>	<b>19</b>	<b>31</b>

SPPS was proved as solid and straightforward technique for peptidomimetic synthesis. During this thesis new methodologies were included in the company know-how, such as the introduction of several secondary amines as C-terminal cappings or the introduction of N-alkylations with longer carbon chains. Furthermore, the use of a high variety of non-natural amino acids was included in the compounds design.

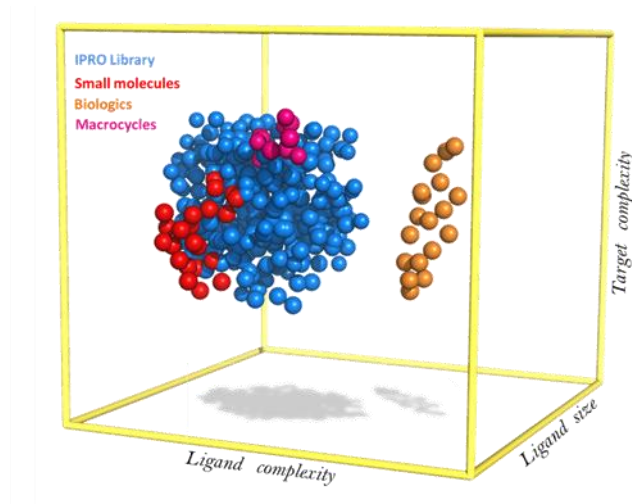


Figure 40. Peptidomimetics are in the chemical space between traditional small molecules and biological macromolecules. Thus, allowing them to target challenging therapeutic proteins that are considered undruggable for small molecules and saving all the complexity of new biologic approaches. Moreover, the high number of different commercially available Fmoc protected amino acids along with the post modifications over the peptide sequence allow a high number of possible variations even for short sequences (3-4 amino acids).

PAMPA was determined to not be the most appropriate biophysical technique to evaluate the permeability by passive diffusion of high hydrophobic and bulky peptidomimetics. We speculate that the material of the PAMPA filter (PVDF) could be underestimating the  $Pe$  values. Indeed, obtained PAMPA data are not aligned with the experimental results when peptides were incubated with cells, and demonstrated that most of them were able to be internalized by cells, as was predicted *in silico*. This argument was also supported when 2017/IP-18-55, the best candidate from the Talin-Vinculin project, was evaluated in PAMPA assay without applying a lipidic barrier. In this terms, the only barrier separating the donor and acceptor compartments was the PVDF plastic filter. Therefore, it was expected a high transport % and almost a null retention % as in these conditions. However, it was found that the compound was highly retained by the plastic filter and as a consequence showed a low transport.

Iproteos Code	$Pe$ (cm/s)	Transport %	<b>Retention %</b>
2017/IP-18-55	$2,4E-07 \pm 1,02E-7$	$2,0 \pm 0,8$	<b><math>83,1 \pm 0,4</math></b>

On the other hand, evaluation of the permeability of the generated compounds by studying their cellular uptake after incubation with SH-SY5Y cells was found to be time consuming and expensive approach. Its scalable application as permeability screening

tool was limited, moreover the protocol was developed in-house, and the information should be validated by other means if a publication was seek.

On the other hand, balance solubility and permeability for those compounds that are bRo5 is a challenging task that can end up in a bottle neck situation. For this reason, the use of cyclodextrins as a solvating excipient was a successful approach that has allowed the evaluation of compounds with low solubility. Hence, allowing the evaluation of the compounds in all the performed projects.

Finally, the use of non-natural amino acids along with selective *N*-alkylations on the compound backbones has led to obtain molecules with a high proteolytic resistance, a pivotal feature for the successful evaluation of drug candidates.

## 7.2. Talin-Vinculin

This mechanosensing based project was the most challenging from all the projects. The lack of any previous reported inhibitor along with the cell viability assay as the only experimental screening technique, at least for the first year of project, challenged the development of this business case. Furthermore, although the role of YAP translocation has been recently demonstrated by Dr. Pere Roca-Cusachs group, the therapeutic application of the PPI disruption has not been validated so far.

Big efforts were done to obtain a potent inhibitor that produced cell death in two relevant cell lines for the target PPI, namely PANC-1 and Talin *-/-*. It was found three hits that showed a cytotoxic effect in cells when were incubated at 60  $\mu$ M. The combination of the most relevant structural motif of the compounds was used to design a second generation of peptidomimetics. This second generation evaluation lead highlighted the compound 2017/IP-18-55, which showed and  $IC_{50} \sim 5 \mu$ M in the same cell conditions. This potency was considered valid to start biophysical and *in vivo* experiments and validate the mechanism of action of the peptidomimetic and its possible therapeutic application.

Traction forces microscopy read out reinforced the predicted mechanism of 2017/IP-18-55 for which the compound was bound to folded Talin and therefore preventing the protein unfolding when forces that normally unfold Talin when are  $>10$  Kpa over the protein<sup>101</sup>. Also, the high stability of the peptide in rat plasma (totally stable after 2 h incubation) and its good solubility (5,1 mM and 2,4 mM in PBS at 15 % and 5 % H $\beta$ -CD respectively) in the presence of cyclodextrins were determinant factors to start SPR, NMR and *in vivo* experiments. Inhibition of YAP's translocation by blocking the mechanosensing cluth as a possible therapeutic solution for solid cancers would open a new complete scenario in the cancer research. Meaning, that 2017/IP-18-55 or a possible more potent derivate from it would be a first-in-class product.

Unfortunately, recently performed SPR and NMR experiments did not find a direct binding of 2017/IP-18-55 to the R3 domain of Talin, which did not allow us to validate the mechanism of action of the compound. In addition, studies in pancreatic cancer mice models did not show a therapeutic activity when the animals were administered with the peptide at different doses.

Altogether, generated results highlight how important is to have a screening technique that can validate the mechanism of action as soon as a hit is found in order to develop a solid optimization process. Furthermore, if reach preclinical phases is seek, the existence of reported inhibitors or a molecule that validates the mechanism of action is highly desirable and additionally can be used as a positive control.

### 7.3. RAD51-BRCA2

This work was done based on the proof-of-concept experiment published by Prof. Sartori's group,<sup>128</sup> in which a BRC4 repeat attached to a cell penetrating peptide is able to disrupt the PPI. This knowledge was included in the computational analysis and allowed to obtain a high success rate, 60 %, from the very beginning. From the 5 synthesized molecules, 3 showed a therapeutic window in which cells were killed only in the presence of Olaparib, a PARP inhibitor.<sup>118</sup> This synergistic effect with Olaparib was the expected effect for a compound able to inhibit the Double Strand Repair mechanism in a clonogenic cell assay. From these 3 molecules, based on the structure of the most potent compound, IP-15-04 (IC<sub>50</sub> of 10-20  $\mu$ M), an optimization process was carried out.

From this optimization process, 4 new candidates were synthesized and 1 of them was found to be active with an approximate IC<sub>50</sub> of 1-2  $\mu$ M in cells. This IC<sub>50</sub> outperforms the best hit of another recent work published by Dr. A. Cavalli's lab.<sup>163</sup> In this work they identified a triazole family and based the design and synthesis of 42 small molecules in this scaffold. The most potent molecule showed an IC<sub>50</sub> up to 10-20  $\mu$ M in cell viability experiment with BxPC3 cells.

This ongoing project has resulted in a very promising work and future experiments will determine the scope of its applicability. As final remarks, it can be extracted that a good structural model is determinant for the high accuracy *in silico* studies. Furthermore, the substitution of a natural peptide for a specifically designed peptidomimetics allows an increase in the binding potency. As the natural peptide had a potency of  $\mu$ M meanwhile the second generation of peptidomimetics showed an IC<sub>50</sub> of 1-2  $\mu$ M.

### 7.4. RAS-Effectors

RAS is well-known therapeutic target involved in many cancer diseases. As a consequence, there is much public information about the protein function and structure that can be downloaded from bibliographic sources and the Protein Database Bank.

Public data allowed us to build an accurate docking model and the design of potent inhibitors that can disrupt the interactions of RAS with their effectors. The evaluation of the RAS/effector signaling protein cascades inhibition was carried out by Prof. Neus Agell's group. On the first round evaluation, four hits were identified when incubated with a pancreatic cell line at 20  $\mu$ M.

In the second round of evaluation, after the optimization of IP-14-01, two additional compounds were identified. Then, supporting the docking model that allowed the designing of the candidates.

The scope of this ongoing project was dependent of the capacity of the most promising peptidomimetics to discriminate between cancer and normal cells. For this reason, the two most potent hits, 2018/IP-14-01 and its derivative analogous IPR-473 were evaluated in a cell proliferation assay with normal and pancreatic cancer cell lines. Although, the parent peptidomimetic was unable to efficiently kill cells, either cancer cell or normal cells, IPR-473 efficiently discriminate at the range of 15-25  $\mu\text{M}$  between healthy and cancer cells, causing cell death to only cancer cell lines. Importantly, all studied cancer cell lines were derived from pancreas cancer, for which the isoform of K-RAS is the more aggressive. Recently, Prof. Kessler's group has published a set of small molecules that bind the SI/II-pocket of RAS at nanomolar range in ITC experiments. The most promising compound, BI-2852, reduced pERK and pAKT levels in a dose-dependent manner in NCI-H358 cells, leading to an antiproliferative effect at  $> 10 \mu\text{M}$ .<sup>164</sup> This work has found that validates the druggability and possible therapeutic effect of the SI/II-pocket and consequently has validated our approach as well, as all peptidomimetics were designed to target a similar area around SI/II-pocket.

The impact of the present project will be determined once the bioavailability of IPR-473 will be evaluated along with its pharmacokinetic properties. Therefore, more experiments will be conducted in order to reach preclinical phase. Additionally, a publication will be sought and the intellectual property of IPR-473 protected.

## 7.5. Retromer-L2 protein

The *in silico* design and synthesis of two peptidomimetics to inhibit the interaction between the Retromer protein and the L2 protein of the HPV was carried out. The computational studies were based on the work done by Dr. Hierro's group, which had previously crystalized the region of the Retromer protein that binds L2 protein of the HPV virus. Concretely, they have crystalized the Retromer bind to a derivate short-peptide of the C-terminal region of L2 protein, Ac-DFYLNH<sub>2</sub> (unpublished data). This interaction is involved in the endosomal release mechanism that HPV virus uses to escape from the lipid capsule once has been internalized inside the cell.

At that point, the maturity of the in-house technology and the distribution of resources were the factors that propitiated that only two candidates were proposed for synthesis and further evaluation. Both compounds resembled the same binding mode of the short-peptide and showed similar docking scores, the Ac-DFYLNH<sub>2</sub> showed a score of -11.4 kcal/mol and IPR-463 and IPR-464 had a score of 11.4 kcal/mol and -9,7 kcal/mol.

Both compounds, IPR-463, 32,7  $\mu\text{M}$ , and IPR-464, 7,7  $\mu\text{M}$ , were more potent than the natural occurring derivate peptide from L2 C-terminal, when were evaluated by ITC.

IPR-464 was the most potent peptidomimetic after the inspection of it predicted binding mode it was speculated that a second round could improve the potency to a nM scale. Likewise, the behavior or IPR-464 will be studied in the upcoming months by Dr. Aitor Hierro's group, evaluating the capacity of the compound to inhibit the HPV release in cells. Furthermore, it will be attempted to co-crystalize the peptidomimetic with the

Retromer and *in vivo* experiments will be sought by Dr. DiMaio's group. In our knowledge, if a therapeutic activity is found for IPR-464, would not only validate the mechanism of release of the HPV, but also would represent the first compound exploiting this mechanism.

