

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN
SCIENZE BIOTECNOLOGICHE, BIOCOMPUTAZIONALI,
FARMACEUTICHE E FARMACOLOGICHE

CICLO 35

Settore concorsuale: 03/DI - CHIMICA E TECNOLOGIE FARMACEUTICHE,
TOSSICOLOGICHE E NUTRACEUTICO - ALIMENTARI

Settore scientifico disciplinare: CHIM/08 – CHIMICA FARMACEUTICA

**INHIBITION AND CHARACTERIZATION OF TWO-METAL
ION ENZYMES TO TREAT CANCER: DNA POLYMERASE η
AND TOPOISOMERASE II α**

Presentata da: Michela Nigro

Coordinatore dottorato:

Prof.ssa Maria Laura Bolognesi

Supervisore:

Dott. Marco De Vivo

Co-supervisore:

Prof.ssa Marinella Roberti

Esame finale anno 2023

Table of contents

<i>Abstract</i>	10
<i>Aim of the project</i>	12
<i>Acknowledgments</i>	13
<i>Chapter 1: Introduction</i>	14
1.1 Cancer treatment and chemoresistance	14
1.2 Two-Metal ion enzymes	16
1.3 DNA topology and replication	19
1.4 DNA damages	20
1.5 Translesion Synthesis	22
1.6 Y-Family polymerases overview	25
1.7 DNA Polymerase η	26
1.7.1 Structure and function	26
1.7.2 Role of DNA polymerase η in chemotherapeutic drug resistance	29
1.8 Small molecules targeting directly Pol η	32
1.9 Topoisomerases	33
1.10 Catalytic mechanism of Topoisomerase II α	34
1.12 Role of Topoisomerase II α in cancer	35
1.11 Structural organization of Topoisomerase II α	37
1.13 Small molecules targeting Topoisomerase II: poisons vs catalytic inhibitors	38
1.14 References	41
<i>Chapter 2: Computer-aided identification, synthesis, and biological evaluation of DNA polymerase η inhibitors for the treatment of cancer</i>	47
2.1 Introduction	48
2.2 Results and discussion	50
2.3 Conclusions	74
2.4 Experimental section	75
2.5 References	110

<i>Chapter 3: Structural studies</i>	116
3.1 The Experimental Approach and Methods	116
3.1.1. Protein expression.....	116
3.1.2. Protein purification.....	117
3.1.3. SDS-PAGE.....	117
3.1.4. Western Blot.....	118
3.1.5. Native Gel Electrophoresis.....	118
3.1.6 Crystallization.....	119
3.1.7 X-ray data processing.....	121
3.1.8 Cryo-EM.....	122
3.1.9 Cryo-EM data collection and analysis.....	123
3.1.10 Thermofluor assay - thermal stability quality control.....	123
3.1.11 Differential scanning fluorimetry.....	124
3.1.12 Isothermal titration calorimetry (ITC).....	125
3.1.13 Dynamic light scattering (DLS).....	125
3.2 Results	126
3.2.2 Protocol optimization of Pol η expression and purification.....	126
3.2.3 Thermofluor characterization and differential scanning calorimetry.....	130
3.2.4 Isothermal titration Calorimetry (ITC).....	132
3.2.5 X-ray crystallography.....	133
3.2.6 Cryo-EM.....	138
3.3 Discussion	141
3.4 References	143
<i>Chapter 4: Hyaluronic acid-based prodrugs for efficient bioavailability</i>	145
4.1 Methods	146
4.1.2 Human cell culture.....	146
4.1.3 Cell proliferation assay.....	146
4.1.4 Internalization of the prodrug.....	147
4.1.5 Immunofluorescence.....	147
4.2 Results and discussions	148
4.2.1 Comparison between the prodrug and drug effects.....	148

4.3 References	158
<i>Chapter 5: Design, synthesis, dynamic docking, biochemical characterization, and in vivo pharmacokinetics studies of novel topoisomerase II poisons with promising antiproliferative activity</i>	159
5.1 Introduction	160
5.2 Results and discussion	163
5.3 Conclusions	175
5.4 Experimental sections	176
5.5 References	191
<i>Chapter 6: Conclusions and future perspectives</i>	197

Table of the figures:

Figure 1 General Mechanism for the 2M Enzymatic Mechanism for RNase H7a. (Taken from: Palermo G, et al., Catalytic metal ions and enzymatic processing of DNA and RNA. <i>Acc Chem Res.</i> 2015.)	17
Figure 2 Nucleotide incorporation mechanism. Scheme of nucleic acid synthesis performed by DNA polymerases. The nucleotide called +1 on the template strand indicates the insertion site. Nucleophile activation (A), nucleotide addition (B), and DNA translocation for nucleic acid polymerization (C), with liberation of a pyrophosphate (PPi) leaving group. Orange indicates the template strand (T) while blue indicates the primer strand (P). (Image taken from Genna V, et al., A Self-Activated Mechanism for Nucleic Acid Polymerization Catalyzed by DNA/RNA Polymerases. <i>J Am Chem Soc.</i> 2016).	18
Figure 3 Scheme of DNA damages. A) normal Watson-Crick base pair, in which adenine is reported in red, guanine in yellow, cytosine in green and thymine in blue. B) The scheme includes the majority of the DNA damages that can be fixed by the DNA polymerase family. C) The main damages repaired by DNA polymerase η , in particular: on the left the cyclobutane pyrimidine (thymine) dimer; in the centre the inter-crosslink between cisplatin and two guanine; on the right 6,4-photoproduct, which can block Pol η	21
Figure 4 Mechanism of translesion synthesis (TLS) by Pol η . The replication fork is stalled by DNA damage, and after a series of events (including the monoubiquitination of PCNA), a polymerase switch provides Pol η with an opportunity to bypass the lesion. Here, only one strand (leading strand) is represented. At the end of the process the lesion remains in the DNA, but the replication of this molecule can proceed.3.....	24
Figure 5 DNA Pol η structure. A) linear representation of all the domain that compose the structure; B) solved structure of the catlit core; C) RIR complex; D) NMR structure of UBZ domain; E) PCNA structure, which interact with Pol η through PIP domain. (Image take from Yang W. An overview of Y-Family DNA polymerases and a case study of human DNA polymerase η . <i>Biochemistry.</i> 2014)	27
Figure 6 Mechanism of TLS mediated by WNR.	29
Figure 7 Mechanism of action of DNA polymerase η that induce the chemoresistance to the Cisplatin drug.(Image taken from Saha P, Mandal T, Talukdar AD, Kumar D, Kumar S, Tripathi PP, Wang QE, Srivastava AK. DNA polymerase eta: A potential pharmacological target for cancer therapy. <i>J Cell Physiol.</i> 2021)	30
Figure 8 Structure of AraC.....	32
Figure 9 A) structure of PNR-7-02; B) structure of autintricarboxylic acid; C) structure of ellagic acid.	32
Figure 10 The catalytic cycle of topoisomerase II. 1) Bind, 2) bent, 3) cleavage, 4) transport, 5) ligation, 6) release, 7) reset. (Taken from Vann KR, et al., Topoisomerase II Poisons: Converting Essential Enzymes into Molecular Scissors. <i>Biochemistry.</i> 2021)	35
Figure 11 Structure of TOPOII. (Kendra R et al., Topoisomerase II Poisons: Converting Essential Enzymes into Molecular Scissors. <i>Biochemistry.</i> 2021)	38
Figure 12 Mechanism of action of TOPOII poisons and inhibitors to treat cancer. (A) TOPOII poisons such as etoposide, which trap TOPOII to 5' termini of DNA DSBs and cause accumulation of TOPOIIcc-blocked DSBs. (B) Merbarone inhibits TOPOII catalytically, stabilizes noncovalent TOPOII-DNA complexes. (image taken from: Uusküla-Reimand L, et al., Untangling the roles of TOP2A and TOP2B in transcription and cancer. <i>Sci Adv.</i> 2022).....	38
Figure 13 Structures of the main TOPOII inhibitors. (image taken from: Nitiss JL. Targeting DNA topoisomerase II in cancer chemotherapy. <i>Nat Rev Cancer.</i> 2009)	39
Figure 14 Structure and potency of rac-ARN18347 and ARN17212.	50

Figure 15 Interactions of ligands with Pol- η active site.³⁴ A) The binding mode of ATP at the transition state as obtained by previous quantum-based computations is represented.³⁴ ATP is shown as yellow sticks, while the ligand pocket is represented as a blue surface. Here, the DNA template and primer (black), and key Pol- η functional residues (blue) are shown as sticks, while the two catalytic metals are highlighted as orange spheres. Dashed lines represent H-bonds and metal coordinations (yellow), as well as π - π stacking (magenta) interactions. B) A representative binding mode of ARN18347 as obtained by molecular docking. The coloring scheme follows that of panel A. C-D) Two representative coordination modes of ARN17212 as obtained by molecular docking. The coloring scheme follows that of panel A. 52

Figure 16 . (A) Interactions of ARN17212 with the Pol η active site. The O4-MgB coordination, H-bond interactions with I48 and the templating base T(0), cation- π interaction with R61, and π - π interaction with the terminal primer residue P(-1) are represented by black, green, blue, and red dashed lines, respectively. (B) Time evolution of the O4-MgB coordination and π interactions of ring B of ARN17212 with R61 and P(-1). The π interactions were monitored by measuring the center-of-mass distances between the rings (CZ atom for R61). 53

Figure 17 Chemical structure of ARN17212 and its possible scaffold exploration. 53

Figure 18 A375. (A) Protein level of PYH2AX after 48h of treatment with cisplatin 1 μ M or 5 μ M, with compounds 21 and 17 (30 μ M), and the co-treatment, was detected by immunoblotting. (B) Relative protein level was quantified with ImageJ (from duplicates mean \pm SD). Results shown here are from one of two experiments with identical results. (C) A375 cells were treated for 24h with Cisplatin 0.5 μ M or 1 μ M, with 21 (20 μ M) and with both cisplatin and compound 21; and localization of PYH2AX was detected by microscopy.(D) splitted channel. Nuclei were stained with Hoechst 33342 (blue). 72

Figure 19 OVCAR3. (A) Protein level of PYH2AX after 48h of treatment with cisplatin 0.5 μ M or 1 μ M, with compounds 21 and 17 (30 μ M), and the co-treatment, was detected by immunoblotting. (B) Relative protein level was quantified (from duplicates mean \pm SD) with ImageJ. Results shown here are from one of two experiments with identical results. (C) OVCAR3 cells were treated for 24h with Cisplatin 0.5 μ M or 1 μ M, with 21 (20 μ M) and with both cisplatin and compound 21; and localization of PYH2AX was detected by microscopy. Below, splitted channels. Nuclei were stained with Hoechst 33342 (blue). 73

Figure 20 Automated pipeline. Here the entire process is described, which integrate the part of the EMBL HTX Lab process, the CrystalDirect technology and the MASSIF-1 beamline at the ESRF synchrotron. 120

Figure 21 Steps for the initial data processing used by autoPROC. 121

Figure 22 PEAQ-ITC. On the left, the instrument that was used thanks to the biophysical platform at the EMBL of Grenoble; on the right the schematic representation of the internal part of the instrument. The reference cell (in blue) is always filled in with deionized water; the sample cell (in pink) is filled in manually with the macromolecule (trying to avoid bubbles formation). The syringe has an automatic mechanism for inject the ligand inside. 125

Figure 23 A) Chromatogram of His-trap column, with the loading and the elution of the protein. On the right the SDS-page with the 1) pellet of the expression, 2) the total lysate after sonication, 3) the supernatant; 4) the flow through of the purification and 5-14) the elution of the protein. B) Chromatogram of the Hi-Trap Heparin column. C) Chromatogram of the gel filtration column, the Size exclusion chromatography 75 increase column. D) SDS-page with 15) the elution from the His-trap, 16-17) the sample after the cleavage, 20-21) the elution from the heparin column and 22-23) the protein after SEC. E) the western blot with the antibody against His-tag,

the signal is present just in the first two lines (24-25), so the cleavage was completed. F) Native-page gel, with 27-28) the enzyme alone, 29-31) the enzyme with the dsDNA. 128

Figure 24 A) SDS-page with 1) total lysate of the protein pellet, 2) supernatant, 3) flow through of the his-trap column, 4-5-6) wash with S2 buffer, 7-8-9-10) 10% of buffer B. The SDS-page on the right is showing the His-trap elution with the gradient of imidazole of the new protocol set up at the EMBL of Grenoble. B) The SDS-page to check the cleavage, 12) pull of the His-trap elution, 13) cleavage, 14) filtered cleavage, 15) pull of the second His-trap elution, 16) cleavage of the second pull after 12 h, 17) cleavage of the second pull after 15 h, 18) filtered cleavage, 19-20-21) second pull cleavage after 18, 19 and 20 h. C) the SDS-page with 22-23) samples loaded in the heparin column, 24-25) Flow through of the heparin, the rest of the samples are the elution of the heparin. D) the SDS-page with all the elution from the injections of the gel filtration column. 129

Figure 25 Thermal Shift Assay. A) The T_m related to two buffers at a fixed concentration with different concentration of salt. B) The T_m changing the buffers. C) The T_m with all the possible buffers with different pH. These data were performed one to check the stability of the sample. 130

Figure 26 ITC curves. A) binding curve between Pol η (in the cell, in complex with the dsDNA and Mg^{2+}) and the dATP (in the syringe); B) binding curve between the enzyme (in the cell in complex with the dsDNA and the dATP) and the Mg^{2+} (in the syringe at 5 mM concentration); C) binding curve between the enzyme (in the cell in complex with dsDNA, the dATP and the Mg^{2+}) and the compound 64 (in the syringe). 133

Figure 27 Coordination of the Mg^{2+} ions in the active site. The two structures reported represent the active site in the presence/absence of the nucleotide. 136

Figure 28 The structure of the active site of the two best structures that we obtain at the EMBL of Grenoble. In blue are reported the density map of the Mg^{2+} ion and of the dAMPNPP (non-hydrolyzable nucleotide). 137

Figure 29 Statistics data of the structure with the two Mg^{2+} ions and the nucleotide. The statistic of the other structure cannot be presented because refinement is not complete. 138

Figure 30 The grid chosen during the screening and the 2D classes. Selected for the 3D reconstruction. 140

Figure 31 Refinement. The graph of the final refinement indicates that the resolution reached is at 8.26 Å. The resolution was very low and it was impossible to fit the structure in this model. 141

Figure 32 HAB (left) boronic acids are introduced via amidation of HA. Boronates bind catechols of quercetin and reversibly release them at low pH (transition of boronic esters from quaternarized to trigonal form followed by hydrolysis), or in the presence of oxidants such as H_2O_2 . (Image taken from Quagliarello et al., 2021).. 146

Figure 33 Internalization of the prodrug of the quercetin. Two different concentrations were used: dark green the 50 μM concentration of the drug bound to the HA; light green the 10 μM concentration of the drug bound to the HA. The reported results are from three independent experiments (mean $\pm SD$). No statistical significance from two-way ANOVA test. 150

Figure 34 A375. Co-treatment using cisplatin (serial dilution, concentrations are reported in the X axis), and a fixed concentration (reported in the brackets) of the prodrugs. 151

Figure 35 A549. Co-treatment using cisplatin (serial dilution, concentrations are reported in the X axis), and a fixed concentration (reported in the brackets) of the prodrugs. 152

Figure 36 OVCAR3. Co-treatment using cisplatin (serial dilution, concentrations are reported in the X axis), and a fixed concentration (reported in the brackets) of the prodrugs. 153

Figure 37 Immunofluorescence to follow the internalization of HA-PBA-Cy3 [20 μM] (only A375) and HA-PBA-QC-Cy3 [50 μM] (red) at 3 different time points. 155

Figure 38 A375. Zoom to check the localization on the treatment with HA-PBA-64-Cy3 [50 μM]. 156

Figure 39 Hybrid topoII poison with the scaffold A (left) was explored to discover ARN16267 as a potent topoII blocker.³⁹ Here, the hybrid scaffold A was expanded to generate structures of type B (right above), with several functional groups introduced in the meta position. Structures of type C (right below) were generated by introducing a spacer between the thiobarbituric core and the mimic E-ring..... 161

Figure 40 Poison activity of the hybrid compounds. Agarose gel electrophoresis of plasmid DNA incubated in the absence (no enzyme) or presence of 1 U topoII containing either 1% DMSO as control vehicle (no enzyme and topoII lanes) or 200 μ M of the compound. Labels are shown above each lane. Numbers in the bottom correspond to the normalized intensity of the linear form (FIII). Plasmid forms are indicated by the arrow-points on the right: supercoiled (FI), nicked (FII), and linear (FIII)..... 170

Figure 41 Interaction diagram of the lead compound 3f bound to the DNA/topoII complex as derived from docking calculations. The computational study flags a cation- π interaction with Arg487, π -stacking with various DNA bases, and an H-bond with T+1, as the main drivers for ligand binding. 171

Figure 42 Dynamic description for the binding of 3f to the in the DNA cleavage/religation active site of topoIIa (PDBID: 5GWK). A) The lead compound in its binding mode. The residues directly interacting with the drug are shown in orange carbons, the rest of the DNA in pink, and the protein in white. The insets zoom into the distances of interest, particularly, those of Arg487 with the E-ring, and the merbarone core with the G+5 and C-1 bases. B) Evolution of the distances of interest over time..... 172

Figure 43 Mouse PK profiles of 3f following intravenous (i.v.) and oral administration (p.o.) at 3 and 10mg/Kg, respectively. The observed and calculated PK parameters following intravenous (I.V.) and oral administration (P.O.). The bioavailability F was calculated to 8% based on the AUC (area under curve) from t=0 to 240 min. 174

Figure 44 Summary of our results. We screened 34 compounds, and we tested only two compounds in three cancer cell lines. Co-treating the cells with compound 64 and cisplatin, we found out that they have a synergistic effect leading to cancer cell death. 198

Abstract

Chemotherapeutic drugs can in many ways disrupt the replication machinery triggering apoptosis in cancer cells: some act directly on DNA and others block the enzymes involved in preparing DNA for replication. Cisplatin-based drugs are common as first-line cancer chemotherapies. These drugs act by forming a covalent interaction with DNA leading to cancer cell death. Another example is etoposide, a molecule that blocks topoisomerase II α leading to the inhibition of dsDNA replication. Despite their efficacy, cancer cells can respond to these treatments over time by overtaking their effects, leading to drug resistance. This phenomenon is referred to as chemoresistance, and represents a major issue when treating cancer.

Chemoresistance events can be triggered by the action of TransLesion DNA Synthesis (TLS) enzymes like DNA polymerase η (Pol η). Pol η is exceptional in its ability to bypass specific DNA damages like cyclobutane pyrimidine dimers (CPDs), which can be caused by UV light. This polymerase helps also to bypass drug-induced damage in cancer cells, allowing DNA replication and cancer cells proliferation even when cisplatin-based chemotherapeutic drugs are in use. For these reasons, Pol η is a promising drug discovery target, whose inhibition would help in overcoming of drug resistance.

This study aims to identify a potent and selective Pol η inhibitor able to improve the efficacy of platinum-based chemotherapeutic drugs. We report the discovery of compound 64 (ARN24964), after an extensive SAR reporting 35 analogs. We evaluated compound 64 on four different cell lines: A375, A549, OVCAR3 and HEK293. Interestingly, the molecule is a Pol η inhibitor able to act synergistically with cisplatin. We could also unravel the mechanism of action of our compound by evaluating the expression level of p γ -H2AX by western blot and immunofluorescence. Moreover, we also synthesized a prodrug form that allowed us to improve its stability and the bioavailability. This compound represents an advanced scaffold featuring good potency, aqueous solubility, microsomal and plasma stability.

In addition to this central theme, this thesis also describes our efforts in developing and characterize a novel hybrid inhibitor/poison for the human topoisomerase II α enzyme, which represents a validated target to treat cancer. In particular, we performed specific assays to study the inhibition of Topoisomerase II α and we evaluated compounds effect on three cancer cell lines. These studies allowed us to identify a compound that is able to inhibit the enzyme with a good pK and a good potency.

Aim of the project

DNA polymerase η is a well known anticancer target. It is very interesting because, even if not essential for the cell survival, it plays a crucial role in the repair of the DNA replication. Indeed, we are interested in finding a new possible Pol η inhibitor. Pol η is involved in the repair mechanism of the DNA damages induced by a largely used chemotherapeutic drug called cisplatin (and all its derivatives), leading to the chemoresistance. This phenomenon is one of the biggest issue of the treatment of several types of cancer disease. Until now, only few inhibitors are reported in literature (for example: PNR-7-02), for this reason we started our SAR campaign to find a novel and potent inhibitor that could be used synergistically with cisplatin to treat cancer.

Moreover, we are interested in the mechanism of action at the atomic level. The second aim of our project is to obtain the co-crystal structure of the ternary complex composed by the enzyme, the dsDNA and the inhibitor. Since there are no crystal structure solved with an inhibitor compound, this result could be incredibly helpful to start a new structural based drug discovery campaign to find a more potent and selective lead compound.

At the same time, as a side project, we studied the inhibition of Topoisomerase II α , an enzyme important in the replication machinery helping in the decatenation of the DNA. This enzyme is also involved in cancer and it has a very similar active site, due to the use of two metal (magnesium ions) ions to open the dsDNA. For this enzyme we want to find a novel inhibitor that could act as a hybrid between poison and inhibitor compound.

Acknowledgments

I would like to start to say thank you to my supervisor, Dr. Marco De Vivo. Who, not only gave me the chance to complete my PhD, but also more importantly gave me the chance to be part of an amazing group that he always led with such positivity and as a perfect mentor for everyone. Thank you for believing in me and always pushing me to do better.

I would like to thank the coordinator of the PhD course, Prof. Maria Laura Bolognesi, who I hope will always remember her XXXV cycle students for being the first. Thank you for always being helpful, available and creating a family environment despite the distance.

Dr. Stefania Girotto, for helping me from day one in a new lab. Thank you for the stimulating discussions, for kindly letting me know my mistakes and, most important, to teach me how to move forward on my own. You are a great source of inspiration.

Dr. Isabella Maria Acquistapace, the postDoc, one of the brightest person I met. This dissertation would not have been like this without your incredible support.

Dr. Nicoletta Brindani and Dr. Federico Munafò, the chemists who were part of the project. Thanks to your crucial work, the project has achieved most of our goals.

Finally, I would like to thank all the facilities that helped and made the project more complete. In particular, the analytical chemistry facilities of IIT, thanks to which we were able to get the DMPK properties of our compounds and made important decisions to take the project forward.

I am very glad to have had the opportunity to spend six months at the EMBL in Grenoble where I met such inspiring scientists and in particular, I would like to thank Dr. Josan Marquez, head of the crystallography facility, for introducing me to this field of biology. For the structural part of the project, I need to thank also the platforms of Grenoble Instruct-ERIC center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-005-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecole Universitaires de Recherche) CBH-EUR-GS (ARN-17-EURE-0003). I thank Caroline Mas for assistance and access to the Biophysical platform that let me use the ITC. I have to say thank you to the IBS institute, which allowed me to use the Prometheus instrument and the entire team of MASSIF-1 beamline led by Dr. Andrew McCarthy that analysed most of my crystals.

Chapter 1: Introduction

1.1 Cancer treatment and chemoresistance

A resolute therapy does not exist for several types of cancer yet. Chemotherapy (drug treatment), surgical operation, radiotherapy and biotherapy, or a combination of these, are the main weapons currently in use. In particular, surgical operation is useful to remove the entire, or part, of the solid tumor, although it is not always possible to use the technique alone¹. Radiotherapy employs radiation, the physical agent used to destroy cancer cells. The type of radiation used is named “ionizing radiation” because it forms ions and deposits energy in the cells of the tissues it passes through. This deposited energy can kill cancer cells or cause genetic changes resulting in cancer cell death². Biotherapy has been developed in recent years: two novel types of immunotherapy have indeed had a marked impact on oncology: checkpoint inhibitory MAbs and chimeric antigen-specific receptor (CAR)-transfected T-cells (CAR-T cells). Immunotherapy has demonstrated to produce long-lasting effects in numerous types of cancer. Antigen-specific immune responses can clearly be effective, even in the late-stage disease. In addition, two other types of biotherapy, antitumor vaccines and oncolytic viruses (OVs), have been developed in recent decades. These types of therapy are physiological and well tolerated⁴.

More than fifty years ago, chemotherapy was introduced as a combination of chemical drugs to treat cancer. Drugs mixtures are very useful in fighting cancer cell resistance because the probability of developing resistance to multiple mechanistically distinct drugs is lower than the probability of resistance of one single drug. Moreover, using a combination of drugs can help to overcome tumor heterogeneity⁵. Chemotherapy acts by inducing cell death and/or inhibiting cancer cell growth and proliferation. For example, there are drugs that act as DNA alkylating agents to damage DNA, or as inhibitor of the DNA replication or antimetabolic agents, which act on microtubules equilibrium. Chemotherapy is still considered one of the main line treatment for many cancer diseases⁶. In the beginning, it was expected that these drugs would be able to kill selectively cancer cells with minimal side effects. Now, it is known that chemotherapy drugs are also useful to reduce the border of solid tumors to make surgical removal easier. This type of

treatment is called neoadjuvant therapy⁷. However, to date is clear that they still lack in selectivity and can quickly incur in chemoresistance.

In the last two decades, a great effort has been put into the search of new therapies leading to a shift from broad-spectrum cytotoxic drugs (e.g., the combination of trametinib and dabrafenib was approved for joint use in 2014, due to the rapid (6–7 months) development of resistance to the sole use of B-Raf inhibitors⁸) to targeted drugs. The targeted drugs can be divided into two main groups: small molecules and macromolecules.

Monoclonal antibodies, polypeptides, antibody-drug conjugates and nucleic acid belongs to the macromolecules group. Even though antibodies are the most selective, they are just able to target antigens on the cell surface and they often require intravenous or subcutaneous dosing due to their high molecular weight⁹. In addition, the production of this kind of molecules is expensive, as well as their controlled-temperature shipping and storage.

Small molecules, instead, have the advantage of a better pharmacokinetics (PK), easier patient compliance and lower costs. Even if monoclonal antibodies have recently been introduced for the treatment of a variety of diseases, the small molecules field is experiencing a huge growth. Indeed, 89 anti-cancer small molecules have been approved in the United States and China¹⁰. These 89 small molecules cover a large spectrum of targets such as kinases, DNA damage repair enzymes, epigenetic regulatory factors, and proteasomes. To date, there still are problems with the response rate and resistance.

A major issue to overcome in cancer therapy is chemoresistance. Various adaptation mechanisms are developed by cancer cells to survive the therapy. Drug resistance can arise through intrinsic and acquired factors. Examples of intrinsic factors are: an alteration in the expression of the target, the decrease in membrane permeability of the drug or a reduction of the binding affinity between the target and the drug. An example of cancer adaptation through the alteration of protein expression could be the acute myeloid leukemia (AML) treatment with cytarabine (AraC). The compound is an anti-cancer drug nucleotide that after multiple phosphorylations can be transformed into cytarabine triphosphate (AraC-triphosphate). AraC has usually no effect on cancer cells, while its phosphorylated form is lethal and damages them. Down-regulation or mutations in the enzymes involved in the compound phosphorylation pathway can reduce the AraC activity and can leads to the development of AraC-resistant cancer cells¹¹. Acquired resistance

is instead enhanced by genetic or environmental factors such as pH, hypoxia, and paracrine signaling interactions with stromal and other tumor cells¹².

The DNA repair pathways are among the most known cancer hallmarks targeted by drugs. In fact, increasing the levels of DNA damage can cause cell-cycle arrest and cell death. Moreover, DNA lesions that continue into the S phase of the cell cycle can block replication fork progression, resulting in the formation of DNA double-strand breaks (DSBs). DSBs are generally considered the most toxic of all DNA lesions¹³ because unrepaired or mis-repaired DNA can bring to genome instability and aberrations, leading to mutation that can affect cellular functions¹⁴.

For all these reasons, it is important to continue to put efforts into finding more drugs that specifically target a protein or an enzyme that is crucial for the survival of the cancer cells. Indeed, genetic therapy is another therapeutical approach that can be tailored to each patient and it is now in fast development. By studying cancer cells mutations, it is possible to create the correct combination of drugs. On the other hand, this implies a more expensive therapy, a possible lack of time and resistance.

In this work, we will focus mostly on the chemotherapy approach to treat cancer, and specifically on the mechanism that induces chemoresistance in cancer cell lines.

1.2 Two-Metal ion enzymes

Metal ions are frequently employed as modulators of biological reactions. Monovalent ions like Na^+ or K^+ are often needed for stabilizing the enzymes surface charge, while divalent ions, that bind with higher affinity, can have structural or catalytic roles¹⁵. Magnesium ions are among the most widespread metal species in cell systems, and they are also the most commonly found metal ion cofactors in enzymes. Due to their ability to create stable complexes with phosphate, they bind nucleotides and they are therefore crucial for DNA and RNA binding¹⁶.

Magnesium ion tends to adopt an octahedral coordination geometry, with a clear preference for oxygen-donor ligands or water. The binding to nucleic acids results either from nonspecific electrostatic attractions between the negatively charged phosphate groups and the metal cation, or from specific ligand coordination¹⁵.

Metalloenzyme, specifically the ones that use magnesium ions, are often capable to cleave or repair DNA and RNA (**Figure 1**). Thanks to X-ray crystallography, it is known that two divalent cations in the catalytic site of polymerases or topoisomerases are present, revealing a recurrent two-metal-ion mechanism for metal-aided enzymatic phosphoryl transfer¹⁷. Mg²⁺-ions enzymes often contain more than one single metal ion in the catalytic site. Studying the metal arrangement, it was shown that it is possible to find one, two or even three metals at the active site.

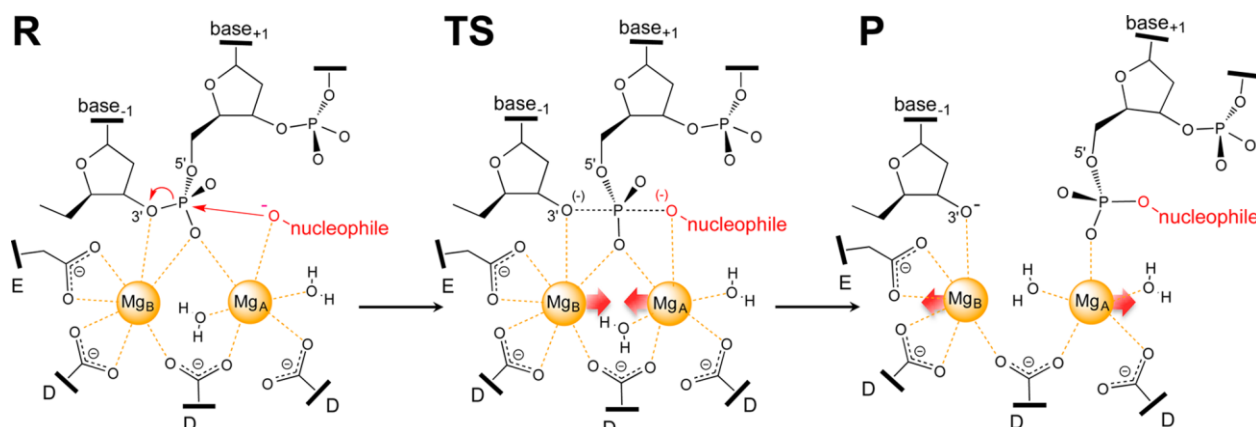


Figure 1 General Mechanism for the 2M Enzymatic Mechanism for RNase H7a. (Taken from: Palermo G, et al., Catalytic metal ions and enzymatic processing of DNA and RNA. *Acc Chem Res.* 2015.)

Specifically, one metal ion favors the formation of the attacking hydroxide ion, while the other helps the exit of the leaving group. Together the two metal ions stabilize the enzymatic transition state (TS).

Usually, the two divalent ions are coordinated, by the backbone of DNA or RNA, by catalytic acidic residues, and by nearby water molecules (*Figure 1*). A specific conserved sequence called “DEDD motif” is responsible for chelating metals. In general, the first mechanism for nucleophile formation is via an Asp residue, part of “DEDD motif”, and this residue can act as a general base for 3'-OH deprotonation. Therefore, this mechanism implies the nucleophile formation and NTP addition.

The role of the metal ions can be summarized in the three main steps:

1. Nucleophilic attack on the scissile phosphodiester;
2. Formation of the transition state showing a penta-coordinated phosphate along the phosphoryl transfer;

3. Cleavage of the scissile bond with subsequent exit of the leaving group.

Recently, the mechanism of action of DNA polymerases, one of the major groups of metalloenzymes, has become better understood.

Nucleic acid polymerization (**Figure 2**) is a key process for genetic stability. It is performed by sets of DNA/RNA polymerases, which are therefore considered drug targets. Polymerases act via the two metal-ion mechanism. The ion is the magnesium, which is needed for incorporating an incoming nucleotide (dNTP) with the release of a pyrophosphate (PPi) group.

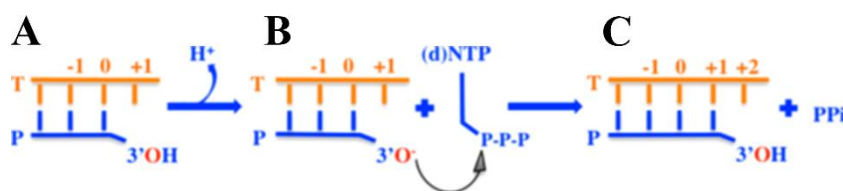


Figure 2 Nucleotide incorporation mechanism. Scheme of nucleic acid synthesis performed by DNA polymerases. The nucleotide called +1 on the template strand indicates the insertion site. Nucleophile activation (A), nucleotide addition (B), and DNA translocation for nucleic acid polymerization (C), with liberation of a pyrophosphate (PPi) leaving group. Orange indicates the template strand (T) while blue indicates the primer strand (P). (Image taken from Genna V, et al., A Self-Activated Mechanism for Nucleic Acid Polymerization Catalyzed by DNA/RNA Polymerases. *J Am Chem Soc.* 2016).

Magnesium ions are also very important for topoisomerase activity. In the eukaryotic topoisomerase II, magnesium ions promote enzyme-substrate interactions through two mechanisms: (i) direct contribution in enzyme-mediated DNA cleavage and (ii) involvement in ATPase reactions as magnesium-ATP substrate¹⁸. It is now clear that divalent metal ions, particularly magnesium, play an essential role in promoting DNA cleavage/elongation activity of topoisomerases and polymerases through both participation in the catalytic events and assistance in the appropriate structural changes. Catalytic Mg²⁺ species are consistently bound to a specific region in the active site.

1.3 DNA topology and replication

The term “DNA topology” was born in 1953 when Watson and Crick proposed the structure of the DNA¹⁹. Indeed, the topology defines how the two complementary strands are entangled²⁰. In particular, DNA topology explained many unanswered questions connected to the replication reaction. In fact, it is now known that the semiconservative replication model²¹ needs the untwisting of the double helix to take place. Actually, these modelling problems were solved only after the discovery of topoisomerases²², enzymes able to modulate DNA topology. Furthermore, the essential importance of DNA topology is emphasized by the strict conservation and necessity of these enzymes across all cell types and some viruses²³.

The 3D structure of the long genome DNA is very complex. It is formed by the combination of four different basis called: adenine, thymine, guanine and cytosine. Specifically, they can be divided in two groups: the pyrimidine, to which the cytosine (C) and thymine (T) belong, and the purine, to which the guanine (G) and the adenine (A) belong. A always pairs with T, and G with C. Accordingly, DNA topology embraces supercoiling, knots, and catenates. Due to the crucial role of DNA, it acts as storage of all genetic information, it needs to be organized in underwinding or overwinding dsDNA which results in positive or negative supercoiled respectively. DNA supercoiling is important to replicate and express only the necessary information. In fact, the two strands of DNA must be separated for the replication to take place. As a result, replication origins and gene promoters are more easily opened and rates of DNA replication and transcription are greatly enhanced^{24, 25}.

The DNA topology can be described as the three-dimensional structure of dsDNA, and also, DNA topology should be an important section of biochemistry and molecular biology because:

- The enzymes involved in the topology act is modulated every time a double helix has to be opened or moved;
- Topoisomerases are the most important enzymes to regulate the topological structure of DNA. They showed to have critical roles in fundamental cellular processes, including replication, transcription, recombination, and mitosis;
- Topoisomerases can also be the targets of very effective anticancer and antibacterial drugs.

Regarding the replication mechanism, it can be divided into three steps: (i) the opening of the double strands to generate the replication fork; (ii) the priming of the template strand; and (iii) the elongation of the new DNA primer. During separation, the two strands of the DNA uncoil in a specific position called the “origin”. Diverse enzymes and proteins then work in conjunction to prepare the strands for replication. Finally, a special enzyme called DNA polymerase organizes the reaction to create a new DNA strand. The main process to initiate DNA replication is organised in two steps. First, a specific initiator enzyme unwinds a short part of the dsDNA. Then, a protein called helicase attaches to and breaks apart the hydrogen bonds between the bases on the DNA strands, leading to the opening of the two strands. As the helicase moves along the DNA molecule, it continues breaking these hydrogen bonds and separating the two chains²⁴.

The DNA polymerase involved in this process, in physiological condition, belong to the A family of polymerase. In particular, mammalian cells possess five different DNA polymerases: α , β , δ , ϵ and γ . Polymerase α is the only eukaryotic polymerase capable of generating triggers for both DNA strands, while polymerase δ is responsible for the elongation of both strands.