## Conclusions





The main goal of this thesis was to apply the Iproteos technology, IPROTech in high scientific impact projects in the field of therapeutic protein-protein interactions. Thus, allowing a validation of the in-house technology efficacy, integration of the generated knowledge into the company core and at the end, enlarge the Iproteos pipeline.

In this scenario this thesis has been focused in accomplish the initially set objectives:

- 1) *Synthesis of high purity (>95 %) novel peptidomimetics that had been designed applying the IPROTech technology in order to disrupt the PPI of interest.*

Over 60 peptidomimetic compounds were designed by using the IPROTech *in silico* tools and manually synthesized by SPPS. Almost all peptide sequences included at least one non-natural amino acid. In addition, highly bulky *N*-terminal cappings were implemented into the lineal sequence. In these terms, secondary amines were also included as C-terminal cappings. Finally, *N*-alkylation of the peptidomimetic amide bonds with longer carbon chains was implemented. Aiming a theoretical reduction of the total structure polarity and an increase of Van der Waals contacts with target protein.

Meaning a unique design of peptide-based structures with a wide variety of sequence modifications that allowing a feasible optimization for a hit compound.

- 2) *The company's focus is targeting intracellular PPIs, therefore permeability across biological barriers is a key asset. Meaning an evaluation of the permeability of the synthesized compounds across biological barriers and/or their capability to be uptake by cells.*

Although PAMPA assay is well-known and accepted technique to evaluate the permeability of compounds by the passive diffusion, its applicability has been proved to be useless for the majority of peptides synthesized in this thesis. All the molecules were predicted *in silico* to be permeable. These predictions were confirmed for all active peptides that have been evaluated *in vitro* (17) and therefore they should cross the cell membrane to reach the target protein. However, their *Pe* value determined by PAMPA indicated the opposite. We concluded that it would be too improbable that the 17 active peptidomimetics were being internalized by active transport mechanisms, and then we conclude that the experimental prediction of the compounds permeability by using PAMPA assay was not the most appropriate approach for the behavior of Iproteos molecules.

To overcome this issue, other biophysical screening techniques were proposed to be implemented, such as IAM chromatography, which predicts measures the capacity of a compounds to cross a column filled with 10-12 branched carbon chain.

- 3) *In order to have permeable compounds, their hydrophobicity and bulky side chains is increased in detriment of solubility. A good solubility is essential for any drug candidate and as a consequence study new approaches to overcome this issue is mandatory to overcome the solubility bottle neck.*

Any modification of the peptide sequence that aimed an increase of solubility will be produce automatically a drastically loss of permeability. For this reason, the use of solubilizing excipients was studied. Cyclodextrins are a family of cyclic oligosaccharides widely used in pharmaceutical formulations, and therefore were selected by its bioavailability administration.

H $\beta$ -Cyclodextrin, the family member chose, showed an outstanding solubilizing potency, by diluting all synthesized compounds that shown a poor solubility in PBS, over 100  $\mu$ M. Thus, allowing to carry out all the planned *in vitro* experiments.

- 4) *Demonstrate the IPROTech efficacy by direct application in relevant scientific projects where their therapeutic mechanism of actions is related with the disruption of a PPI.*

First, the technology was applied to disrupt the Talin/Vinculin PPI, a very challenging target because its mechanosensing mechanism. A molecule, 2017/IP-18-55 reached *in vivo* studies although unfortunately a therapeutic effect was not observed.

For the other three projects, high potent inhibitors were found. At the present moment, the three projects are still ongoing and potentially they could reach preclinical phases.

- 5) *Integration of all generated data into the company know-how in order to optimize the IPROTech efficacy.*

From the first to the last started project new features have been incorporated into the computational platform and in the peptide designing process thanks to the experimental data obtained. That has conducted to a substantial increase of the docking results accuracy.

A relevant example is the project to find inhibitors of the Retromer/L2 interaction for which only two peptides were synthesized and both of them were more potent than the L2 binding motif derivate peptide. Equally important is the fact that for all the four targeted PPI more than one hit was identified on the first screening evaluation, highlighting the potential of the technology and the application of peptidomimetics for these challenging targets.

- 6) *Include new business cases in the company's pipeline to diversify the number of projects.*

Fruitful success in the pharmaceutical field for small companies relies in the development of a solid pipeline in which different molecules are studied for their possible prescription as a drug treatment for a given diseases. In this regard, Talin/Vinculin inhibitors failed when studied *in vivo*, being until date, the most advanced project. Fortunately, the high potency and *in vitro* activity of IPR-469 and IPR-468 (RAD51/BRCA2), IP-14-01 and IPR-473 (RAS/Effectors) and IPR-464 (Retromer/L2), still opens the door to reach preclinical phases.

## *Materials and Methods*



## 1.1. General protocol for computational studies

All virtual candidates were designed and selected for synthesis by applying molecular modelling techniques. A first evaluation of the therapeutic target published information (crystallographic or NMR structures, reported inhibitors and PPI binding site) determines the execution of the following steps.

### 1.1.1 Docking general protocol

Once the active site was defined by means of a reported structure or elucidate after applying IPRO Filter in-house technology, a high-throughput docking analysis assay was performed. Typically, the size of the virtual library comprises 50K-100K peptidomimetic structures. All compounds and the protein under evaluation are protonated at physiological pH (7.4) and transformed to a PDBQT format by using MGLTools. Autogrid was used to generate the grid size<sup>165</sup>. Usually, the box was set to 50 x 50 x 50 points with a spacing of 0.375 Å and centered into the active site. Then, docking exhaustivity conformation search of 32 or 8 (based on the particular computational resources) in an energy window of 20 kcal/mol. No protein residue flexibility was allowed. Seed value number was set to 31415 for reproducibility issues.

After performing the docking evaluation, Ramachandran analysis to prime selection was applied. Briefly, this post-filtering analysis is devoted to identify those sequences able to adopt suitable phi/psi dihedral angles. If not, docking conformation is rejected as plausibly classified as a false positive docking response. This filtering algorithm, supported by evidences found in peptide-docking benchmarks, reduces the number of potential docking solutions in a factor of 10-20x.

Additionally, a contact post-filtering algorithm profiling the number of inter- and intramolecular hydrogen bond contacts (intra-HB) formed between the protein and the ligand was applied. All conformations with and intra-HB lower or equal than 1 were retained for stability verification. This algorithm is devoted to reject false docking solutions where docking engine has primed low-entropically solutions rather than extended peptidomimetic conformations.

Because of large peptidomimetic flexibility, stochastic result verification was conducted by computing the root mean square deviation (RMSD) between the first two top-ranked poses of each sequence, retaining only those sequences where RMSD between former and second accepted pose (after applied post-filters) was lower than 3 Å.

To evaluate suitable score threshold selection, histogram score analysis was computed. The average score was of -7.5 kcal/mol. Compounds with score larger than averaged value were selected for visual inspection (< 200), rejecting those situations where unrealistic bioactive conformations were found (e.g. close orientation of positive charged residues to hydrogen bond donors, close orientation of negatively charged residues to hydrogen bond acceptors and insertion of polar residues into highly hydrophobic clefts).

### 1.1.2. IPRO permeability general protocol

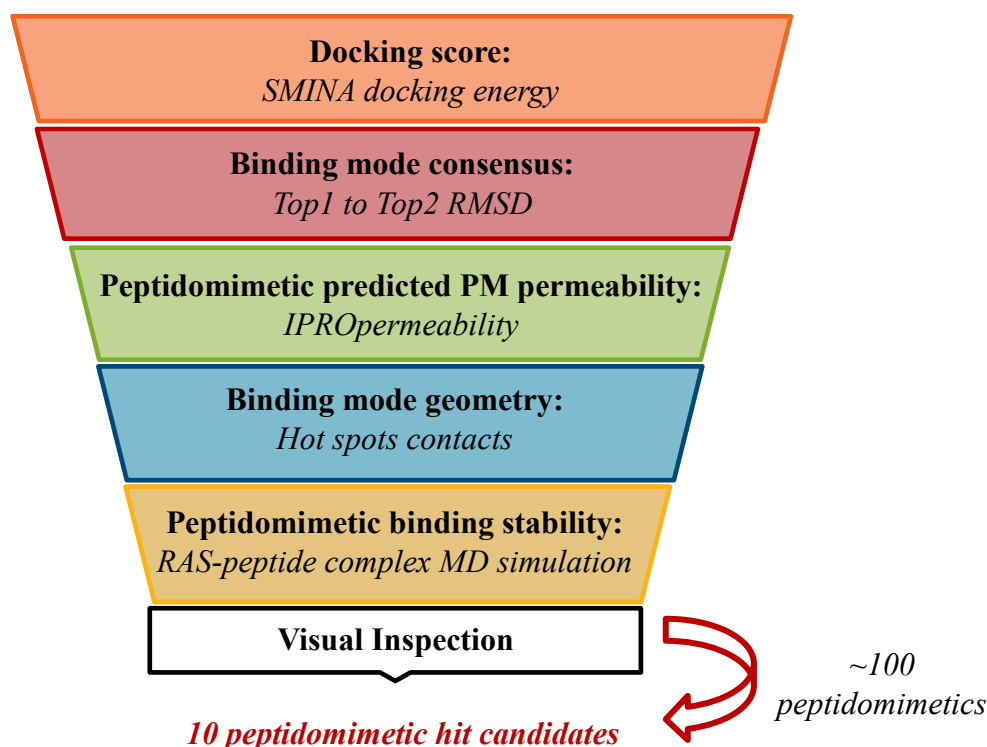
Retained docked conformations were evaluated by IPRO Permeability tool. The simulation protocol was conducted with Sander module of AmberTools package.

Each sequence structure was subjected to below sequential workflow:

- Generation of random linear extended conformation of hit sequences. This procedure assures sub-sequent steps are not biased by docking conformation.
- Linear random peptidomimetic conformation subjected to conformational landscape exploration to identify the 25 most representative conformations of the molecule conformational ensemble. Each conformation was profiled in terms of Ramachandran verification to ensure unbiased effects along the simulation.
- Generation of coordinate, topological and configuration files required to conduct implicit solvent conditions for each conformation. A total time of 2 nanoseconds (ns) was explored, yielding an accumulated simulation time of 50 ns.
- Post-analysis evaluation. The former step is conducted by cpptraj module<sup>166</sup> of AmberTools to extract from each conformation the sampled geometries in PDB format. Secondly, each conformation is evaluated by ICM-browser software<sup>70</sup> to compute individual polar accessible surface area (pASA).
- Statistical analysis was completed with calculation of averaged pASA and standard deviation.

On the following figure can be found a summary of the main steps to obtain *de novo* designed peptidomimetic structures to target the protein under study.

~50K-100K  
input sequences



## 1.2. General protocol for peptide synthesis

All compounds were synthesized by means of SPPS following an Fmoc/*t*Bu strategy. Syntheses were performed on a 100  $\mu$ mol-scale/each, using L-, D- and non-natural amino acids. Syntheses were done manually in polypropylene syringes with intermittent manual stirring during the coupling process to mix reagents. Solvents and soluble reagents were removed by suction.

### 1.2.1. Resin initial conditioning and coupling of the first amino acid

Resin selection was determined by the C-terminal moiety of the compounds. For those compounds having a carboxylic acid group or a C-terminal capping, the 2-chlorotrityl chloride resin<sup>167</sup> was used. On the other hand, H-Rink Amide Chemmatrix resin<sup>168</sup> was the used for peptides sequences having a carboxamide moiety at C-terminal.

### 1.2.2. 2-chlorotrytil chloride resin conditioning: coupling of the first amino acid

Resin was initially washed with DCM and DMF (5 x 1 min DCM, 5 x 1 min DMF and 1 x 5 min DCM). After that, 0.6 equivalents of the first amino acid were mixed with few drops of DCM and added to the resin and 5 equivalent of DIEA were added in two portions, first, 1/3 part of the amount of DIEA was added and allowed to react for 10 min, then the remaining amount of DIEA (2/3) was added and allowed to react for 50 min. Next, the Fmoc group was removed by using a mixture of 20 % piperidine in DMF (4 mL/g resin, 2 x 1 min and 1 x 10 min). The piperidine mixtures and DMF washes were



collected in a volumetric flask, the volume of which was completed with DMF. After this, its UV spectroscopy was measured in order to calculate the functionalization of the resin according to equation 1.

$$(1) \quad Z = \frac{A \cdot X}{\epsilon \cdot Y \cdot l}$$

A = Absorbance, X = Volume of solvent (mL),  $\epsilon$  = Molar coefficient ( $5800 \text{ L} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$ ), Y = Resin weight (g), l = Length of the cell (cm) and Z = Loading of the resin.

#### 1.2.3. H-Rink Amide Chemmatrix Resin conditioning and coupling of the first amino acid

H-Rink Amide Chemmatrix was washed with DMF and DCM (5 x 1 min DCM, 1 x 60 min DCM, 1 x 5 min DMF and 1 x 20 min DMF) before the addition of the first amino acid. After that, 3 equivalents of the amino acid were added after 3 min of pre-activation in a mixture with 3 equivalents of Oxyma pure and 8 equivalents of DIC and few drops of DMF and let it react overnight. Oxyma was used as it allows performing an overnight reaction without the risk of racemization. The extend of coupling of this first amino acid was evaluated using the Kaiser's test.<sup>169</sup>

#### 1.2.4. Peptide chain elongation

Amino acid couplings were performed following the next steps. Wash of the resin (5 x 1 min DMF), Fmoc removal of the amino acid already anchored onto the resin (2 x 1 min 20 % piperidine in DMF and 1 x 10 min 20 % piperidine in DMF), wash (5 x 1 min DMF and 5 x 1 min DCM), colorimetric test (Kaiser<sup>169</sup> or Chloranil<sup>88</sup>), coupling of the amino acid (4 equivalents of Fmoc-Aa-OH, 8 equivalents DIEA, 4 equivalents of TBTU and few drops of DMF. All the mixture is pre-activated during 3 min, then was added to the resin for 75 min), washed (5 x 1 min DMF and 5 x 1 min DCM).

Finally, a colorimetric test (Kaiser and Chloranil test) was applied to evaluate the extension of the coupling reaction. In those cases where the reaction was incomplete, the recoupling was performed under the standard conditions of coupling.

#### 1.2.5. Amino acid N-alkylation

The process for N-methylation of amino acids used was the method described by Miller *et al.*,<sup>170</sup> which is divide into the three following steps (these steps are performed after Fmoc removal of the amino acid anchored onto the resin which was going to be N-alkylated):

1. Protection and activation of the amino group with *o*-NBS: 4 equivalents of *o*-NBS, 3 equivalents of 2,3,5-Collidine and few drops of DMF (1 x 30 min and 2 x 20 min).
2. Deprotonation and N-methylation: 3 equivalents of 1,8-Diazabicyclo[5.4.0]undec-7-ec in DMF added to the resin (5 min), after that 10 equivalents of Dimethylsulfate were added on the resin (10 min). This treatment was repeated 2 times more.
3. *o*-NBS removal: two treatments with 10 equivalents of  $\beta$ -mercaptoethanol, 5 equivalents of 1,8-Diazabicyclo[5.4.0]undec-7-ec and the minimum volume possible of DMF for 2 times (1 x 10 min and 1 x 40 min)

#### 1.2.6. Cleavage of the peptide from the resin

If lateral-side chain protecting groups were need to be preserved after cleavage, the following protocol was applied. Otherwise, this step was skipped and the same protocol described in *Remove of protecting groups (1.2.9.)* was applied.

The peptide-resin was treated with a mixture of 5 % of TFA in DCM (3 x 15 min, 6 mL). The mixture treatment and DCM washes were removed by suction and collected in a round bottom flask and combined in order to obtain the cleaved peptide from the resin. Then, the solvent from the collected mixture was evaporated under vacuum until dryness. Finally, the solid residue was diluted with ACN/H<sub>2</sub>O solution (50:50) and lyophilized. This acidolytic treatment allowed to obtain the side-chain protected version of the peptide.

To obtain a head-to-tail compound, the next step was performed

#### 1.2.7. Lineal peptide chain head-to-tail cyclization

Estimating a synthetic yield of 70 %, the crude obtained after the lyophilization was dissolved in DMF to obtain a 0.005 M concentration. After that, 8 equivalents of NaHCO<sub>3</sub> and 2 equivalents of DPPA were added. The mixture was stirred for 72 h at room temperature under constant stirring in a closed flask. Afterwards, the solvent was evaporated under vacuum.

On the other hand, if the final product contains a C-terminal capping group the following procedure was applied.

#### 1.2.8. C-terminal capping

The crude of synthesis was dissolved in the minimum volume possible of DCM and 3 equivalents of the capping group, 3 eq. of HOAt and 3 eq. of EDC·Cl were added. The mixture was allowed to react for 3 h at room temperature under constant stirring. Once the reaction was completed, an extraction was performed by means of washes 3 times with each saturated solution, NaHCO<sub>3</sub>, NH<sub>4</sub>Cl and NaCl. The organic layer is then dried with Na<sub>2</sub>SO<sub>4</sub> anhydride, and the solvent was evaporated under vacuum, then was dissolved with a mixture of ACN:H<sub>2</sub>O (50:50) and lyophilized.

#### 1.2.9. Remove of the lateral chains protecting groups

Lateral chains protecting groups were removed through and acidic mixture treatment. For the peptides generated in this project, side-chains were removed using a mixture of TFA 95 %: TIS 2.5 %: H<sub>2</sub>O (15 min x 3).

#### 1.2.10. Peptide purification

Compounds were purified using Semipreparative RP-HPLC (Waters) or Teledyne ISCO system for crudes with low impurities. The crude to be purified was dissolved in ACN:H<sub>2</sub>O (using the lowest amount as possible of ACN:H<sub>2</sub>O).

The column used was C18. The gradient used was G0100t30min. Solvents: ACN with 0.05 TFA % and H<sub>2</sub>O with 0.1 % TFA. Flow rate= 16 mL/min. Detection = 220 nm.

Collected fractions of interest were analyzed by analytical HPLC and HPLC/MS combined and lyophilized.

For difficult peptide crudes, instead a gradient of G0100t30min, a more narrowed gradient was used to obtain a better separation of the peak corresponding to the desired product from the impurities.

#### 1.2.11. Colorimetric tests

The Kaiser test, also known as ninhydrin test, is used during solid-phase peptide chain assembly to monitor the deprotection and coupling extends due to a color switch, from yellow to blue in case of primary amines presence. The method presents a high sensitiveness and a negative test assures a coupling rate higher than 99 %.

For the test a few dried peptide-resin beads (previously washed with DCM) were transferred to small glass tube. Afterward 6 drops of reagent solution A and 2 drops of reagent solution B were added and the final mixture was heated at 100 °C for 3 min.

Reagent solution A: In 100 mL of absolute EtOH, 400 g of phenol were added and heated until complete dissolution. Then, 65 mg of KCN were dissolved in 100 mL of H<sub>2</sub>O and 20 mL of this mixture were added to 1000 mL of freshly distilled pyridine over ninhydrin. Both solutions were stirred for 45 min with 40 g of Amberlite MB-3 ion exchange resin, filtered, and combined.

Reagent solution B: 2.5 g of ninhydrin are dissolved in 50 mL of absolute EtOH and the final solution was kept in flask protect from light.

The chloranil test allows the detection of secondary amines between SPPS steps and is used to evaluate couplings onto proline or *N*-methylated residues due to a color switch from yellow, amber or brown to blue (indicative of the presence of free secondary amines).

For the test a few dried peptide-resin beads were transferred to small glass tube (previously washed with DCM). Afterward, 20 µl of a saturated chloranil solution and 200 µl of acetone were added. The mixture was allowed to react for 10 min at room temperature.

### **1.3. Liquid chromatography**

#### 1.3.1. ISCO

Crudes to be purified were dissolved in ACN (using the lowest amount as possible of ACN), mixed with silica C18, using a Gold C18 column aq. Gold (20-40 µm, 100 Å). Finally solvent was removed by evaporation until dryness. After which, the obtained solid was loaded in solid form into the Teledyne Combi flash ISCO RF system. Purifications were run using 30 min gradient of ACN with 0.1 TFA% in H<sub>2</sub>O 0.1% TFA. Flow rate = 15 mL/min. Detection = 220 nm.

#### 1.3.2. Semipreparative-HPLC

UV chromatograms were record with Waters 2998 photodiode array detector module. Detection was performed at 220 nm. The crude products obtained were purified by reverse-phase column chromatography, Aeris 5 µm, PEPTIDE XB-C18 100 Å, LC

COLUMN 250 x 10.0 mm. Mobile phase: H<sub>2</sub>O (0.1 % TFA) and ACN (0.1 % TFA), flow 16 mL/min.

### 1.3.3. Analytical HPLC

HPLC chromatograms were recorded on a Waters Alliance 2695 separation module equipped with a 2996 photodiode array detector (PDA) and a Sunfire C18 column (100 x 4.6 mm x 5 μm, 100 Å, Waters), and Empower software. Flow rate: 1.6 mL/min, mobile phase: H<sub>2</sub>O (0.1 % TFA) and ACN (0.1 % TFA). Detection was performed at 220 nm.

### 1.3.4. Analytical HPLC-MS

HPLC-MS chromatograms were recorded on a Waters Alliance 2796 separation module system equipped with a Waters 2996 photodiode array detector, quadruple 3100 Mass Detector and a Sunfire C18 column (2.1 x 100 mm x 3.5 μm, 100 Å, Waters), and Masslynx software. Flow rate: 0.3 ml/min, mobile phase: H<sub>2</sub>O (0.1 % Formic Acid) and ACN (0.1 % Formic Acid).

### 1.3.5. Analytical UPLC-MS

UPLC-MS chromatograms were recorded using a UPLC ACQUITY H coupled to ACQUITY UPLC PhotodiodeArray Detector, ESI-MS micromass ZQ (Waters) and Masslynx software (Waters) using a Sunfire C18 column (100 mm x 2.1 mm, 3.5 μm, Waters); flow rate = 0.3 mL/min, solvents H<sub>2</sub>O (0.1 % formic acid) and ACN (0.07 % formic acid).

### 1.3.6. MALDI-TOF

Mass spectra were obtained using a MALDI-TOF Applied Biosystem 4700 with a N<sub>2</sub> laser of 337 nm.

Matrix: 20 mg/mL ACH in 50 % / 49.9 % H<sub>2</sub>O / 0.1 % TFA

Sample preparation: a mix of peptide solution (0.5 μL) and matrix (0.5 μL) was placed on a MALDI-TOF plate and dried by air.

## **1.4. Physicochemical properties experiments**

### 1.4.1. Solubility in H<sub>2</sub>O or aqueous buffers

The solubility of synthesized compounds was evaluated in H<sub>2</sub>O with different percentages of solubilizing excipients (DMSO, methanol or Tween20) or in the aqueous buffer that was later implemented in the corresponding *in vitro* assays.