1.4 DNA damages

One of the major causes of DNA damages is sunlight, in particular its ultraviolet radiation (UV) component. The UV can be divided in UV-A (wavelengths between 315-400 nm), UV-B (wavelengths between 280-315 nm) and UV-C (wavelengths below 280 nm). Of these, only UV-A and part of UV-B reach the earth surface. It is known that UV can penetrate just the outer-skin layers. Specifically, UV-B can reach the epidermal while the UV-A can penetrate deeper and reach the dermis. For these reasons UV-B are classified among the exogenous factor causing skin cancer, while UV-A are involved in skin aging processes. Human skin is constantly exposed to UV light; indeed it develops different type of protection (i.e. melanin synthesis) and repair mechanisms to avoid damages of the DNA. An excessive exposure to sunlight could lead to a failure of these defense mechanisms. In fact, UV rays can cause DNA rearrangements called photoproducts, which are well known to be mutagenic. Frequently, UV rays acts on two adjacent pyrimidine bases that become a dimer. This type of mutation is called cyclobutane pyrimidine dimer (CPD)²⁶(**Figure 3**).

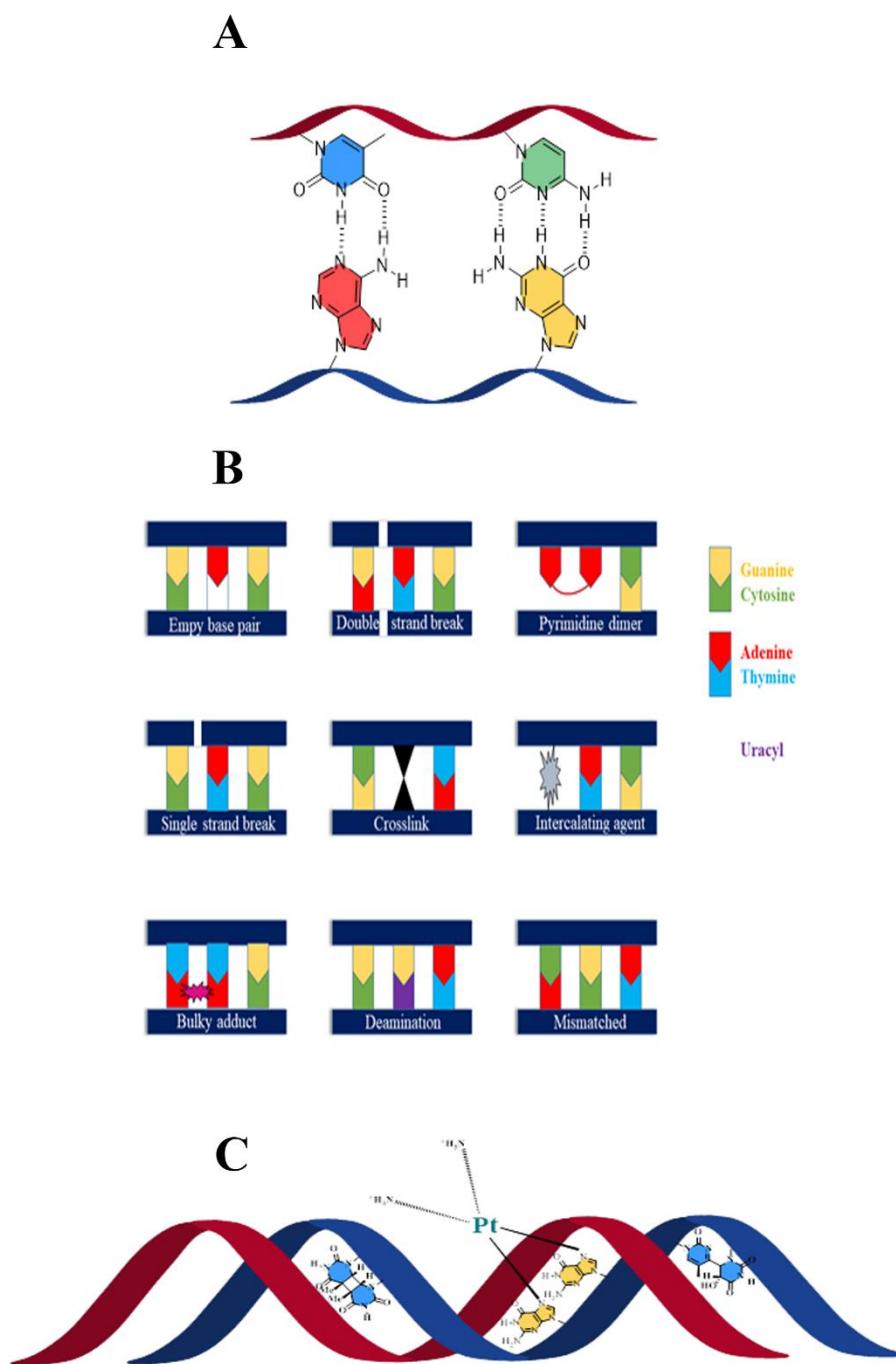


Figure 3 Scheme of DNA damages. A) normal Watson-Crick base pair, in which adenine is reported in red, guanine in yellow, cytosine in green and thymine in blue. B) The scheme includes the majority of the DNA damages that can be fixed by the DNA polymerase family. C) The main damages repaired by DNA polymerase η , in particular: on the left the cyclobutane pyrimidine (thymine) dimer; in the centre the inter-crosslink between cisplatin and two guanine; on the right 6,4-photoproduct, which can block Pol η .

DNA repair mechanisms are dramatically important. In certain human syndromes characterized by a lack of DNA repair processes, such as xeroderma pigmentosum (XP), the clinical effects of the mutations are intensified. Most of XP patients are known to have a deficiency in nucleotide excision repair (NER), a repair system that corrects DNA lesions induced by the UV rays. Some XP patients can instead present defects in the replication of UV-damaged DNA. The major clinical symptom of XP patients is the higher frequency of tumors in the visible areas of skin, indicating a strong relationship between DNA mutations, genetic instability and cancer. Moreover, XP cells were sensitive to UV and incapable to carry out unscheduled DNA synthesis (UDS). UDS, indeed, is part of the NER pathway. This exceptional discovery provided the first evidence that a cell defect in DNA metabolism is associated with a clinical phenotype. Many experiments helped to classify XP in seven different complementation groups (XP-A to XP-G), matching to seven mutated genes. Years later, however, the cells of some XP patients were demonstrated as NER proficient and capable of UDS. These cells were not able to replicate DNA in presence of UV-damages, confirming that other problems affected these cells. Indeed, it was understood that these cells were different and now were referred to as XP variants (XP-V). In conclusion, thanks to the identification of the mechanism of action of the DNA polymerase η , it was confirmed that XP-V patients were deficient in TransLesion Synthesis (TLS) DNA repair system³.

1.5 Translesion Synthesis

Translesion synthesis (TLS) is a conserved process, during the replication, from bacteria to humans²⁷, which involves several enzymes, known as TLS enzymes. Currently, 17 human DNA polymerases have been structurally and biochemically characterized. They can be classified into A, B, X, Y-families according to their sequence homology and structural similarities²⁸. Specifically, TLS is the process by which a DNA lesion is bypassed by the incorporation of a nucleotide opposite to the lesion. Many DNA lesions cannot be used as a template by the highly stringent replicative DNA polymerases, which are optimized to replicate the entire genome with high accuracy and efficiency.

The most common TLS Pols belong to the Y-family of DNA polymerases, which include Pol η , Pol ι , Pol κ , and Rev1, and B-family DNA polymerases Pol ζ . The remaining two families of Pols, A and X, also shows TLS activity. Actually, they contribute to DNA repair pathways through base

excision repair (BER) and non-homologous end joining (NHEJ)²⁹. Additionally, DNA primase/polymerase PrimPol, the most recent DNA polymerase discovered, has the capability of bypassing several of DNA lesions. Translesion synthesis is supposed to take place via two non-mutually exclusive processes. Firstly, TLS Pols are recruited at a replication fork, and secondly, they need to fill post-replicative gaps³⁰. In the first part of the mechanism, several enzymes are needed for the switching processes to occur: i) the replicative polymerase, which should be detached from the stalled replication fork; (ii) one or two TLS polymerases, that should bind to the replication terminus for nucleotide insertion and extension of the dsDNA; (iii) the replicative polymerase, which can intervene, after the dissociation of TLS pols, to continue the replication mechanism as usual (**Figure 4**). In conclusion, the pathway doesn't require many switching events³¹.

The TLS is one of the main pathways of the more mutagenic branch of DNA damage tolerance (DDT) processes. These mechanisms are very important; indeed cells that lack of DDT can incur in the risk of replication fork collapse, translocations, chromosome aberrations, and finally cell death. DDT uses a quite different approach from DNA repair. Rather than fixing DNA to its original sequence and structure, following DDT action, the lesion still exists. Since the main function of damage tolerance is to provisionally bypass a DNA lesion and not to regenerate the exact sequence, DDT mechanisms are adjusted to allow cell survival by promoting the completion of DNA replication to the detriment of accuracy of the genetic information. Consequently, it is not surprising that DNA damage tolerance often functions in a mutagenic manner.

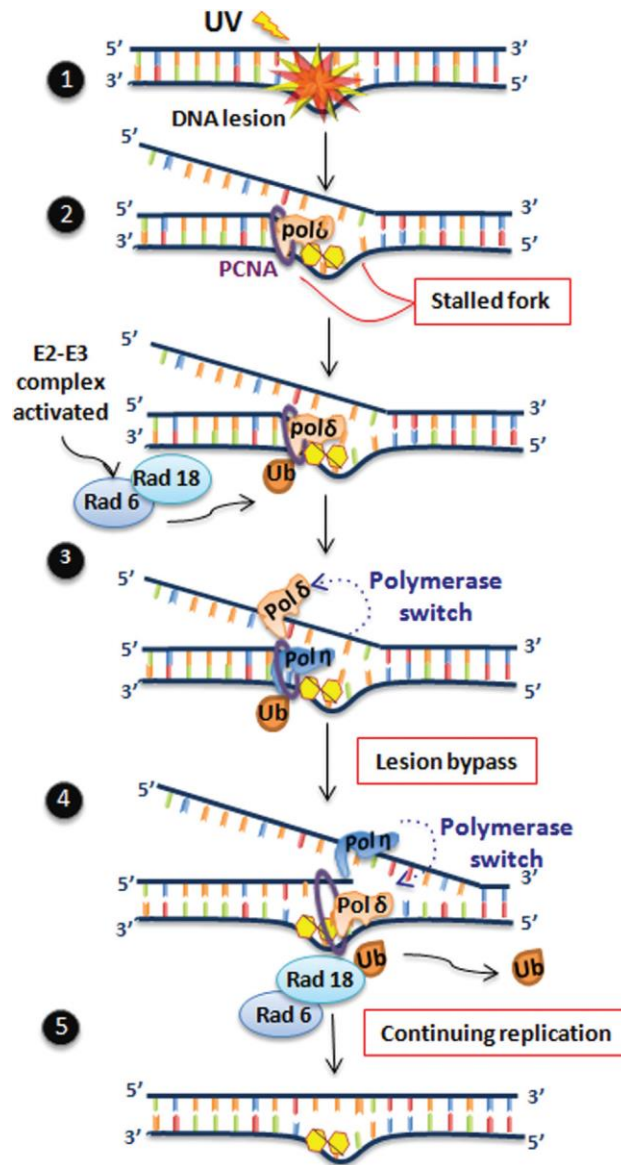


Figure 4 Mechanism of translesion synthesis (TLS) by Pol η . The replication fork is stalled by DNA damage, and after a series of events (including the monoubiquitination of PCNA), a polymerase switch provides Pol η with an opportunity to bypass the lesion. Here, only one strand (leading strand) is represented. At the end of the process the lesion remains in the DNA, but the replication of this molecule can proceed.³

1.6 Y-Family polymerases overview

Y-Family polymerases (Pols) are composed of six major groups, which are differentiated according to their amino acid sequence. *E. coli* pol IV (called DinB), Pol V and the four human polymerases Pol η , ι , κ , and Rev1 belongs to this family of Pols^{32,33}. All Y-Family polymerases comprise two functional parts: the catalytic portion of 350–500 residues and a regulatory portion from 10 (in DinB, Dbh and Dpo4) to 600 residues (in Rev1). At first, Pol η and ι were thought to be homologues, but it was realized that the two enzymes shared little in functional similarity. Inactivation of *POLH* (the gene encoding Pol η) in humans leads to a genetic disease that includes extreme UV sensitivity and predisposition to skin cancers. The function of Pol ι remains instead still unclarified, as Pol ι $-/-$ and Pol ι $+/+$ mice are comparable in all aspects of growth.

Usually, the Y-Family is involved in the first step of the TLS. Pol ζ (B-family)³⁴ takes over the primer extension as soon as the error on the DNA is repaired. The errors induced by UV that are usually encountered on DNA are two: *cis-syn* pyrimidine dimers (**Figure 3**) cyclobutane pyrimidine dimers, or CPD) and 6-4 photoproducts. Pyrimidine dimer is a frequent lesion that affects the single DNA strand. In particular, two adjacent pyrimidines are connected by a cyclobutene ring (**Figure 3**) in an intrastrand cross-link. CPD are different from the less frequent 6,4-photoproduct, which is also a pyrimidine dimer, but without the cyclobutene ring³⁵.

Because 6-4 photoproducts are efficiently removed by nucleotide excision repair (NER), they rarely arrive to the S phase. Rev1, which is the polymerase capable to perform NER, is a template-independent deoxycytidyl transferase and has more than 600 residues outside of the catalytic region. The non-catalytic regions of Rev1 seems to have a role in regulating other Y-Family polymerases and contribute to TLS more than its catalytic activity.

Another type of mutagenic factors are the polycyclic aromatic hydrocarbons (PAHs), pollutants generated through the incomplete combustion of fossil fuel and biomass. They can form DNA base adducts identified as carcinogens³⁶. An example of PAHs-induced DNA damage is benzo[a]pyrene. Dpo4 and eukaryotic Pol κ , are able to bypass these DNA bulky adducts. In line with this, the absence of Pol κ in cell cultures causes a slightly increase in UV sensitivity: Pol κ has been shown to be involved in excision repair of UV lesions.

Another mutation is 7,8-dihydro-8-oxoguanine (8-oxo-G). This lesion is oxidative and promutagenic. Replicative DNA polymerases tend to incorporate dATP opposite an 8-oxo-G, resulting in a G to T transversion mutation³⁷. Interestingly, Dpo4, Pol η and ι , and presumably Rev1 are capable to bypass 8-oxo-G with correct dCTP incorporation.

Although the Y family of polymerases is characterized by an enlarged site for replicating bases, each of the member differs in the shape and size of this site according to the lesion it has to repair. Pol η accommodates and correctly bypasses *cis-syn* pyrimidine dimers and also a 1,2-intrastrand d(GpG)–cisplatin crosslink³⁸, but, unfortunately, it is blocked by 6-4 photoproducts, BPDE–dG, and other bulky adducts. Similarly, only Pol κ , allows the accommodation of the minor groove of BPDE–dG adduct and accurately incorporates dCTP, to correct the lesion. Pol η relies on two exclusively conserved residues (Arg61 and Gln38 in humans) for bypassing *cis-syn* pyrimidine dimers. Nevertheless, these two residues also play a key role in the efficient misincorporation of dGTP opposite a dT, mainly when the mismatched dT:dGTP replicating base pair is immediately led by an AT base pair, in the DNA sequence known as the WA motif.

1.7 DNA Polymerase η

DNA polymerase η (Pol η), member of the Y-Family of Pols, is one of the most studied enzymes of TLS. Pol η stands out for its lesion-bypass specificity and efficiency. It plays an important role in supporting the DNA elongation in presence of lesions not only in physiological condition, but also in cancer environment when chemotherapeutic drug such as cisplatin are in use.

1.7.1 Structure and function

The DNA Pol η structure is formed by 713 amino acidic residues and is encoded by *POLH* gene located on human chromosome 6p21.1–6p12. It shares with all the other families of polymerases the catalytic core composed of the finger (res 16–87), palm (res 1–18 and 88–239) and thumb (res 240–306) subdomains. It also includes an additional domain, specific of the Y-Family, called the little finger (res 307–432) domain. The “little finger” domain (LF), also known as the polymerase-associated domain (PAD), has more sequence variability than the catalytic core³⁹.

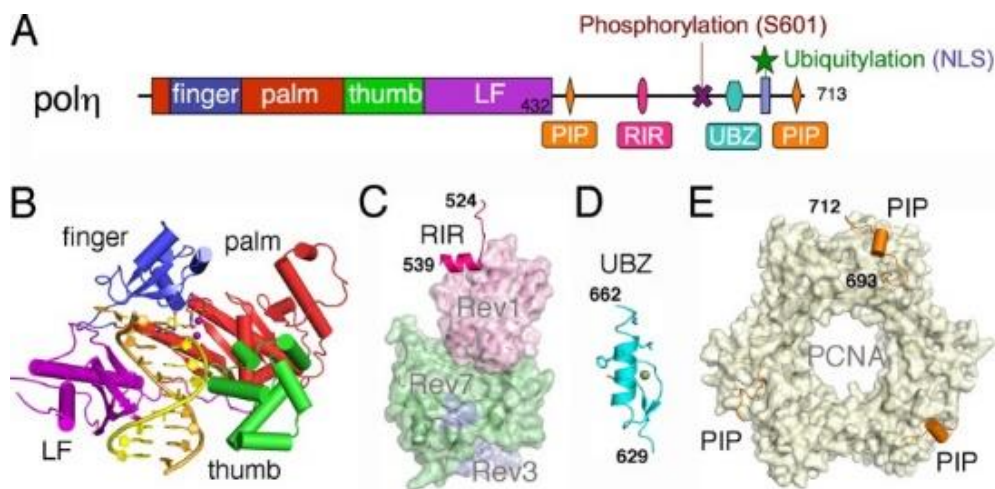


Figure 5 DNA Pol η structure. A) linear representation of all the domain that compose the structure; B) solved structure of the catalytic core; C) RIR complex; D) NMR structure of UBZ domain; E) PCNA structure, which interact with Pol η through PIP domain. (Image taken from Yang W. An overview of Y-Family DNA polymerases and a case study of human DNA polymerase η . *Biochemistry*. 2014)

Several three-dimensional structures of the catalytic region of a number of Y-Family polymerases from bacteria and archaea to humans have already been solved. The DNA substrate is bound between the thumb and little finger domain. Two metal ions or three as recently has been proposed, are contributing to the catalytic function, which can be either Mg^{2+} or Mn^{2+} , commonly referred to as Mg/MnA, B, and now also C⁴⁰. The function of the first two catalytic ions is largely recognized since these ions are identified in other DNA polymerases across different species and are related to different variations in quantum effects during catalysis⁴¹. In the catalytic core of human Pol η , these ions are coordinated by two conserved Asp residues (Asp15 and Asp115), a Met14 residue for MgA, and a Glu116 residue for MgB⁴². Although Pol η can use either Mg^{2+} or Mn^{2+} , Gao and Yang showed that it has a preference for Mg^{2+} . The third metal ion has been confirmed to be transient by nature; it aids in the protonation of the pyrophosphate, which is released during the incorporation of new dNTPs into the DNA double strands and neutralizes the protein's negative charge transition state during the elongation. Moreover, the third metal ion appears to act synchronically with a conserved Arg residue (Arg61). In this sense, Arg61 interacts through hydrogen bonds with the incoming dNTP α -phosphate and will, in the end, undergo a

conformational change, diverge from the site, and be replaced by the third metal ion and a water molecule⁴³.

The C-terminal region of Y-Family polymerases usually contains a PCNA interacting peptide (PIP), a Rev1 interacting region (RIR), an ubiquitin-binding module (UBM) and an ubiquitin binding zinc domain (UBZ). Rev1 polymerase C-terminal is instead characterized by N-terminal BRCT domain and a C-terminal protein binding domain³⁴ (**Figure 5**).

The regulatory region is fundamental for the function of the enzyme. Indeed, UV-induced DNA damage leads to the formation of the RAD6-RAD18 complex, which induces mono-ubiquitination of PCNA at lysine 164. This is a critical trigger to recruit Pol η at the site of the damaged replication, thanks to the recognition of the PIP motif. Pol η directly interacts with mono-ubiquitinated PCNA (ub-PCNA) and starts the progression of the stalled replication fork, after the switching with the replicative Pol⁴⁴. Furthermore, ub-PCNA interacts specifically with Pol η or other TLS polymerase and not with replicative polymerases. This provides an additional advantage to TLS polymerases that does not have to compete with replicative polymerases in bypassing the damages⁴⁴. Nevertheless, the requirement of PCNA mono-ubiquitination for Pol η -mediated TLS is not well characterized. Some reports suggest that Pol η can bypass UV-induced CPD lesions independently from the ub-PCNA presence. Other studies reflect that mono-ubiquitination is crucial to bypass UV induced lesions⁴⁵. Moreover, additional studies have shown that Pol η exhibits a similar affinity for mono-ubiquitinated and not ubiquitinated PCNA⁴⁶. Therefore, it is still unclear if the mono-ubiquitination of PCNA mainly enhances the recruitment of TLS polymerases by increasing their affinity to the damage site and facilitating the polymerase switch “in” or “out” between high fidelity polymerases and TLS polymerases.

To summarize, in the case of the CPD lesion, Pol η can perform both nucleotide incorporation and extension activity⁴⁷. 6-4 photoproducts can instead be bypassed through the cooperative work of two TLS polymerases: Pol η and Pol ζ . Specifically, Pol η performs the incorporation, and Pol ζ performs the extension. In the case of Pol η deficiency or mutational inactivation, UV-induced lesions can be bypassed by backup TLS polymerases belonging to the Y-family⁴⁸. The number of TLS polymerases required to bypass a specific lesion depends on the type of dNTP that should be incorporated opposite to the damaged site.

An important player in the TLS pathway is WRN⁴⁹ (**Figure 6**). A member of the RecQ helicase family, it plays a role in DNA unwinding, which can help the mechanisms exhibited by Pol η to bypass different types of DNA lesions. The presence of genetic defects in the WRN gene, results in the Werner syndrome. WRN can physically interact and embrace PCNA. In the literature, it is shown that WRN regulates TLS through modulation of ub-PCNA. The DNA damages induced by chemotherapy/radiation leads to phosphorylation of WRN by ATM/NBS1, and this phosphorylation event induces its degradation. This degradation triggers the release of PCNA from WRN bonding, and free PCNA could be monoubiquitinated through the Rad6/Rad18 complex. Ub-PCNA recruits Pol η and activates the TLS mechanism for lesion bypass.

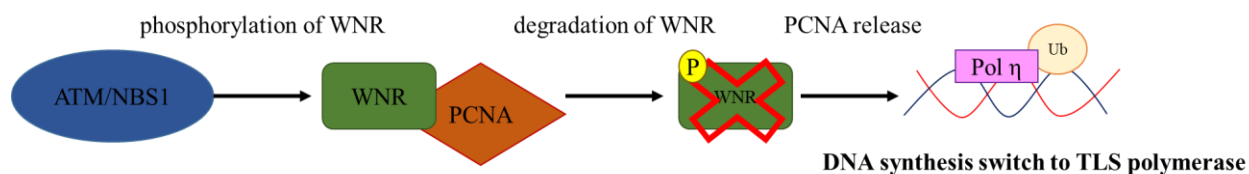


Figure 6 Mechanism of TLS mediated by WRN.

1.7.2 Role of DNA polymerase η in chemotherapeutic drug resistance

DNA polymerase η is also able to perform translesion synthesis on damages introduced by therapeutic agents. Indeed, it can repair or bypass N-acetyl-2-aminofluorene-modified guanine, 8-oxoG, O6 - methylguanine (O6 MeG), and cisplatin intrastrand crosslinked guanosines (Pt-GG)⁵⁰. For example, O6 -methylguanine is a mispairing lesion that causes GC \rightarrow AT transition mutations. This type of damage strongly blocks replicative DNA polymerases. Nevertheless, human Pol η is able to repair the lesion quite efficiently by inserting a C or a T opposite to it⁵¹.

The ability of Pol η to bypass DNA lesions helps the cells to continue the elongation, thus providing an alternative pathway in order to survive. Indeed, the DNA duplication process is very important for cell proliferation. Cancer cells have higher proliferation rates than healthy tissues. For this reason, anticancer agents are often designed to target the DNA replication pathway. So far, many anticancer therapeutic drugs are designed to inhibit DNA elongation.

It is known that Pol η can accept abnormal DNA structures. The enzyme contributes by helping the replication machinery to overcome these anticancer agents, therefore reducing their activity

towards cancer cells. This mechanism, thanks to which the cancer cells can survive, leads to drug resistance.

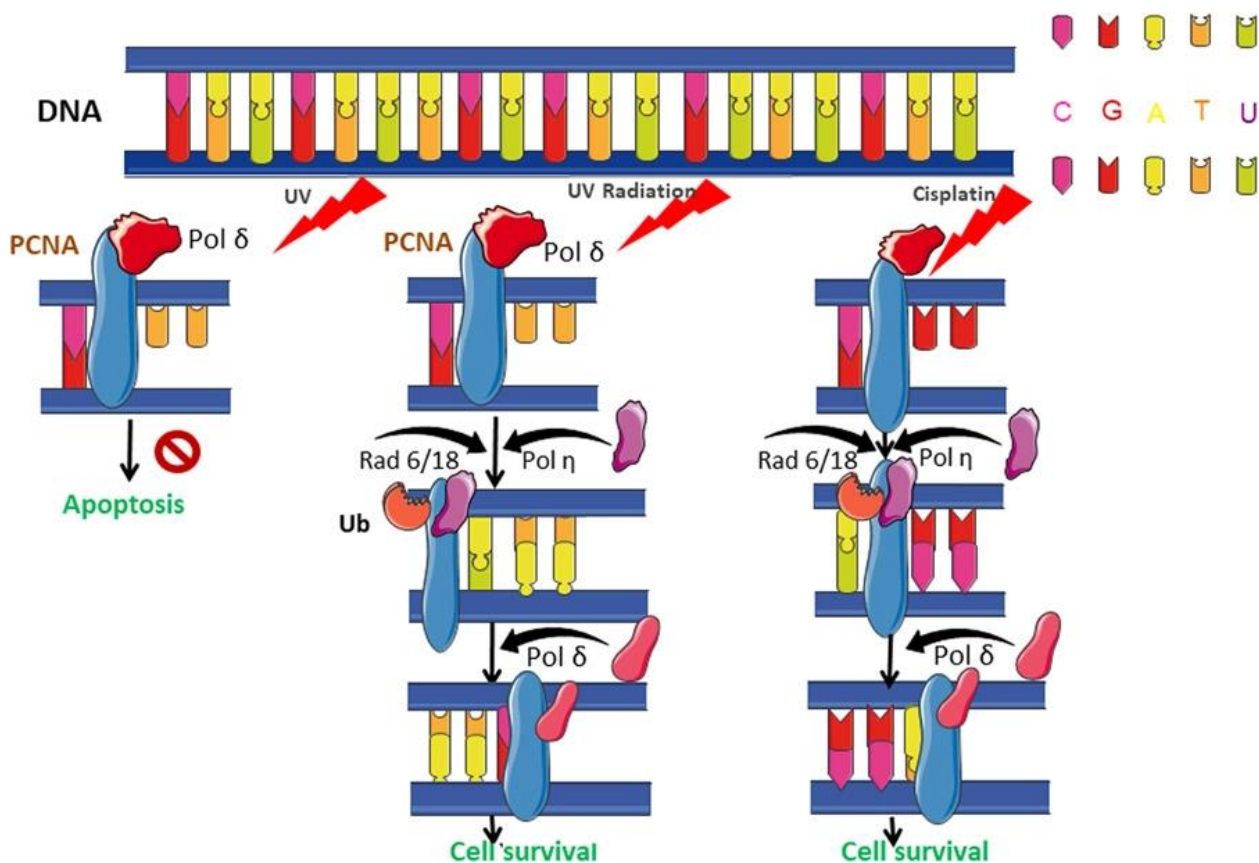


Figure 7 Mechanism of action of DNA polymerase η that induce the chemoresistance to the Cisplatin drug.(Image taken from Saha P, Mandal T, Talukdar AD, Kumar D, Kumar S, Tripathi PP, Wang QE, Srivastava AK. DNA polymerase eta: A potential pharmacological target for cancer therapy. *J Cell Physiol.* 2021)

The most extensively studied anticancer compounds that interact with Pol η are cisplatin and its analogs. Cisplatin and the Pt-derivatives oxaliplatin and carboplatin are the most widely clinically used compounds for cancer treatments. Cisplatin and carboplatin, in particular, are effective against solid tumors such as small-cell lung, ovarian, head, and neck cancers. Oxaliplatin is usually used for the treatment of primary advanced colorectal cancer and cisplatin resistant ovarian cancers²⁶. Although cisplatin and its analogs are widely used clinically, a major limitation for these compounds is the induced drug resistance. It is believed that they act through the formation of DNA adducts leading to the stall of the replication fork, stop DNA chain elongation, and ultimately cause cancer cell death. Cisplatin forms covalent complexes with DNA and leads to 1,2-intrastrand

GG cross links (65%) (**Figure 3**), and 1,2-intrastrand AG cross links (25%). In addition, other forms of minor lesions have also been reported, such as the interstrand GG crosslinks and other monoadducts.

DNA repair mechanisms have been studied for their potential implication on the cisplatin-induced resistance. Among these DNA repair mechanisms, NER has been shown to be the main DNA repair pathway in action. However, the potential involvement of translesion DNA polymerases in bypassing crosslinked DNA has also been investigated. Biochemical analysis on the interactions between cisplatin-crosslinked DNA template and several translesion synthesis polymerases has shown that Pol η has better efficiency in translesion synthesis past Pt-GG adducts as compared to other TLS polymerases. Indeed, Pol η is three time more efficient (kcat/Km) than pol β in bypassing cisplatin lesion (Pt-AGG).

Thanks to the co-crystal structure of a yeast Pol η in complex with cisplatin-crosslinked GG-containing DNA template (PDB: 2R8J), it is possible to understand its mechanism of action. In detail, it was highlighted that Pol η correctly selects dCTPs by recognising the complementarity of hydrogen bonding between 3'OH group of dCTP and the backbone of Pol η at Phe35, instead of the induced fit mechanism employed in high fidelity polymerases³⁸. In addition to the structural and biochemical characterization, a study in cells showed a 2–3-fold higher cisplatin-induced mutation frequency in fibroblast cells derived from XP-V patients (lacking DNA repair); as compared to (i) the isogenic XP-V cells complemented with Pol η expression or (ii) wild-type human fibroblasts. Moreover, cells that express Pol η are less sensitive to cisplatin treatment. This strongly suggested that Pol η has a fundamental role in the error-free translesion synthesis across the cisplatin adducts. A recent study has in fact demonstrated that the Pol η expression level is correlated with the cytotoxicity of cisplatin in non-small cell lung cancer. This study also suggested that the expression level of Pol η could be used as a predictive but not a prognostic marker in non-small cell lung cancer patient candidate to platinum-based chemotherapy.

In addition to cisplatin and platinum-based compounds, Pol η has also been shown to affect the action of anticancer nucleoside analogs. Nucleoside analogs are another important class of DNA targeting anticancer agents that have long been used for cancer treatment. For example, AraC is used to treat leukemia and gemcitabine, and b-D-20,20 - difluorodeoxycytidine (dFdC) is used to treat various types of cancer, including non-small cell lung cancer, breast, and pancreatic cancer.

Similar to cisplatin, the acquired drug resistance is also a limitation for nucleoside analogs. Structural and mechanistic studies showed that the arabinose sugar moiety of AraC (**Figure 8**) and the di-fluoro group on the 20 position of the sugar moiety of dFdC alter the DNA structure, which significantly reduced the extension efficiency for replicative DNA polymerases. In comparison, Pol η efficiently extends from either AraC or gemcitabine at the 30 termini of DNA. In addition, Pol η also replicates across both AraC and dFdC sites in the template DNA, which was shown to block DNA polymerases²⁷.

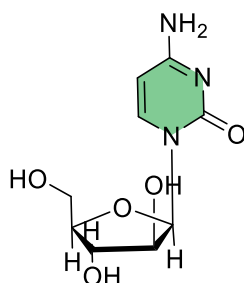


Figure 8 Structure of AraC.

1.8 Small molecules targeting directly Pol η

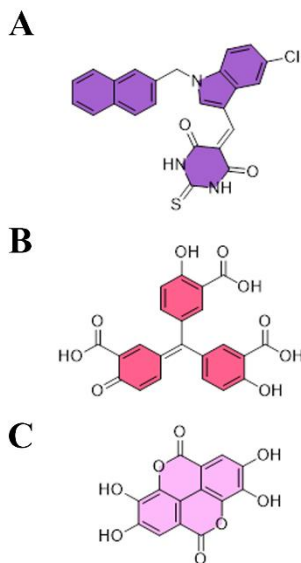


Figure 9 A) structure of PNR-7-02; B) structure of autintrinsic acid; C) structure of ellagic acid.

To date, a panel of small molecules inhibitors of Y family of polymerase have been developed. Among these autintrinsic acid (IC_{50} = 77-99 nM) and ellagic acid (IC_{50} = 62-81 nM) (**Figure 9**, panel B and C) were identified as potent and nonselective inhibitors of TLS pols family. Another class was identified by Eoff *et al.* (2018), which is an indole thioarbituric acid series of allosteric inhibitors of Pol η . The most potent is the indole thioarbituric acid derivative ITBA-12, identified as hit compound. ITBA-12 was evaluated against different DNA polymerases demonstrating that the compound could also inhibit other Y-family Pols. An additional SAR from the ITBA scaffold have been designed, from which PNR-7-02 (*Figure 9*, panel A) was identified as a slightly more potent inhibitor of Pol η (IC_{50} = 8 μ M)²⁸.

1.9 Topoisomerases

DNA topology is crucial for cell survival.⁵² To avoid detrimental consequences, torsional stress as well as knots and tangles in DNA must be managed. Both intramolecular knots and tangles can block essential cellular processes: the knots make double strand division impossible; the tangles can be an obstacle to chromosomes segregation. Knots, that are links within a single DNA segment, can be generated during recombination, and tangles, that are links between different DNA segments, during replication.

Topoisomerases are the enzymes involved in the regulation of the DNA topology. They act by generating transient breaks in the double strands. Topoisomerases can be divided into two groups: type I and type II, based on the number of DNA strands that they can cleave and ligate. Type I topoisomerase are able to transiently cleave single strand DNA to modify the topology of supercoiled DNA. Type II topoisomerase are able to cleave both the strands of DNA in order to modulate underwinding and overwinding and remove tangles and knots^{24,53}. Specifically, DNA can be both negatively supercoiled (left-handed or under-winding) or positively supercoiled (right-handed or over-winding). Usually is negatively supercoiled.

Type I enzymes can be divided in two other subclasses: type IA and type IB. Type I topoisomerases, after creating transient ssDNA breaks, makes possible: (i) the independent passage of the opposite intact strand through the break (type IA); or (2) the controlled rotation of the helix around the break (type IB). Also, type IA requires divalent metal ions for their activity, and they covalently bind the 5'-terminal phosphate of the broken DNA strand. Type IB enzymes, instead, do not need divalent metal ions and they covalently bind the 3'-terminal phosphate⁵⁴. Moreover, thanks to the ability of both, they can resolve inter-linked ssDNA molecules. Nevertheless, type I topoisomerases are not able to unknot or untangle duplex molecules. In eukaryotes there are two nuclear type I topoisomerases: topoisomerase I (a type IB) and topoisomerase III (a type IA)⁵⁴. Nuclear topoisomerase I acts by relaxing underwound or overwound DNA. It also removes positive super helical twists that accumulate before replication and transcription complexes. It is known from the literature that levels of topoisomerase I are high in proliferating tissues and cancer cells^{54,24,55}. Topoisomerase III is a type IA enzyme. Most eukaryotes express two isoforms of this enzyme: topoisomerase IIIa and b. Topoisomerase IIIa

works, together with the RecQ helicase, to resolve recombination intermediates. Less it is understood about the cellular functions of topoisomerase IIIb^{24,25}.

In contrast to type IA and type IB enzymes, type II α and II β DNA topoisomerases need ATP and divalent metal ions to catalyse the transport of one intact DNA double helix through another. Type IIA enzymes share several common mechanistic features with the less-extensively studied type IIB enzymes, but there are distinct structural differences between the two subfamilies. Human topoisomerase II α and II β are encoded by two separate genes (17q21–22 and 3p24, respectively) and diverge in molecular mass (170 and 180 kDa, respectively). They share extensive sequence identity (70%) and show similar enzymatic properties but differ significantly in their functions, cellular regulation, and expression. The enzyme topoisomerase II α is expressed in correlation with cellular growth and is crucial for the survival of proliferating cells. Though the α isoform is practically non-existent in quiescent and differentiated tissues, rapidly proliferating cells can contain ~500,000 molecules. Topoisomerase II α acts on replication forks and remains closely bound to chromosomes during mitosis. The enzyme has a role in DNA transcription and recombination. It is important for fork conjunction and termination of replication, and it is required for proper chromosome organization and segregation. In contrast to topoisomerase II α , topoisomerase II β it is essential for neuronal development. High levels of the β isoform are found in most cell types, independent of proliferation status. Topoisomerase II β seems to play an important role in the transcription of hormonally and developmentally regulated genes²⁵.