First, 3 triplicates of the compounds under study were dissolved at high concentration, typically 1-10 mM in aqueous medium, vortexed for 2 min at maximum speed and sonicated for 60 min. Then, samples were constantly agitated for 48 h to let them reach the solubility equilibrium. Once in equilibrium, samples were centrifuged for 15 min at 15000 rpm at 25°C. Next, the supernatant was filtered by using Millipore filters (0.45 μM pVDF hydrophilic).

As control, 3 triplicates of each compound were diluted to 1-10 mM (or to the expected final concentration of samples in aqueous solution) with ACN/H<sub>2</sub>O (1:1).

Finally, controls and samples were injected in the HPLC for peak area quantification. For each one, the average peak area corresponding to the peptide under study was integrated and compared with the average area of the control peaks.

#### 1.4.2. Solubility in Cyclodextrin buffer

The vehicle preparation was performed following the manufacturer guidelines. A pH 7.4 phosphate buffer was prepared by adding 3.8 g of  $\text{Na}_2\text{HPO}_4 \cdot 12 \text{H}_2\text{O}$  + 3.8 g of  $\text{NaH}_2\text{PO}_4$  in 100 mL of  $\text{H}_2\text{O}$ . The buffer was magnetically stirred until completed dissolution of salts.

The 30 % HP $\beta$ -cyclodextrin buffer was obtained when 30 g of HP $\beta$ -cyclodextrin were diluted in 100 mL of the previously described pH 7.4 phosphate buffer. Solution was magnetically stirred until completed dissolution. Finally, the solution was sterilized by filtration using a Millipore filter (0.45  $\mu\text{M}$  PVDF hydrophilic).

In order to adjust the volumes to the experimental needs, this protocol was scaled down 10-folds by diluting 3 g HP $\beta$ -Cyclodextrin in 10 mL of the pH 7.4 phosphate buffer. The same protocol was followed when 5 % and 15 % HP $\beta$ -Cyclodextrin solutions were prepared by diluting 0.5 g and 1.5 g instead of 3 g of HP $\beta$ -Cyclodextrin.

Once the desired HP $\beta$ -Cyclodextrin solution was prepared, compound's solubility was evaluated following the same procedure described in the *Solubility in H<sub>2</sub>O or aqueous buffers* section (1.4.1).

#### 1.4.3. Solubility in cell culture medium

Compounds solubility was studied in cell culture medium at the desired DMSO percentage. This percentage corresponds to the *in vitro* experiments final concentration of compound under study, when this one was incubated with cells. In order to extrapolate the results, the same dilution protocol applied in cell experiments was also implemented to evaluate compound solubility.

The peptide solubility refers to the unbound peptide present in the cell culture medium, as the peptide could interact and get stick on medium proteins.

Hence, 3 triplicates of peptide stock 10 mM solution in DMSO were diluted in DMEM at 1-5 % of DMSO (depending on the *in vitro* experiments) to the same final concentration used in the cell experiment, typically 20-60  $\mu\text{M}$ . The mixture vortexed for 2 min at maximum speed, sonicated for 5 min and let to reach the equilibrium (48h) at 37 °C under constant agitation. Afterwards, the samples were centrifuged at 15000 rpm for 5 min at 25 °C. Next, the supernatant (100  $\mu\text{L}$ ) was diluted with cold acetonitrile (200  $\mu\text{L}$ ) and centrifuged again at 15000 rpm at 25 °C for 5 min. The final content of the compound in the supernatant was analyzed by HPLC.

As control, the compound was diluted to the stock solution in DMSO concentration. Later, same dilutions applied in DMEM are also performed to control sample in triplicate with ACN/H<sub>2</sub>O 1:1 instead of DMEM. In this case, no centrifugation is required for controls as the peptide is fully soluble in ACN/H<sub>2</sub>O and then, there is no presence of precipitates. Finally, the triplicates were injected in the HPLC.

The average peak area corresponding to the peptide under study in cell culture medium was integrated and compared with the average area of the control peaks.

#### 1.4.4. Molar coefficient extinction evaluation ( $\epsilon$ )

To measure  $\epsilon$  at  $\lambda = 220$  nm, compounds of interest were diluted in ACN/H<sub>2</sub>O (1:1). Calibration curves at 100, 50, 20 and 10  $\mu$ M were determined for each compound by using triplicates for each point. The molar extinction coefficient was calculated according to the Beer-Lambert Law<sup>171</sup>:

$$\epsilon = \frac{A}{C \cdot l}$$

Where;

A = Compound absorbance at 220 nm for a given concentration

C = Compound concentration in ACN/H<sub>2</sub>O

l = Distance that light travels through the sample solution

### **1.5. Biophysical properties evaluation**

#### 1.5.1. Stability in rat plasma

Rat plasma was allowed to thaw for 1 h in ice bath and afterwards mixed 1:1 with a PBS solution, pH= 7.4. Prior to start the experiment, the rat plasma/PBS mixture was warmed to 37 °C for 5 min.

Next, the compounds under study and Benfluorex (positive control which is totally degraded in 30 min) were diluted to 20  $\mu$ M with the previously warmed mixture of rat plasma/PBS with a final percentage of 5 % DMSO. Samples, two duplicates for each compound, were agitated at 40 rpm at 37 °C for 2 h. At selected time points, 100  $\mu$ L of the sample were harvested and mixed with 300  $\mu$ L of cold ACN, to precipitate plasma protein and these precipitated samples were cooled down for 30 min in an ice bath. After that, precipitated samples were centrifuged at 15000 rpm for 15 min at 4 °C. Next, the content of compound in the supernatant (100  $\mu$ L) was analyzed by HPLC. The ratio between the HPLC peak area of the compound at time zero and the area at a given time is proportional to the compound degradation.

For Benfluorex time points were at 0, 5, 15 and 30 min, as its stability in Rat plasma is very low. For the peptidomimetics under study, time points were taken between 0 and 120 min depending of each compound. For those compounds that was expected a high stability the number of time points were spaced, whereas expect low stable compounds were evaluated in a reduced time window. Rat plasma activity has been reported to be inactive after 2 h of incubation time, for this reason no points were taken after that time.

### 1.5.2. Stability in Human serum

Human serum was allowed to thaw for 1 hour in ice bath; afterwards it was mixed 9:1 with HBSS buffer solution.

Next, compounds of interest and Acyl Carrier protein, which is a positive control that is totally degraded in serum proteases in 30 min at the experimental conditions, also known as Acp-peptide (H-Val-Gln-Ala-Ala-Ile-Asp-Tyr-Ile-Asn-Glu-OH), were diluted to a final concentration of 150  $\mu\text{M}$  and incubated at 37°C under agitation (40 rpm). At several time points, 50  $\mu\text{L}$  of each sample were harvested and mixed with 200  $\mu\text{L}$  of cold MeOH (4°C) to precipitate serum proteins. The precipitated mixture was kept cold for 30 min in an ice bath. After that, the samples were centrifuged at 15000 rpm for 1h at 4 °C. Later, the supernatant (100  $\mu\text{L}$ ) was injected in to the HPLC for the quantification of the compound peak area. The difference between the area at time 0 and the area at a given time was used for the calculation of the compound degradation.

For Acp-peptide time points were at 0, 5, 15 and 30 min, as its stability in human serum is very low. For peptides under study, time points were between 0 and 1400 min depending of each compound, exponentially spacing the time points for those compounds expected to have a high stability.

### 1.5.3. Parallel Artificial Permeability Assay (PAMPA)

PAMPA was used to determine the capacity of compounds to cross biological barriers. The effective permeability ( $Pe$ ) of the compounds was measured at an initial concentration of 50  $\mu\text{M}$  in PAMPA system solution buffer with 5 % DMSO. The buffer solution was prepared from a standard one, commercialized by pION, and following manufacturer's instructions. Then, 5 mL of pION buffer were diluted in 19,5 ml of H<sub>2</sub>O at pH= 7.4. Compounds of interest and propranolol (a positive control which is a permeable by passive diffusion small molecule) were dissolved in buffer solution to the final concentration. PAMPA is composed for two plates; donor, which is filled with 195  $\mu\text{L}$  of diluted samples, and acceptor, where a mixture of phospholipids (20 mg/mL diluted to 1 % in dodecane) are poured over a polymeric filter. After that, in the same compartment 195  $\mu\text{L}$  of PAMPA system solution containing 5 % DMSO are added. Finally, both plates are assembled, forming a sandwich and ensuring that the underside of the membrane is in contact with the buffer of the donor plates. Incubation is performed in a saturated humidity atmosphere (Gut-BOX®, PiON) for 16 h under magnetic agitation (UWL 25  $\mu\text{m}$ ) at room temperature. After incubation, the content of the acceptor and donor wells, as wells as time zero samples are analyzed by HPLC or UPLC/MS.

Note. The phospholipid mixture used was a GIT-0 Lipid Solution (PIN: 110669) from pION.

## 1.6. Cell-based experiments

### 1.6.1. Cell drug internalization assay

The peptidomimetics of interest were incubated by triplicate with SH-S5YS, Talin -/- (Mouse Embryonic Fibroblast Talin 1 -/-, which overexpresses Talin-2) and/or Panc-1 (pancreatic cancer cell line) cells at 60  $\mu\text{M}$  concentration with 5 % DMSO for 2 h (time used in the in vitro efficacy studies). Cells were previously incubated 24 h until reach 100 % of confluence in a 96-well/plate, in these confluence cells is where the experiment was carried out. After incubation, cells were washed 5 times with PBS. Then, cells were trypsinized and placed into 0.5 mL Eppendorf with 100  $\mu\text{L}$  of PBS. Afterwards, 3 cleaning steps more with PBS were carried out. Next, cells were lysed using a 20 mM Tris, 100 mM NaCl and 1 mM Ethylenediaminetetraacetic acid (EDTA 0.5 % Triton X-100) buffer for 2 min at 37 °C and 5 %  $\text{CO}_2$ .

The mixture was centrifuged at 15000 rpm for 15 min at 25 °C. The supernatant was lyophilized overnight and resuspended by adding 100  $\mu\text{l}$  of ACN/H<sub>2</sub>O 1:1.

After that, the mixture was injected in UPLC/MS. The concentration of each peptidomimetic was calculated from its corresponding calibration curve. Then, for each compound a calibration curve was obtained by injecting in UPLC/MS a compound standard solution at 1, 5, 10 and 20  $\mu\text{M}$  in ACN/H<sub>2</sub>O.

The percentage of internalization was calculated by comparing the amount of compound after the 2 h incubation and washes and the initial amount of compound added to the cells (time 0).

### 1.6.2. Cell proliferation assay

*In vitro* efficacy of synthesized compounds for the Talin-Vinculin project (Chapter 3) were evaluated in-house as well, through a cell proliferation assay in neuroblastoma derived cell line, SH-SY5Y.

Cells were seeded onto 96-well plates (10.000 cells/well to reach 80-90 % of confluence) and incubated for 24 h in a humidified atmosphere of 37°C with 5 %  $\text{CO}_2$ . After this time, cell culture (DMEM) was replaced with 200  $\mu\text{L}$  of fresh DMEM at 5 % of DMSO containing 60  $\mu\text{M}$  of the respective compound under study and incubated again for 1 hour.