1.10 Catalytic mechanism of Topoisomerase II α

The catalytic mechanism of Topoisomerase II α can be easily explained in seven steps (**Figure 10**)^{25,52,55}.

1. Bind: the enzyme starts the mechanism through the binding to two DNA segments crossing over each other. The enzyme binds first, through a transesterification reaction, the dsDNA that will be cut (G-segment). The second dsDNA (T-segment) will be moved through the cleaved G-segment.

2. Bent: the enzyme bents the DNA sequence of the G-segment, this sequence is distorted to an angle of about 150° and can be used as catalytic site. In this step the presence of divalent metal ions is required.
3. Cleavage: the cleavage of the G-segment is possible thanks to the two tyrosine residues that can perform a nucleophilic attack.
4. Transport: after the binding with two ATP molecules, the enzyme can close the N-gate and force the T-segment to pass through the open gate DNA, one ATP molecule is hydrolysed.
5. Ligation: the G-segment is ligated and the T-segment can be released.
6. Release: After the second ATP hydrolysis, the G-segment can be released.
7. Reset: the enzyme conformation is reset and is now ready to take another DNA crossover.

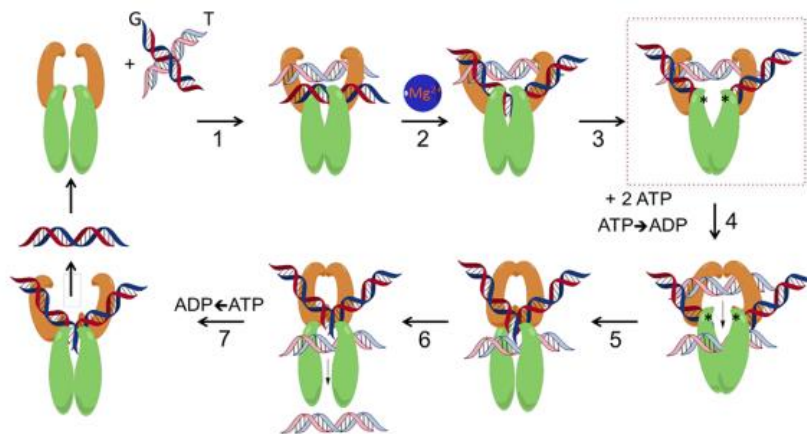


Figure 10 The catalytic cycle of topoisomerase II. 1) Bind, 2) bent, 3) cleavage, 4) transport, 5) ligation, 6) release, 7) reset. (Taken from Vann KR, et al., *Topoisomerase II Poisons: Converting Essential Enzymes into Molecular Scissors*. *Biochemistry*. 2021)

1.12 Role of Topoisomerase II α in cancer

The DNA topoisomerase II α (TOPOII α) enzymatic activity is essential for cell survival and, more important, the dsDNA breaks catalyzed by the enzyme are crucial in proliferating cells. Specifically, the entanglements, which occur during DNA repair, recombination, replication, transcription and segregation, can be fixed by letting the passage of dsDNA segment through these openings. TOP2 α interfacial inhibitors/poisons are widely exploited as anticancer drugs. The anticancer drugs in use, apply their cytotoxic effects by hampering the reversal of the TOPOII α -

DNA covalent cleavage complexes, which subsequently leads to the accumulation of DNA breaks and ultimately cell death⁵⁶.

TOPOII α poisons are normally used as chemotherapeutic agents in adults and pediatric patients to treat a wide variety of solid tumors, leukemias, and lymphomas. Indeed, cisplatin/etoposide is already in use as first-line treatment for small cell lung cancer; while other drugs are used in combination with doxorubicin and epirubicin as a preoperative/adjuvant therapy regimen for breast cancer; also, to treat acute myeloid leukemia (AML), daunorubicin and mitoxantrone are used.

Although TOPOII α poisons are extensively utilized, the major problem that affects therapies targeting this enzyme is the acquired chemoresistance, which decrease the drugs efficacy. To date, many chemoresistance mechanisms have been understood, however acquired resistance to TOPOII α poisons is frequently associated with decreased of its expression levels or altered sub-cellular localization. It is known that the cytotoxic activity of these drugs rely on the formation of TOPOII α -DNA covalent cleavage complexes.

The currently used TOPOII-targeted anticancer drugs are dual inhibitors of TOPOII α and β . So far, TOPOII β rather than TOPOII α is known to be implicated in the adverse effects of Top2-targeted drugs. An example is the therapy-related acute myeloid leukemia (t-AML) which results from balanced chromosome translocations present in several patients following etoposide and mitoxantrone treatments. A recent study also showed that TOPOII β poisoning is involved in the cardiotoxicity of anthracyclines. Despite the toxic side effects linked to TOPOII β , the focus is always in finding a selective and potent drug targeting TOPOII α . These because this enzyme is highly expressed and sometimes amplified in tumors such as breast and colon cancers along with ERB2. However, TOPOII β is also acting as a housekeeping gene with broad roles at promoter sites. This conclusion questions the validity of further developing the TOPOII β specific inhibitor R(+XK469).

The drug discovery pipeline of TOPOII α -specific inhibitors is simplified by the easily achievable enzyme, which is sold readily available for biochemical assays. In addition, rational drug development could take into consideration the prior knowledge that anthracyclines tend to be TOPOII α -specific whereas demethylepipodophyllotoxin derivatives are very effective against TOPOII β . This is particularly significant nowadays, with the availability of drug-enzyme-DNA

co-crystal structures that can be used to rationalize the chemical design of TOPOII α -specific drugs⁵⁷.

1.11 Structural organization of Topoisomerase II α

Type II topoisomerases are divided into two distinct subfamilies according to the structure and the mechanism: II α and II β . II α family includes eukaryotic TOPOII, viral Top2s, bacterial DNA gyrase and Top4. Structural and biochemical information are known for the II α enzymes: they have a similar architecture, with a twofold symmetric subunit composition. Eukaryotic TOPOII are large homodimers, with each protomer (~150 kDa) formed by an N-terminal and C-terminal domains. In particular, the first one is composed by: GHKL (Gyrase, Hsp90, Histidine Kinase, MutL) also known as ATPase domain, followed by a structure-relaying transducer domain, a Rossmann-like metal-chelating TOPRIM (topoisomerase-primase) domain, a winged-helix DNA-binding (WHD) domain, a tower/shoulder domain, a coiled-coil region that links to the main dimerization domain. The second is a less conserved and largely disordered C-terminal domain (CTD), it contains nuclear localization signals and post-transcriptional modification sites. The CTR is largely disordered when it is not bound to DNA, and it is poorly conserved among the type II enzymes. The domain is not necessary for enzyme function, but it contains protein elements that allow topoisomerase II α to discern the geometry of DNA supercoils and to remove positive supercoils ~10 times faster than it does with negative supercoils. The active site is where the tyrosine is located, specifically in the WHD domain of the cleavage core. Despite the high degree of similarity throughout the ATPase domain and the cleavage core, a key distinction between the bacterial and eukaryotic enzymes is that the CTDs of GyrA and Top4C adopt a unique β -pinwheel fold with DNA-binding and/or DNA-wrapping activity. However, the C-terminal tails of the

eukaryotic enzymes are likely unstructured and have been implicated in the regulation of cellular localization and enzymatic activity^{55,25}.

1.13 Small molecules targeting Topoisomerase II: poisons vs catalytic inhibitors

The drugs targeting DNA topoisomerase II that have already been in use can be divided into two groups according to their mechanism of action (**Figure 12**): catalytic inhibitors and poisons. Catalytic inhibitors act, for example, through the inhibition of the DNA binding and/or by blocking the ATP-binding site, hampering ATP hydrolysis which is needed for TOPOII function.

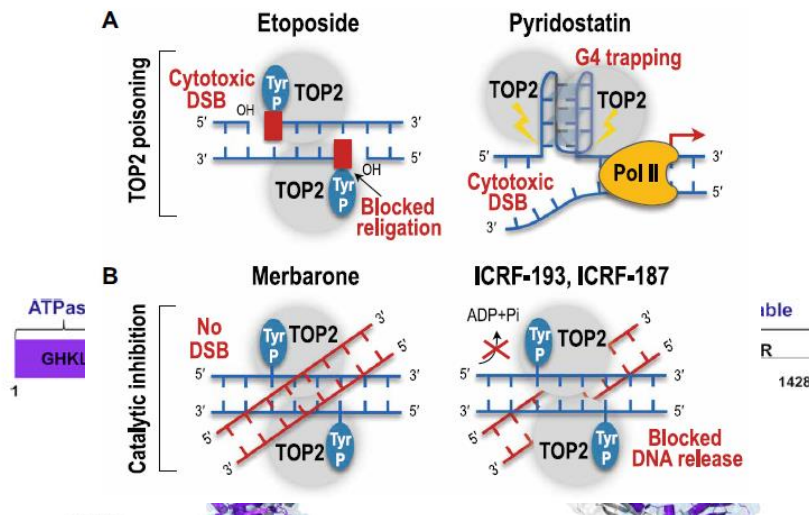


Figure 12 Mechanism of action of TOPOII poisons and inhibitors to treat cancer. (A) TOPOII poisons such as etoposide, which trap TOPOII to 5' termini of DNA DSBs and cause accumulation of TOPOIIcc-blocked DSBs. (B) Merbarone inhibits TOPOII catalytically, stabilizes noncovalent TOPOII-DNA complexes. (image taken from: Uusküla-Reimand L, et al.,. Untangling the roles of TOP2A and TOP2B in transcription and cancer. *Sci Adv.* 2022)



Figure 11 Structure of TOPOII. (Kendra R et al., *Topoisomerase II Poisons: Converting Essential Enzymes into Molecular Scissors.* *Biochemistry.* 2021)

One notable example is merbarone, which was one of the first and most promising topoII inhibitors⁵⁸.

In the group of TOPOII poisons is possible to find etoposide and doxorubicin, which form ternary complexes with TOPOII and DNA and can trap TOPOII to 5' termini of DNA DSBs and inhibit the uptake of broken DNA strands⁵⁹. Etoposide has been considered a poor DNA intercalator. However, it is now clear that the function of etoposide is achieved through the formation of a DNA:TOPOII complex that shows extensive interactions of etoposide with DNA. These class of drugs “poison” TOPOII hindering normal cellular functions and starting DNA damage repair

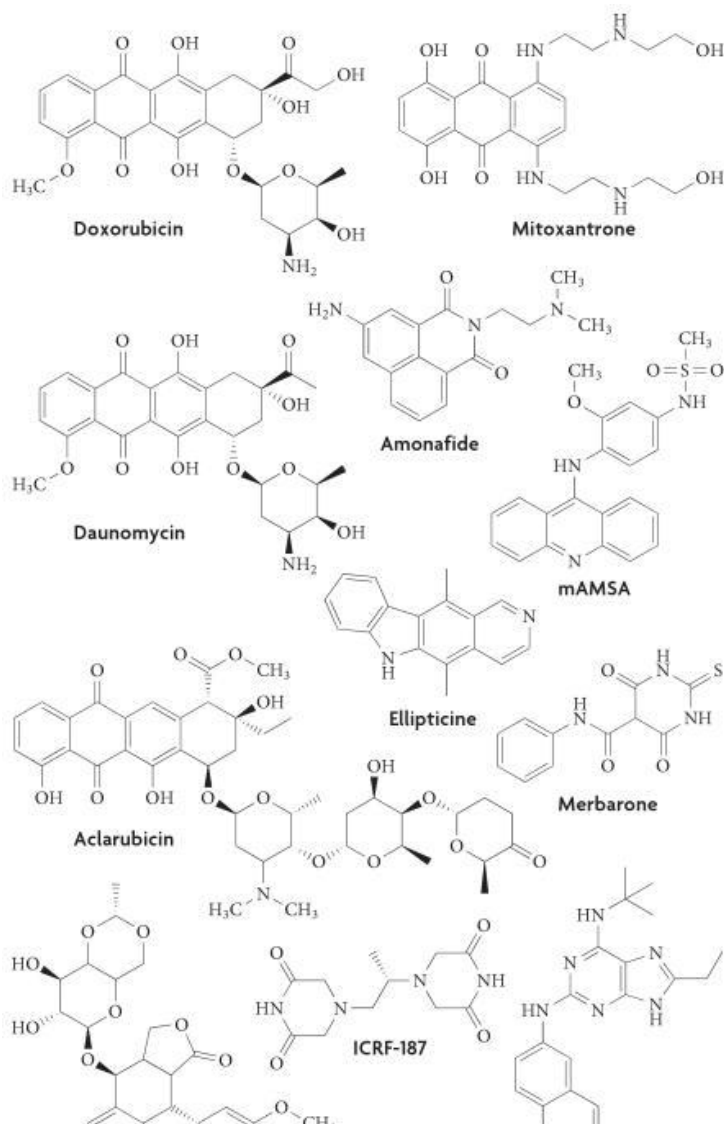


Figure 13 Structures of the main TOPOII inhibitors. (image taken from: Nitiss JL. Targeting DNA topoisomerase II in cancer chemotherapy. *Nat Rev Cancer*. 2009)

(DDR) and repair pathways. The DNA ends bound to the trapped topoisomerase are processed, leading to the accumulation of DSBs. If DSBs are not continuously re-bound, they become 'naked'

and, therefore, cytotoxic. To date, TOPOII poisons can act on both TOPOII α and β . Unfortunately, this specific behaviour has been associated with the development of severe side effects including cardiotoxicity, chromosomal translocations, and secondary malignancies.

On the other hand, TOPOII inhibitors (Error! Reference source not found.) that impairs enzymatic activity by blocking noncovalent catalytic steps in the reaction cycle leads to different biochemical and cellular consequences⁵⁹. An example is merbarone, a thiobarbituric derivative (6-hydroxy-4-oxo-N- phenyl-2-thioxo-1H-pyrimidine-5-carboxamide) that acts on the catalytic activity of TOPOII. In particular, it can block the cleavage of the dsDNA. Notably, merbarone underwent clinical trials as a chemotherapy drug. These trials failed because merbarone displayed nephrotoxicity issues and did not generate the expected efficacy. Other two examples are ICRF-193 and ICRF-187, which limit the ATP hydrolysis step of the catalytic cycle. These inhibitors can block the reopening of the ATPase domain, and keep the protein in the dimerized form, impeding the enzyme turnover and the end of the cycle. Moreover, ICRF-187 is also used in combination with doxorubicin in the clinic to reduce cardiotoxicity and heart failure in patients receiving anthracycline chemotherapy^{59,60}. Recently, also pyridostatin (PDS) and CX5461 were discovered through screenings of genotoxic agents. They resulted to be highly potent poisons. Both inhibitors are small-molecule G4 ligands able to stabilize four-stranded G-quadruplex and R-loop DNA secondary structures by trapping TOPOII. This leads to DNA DSBs and cytotoxicity. Compared to the other TOPOII poisons that non-specifically target both TOPOII α and TOPOII β , PDS acts specifically on TOPOII α . CX5461, instead, is specific for TOPOII β . Therefore, PDS and CX5461 are very promising candidates for anticancer drug development, with CX5461 currently undergoing phase 1 clinical trials for treatment of breast, ovary, pancreas, and prostate cancers⁶⁰. Nevertheless, we developed a new class of merbarone derivatives with a good potency and good pharmacokinetics properties to improve the efficacy of these anticancer already present in literature. In particular, this new class has an hybrid behaviour between the poison and inhibitor. Thanks to this study it was possible to deeply understand the moiety of the small molecules able to drive the antiproliferative activity and the inhibition of TOPOII.

1.14 References

- (1) Wyld, L.; Audisio, R. A.; Poston, G. J. The evolution of cancer surgery and future perspectives. *Nat Rev Clin Oncol* **2015**, *12* (2), 115-124. DOI: 10.1038/nrclinonc.2014.191
- (2) Baskar, R.; Lee, K. A.; Yeo, R.; Yeoh, K. W. Cancer and radiation therapy: current advances and future directions. *Int J Med Sci* **2012**, *9* (3), 193-199. DOI: 10.7150/ijms.3635
- (3) Menck CFM, M. V. DNA repair diseases: What do they tell us about cancer and aging. *Genetics Molecular Biology* **2014**, *37*, 220-233. DOI: 10.1590/S1415-4757201400020000.
- (4) Schirmacher, V. From chemotherapy to biological therapy: A review of novel concepts to reduce the side effects of systemic cancer treatment (Review). *Int J Oncol* **2019**, *54* (2), 407-419. DOI: 10.3892/ijo.2018.4661
- (5) Pomeroy, A. E.; Schmidt, E. V.; Sorger, P. K.; Palmer, A. C. Drug independence and the curability of cancer by combination chemotherapy. *Trends Cancer* **2022**, *8* (11), 915-929. DOI: 10.1016/j.trecan.2022.06.009
- (6) Falzone, L.; Salomone, S.; Libra, M. Evolution of Cancer Pharmacological Treatments at the Turn of the Third Millennium. *Front Pharmacol* **2018**, *9*, 1300. DOI: 10.3389/fphar.2018.01300
- (7) Saini, K. S.; Twelves, C. Determining lines of therapy in patients with solid cancers: a proposed new systematic and comprehensive framework. *Br J Cancer* **2021**, *125* (2), 155-163. DOI: 10.1038/s41416-021-01319-8
- (8) Block, K. I.; Gyllenhaal, C.; Lowe, L.; Amedei, A.; Amin, A.; Amin, A.; Aquilano, K.; Arbiser, J.; Arreola, A.; Arzumanyan, A.; et al. Designing a broad-spectrum integrative approach for cancer prevention and treatment. *Semin Cancer Biol* **2015**, *35* Suppl (Suppl), S276-S304. DOI: 10.1016/j.semcancer.2015.09.007
- (9) Ryman, J. T.; Meibohm, B. Pharmacokinetics of Monoclonal Antibodies. *CPT Pharmacometrics Syst Pharmacol* **2017**, *6* (9), 576-588. DOI: 10.1002/psp4.12224
- (10) Bedard, P. L.; Hyman, D. M.; Davids, M. S.; Siu, L. L. Small molecules, big impact: 20 years of targeted therapy in oncology. *The Lancet* **2020**, *395* (10229), 1078-1088. DOI: 10.1016/s0140-6736(20)30164-1.
- (11) Mansoori, B.; Mohammadi, A.; Davudian, S.; Shirjang, S.; Baradaran, B. The Different Mechanisms of Cancer Drug Resistance: A Brief Review. *Adv Pharm Bull* **2017**, *7* (3), 339-348. DOI: 10.15171/apb.2017.041

- (12) Gatenby, R. A.; Gillies, R. J.; Brown, J. S. Evolutionary dynamics of cancer prevention. *Nat Rev Cancer* **2010**, *10* (8), 526-527. DOI: 10.1038/nrc2892
- (13) Helleday, T.; Petermann, E.; Lundin, C.; Hodgson, B.; Sharma, R. A. DNA repair pathways as targets for cancer therapy. *Nat Rev Cancer* **2008**, *8* (3), 193-204. DOI: 10.1038/nrc2342
- (14) Alhmoud, J. F.; Woolley, J. F.; Al Moustafa, A. E.; Malki, M. I. DNA Damage/Repair Management in Cancers. *Cancers (Basel)* **2020**, *12* (4). DOI: 10.3390/cancers12041050
- (15) Sreedhara, A.; Cowan, J. A. Structural and catalytic roles for divalent magnesium in nucleic acid biochemistry. *Biometals* **2002**, *15* (3), 211-223. DOI: 10.1023/a:1016070614042
- (16) Sissi, C.; Palumbo, M. Effects of magnesium and related divalent metal ions in topoisomerase structure and function. *Nucleic Acids Research* **2009**, *37* (3), 702-711. DOI: 10.1093/nar/gkp024.
- (17) Genna, V.; Gaspari, R.; Dal Peraro, M.; De Vivo, M. Cooperative motion of a key positively charged residue and metal ions for DNA replication catalyzed by human DNA Polymerase- ϵ . *Nucleic Acids Res* **2016**, *44* (6), 2827-2836. DOI: 10.1093/nar/gkw128
- (18) Deweese, J. E.; Osheroff, N. The use of divalent metal ions by type II topoisomerases. *Metallomics* **2010**, *2* (7), 450-459. DOI: 10.1039/c003759a
- (19) Watson, J. D.; Crick, F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **1953**, *171* (4356), 737-738. DOI: 10.1038/171737a0
- (20) Neuman, K. C. Single-molecule measurements of DNA topology and topoisomerases. *J Biol Chem* **2010**, *285* (25), 18967-18971. DOI: 10.1074/jbc.R109.092437
- (21) Meselson, M.; Stahl, F. W. The Replication of DNA in Escherichia Coli. *Proc Natl Acad Sci U S A* **1958**, *44* (7), 671-682. DOI: 10.1073/pnas.44.7.671
- (22) Wang, J. C. Interaction between DNA and an Escherichia coli protein omega. *J Mol Biol* **1971**, *55* (3), 523-533. DOI: 10.1016/0022-2836(71)90334-2
- (23) Champoux, J. J. DNA topoisomerases: structure, function, and mechanism. *Annu Rev Biochem* **2001**, *70*, 369-413. DOI: 10.1146/annurev.biochem.70.1.369
- (24) Deweese, J. E.; Osheroff, M. A.; Osheroff, N. DNA Topology and Topoisomerases: Teaching a "Knotty" Subject. *Biochem Mol Biol Educ* **2008**, *37* (1), 2-10. DOI: 10.1002/bmb.20244
- (25) Vann, K. R.; Oviatt, A. A.; Osheroff, N. Topoisomerase II Poisons: Converting Essential Enzymes into Molecular Scissors. *Biochemistry* **2021**, *60* (21), 1630-1641. DOI: 10.1021/acs.biochem.1c00240.

- (26) Reichrath, J.; Rass, K. Ultraviolet damage, DNA repair and vitamin D in nonmelanoma skin cancer and in malignant melanoma: an update. *Adv Exp Med Biol* **2014**, *810*, 208-233. DOI: 10.1007/978-1-4939-0437-2_12
- (27) Chou, K. M. DNA polymerase eta and chemotherapeutic agents. *Antioxid Redox Signal* **2011**, *14* (12), 2521-2529. DOI: 10.1089/ars.2010.3673
- (28) Hubscher, U.; Maga, G.; Spadari, S. Eukaryotic DNA polymerases. *Annu Rev Biochem* **2002**, *71*, 133-163. DOI: 10.1146/annurev.biochem.71.090501.150041
- (29) Sale, J. E. Translesion DNA synthesis and mutagenesis in eukaryotes. *Cold Spring Harb Perspect Biol* **2013**, *5* (3), a012708. DOI: 10.1101/cshperspect.a012708 .
- (30) Sale, J. E.; Lehmann, A. R.; Woodgate, R. Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nat Rev Mol Cell Biol* **2012**, *13* (3), 141-152. DOI: 10.1038/nrm3289
- (31) Zhao, L.; Washington, M. T. Translesion Synthesis: Insights into the Selection and Switching of DNA Polymerases. *Genes (Basel)* **2017**, *8* (1). DOI: 10.3390/genes8010024
- (32) Boiteux, S.; Jinks-Robertson, S. DNA repair mechanisms and the bypass of DNA damage in *Saccharomyces cerevisiae*. *Genetics* **2013**, *193* (4), 1025-1064. DOI: 10.1534/genetics.112.145219
- (33) Ohmori, H.; Friedberg, E. C.; Fuchs, R. P.; Goodman, M. F.; Hanaoka, F.; Hinkle, D.; Kunkel, T. A.; Lawrence, C. W.; Livneh, Z.; Nohmi, T.; et al. The Y-family of DNA polymerases. *Mol Cell* **2001**, *8* (1), 7-8. DOI: 10.1016/s1097-2765(01)00278-7
- (34) Yang, W. An overview of Y-Family DNA polymerases and a case study of human DNA polymerase eta. *Biochemistry* **2014**, *53* (17), 2793-2803. DOI: 10.1021/bi500019s
- (35) Kuzminov, A. Inhibition of DNA synthesis facilitates expansion of low-complexity repeats: is strand slippage stimulated by transient local depletion of specific dNTPs? *Bioessays* **2013**, *35* (4), 306-313. DOI: 10.1002/bies.201200128
- (36) Phillips, D. H. Fifty years of benzo(a)pyrene. *Nature* **1983**, *303* (5917), 468-472. DOI: 10.1038/303468a0
- (37) van Loon, B.; Markkanen, E.; Hubscher, U. Oxygen as a friend and enemy: How to combat the mutational potential of 8-oxo-guanine. *DNA Repair (Amst)* **2010**, *9* (6), 604-616. DOI: 10.1016/j.dnarep.2010.03.004

- (38) Alt, A.; Lammens, K.; Chiocchini, C.; Lammens, A.; Pieck, J. C.; Kuch, D.; Hopfner, K. P.; Carell, T. Bypass of DNA lesions generated during anticancer treatment with cisplatin by DNA polymerase η . *Science* **2007**, *318* (5852), 967-970. DOI: 10.1126/science.1148242
- (39) Saha, P.; Mandal, T.; Talukdar, A. D.; Kumar, D.; Kumar, S.; Tripathi, P. P.; Wang, Q. E.; Srivastava, A. K. DNA polymerase η : A potential pharmacological target for cancer therapy. *J Cell Physiol* **2021**, *236* (6), 4106-4120. DOI: 10.1002/jcp.30155
- (40) Wang, J.; Smithline, Z. B. Crystallographic evidence for two-metal-ion catalysis in human pol η . *Protein Sci* **2019**, *28* (2), 439-447. DOI: 10.1002/pro.3541
- (41) Yang, W.; Lee, J. Y.; Nowotny, M. Making and breaking nucleic acids: two-Mg²⁺-ion catalysis and substrate specificity. *Mol Cell* **2006**, *22* (1), 5-13. DOI: 10.1016/j.molcel.2006.03.013
- (42) Gregory, M. T.; Park, G. Y.; Johnstone, T. C.; Lee, Y. S.; Yang, W.; Lippard, S. J. Structural and mechanistic studies of polymerase η bypass of phenanthriplatin DNA damage. *Proc Natl Acad Sci U S A* **2014**, *111* (25), 9133-9138. DOI: 10.1073/pnas.1405739111
- (43) Feltes, B. C.; Menck, C. F. M. Current state of knowledge of human DNA polymerase η protein structure and disease-causing mutations. *Mutat Res Rev Mutat Res* **2022**, *790*, 108436. DOI: 10.1016/j.mrrev.2022.108436
- (44) Kannouche, P.; Fernandez de Henestrosa, A. R.; Coull, B.; Vidal, A. E.; Gray, C.; Zicha, D.; Woodgate, R.; Lehmann, A. R. Localization of DNA polymerases η and ι to the replication machinery is tightly co-ordinated in human cells. *EMBO J* **2003**, *22* (5), 1223-1233. DOI: 10.1093/emboj/cdf618
- (45) Temviriyankul, P.; van Hees-Stuivenberg, S.; Delbos, F.; Jacobs, H.; de Wind, N.; Jansen, J. G. Temporally distinct translesion synthesis pathways for ultraviolet light-induced photoproducts in the mammalian genome. *DNA Repair (Amst)* **2012**, *11* (6), 550-558. DOI: 10.1016/j.dnarep.2012.03.007
- (46) Haracska, L.; Unk, I.; Prakash, L.; Prakash, S. Ubiquitylation of yeast proliferating cell nuclear antigen and its implications for translesion DNA synthesis. *Proc Natl Acad Sci U S A* **2006**, *103* (17), 6477-6482. DOI: 10.1073/pnas.0510924103
- (47) Johnson, R. E.; Prakash, S.; Prakash, L. Requirement of DNA polymerase activity of yeast Rad30 protein for its biological function. *J Biol Chem* **1999**, *274* (23), 15975-15977. DOI: 10.1074/jbc.274.23.15975

- (48) Johnson, R. E.; Haracska, L.; Prakash, S.; Prakash, L. Role of DNA polymerase eta in the bypass of a (6-4) TT photoproduct. *Mol Cell Biol* **2001**, *21* (10), 3558-3563. DOI: 10.1128/MCB.21.10.3558-3563.2001
- (49) Kobayashi, J.; Okui, M.; Asaithamby, A.; Burma, S.; Chen, B. P.; Tanimoto, K.; Matsuura, S.; Komatsu, K.; Chen, D. J. WRN participates in translesion synthesis pathway through interaction with NBS1. *Mech Ageing Dev* **2010**, *131* (6), 436-444. DOI: 10.1016/j.mad.2010.06.005
- (50) Prakash, S.; Johnson, R. E.; Prakash, L. Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annu Rev Biochem* **2005**, *74*, 317-353. DOI: 10.1146/annurev.biochem.74.082803.133250
- (51) Havener, J. M.; Nick McElhinny, S. A.; Bassett, E.; Gauger, M.; Ramsden, D. A.; Chaney, S. G. Translesion synthesis past platinum DNA adducts by human DNA polymerase mu. *Biochemistry* **2003**, *42* (6), 1777-1788. DOI: 10.1021/bi0270079
- (52) Deweese, J. E.; Osheroff, N. The DNA cleavage reaction of topoisomerase II: wolf in sheep's clothing. *Nucleic Acids Res* **2009**, *37* (3), 738-748. DOI: 10.1093/nar/gkn937
- (53) Liu, Z.; Deibler, R. W.; Chan, H. S.; Zechiedrich, L. The why and how of DNA unlinking. *Nucleic Acids Res* **2009**, *37* (3), 661-671. DOI: 10.1093/nar/gkp041
- (54) Wang, J. C. DNA topoisomerases. *Annu Rev Biochem* **1996**, *65*, 635-692. DOI: 10.1146/annurev.bi.65.070196.003223
- (55) Chen, S. H.; Chan, N. L.; Hsieh, T. S. New mechanistic and functional insights into DNA topoisomerases. *Annu Rev Biochem* **2013**, *82*, 139-170. DOI: 10.1146/annurev-biochem-061809-100002
- (56) Elton, T. S.; Ozer, H. G.; Yalowich, J. C. Effects of DNA topoisomerase IIalpha splice variants on acquired drug resistance. *Cancer Drug Resist* **2020**, *3* (2), 161-170. DOI: 10.20517/cdr.2019.117
- (57) Pommier, Y. Drugging topoisomerases: lessons and challenges. *ACS Chem Biol* **2013**, *8* (1), 82-95. DOI: 10.1021/cb300648v
- (58) Arencibia, J. M.; Brindani, N.; Franco-Ulloa, S.; Nigro, M.; Kuriappan, J. A.; Ottonello, G.; Bertozzi, S. M.; Summa, M.; Giroto, S.; Bertorelli, R.; et al. Design, Synthesis, Dynamic Docking, Biochemical Characterization, and in Vivo Pharmacokinetics Studies of Novel Topoisomerase II

Poisons with Promising Antiproliferative Activity. *J Med Chem* **2020**, *63* (7), 3508-3521. DOI: 10.1021/acs.jmedchem.9b01760

(59) Nitiss, J. L. Targeting DNA topoisomerase II in cancer chemotherapy. *Nat Rev Cancer* **2009**, *9* (5), 338-350. DOI: 10.1038/nrc2607

(60) Uuskula-Reimand, L.; Wilson, M. D. Untangling the roles of TOP2A and TOP2B in transcription and cancer. *Sci Adv* **2022**, *8* (44), eadd4920. DOI: 10.1126/sciadv.add4920

Chapter 2: Computer-aided identification, synthesis, and biological evaluation of DNA polymerase η inhibitors for the treatment of cancer

Federico Munafò,^{1,#} Michela Nigro,^{1,#} Nicoletta Brindani,^{1,#} Jacopo Manigrasso,¹

Inacrist Geronimo,¹ Giuliana Ottonello,² Andrea Armirotti,² Marco De Vivo^{1,*}

#Equally contributed.

This manuscript has been published on the European Journal of Medicinal Chemistry in the 2023 Jan 2;248:115044. doi: 10.1016/j.ejmech.2022.115044. PMID: 36621139.

Supporting information:

<https://ars.els-cdn.com/content/image/1-s2.0-S0223523422009461-mmc1.pdf>

Please note that the compound **21** will correspond to compound 64 (ARN24964), and compound **17** will correspond to compound 61 (ARN24961).

2.1 Introduction

DNA-damaging chemotherapeutics, such as cisplatin, are often used to treat cancer patients. These agents generate DNA lesions that cannot be handled by high-fidelity replicative DNA polymerases. Thus, such chemotherapeutics block DNA synthesis, ultimately leading to cancer cell death. However, cancer cells are equipped with a DNA repair machinery that can remove such lesions before they become toxic.¹⁻⁴ At the clinical level, this translates in the development of resistance to the pharmacological treatment. In turn, this problem requires either an increase in drug dosage, with a rise of possible side effects, or a change in drug regimen. The development of chemoresistance thus limits both the short and long-term effectiveness of genotoxic agents.⁵⁻⁷

Chemoresistance involves different mechanisms, such as translesion DNA synthesis (TLS) and homologous repair (HR) systems, which in normal conditions allow the normal replication process.^{1,8} TLS involves several specialized DNA polymerases belonging to the Y-family, which are characterized by a lack of 3'-5' proofreading activity, a low fidelity, a low processivity, and low catalytic efficiency. On the other hand, these polymerases have a spacious active site that can accommodate a wide range of DNA damage types and thereby bypass the lesions induced by chemotherapies.⁹⁻¹² This particular class has indeed a special role assisting the replication across damaged regions, which has been largely characterized over the years.¹³⁻¹⁷

In this study, we target DNA Polymerase η (Pol η), which is exceptional in its ability to bypass specific DNA damages like cyclobutane pyrimidine dimers (CPDs), often caused by UV light.^{18,19} This specific activity of Pol η is therefore beneficial in healthy cells. However, the ability of such Pol η to bypass lesions along the DNA is also known to be involved in chemoresistance. For example, Pol η allows DNA replication even when DNA is damaged by anticancer drugs like cisplatin and its derivatives.²⁰⁻²² As a matter of fact, cancer patients undergoing cisplatin or related chemotherapies at times develop elevated Pol η expression, which can lead to chemo-resistance issues.²³⁻²⁵ Several studies have identified multiple instances where Pol η was shown to contribute to cancer progression and/or treatment. It has been demonstrated that high expression of Pol η is directly correlated to cisplatin resistance in the case of ovarian cancer, and inversely correlated to survival of patients treated with platinum drugs in the cases of lung and gastric adenocarcinomas.^{20,23,24} Also, low expression of Pol η was associated with positive response rate in neck mucosal-derived squamous cell carcinoma patients treated with cisplatin.²⁵ Thus, inhibition of Pol η is now

considered a promising strategy to render platinum-based treatments more effective by limiting the emergence of chemo-resistant tumor cells.^{15, 17, 26}

In this regard, in 2009, the aurintricarboxylic acid (ATCA) and the ellagic acid were identified as Pol η inhibitors. These compounds exhibit very promising nanomolar IC₅₀ values for Pol η inhibition (75 nM and 62 nM for ATCA and ellagic acid, respectively).²⁷ However, they have been shown not to be target specific, being potent inhibitor also of other nucleic-acid processing enzymes such as topoisomerases and ribonuclease. In particular, ATCA polymerizes in aqueous environment forming a stable free radical that inhibits protein-nucleic acid interactions.²⁸⁻³⁰ To the best of our knowledge, the most potent class of Pol η inhibitors is structurally related to the *N*-benzoyl indolylbarbituric acid (ITBA).^{31, 32} Steady-state kinetic experiments showed that ITBA compounds likely act through an allosteric mechanism.³² This hypothesis was further confirmed by molecular docking studies suggesting that ITBA compounds act by binding to a pocket between the finger and little finger domains in Pol η .³² Notably, one of the most potent compounds belonging to this class is ITBA-16, with an IC₅₀ of $16 \pm 3 \mu\text{M}$.³² The synthesis of close derivatives led to the generation of a second class of inhibitors with IC₅₀ in the low micromolar range and to the identification of PNR-02, which exhibited an IC₅₀ of $8 \mu\text{M}$.³¹ More recently, X-ray crystallography was used to screen a library of approximately 300 fragments, resulting in the identification of few compounds active in the submillimolar range against Pol η , with IC₅₀ values ranging from $230 \mu\text{M}$ to 8mM .³³

Our aim was the identification of new Pol η inhibitors, in the search of new anticancer therapies capable to overcome drug resistance to platinum-containing drugs. Also, several recent co-crystal structures of Pol η , solved with either undamaged DNA or platinum adducts at its metal-aided catalytic center, have opened to the possibility of structure-based drug design.²² Based on these premises, here we present our computer-aided structure-based effort to design, synthesize and develop a new class of Pol η inhibitors. We report 35 derivatives based on our initial chromone hit. Through an extensive structure-activity relationship and characterization of our compounds performed via molecular modeling and simulations, and *in vitro* experiments, we have identified the most interesting compound, ARN24964, which represents a promising scaffold to pursue further optimization efforts toward the identification of a drug candidate.

2.2 Results and discussion

Static and molecular dynamics-based docking for hit identification. To identify compounds targeting the two-metal-ion (Mg^{2+}) active site of Pol- η , we first performed a virtual screening campaign of the D3 internal library, which contains ~15,000 drug-like molecules readily available for experimental testing (see methods). Based on docking scores and poses, twenty molecules were selected for *in vitro* testing as Pol η inhibitors. Toward this end, we used a gel-based DNA primer elongation assay (see methods). In this way, we identified the compound ARN18347 (racemic mixture, *rac*), which completely inhibits Pol η at 100 μ M and showed a 22.8% of inhibition at 1 μ M (**Figure 14**).