The volume per well was 200  $\mu\text{L}$  with a final DMSO concentration of 5 %. In all assays two more conditions were tested as controls: positive control (cells treated with 5 % DMSO), and negative control (0.5 % Triton X-100 in DMEM with 5 % DMSO). Cell viability was measured using 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT). After the correspondent incubation time, cell were washed with PBS (dir N°) and a 180  $\mu\text{L}$  of DMEM and 20  $\mu\text{L}$  of MTT solution (5 mg/mL in phosphate buffered saline (PBS)) were added to the cells, which were kept in the incubator for an additional hour.. Then, cells were washed with PBS three times and 150  $\mu\text{L}$  dimethyl sulfoxide (DMSO) was add to each well of the plate. Cells were incubated for 30 minutes under constant orbital agitation (90 r.p.m.) in order to solubilize the formazan crystals formed. Absorbance intensity was read at a wavelength of 595 nm in Synergy HTX Absorbance



microplate reader. The difference between the absorbance of not treated cells (considered as 100 % cell viability) with those incubated with the compounds allows obtaining the percentage of cell proliferation.

## References



1. Kinch, M. S., Haynesworth, A., Kinch, S. L. & Hoyer, D. An overview of FDA-approved new molecular entities: 1827–2013. *Drug Discov. Today* 19, 1033–1039 (2014).
2. Schmid, E. F., Smith, D. A. & Smith, D. A. Is declining innovation in the pharmaceutical industry a myth? *Drug Discv Today*. 10, 1031–1039 (2005).
3. Mullard, A. 2018 FDA drug approvals. *Nat. Rev. Drug Discov.* 18, 85–89 (2019).
4. <https://cen.acs.org/pharmaceuticals/drug-development/FDA-approved-record-number-drugs/97/web/2019/01>.
5. Bakail, M. & Ochsenein, F. Comptes Rendus Chimie Targeting protein e protein interactions , a wide open fi eld for drug design Inhiber les interactions prot eine e prot eine , un large champ ouvert pour le d eveloppement de nouveaux m edicaments. *Comptes rendus - Chim.* 19, 19–27 (2016).
6. Reth, M. C O M M E N T A R Y Matching cellular dimensions with molecular sizes. *Nat. Publ. Gr.* 14, 765–767 (2013).
7. Nelson, A. L., Dhimolea, E. & Reichert, J. M. Development trends for human monoclonal antibody therapeutics. *Nat Rev Drug Discov* 9, 767–74 (2010).
8. Chang, H. & Sun, Z. Increasing stability of antibody via antibody engineering: Stability engineering on an anti-hVEGF. 2620–2630 (2014).
9. Spires, T. L., Johannes, J., Dietmar, A. & Thal, R. Interactions of pathological proteins in neurodegenerative diseases. *Acta Neuropathol.* 134, 187–205 (2017).
10. Chang, A. N. *et al.* Structural and protein interaction effects of hypertrophic and dilated cardiomyopathic mutations in alpha-tropomyosin. *Front Physiol.* 5, 1–9 (2014).
11. London, N., Movshovitz-attias, D. & Schueler-furman, O. The Structural Basis of Peptide-Protein Binding Strategies. *Struct. Des.* 18, 188–199 (2010).
12. Henninot, A., Collins, J. C. & Nuss, J. M. The Current State of Peptide Drug Discovery: Back to the Future? *J. Med. Chem.* acs.jmedchem.7b00318 (2017).
13. A. D. McNaught and A. Wilkinson. *IUPAC. Compendium of Chemical Terminology, 2nd ed. (the 'Gold Book').* (1997).
14. Sheffler, W. & Baker, D. RosettaHoles : Rapid assessment of protein core packing for structure prediction , refinement , design , and validation. *Protein Sci.* 18, 229–239 (2009).
15. Gaynes, R. The Discovery of Penicillin — New Insights After More Than 75 Years of Clinical Use. *Emergin Infect. Dis.* 23, 849–853 (2017).

16. Lau, J. L. & Dunn, M. K. Therapeutic peptides : Historical perspectives , current development trends , and future directions. *Bioorg. Med. Chem.* 26, 2700–2707 (2018).
17. Werle, M. Strategies to improve plasma half life time of peptide and protein drugs. *Amino Acids.* 351–367 (2006).
18. Strömstedt, A. A., Pasupuleti, M., Schmidtchen, A. & Malmsten, M. Evaluation of Strategies for Improving Proteolytic Resistance of Antimicrobial Peptides by Using Variants of EFK17, an Internal Segment of LL-37. *Antimicrob. Agents Chemother.* 53, 593 LP-602 (2009).
19. Böttger, R., Hoffmann, R. & Knappe, D. Differential stability of therapeutic peptides with different proteolytic cleavage sites in blood, plasma and serum. *PLoS One* 12, e0178943 (2017).
20. Darwich, A. S., Neuhoff, S. & Jamei, M. Interplay of Metabolism and Transport in Determining Oral Drug Absorption and Gut Wall Metabolism : A Simulation Assessment Using the ‘ Advanced Dissolution , Absorption , Metabolism ( ADAM )’ Model. 450, 716–729 (2010).
21. Greenwald, R. B. PEG drugs : an overview. *J. Control. Release* 74, 159–171 (2001).
22. Grassl, S. M. 16 - Mechanisms of Carrier-Mediated Transport: Facilitated Diffusion, Cotransport, and Countertransport. in (ed. Sperelakis, N. B. T.-C. P. S. B. (Third E.)) 249–259 (Academic Press, 2001).
23. Hervé, F., Ghinea, N. & Scherrmann, J.-M. CNS delivery via adsorptive transcytosis. *AAPS J.* 10, 455–472 (2008).
24. Raucher, D. & Ryu, J. S. Cell-penetrating peptides: Strategies for anticancer treatment. *Trends in Molecular Medicine* 21, 560–570 (2015).
25. Guidotti, G., Brambilla, L. & Rossi, D. Cell-Penetrating Peptides: From Basic Research to Clinics. *Trends in Pharmacological Sciences* 38, 406–424 (2017).
26. Franc, I., Lipinski, A. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev.* 23, 3-26 (1997).
27. Ñ, C. A. L. Lead profiling Lead- and drug-like compounds : the rule-of-five revolution. *Drug Discov Today Technol*, 1, 337–341 (2004).
28. Laupacis, A., Keown, P. A., Ulan, R. A., Mckenzie, N. & Stiller, C. R. Therapeutic Review Cyclosporin A : a powerful immunosuppressant. *Canc Med Assoc J.* 126, 1041–1046 (1982).
29. Tyagi, M. *et al.* Toward the Design of Molecular Chameleons : Flexible Shielding of. *Org. Lett.* 21, 5737-5742 (2018).

30. Whitty, A. *et al.* Quantifying the Chameleonic Properties of Macrocycles and other High Molecular Weight Drugs. *Drug Discov Today*. 21, 712–717 (2018).
31. Uhlig, T. *et al.* The emergence of peptides in the pharmaceutical business : From exploration to exploitation. *EUPROT* 4, 58–69 (2014).
32. Oesterlin, C. Solid Phase Peptide Synthesis . I . Tetrapeptide1. *J Am Chem Soc*. 85, 2149-2154 (1963).
33. Erak, M., Bellmann-sickert, K., Els-heindl, S. & Beck-sickinger, A. G. Bioorganic & Medicinal Chemistry Peptide chemistry toolbox – Transforming natural peptides into peptide therapeutics. *Bioorg Med Chem*. 26, 2759–2765 (2018).
34. Pan, D. *et al.* A general strategy for developing cell-permeable photo-modulatable organic fluorescent probes for live-cell super-resolution imaging. *Nat. Commun*. 5, 1–8 (2014).
35. Qiu, X. *et al.* Single-Cell Resolution Imaging of Retinal Ganglion Cell Apoptosis In Vivo Using a Cell-Penetrating Caspase- Activatable Peptide Probe. *PLOS ONE*. 9, 1–10 (2014).
36. Pasqualini, R., Ruoslathi, E. organ targeting in vivo using phage display peptide libraries. *Nature*. 380, 364-366 (1996).
37. Passioura, T. & Suga, H. A RaPID way to discover nonstandard macrocyclic peptide modulators of drug targets. *Chem Commun*. 53, (2017).
38. Daniel R., Stroik et al. Targeting protein-protein interactions for therapeutic discovery via FRET-based high-throughput screening in living cells. *Sci Reports*.12560 (2018)
39. Di, L. Strategic Approaches to Optimizing Peptide ADME Properties. *AAPS J*. 17, 134–143 (2015).
40. Acids, N. A., Bonds, P., Gentilucci, L., Marco, R. De & Cerisoli, L. Chemical Modifications Designed to Improve Peptide Stability : Incorporation of Non-Natural Amino Acids, Pseudo-Peptide bonds, and Cyclization. *Cur. Pharm. Design*. 28, 3185–3203 (2010).
41. Yin, N., Brimble, M. A., Harris, P. W. R. & Wen, J. Enhancing the Oral Bioavailability of Peptide Drugs by using Chemical Modification and Other Approaches. *Med. Chem. (Los. Angeles)*. 4, 763–769 (2014).
42. Article, E. *et al.* Stapled peptides as a new technology to investigate protein–protein interactions in human platelets. *Chem. Sci*. 9, 4638–4643 (2018).
43. Checco, J. W. *et al.*  $\alpha/\beta$ -Peptide Foldamers Targeting Intracellular Protein-Protein Interactions with Activity in Living Cells. *J Am Chem Soc*. 137, 11365–11375 (2016).
44. Liu, T. *et al.* Enhancing protein stability with extended disulfide bonds. *Proc Natl Acad Sci*. 113, 5910-5 (2016).

45. Biron, E. *et al.* Improving oral bioavailability of peptides by multiple N-methylation: Somatostatin analogues. *Angew. Chemie - Int. Ed.* 47, 2595–2599 (2008).
46. Dong, Q. G. *et al.* Improvement of enzymatic stability and intestinal permeability of deuterohemin-peptide conjugates by specific multi-site N-methylation. *Amino Acids* 43, 2431–2441 (2012).
47. Chatterjee, J., Rechenmacher, F. & Kessler, H. N-Methylation of peptides and proteins: An important element for modulating biological functions. *Angewandte Chemie - International Edition* 52, 254–269 (2013).
48. Linde, Y. *et al.* Structure-Activity Relationship and Metabolic Stability Studies of Backbone Cyclization and N-Methylation of Melanocortin Peptides. *Biopolymers* 90, 671–682 (2008).
49. Bumpust, F. M. Amino acid side chain conformation in angiotensin II and analogs : Correlated results of circular dichroism and <sup>1</sup>H nuclear magnetic resonance Biochemistry : Correlated results of circular dichroism and <sup>1</sup>H nuclear magnetic resonance. *Proc Natl Acad Sci.* 77, 82–86 (1980).
50. Laufer, B. *et al.* The Impact of Amino Acid Side Chain Mutations in Conformational Design of Peptides and Proteins. *Chemistry.* 5385–5390 (2010).
51. Miller, S. C. & Scanlan, T. S. Site-selective N-methylation of peptides on solid support. *J. Am. Chem. Soc.* 119, 2301–2302 (1997).
52. Olsen, C. A., Witt, M., Hansen, S. H., Jaroszewski, W. & Franzyk, H. Fukuyama – Mitsunobu alkylation in amine synthesis on solid phase revisited : N-alkylation with secondary alcohols and synthesis of curtatoxins. *Tetrahedron.* 61, 6046–6055 (2005).
53. Yan, T., Feringa, B. L. & Barta, K. Direct N-alkylation of unprotected amino acids with alcohols. *Sci Advances.* 3, 1–8 (2017).
54. Blout, E. R. Cyclic peptides: Past, present, and future. *Biopolymers* 20, 1901–1912 (1981).
55. Dougherty, P. G., Qian, Z. & Pei, D. Macrocycles as protein–protein interaction inhibitors. *Biochem. J.* 474, 1109–1125 (2017).
56. Villar, E. A. *et al.* How proteins bind macrocycles. *Nat Chem Biol.* 10, (2014).
57. Swaminathan. P and Francine F. Foss. Romidepsin for the Treatment of Peripheral T-Cell Lymphoma. *Oncologist.* 20, 1084–1091 (2015).
58. Scorr, J. K. & Smith, G. P. Searching for Peptide Ligands Epitope Library with an Epitope Library. *Science.* 416, 386-390 (1990).
59. Heinis, C., Rutherford, T., Freund, S. & Winter, G. Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat Chem Biol.* 5, 502–507 (2009).