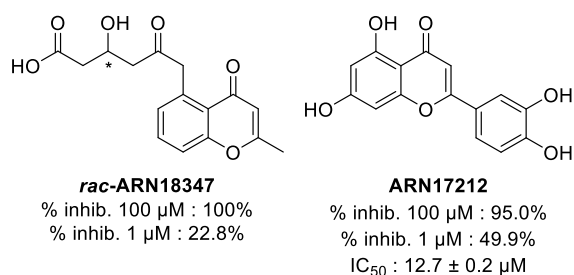


Figure 14 Structure and potency of *rac*-ARN18347 and ARN17212.

Notably, the predicted binding pose of ARN18347 at the catalytic site (including both ions, MgA-MgB) is very similar to the pose of the Pol η 's endogenous substrate ATP, at the transition state during catalysis (**Figure 15 A-B**) – as identified by previous computational studies from our group.³⁴ Specifically, we found that the aromatic scaffold of ARN18347 establishes a π - π interaction with the 3' end of the DNA primer strand, well mimicking the interaction formed by the ATP nucleobase when bound at the catalytic site.³⁵ In this orientation, the side-chain of ARN18347 could coordinate with MgB, via the hydroxyl and carboxyl groups, similarly to what observed for the α , β , and γ phosphates of the ATP substrate, as showed in Pol η structures (e.g., PDBid: 4ECR).³⁵ Based on such structural features for the binding of ARN18347, we then performed an atom-pair Tanimoto similarity screen of the D3 library (see methods) to identify analogous compounds. This search led to the identification of a naturally occurring flavone that, we noted, has a more easily tractable chemical scaffold. This was ARN17212, a flavone known as luteolin. As compared to ARN18347, ARN17212 returned a better inhibition at 1 μ M, with an IC₅₀ of 12.7 \pm 0.2 μ M (**Figure 14**). This was therefore elected our hit compound for further exploration.

ARN17212 belongs to a group of natural compounds, called flavonoids, which are widely diffused in plants and are known to have several beneficial effects on human health.³⁶⁻³⁸ Chemically, ARN17212 features a 1,4-benzopyrone (or chromone) core, which is formed by fused rings A and C. It bears an additional ring in position 2, called ring B. Furthermore, ring A is characterized by a 5,7-dihydroxy substitution pattern and ring B is characterized by a 3',4'-dihydroxy substitution pattern.

We performed extensive static and dynamic (molecular dynamics MD-based) docking analyses of ARN17212 binding at the catalytic reaction center of Pol η , containing both MgA and MgB ions. These computations allowed us to identify a series of different and congeneric binding modes. Two exemplifying poses are reported in **Figure 15** C-D. In all the cases, ARN17212 establishes close contacts with both the catalytic metals, MgA-MgB, similarly to what observed for the endogenous ATP substrate in existing Pol- η structures (e.g., PDBid: 4ECR).³⁵ Specifically, we found that ARN17212 can assume multiple orientations to bind the metal cluster via two different coordination modes. That is, mode 1, where the chromone moiety coordinates both MgA and MgB (**Figure 15C**); and mode 2, where both 3'- and 4'-hydroxyl groups of the ring B coordinate only the MgB (**Figure 15D**).

To determine whether one coordination mode of ARN17212 was preferred over the other, we also performed unbiased MD simulations of the Pol η /DNA/ARN17212 ternary complex (**Figure 16**). Binding mode 2 proved to be unstable as ARN17212 dissociated from MgB after only approximately 0.5 ns of MD simulation. On the other hand, in binding mode 1, only the coordination between the 4-oxo and MgB remained stable throughout the 500-ns MD simulation (**Figure 16**). In addition to this coordination, ARN17212 was also stabilized by the multiple interactions of ring B with the target. Namely, the H-bonding of 4'-hydroxyl with I48 and the templating base T(0) (18.6% and 11.2% occupancy, respectively), H-bonding of 3'-hydroxyl with T(0) (20.9% occupancy), cation- π interaction with R61, and π - π interaction with the terminal primer residue P(-1).

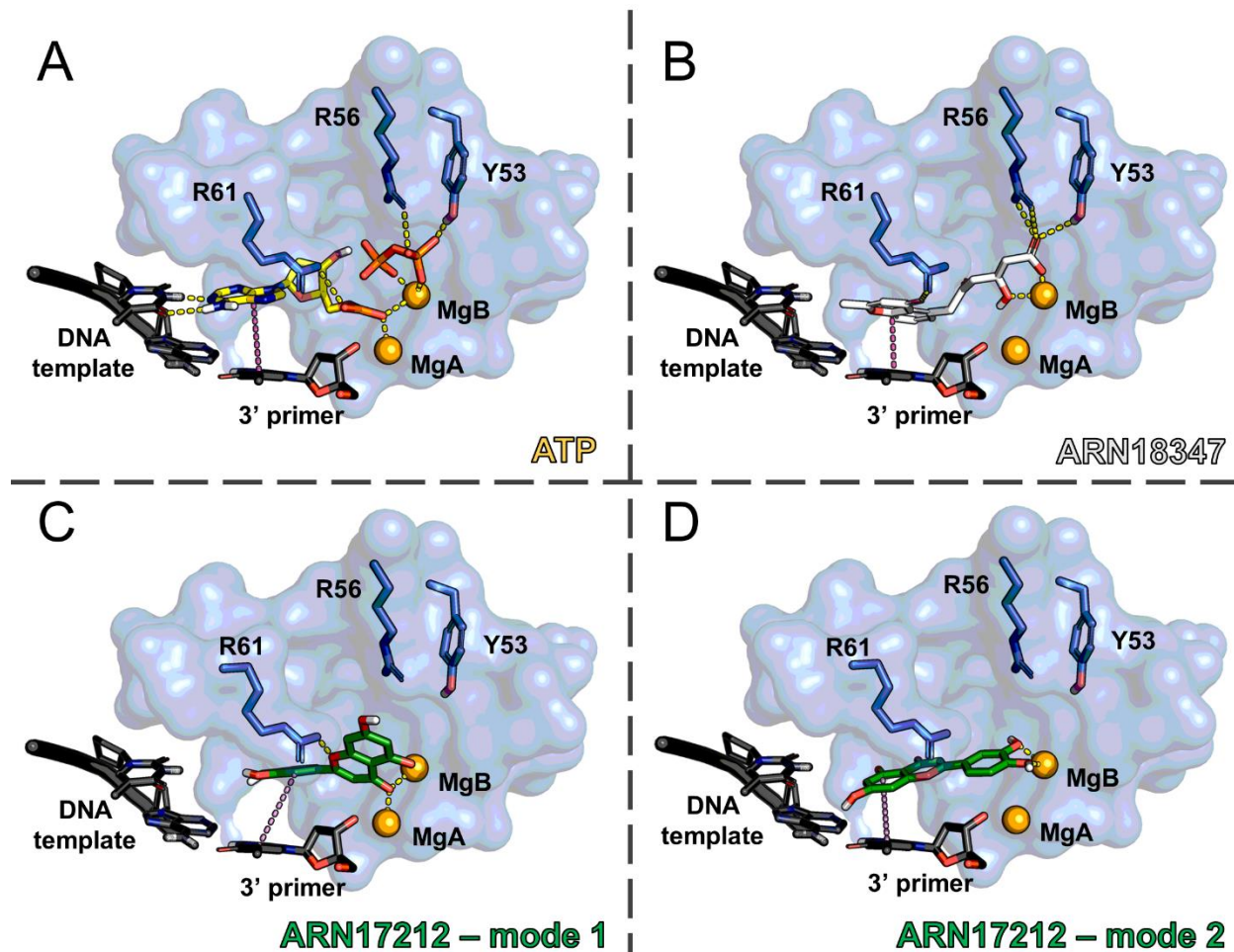


Figure 15 Interactions of ligands with Pol- η active site.³⁴ A) The binding mode of ATP at the transition state as obtained by previous quantum-based computations is represented.³⁴ ATP is shown as yellow sticks, while the ligand pocket is represented as a blue surface. Here, the DNA template and primer (black), and key Pol- η functional residues (blue) are shown as sticks, while the two catalytic metals are highlighted as orange spheres. Dashed lines represent H-bonds and metal coordinations (yellow), as well as π - π stacking (magenta) interactions. B) A representative binding mode of ARN18347 as obtained by molecular docking. The coloring scheme follows that of panel A. C-D) Two representative coordination modes of ARN17212 as obtained by molecular docking. The coloring scheme follows that of panel A.

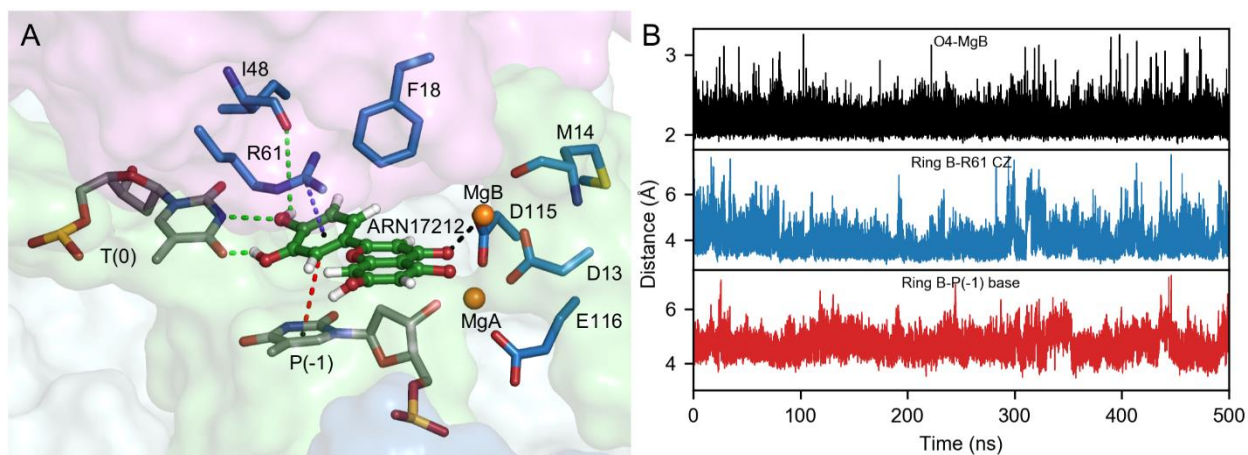


Figure 16 . (A) Interactions of ARN17212 with the Pol η active site. The O4–MgB coordination, H-bond interactions with I48 and the templating base T(0), cation- π interaction with R61, and π - π interaction with the terminal primer residue P(-1) are represented by black, green, blue, and red dashed lines, respectively. (B) Time evolution of the O4–MgB coordination and π interactions of ring B of ARN17212 with R61 and P(-1). The π interactions were monitored by measuring the center-of-mass distances between the rings (CZ atom for R61).

Design and optimization of ARN17212 new analogs. We have explored the chemical scaffold of the hit compound ARN17212, **1**, ($IC_{50} = 12.7 \pm 0.2 \mu M$) through 34 differently substituted flavones, exploring the sequential functionalization of each ring, namely A, B, and C (**Figure 17**).

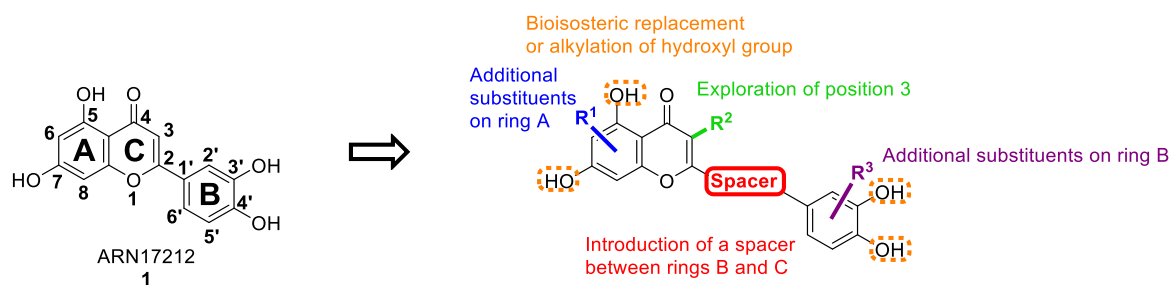


Figure 17 Chemical structure of ARN17212 and its possible scaffold exploration.

At first, we explored the role of hydroxyl groups on ring A generating compounds **2**, **4** and **6** (Table 1). We also considered possible substitutions with a methoxy and fluorine group in compounds **3** and **5**. With a similar rationale, we then moved our attention to ring B, where we tested the impact of removal of one or more hydroxyl group(s) in 3' and 4' positions, with compounds **7**, **8** and **9**. We also considered their substitution with benzodioxole in compounds **10** and **11**, and a salicylic moiety in compound **12**. Moreover, we investigated the substitution of the 2' position in ring B,

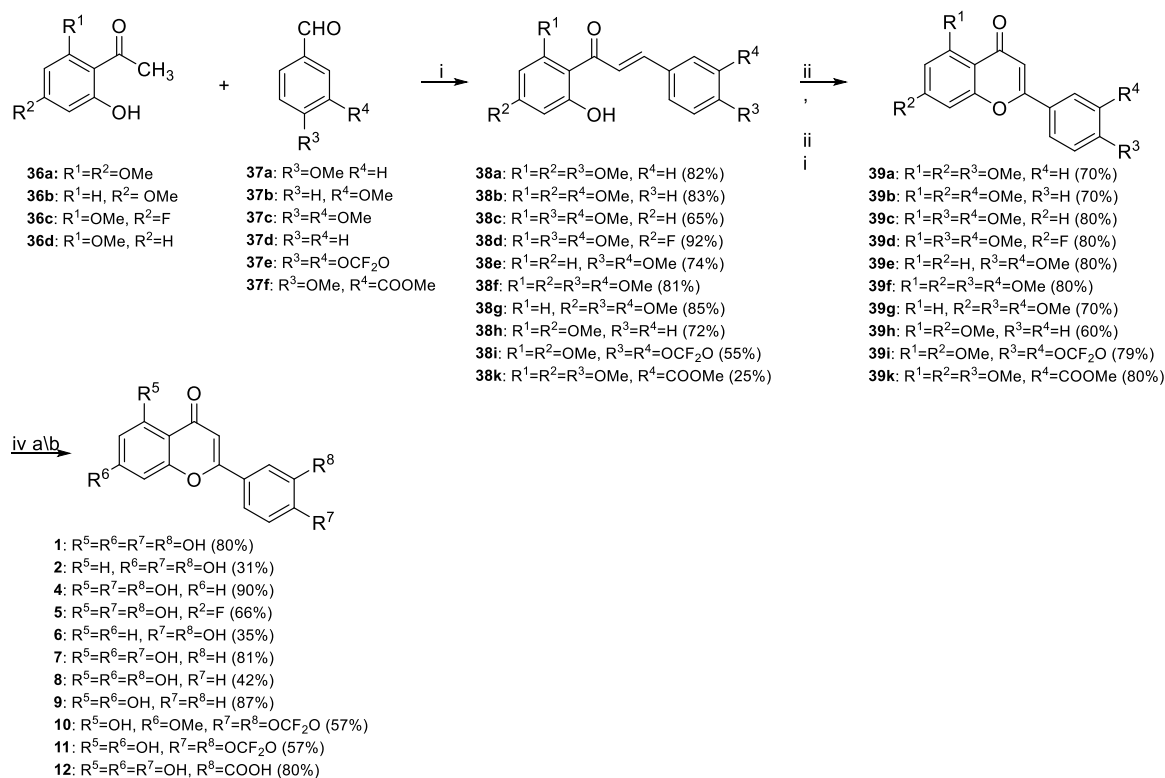
adding lipophilic groups combined with modifications in ring A, with compounds **13-16**. We also evaluated the insertion of an additional pendant in 3 position of ring C as for derivatives **17-24**. We further expanded this panel of in-house synthesized flavones with additional seven close analogs retrieved from our internal library, i.e. compounds **25-31** (Table 2), which include some well-known natural polyphenolic compounds such as quercetin, **26**. Finally, we synthesized a set of four analogs bearing a spacer with a different degree of flexibility between rings B and C, in compounds **32-35** (Table 3). In this way, we were able to test differently substituted flavones and explore the chemical structure of the initial hit **1**, as discussed hereafter.

Chemistry. Most of the synthesized flavones were obtained through a 3-step synthetic route, consisting by a starting aldol condensation followed by an oxidative cyclization as key steps for the formation of the flavonoid core (Schemes 1, 2, 6). Some exceptions are represented by compounds **3**, **17-21**, and **32-35**. In particular, **3** was prepared starting directly from **1** (Scheme 4), **17-21** were obtained starting from commercially available rutin (Scheme 5), and for **32-35** the chromone core was introduced with different synthetic approaches based on the nature of the spacer (Scheme 7).

As depicted in Scheme 1, analogues **1**, **2**, and **4-12** were obtained through a 3-step synthetic route, consisting by an aldol condensation between 2-hydroxyacetophenone **36a-d** and suitable substituted benzaldehyde **37a-f**, with generation of chalcone intermediates **38a-k** in 65-86% yields. Chalcones underwent to an oxidative cyclization with catalytic amounts of iodine in DMSO at high temperature.

Except for compounds **5**, **10** and **11**, we observed the formation of small amounts of a 3-iodinated byproduct, which formed an inseparable mixture with the desired product, possibly due to π - π stacking interactions between the two aromatic flat molecules. To circumvent this problem, the crude mixture was treated with sodium formate in the presence of tetrakis(triphenylphosphine)palladium(0) in DMF at 100 °C, obtaining pure compounds **39a-k** in 60-80% yields over two steps. Then final deprotection with pyridium chloride at 190 °C gave final product **1**, **2** and **4**, **5**, **7-9**, **11**, **12** with yields that ranged from 30% to 87%. A different procedure with boron tribromide in DCM was used for the obtainment of compounds **6** and **10** from related direct precursors **39e** and **39i** with 35-57% yields.

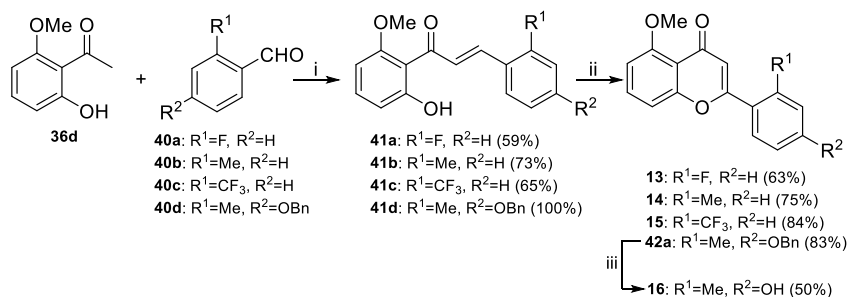
Scheme 1^a. Synthesis of compounds **1**, **2** and **4-12**.



a) (i) KOH, MeOH, r.t. (ii) I₂ cat., DMSO, 135 °C, (iii) NaHCO₂, Pd(PPh₃)₄, DMF dry, 100 °C, (iv) (a) pyridinium chloride, 190 °C (b) BBr₃, DCM dry, -78 °C/40 °C.

The synthesis of compounds **13-16** is shown in Scheme 2. The same sequential transformation of aldol condensation, iodine-mediated cyclization and final deprotection starting from acetophenone **36d** and benzaldehydes **40a-d** gave access to final compounds **13-16**, with nice 20-69% overall yield.

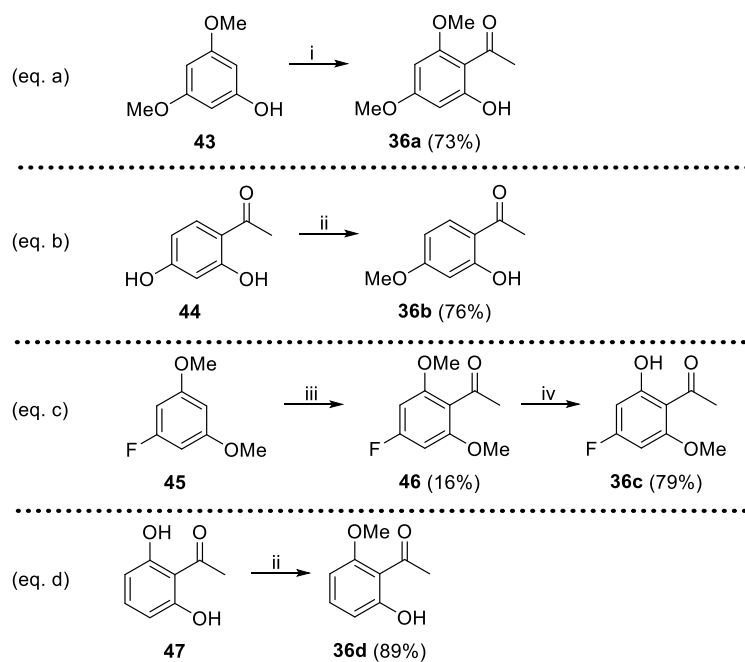
Scheme 2^a. Synthesis of compounds **13-16**.



a) (i) KOH, MeOH, r.t. (ii) I₂ cat., DMSO, 135 °C (iii) Et₃SiH, Pd/C, DCM/MeOH 1:1, r.t.

In particular, 2-hydroxyacetophenones **36a-d**, used for the preparation of compounds **1**, **2** and **4-16**, were prepared with different synthetic routes outlined in Scheme 3 (eq. a-d). The Fries rearrangement between 3,5-dimethoxyphenol **43** and ethyl acetate in the presence of boron trifluoride etherate and acetic acid afforded the desired product **36a** in a 73% yield (eq. a). Monomethylation with iodomethane and potassium carbonate of **44** and **47** produced desired compounds **36b** and **36d** in 89% and 76% yields, respectively (eq. b, d). Lastly, Friedel-Craft acylation of **45** with acetyl chloride and aluminum trichloride afforded compound **46** as a minor isomer in a yield of 17%. The subsequent monodemethylation with boron tribromide in DCM afforded the desired product **36c** in 79% yield (eq. c).

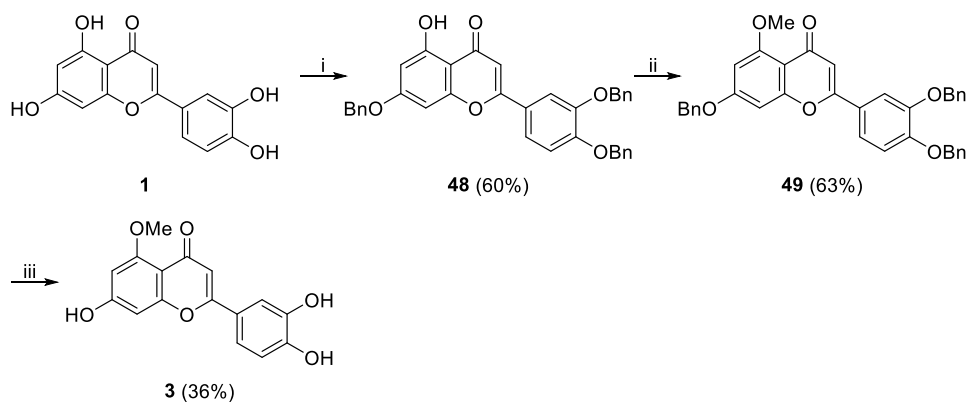
Scheme 3^a. Synthesis of compounds 36a-d.



a) (i) $\text{BF}_3 \cdot \text{Et}_2\text{O}$, AcOH, EtOAc dry, 85 °C (ii) MeI, K_2CO_3 , acetone, reflux, (iii) CH_3COCl , AlCl_3 , DCM, r.t. (iv) BBr_3 , DCM dry, -60°C/-20 °C.

As shown in Scheme 4, in order to obtain analogue **3** with three free hydroxyl group in position 7, 3' and 4' and the methoxy group in position 5, we exploited a different strategy of selective protection/deprotection steps starting from our hit compound **1**. Hydroxyl groups on positions 7, 3' and 4' of compound **1** were selectively protected with benzyl groups in a good 60% yield. Notably, benzylation of the 5-OH group was not observed. At this point, hydroxyl group on position 5 of ring A was methylated in a 63% yield by treating the tribenzylated intermediate **48** with iodomethane and potassium carbonate in DMF at room temperature. Finally, removal of three benzyl groups of **49** with triethyl silane and palladium on carbon in a 1:1 mixture of DCM and methanol afforded the desired product **3** in 36% yield.

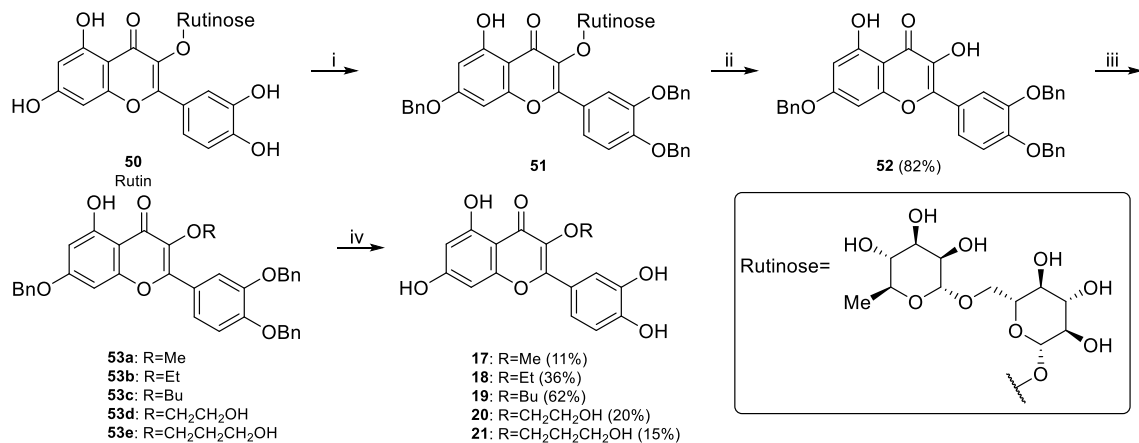
Scheme 4^a. Synthesis of compound **3**.



a) (i) BnBr, K₂CO₃, DMF, r.t. (ii) MeI, K₂CO₃, DMF, r.t. (iii) Et₃SiH, Pd/C, DCM/MeOH 1:1, r.t.

As depicted in Scheme 5, the insertion of alkoxy chain of compounds **17-21** was allowed by a synthetic strategy that started from the commercially available glycoside **50** (rutin). Rutin possesses the oxygen in position 3 already installed as summarized in Scheme 5. Benzylation of rutin with benzyl bromide in the presence of potassium carbonate in DMF at room temperature, followed by the hydrolysis of glycosidic bond with HCl conc. in ethanol under reflux, led mainly to the formation of tetrabenzylated product **52** with unmasked hydroxyl group in 3 position with a yield of 85%. This was alkylated with the proper alkyl bromide in the presence of potassium carbonate in DMF at room temperature. Lastly, cleavage of the benzyl group with triethylsilane and palladium on carbon in a 1:1 mixture of DCM and methanol afforded final compounds **17-21** in 15-63% yields over two steps.

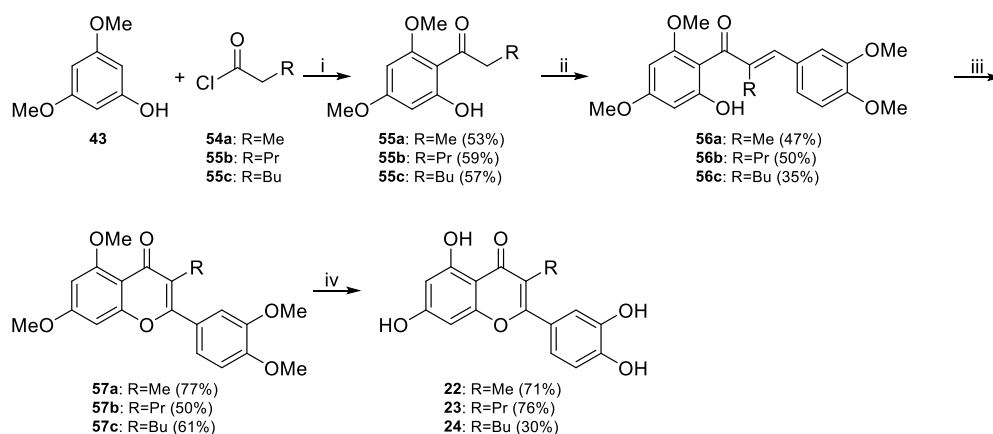
Scheme 5^a. Synthesis of compounds **17-21**.



a) (i) BnBr, K₂CO₃, DMF, r.t. (ii) HCl 37%, EtOH, reflux (iii) RBr, K₂CO₃, DMF, r.t., (iv) Et₃SiH, Pd/C, DCM/MeOH 1:1, r.t.

The preparation of compounds **22-24** with several alkyl chains in 3 position started from 3,5-dimethoxyphenol **43** and proper acyl chlorides **54a-c** with the specific chain of final target compound, already embedded (Scheme 6). The sequential transformation of Friedel-Craft acylation of **43**, in the presence of aluminum trichloride as Lewis acid in DCM, afforded intermediates **55a-c** in 57-59% yield, which underwent to aldol condensation with vetraldehyde in basic conditions (KOH in MeOH), giving chalcone intermediates **56a-c** in 50-60% yields. Then, the oxidative iodine mediated cyclization of **56a-c** produced intermediates **57a-c** in good 60-70% yields. Notably, the alkyl chain on position 3 prevented the formation of the 3-iodinated byproduct. Lastly, desired final products **22-24** were obtained in 30-76% yields through demethylation of **57a-c** with pyridinium chloride at 190 °C.

Scheme 6^a. Synthesis of compounds **22-24**.



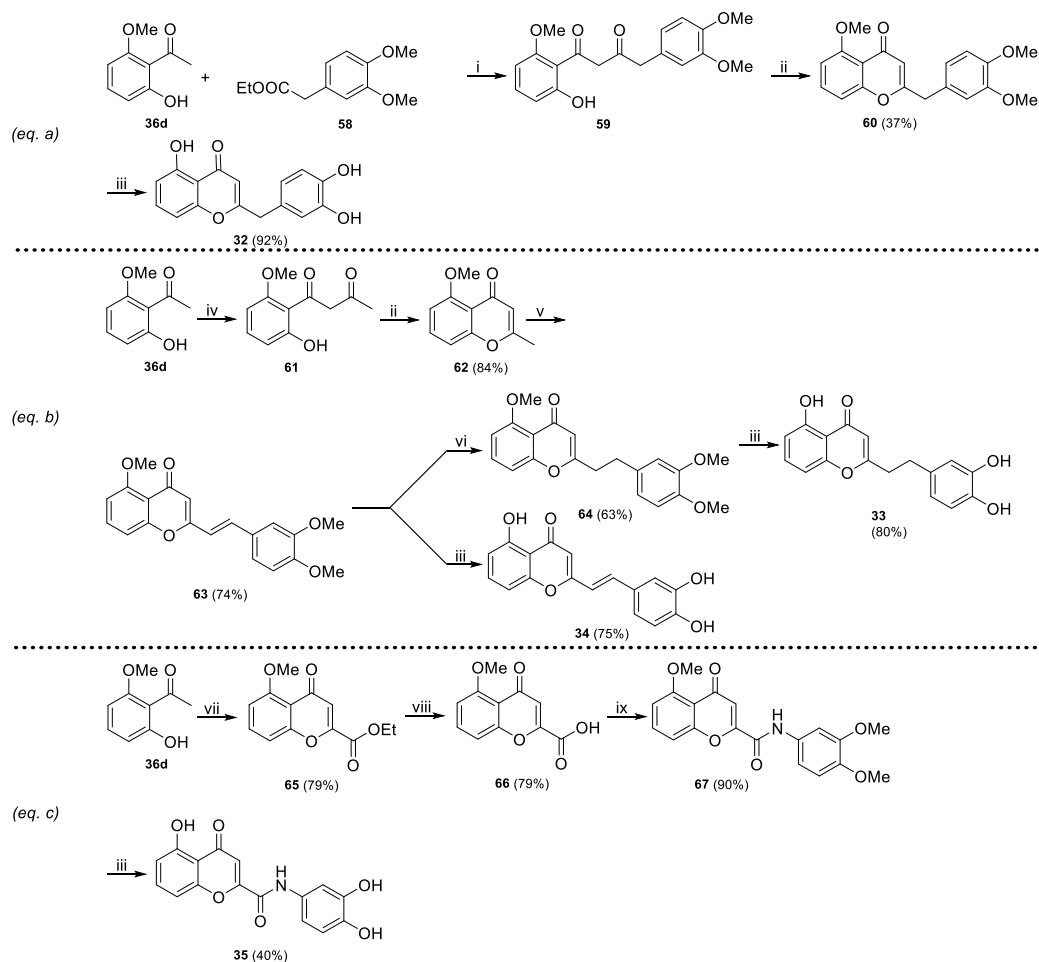
a) (i) AlCl_3 , DCM dry, r.t. (ii) vetraldehyde, KOH, MeOH, r.t., (iii) I_2 cat., DMSO, 135 °C (iv) pyridinium chloride, 190 °C.

Compounds **32-35** were synthesized as summarized in Scheme 7. As we previously mentioned, this set of compounds was prepared with different synthetic approaches depending on the nature of the spacer. The insertion of methylene spacer in compound **32** was allowed through a 3-step synthetic route (Scheme 7, eq a). The initial Claisen condensation between **36d** and **58** gave the 1,3-diketone **59**, that was subsequently subjected to dehydrative cyclization to afford **60** in a 37% yield over two steps. Finally, deprotection with boron tribromide gave the desired compound **32** in a 92% yield (eq. a). The elongation of the spacer with one more carbon atom in **33** and **34** was obtained according to the previous synthetic strategy, with minor modifications. The Claisen condensation between **36d** and ethyl acetate, in presence of sodium hydride as a base, gave the 1,3-diketone intermediate **61**, which was then cyclized to **62** with HCl 37% in methanol in a 84% yield over two steps.

The next vinylogous aldol condensation of **62** with vetraldehyde in basic condition (sodium ethoxide, ethanol) afforded divergent intermediate (*E*)-**63** in a stereoselective fashion (10:1 ratio *E/Z* isomers). The reduction of the ethylene of **63**, with ammonium acetate and palladium on carbon, gave **64** in 63% yield, which after final deprotection with boron tribromide led to the final compound **33**, with a completely saturated hydrocarbon spacer (eq. b). On the other hand, the direct deprotection of **62** with boron tribromide led mainly to the formation of the final compound **32**, with a vinyl spacer (eq. b).

Compound **35** was prepared by amide coupling between the proper chromone-2-carboxylic acid **66** and aniline derivatives as a key step. Chromone-2-carboxylic acid derivative **66** was obtained in 72% yield by reacting **36d** with diethylmalonate and sodium ethoxide followed by ester hydrolysis with potassium carbonate. Amide coupling between **66** and 3,4-dimethoxyaniline was achieved using HATU and DIPEA. Lastly, deprotection with boron tribromide led to the formation of final compound **35** with a 35% yield over two steps (eq. c).

Scheme 7^a. Synthesis of compounds **32-35**.



a) (i) NaH, THF dry, reflux, (ii) HCl 37%, MeOH dry, r.t., (iii) BBr₃, DCM dry, -78°C/40 °C, (iv) NaH, EtOAc, THF dry, 70°C (v) vetraldehyde, EtONa, EtOH, 50 °C, (vi) NH₄COOH, Pd/C, MeOH dry, reflux, (vi) NaH, THF dry, reflux, (vii) NaH, diethyl malonate, NaOEt, EtOH, reflux, (viii) K₂CO₃, EtOH, THF, 50 °C, (ix) 3,4-dimethoxyalanine, HATU, DIPEA, DCM, DMF, r.t.

Structure-activity relationship of our Pol η inhibitors.

Ring A. On ring A, deletion of hydroxyl group on position 5, as in **2**, resulted in a 2-fold loss of activity. In fact, **2** has an IC₅₀ of 27.0 ± 7.2 μM (as compared to 12.7 ± 0.2 μM of **1**). However, when a methoxy group was inserted in position 5, as in **3**, the IC₅₀ was 10.2 ± 1.8 μM. This may be due to the increase of lipophilicity, which could favor the establishment of Van der Waals interactions with the receptor. On the contrary, deletion of hydroxyl group on position 7, as in **4**, led to an IC₅₀ of 8.7 ± 1.4 μM. This data prompted us to further investigate this position. The

replacement of the hydrogen in position 7, substituted with a fluorine in compound **5**, returned an IC_{50} of $13.8 \pm 4.1 \mu\text{M}$, confirming that a hydroxyl group in position 7 was not crucial for the inhibitory activity. As expected, deletion of both hydroxyl groups, as in **6**, dropped the activity to an IC_{50} of $30.5 \pm 7.6 \mu\text{M}$. Lastly, an additional hydroxyl group on position 6, as in **25**, maintained the activity with an IC_{50} of $11.8 \pm 2.8 \mu\text{M}$, even in the absence of the catechol moiety on ring B.

This indicates that the presence of contiguous hydroxyl groups on the ring A favors the coordination of the catalytic metal ions, thus stabilizing the binding of the inhibitor, as also supported by docking calculation (Figure S1, supporting info).

These data suggest that a hydroxyl or a methoxy group in position 5 is needed to retain the activity, whereas a hydroxyl group in position 7 could be removed or substituted with a fluorine without affecting potency. Finally, an additional hydroxyl group in position 6 is tolerated.

Ring B. Shifting our attention to ring B, we investigated the role of each individual hydroxyl group, generating compounds **7**, with one hydroxyl group in position 4', and **8**, with one hydroxyl group in 3'. Compound **7** turned out 2-fold worse in activity compared to **1**, with an IC_{50} of $22.2 \pm 8.0 \mu\text{M}$, while **8** displayed a 2.5-fold increase in activity, with an IC_{50} of $5.0 \pm 0.8 \mu\text{M}$, suggesting a crucial role of the hydroxyl group in position 3', whereas the one on position 4' appears less relevant.

Removal of the catechol motif, as in **9**, retained the activity, with an IC_{50} of $14.7 \pm 7.0 \mu\text{M}$. We then investigated the replacement of the catechol group with different functionalities: first, we masked the two contiguous hydroxyl groups, as in the difluorobenzodioxole analogues **10** and **11**, and we then also introduced the salicylic motif in **12**, mimicking the characteristic motif in the aurintricarboxylic acid (ATCA), a well-known potent and non-selective inhibitor of DNA polymerases ($IC_{50} = 0.4 \mu\text{M}$, a result obtained with a different assay called Gel-based DNA primer elongation assay).²⁷ Such modifications in **10** and **11** annihilated the activity ($IC_{50} > 100 \mu\text{M}$). The modification in **12** maintained the activity, returning an IC_{50} of $7.7 \pm 5.5 \mu\text{M}$.

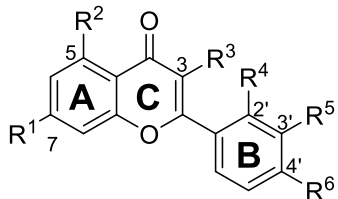
We then designed and synthesized four additional compounds bearing the methoxy group on position 5 on ring A, and different substitution patterns on ring B. Toward this end, we introduced lipophilic groups in position 2', in absence of the catechol moiety, generating ortho fluorine, methyl and trifluoromethyl analogues, i.e. compounds **13**, **14** and **15**, respectively. All compounds

were inactive, with an $IC_{50} > 100 \mu\text{M}$. Even when the hydroxyl group on position 4' was maintained as in **16**, the activity was not restored ($IC_{50} = 98.3 \pm 3.0 \mu\text{M}$).

Ring C. In this case, we started testing quercetin **26**, which has one additional hydroxyl group in position 3 of ring C, and which returned an $IC_{50} = 15.8 \pm 1.9 \mu\text{M}$, thus comparable to the selected hit **1** ($IC_{50} = 12.7 \pm 0.2 \mu\text{M}$). Thus, we further investigated position 3 of such ring and evaluated the effect of different alkoxy chains. Introduction of a methoxy group, as in **17**, retained the potency leading to an IC_{50} of $10.8 \pm 5.0 \mu\text{M}$. Elongation of the alkoxy chain with ethoxy, and propoxy groups, as in **18** and **19**, had a negative effect, with an almost 3-fold and 5-fold decrease in activity, respectively (IC_{50} of $33.5 \pm 1.8 \mu\text{M}$ for **18** and $62.4 \pm 13.6 \mu\text{M}$ for **19**).

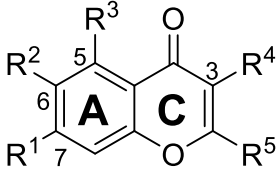
Since extending the length of the lipophilic chain resulted detrimental for the activity of these derivatives, we tried then to introduce a polar hydroxyl group at the terminal position of the alkoxy groups, thus generating the hydroxyalkoxy derivatives **20** and **21**. When the ethoxy chain was added with an additional hydroxyl group, as in **20**, activity dropped to an IC_{50} of $87.5 \pm 0.3 \mu\text{M}$. On the contrary, conversion of the propoxy chain into a 3-hydroxypropoxy chain, as in **21**, restored the activity, with an IC_{50} of $14.7 \pm 8.7 \mu\text{M}$. The role of the oxygen on position 3 was then investigated introducing simple alkyl chains. The resulting activity data suggest that such oxygen atom plays a crucial role for inhibition. In fact, compound **22**, with a methyl group, had an IC_{50} of $84.6 \pm 3.7 \mu\text{M}$, compound **23**, with a *n*-propyl chain, had an IC_{50} of $87.1 \pm 0.3 \mu\text{M}$, and compound **24**, bearing a *n*-butyl chain, had an $IC_{50} > 100 \mu\text{M}$. Finally, we found that isoflavones are less potent, as suggested by the comparison of flavone **7** (IC_{50} of $22.2 \pm 8.0 \mu\text{M}$), with its corresponding isoflavone **27**, which returned an IC_{50} of $92.4 \pm 10.7 \mu\text{M}$. Notably, MD simulations of compound **21** showed a stable binding pose (Figure S3), comparable to the initial hit (**Figure 14**).


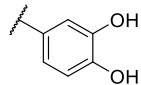
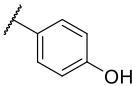
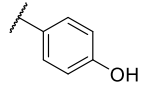
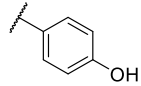
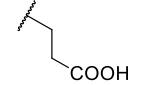
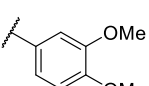
Table 1 IC₅₀ values of compounds **1-24**.



Compound	R ¹	R ²	R ³	R ⁴	R ⁵	R ⁶	IC ₅₀ (μM) ^a
1 ARN17212	OH	OH	H	H	OH	OH	12.7 ± 0.2
2	OH	H	H	H	OH	OH	27.0 ± 7.2
3	OH	OMe	H	H	OH	OH	10.2 ± 1.8
4	H	OH	H	H	OH	OH	8.7 ± 1.4
5	F	OH	H	H	OH	OH	13.8 ± 4.1
6	H	H	H	H	OH	OH	30.5 ± 7.6
7	OH	OH	H	H	H	OH	22.2 ± 8.0
8	OH	OH	H	H	OH	H	5.0 ± 0.8
9	OH	OH	H	H	H	H	14.8 ± 7.3
10	OMe	OH	H	H	OCF ₂ O		>100
11	OH	OH	H	H	OCF ₂ O		>100
12	OH	OH	H	H	COOH	OH	7.7 ± 5.5
13	H	OMe	H	F	H	H	> 100
14	H	OMe	H	Me	H	H	> 100
15	H	OMe	H	CF ₃	H	H	> 100
16	H	OMe	H	Me	H	OH	98.3 ± 3.0
17	OH	OH	OMe	H	OH	OH	10.8 ± 5.0
18	OH	OH	OEt	H	OH	OH	33.5 ± 1.8
19	OH	OH	OBu	H	OH	OH	62.4 ± 13.6
20	OH	OH	OCH ₂ CH ₂ OH	H	OH	OH	87.5 ± 0.3
21	OH	OH	OCH ₂ CH ₂ CH ₂ OH	H	OH	OH	14.7 ± 8.7
22	OH	OH	Me	H	OH	OH	84.6 ± 3.7
23	OH	OH	Pr	H	OH	OH	87.1 ± 0.3
24	OH	OH	Bu	H	OH	OH	> 100

a) Experiments have been done in duplicate

Table 2. IC₅₀ values of compounds **25-31**.


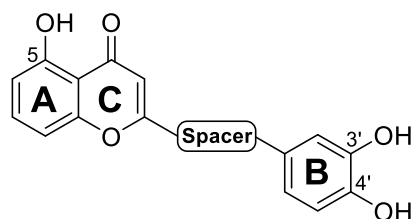
Compound	R ¹	R ²	R ³	R ⁴	R ⁵	IC ₅₀ (μM) ^a
25 Baicalein	OH	OH	OH	H		11.8 ± 2.8
26 Quercetin	OH	H	OH	OH		15.8 ± 1.9
27 Genestein	OH	H	OH		H	92.4 ± 10.7
28 Kaempferol	OH	H	OH	OH		12.7 ± 0.7
29 Isokaempferide	OH	H	OH	OMe		16.1*
30	OH	H	OH	H		> 100
31 Sinesetin	OMe	OMe	OMe	H		> 100

a) Experiments have been done in duplicate ^{b)} data was generated through one replicate (n = 1).

Exploring a spacer between ring B and C. We also investigated a small set of compounds bearing a spacer between ring B and C (**32-35**, Table 3). Such spacer spanned different lengths and degrees of structural flexibility. Since demethylation of the hydroxyl group on position 7 needed harsh conditions (typically pyridium chloride at 190 °C) and hydroxyl group on position 7 is not relevant for the activity (compound **4**, Table 1), we designed and synthesized analogs without the hydroxyl group in position 7. In this way, we were able to greatly streamline the synthesis using milder final deprotection procedures. Notably, our data (see above) also suggest that removal of the hydroxyl group on position 7 has no significant effect on potency, thus such analogs can be compared to **1**.

Overall, we found that the introduction of an aliphatic spacer drops the inhibitory activity, regardless of the length and degree of flexibility. Compound **32**, with a methylene spacer, was inactive ($IC_{50} > 100 \mu M$). Activity was not restored even if the spacer was elongated, as in **33**, or if the flexibility was reduced through a double carbon-carbon bond, as in **34**. On the contrary, compound **35**, characterized by a rigid amide spacer, had an IC_{50} of $26.5 \pm 6.9 \mu M$.

Table 3 IC_{50} values of compounds **32-35**.



Compound	spacer	IC_{50} (μM) ^a
32	-CH ₂ -	>100
33		>100
34		>100
35		26.5 ± 6.9

a) Experiments have been done in duplicate

In summary, the hydroxyl group of **1** on position 5 of ring A is crucial and could not be removed, but it could be methylated, as proved by **3**. On the other hand, the hydroxyl group on position 7 is less essential and it could be removed or substituted with a bioisoster fluorine without affecting the potency, as supported by **4** and **5**. Moreover, an additional hydroxyl group on position 6 is tolerated, as proved by **25**. On ring B, the catechol motif of **1** is not necessary for the activity, but a hydroxyl group on position 3' could not be removed, as supported by **7**. On ring C, the introduction of a hydroxylalkoxy group in position 3 with a proper balance between the lipophilic carbon chain and the polar hydroxyl group, as in **21**, maintained the activity and it opens the way to further investigations. Interestingly, the strategy of the insertion of an amide spacer between rings B and C led to active compound **35**.

DMPK proprieties of our lead compounds. After the design, synthesis, and evaluation of the new flavonoid derivatives for their inhibitory activity against Pol η *in vitro*, we selected the most

potent and promising inhibitors: the hit **1** and analogs **3-5**, **7**, **8**, **17**, **21**, **25**, and **26** for further evaluation in term of drug like properties (kinetic solubility, plasma and microsomal stability). The effect of different substituents on positions 5 and 7 of ring A was studied with **3-5**. Compound **25** was studied to evaluate if three contiguous hydroxyl groups on ring A could be detrimental for the stability. The effect of each hydroxyl group in specific position of ring B was evaluated with **7** and **8**. The effect of a hydroxyl group, an alkoxy chain, and a hydroxyalkoxy chain on position 3 of ring C was evaluated with **17**, **21**, and **26**.

We assessed the metabolic stability using mouse serum and mouse liver microsomes, and their kinetic solubility (S_k) in neutral buffer (pH 7.4). Natural flavonoids **1**, **4**, **7**, **8**, **25**, **26** functionalized with only hydroxyl groups had a poor kinetic solubility in the range 6-21 μM . Indeed, removal of one hydroxyl group slightly decreased solubility as for compound **4**, **7**, and **8** ($S_k = 13 \pm 3 \mu\text{M}$, $9 \pm 3 \mu\text{M}$ and $6 \pm 1 \mu\text{M}$ respectively, Table 4) compared to solubility of the hit **1** ($S_k = 21 \pm 4 \mu\text{M}$, Table 4). Unexpectedly, replacement of hydroxyl group on position 7 on ring A with fluorine, as in **5**, annihilated the solubility ($< 1 \mu\text{M}$). Notably, derivatives **17** and **21**, with methoxy and hydroxyethoxy groups in 3 position, displayed an excellent solubility in aqueous buffer (pH 7.4), greater than 200 μM (Table 4), while compound **3** with the methoxy group in 5 position showed an acceptable kinetic solubility of $87 \pm 1 \mu\text{M}$ (Table 4) suggesting that the introduction of a group with a balance of lipophilicity and polarity could help the solubility.

Most derivatives had excellent plasma and microsomal stability, with half-life values greater than 120 and 60 min, respectively (compounds **1**, **3-5**, **7**, **17**, Table 4). Notably, introduction of an additional hydroxyl group on position 6 of ring A, as in **25**, did not affect plasma stability and displayed an acceptable microsomal stability with a half-life value of 41 ± 2 min. Microsomal stability dropped for compounds **7** and **8** with half-life values of 8 ± 1 min and 11 ± 1 min, respectively, suggesting that compounds with a single hydroxyl group on ring B are rapidly degraded by liver metabolism. Notably, these modifications did not affect plasmatic solubility. On the other hand, introduction of a hydroxyl group on position 3 of ring C, as in **26**, did not affect the microsomal stability, but it displayed a very low plasma stability, with a half-life value of 7 ± 2 min. As previously mentioned, this was restored to $t_{1/2} > 120$ min when the hydroxyl group was alkylated, as in **17** and **21**. On the other hand, baicalein, **25**, with three contiguous hydroxyl groups on ring A, displayed an acceptable microsomal stability of 41 min (Table 4), and good plasmatic stability was maintained ($t_{1/2} = 120$ min, Table 4). In conclusion, our hit compound **1** had an

excellent metabolic stability profile, but a low kinetic solubility at physiological pH. Notably, among the **9** evaluated analogs, **17** and **21** have a much-improved kinetic solubility compared to **1** and retain excellent metabolic stability values. Considering the good IC₅₀ values and the overall drug-like profile, compounds **17** and **21** were selected for further evaluation in cell-based assays.

Table 4. Kinetic solubility, plasma stability, and microsomal stability of selected most potent inhibitors.

Compound	Kinetic solubility (μM)	t _{1/2} plasma (min)	t _{1/2} microsomes (min)
1	21 ± 4	>120	>60
ARN17212			
3	87 ± 1	>120	>60
4	9 ± 1	>120	>60
5	< 1	>120	>60
7	13 ± 3	>120	8 ± 1
8	6 ± 1	>120	11 ± 1
17	229 ± 10	>120	>60
21	231 ± 15	90 ± 21	>60
25	12 ± 3	>120	41 ± 2
26	16 ± 5	7 ± 2	>60

In addition, we underline that inhibitors **21** and **17** exhibited a potency in the low micro-molar range, comparable to PNR-02 (a derivative of ITBA).^{31, 32} Notably, this new class of Pol η inhibitors is also more potent than the 300 fragments identified through the high-throughput X-ray crystallography approaches.³³ Also, our efforts to modify a flavonoid scaffold resulted in an important improvement of the drug-like properties compared to natural flavonoid source. Compounds **21** and **17** have a good aqueous solubility major than 200 μM, much higher than the initial hit **1**. In particular, derivative **17** showed a remarkable plasma and microsomal stability, major than 120 and 60 minutes, respectively (Table 4).

Antiproliferative activity evaluation of the best compounds, 17 and 21. The antiproliferative activity of compounds **17** and **21** was assessed over 48 h in three cancer cell lines (i) A549, non small cell lung cancer; (ii) A375, malignant melanoma; and (iii) OVCAR3, ovarian

adenocarcinoma. The cytotoxic effect was also measured on a normal line (iv) HEK-293, kidney epithelial cells (Figure S3).

Compound **17** was cytotoxic with a LD_{50} of $60 \pm 5 \mu\text{M}$ for A549, $29 \pm 13 \mu\text{M}$ for A375 and $5.2 \pm 1.3 \mu\text{M}$ for OVCAR3. On the normal cell line, compound **17** was instead not cytotoxic at 48 h, with an effect only when the treatment lasted 72 h (LD_{50} of $30.5 \pm 7.8 \mu\text{M}$). Compound **21** showed no cytotoxicity in all cell lines, with a $LD_{50} > 100 \mu\text{M}$. After the assessment of the compound's antiproliferative activity, we proceeded to check cell viability in co-treatments studies of **17** and **21** with cisplatin. The LD_{50} for a treatment with cisplatin alone (0-50 μM) was: $29.8 \pm 5.2 \mu\text{M}$ for the A549 cell lines; $6.9 \pm 2.9 \mu\text{M}$ for the A375; $3.3 \pm 0.8 \mu\text{M}$ for OVCAR3; $4.3 \pm 0.9 \mu\text{M}$ for the HEK-293. Co-treating the three cancer cell lines with either one of the two inhibitors, the LD_{50} value decreased from 3- to 17-fold depending on the inhibitor and the concentration used, as shown in Table 5. To further investigate the mechanism of action, we calculated the combination index (CI) value for all the treatments using the Chou-Talalay method.³⁹ The CI allows to quantitatively classify drug combination additive effect ($0.8 < CI < 1.2$), synergism ($CI < 0.8$), and antagonism ($CI > 1.2$). The inferred CI (Table 6) was 0.66 and 0.27 for co-treatment with 35 μM and 50 μM , respectively, of compound **21** in the A375 cancer cell lines. This indicates a synergistic effect between cisplatin and compound **21**. Notably, such a synergistic effect was not limited to the A375 cell line. The same effect was obtained in A549 (CI= 0.78 and 0.6) and OVCAR3 (CI= 0.52 and 0.50) cell lines. Interestingly, compound **17** turned out the most potent analog (Table 6), exhibiting an additive effect.

Subsequently, we tested the same co-treatment on normal cell line HEK-293. Specifically, we treated cells with cisplatin (0-50 μM) and 30 μM of compounds **17** or **21**. The LD_{50} after 48 h of treatment did not decrease suggesting that the compounds have no effect on normal cells viability. Finally, we monitored the phosphorylation of the histone protein $\gamma\text{-H2AX}$ to form $p\gamma\text{-H2AX}$ by western blot and immunofluorescence analysis. Indeed, $p\gamma\text{-H2AX}$ accumulation is directly proportional to the extent of DNA double strand breaks. This allowed us to follow the DNA damages caused by the co-treatment. Using western blot, we treated such cells with 1 μM or 5 μM of cisplatin (A375) and 0,5 μM and 1 μM of cisplatin (OVCAR3) and 30 μM of compound **17** or **21** for 48 h. Moreover, we performed also the co-treatment using one fixed concentration of cisplatin (5 μM) and different concentrations of compound **21** (20 and 50 μM). Results are reported in the S.I. **Figure S4**. For the immunofluorescence, we treated cells with 0.5 μM or 1 μM of

cisplatin and with (or without) 20 μM of compound **21** for 24 h. From Figure 16 and 17, cisplatin alone induces a high level of DNA damage, but when it was used together with compounds **21** or **17**, the accumulation of $\text{p}\gamma\text{-H2AX}$ increases > 2-fold in the A375 cell line (no statistical significance for the one-way ANOVA test for the A375 cell lines, while is significant for the OVCAR3 $**p<0.0015$).

More interestingly, compound **21** is considered the most promising inhibitor of this class due to its inability to induce, when used alone, DNA damage as proven by unchanged $\text{p}\gamma\text{-H2AX}$ extent. On the contrary, compound **17** increased $\text{p}\gamma\text{-H2AX}$ accumulation, suggesting that it may be also involved in other mechanisms related to DNA damage.

In summary, two compounds were characterized for their cytotoxicity in cancer and normal cell lines. They showed to be moderately or not cytotoxic in normal and cancer cell line. Moreover, compound **21** and **17** proved to be able to influence the cisplatin effect synergistically or additively in cancer cell line, respectively. Interestingly, compound **21** showed to lead to a 2-fold increase in $\gamma\text{-H2AX}$ accumulation only when administered in combination with cisplatin. Notably, these results also confirmed a role of Pol η in the repair of DNA damage caused by cisplatin. The absence of the same effect in the normal cell line is a further confirmation that in a non-cancer system the expression levels of Pol η are not significant, probably because of the very low levels of DNA damage.

Table 5. LD₅₀ values of the cisplatin co-treatment with compounds 21 and 17 in cancer cell lines.

LD ₅₀ (μM)	[Cisplatin]	[Compound 21]					[Compound 17]			
	0-50 μM ^a	10 μM	35 μM	50 μM	75 μM	100 μM	5 μM	10 μM	20 μM	30 μM
A549	29.8 \pm 5.2	/	/	19.7 \pm 0.5	/	9.4 \pm 4.8	/	/	11.9 \pm 4.3	7.4 \pm 4.2
A375	6.9 \pm 2.6	/	2.6 \pm 0.03	1 \pm 0.6	/	/	/	2 \pm 0.6	0.5 \pm 0.3	0.4 \pm 0.2
OVCAR3	3.3 \pm 0.8	4.8 \pm 1.3	/	2.8 \pm 0.01	0.27 \pm 0.24	/	4.8 \pm 1.3	5.8 \pm 0.9	/	/
HEK293	17.8 \pm 9.8	/	14.8 \pm 7.1	/	/	/	/	/	/	13.9 \pm 4.7

^a Serial dilution of Cisplatin from 50 μM to 0 μM .

Table 6. Combination Index (CI) values of the cisplatin co-treatment with compounds 21 and 17 in cancer cell lines.

CI		[21] μM				[17] μM		
		35	50	75	100	10	20	30
A375	Cisplatin 12.5 μM	0.78	0.23			1.48		
	Cisplatin 6 μM	0.66	0.27			0.96		
A549	Cisplatin 12.5 μM		0.8		0.58		1.24	0.94
	Cisplatin 6 μM		0.78		0.60		1.03	0.96
OVCAR3	Cisplatin 12.5 μM		0.78	0.72		2.33		
	Cisplatin 6 μM		0.52	0.5		1.15		

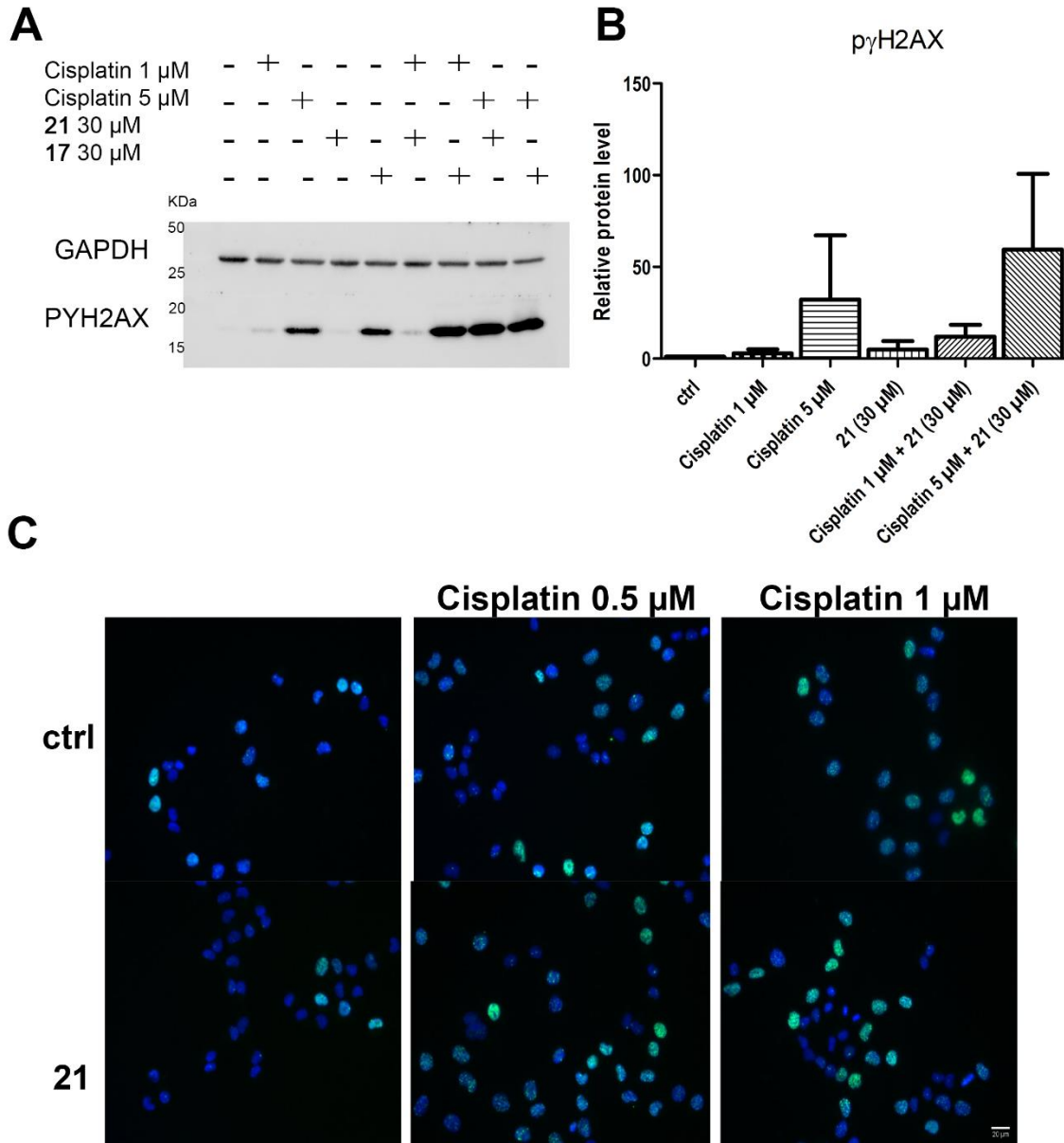


Figure 18 A375. (A) Protein level of PYH2AX after 48h of treatment with cisplatin 1 μ M or 5 μ M, with compounds 21 and 17 (30 μ M), and the co-treatment, was detected by immunoblotting. (B) Relative protein level was quantified with ImageJ (from duplicates mean \pm SD). Results shown here are from one of two experiments with identical results. (C) A375 cells were treated for 24h with Cisplatin 0.5 μ M or 1 μ M, with 21 (20 μ M) and with both cisplatin and compound 21; and localization of PYH2AX was detected by microscopy. (D) splitted channel. Nuclei were stained with Hoechst 33342 (blue).

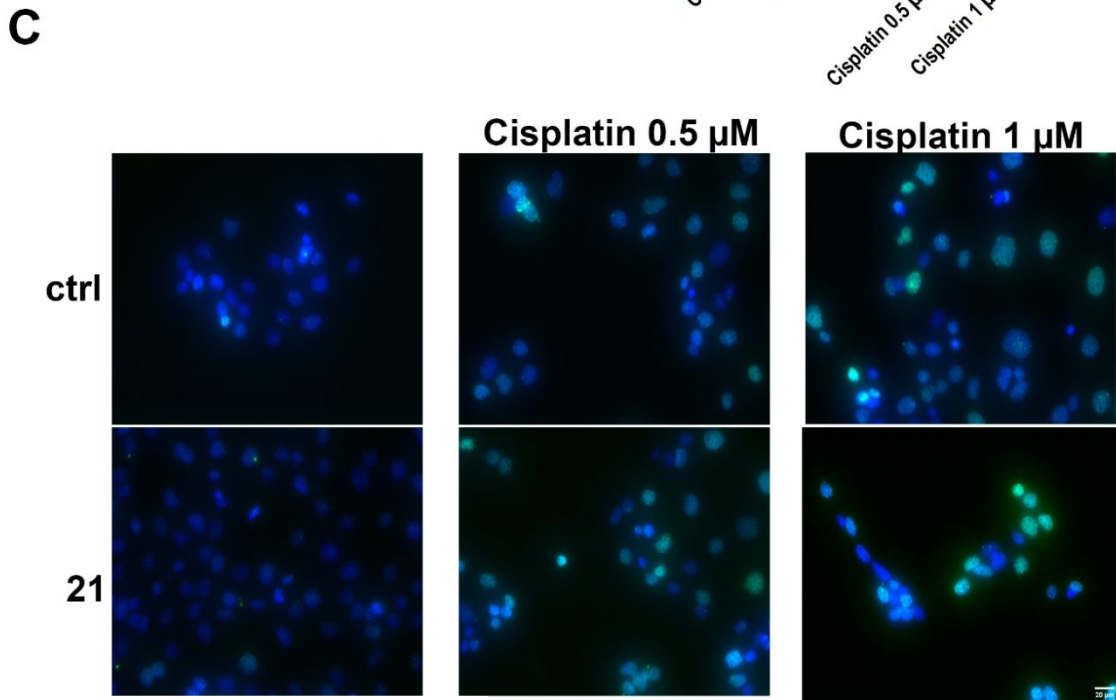
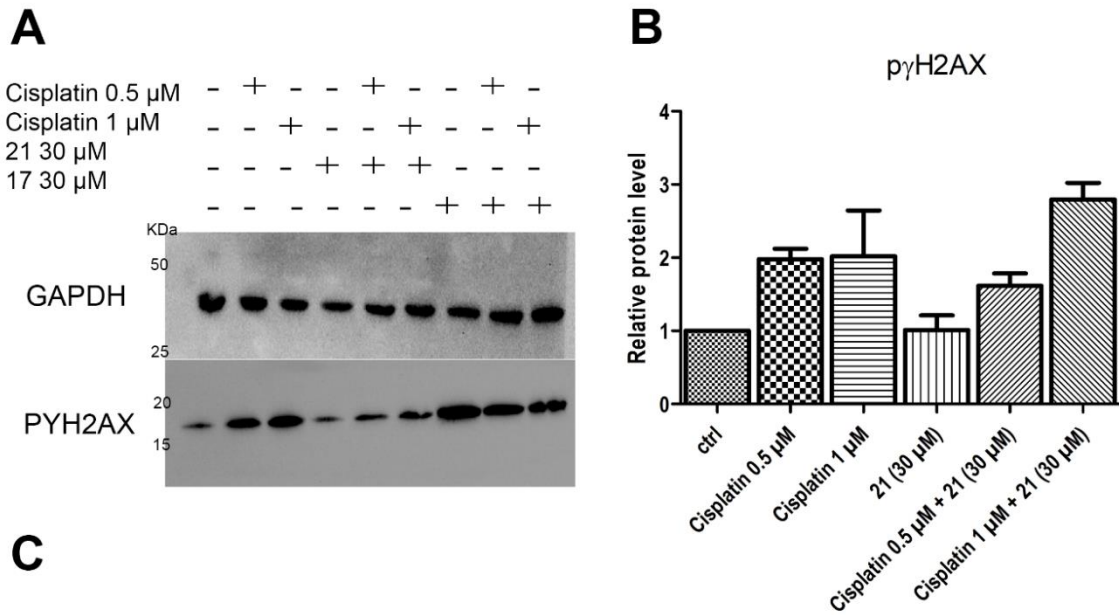


Figure 19 OVCAR3. (A) Protein level of PYH2AX after 48h of treatment with cisplatin 0.5 μ M or 1 μ M, with compounds 21 and 17 (30 μ M), and the co-treatment, was detected by immunoblotting. (B) Relative protein level was quantified (from duplicates mean \pm SD) with ImageJ. Results shown here are from one of two experiments with identical results. (C) OVCAR3 cells were treated for 24h with Cisplatin 0.5 μ M or 1 μ M, with 21 (20 μ M) and with both cisplatin and compound 21; and localization of PYH2AX was detected by microscopy. Below, splitted channels. Nuclei were stained with Hoechst 33342 (blue).

2.3 Conclusions

Elevated expression of the translesion DNA synthesis enzymes like Pol η is an important factor in the drug resistance developed by patients to DNA-damaging chemotherapeutics. Therefore, inhibition of Pol η is an attractive strategy for the development of drugs to overcome chemoresistance. Pol η is overexpressed in several cancer types, decreasing the effectiveness of platinum-based drugs. The development of Pol η inhibitors could slow the development of drug resistance and increase the efficacy of the treatment.

Toward this end, starting with a virtual screening campaign followed by experimental testing, we identified ARN18347 as our initial hit compound. Then, a subsequent similarity search led to the identification of ARN17212, which had a comparable potency, but a more easily tractable chemical scaffold of the initial hit ARN18347. Several ARN17212 analogs have been synthesized and evaluated *in vitro* in terms of inhibitory potency of Pol η and DMPK properties (aqueous kinetic solubility, microsomal stability, and plasma stability). Specifically, 35 newly synthesized analogues have been generated, and compound **21** (ARN24964) stands out as the most potent and drug-like inhibitor of Pol η (IC_{50} of $14.7 \pm 8.7 \mu\text{M}$), with a good DMPK overall profile. In addition, ARN24964 was able to enhance the effectiveness of the drug cisplatin, with a synergistic effect. This effect was limited to Pol η -proficient cancer cells (A375, A549, and OVCAR3), as it did not sensitize healthy cells (HEK-293). Analysis of YH2AX expression further validated the ability of ARN24964 to increase the genotoxic effect of cisplatin.

In conclusion, we have discovered a novel small-molecule inhibitor of Pol η , ARN24964, which is capable to potentiate the cytotoxic effect of cisplatin on cultured cancer cell lines. This compound represents an advanced scaffold featuring good potency, aqueous solubility, microsomal stability and plasma stability, thus representing a solid starting point for further developments in a scenario where there are only a few compounds designed to inhibit such a key enzyme and overcome resistance to platinum-based chemotherapeutics. Thus, taken together, these results endorse a further exploration of this new set of Pol η inhibitors. We foresee a further expansion of this chemical class to tune its pharmacokinetics and biodistribution, considering also specific drug delivery procedures, as previously performed for the well-known flavonoid quercetin.⁴⁰

2.4 Experimental section

Chemistry. General considerations. All the commercially available reagents and solvents were used as purchased from vendors without further purification. Dry solvents were purchased from Sigma-Aldrich. Automated column chromatography purifications were done using a Teledyne ISCO apparatus (CombiFlash[®] Rf) with pre-packed silica gel columns of different sizes (from 4 g up to 80 g) and mixtures of increasing polarity of cyclohexane and ethyl acetate (EtOAc) or dichloromethane (DCM) and methanol (MeOH). NMR experiments were run on a Bruker Avance III 400 system (400.13 MHz for ¹H, and 100.62 MHz for ¹³C), equipped with a BBI probe and Z-gradients. Spectra were acquired at 300 K, using deuterated dimethylsulfoxide (DMSO-d₆) or deuterated chloroform (CDCl₃) as solvents. For ¹H-NMR, data are reported as follows: chemical shift, multiplicity (s= singlet, d= doublet, dd= double of doublets, ddd= doublet of doublet of doublets, t= triplet, td= triplet of doublets, q= quartet, p= quintet, m= multiplet), coupling constants (Hz) and integration. UPLC/MS analyses were run on a Waters ACQUITY UPLC/MS system consisting of a SQD (single quadrupole detector) mass spectrometer equipped with an electrospray ionization interface and a photodiode array detector. The PDA range was 210–400 nm. Analyses were performed on an ACQUITY UPLC BEH C18 column (100x2.1mmID, particle size 1.7 μm) with a VanGuard BEH C18 pre-column (5x2.1 mmID, particle size 1.7 μm). Mobile phase was 10 mM NH₄OAc in H₂O at pH 5 adjusted with CH₃COOH (A) and 10 mM NH₄OAc in CH₃CN–H₂O (95:5) at pH 5.0. Two types of gradients were applied depending on the analysis, gradient 1 (5 % to 100 % mobile phase B in 3 min) or gradient 2 (50 % to 100 % mobile phase B in 3 min). Electrospray ionization in positive and negative mode was applied. ESI was applied in positive and negative mode. All tested compounds showed ≥ 95% purity by UPLC/MS analysis. Compound **1** was synthesized following procedure described by Ji Zhang and co-authors.⁴¹

General procedure A. Aldol condensation. The appropriate ketone of type **36** or **55** (1.0 eq.) and benzaldehyde of type **37** or **40** (1.1 eq) were added to a solution of KOH (20.0 eq.) in MeOH (0.12 M). The reaction mixture was stirred at room temperature for 72 hours. After complete conversion of starting materials, the reaction mixture was acidified to pH=5 with HCl 1M and extracted with EtOAc (3x5 mL). Combined organic layers were dried over MgSO₄ and concentrated under vacuum. The product was purified by silica or trituration with EtOH yielding the pure intermediate.

General procedure B. Oxidative cyclization. The appropriate chalcone of type **38**, **41** or **56** (1 eq.) was dissolved in DMSO (0.3 M) and heated at 135 °C under argon atmosphere. I₂ (0.05 eq.) was added and the reaction mixture was stirred until full conversion of starting material. After reaction completion, the reaction mixture was cooled to room temperature and Na₂S₂O₃ 1N was added to quench the iodine. The crude product was filtered, washed with water and purified by flash chromatography yielding the pure compound.

General procedure C. Method A for aryl methyl ether cleavage. The appropriate compound of type **39** or **57** (1 eq.) was treated with pyridinium chloride (10 eq.) and the reaction mixture was heated at 190 °C under argon atmosphere until total conversion of starting material. After reaction completion, reaction mixture was cooled down to room temperature and added with water. The crude product was filtered, washed with water and purified by silica.

General procedure D. Method B for aryl methyl ether cleavage. The appropriate compound of type **39**, **59**, **62**, **63** or **66** (1 eq.) was dissolved in DCM dry (0.05 M) and cooled to 0 °C. A 1M solution of BBr₃ in DCM (1.5 eq for each methoxy group) was added and the reaction mixture was allowed to warm to room temperature and stirred until complete conversion of starting material under argon atmosphere. After reaction completion, the reaction was quenched with MeOH (10 mL) and concentrated under reduced pressure. The crude product was rised in MeOH (15 mL), concentrated under vacuum and purified by flash chromatography or crystallization.