60. Goto, Y., Katoh, T. & Suga, H. Flexizymes for genetic code reprogramming. *Nat Protocol.* 6, 779–790 (2011).
61. Mclean, J. A. & Lybrand, T. P. Distance Geometry Protocol to Generate Conformations of Natural Products to Structurally Interpret Ion Mobility-Mass Spectrometry Collision Cross Sections. *J. Physical Chem. B* 13812–13820 (2014).
62. Kamenik, A. S., Lessel, U., Fuchs, J. E., Fox, T. & Liedl, K. R. Peptidic Macrocycles - Conformational Sampling and Thermodynamic Characterization. *J. Chem. Inf. Model.* 58, 982–992 (2018).
63. Salomon-ferrer, R., Case, D. A. & Walker, R. C. An overview of the Amber biomolecular simulation package. *Comput. Mol. Sci.* 3, 198–210 (2013).
64. Boyle, N. M. O. *et al.* Open Babel : An open chemical toolbox. *J Cheminform.* 1–14 (2011).
65. Halgren, T. A. Identifying and Characterizing Binding Sites and Assessing Druggability. *Inf. Model.* 377–389 (2009).
66. Kozakov, D. *et al.* The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nat. Protoc.* 10, 733–755 (2015).
67. Guvench, O. & Mackerell, A. D. Computational Fragment-Based Binding Site Identification by Ligand Competitive Saturation. *PLoS Computational Biology.* 5, (2009).
68. Seco, J., Luque, F. J. & Barril, X. Binding Site Detection and Druggability Index from First Principles. *J Med Chem.* 52, 2363–2371 (2009).
69. Fichthorn, K. A., Weinberg, W. H. & Fichthorn, K. A. Theoretical foundations of dynamical Monte Carlo simulations Theoretical foundations of dynamical Monte Carlo simulations. *J. Chem. Phys.* 95, 1090–1096 (1991).
70. Neves, Marco A. C., Totrov, Maxim, Abagyan, R. Docking and scoring with ICM: the benchmarking results and strategies for improvement. *J Med Chem.* 26, 675–686 (2012).
71. Jones, G. *et al.* Development and Validation of a Genetic Algorithm for Flexible Docking. *J Mol Biol.* 267, 727-748 (1997).
72. Friesner, R. A. *et al.* Glide : A New Approach for Rapid , Accurate Docking and Scoring . 1 . Method and Assessment of Docking Accuracy. *J Med Chem.* 47, 1739–1749 (2004).
73. Vilar, S., Cozza, G. & Moro, S. Medicinal Chemistry and the Molecular Operating Environment ( MOE ): Application of QSAR and Molecular Docking to Drug Discovery. *Curr Topics Med Chem.* 8, 1555–1572 (2008).



74. Ruiz-carmona, S. *et al.* rDock : A Fast , Versatile and Open Source Program for Docking Ligands to Proteins and Nucleic Acids. *PLoS Comput Biol.* 10, 1–7 (2014).
75. Olson, O. T. and A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J Comput Chem.* 31, 455–461 (2011).
76. Rentsch, R. & Renard, B. Y. Docking small peptides remains a great challenge : an assessment using AutoDock Vina. *Brief Bioinform.* 16, 1045–1056 (2015).
77. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of Polypeptide Chain Configurations. *J. Mol. Biol.* 7, 95–99 (1963).
78. Hauser, A. S. & Windshügel, B. LEADS-PEP: A Benchmark Data Set for Assessment of Peptide Docking Performance. *J. Chem. Inf. Model.* 56, 188–200 (2016).
79. Koes, D. R., Baumgartner, M. P. & Camacho, C. J. Lessons Learned in Empirical Scoring with smina from the CSAR 2011 Benchmarking Exercise. *J Chem Inf Model.* 53, 1893–1904 (2014).
80. Wang, R. & Wang, S. How Does Consensus Scoring Work for Virtual Library Screening? An Idealized Computer Experiment. *J. Chem. Inf. Comput. Sci.* 41, 1422–1426 (2001).
81. Prasanna, S. & Doerksen, R. J. Topological Polar Surface Area : A Useful Descriptor in 2D-QSAR. *Curr Med Chem.* 1848, 21–41 (2009).
82. Over, B. *et al.* Structural and conformational determinants of macrocycle cell permeability. *Nat Chem Biol.* 12, 1065–1074 (2016).
83. Ertl, P., Rohde, B. & Selzer, P. Fast Calculation of Molecular Polar Surface Area as a Sum of Fragment-Based Contributions and Its Application to the Prediction of Drug Transport Properties. *J Med Chem.* 3714–3717 (2000).
84. Rezai, T., Yu, B., Millhauser, G. L., Jacobson, M. P. & Lokey, R. S. Testing the conformational hypothesis of passive membrane permeability using synthetic cyclic peptide diastereomers. *J. Am. Chem. Soc.* 128, 2510–2511 (2006).
85. Palm, K. *et al.* Evaluation of Dynamic Polar Molecular Surface Area as Predictor of Drug Absorption : Comparison with Other Computational and Experimental Predictors. *J Med Chem.* 5382–5392 (1998).
86. Bergström, C. A. S., Charman, W. N. & Porter, C. J. H. Computational prediction of formulation strategies for beyond-rule-of-5 compounds. *Adv. Drug Deliv. Rev.* 101, 6–21 (2016).
87. Palomo, J. M. Solid-phase peptide synthesis : an overview focused on the preparation of biologically relevant peptides. *RSC Adv.* 4, 32658–32672 (2014).

88. Vojkovsky, T. Detection of secondary amines on solid phase. *Pept. Res.* 8, 236–237 (1995).
89. Kalepu, S. & Nekkanti, V. Insoluble drug delivery strategies : review of recent advances and business prospects. *Acta Pharm. Sin. B* 5, 442–453 (2015).
90. Jansook, P., Ogawa, N. & Loftsson, T. Cyclodextrins : structure , physicochemical properties and pharmaceutical applications. *Int. J. Pharm.* 535, 272–284 (2018).
91. Williams, H. D., Trevaskis, N. L., Charman, S. A., Shanker, R. M. & Charman, W. N. Strategies to Address Low Drug Solubility in Discovery and Development. *Pharmacol Rev.* 65, 315–499 (2013).
92. Gould, S. & Scott, R. C. 2-Hydroxypropyl- $\beta$ -cyclodextrin (HP- $\beta$ -CD): A toxicology review. *Food Chem Toxicol.* 43, 1451–1459 (2005).
93. Brewster, M. E. & Loftsson, T. Cyclodextrins as pharmaceutical solubilizers. *Adv Drug Deliv Rev.* 59, 645–666 (2007).
94. Singh, S. S. Preclinical Pharmacokinetics : An Approach Towards Safer and Efficacious Drugs. *Curr Drug Metab.* 7, 165–182 (2006).
95. Zhang, D., Luo, G., Ding, X. & Lu, C. Preclinical experimental models of drug metabolism and disposition in drug discovery and development. *Acta Pharmaceutica Sinica B.* 2, 549–561 (2012).
96. Di, L., Kerns, E. H., Hong, Y. & Chen, H. Development and application of high throughput plasma stability assay for drug discovery. *Int. J. Pharm.* 297, 110–119 (2005).
97. Avdeef, A. *et al.* PAMPA — Critical Factors for Better Predictions of Absorption. *J Pharm Sci.* 96, 2893–2909 (2007).
98. Elosegui-artola, A., Trepate, X. & Roca-cusachs, P. Control of Mechanotransduction by Molecular Clutch Dynamics. *Trends Cell Biol.* 28, 356–367 (2018).
99. Yan, J., Yao, M., Gault, B. T. & Sheetz, M. P. Talin Dependent Mechanosensitivity of Cell Focal Adhesions. *Cellular and Molecular Bioengineering* 8, 151–159 (2015).
100. Ringer, P. Weibi, A. Cost, Al. Freikamp, A. Sabass, B. Mehlich, A. Tramier, M. Rief, M and Grashoff, C. Multiplexing molecular tension sensors reveals piconewton force gradient across talin-1. *Nat Methods.* 14, 1090-1096 (2017).
101. Elosegui-Artola, A. *et al.* Mechanical regulation of a molecular clutch defines force transmission and transduction in response to matrix rigidity. *Nat. Cell Biol.* 18, 540–548 (2016).
102. Klapholz, B. & Brown, N. H. Talin – the master of integrin adhesions. *J. Cell Sci.* jcs.190991 (2017).

103. Elosegui-Artola, A. *et al.* Force Triggers YAP Nuclear Entry by Regulating Transport across Nuclear Pores. *Cell*. 171, 1397-1410 (2017).
104. Moroishi, T., Hansen, C. G. & Guan, K.-L. The emerging roles of YAP and TAZ in cancer. *Nat. Rev. Cancer* 15, 73–79 (2015).
105. Zhao, B. *et al.* Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control. *Genes Dev.* 21, 2747–2761 (2007).
106. Goult, B. T. *et al.* RIAM and Vinculin Binding to Talin Are Mutually Exclusive and Regulate Adhesion Assembly and Turnover. *J Biol Chem.* 288, 8238–8249 (2013).
107. Baxter, N. J., Zacharchenko, T., Barsukov, I. L. & Williamson, M. P. Pressure-Dependent Chemical Shifts in the R3 Domain of Talin Show that It Is Thermodynamically Poised for Binding to Either Vinculin or RIAM. *Structure*. 25, 1856–1866.e2 (2017).
108. Yao, M. Benjamin, T. Goult. Hu Chen. Peiwen, C. Sheetz, M. and Yan, J. Mechanical activation of vinculin binding to talin locks talin in an unfolded conformation. *Sci. Rep.* 4, (2014).
109. Cell scientist to watch – Pere Roca-Cusachs. *J. Cell Sci.* 131, jcs222596 (2018).
110. Klein, M. A. Stabilized helical peptides: a strategy to target protein-protein interactions. *ACS Med. Chem. Lett.* 5, 838–839 (2014).
111. Kovalevich, J. & Langford, D. Considerations for the Use of SH-SY5Y Neuroblastoma Cells in Neurobiology BT - Neuronal Cell Culture: *Methods and Protocols*. in (eds. Amini, S. & White, M. K.) 9–21 (Humana Press, 2013).
112. Ackerman, S. & Horton, W. Chapter 2.4 - Effects of Environmental Factors on DNA: Damage and Mutations. in *Green Chemistry* (eds. Török, B. & Dransfield, T. B. T.-G. C.) 109–128 (Elsevier, 2018).
113. Heitzer, E. & Tomlinson, I. Replicative DNA polymerase mutations in cancer. *Curr. Opin. Genet. Dev.* 24, 107–113 (2014).
114. Aguilera, A. & Gómez-gonzález, B. Genome instability : a mechanistic view of its causes and consequences. *Nat. Rev. Genet.* 9, (2008).
115. Hoeijmakers, J. H. J. Genome maintenance mechanisms for preventing cancer. *Insight Rev. Artic.* 411, 366–374 (2001).
116. Shailani, A., Kaur, R. P. & Munshi, A. A comprehensive analysis of BRCA2 gene: focus on mechanistic aspects of its functions, spectrum of deleterious mutations, and therapeutic strategies targeting BRCA2-deficient tumors. *Med. Oncol.* 35, 1–10 (2018).