General procedure E. Benzyl deprotection. Appropriate benzylated compound of type **42**, **49** or **53** was dissolved in a 1:1 mixture MeOH/DCM (0.04 M) under an argon atmosphere. Pd/C (20% w/w) and triethylsilane (6 eq. for each benzyl group) were added to the solution. The reaction mixture was stirred at 40°C until complete conversion of starting material. Then, the reaction mixture was filtered over a bed of celite and concentrated under vacuum. The crude product was rised in EtOAc and the organic phase was washed with water, dried over MgSO₄ and concentrated under vacuum. The crude product was purified by silica.

General procedure F. Claisen condensation. A solution of ketone **36d** (1 eq.) in THF dry (0.3 M) was added to a suspension of NaH 60% dispersion in mineral oil (4 eq.) in THF dry (1.2 M) under argon, followed by the addition of appropriate ester (2 eq.) at room temperature. Then

reaction mixture stirred at reflux under argon until complete consumption of starting material. The reaction mixture was quenched by pouring into ice and further acidified until pH 6 with HCl 2M aq and then extracted with EtOAc. Collected organic layers were dried over Na₂SO₄ and concentrated under vacuum. The crude was used as such without further purification.

General procedure G. Dehydrative Cyclization. The appropriate diketone of type **58** or **60** (1 eq.) was treated with a 9:1 mixture MeOH/HCl 37% (0.15 M). Reaction mixture stirred at room temperature until complete consumption of starting material. Then the solvent was removed under vacuum, the residue was rised with EtOAc and washed with NaHCO₃ sat. sol. Organic layer was divided, dried over Na₂SO₄ and concentrated under vacuum. Purification by silica afforded the pure desired product.

2-(3,4-dihydroxyphenyl)-7-hydroxy-4H-chromen-4-one (2). Title compound was synthesized following the general procedure **C** previously described using intermediate **39b** (139 mg, 0.44 mmol) and pyridinium chloride (509 mg, 4.4 mmol). Purification by silica (elution by gradient from 100:0 to 90:10 DCM/MeOH) afforded of pure compound **2** (37 mg, 31% yield). UPLC/MS Rt = 1.29 min (gradient 1), MS (ESI) m/z 270.9 [M+H]⁺. [M+H]⁺ calculated for C₁₅H₁₁O₅: 271.0. HRMS (AP-ESI) m/z calculated for C₁₅H₁₁O₅ [M + H]⁺ 271.0606, found 271.0594. ¹H NMR (400 MHz, DMSO-d₆) δ 7.84 (d, *J* = 8.5 Hz, 1H), 7.48 – 7.30 (m, 2H), 6.91 – 6.82 (m, 3H), 6.57 (s, 1H).

2-(3,4-dihydroxyphenyl)-7-hydroxy-5-methoxy-4H-chromen-4-one (3). Title compound was synthesized following the general procedure **E** previously described using intermediate **49** (67 mg, 0.12 mmol), Pd/C 20% w/w (30 mg), Et₃SiH (0.46 mL, 2.2 mmol) in a 1:1 mixture MeOH/DCM (3 mL). Trituration with cyclohexane (2 mL) and EtOAc (2 mL) afforded the pure compound **3** (13 mg, 36%). UPLC/MS Rt = 1.20 min (gradient 1), MS (ESI) m/z: 301.1 [M+H]⁺. [M+H]⁺ calculated for C₁₆H₁₃O₆: 301.1. HRMS (AP-ESI) m/z calculated for C₁₆H₁₃O₆ [M + H]⁺ 301.0712, found 301.0703. ¹H NMR (400 MHz, DMSO-d₆) δ 7.34 – 7.27 (m, 2H), 6.89 – 6.82 (m, 1H), 6.48 (d, *J* = 2.1 Hz, 1H), 6.39 (s, 1H), 6.36 (d, *J* = 2.1 Hz, 1H), 3.78 (s, 3H). ¹³C NMR (151 MHz, DMSO-d₆) δ 175.6 (CO), 162.4 (Cq), 160.6 (Cq), 160.0 (Cq), 159.0 (Cq), 148.8 (Cq), 145.7 (Cq), 121.8 (Cq), 118.1 (CH), 116.0 (CH), 112.9 (CH), 107.2 (Cq), 106.0 (CH), 96.4 (CH), 95.1 (CH), 55.9 (CH₃).

2-(3,4-dihydroxyphenyl)-5-hydroxy-4H-chromen-4-one (4). Title compound was synthesized following the general procedure **C** previously described using intermediate **39c** (60 mg, 0.19 mmol) and pyridinium chloride (219 mg, 1.9 mmol). Purification by silica (elution by gradient from 98:2 to 80:20 DCM/MeOH) afforded of pure compound **4** (46 mg, 90% yield). UPLC/MS Rt = 1.79 min (gradient 1), MS (ESI) m/z 271.0 [M+H]⁺. [M+H]⁺ calculated for C₁₅H₁₁O₅: 271.0. HRMS (AP-ESI) m/z calculated for C₁₅H₁₁O₅ [M + H]⁺ 271.0606, found 271.0595. ¹H NMR (400 MHz, DMSO-d₆) δ 7.56 (t, *J* = 8.3 Hz, 1H), 7.40 (dd, *J* = 8.5, 2.5 Hz, 1H), 7.27 (d, *J* = 2.5 Hz, 1H), 7.06 (d, *J* = 8.3 Hz, 1H), 6.71 (d, *J* = 8.2 Hz, 1H), 6.60 (s, 1H), 6.56 (d, *J* = 8.5 Hz, 1H). ¹³C NMR (101 MHz, DMSO-d₆) δ 182.4 (CO), 165.5 (Cq), 160.1 (Cq), 155.8 (Cq), 154.6 (Cq), 147.2 (Cq), 135.3 (CH), 119.8 (CH), 118.1 (Cq), 116.1 (CH), 112.2 (CH), 110.8 (CH), 110.0 (Cq), 107.0 (CH), 102.1 (CH).

2-(3,4-dihydroxyphenyl)-7-fluoro-5-hydroxy-4H-chromen-4-one (5). Title compound was synthesized following the general procedure **D** previously described using intermediate **39d** (100 mg, 0.3 mmol), BBr₃ 1M in DCM (1.4 mL, 1.4 mmol) in anhydrous DCM (6 mL). Crystallization with EtOH (1.5 mL) afforded pure compound **5** (56 mg, 66% yield). UPLC/MS Rt = 1.99 min (gradient 1), MS (ESI) m/z 287.1 [M-H]⁻. [M-H]⁻ calculated for C₁₅H₇FO₅: 287.2. HRMS (AP-ESI) m/z calculated for C₁₅H₉FO₅ [M + H]⁺ 289.0512, found 289.0512. ¹H NMR (400 MHz, DMSO-d₆) δ 13.24 (s, 1H), 7.47 (dd, *J* = 8.4, 2.3 Hz, 1H), 7.44 (d, *J* = 2.3 Hz, 1H), 7.08 (dd, *J* = 10.0, 2.4 Hz, 1H), 6.91 (d, *J* = 8.3 Hz, 1H), 6.84 (s, 1H), 6.72 (dd, *J* = 10.9, 2.3 Hz, 1H). ¹³C-NMR (101 MHz, DMSO- d₆) δ 182.6 (CO), 166.4 (d, ¹*J*_{C,F} = 248.9 Hz, CF), 165.4 (Cq), 162.4 (d, ⁴*J*_{C,F} = 16.5 Hz, Cq), 157.3 (d, ⁴*J*_{C,F} = 17.6 Hz, Cq), 150.7 (Cq), 146.3 (Cq), 121.5 (Cq), 119.9 (CH), 116.5 (CH), 114.1 (CH), 107.8 (CH), 103.8 (Cq), 99.7 (d, ³*J*_{C,F} = 25.3 Hz, CH), 95.7 (d, ³*J*_{C,F} = 27.0 Hz, CH). (565 MHz): δ – 98.9 MHz.

2-(3,4-dihydroxyphenyl)-4H-chromen-4-one (6). Title compound was synthesized following the general procedure **C** previously described using intermediate **39e** (157 mg, 0.56 mmol) and pyridinium chloride (644 mg, 5.6 mmol). Purification by silica (elution by gradient from 100:0 to 96:4 DCM/MeOH) afforded of pure compound **6** (50 mg, 35% yield). UPLC/MS Rt = 1.56 min (gradient 1), MS (ESI) m/z 255.0 [M+H]⁺. [M+H]⁺ calculated for C₁₅H₁₁O₄: 255.1. HRMS (AP-ESI) m/z calculated for C₁₅H₁₁O₄ [M + H]⁺ 255.0657, found 255.0647. ¹H NMR (400 MHz, DMSO-d₆) δ 8.02 (dd, *J* = 7.9, 1.6 Hz, 1H), 7.79 (ddd, *J* = 8.7, 7.1, 1.7 Hz, 1H), 7.70 (dd, *J* = 8.4,

1.0 Hz, 1H), 7.49 – 7.42 (m, 3H), 6.91 (dd, $J = 8.9$ Hz, 1H), 6.75 (s, 1H). ^{13}C NMR (101 MHz, DMSO- d_6) δ 176.8 (CO), 163.3 (Cq), 155.6 (Cq), 149.7 (Cq), 145.8 (Cq), 134.0 (CH), 125.3 (CH), 124.8 (CH), 123.3 (CH), 121.8 (CH), 118.8 (CH), 118.3 (CH), 116.0 (CH), 113.3 (CH), 104.8 (CH).

5,7-dihydroxy-2-(4-hydroxyphenyl)-4H-chromen-4-one (7). Title compound was synthesized following the general procedure **C** previously described using intermediate **39f** (136 mg, 0.43 mmol) and pyridinium chloride (497 mg, 4.3 mmol). Purification by silica (elution by gradient from 100:0 to 90:10 DCM/MeOH) afforded of pure compound **7** (88 mg, 81% yield). UPLC/MS $R_t = 1.61$ min (gradient 1), MS (ESI) m/z 271.2, $[\text{M}+\text{H}]^+$. $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{15}\text{H}_{11}\text{O}_5$: 271.0. HRMS (AP-ESI) m/z calculated for $\text{C}_{15}\text{H}_{11}\text{O}_5$ $[\text{M} + \text{H}]^+$ 271.0606, found 271.0599. ^1H NMR (400 MHz, DMSO- d_6) δ 12.95 (s, 1H), 7.95 – 7.87 (m, 2H), 6.96 – 6.88 (m, 2H), 6.75 (s, 1H), 6.44 (d, $J = 2.1$ Hz, 1H), 6.15 (d, $J = 2.1$ Hz, 1H). ^{13}C NMR (101 MHz, DMSO- d_6) δ 182.1 (CO), 164.9 (Cq), 164.1 (Cq), 161.9 (Cq), 161.7 (Cq), 128.9 (2C, CH), 121.6 (Cq), 116.4 (2C, CH), 104.0 (CH), 103.3 (CH), 100.0 (CH), 99.4 (CH), 94.5 (CH).

5,7-dihydroxy-2-(3-hydroxyphenyl)-4H-chromen-4-one (8). Title compound was synthesized following the general procedure **C** previously described using intermediate **39g** (180 mg, 0.58 mmol) and pyridinium chloride (671 mg, 5.8 mmol). Purification by silica (elution by gradient from 100 to 90:10 DCM/MeOH) afforded of pure compound **8** (57 mg, 42% yield). UPLC/MS $R_t = 1.70$ min (gradient 1), MS (ESI) m/z 271.1, $[\text{M}+\text{H}]^+$. $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{15}\text{H}_{11}\text{O}_5$: 271.0. HRMS (AP-ESI) m/z calculated for $\text{C}_{15}\text{H}_{11}\text{O}_5$ $[\text{M} + \text{H}]^+$ 271.0606, found 271.0594. ^1H NMR (400 MHz, DMSO- d_6) δ 12.82 (s, 1H), 7.48 (dd, $J = 7.6, 1.7$ Hz, 1H), 7.43 – 7.32 (m, 2H), 7.00 (dd, $J = 8.0, 2.5$ Hz, 1H), 6.83 (d, $J = 2.1$ Hz, 1H), 6.48 (d, $J = 2.1$ Hz, 1H), 6.21 (d, $J = 2.1$ Hz, 1H). ^{13}C NMR (101 MHz, DMSO- d_6) δ 181.8 (CO), 164.6 (Cq), 163.3 (Cq), 161.5 (Cq), 157.9 (Cq), 157.5 (Cq), 132.0 (Cq), 130.3 (CH), 119.1 (CH), 117.3 (CH), 112.8 (CH), 105.1 (CH), 103.9 (Cq), 99.1 (CH), 94.1 (CH).

5,7-dihydroxy-2-phenyl-4H-chromen-4-one (9). Title compound was synthesized following the general procedure **C** previously described using intermediate **39h** (94 mg, 0.33 mmol) and pyridinium chloride (381 mg, 3.3 mmol). Purification by silica (elution by gradient from 100:0 to 90:10 DCM/MeOH) afforded of pure compound **9** (70 mg, 87% yield). UPLC/MS $R_t = 1.99$ min

(gradient 1), MS (ESI) m/z 255.2, $[M+H]^+$. $[M+H]^+$ calculated for $C_{15}H_{11}O_4$: 255.1. HRMS (AP-ESI) m/z calculated for $C_{15}H_{11}O_4$ $[M + H]^+$ 255.0657, found 255.0645. 1H NMR (400 MHz, DMSO- d_6) δ 12.81 (s, 1H), 8.18 – 7.94 (m, 2H), 7.69 – 7.50 (m, 3H), 6.96 (s, 1H), 6.51 (d, $J = 2.0$ Hz, 1H), 6.20 (d, $J = 2.0$ Hz, 1H). ^{13}C NMR (101 MHz, DMSO- d_6) δ 182.3 (CO), 165.1 (Cq), 163.6 (Cq), 161.9 (Cq), 157.9 (Cq), 132.4 (CH), 131.2 (Cq), 129.6 (2C, CH), 126.8 (2C, CH), 105.6 (CH), 104.4 (Cq), 99.5 (CH), 94.6 (CH).

2-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)-5-hydroxy-7-methoxy-4H-chromen-4-one (10). Title compound was synthesized according to general procedure **D** previously described using intermediate **39i** (100 mg, 0.27 mmol), BBr_3 1M in DCM (0.42 mL, 0.42 mmol) in anhydrous DCM (10 mL). Crystallization with EtOH (1.5 mL) afforded pure compound **10** (56 mg, 57% yield). UPLC/MS $R_t = 2.67$ min (gradient 1), MS (ESI) m/z 347.0, $[M-H]^-$. $[M-H]^-$ calculated for $C_{17}H_9F_2O_6$: 347.3. HRMS (AP-ESI) m/z calculated for $C_{15}H_{11}O_4$ $[M + H]^+$ 349.0523, found 349.0513. 1H NMR (400 MHz, $CDCl_3$) δ 7.64 (dd, $J = 8.4, 1.8$ Hz, 1H), 7.57 (d, $J = 1.8$ Hz, 1H), 7.17 (d, $J = 8.5$ Hz, 1H), 6.61 (d, $J = 0.9$ Hz, 1H), 6.55 (d, $J = 2.2$ Hz, 1H), 6.38 (d, $J = 2.3$ Hz, 1H), 3.95 (s, 3H), 3.91 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 182.2 (CO), 165.9 (Cq), 162.4 (Cq), 165.4 (Cq), 157.7 (Cq), 146.2 (Cq), 144.6 (Cq), 127.8 (Cq), 122.9 (CH), 110.2 (CH), 107.7 (CH), 106.1 (CH), 105.7 (CH), 98.5 (CH), 92.9 (CH), 56.0 (CH_3). ^{19}F (565 MHz): $\delta - 47.9$ MHz.

2-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)-5,7-dihydroxy-4H-chromen-4-one (11). Title compound was synthesized following the general procedure **C** previously described using intermediate **39i** (100 mg, 0.27 mmol) and pyridinium chloride (324, 2.8 mmol). Crystallization with EtOH (1.5 mL) afforded pure compound **11** (28 mg, 30% yield). UPLC/MS $R_t = 2.28$ min (gradient 1), MS (ESI) m/z 333.1, $[M-H]^-$. $[M-H]^-$ calculated for $C_{16}H_6F_2O_6$: 333.2. HRMS (AP-ESI) m/z calculated for $C_{16}H_6F_2O_6$ $[M + H]^+$ 335.0367, found 335.0356. 1H NMR (400 MHz, DMSO- d_6) δ 12.71 (s, 1H), 8.12 (d, $J = 1.8$ Hz, 1H), 7.94 (dd, $J = 8.6, 1.8$ Hz, 1H), 7.57 (d, $J = 8.6$ Hz, 1H), 6.96 (s, 1H), 6.51 (d, $J = 2.1$ Hz, 1H), 6.19 (d, $J = 2.1$ Hz, 1H). ^{13}C NMR (101 MHz, DMSO- d_6) δ 182.2 (CO), 165.1 (Cq), 162.0 (Cq), 161.8 (Cq), 157.8 (Cq), 145.5 (Cq), 143.9 (Cq), 131.7 (t, $J = 253.9$ Hz, CF_2), 128.0 (CH), 124.2 (CH), 111.1 (CH), 108.8 (CH), 105.9 (CH), 104.3 (Cq), 99.6 (CH), 94.7 (CH). ^{19}F (565 MHz): $\delta - 47.9$ MHz.

5-(5,7-dihydroxy-4-oxo-4H-chromen-2-yl)-2-hydroxybenzoic acid (12). Title compound was synthesized following the general procedure **C** previously described using intermediate **39k** (70

mg, 0.19 mmol) and pyridinium chloride (220 mg, 1.9 mmol). Crystallization with EtOH (1.5 mL) afforded pure compound **12** (28 mg, 30% yield). UPLC/MS Rt = 1.32 min (gradient 1), MS (ESI) m/z 313.0, $[M-H]^-$. $[M-H]^-$ calculated for $C_{16}H_9O_7$: 313.2. HRMS (AP-ESI) m/z calculated for $C_{16}H_{11}O_7$ $[M+H]^+$ 315.0505, found 315.0492. 1H NMR (400 MHz, DMSO- d_6) δ 12.88 (s, 1H), 10.86 (s, 1H), 8.39 (d, $J = 2.4$ Hz, 1H), 8.18 (dd, $J = 8.8, 2.5$ Hz, 1H), 7.11 (d, $J = 8.8$ Hz, 1H), 6.86 (s, 1H), 6.50 (d, $J = 2.1$ Hz, 1H), 6.21 (d, $J = 2.1$ Hz, 1H). ^{13}C NMR (101 MHz, DMSO- d_6) δ 181.7 (Cq), 171.1 (Cq), 164.3 (Cq), 163.7 (Cq), 162.4 (Cq), 161.5 (Cq), 157.3 (Cq), 133.25 (CH), 128.6 (CH), 121.6 (Cq), 118.2 (CH), 114.0 (Cq), 104.0 (Cq), 103.8 (CH), 98.9 (CH), 94.0 (CH).

2-(2-fluorophenyl)-5-methoxy-4H-chromen-4-one (13). Title compound was synthesized according to general procedure **B** previously described using **41a** (96 mg, 0.40 mmol) and iodine (5 mg, 0.02 mmol) in DMSO (1.3 mL). Purification by silica (elution by gradient from 100:0 to 50:50 cyclohexane/EtOAc) afforded pure compound **13** (79 mg, 63% yield). UPLC/MS Rt = 1.96 min (gradient 1), MS (ESI) m/z: 271.1 $[M+H]^+$. $[M+H]^+$ Calculated for $C_{16}H_{12}FO_3$: 271.1. HRMS (AP-ESI) m/z calculated for $C_{16}H_{12}FO_3$ $[M+H]^+$ 271.0770, found 271.0767. 1H NMR (400 MHz, DMSO- d_6) δ 8.01 (ddd like td, $J = 7.9, 1.8$ Hz, 1H), 7.72 (dd like t, $J = 8.4$ Hz, 1H), 7.68 – 7.62 (m, 1H), 7.49 – 7.39 (m, 2H), 7.24 (d, $J = 8.4$ Hz, 1H), 7.02 (d, $J = 8.3$ Hz, 1H), 6.61 (s, 1H), 3.87 (s, 3H). ^{13}C NMR (101 MHz, DMSO- d_6) δ 176.1 (CO), 161.0 (Cq), 158.4 (d, $^1J_{C,F} = 143.7$ Hz, CF), 158.4 (Cq), 156.1 (d, $^3J_{C,F} = 3.4$ Hz, Cq), 134.6 (CH), 133.5 (d, $^3J_{C,F} = 9.1$ Hz, CH), 129.3 (CH), 125.2 (d, $^4J_{C,F} = 3.6$ Hz, CH), 119.3 (d, $^2J_{C,F} = 10.0$ Hz, Cq), 116.9 (d, $^2J_{C,F} = 22.1$ Hz, CH), 113.4 (CH), 113.0 (d, $^4J_{C,F} = 10.1$ Hz, CH), 110.0 (CH), 107.4 (CH), 56.2 (CH₃). ^{19}F NMR (565 MHz, DMSO- d_6) δ -148.50.

5-methoxy-2-(o-tolyl)-4H-chromen-4-one (14). Title compound was synthesized according to general procedure **B** previously described using **41b** (118 mg, 0.70 mmol) and iodine (8 mg, 0.03 mmol) in DMSO (2.3 mL). Purification by silica (elution by gradient from 100:0 to 50:50 cyclohexane/EtOAc) afforded pure compound **14** (79 mg, 75% yield). UPLC/MS Rt = 2.01 min (gradient 1), MS (ESI) m/z: 267.1 $[M+H]^+$. $[M+H]^+$ Calculated for $C_{17}H_{15}O_3$: 267.1. HRMS (AP-ESI) m/z calculated for $C_{17}H_{15}O_3$ $[M+H]^+$ 267.1021, found 267.1017. 1H NMR (400 MHz, DMSO- d_6) δ 7.69 (dd like t, $J = 8.4$ Hz, 1H), 7.59 (dd, $J = 7.7, 1.4$ Hz, 1H), 7.46 (ddd like td, $J = 7.5, 1.4$ Hz, 1H), 7.41 – 7.31 (m, 2H), 7.18 (d, $J = 8.3$ Hz, 1H), 7.01 (d, $J = 8.3$ Hz, 1H), 6.33 (s, 1H), 3.87 (s, 3H), 2.43 (s, 3H). ^{13}C NMR (101 MHz, DMSO- d_6) δ 176.2 (CO), 162.3 (Cq), 159.2

(Cq), 157.8 (Cq), 136.4 (Cq), 134.4 (CH), 131.8 (CH), 131.1 (CH), 130.7 (Cq), 129.1 (CH), 126.3 (CH), 113.6 (CH), 112.7 (Cq), 109.9 (CH), 107.3 (CH), 56.2 (CH₃), 20.1 (CH₃).

5-methoxy-2-(2-(trifluoromethyl)phenyl)-4H-chromen-4-one (15). Title compound was synthesized according to general procedure **B** previously described using **41c** (126 mg, 0.35 mmol) and iodine (5 mg, 0.02 mmol) in DMSO (1.2 mL). Purification by silica (elution by gradient from 100:0 to 94:6 DCM/MeOH) afforded pure compound **15** (108 mg, 84% yield). UPLC/MS Rt = 2.13 min (gradient 1), MS (ESI) m/z: 321.1 [M+H]⁺. [M+H]⁺ Calculated for C₁₇H₁₂F₃O₃: 321.1. HRMS (AP-ESI) m/z calculated for C₁₇H₁₂F₃O₃ [M+H]⁺ 321.0738, found 321.0744. ¹H NMR (400 MHz, DMSO-d₆) δ 8.00 – 7.93 (m, 1H), 7.89 – 7.77 (m, 3H), 7.71 (dd like t, *J* = 8.4 Hz, 1H), 7.07 (d, *J* = 8.4 Hz, 1H), 7.04 (d, *J* = 8.4 Hz, 1H), 6.42 (s, 1H), 3.88 (s, 3H). ¹³C NMR (101 MHz, DMSO-d₆) δ 176.0 (CO), 160.8 (Cq), 159.3 (Cq), 157.1 (Cq) 134.8 (CH), 133.1 (CH), 131.5 (CH), 130.4 (Cq), 127.1 (CH), 127.0 (q, ²*J*_{C,F} = 5.1 Hz, Cq), 126.8 (Cq), 123.7 (q, ¹*J*_{C,F} = 273.8 Hz, CF₃), 113.6 (Cq), 113.1 (CH), 109.7 (CH), 107.6 (CH), 56.2 (CH₃). ¹⁹F NMR (565 MHz, DMSO-d₆) δ -47.90.

2-(4-hydroxy-2-methylphenyl)-5-methoxy-4H-chromen-4-one (16). The title compound **16** was prepared following general procedure **E** using compound **42a** (181 mg, 0.49 mmol), Pd/C (40 mg), Et₃SiH (0.5 mL, 2.35 mmol) in a 1:1 mixture MeOH/DCM (12 mL). Purification by silica (elution by gradient from 100:0 to 0:100 cyclohexane/EtOAc) afforded of pure compound **16** (69 mg, 50% yield). UPLC/MS Rt = 1.64 min (gradient 1), MS (ESI) m/z: 283.1 [M+H]⁺. [M+H]⁺ calculated for C₁₇H₁₅O₄: 283.1. HRMS (AP-ESI) m/z calculated for C₁₇H₁₅O₄ [M+H]⁺ 283.0970, found 283.0965. ¹H NMR (400 MHz, DMSO-d₆) δ 7.66 (dd like t, *J* = 8.4 Hz, 1H), 7.44 (d, *J* = 8.3 Hz, 1H), 7.15 (dd, *J* = 8.4, 0.9 Hz, 1H), 6.98 (dd, *J* = 8.4, 0.9 Hz, 1H), 6.77 – 6.68 (m, 2H), 6.23 (s, 1H), 3.86 (s, 3H), 2.38 (s, 3H). ¹³C NMR (101 MHz, DMSO-d₆) δ 176.3 (CO), 163.2 (Cq), 159.6 (Cq), 159.2 (Cq), 157.8 (Cq), 138.4 (Cq), 134.1 (CH), 130.9 (CH), 122.4 (Cq), 117.9 (CH), 113.6 (Cq), 113.3 (CH), 111.6 (CH), 109.9 (CH), 107.2 (CH), 56.2 (CH₃), 20.6 (CH₃).

2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-3-methoxy-4H-chromen-4-one (17). Title compound was synthesized according to general procedure **E** previously described using **53a** (176 mg, 0.30 mmol), Pd/C 20% w/w (50 mg) and, Et₃SiH (1.12 mL, 5.4 mmol) in a 1:1 mixture MeOH/DCM (7.5 mL). Purification by silica (elution by gradient from 100:0 to 50:50 cyclohexane/EtOAc) afforded pure compound **17** (10 mg, 11% yield over two steps). UPLC/MS Rt = 1.52 min (gradient

1), MS (ESI) m/z : 317.0 $[M+H]^+$. $[M+H]^+$ calculated for $C_{16}H_{13}O_7$: 317.1. HRMS (AP-ESI) m/z calculated for $C_{16}H_{13}O_7$ $[M+H]^+$ 317.0661, found 317.0656. 1H NMR (400 MHz, DMSO- d_6) δ 12.70 (s, 1H), 7.54 (d, $J = 2.3$ Hz, 1H), 7.44 (dd, $J = 8.4, 2.3$ Hz, 1H), 6.90 (d, $J = 8.5$ Hz, 1H), 6.39 (d, $J = 2.0$ Hz, 1H), 6.18 (d, $J = 2.0$ Hz, 1H), 3.77 (s, 3H). ^{13}C NMR (101 MHz, DMSO- d_6) δ 177.9 (CO), 164.2 (Cq), 161.3 (Cq), 156.4 (Cq), 155.6 (Cq), 148.8 (Cq), 145.3 (Cq), 137.7 (Cq), 120.8 (Cq), 120.6 (CH), 115.8 (CH), 115.4 (CH), 104.2 (Cq), 98.6 (CH), 93.6 (CH), 59.7 (CH₃).

2-(3,4-dihydroxyphenyl)-3-ethoxy-5,7-dihydroxy-4H-chromen-4-one (18). Title compound was synthesized according to general procedure **E** previously described using **53b** (138 mg, 0.23 mmol), Pd/C 20% w/w (50 mg) and, Et₃SiH (0.44 mL, 2.1 mmol) in a 1:1 mixture MeOH/DCM (5.8 mL). Purification by silica (elution by gradient from 100:0 to 50:50 cyclohexane/EtOAc) afforded pure compound **18** (23 mg, 36% yield over two steps). UPLC/MS Rt = 1.64 min (gradient 1), MS (ESI) m/z : 331.0 $[M+H]^+$. $[M+H]^+$ calculated for $C_{17}H_{15}O_7$: 331.1. HRMS (AP-ESI) m/z calculated for $C_{17}H_{15}O_7$ $[M+H]^+$ 331.0817, found 331.0809. 1H NMR (400 MHz, DMSO- d_6) δ 12.74 (s, 1H), 7.57 (d, $J = 2.3$ Hz, 1H), 7.47 (dd, $J = 8.4, 2.2$ Hz, 1H), 6.89 (d, $J = 8.5$ Hz, 1H), 6.40 (d, $J = 2.1$ Hz, 1H), 6.18 (d, $J = 2.1$ Hz, 1H), 4.01 (q, $J = 7.0$ Hz, 2H), 1.26 (t, $J = 7.0$ Hz, 3H). ^{13}C NMR (101 MHz, DMSO- d_6) δ 178.1 (CO), 164.2 (Cq), 161.3 (Cq), 156.4 (Cq), 155.9 (Cq), 148.7 (Cq), 145.2 (Cq), 136.6 (Cq), 121.0 (CH), 120.7 (Cq), 115.7 (CH), 115.5 (CH), 104.1 (Cq), 98.5 (CH), 93.6 (CH), 67.7 (CH₂), 15.3 (CH₃).

3-butoxy-2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-4H-chromen-4-one (19). Title compound was synthesized according to general procedure **E** previously described using **53c** (314 mg, 0.5 mmol), Pd/C 20% w/w (50 mg) and, Et₃SiH (0.94 mL, 4.5 mmol) in a 1:1 mixture MeOH/DCM (12.5 mL). Purification by silica (elution by gradient from 100:0 to 50:50 cyclohexane/EtOAc) afforded pure compound **19** (110 mg, 62% yield over two steps). UPLC/MS Rt = 1.64 min (gradient 1), MS (ESI) m/z : 359.1 $[M+H]^+$. $[M+H]^+$ calculated for $C_{19}H_{19}O_7$: 359.1. HRMS (AP-ESI) m/z calculated for $C_{19}H_{19}O_7$ $[M+H]^+$ 359.1130, found 359.1126. 1H NMR (400 MHz, DMSO- d_6) δ 12.74 (s, 1H), 7.52 (d, $J = 2.2$ Hz, 1H), 7.44 (dd, $J = 8.4, 2.2$ Hz, 1H), 6.88 (d, $J = 8.4$ Hz, 1H), 6.44 – 6.36 (m, 1H), 6.20 – 6.15 (m, 1H), 3.92 (t, $J = 6.6$ Hz, 2H), 1.68 – 1.56 (m, 2H), 1.44 – 1.30 (m, 2H), 0.86 (t, $J = 7.4$ Hz, 3H). ^{13}C NMR (101 MHz, DMSO- d_6) δ 178.0 (CO), 164.2 (Cq), 161.3 (Cq), 156.4 (Cq), 156.0 (Cq), 148.6 (Cq), 145.2 (Cq), 136.8 (Cq), 121.0 (CH),

120.7 (Cq), 115.6 (CH), 115.6 (CH), 104.2 (Cq), 98.5 (CH), 93.5 (CH), 71.7 (CH₂), 31.5 (CH₂), 18.6 (CH₂), 13.7 (CH₃).

2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-3-(2-hydroxyethoxy)-4H-chromen-4-one (20). Title compound was synthesized according to general procedure **E** previously described using **53d** (185 mg, 0.3 mmol), Pd/C 20% w/w (50 mg) and, Et₃SiH (0.56 mL, 2.7 mmol) in a 1:1 mixture MeOH/DCM (7.5 mL). Purification by silica (elution by gradient from 100:0 to 0:100 cyclohexane/EtOAc) afforded pure compound **20** (21 mg, 20% yield over two steps). UPLC/MS Rt = 1.40 min (gradient 1), MS (ESI) m/z: 345.1 [M-H]⁻. [M-H]⁻ calculated for C₁₇H₁₃O₈: 345.1. HRMS (AP-ESI) m/z calculated for C₁₇H₁₅O₈ [M+H]⁺ 347.0766, found 347.0757. ¹H NMR (400 MHz, DMSO-d₆) δ 12.69 (s, 1H), 7.57 (d, *J* = 8.2 Hz, 2H), 6.86 (d, *J* = 8.2 Hz, 1H), 6.36 (d, *J* = 2.1 Hz, 1H), 6.15 (d, *J* = 2.1 Hz, 1H), 3.99 (t, *J* = 5.3 Hz, 2H), 3.65 (t, *J* = 5.3 Hz, 2H). ¹³C-NMR (101 MHz, DMSO-d₆) δ 177.9 (CO), 164.8 (Cq), 161.2 (Cq), 156.4 (Cq), 155.6 (Cq), 148.8 (Cq), 145.2 (Cq), 136.8 (Cq), 121.1 (CH), 120.9 (Cq), 115.7 (CH), 115.4 (CH), 103.9 (Cq), 98.7 (CH), 93.6 (CH), 73.7 (CH₂), 60.2 (CH₂).

2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-3-(3-hydroxypropoxy)-4H-chromen-4-one (21). Title compound was synthesized according to general procedure **E** previously described using **53e** (378 mg, 0.6 mmol), Pd/C 20% w/w (70 mg) and, Et₃SiH (0.56 mL, 2.7 mmol) in a 1:1 mixture MeOH/DCM (7.5 mL). Purification by silica (elution by gradient from 100:0 to 40:60 cyclohexane/EtOAc) afforded pure compound **21** (33 mg, 15% yield over two steps). UPLC/MS Rt = 1.40 min (gradient 1), MS (ESI) m/z: 361.1 [M+H]⁺. [M+H]⁺ calculated for C₁₈H₁₇O₈: 361.1. HRMS (AP-ESI) m/z calculated for C₁₈H₁₇O₈ [M+H]⁺ 361.0923, found 361.0923. ¹H NMR (400 MHz, DMSO-d₆) δ 12.72 (s, 1H), 7.53 (d, *J* = 2.2 Hz, 1H), 7.46 (dd, *J* = 8.4, 2.2 Hz, 1H), 6.89 (d, *J* = 8.4 Hz, 1H), 6.39 (d, *J* = 2.1 Hz, 1H), 6.18 (d, *J* = 2.1 Hz, 1H), 4.01 (t, *J* = 6.7 Hz, 2H), 3.49 (t, *J* = 6.4 Hz, 2H), 1.81 (p, *J* = 6.5 Hz, 2H). ¹³C NMR (101 MHz, DMSO-d₆) δ 178.0 (CO), 164.1 (Cq), 161.3 (Cq), 156.4 (Cq), 156.0 (Cq), 148.6 (Cq), 145.2 (Cq), 136.8 (Cq), 121.0 (CH), 120.8 (Cq), 115.6 (CH), 115.5 (CH), 104.2 (Cq), 98.6 (CH), 93.6 (CH), 69.7 (CH₂), 57.7 (CH₂), 32.9 (CH₂).

2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-3-methyl-4H-chromen-4-one (22). Title compound was synthesized following the general procedure **C** previously described using intermediate **57a** (59 mg, 0.16 mmol) and pyridinium chloride (185 mg, 1.6 mmol). Purification by silica (elution

by gradient from 100:0 to 80:20 DCM/MeOH) afforded pure compound **22** (34 mg, 71% yield). UPLC/MS Rt = 1.62 min (gradient 1), MS (ESI) m/z: 301.0 [M+H]⁺. [M+H]⁺ calculated for C₁₆H₁₃O₆: 301.1. HRMS (AP-ESI) m/z calculated for C₁₆H₁₃O₆ [M+H]⁺ 301.0712, found 301.0707. ¹H NMR (400 MHz, DMSO-d₆) δ 13.06 (s, 1H), 7.09 (d, *J* = 2.2 Hz, 1H), 7.02 (dd, *J* = 8.2, 2.2 Hz, 1H), 6.88 (d, *J* = 8.2 Hz, 1H), 6.30 (d, *J* = 2.1 Hz, 1H), 6.16 (t, *J* = 2.6 Hz, 1H), 2.01 (s, 3H). ¹³C NMR (101 MHz, DMSO-d₆) δ 182.2 (CO), 164.9 (Cq), 161.9 (Cq), 161.8 (Cq), 157.7 (Cq), 148.50 (Cq), 145.7 (Cq), 123.6 (Cq), 121.5 (CH), 116.5 (CH), 115.9 (CH), 113.8 (Cq), 103.1 (Cq), 99.1 (CH), 93.8 (CH), 11.3 (CH₃).