117. Lord, C. J., Tutt, A. N. J. & Ashworth, A. Synthetic Lethality and Cancer Therapy: Lessons Learned from the Development of PARP Inhibitors. *Annu. Rev. Med.* 66, 455–470 (2015).
118. Ohmoto, A. & Yachida, S. Current status of poly(ADP-ribose) polymerase inhibitors and future directions. *Onco. Targets. Ther.* 10, 5195–5208 (2017).
119. Moyo, K. *et al.* Small molecule inhibitors targeting DNA repair and DNA repair deficiency in research and cancer therapy. *Cell Chem Biol.* 4, 1–19 (2016).
120. Nieborowska-Skorska, M. *et al.* Gene expression and mutation-guided synthetic lethality eradicates proliferating and quiescent leukemia cells. *J. Clin. Invest.* 127, 2392–2406 (2017).
121. Lord, C. J. & Ashworth, A. RAD51, BRCA2 and DNA repair: A partial resolution. *Nat. Struct. Mol. Biol.* 14, 461–462 (2007).
122. Prakash, R., Zhang, Y., Feng, W. & Jasin, M. Homologous Recombination and Human Health: The Roles of BRCA1, BRCA2, and Associated Proteins. *Perspect. Biol.* 1–29 (2015). 0
123. Cruz, C. *et al.* RAD51 foci as a functional biomarker of homologous recombination repair and PARP inhibitor resistance in germline BRCA-mutated breast cancer. *Ann. Oncol.* 29, 1203–1210 (2018).
124. Sullivan-reed, K. *et al.* Simultaneous Targeting of PARP1 and RAD52 Triggers Dual Synthetic Lethality in BRCA-Deficient Tumor Cells. *Cell Rep.* 23, 3127–3136 (2018).
125. Esashi, F., Galkin, V. E., Yu, X., Egelman, E. H. & West, S. C. Stabilization of RAD51 nucleoprotein filaments by the C-terminal region of BRCA2. *Nat. Struct. Mol. Biol.* 14, 468–474 (2007).
126. Roberti, M. *et al.* Rad51/Brca2 disruptors inhibit homologous recombination and synergize with olaparib in pancreatic cancer cells. *Eur. J. Med. Chem.* 165, 80–92 (2019).
127. Scott, D. E. *et al.* Small-molecule inhibitors that target protein-protein interactions in the RAD51 family of recombinases. *ChemMedChem* 10, 296–303 (2015).
128. Trenner, A., Godau, J. & Sartori, A. A. A Short BRCA2-Derived Cell-Penetrating Peptide Targets RAD51 Function and Confers Hypersensitivity toward PARP Inhibition. *Mol. Cancer Ther.* 17, 1392 LP-1404 (2018).
129. Franken, N. A. P., Rodermond, H. M., Stap, J., Haveman, J. & Bree, C. Van. Clonogenic assay of cells in vitro. *Nat. Protoc.* 1, 2315–2319 (2006).
130. Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. *Nature* 300, 149–152 (1982).

131. Cox, A. D., Fesik, S. W., Kimmelman, A. C., Luo, J. & Der, C. J. Drugging the undruggable RAS: Mission Possible? *Nat. Rev. Drug Discov.* 13, 828–851 (2014).
132. Stephen, A. G., Esposito, D., Bagni, R. G. & McCormick, F. Dragging ras back in the ring. *Cancer Cell* 25, 272–281 (2014).
133. Scheffzek, K. *et al.* The Ras-RasGAP Complex : Structural Basis for GTPase Its in Activation and Oncogenic Ras Mutants. *Adv. Sci.* 277, 333–338 (2010).
134. Maurer, T. *et al.* Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity. *Proc. Natl. Acad. Sci.* 109, 5299–5304 (2012).
135. Sun, Q. *et al.* Discovery of small molecules that bind to K-Ras and inhibit Sos-mediated activation. *Angew. Chemie - Int. Ed.* 51, 6140–6143 (2012).
136. Ahearn, I. M., Haigis, K., Bar-Sagi, D. & Philips, M. R. Regulating the regulator: Post-translational modification of RAS. *Nat. Rev. Mol. Cell Biol.* 13, 39–51 (2012).
137. Kohl, N. E. *et al.* Selective Inhibition of ras-Dependent Transformation by a Farnesyltransferase Inhibitor. *Science* (80-. ). 260, 1934–1937 (1993).
138. Lerner, E. C. *et al.* Ras C AAX Peptidomimetic FTI-277 Selectively Blocks Oncogenic Ras Signaling by Inducing Cytoplasmic Accumulation of Inactive Ras-Raf Complexes. *J Biol Chem.* 270, 26802–26806 (1995).
139. E. Van Cutsem, H. van de Velde, P. Karasek, H. Oettle, W.L. Vervenne, A. Szawlowski, P. Schoffski, S. Post, C. Verslype, H. Neumann, H. Safran, Y. Humblet, J. Perez Ruixo, Y. Ma, and D. V. H. Phase III Trial of Gemcitabine Plus Tipifarnib Compared With Gemcitabine Plus Placebo in Advanced Pancreatic Cancer. *J. Clin. Oncol.* 22, (2004).
140. S. Rao, D. Cunningham, A. de Gramont, W. Scheithauer, M. Smakal, Y. Humblet, G. Kourteva, T. Iveson, T. Andre, J. Dostalova, A. Illes, R. Belly, J.J. Perez-Ruixo, Y.C. Park, and P. A. P. Phase III Double-Blind Placebo-Controlled Study of Farnesyl Transferase Inhibitor R115777 in Patients With Refractory Advanced Colorectal Cancer. *J. Clin. Oncol.* 22, (2004).
141. Berndt, N., Hamilton, A. D. & Sebt, S. M. Targeting protein prenylation for cancer therapy. *Nat Rev Cancer.* 11, (2011).
142. Cox, A. D., Der, C. J. & Philips, M. R. Targeting RAS Membrane Association : Back to the Future for Anti-RAS Drug Discovery ? *Clin. Cancer Res.* 1819–1828 (2015).
143. Zeitouni, D., Pylayeva-gupta, Y., Der, C. J. & Bryant, K. L. KRAS Mutant Pancreatic Cancer : No Lone Path to an Effective Treatment. *Cancers.* 8, 1–22 (2016).
144. Bollag, G. *et al.* Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature* 467, 596–599 (2010).

145. Liszkay, G. *et al.* Combined Vemurafenib and Cobimetinib in BRAF-Mutated Melanoma James. *N Engl J Med.* 13, 1867–1876 (2014).
146. Sullivan, R. J. *et al.* First-in-Class ERK1 / 2 Inhibitor Ulixertinib ( BVD-523 ) in Patients with MAPK Mutant Advanced Solid Tumors : Results of a Phase I Dose-Escalation and Expansion Study. *Cancer Discov.* 8, (2018).
147. McCormick, F. Progress in targeting RAS with small molecule drugs. *Biochem J.* 21, 365–374 (2019).
148. Margolis, B. & Skolnik, E. Y. Activation of RAS by Receptor Tyrosine Kinases. *J. Am. Soc. Nephrol* 5, 1288–1299 (1994).
149. A. Robert MacLeod, J. R. and M. S. Regulation of DNA Methylation by the RAS Signaling Pathway. *J. Biol. Chem.* 270, 11327–11337 (1995).
150. Soman, G. *et al.* MTS dye based colorimetric CTLL-2 cell proliferation assay for product release and stability monitoring of Interleukin-15: Assay qualification, standardization and statistical analysis. *J. Immunol. Methods* 348, 83–94 (2009).
151. Bray, F., Ren, J., Masuyer, E. & Ferlay, J. Global estimates of cancer prevalence for 27 sites in the adult population in 2008. *Int. J. Pharm.* 132, 1133–1145 (2008).
152. Kumar, C. S., Dey, D., Ghosh, S. & Banerjee, M. Breach : Host Membrane Penetration and Entry by Nonenveloped Viruses. *Trends Microbiol.* 26, 525–537 (2018).
153. Harden, M. E. & Munger, K. Mutation Research / Reviews in Mutation Research Human papillomavirus molecular biology. *Mutat. Res. Mutat. Res.* 772, 3–12 (2017).
154. Day, P. M., Thompson, C. D., Schowalter, R. M., Lowy, D. R. & Schiller, J. T. Identification of a Role for the trans -Golgi Network in Human Papillomavirus 16 Pseudovirus Infection. *Cold Spring Harb Perspect Biol* 87, 3862–3870 (2013).
155. Lucas, M. & Hierro, A. Retromer. *Curr. Biol.* 687–689 (2017).
156. Burd, C. & Cullen, P. J. Retromer : A Master Conductor of Endosome Sorting. *Cold Spring Harb Perspect Biol* 6, 1–13 (2014).
157. Siddiq, A., Broniarczyk, J. & Banks, L. Papillomaviruses and Endocytic Trafficking. *Int. J. Mol. Sci.* 19, 1–24 (2018).
158. Popa, A., Zhang, W., Harrison, M. S., Goodner, K. & Kazakov, T. Direct Binding of Retromer to Human Papillomavirus Type 16 Minor Capsid Protein L2 Mediates Endosome Exit during Viral Infection. *PLoS Pathog* 11, 1–21 (2015).
159. Zhang, P. *et al.* Cell-Penetrating Peptide Mediates Intracellular Membrane Passage of Human Papillomavirus L2 Protein to Trigger Retrograde Trafficking Article Cell-Penetrating Peptide Mediates Intracellular Membrane Passage of Human Papillomavirus L2 Protein to Trigger Retrograde Trafficking. *Cell* 174, 1465–1476.e13 (2018).

160. Lucas, M. *et al.* Structural Mechanism for Cargo Recognition by the Retromer Complex Article Structural Mechanism for Cargo Recognition by the Retromer Complex. *Cell* 167, 1623–1635 (2016).
161. Pierce, M. M., Raman, C. S. & Nall, B. T. Isothermal Titration Calorimetry of Protein – Protein Interactions. *Methods* 19, 213–221 (1999).
162. Psyri, A. & DiMaio, D. Human papillomavirus in cervical and head-and-neck cancer. *Nat. Clin. Pract. Oncol.* 5, 24 (2008).
163. Roberti, M. *et al.* European Journal of Medicinal Chemistry Rad51 / BRCA2 disruptors inhibit homologous recombination and synergize with olaparib in pancreatic cancer cells. *Eur. J. Med. Chem.* 165, 80–92 (2019).
164. Kessler, D. *et al.* Drugging an undruggable pocket on KRAS. *Proc. Natl. Acad. Sci.* 116, 15823 LP-15829 (2019).
165. Morris, G. M., Huey, R. & Olson, A. J. Using AutoDock for Ligand-Receptor Docking. *Curr. Protoc. Bioinforma.* 24, 1–8 (2008).
166. Roe, D. R. & Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* 9, 3084–3095 (2013).
167. Barlos, K., Gatos, D. & Stavropoulos, G. 2-Chlorotriyl Chloride Resin — Studies on Anchoring of Fmoc-Amino Acids and Peptide Cleavage. *Int. J. Pept. Protein Res.* 37, 513–520 (1991).
168. García-Martín, F. *et al.* ChemMatrix, a Poly(ethylene glycol)-Based Support for the Solid-Phase Synthesis of Complex Peptides. *J. Comb. Chem.* 8, 213–220 (2006).
169. Kaiser, E., Colecott, R. L., Bossinger, C. D. & Cook, P. I. Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem.* 34, 595–598 (1970).
170. Miller, S. C., Scanlan, T. S. & Francisco, S. Site-Selective N-Methylation of Peptides on Solid Support Department of Pharmaceutical Chemistry University of California Conformational restriction of peptides can yield useful information about their bioactive conformation. 1 N -Methyl amino acid sub. 7863, 2301–2302 (1997).
171. Mayerhöfer, T. G. & Popp, J. Beer ' s Law – Why Absorbance Depends ( Almost ) Linearly on Concentration. *ChemPhysChem* 20, 511–515 (2019).