2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-3-propyl-4H-chromen-4-one (23). Title compound was synthesized following the general procedure **C** previously described using intermediate **57b** (54 mg, 0.14 mmol) and pyridinium chloride (185 mg, 1.6 mmol). Purification by silica (elution by gradient from 100:0 to 94:6 DCM/MeOH) afforded pure compound **23** (35 mg, 76% yield). UPLC/MS Rt = 1.93 min (gradient 1), MS (ESI) m/z: 329.0 [M+H]⁺. [M+H]⁺ calculated for C₁₈H₁₇O₆: 329.1. HRMS (AP-ESI) m/z calculated for C₁₈H₁₇O₆ [M+H]⁺ 329.1025, found 329.1017. ¹H NMR (400 MHz, DMSO-d₆) δ 13.09 (s, 1H), 7.02 (d, *J* = 2.1 Hz, 1H), 6.94 (dd, *J* = 8.2, 2.1 Hz, 1H), 6.89 (d, *J* = 8.2 Hz, 1H), 6.29 (d, *J* = 2.1 Hz, 1H), 6.17 (d, *J* = 2.1 Hz, 1H), 2.45 – 2.36 (m, 2H), 1.57 – 1.43 (m, 2H), 0.85 (t, *J* = 7.3 Hz, 3H). ¹³C NMR (101 MHz, DMSO-d₆) δ 182.2 (CO), 164.7 (Cq), 162.8 (Cq), 161.9 (Cq), 157.8 (Cq), 148.2 (Cq), 145.7 (Cq), 123.8 (Cq), 120.9 (CH), 118.7 (Cq), 116.2 (CH), 116.0 (CH), 103.5 (Cq), 99.0 (CH), 93.8 (CH), 27.1 (CH₂), 22.2 (CH₂), 14.5 (CH₃).

3-butyl-2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-4H-chromen-4-one (24). Title compound was synthesized following the general procedure **C** previously described using intermediate **57c** (48 mg, 0.12 mmol) and pyridinium chloride (139 mg, 1.2 mmol). Purification by silica (elution by gradient from 100:0 to 90:10 DCM/MeOH) afforded of pure compound **24** (10 mg, 30% yield). UPLC/MS Rt = 2.07 min (gradient 1), MS (ESI) m/z: 343.2 [M+H]⁺. [M+H]⁺ calculated for C₁₉H₁₉O₆: 343.1. HRMS (AP-ESI) m/z calculated for C₁₉H₁₉O₆ [M+H]⁺ 343.1181, found 343.1167. ¹H NMR (400 MHz, DMSO-d₆) δ 13.07 (s, 1H), 7.00 (d, *J* = 2.1 Hz, 1H), 6.92 (dd, *J* = 8.2, 2.1 Hz, 1H), 6.87 (d, *J* = 8.1 Hz, 1H), 6.25 (d, *J* = 2.1 Hz, 1H), 6.13 (d, *J* = 2.1 Hz, 1H), 2.46 – 2.38 (m, 2H), 1.52 – 1.40 (m, 2H), 1.32 – 1.18 (m, 2H), 0.82 (t, *J* = 7.3 Hz, 3H). ¹³C NMR (101 MHz, DMSO-d₆) δ 182.1 (CO), 165.4 (Cq), 162.6 (Cq), 161.9 (Cq), 157.8 (Cq), 148.3 (Cq), 145.7

(Cq), 123.8 (Cq), 120.8 (CH), 118.7 (CH), 116.2 (CH), 116.0 (CH), 103.3 (Cq), 99.2 (CH), 93.8 (CH), 31.0 (CH₂), 24.7 (CH₂), 22.6 (CH₂), 14.1 (CH₃).

2-(3,4-dihydroxybenzyl)-5-hydroxy-4H-chromen-4-one (32). Compound **32** was prepared according to general procedure **D** using intermediate **59** (150 mg, 0.46 mmol), BBr₃ 1M in DCM (2.1 mL, 2.06 mmol) in anhydrous DCM (6 mL). The crude was purified by silica gel flash chromatography (elution by gradient from 100:0 to 90:10 DCM/EtOAc) to yield product **32** (120 mg, 92% yield). UPLC/MS Rt = 1.77 min (gradient 1), MS (ESI) m/z: 285.0 [M+H]⁺. [M+H]⁺ calculated for C₁₆H₁₃O₅: 285.3. HRMS (AP-ESI) m/z calculated for C₁₆H₁₃O₅ [M+H]⁺ 285.0763, found 285.0752. ¹H NMR (400 MHz, DMSO-d₆) δ 12.60 (s, 1H), 8.81 (s, 1H), 7.61 (dd, *J* = 8.4, 8.4 Hz, 1H), 6.99 (dd, *J* = 8.5, 0.9 Hz, 1H), 6.78 (dd, *J* = 8.3, 0.9 Hz, 1H), 6.72 (d, *J* = 2.2 Hz, 1H), 6.70 (d, *J* = 8.0 Hz, 1H), 6.60 (dd, *J* = 8.0, 2.1 Hz, 1H), 6.24 (s, 1H), 3.85 (s, 2H). ¹³C NMR (101 MHz, DMSO-d₆) δ 183.1 (CO), 171.2 (Cq), 159.9 (Cq), 156.2 (Cq), 145.4 (Cq), 144.6 (Cq), 135.8 (CH), 125.8 (Cq), 120.0 (CH), 116.4 (CH), 115.8 (CH), 110.9 (CH), 109.8 (Cq), 108.4 (CH), 107.2 (CH), 38.7 (recovered from HSQC, CH₂).

2-(3,4-dihydroxyphenethyl)-5-hydroxy-4H-chromen-4-one (33). Compound **33** was prepared according to general procedure **D** using intermediate **63** (64 mg, 0.19 mmol), BBr₃ 1M in DCM (0.86 mL, 0.68 mmol) in anhydrous DCM (3.8 mL). The crude was purified by silica gel (elution by gradient from 100:0 to 85:15 DCM/EtOAc) to yield product **34** (45 mg, 80% yield). UPLC/MS Rt: 1.89 min (gradient 1), MS (ESI) m/z 299.0, [M+H]⁺. [M+H]⁺ calculated for C₁₇H₁₅O₅: 299.3. HRMS (AP-ESI) m/z calculated for C₁₇H₁₅O₅ [M+H]⁺ 299.0919, found 299.0915. ¹H NMR (400 MHz, DMSO-d₆) δ 12.63 (br. s, 1H), 7.63 (dd, *J* = 8.4, 8.4 Hz, 1H), 7.04 (dd, *J* = 8.5, 0.9 Hz, 1H), 6.78 (dd, *J* = 8.2, 0.9 Hz, 1H), 6.62 (d, *J* = 5.1 Hz, 1H), 6.60 (bs, 1H), 6.47 (dd, *J* = 8.0, 2.1 Hz, 1H), 6.26 (s, 1H), 2.91 (ddd, *J* = 8.7, 6.8, 1.7 Hz, 2H), 2.84 (ddd, *J* = 8.4, 6.7 Hz, 1.8 Hz, 2H). ¹³C-NMR (101 MHz, DMSO-d₆) δ 183.0 (CO), 171.3 (Cq), 159.9 (Cq), 156.3 (Cq), 145.2 (Cq), 143.7 (Cq), 135.8 (CH), 130.6 (Cq), 118.9 (CH), 115.7 (CH), 115.5 (CH), 110.8 (CH), 109.8 (Cq), 108.4 (CH), 107.3 (CH), 35.3 (CH₂), 31.4 (CH₂).

(E)-2-(3,4-dihydroxystyryl)-5-hydroxy-4H-chromen-4-one (34). Compound **34** was prepared according to general procedure **D** using: intermediate **62** (89 mg, 0.26 mmol), BBr₃ (1M in DCM) (1.2 mL, 1.17 mmol) in anhydrous DCM (5.2 mL). The crude was purified by silica gel (elution by gradient from 100:0 to 80:20 DCM/EtOAc) to yield product **17** (53 mg, 66% yield). UPLC/MS

Rt: 1.94 min (gradient 1), MS (ESI) m/z 297.0, $[M+H]^+$. $[M+H]^+$ calculated for $C_{17}H_{13}O_5$: 297.3. HRMS (AP-ESI) m/z calcd for $C_{17}H_{13}O_5$ $[M + H]^+$ 297.0763, found 297.0754. 1H NMR (400 MHz, $CDCl_3$) δ 12.80 (s, 1H), 7.64 (t, $J = 8.4$ Hz, 1H), 7.59 (d, $J = 16.0$ Hz, 1H), 7.15 – 7.09 (m, 2H), 7.05 (dd, $J = 8.2, 2.0$ Hz, 1H), 6.88 (d, $J = 16.0$ Hz, 1H), 6.81 (d, $J = 8.1$ Hz, 1H), 6.78 (d, $J = 8.1$ Hz, 1H), 6.47 (s, 1H). ^{13}C NMR (101 MHz, $DMSO-d_6$) δ 182.8 (Cq), 164.1 (Cq), 159.9 (Cq), 155.7 (Cq), 148.4 (Cq), 145.7 (Cq), 138.6 (CH), 135.7 (CH), 126.4 (Cq), 121.3 (CH), 116.1 (CH), 115.9 (CH), 114.5 (CH), 110.7 (CH), 110.1 (Cq), 107.3 (CH), 107.2 (CH).

***N*-(3,4-dihydroxyphenyl)-5-hydroxy-4-oxo-4H-chromene-2-carboxamide (35)**. Compound **35** was prepared according to general procedure **D** using intermediate **66** (45 mg, 0.13 mmol), BBr_3 1M in DCM (0.59 mL, 0.59 mmol) in anhydrous DCM (2.6 mL). The crude was purified by silica (elution by gradient from 100:0 to 98:2 DCM/MeOH) to yield product **35** (16 mg, 40% yield). UPLC/MS Rt: 1.65 min (gradient 1), MS (ESI) m/z 312.0, $[M-H]^-$. $[M-H]^-$ calculated for $C_{16}H_{10}NO_6$: 312.3. HRMS (AP-ESI) m/z calcd for $C_{16}H_{11}NO_6$ $[M + H]^+$ 314.0665, found 314.0662. 1H NMR (400 MHz, $DMSO-d_6$) δ 12.30 (br. s, 1H), 10.48 (br. s, 1H), 9.01 (br. s, 1H), 7.77 (dd, $J = 8.4, 8.4$ Hz, 1H), 7.30 (d, $J = 2.5$ Hz, 1H), 7.26 (d, $J = 8.4$ Hz, 1H), 7.03 (dd, $J = 8.5, 2.5$ Hz, 1H), 6.95 (s, 1H), 6.88 (d, $J = 8.3$ Hz, 1H), 6.74 (d, $J = 8.5$ Hz, 1H). ^{13}C NMR (101 MHz, $DMSO-d_6$) δ 183.5 (CO), 159.7 (CONH), 157.2 (Cq), 156.3 (Cq), 155.4 (Cq), 145.0 (Cq), 142.9 (Cq), 136.7 (CH), 129.2 (Cq), 115.2 (CH), 112.5 (CH), 111.4 (CH), 110.9 (CH), 109.7 (CH), 109.6 (Cq), 108.1 (CH).

1-(2-hydroxy-4,6-dimethoxyphenyl)ethan-1-one (36a). CH_3COOH (1.5 mL, 27.24 mmol) and $BF_3 \cdot Et_2O$ (6.7 mL, 54.49 mmol) were added to a solution of 3,5-dimethoxyphenol **43** (3.0 g, 19.46 mmol) in anhydrous EtOAc (13 mL). The reaction mixture was refluxed for 3 hours under an argon atmosphere. Then, it was quenched with water (20 mL) and extracted with EtOAc (100 mL). Combined organic layers were washed with $NaHCO_3$ (25 mL x2), dried over $MgSO_4$ and concentrated under vacuum. Purification by silica gel chromatography (elution by gradient from 100:0 to 90:10 cyclohexane/EtOAc) afforded the pure product as a white show solid (2.8 g, 73% yield). UPLC/MS: Rt = 2.07 min (gradient 1). MS (ESI) m/z : 197.0 $[M + H]^+$, $[M + H]^+$ calculated for $C_{10}H_{13}O_4$: 197.1. 1H NMR (400 MHz, $CDCl_3$) δ 14.02 (s, 1H), 6.06 (d, $J = 2.3$ Hz, 1H), 5.92 (d, $J = 2.4$ Hz, 1H), 3.85 (s, 3H), 3.82 (s, 3H), 2.61 (s, 3H).

1-(2-hydroxy-4-methoxyphenyl)ethan-1-one (36b). K₂CO₃ (245 mg, 1.84 mmol) and MeI (0.21 mL, 3.45 mmol) were added 2,5-dihydroxyacetophenone **44** (350 mg, 2.30 mmol) in anhydrous acetone (8.0 mL). The reaction mixture was refluxed for 210 minutes under an argon atmosphere. Then, the solvent was removed under vacuum and the residue was raised with DCM (10 mL) and HCl 0.5 M (5 mL). Organic phase was collected, dried over MgSO₄ and concentrated under vacuum. Purification by silica gel chromatography (elution by gradient from 60:40 to 40:60 cyclohexane/EtOAc) afforded the pure product as a colorless oil (289 mg, 76% yield). UPLC/MS: Rt = 1.86 min (gradient 1). MS (ESI) *m/z*: 166.9 [M + H]⁺, [M + H]⁺ calculated for C₉H₁₀O₃: 167.1. ¹H NMR (400 MHz, CDCl₃) δ 12.74 (s, 1H), 7.63 (d, *J* = 8.7 Hz, 1H), 6.48 – 6.40 (m, 2H), 3.84 (s, 3H), 2.56 (s, 3H).

1-(4-fluoro-2-hydroxy-6-methoxyphenyl)ethan-1-one (36c). To a solution of **46** (239 mg, 1.2 mmol) in anhydrous DCM (2.5 mL) was added 1M BBr₃ in DCM (1.2 mL, 1.2 mmol) at – 78 °C. The solution was warmed to room temperature and stirred for 30 minutes under an argon atmosphere. Then, reaction mixture was quenched with water (5 mL) and extracted with DCM (10 mL). The organic layer was washed with water (4 mL x2), dried over anhydrous MgSO₄ and evaporated *in vacuo* to afford the title compound (175 mg, 79% yield). UPLC/MS: Rt = 2.04 min (gradient 1). MS (ESI) *m/z*: 185.0 [M + H]⁺, [M + H]⁺ calculated for C₉H₁₀FO₃: 185.1. ¹H NMR (400 MHz, CDCl₃) δ 13.75 (d, *J* = 1.4 Hz, 1H), 6.27 (dd, *J* = 10.2, 2.4 Hz, 1H), 6.12 (dd, *J* = 11.0, 2.4 Hz, 1H), 3.89 (s, 3H), 2.64 (s, 3H).

1-(2-hydroxy-6-methoxyphenyl)ethan-1-one (36d). To a solution of 2,6-dihydroxyacetophenone **47** (2000 mg, 13.16 mmol) in anhydrous acetone (18.0 mL), K₂CO₃ (1453 mg, 10.53 mmol) and MeI (0.90 mL, 14.48 mmol) were added. The reaction mixture was refluxed for 3 hours under argon atmosphere. Then, the solution was concentrated under concentrated under vacuum and raised with DCM (25 mL) and water (25 mL). Organic layer was divided, washed with water (10 mL), dried over magnesium sulfate and concentrated under vacuum. Purification by silica gel chromatography (elution by isocratic 80:20 cyclohexane/EtOAc) afforded the pure product as a pale yellow solid (1936 mg, 89% yield). UPLC/MS: Rt = 1.97 min (gradient 1). MS (ESI) *m/z*: 166.9 [M+H]⁺, [M+H]⁺ calculated for C₉H₁₁O₃, 167.1. ¹H NMR (400 MHz, CDCl₃) δ 13.23 (s, 1H), 7.34 (dd, *J* = 8.3 Hz, 1H), 6.57 (dd, *J* = 8.4, 1.0 Hz, 1H), 6.39 (dd, *J* = 8.3, 0.9 Hz, 1H), 3.90 (s, 3H), 2.67 (s, 3H).

(E)-3-(3,4-dimethoxyphenyl)-1-(2-hydroxy-4-methoxyphenyl)prop-2-en-1-one (38b). The title compound was synthesized according to general procedure **A** using **36b** (250 mg, 1.5 mmol), 3,4-dimethoxybenzaldehyde **37c** (274 mg, 1.65 mmol) and KOH (1685 mg, 30.0 mmol) in anhydrous MeOH (12.5 mL). Trituration with ethanol (3 mL) afforded the pure compound **38b** (404 mg, 85% yield). UPLC/MS: Rt = 2.39 min (gradient 1). MS (ESI) m/z : 315.1 $[M+H]^+$, $[M+H]^+$ calculated for $C_{18}H_{19}O_5$, 315.1. 1H NMR (400 MHz, DMSO- d_6) δ 13.63 (s, 1H), 8.29 (d, J = 9.0 Hz, 1H), 7.89 (d, J = 15.4 Hz, 1H), 7.79 (d, J = 15.3 Hz, 1H), 7.57 (d, J = 2.0 Hz, 1H), 7.42 (dd, J = 8.3, 2.0 Hz, 1H), 7.04 (d, J = 8.4 Hz, 1H), 6.57 (dd, J = 9.0, 2.5 Hz, 1H), 6.51 (d, J = 2.5 Hz, 1H), 3.86 (s, 3H), 3.85 (s, 3H), 3.83 (s, 3H).

(E)-3-(3,4-dimethoxyphenyl)-1-(2-hydroxy-6-methoxyphenyl)prop-2-en-1-one (38c). The title compound was synthesized according to general procedure **A** using **36d** (250 mg, 1.5 mmol), 3,4-dimethoxybenzaldehyde **37a** (274 mg, 1.65 mmol) and KOH (1685 mg, 30.0 mmol) in anhydrous MeOH (12.5 mL). Trituration with EtOH (3 mL) afforded the pure compound **38c** (304 mg, 65% yield). UPLC/MS: Rt = 2.29 min (gradient 1). MS (ESI) m/z : 315.2 $[M+H]^+$, $[M+H]^+$ calculated for $C_{18}H_{19}O_5$: 315.1. 1H -NMR (400 MHz, DMSO- d_6) δ 10.29 (s, 1H), 7.31 – 7.18 (m, 4H), 7.07 (d, J = 15.9 Hz, 1H), 6.98 (d, J = 8.3 Hz, 1H), 6.56 (d, J = 12.9 Hz, 1H), 6.53 (d, J = 12.9 Hz, 1H), 3.80 (s, 3H), 3.79 (s, 3H), 3.73 (s, 3H).

(E)-3-(3,4-dimethoxyphenyl)-1-(4-fluoro-2-hydroxy-6-methoxyphenyl)prop-2-en-1-one (38d). Title compound was synthesized according to general procedure **A** previously described using **36c** (150 mg, 0.46 mmol), 3,4-dimethoxybenzaldehyde **37a** (84 mg, 0.51 mmol), KOH (515 mg, 9.2 mmol) in anhydrous MeOH (3.9 mL). The crude was purified by trituration with EtOH (1.5 mL) to yield title intermediate **38d** (120 mg, 92% yield). UPLC/MS Rt = 2.47 min (gradient 1), MS (ESI) m/z 333.5, $[M+H]^+$. $[M+H]^+$ calculated for $C_{18}H_{19}FO_5$: 333.3. 1H -NMR (400 MHz, $CDCl_3$) δ 13.90 (d, J = 1.4 Hz, 1H), 7.81 (d, J = 15.5 Hz, 1H), 7.72 (d, J = 15.5 Hz, 1H), 7.23 (dd, J = 8.3, 2.0 Hz, 1H), 7.13 (d, J = 2.0 Hz, 1H), 6.91 (d, J = 8.3 Hz, 1H), 6.32 (dd, J = 10.1, 2.4 Hz, 1H), 6.16 (dd, J = 11.0, 2.4 Hz, 1H), 3.95 (s, 3H), 3.94 (s, 3H), 3.94 (s, 3H).

(E)-3-(3,4-dimethoxyphenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one (38e). Title compound was synthesized according to general procedure **A** previously described using **36e** (250 mg, 1.84 mmol), 3,4-dimethoxybenzaldehyde **37c** (336 mg, 2.02 mmol), KOH (2064 mg, 36.8 mmol) in anhydrous MeOH (15 mL). The crude was purified by trituration with EtOH (2.0 mL) to yield title intermediate **38e** (385 mg, 74% yield). UPLC/MS Rt = 2.38 min (gradient 1), MS (ESI) m/z 285.1

[M+H]⁺. [M+H]⁺ calculated for C₁₇H₁₇O₄: 285.1. ¹H NMR (400 MHz, DMSO-d₆) δ 12.77 (s, 1H), 8.29 (dd, *J* = 8.0, 1.6 Hz, 1H), 7.93 (d, *J* = 15.4 Hz, 1H), 7.82 (d, *J* = 15.4 Hz, 1H), 7.60 – 7.53 (m, 2H), 7.44 (dd, *J* = 8.3, 2.0 Hz, 1H), 7.07 – 6.97 (m, 3H), 3.87 (s, 3H), 3.83 (s, 3H).

(E)-1-(2-hydroxy-4,6-dimethoxyphenyl)-3-(4-methoxyphenyl)prop-2-en-1-one (38f). Title compound was synthesized according to general procedure **A** previously described using **36a** (250 mg, 1.28 mmol), p-methoxybenzaldehyde **37a** (192 mg, 1.41 mmol), KOH (1431 mg, 25.5 mmol) in anhydrous MeOH (11 mL). The crude was purified by trituration with EtOH (2.0 mL) to yield title intermediate **38f** (332 mg, 82% yield). UPLC/MS Rt = 2.61 min (gradient 1), MS (ESI) m/z 315.3 [M+H]⁺. [M+H]⁺ calculated for C₁₈H₁₉O₅: 315.1. ¹H NMR (400 MHz, DMSO-d₆) δ 13.54 (s, 1H), 7.73 – 7.66 (m, 2H), 7.65 (s, 2H), 7.05 – 6.97 (m, 2H), 6.15 (d, *J* = 2.3 Hz, 1H), 6.12 (d, *J* = 2.3 Hz, 1H), 3.89 (s, 3H), 3.82 (s, 6H).

(E)-1-(2-hydroxy-4,6-dimethoxyphenyl)-3-(3-methoxyphenyl)prop-2-en-1-one (38g). Title compound was synthesized according to general procedure **A** previously described using **36a** (150 mg, 1.10 mmol), m-methoxybenzaldehyde **37b** (164 mg, 1.20 mmol), KOH (1234 mg, 22.0 mmol) in anhydrous MeOH (9 mL). The crude was purified by trituration with EtOH (2.0 mL) to yield title intermediate **38g** (245 mg, 83% yield). UPLC/MS Rt = 2.63 min (gradient 1), MS (ESI) m/z 315.2 [M+H]⁺. [M+H]⁺ calculated for C₁₈H₁₉O₅: 315.1. ¹H NMR (400 MHz, DMSO-d₆) δ 13.31 (s, 1H), 7.74 (d, *J* = 15.7 Hz, 1H), 7.60 (d, *J* = 15.7 Hz, 1H), 7.41 – 7.33 (m, 1H), 7.32 – 7.28 (m, 1H), 7.26 (dd like t, *J* = 2.0 Hz, 1H), 7.01 (ddd, *J* = 8.2, 2.6, 1.0 Hz, 1H), 6.16 (d, *J* = 2.4 Hz, 1H), 6.13 (d, *J* = 2.3 Hz, 1H).

(E)-1-(2-hydroxy-4,6-dimethoxyphenyl)-3-phenylprop-2-en-1-one (38h). Title compound was synthesized according to general procedure **A** previously described using **36a** (150 mg, 1.10 mmol), benzaldehyde **37d** (90 mg, 1.20 mmol), KOH (1234 mg, 22.0 mmol) in anhydrous MeOH (10 mL). The crude was purified by trituration with EtOH (2.0 mL) to yield title intermediate **38h** (158 mg, 72% yield). UPLC/MS Rt = 2.63 min (gradient 1), MS (ESI) m/z 285.1 [M+H]⁺. [M+H]⁺ calculated for C₁₇H₁₇O₄: 285.1. ¹H-NMR (400 MHz, DMSO-d₆) δ 13.38 (s, 1H), 7.76 (d, *J* = 15.7 Hz, 1H), 7.74 – 7.70 (m, 2H), 7.64 (d, *J* = 15.7 Hz, 1H), 7.45 (dd, *J* = 5.0, 2.0 Hz, 3H), 6.23 – 6.08 (m, 2H), 3.90 (s, 3H), 3.83 (s, 3H).

(E)-3-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)-1-(2-hydroxy-4,6-dimethoxyphenyl)prop-2-en-1-one (38i). Title compound was synthesized according to general procedure **A** previously

described using **36a** (200 mg, 1.0 mmol), 2,2-Difluoro-1,3-benzodioxole-5-carboxaldehyde **37e** (210 mg, 1.1 mmol), KOH (1680 mg, 20 mmol) in anhydrous MeOH (9 mL). The crude was purified by trituration with EtOH (3.0 mL) to yield title intermediate **38i** (200 mg, 55% yield). UPLC/MS Rt = 1.98 min (gradient 1), MS (ESI) m/z 365.0 [M+H]⁺. [M+H]⁺ calculated for C₁₈H₁₆F₂O₆: 365.3. ¹H NMR (400 MHz, CDCl₃) δ 14.19 (s, 1H), 7.79 (d, *J* = 15.5 Hz, 1H), 7.70 (d, *J* = 15.6 Hz, 1H), 7.32 (d, *J* = 1.6 Hz, 1H), 7.30 (dd, *J* = 8.8, 1.7 Hz, 1H), 7.08 (d, *J* = 8.7 Hz, 1H), 6.11 (d, *J* = 2.3 Hz, 1H), 5.97 (d, *J* = 2.4 Hz, 1H), 3.92 (s, 3H), 3.84 (s, 3H).

Methyl (E)-5-(3-(2-hydroxy-4,6-dimethoxyphenyl)-3-oxoprop-1-en-1-yl)-2-methoxybenzoate (38k). A mixture of aldehyde **37f** (200 mg, 1 mmol) and pyrrolidine (0.03 mL, 0.31 mmol) in DCM (4 mL) stirred on air at 30°C for 10 minutes. After that compound **36a** (195 mg, 1 mmol) was added and the reaction mixture stirred for 20 h. Then the solvent was removed under vacuum. Purification by silica (elution by gradient from 100:0 to 80:20 cyclohexane/EtOAc) gave pure compound **38k** (93 mg, 25% yield). UPLC/MS Rt = 2.43 min (gradient 1), MS (ESI) m/z 373.1 [M+H]⁺. [M+H]⁺ calculated for C₂₀H₂₁O₇: 373.4. ¹H NMR (400 MHz, CDCl₃) δ 8.08 (d, *J* = 2.4 Hz, 1H), 7.83 (d, *J* = 15.6 Hz, 1H), 7.74 (d, *J* = 15.6 Hz, 1H), 7.70 (dd, *J* = 8.7, 2.4 Hz, 1H), 7.02 (d, *J* = 8.7 Hz, 1H), 6.11 (d, *J* = 2.3 Hz, 1H), 5.97 (d, *J* = 2.4 Hz, 1H), 3.96 (s, 3H), 3.93 (s, 6H), 3.84 (s, 3H).

2-(3,4-dimethoxyphenyl)-7-methoxy-4H-chromen-4-one (39b). Compound **39b** was synthesized according to general procedure **B** using chalcone **38b** (200 mg, 0.64 mmol) and I₂ (7.6 mg, 0.03 mmol) in DMSO (2.1 mL). The crude was purified by trituration with EtOH (1.5 mL) to yield title intermediate **39b** (139 mg, 70% yield). UPLC/MS Rt = 1.96 min (gradient 1), MS (ESI) m/z 313.4 [M+H]⁺. [M+H]⁺ calculated for C₁₈H₁₇O₅: 313.1. ¹H NMR (400 MHz, DMSO-d₆) δ 7.94 (d, *J* = 8.8 Hz, 1H), 7.72 (dd, *J* = 8.5, 2.2 Hz, 1H), 7.60 (d, *J* = 2.2 Hz, 1H), 7.34 (d, *J* = 2.4 Hz, 1H), 7.15 (d, *J* = 8.6 Hz, 1H), 7.07 (dd, *J* = 8.8, 2.4 Hz, 1H), 6.96 (s, 1H), 3.94 (s, 3H), 3.90 (s, 3H), 3.87 (s, 3H).

2-(3,4-dimethoxyphenyl)-5-methoxy-4H-chromen-4-one (39c). Compound **39c** was synthesized according to general procedure **B** using chalcone **38c** (202 mg, 0.64 mmol), I₂ (7.6 mg, 0.03 mmol) in DMSO (2.1 mL). The crude was purified by trituration with EtOH (1.5 mL) to yield title intermediate **39b** (160 mg, 80% yield). UPLC/MS Rt = 1.82 min (gradient 1), MS (ESI) m/z 313.2 [M+H]⁺. [M+H]⁺ calculated for C₁₈H₁₇O₅: 313.1. ¹H NMR (400 MHz, DMSO-d₆) δ

7.73 – 7.62 (m, 2H), 7.55 (d, $J = 2.2$ Hz, 1H), 7.32 – 7.25 (m, 1H), 7.12 (d, $J = 8.6$ Hz, 1H), 6.98 (d, $J = 8.3$ Hz, 1H), 6.85 (s, 1H), 3.88 (s, 3H), 3.86 (s, 3H), 3.84 (s, 3H).

2-(3,4-dimethoxyphenyl)-7-fluoro-5-methoxy-4H-chromen-4-one (39d). Compound **39d** was synthesized according to general procedure **B** using chalcone **38d** (150 mg, 0.45 mmol) and I_2 (6.0 mg, 0.02 mmol) in DMSO (1.5 mL). The crude was purified by trituration with EtOH (1.0 mL) to yield title intermediate **39b** (120 mg, 80% yield). UPLC/MS Rt = 1.93 min (gradient 1), MS (ESI) m/z 331.1 $[M+H]^+$. $[M+H]^+$ calculated for $C_{18}H_{16}FO_5$: 331.3. 1H NMR (400 MHz, $CDCl_3$) δ 7.50 (dd, $J = 8.5, 2.2$ Hz, 1H), 7.31 (d, $J = 2.1$ Hz, 1H), 6.97 (d, $J = 8.5$ Hz, 1H), 6.82 (dd, $J = 9.0, 2.4$ Hz, 1H), 6.64 (s, 1H), 6.57 (dd, $J = 11.1, 2.4$ Hz, 1H), 3.99 (s, 3H), 3.97 (s, 3H), 3.95 (s, 3H).

2-(3,4-dimethoxyphenyl)-4H-chromen-4-one (39e). Compound **39e** was synthesized according to general procedure **B** using chalcone **38e** (200 mg, 0.70 mmol) and I_2 (7.8 mg, 0.04 mmol) in DMSO (2.3 mL). The crude was purified by trituration with EtOH (1.0 mL) to yield title intermediate **39e** (157 mg, 80% yield). UPLC/MS Rt = 1.94 min (gradient 1), MS (ESI) m/z 283.2 $[M+H]^+$. $[M+H]^+$ calculated for $C_{17}H_{15}O_4$: 283.1. 1H NMR (400 MHz, $DMSO-d_6$) δ 8.04 (dd, 1H), 7.86 – 7.78 (m, 2H), 7.73 (dd, $J = 8.5, 2.2$ Hz, 1H), 7.62 (d, $J = 2.2$ Hz, 1H), 7.50 (ddd, $J = 8.1, 6.3, 1.9$ Hz, 1H), 7.15 (d, $J = 8.6$ Hz, 1H), 7.05 (s, 1H), 3.90 (s, 3H), 3.86 (s, 3H).

5,7-dimethoxy-2-(4-methoxyphenyl)-4H-chromen-4-one (39f). Compound **39f** was synthesized according to general procedure **B** using chalcone **38f** (200 mg, 0.64 mmol) and I_2 (7.6 mg, 0.03 mmol) in DMSO (2.1 mL). The crude was purified by trituration with EtOH (1.0 mL) to yield title intermediate **39e** (136 mg, 70% yield). UPLC/MS Rt = 1.98 min (gradient 1), MS (ESI) m/z 313.4 $[M+H]^+$. $[M+H]^+$ calculated for $C_{18}H_{17}O_5$: 313.1. 1H NMR (400 MHz, $DMSO-d_6$) δ 8.03 – 7.95 (m, 2H), 7.14 – 7.05 (m, 2H), 6.85 (d, $J = 2.3$ Hz, 1H), 6.66 (s, 1H), 6.50 (d, $J = 2.3$ Hz, 1H), 3.90 (s, 3H), 3.85 (s, 3H), 3.83 (s, 3H).

5,7-dimethoxy-2-(3-methoxyphenyl)-4H-chromen-4-one (39g). Compound **39g** was synthesized according to general procedure **B** using chalcone **38g** (245 mg, 0.78 mmol) and I_2 (7.6 mg, 0.03 mmol) in DMSO (2.6 mL). The crude was purified by trituration with EtOH (1.5 mL) to yield title intermediate **39g** (185 mg, 70% yield). UPLC/MS Rt = 2.04 min (gradient 1), MS (ESI) m/z 313.0 $[M+H]^+$. $[M+H]^+$ calculated for $C_{18}H_{17}O_5$: 313.1. 1H NMR (400 MHz, $DMSO-d_6$) δ 7.62 (dd, $J = 8.0, 1.1$ Hz, 1H), 7.58 – 7.53 (m, 1H), 7.46 (t, $J = 8.0$ Hz, 1H), 7.22 – 7.10 (m, 1H),

6.89 (d, $J = 2.3$ Hz, 1H), 6.81 (s, 1H), 6.51 (d, $J = 2.3$ Hz, 1H), 3.90 (s, 3H), 3.86 (s, 3H), 3.84 (s, 3H).

5,7-dimethoxy-2-phenyl-4H-chromen-4-one (39h). Compound **39h** was synthesized in two steps. In step 1, general procedure **B** was applied using chalcone **38h** (155 mg, 0.24 mmol) and I₂ (7.6 mg, 0.03 mmol) in DMSO (0.8 mL) obtaining the desired compound **39h** (146 mg) in a mixture with the corresponding 3-iodinated byproduct (20 mg). In step 2, bis(triphenylphosphine)palladium chloride (3.5 mg, 0.005 mmol) and HCOONa (6.8 mg, 0.1 mmol) were added to a solution of the crude dissolved in anhydrous DMF (1.0 mL) and the reaction mixture was stirred at 100 °C for 2 hours. Then, the reaction mixture was diluted with water (5 mL) and DCM (5 mL). Organic phase was collected, dried over MgSO₄ and concentrated at low pressure. Purification by silica gel chromatography (elution by gradient from 100:0 to 98:2 DCM/MeOH) afforded the pure product as a white show solid (94 mg, 60% yield over 2 steps). UPLC/MS Rt = 1.97 min (gradient 1), MS (ESI) m/z 283.0 [M+H]⁺. [M+H]⁺ calculated for C₁₇H₁₅O₄: 283.1. ¹H NMR (400 MHz, DMSO-d₆) δ 8.10 – 8.01 (m, 2H), 7.62 – 7.50 (m, 3H), 6.88 (d, $J = 2.3$ Hz, 1H), 6.77 (s, 1H), 6.52 (d, $J = 2.3$ Hz, 1H), 3.91 (s, 3H), 3.84 (s, 3H).

2-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)-5,7-dimethoxy-4H-chromen-4-one (39i). Compound **39i** was synthesized according to general procedure **B** using chalcone **38i** (180 mg, 0.51 mmol) and iodine (6.4 mg, 0.02 mmol) in DMSO (1.7 mL). The crude was purified by trituration with EtOH (1.0 mL) to yield title intermediate **39i** (140 mg, 79% yield). UPLC/MS Rt = 1.06 min (gradient 2), MS (ESI) m/z 363.3 [M+H]⁺. [M+H]⁺ calculated for C₁₈H₁₃F₂O₆: 363.3. ¹H NMR (400 MHz, DMSO-d₆) δ 7.64 (dd, $J = 8.4, 1.8$ Hz, 1H), 7.57 (d, $J = 1.8$ Hz, 1H), 7.17 (d, $J = 8.5$ Hz, 1H), 6.61 (d, $J = 0.9$ Hz, 1H), 6.55 (d, $J = 2.2$ Hz, 1H), 6.38 (d, $J = 2.3$ Hz, 1H), 3.95 (s, 3H), 3.91 (s, 3H).

methyl 5-(5,7-dimethoxy-4-oxo-4H-chromen-2-yl)-2-methoxybenzoate (39k). Compound **39k** was synthesized according to general procedure **B** using chalcone **38k** (150 mg, 0.40 mmol) and iodine (2.0 mg, 0.02 mmol) in DMSO (1.4 mL). The crude was purified by trituration with EtOH (1.0 mL) to yield title intermediate **39k** (120 mg, 79% yield). UPLC/MS Rt = 1.83 min (gradient 2), MS (ESI) m/z 371.0 [M+H]⁺. [M+H]⁺ calculated for C₂₀H₁₉O₇: 371.3. ¹H NMR (400 MHz, CDCl₃) δ 8.35 (d, $J = 2.4$ Hz, 1H), 7.97 (dd, $J = 8.9, 2.5$ Hz, 1H), 7.09 (d, $J = 8.9$ Hz, 1H), 6.74 (s,

1H), 6.60 (d, $J = 2.3$ Hz, 1H), 6.39 (d, $J = 2.3$ Hz, 1H), 3.99 (s, 3H), 3.96 (s, 3H), 3.95 (s, 3H), 3.93 (s, 3H).