*Product characterization*





## 2.1. Talin-Vinculin inhibitors

Code	Formula:	Calculated mass:	Purity:	Gradient ACN/H <sub>2</sub> O:	Retention time (min):	Mass Identification:
2018/IP-15-01	C <sub>54</sub> H <sub>64</sub> N <sub>6</sub> O <sub>5</sub>	877.12 g/mol	>99 %	G0100 G40100	HPLC (4 min) 3.1 min 1.8 min.	[M+H] <sup>+</sup> =877.7 Da [M+2H] <sup>+</sup> =439.7 Da
2018/IP-15-02	C <sub>36</sub> H <sub>58</sub> N <sub>6</sub> O <sub>5</sub>	654.88 g/mol	90 %	G0100 G40100	HPLC (4 min) 2.7 min 1.4 min.	[M+H] <sup>+</sup> = 655.6 Da
2018/IP-15-03	C <sub>40</sub> H <sub>60</sub> N <sub>6</sub> O <sub>5</sub>	704.9 g/mol	>99 %	G0100 G40100	HPLC ( 4 min) 2.7 min 1.2 min	[M+H] <sup>+</sup> = 705.7 Da
2018/IP-15-04	C <sub>38</sub> H <sub>45</sub> N <sub>5</sub> O <sub>5</sub>	651.79 g/mol	>99 %	G0100 G40100	HPLC( 4 min) 3.0 min 1.7 min	[M+H] <sup>+</sup> = 652.5 Da
2018/IP-15-05	C <sub>41</sub> H <sub>53</sub> N <sub>5</sub> O <sub>4</sub>	679.89 g/mol	>99 %	G0100 G40100	HPLC (4 min) 3.0 min 1.7 min	[M+H] <sup>+</sup> = 680.6 Da
IPR-465	C <sub>39</sub> H <sub>47</sub> N <sub>5</sub> O <sub>5</sub>	665.82 g/mol	>99 %	G0100 G40100	HPLC (4 min) 3.1 min 1.8 min	[M+H] <sup>+</sup> = 666.3 Da
IPR-467	C <sub>49</sub> H <sub>51</sub> N <sub>5</sub> O <sub>4</sub>	773.96 g/mol	98 %	G0100 G40100	HPLC (4 min) 2.9 min 1.7 min	[M+H] <sup>+</sup> = 774.4 Da
IPR-468	C <sub>48</sub> H <sub>50</sub> N <sub>6</sub> O <sub>5</sub>	790.95 g/mol	92 %	G0100 G40100	HPLC ( 4 min) 2.7 min 1.3 min	[M+H] <sup>+</sup> = 791.4 Da
IPR-469	C <sub>44</sub> H <sub>47</sub> N <sub>5</sub> O <sub>4</sub>	709.88 g/mol	> 99 %	G0100 G40100	HPLC ( 4 min) 2.9 min 1.5 min.	[M+H] <sup>+</sup> = 710.3 Da [M+2H] <sup>+</sup> = 355.8 Da

### 2.3. RAD51-BRCA2 inhibitors

Code	Formula:	Calculated mass:	Purity:	Gradient ACN/H <sub>2</sub> O:	Retention time (min):	Mass Identification:
2018/IP-15-01	C <sub>54</sub> H <sub>64</sub> N <sub>6</sub> O <sub>5</sub>	877.12 g/mol	>99 %	G0100 G40100	HPLC (4 min) 3.1 min 1.8 min.	[M+H] <sup>+</sup> =877.7 Da [M+2H] <sup>+</sup> =439.7 Da
2018/IP-15-02	C <sub>36</sub> H <sub>58</sub> N <sub>6</sub> O <sub>5</sub>	654.88 g/mol	90 %	G0100 G40100	HPLC (4 min) 2.7 min 1.4 min.	[M+H] <sup>+</sup> = 655.6 Da
2018/IP-15-03	C <sub>40</sub> H <sub>60</sub> N <sub>6</sub> O <sub>5</sub>	704.9 g/mol	>99 %	G0100 G40100	HPLC ( 4 min) 2.7 min 1.2 min	[M+H] <sup>+</sup> = 705.7 Da
2018/IP-15-04	C <sub>38</sub> H <sub>45</sub> N <sub>5</sub> O <sub>5</sub>	651.79 g/mol	>99 %	G0100 G40100	HPLC( 4 min) 3.0 min 1.7 min	[M+H] <sup>+</sup> = 652.5 Da
2018/IP-15-05	C <sub>41</sub> H <sub>53</sub> N <sub>5</sub> O <sub>4</sub>	679.89 g/mol	>99 %	G0100 G40100	HPLC (4 min) 3.0 min 1.7 min	[M+H] <sup>+</sup> = 680.6 Da
IPR-465	C <sub>39</sub> H <sub>47</sub> N <sub>5</sub> O <sub>5</sub>	665.82 g/mol	>99 %	G0100 G40100	HPLC (4 min) 3.1 min 1.8 min	[M+H] <sup>+</sup> = 666.3 Da
IPR-467	C <sub>49</sub> H <sub>51</sub> N <sub>5</sub> O <sub>4</sub>	773.96 g/mol	98 %	G0100 G40100	HPLC (4 min) 2.9 min 1.7 min	[M+H] <sup>+</sup> = 774.4 Da
IPR-468	C <sub>48</sub> H <sub>50</sub> N <sub>6</sub> O <sub>5</sub>	790.95 g/mol	92 %	G0100 G40100	HPLC ( 4 min) 2.7 min 1.3 min	[M+H] <sup>+</sup> = 791.4 Da
IPR-469	C <sub>44</sub> H <sub>47</sub> N <sub>5</sub> O <sub>4</sub>	709.88 g/mol	> 99 %	G0100 G40100	HPLC ( 4 min) 2.9 min 1.5 min.	[M+H] <sup>+</sup> = 710.3 Da [M+2H] <sup>+</sup> = 355.8 Da

### 2.3. RAS-Effectors inhibitors

Code	Formula:	Calculated mass:	Purity:	Gradient ACN/H <sub>2</sub> O:	Retention time (min):	Mass Identification:
2018/IP-14-01	C <sub>43</sub> H <sub>50</sub> N <sub>6</sub> O <sub>5</sub>	730.89 g/mol	> 99 %	G0100 G40100	HPLC (4 min) 2.9 min 1.6 min	[M+H] <sup>+</sup> = 731.6 Da [M+2H] <sup>+</sup> = 366.6 Da
2018/IP-14-02	C <sub>42</sub> H <sub>51</sub> N <sub>5</sub> O <sub>4</sub>	689.88 g/mol	> 99 %	G0100 G40100	HPLC (4 min) 2.9 min 1.5 min	[M+H] <sup>+</sup> = 690.6 Da [M+2H] <sup>+</sup> = 346.1 Da
2018/IP-14-03	C <sub>42</sub> H <sub>60</sub> N <sub>6</sub> O <sub>5</sub>	728.96 g/mol	> 99 %	G0100 G40100	HPLC (4 min) 3.0 min 1.7 min	[M+H] <sup>+</sup> = 729.7 Da [M+2H] <sup>+</sup> = 365.7 Da
2018/IP-14-04	C <sub>54</sub> H <sub>68</sub> N <sub>6</sub> O <sub>5</sub>	881.15 g/mol	98 %	G0100 G40100	HPLC (4 min) 3.3 min 2.3 min	[M+H] <sup>+</sup> = 881.8 Da [M+2H] <sup>2+</sup> = 441.7 Da
2018/IP-14-05	C <sub>42</sub> H <sub>59</sub> N <sub>5</sub> O <sub>5</sub>	713.95 g/mol	99 %	G0100 G40100	HPLC (4 min) 3.1 min 2.0 min	[M+H] <sup>+</sup> = 714.6 Da [M+2H] <sup>2+</sup> = 358.1 Da
2018/IP-14-06	C <sub>49</sub> H <sub>71</sub> N <sub>7</sub> O <sub>5</sub>	838.13 g/mol	99 %	G0100 G40100	HPLC (4 min) 3.2 min 2.0 min	[M+H] <sup>+</sup> = 838.7 Da [M+2H] <sup>2+</sup> = 420.2 Da
2018/IP-14-07	C <sub>46</sub> H <sub>63</sub> N <sub>5</sub> O <sub>4</sub>	750.02 g/mol	80 %	G0100 G40100	HPLC (4min) 3.1 min 2.0 min	[M+H] <sup>+</sup> = 750.7 Da [M+2H] <sup>2+</sup> = 376.1 Da

2018/IP-14-08	$C_{41}H_{58}N_6O_5$	714.94 g/mol	> 99 %	G0100 G40100	HPLC (4 min) 3.0 min 1.7 min	$[M+H]^+ = 715.7$ Da $[M+2H]^{2+} = 358.7$ Da
2018/IP-14-09	$C_{44}H_{61}N_5O_5$	739.98 g/mol	> 99 %	G0100 G40100	HPLC (4 min) 3.4 min 2.4 min	$[M+H]^+ = 740.7$ Da
IPR-471	$C_{50}H_{66}N_6O_6$	847.10 g/mol	98 %	G0100 G40100	HPLC (4 min) 3.6 min 2.8 min	$[M+H]^+ = 847.5$ Da
IPR-472	$C_{51}H_{58}N_6O_5$	835.04 g/mol	91 %	G0100 G40100	HPLC (4 min) 3.2 min 2.0 min	$[M+H]^+ = 835.5$ Da $[M+2H]^{2+} = 418.2$ Da
IPR-473	$C_{47}H_{58}N_6O_5$	787.00 g/mol	95 %	G0100 G40100	3.1 min 1.9 min.	$[M+H]^+ = 787.4$ Da
IPR-474	$C_{49}H_{59}N_5O_6$	814.02 g/mol	> 99 %	G0100 G40100	3.6 min 2.7 min.	$[M+H]^+ = 814.5$ Da

#### 2.4. Retrome-L2 inhibitors

Code	Formula:	Calculated mass:	Purity:	Gradient ACN/H <sub>2</sub> O:	Retention time (min):	Mass Identification:
IPR-463	C <sub>49</sub> H <sub>67</sub> N <sub>7</sub> O <sub>6</sub>	850.10 g/mol	98 %	G0100 G40100	HPLC (4 min) 2.7 min 1.5 min	[M+H] <sup>+</sup> = 850.6 Da [M+2H] <sup>2+</sup> = 425.9 Da
IPR-464	C <sub>40</sub> H <sub>57</sub> N <sub>7</sub> O <sub>6</sub>	731.92 g/mol	> 99 %	G0100 G40100	HPLC (4 min) 2.5 min 0.8 min	[M+H] <sup>+</sup> = 732.4 Da [M+2H] <sup>2+</sup> = 366.9 Da