(E)-3-(2-fluorophenyl)-1-(2-hydroxy-6-methoxyphenyl)prop-2-en-1-one (41a). Title compound was synthesized according to general procedure **A** previously described using **36d** (100 mg, 0.60 mmol), 2-fluorobenzaldehyde **40a** (82.0 mg, 0.66 mmol) and KOH (675 mg, 12 mmol) in anhydrous MeOH (5 mL). The crude was purified by trituration with EtOH (2.0 mL) to yield title intermediate **41a** (96 mg, 59% yield). UPLC/MS Rt = 2.52 min (gradient 1), MS (ESI) m/z: 273.1 [M+H]⁺. [M+H]⁺ Calculated for C₁₆H₁₅FO₃ : 273.1. ¹H NMR (400 MHz, CDCl₃) δ 13.10 (s, 1H), 7.98 (d, $J = 15.8$ Hz, 1H), 7.88 (d, $J = 15.8$ Hz, 1H), 7.61 (ddd like td, $J = 7.6, 1.8$ Hz, 1H), 7.42 – 7.32 (m, 2H), 7.19 (ddd like td, $J = 7.5, 1.1$ Hz, 1H), 7.15 (d, $J = 1.1$ Hz, 1H), 6.62 (dd, $J = 8.4, 1.0$ Hz, 1H), 6.44 (dd, $J = 8.4, 1.0$ Hz, 1H), 3.95 (s, 3H).

(E)-1-(2-hydroxy-6-methoxy-phenyl)-3-(o-tolyl)prop-2-en-1-one (41b). Title compound was synthesized according to general procedure **A** previously described using **36d** (100 mg, 0.60 mmol), 2-metilbenzaldehyde **40b** (0.08 mL, 0.66 mmol) and KOH (675 mg, 12 mmol) in anhydrous MeOH (5 mL). Purification by silica (elution by gradient from 100:0 to 95:5 cyclohexane/EtOAc) afforded of pure compound **41b** (118 mg, 73% yield). UPLC/MS Rt = 2.60 min (gradient 1), MS (ESI) m/z: 269.1 [M+H]⁺. [M+H]⁺ Calculated for C₁₇H₁₈O₃: 269.1. ¹H NMR (400 MHz, CDCl₃) δ 13.15 (s, 1H), 8.11 (d, $J = 15.5$ Hz, 1H), 7.79 (d, $J = 15.5$ Hz, 1H), 7.69 – 7.62 (m, 1H), 7.37 (dd like t, $J = 8.3$ Hz, 1H), 7.33 – 7.21 (m, 3H), 6.62 (d, $J = 8.3$ Hz, 1H), 6.44 (d, $J = 8.3$ Hz, 1H), 3.95 (s, 3H), 2.51 (s, 3H).

(E)-1-(2-hydroxy-6-methoxyphenyl)-3-(2-(trifluoromethyl)phenyl)prop-2-en-1-one (41c). Title compound was synthesized according to general procedure **A** previously described using **36d** (100 mg, 0.60 mmol), 2-(trifluoromethyl)benzaldehyde **40c** (0.09 mL, 0.66 mmol) and KOH (675 mg, 12 mmol) in anhydrous MeOH (5 mL). The crude was purified by trituration with EtOH (2.0 mL) to yield title intermediate **41c** (126 mg, 65% yield). UPLC/MS Rt = 2.64 min (gradient 1), MS (ESI) m/z: 323.1 [M+H]⁺. [M+H]⁺ Calculated for C₁₇H₁₅F₃O₃: 323.1. ¹H NMR (400 MHz, CDCl₃) δ 13.01 (s, 1H), 8.17 – 8.07 (m, 1H), 7.81 (d, $J = 1.9$ Hz, 1H), 7.78 (d, $J = 6.0$ Hz, 1H), 7.76 – 7.70 (m, 1H), 7.60 (t, $J = 7.6$ Hz, 1H), 7.49 (t, $J = 7.7$ Hz, 1H), 7.38 (t, $J = 8.3$ Hz, 1H), 6.63 (dd, $J = 8.4, 1.0$ Hz, 1H), 6.43 (dd, $J = 8.3, 1.0$ Hz, 1H), 3.94 (s, 3H).

(E)-3-(4-(benzyloxy)-2-methylphenyl)-1-(2-hydroxy-6-methoxyphenyl)prop-2-en-1-one

(41d). Title compound was synthesized according to general procedure **A** previously described using **36d** (100 mg, 0.60 mmol), 4-benzyloxy-2-methyl-benzaldehyde **40d** (149 mg, 0.66 mmol) and KOH (675 mg, 12 mmol) in anhydrous MeOH (5 mL). The crude was purified by trituration with EtOH (2.0 mL) to yield title intermediate **41c** (224 mg, quantitative yield). UPLC/MS Rt = 2.21 min (gradient 2), MS (ESI) m/z : 375.1 $[M+H]^+$. $[M+H]^+$ calculated for C₂₄H₂₃O₄: 375.1. ¹H NMR (400 MHz, DMSO-d₆) δ 7.71 (d, J = 8.6 Hz, 1H), 7.58 (d, J = 15.8 Hz, 1H), 7.48 – 7.30 (m, 5H), 7.26 (dd like t, J = 8.3 Hz, 1H), 7.01 (d, J = 15.8 Hz, 1H), 6.97 – 6.85 (m, 2H), 6.63 – 6.47 (m, 2H), 5.14 (s, 2H), 3.75 (s, 3H), 2.27 (s, 3H).

2-(4-(benzyloxy)-2-methylphenyl)-5-methoxy-4H-chromen-4-one (42a). Title compound was synthesized according to general procedure **B** previously described using **41d** (224 mg, 0.59 mmol) and iodine (7 mg, 0.03 mmol) in DMSO (2.0 mL). Purification by silica (elution by gradient from 100:0 to 20:80 cyclohexane/EtOAc) afforded pure compound **42a** (181 mg, 83% yield). UPLC/MS Rt = 1.39 min (gradient 1), MS (ESI) m/z : 373.1 $[M+H]^+$. $[M+H]^+$ Calculated for C₂₄H₂₂O₄: 373.1. ¹H NMR (400 MHz, DMSO-d₆) δ 7.67 (dd like t, J = 8.4 Hz, 1H), 7.55 (d, J = 8.6 Hz, 1H), 7.49 – 7.44 (m, 2H), 7.44 – 7.37 (m, 2H), 7.37 – 7.30 (m, 1H), 7.16 (dd, J = 8.5, 0.9 Hz, 1H), 7.05 (d, J = 2.6 Hz, 1H), 7.02 – 6.95 (m, 2H), 6.28 (s, 1H), 5.18 (s, 2H), 3.87 (s, 3H), 2.43 (s, 3H).

1-(4-fluoro-2,6-dimethoxyphenyl)ethan-1-one (46). To a suspension of AlCl₃ (2048 mg, 15.4 mmol) in anhydrous DCM (10 mL) at 0° C were added 1-fluoro-3,5-dimethoxy-benzene **45** (1.0 mL, 7.7 mmol) and CH₃COCl (0.6 mL, 8.5 mmol). The reaction mixture was stirred at 0°C for 1 hour under an argon atmosphere. Then, reaction mixture was diluted with water (15 mL) and extracted with DCM (3x30 mL). Combined organic layers were dried over MgSO₄ and concentrated under vacuum. Purification by silica gel chromatography (elution by gradient from 100:0 to 95:5 cyclohexane/EtOAc) afforded the pure product as a white solid (239 mg, 16% yield). UPLC/MS: Rt = 1.77 min (gradient 1). MS (ESI) m/z : 199.0 $[M + H]^+$, $[M + H]^+$ calculated for C₁₀H₁₂FO₃: 199.1. ¹H NMR (400 MHz, CDCl₃) δ 6.29 (d, J = 10.7 Hz, 2H), 3.78 (s, 6H), 2.46 (s, 3H).

7-(benzyloxy)-2-(3,4-bis(benzyloxy)phenyl)-5-hydroxy-4H-chromen-4-one (48). K₂CO₃ (198 mg, 1.42 mmol) and benzyl bromide (0.12 mL, 1.02 mmol) were added to a solution of luteolin **1** (100 mg, 0.34 mmol) in dry DMF (6 mL). Reaction mixture was stirred under an argon atmosphere

for 4 hours at 0 °C. After reaction completion, reaction mixture was diluted with water (10 mL) and extracted with EtOAc (2x20 mL). Combined organic layers were dried over MgSO₄ and concentrated under vacuum. Purification by silica (elution by gradient from 100:0 to 50:50 cyclohexane/EtOAc) afforded pure compound **48** (113 mg, 60% yield). UPLC/MS Rt = 2.70 min (gradient 2), MS (ESI) m/z: 557.1 [M+H]⁺. [M+H]⁺ calculated for C₃₆H₂₉O₆: 557.2. ¹H NMR (400 MHz, CDCl₃) δ 12.77 (s, 1H), 7.52 – 7.28 (m, 17H), 7.02 (d, *J* = 8.5 Hz, 1H), 6.52 (d, *J* = 2.3 Hz, 1H), 6.50 (s, 1H), 6.44 (d, *J* = 2.2 Hz, 1H), 5.25 (s, 2H), 5.24 (s, 2H), 5.15 (s, 2H).

7-(benzyloxy)-2-(3,4-bis(benzyloxy)phenyl)-5-methoxy-4H-chromen-4-one (49). MeI (0.02 mL, 0.22 mmol) and K₂CO₃ (30 mg, 0.22 mmol) were added to a solution of **48** (60 mg, 0.11 mmol) in anhydrous DMF (1.5 mL). The reaction mixture was stirred under an argon atmosphere at 75 °C for 3 hours. After reaction completion, the reaction mixture was diluted with water (5 mL) and extracted with DCM (10 mL x2). Combined organic phases were dried over MgSO₄ and concentrated under vacuum. Purification by silica (elution by gradient from 100:0 to 0:100 cyclohexane/EtOAc) afforded pure compound **49** (39 mg, 63%). UPLC/MS Rt = 2.23 min (gradient 2), MS (ESI) m/z: 571.1 [M+H]⁺. [M+H]⁺ calculated for C₃₇H₃₁O₆: 570.0. ¹H NMR (400 MHz, CDCl₃) δ 7.52 – 7.28 (m, 17H), 7.01 (d, *J* = 8.4 Hz, 1H), 6.60 (d, *J* = 2.3 Hz, 1H), 6.58 (s, 1H), 6.46 (d, *J* = 2.3 Hz, 1H), 5.23 (d, *J* = 3.9 Hz, 4H), 5.16 (s, 2H), 3.94 (s, 3H).

7-(benzyloxy)-2-(3,4-bis(benzyloxy)phenyl)-5-hydroxy-3-(((2R,3S,4R,5R,6S)-3,4,5-trihydroxy-6-(((2S,3S,4S,5S,6R)-3,4,5-trihydroxy-6-methyltetrahydro-2H-pyran-2-yl)oxy)methyl)tetrahydro-2H-pyran-2-yl)oxy)-4H-chromen-4-one (51). To a solution of rutin **50** (2000 mg, 2.96 mmol) in DMF (20 mL) were added K₂CO₃ (1716 mg, 12.44 mmol) and benzoyl bromide (2.8 mL, 23.68 mmol). The reaction mixture was stirred overnight at room temperature. Then it was diluted with EtOAc (60 mL). The organic phase was divided, washed with water (2x50 mL), dried over MgSO₄ and concentrated under vacuum yielding of crude **51**, which was used without further purification for the next step (2000 mg). UPLC/MS Rt = 2.44 min (gradient 1), MS (ESI) m/z: 881.3 [M+H]⁺. [M+H]⁺ calculated for C₄₈H₄₉O₁₆: 881.3.

7-(benzyloxy)-2-(3,4-bis(benzyloxy)phenyl)-3,5-dihydroxy-4H-chromen-4-one (52). Intermediate **51** (2000 mg, 2.27 mmol) was dissolved in EtOH (14 mL) and added with HCl 37% (2 mL). The reaction mixture was refluxed for 2 hours. After complete conversion of starting material, the reaction mixture was cooled to room temperature and filtered. The precipitate was

washed with water (5 mL) and cold MeOH (5 mL) yielding pure **52** (1394 mg, yield: 82% over two steps). UPLC/MS Rt = 2.55 min (gradient 2), MS (ESI) m/z: 573.2 [M+H]⁺. [M+H]⁺ calculated for C₃₆H₂₉O₇: 573.2. ¹H NMR (400 MHz, DMSO-d₆) δ 11.70 (s, 1H), 7.85 (d, *J* = 2.1 Hz, 1H), 7.77 (dd, *J* = 8.5, 2.0 Hz, 1H), 7.54 – 7.29 (m, 15H), 7.04 (d, *J* = 8.7 Hz, 1H), 6.52 (d, *J* = 2.2 Hz, 1H), 6.45 (d, *J* = 2.2 Hz, 1H), 5.25 (s, 4H), 5.15 (s, 2H).

7-(benzyloxy)-2-(3,4-bis(benzyloxy)phenyl)-5-hydroxy-3-methoxy-4H-chromen-4-one (53a).

K₂CO₃ (175 mg, 0.30 mmol) and MeI (0.03 mL, 0.45 mmol) were added to a solution of intermediate **52** (175 mg, 0.30 mmol) in anhydrous DMF (10 mL) and the reaction mixture was stirred under an argon atmosphere for 3 hours at room temperature. After reaction completion, the reaction mixture was poured into water and extracted with EtOAc (20 mL x2). Combined organic layer was divided, dried over MgSO₄ and concentrated under vacuum yielding crude **53a**, that was used for the next step without further purification (176 mg). UPLC/MS Rt = 2.66 min (gradient 2), MS (ESI) m/z: 587.1 [M+H]⁺. [M+H]⁺ calculated for C₃₇H₃₁O₇: 587.2.

7-(benzyloxy)-2-(3,4-bis(benzyloxy)phenyl)-3-ethoxy-5-hydroxy-4H-chromen-4-one (53b).

K₂CO₃ (95 mg, 0.69 mmol) and iodoethane (0.03 mL, 0.34 mmol) were added to a solution of intermediate **52** (132 mg, 0.23 mmol) in anhydrous DMF (15 mL) and the reaction mixture was stirred under an argon atmosphere for 3 hours at room temperature. After reaction completion, the reaction mixture was poured into water and extracted with ethyl acetate (20 mL x2). Combined organic layer was divided, dried over MgSO₄ and concentrated under vacuum yielding crude **53b**, that was used for the next step without further purification (138 mg). UPLC/MS Rt = 2.70 min (gradient 2), MS (ESI) m/z: 601.0 [M+H]⁺. [M+H]⁺ calculated for C₃₈H₃₃O₇: 601.2.

7-(benzyloxy)-2-(3,4-bis(benzyloxy)phenyl)-3-butoxy-5-hydroxy-4H-chromen-4-one (53c).

K₂CO₃ (207 mg, 1.50 mmol) and 1-bromobutane (0.07 mL, 0.70 mmol) were added to a solution of intermediate **52** (300 mg, 0.50 mmol) in anhydrous DMF (10 mL) and the reaction mixture was stirred under an argon atmosphere for 3 hours at room temperature. After reaction completion, the reaction mixture was poured into water and extracted with EtOAc (20 mL x2). Combined organic layer was divided, dried over MgSO₄, filtered and concentrated under vacuum yielding crude **53c**, that was used for the next step without further purification (314 mg). UPLC/MS Rt = 2.72 min (gradient 2), MS (ESI) m/z: 629.0 [M+H]⁺. [M+H]⁺ calculated for C₄₀H₃₇O₇: 629.2.

7-(benzyloxy)-2-(3,4-bis(benzyloxy)phenyl)-5-hydroxy-3-(2-hydroxyethoxy)-4H-chromen-4-one (53d). K₂CO₃ (124 mg, 0.9 mmol) and 2-bromoethanol (0.03 mL, 0.45 mmol) were added to a solution of intermediate **52** (180 mg, 0.30 mmol) in anhydrous DMF (10 mL) and the reaction mixture was stirred overnight under an argon atmosphere at room temperature. After reaction completion, the reaction mixture was poured into water and extracted with EtOAc (20 mL x2). Combined organic layer was divided, dried over MgSO₄, filtered and concentrated under vacuum yielding crude **53d**, that was used for the next step without further purification (185 mg). UPLC/MS Rt = 2.58 min (gradient 2), MS (ESI) m/z: 615.0 [M-H]⁻. [M-H]⁻ calculated for C₃₈H₃₁O₈: 615.2.

7-(benzyloxy)-2-(3,4-bis(benzyloxy)phenyl)-5-hydroxy-3-(3-hydroxypropoxy)-4H-chromen-4-one (53e). K₂CO₃ (248 mg, 0.18 mmol) and 3-bromo-1-propanol (0.14 mL, 0.9 mmol) were added to a solution of intermediate **52** (350 mg, 0.60 mmol) in anhydrous DMF (10 mL) and the reaction mixture was stirred under an argon atmosphere for three hours at room temperature. After reaction completion, the reaction mixture was poured into water and extracted with EtOAc (20 mL x2). Combined organic layer was divided, dried over MgSO₄, filtered and concentrated under vacuum yielding crude **53e**, that was used for the next step without further purification (378 mg). UPLC/MS Rt = 2.40 min (gradient 2), MS (ESI) m/z: 631.0 [M+H]⁺. [M+H]⁺ calculated for C₃₉H₃₅O₈: 631.2.

1-(2-hydroxy-4,6-dimethoxyphenyl)propan-1-one (55a). BBr₃ 1M in DCM (2.14 mL, 2.14 mmol) was slowly added to a solution of 3,5-dimethoxyphenol **43** (300 mg, 1.95 mmol) in anhydrous DCM (3 mL) and the reaction mixture was stirred for 5 minutes at room temperature under an argon atmosphere. Then, propanoyl chloride **54a** (0.19 mL, 2.14 mmol) was added dropwise and the reaction mixture was stirred at room temperature for 6 hours. After complete conversion of starting material, reaction mixture was cooled to 0 °C, acidified with HCl 1M and the organic phase was separated. Aqueous phase was extracted with DCM (15 mL x2) and combined organic layers were dried over MgSO₄, filtered, and concentrated under vacuum. Trituration with methanol (3 mL) gave the pure product **55a** (214 mg, 53% yield). UPLC/MS Rt = 2.30 min (gradient 1), MS (ESI) m/z: 211.1 [M+H]⁺. [M+H]⁺ calculated for C₁₁H₁₄O₄: 211.1. ¹H NMR (400 MHz, CDCl₃) δ 14.08 (s, 1H), 6.07 (d, *J* = 2.4 Hz, 1H), 5.92 (d, *J* = 2.4 Hz, 1H), 3.85 (s, 3H), 3.81 (s, 3H), 3.02 (q, *J* = 7.2 Hz, 2H), 1.15 (t, *J* = 7.2 Hz, 3H).

1-(2-hydroxy-4,6-dimethoxyphenyl)pentan-1-one (55b). AlCl₃ (780 mg, 5.85 mmol) was added to a solution of 3,5-dimethoxyphenol **43** (300 mg, 1.95 mmol) in anhydrous DCM (5 mL) and the reaction mixture was stirred at room temperature for 30 minutes under an argon atmosphere. Then, valeryl chloride **54b** (0.26 mL, 2.14 mmol) was slowly added and the reaction mixture was stirred at room temperature for 1 hour. After complete conversion of starting materials, reaction mixture was cooled to 0 °C and poured into water (5 mL). Organic phase was collected and aqueous phase was extracted with DCM (15 mL x2). Combined organic layers were dried over MgSO₄, filtered, and concentrated under vacuum. Purification by silica (elution by gradient from 100:0 to 90:10 cyclohexane/EtOAc) afforded pure compound **55b** (275 mg, 59% yield). UPLC/MS Rt = 2.64 min (gradient 1), MS (ESI) m/z: 239.0 [M+H]⁺. [M+H]⁺ calculated for C₁₃H₁₉O₄: 239.1. ¹H NMR (400 MHz, DMSO-d₆) δ 13.56 (s, 1H), 6.11 (d, *J* = 2.3 Hz, 1H), 6.08 (d, *J* = 2.3 Hz, 1H), 3.85 (s, 3H), 3.80 (s, 3H), 2.92 (t, *J* = 7.4 Hz, 2H), 1.55 (p, *J* = 7.4 Hz, 3H), 1.40 – 1.25 (m, 3H), 0.90 (t, *J* = 7.3 Hz, 3H).

1-(2-hydroxy-4,6-dimethoxyphenyl)hexan-1-one (55c). AlCl₃ (780 mg, 5.85 mmol) was added to a solution of 3,5-dimethoxyphenol **43** (300 mg, 1.95 mmol) in anhydrous DCM (5 mL) and the reaction mixture was stirred at room temperature for 30 minutes under an argon atmosphere. Then, hexanoyl chloride **54c** (0.19 mL, 1.95 mmol) was slowly added and the reaction mixture was stirred at room temperature for 2 hours. After complete conversion of starting materials, reaction mixture was cooled to 0 °C and poured into water (5 mL). Organic phase was collected and aqueous phase was extracted with DCM (15 mL x2). Combined organic layers were dried over MgSO₄, filtered, and concentrated under vacuum. Purification by silica (elution by gradient from 100:0 to 90:10 cyclohexane/EtOAc) afforded pure compound **55c** (282 mg, 57% yield). UPLC/MS Rt = 1.89 min (gradient 2), MS (ESI) m/z: 253.2 [M+H]⁺. [M+H]⁺ calculated for C₁₄H₂₁O₄: 253.0. ¹H NMR (400 MHz, CDCl₃) δ 14.10 (s, 1H), 6.07 (d, *J* = 2.4 Hz, 1H), 5.92 (d, *J* = 2.4 Hz, 1H), 3.85 (s, 3H), 3.82 (s, 3H), 3.01 – 2.93 (m, 2H), 1.66 (p, *J* = 7.3 Hz, 2H), 1.39 – 1.29 (m, 4H), 0.95 – 0.87 (m, 3H).

(E)-3-(3,4-dimethoxyphenyl)-1-(2-hydroxy-4,6-dimethoxyphenyl)-2-methylprop-2-en-1-one (56a). The title compound was synthesized according to general procedure **A** using **55a** (100 mg, 0.48 mmol), 3,4-dimethoxybenzaldehyde **37a** (87 mg, 0.53 mmol) and KOH (539 mg, 9.6 mmol) in anhydrous MeOH (4.0 mL). Purification by silica (elution by gradient from 100:0 to 70:30

cyclohexane/EtOAc) afforded pure compound **56a** (80 mg, 47% yield). UPLC/MS Rt = 2.12 min (gradient 1), MS (ESI) m/z: 359.1 [M+H]⁺. [M+H]⁺ calculated for C₂₀H₂₃O₆: 359.1. ¹H NMR (400 MHz, DMSO-d₆) δ 7.11 (d, *J* = 1.6 Hz, 1H), 7.00 (s, 2H), 6.97 (s, 1H), 6.12 (d, *J* = 2.2 Hz, 1H), 6.07 (d, *J* = 2.1 Hz, 1H), 3.78 (s, 3H), 3.76 (s, 3H), 3.74 (s, 3H), 3.66 (s, 3H), 2.09 (d, *J* = 1.3 Hz, 3H).

(E)-2-(3,4-dimethoxybenzylidene)-1-(2-hydroxy-4,6-dimethoxyphenyl)pentan-1-one (56b).

The title compound was synthesized according to general procedure **A** using **55b** (150 mg, 0.60 mmol), 3,4-dimethoxybenzaldehyde **37a** (109 mg, 0.66 mmol) and KOH (707 mg, 12.6 mmol) in anhydrous MeOH (5.0 mL). Purification by silica (elution by gradient from 100:0 to 30:70 cyclohexane/EtOAc) afforded pure compound **56b** (116 mg, 50% yield). UPLC/MS Rt = 2.40 min (gradient 1), MS (ESI) m/z: 387.2 [M+H]⁺. [M+H]⁺ calculated for C₂₂H₂₇O₆: 387.1.

¹H NMR (400 MHz, DMSO-d₆) δ 9.70 (s, 1H), 7.04 (s, 1H), 7.01 (d, *J* = 8.1 Hz, 1H), 6.97 – 6.89 (m, 2H), 6.12 (d, *J* = 2.1 Hz, 1H), 6.07 (d, *J* = 2.1 Hz, 1H), 3.78 (s, 3H), 3.76 (s, 3H), 3.74 (s, 3H), 3.65 (s, 3H), 1.57 – 1.44 (m, 2H), 1.17 (t, *J* = 7.1 Hz, 2H), 0.96 (t, *J* = 7.3 Hz, 3H).

(E)-2-(3,4-dimethoxybenzylidene)-1-(2-hydroxy-4,6-dimethoxyphenyl)hexan-1-one (56c).

The title compound was synthesized according to general procedure **A** using **55c** (150 mg, 0.60 mmol), 3,4-dimethoxybenzaldehyde **37a** (109 mg, 0.66 mmol) and KOH (707 mg, 12.6 mmol) in anhydrous MeOH (5.0 mL). Purification by silica (elution by gradient from 100:0 to 60:40 cyclohexane/EtOAc) afforded pure compound **56c** (81 mg, 35% yield). UPLC/MS Rt = 2.50 min (gradient 1), MS (ESI) m/z: 401.0 [M+H]⁺. [M+H]⁺ calculated for C₂₃H₂₉O₆: 401.2. ¹H NMR (400 MHz, DMSO-d₆) δ 9.70 (s, 1H), 7.06 – 6.97 (m, 2H), 6.94 (d, *J* = 2.0 Hz, 1H), 6.92 (d, *J* = 2.3 Hz, 1H), 6.12 (d, *J* = 2.1 Hz, 1H), 6.07 (d, *J* = 2.1 Hz, 1H), 3.78 (s, 3H), 3.76 (s, 3H), 3.74 (s, 3H), 3.65 (s, 3H), 2.62 – 2.53 (m, 2H), 1.51 – 1.42 (m, 2H), 1.42 – 1.31 (m, 2H), 0.90 (t, *J* = 7.2 Hz, 3H).

2-(3,4-dimethoxyphenyl)-5,7-dimethoxy-3-methyl-4H-chromen-4-one (57a).

Title compound was synthesized according to general procedure **B** previously described using **56a** (77 mg, 0.22 mmol) and iodine (2.7 mg, 0.01 mmol) in DMSO (0.7 mL). Trituration with EtOH (1 mL) afforded pure compound **57a** (59 mg, 77% yield). UPLC/MS Rt = 1.96 min (gradient 1), MS (ESI) m/z: 357.0 [M+H]⁺. [M+H]⁺ calculated for C₂₀H₂₁O₆: 357.1. ¹H NMR (400 MHz, DMSO-d₆) δ 7.28 –

7.20 (m, 2H), 7.14 – 7.08 (m, 1H), 6.67 (d, $J = 2.3$ Hz, 1H), 6.48 (d, $J = 2.3$ Hz, 1H), 3.87 (s, 3H), 3.86 – 3.80 (m, 9H), 1.96 (s, 3H).

2-(3,4-dimethoxyphenyl)-5,7-dimethoxy-3-propyl-4H-chromen-4-one (57b). Title compound was synthesized according to general procedure **B** previously described using **56b** (116 mg, 0.30 mmol) and iodine (4.0 mg, 0.01 mmol) in DMSO (1.0 mL). Purification by silica (elution by gradient from 75:15 to 50:50 cyclohexane/EtOAc) afforded pure compound **57b** (54 mg, 50% yield). UPLC/MS Rt = 2.23 min (gradient 1), MS (ESI) m/z: 385.0 [M+H]⁺. [M+H]⁺ calculated for C₂₂H₂₅O₆: 385.2. ¹H NMR (400 MHz, DMSO-d₆) δ 7.20 – 7.15 (m, 2H), 7.13 (s, 1H), 6.64 (d, $J = 2.3$ Hz, 1H), 6.48 (d, $J = 2.3$ Hz, 1H), 3.86 (s, 3H), 3.84 (s, 3H), 3.83 (s, 3H), 3.81 (s, 3H), 2.38 – 2.29 (m, 2H), 1.53 – 1.39 (m, 2H), 0.82 (t, $J = 7.3$ Hz, 3H).

3-butyl-2-(3,4-dimethoxyphenyl)-5,7-dimethoxy-4H-chromen-4-one (57c). Title compound was synthesized according to general procedure **B** previously described using **56c** (81 mg, 0.20 mmol) and iodine (4 mg, 0.01 mmol) in DMSO (0.7 mL). Trituration with EtOH (1 mL) afforded pure compound **57c** (48 mg, 61% yield). UPLC/MS Rt = 2.38 min (gradient 1), MS (ESI) m/z: 399.0 [M+H]⁺. [M+H]⁺ calculated for C₂₃H₂₇O₆: 399.2. ¹H NMR (400 MHz, DMSO-d₆) δ 7.20 – 7.14 (m, 2H), 7.14 – 7.08 (m, 1H), 6.64 (d, $J = 2.3$ Hz, 1H), 6.47 (d, $J = 2.3$ Hz, 1H), 3.85 (s, 3H), 3.83 (s, 3H), 3.83 (s, 3H), 3.81 (s, 3H), 2.36 (t, $J = 7.9$ Hz, 2H), 1.42 (p, $J = 7.5$ Hz, 2H), 1.23 (h, $J = 7.3$ Hz, 2H), 0.81 (t, $J = 7.3$ Hz, 3H).

4-(3,4-dimethoxyphenyl)-1-(2-hydroxy-6-methoxyphenyl)butane-1,3-dione (58). Compound **58** was prepared following general procedure **F** using **36d** (200 mg, 1.2 mmol), ethyl 2-(3,4-dimethoxyphenyl)acetate **57** (538 mg, 2.4 mmol), NaH 60% dispersion in mineral oil (192 mg, 4.8 mmol) in THF dry (6 mL). The crude was used as such without further purification. UPLC/MS Rt = 1.85 min (gradient 1), MS (ESI) m/z: 345.0 [M+H]⁺. [M+H]⁺ calculated for C₁₉H₂₁O₆: 345.4.

2-(3,4-dimethoxybenzyl)-5-methoxy-4H-chromen-4-one (59). Compound **59** was prepared following general procedure **G** using: crude **58** (412 mg), HCl 37% (1 mL) in MeOH (7 mL). Reaction mixture stirred at room temperature for 20 hours. Purification by silica (elution by gradient from 100:0 to 50:50 DCM/EtOAc) afforded the pure title compound **59** as white powder (145 mg, yield 37 % over 2 steps). UPLC/MS Rt = 1.80 min (gradient 1), MS (ESI) m/z: 327.1 [M+H]⁺. [M+H]⁺ calculated for C₁₉H₁₉O₅: 327.3. ¹H NMR (400 MHz, CDCl₃) δ 7.50 (dd, $J = 8.41$

Hz, 1H), 6.96 (d, $J = 8.35$ Hz, 1H), 6.83 (br s, 2H), 6.79-6.77 (m, 2H), 6.06 (s, 1H), 3.95 (s, 3H), 3.97 (s, 3H), 3.86 (s, 3H), 3.48 (s, 2H).

1-(2-hydroxy-6-methoxyphenyl)butane-1,3-dione (60). Compound **60** was prepared following general procedure **F** using **36d** (1383 mg, 8.3 mmol), NaH 60% dispersion in mineral oil (1328 mg, 33.2 mmol) in a 5:1 mixture THF/EtOAc (24 mL). The obtained crude was used as such, without further purifications. UPLC/MS Rt = 1.85 min (gradient 1), MS (ESI) m/z: 345.0 [M+H]⁺. [M+H]⁺ calculated for C₁₉H₂₁O₆: 345.4.

5-methoxy-2-methyl-4H-chromen-4-one (61). Compound **61** was prepared following general procedure **G** using: crude **60** (1480 g), HCl 37% (1 mL), in MeOH (20 mL). Purification by typical silica gel (elution by gradient from 100:0 to 75:25 cyclohexane/EtOAc) afforded the pure title compound **61** as white powder (1.18 g, 84 % yield over 2 steps). UPLC/MS Rt = 1.43 min (gradient 1), MS (ESI) m/z: 190.9 [M+H]⁺. [M+H]⁺ calculated for C₁₁H₁₁O₃: 191.2. ¹H NMR (400 MHz, CDCl₃) δ 7.51 (dd, $J = 8.4$ Hz, 1H), 6.98 (dd, $J = 8.5, 1.0$ Hz, 1H), 6.78 (dd, $J = 8.4, 1.0$ Hz, 1H), 6.10 (d, $J = 0.9$ Hz, 1H), 3.97 (s, 3H), 2.31 (d, $J = 0.7$ Hz, 3H).

(E)-2-(3,4-dimethoxystyryl)-5-methoxy-4H-chromen-4-one (62). Sodium ethoxide (55 mg, 0.31 mmol) was added to a solution of compound **61** (100 mg, 0.62 mmol) in EtOH (2 mL), followed by addition of a solution of 3,4-dimethoxybenzaldehyde **36a** (110 mg, 0.62 mmol) in EtOH (1 mL). Reaction mixture stirred at 50°C for 2 hours. After that, HCl 2M aq was added to reaction mixture until pH=4, the resulted precipitate was filtered and washed with water (0.5 mL) and dried under vacuum to afford pure title compound **62** (140 mg, 74% yield). UPLC/MS Rt: 2.11 min (gradient 1), MS (ESI) m/z 339.0 [M+H]⁺. [M+H]⁺ calculated for C₂₀H₁₈O₅: 339.1. ¹H NMR (400 MHz, CDCl₃) δ 7.57 (dd, $J = 8.4$ Hz, 1H), 7.51 (d, $J = 16.0$ Hz, 1H), 7.15 (dd, $J = 8.4, 2.0$ Hz, 1H), 7.10 (dd, $J = 8.4, 1.0$ Hz, 1H), 7.09 (d, $J = 2.3$ Hz, 1H), 6.90 (d, $J = 8.3$ Hz, 1H), 6.82 (dd, $J = 8.4, 0.9$ Hz, 1H), 6.61 (d, $J = 16.0$ Hz, 1H), 6.38 (s, 1H), 3.99 (s, 3H), 3.97 (s, 3H), 3.93 (s, 3H).

2-(3,4-dimethoxyphenethyl)-5-methoxy-4H-chromen-4-one (63). Ammonium formate (21 mg, 0.30 mmol) and Pd(OH)₂/C (15 mg, 15% m/m) were added to a solution of intermediate **62** (100 mg, 0.29 mmol) in MeOH (3 mL). Reaction mixture stirred at 80°C for 4 hours. After that the mixture was filtered over a pad of celite, rinsed with MeOH (10 mL), and concentrated under vacuum. Purification by silica (elution by gradient from 100:0 to 60:40 DCM/EtOAc) afforded pure title compound **63** (64 mg, 63% yield). UPLC/MS Rt: 1.89 min

(gradient 1), MS (ESI) m/z 341.0 $[M+H]^+$. $[M+H]^+$ Calculated for $C_{20}H_{21}O_5$: 341.4. 1H NMR (400 MHz, $CDCl_3$) δ 7.52 (dd, $J = 8.4, 8.4$ Hz, 1H), 6.99 (dd, $J = 8.4, 0.9$ Hz, 1H), 6.81 – 6.76 (m, 2H), 6.75 – 6.69 (m, 2H), 6.08 (s, 1H), 3.97 (s, 3H), 3.85 (s, 3H), 3.82 (s, 3H), 2.97 (dd, $J = 9.4, 7.3$ Hz, 2H), 2.84 (dd, $J = 9.3, 7.2$ Hz, 2H).

ethyl 5-methoxy-4-oxo-4H-chromene-2-carboxylate (64). NaOEt (210 mg, 3.0 mmol) was dissolved in absolute EtOH (4 mL). A mixture of diethyl oxalate (310 mg, 2.1 mmol) and **36d** (100 mg, 0.6 mmol) in absolute EtOH (2 mL) was slowly added to NaOEt solution. The solution was refluxed for 2 hours until complete consumption of starting material. Then the mixture was allowed to cool to room temperature and neutralized with HCl 2M. The mixture was extracted with EtOAc (5 mL x3), collected organic layer were dried over Na_2SO_4 , filtered, and concentrated under vacuum. Purification by silica (elution by gradient from 100:0 to 75:25 cyclohexane/EtOAc) afforded pure title compound **64** (120 mg, 79% yield). UPLC/MS Rt: 1.50 min (gradient 1), MS (ESI) m/z 249.0, $[M+H]^+$. $[M+H]^+$ Calculated for $C_{13}H_{13}O_5$: 249.3. 1H NMR (400 MHz, $CDCl_3$) δ 7.40 (dd, $J = 8.4, 8.4$ Hz, 1H), 6.60 – 6.58 (m, 2H), 6.42 (d, $J = 8.4$ Hz, 1H), 4.38 (q, $J = 7.2$ Hz, 2H), 3.93 (s, 3H), 1.34 (t, $J = 7.2$ Hz, 3H).

5-methoxy-4-oxo-4H-chromene-2-carboxylic acid (65). K_2CO_3 (100 mg, 0.60 mmol) was added to a solution of intermediate **64** (100 mg, 0.40 mmol) in a 3:1 mixture THF/EtOH (4 mL). Reaction mixture stirred for 6 hours at 50°C until complete consumption of starting material, then HCl 2M aq. was added until pH=5, and the mixture was extracted with EtOAc (4 mL x3). Collected organic layers were dried over Na_2SO_4 , filtered and concentrated under vacuum to yield pure compound **65** (70 mg, 79% yield). UPLC/MS Rt: 0.80 min (gradient 1), MS (ESI) m/z 220.9, $[M+H]^+$. $[M+H]^+$ Calculated for $C_{11}H_9O_5$: 221.2. 1H NMR (400 MHz, $DMSO-d_6$) δ 7.72 (dd, $J = 8.4, 8.4$ Hz, 1H), 7.18 (d, $J = 8.4$ Hz, 1H), 7.02 (d, $J = 8.3$ Hz, 1H), 3.86 (s, 3H).

N-(3,4-dimethoxyphenyl)-5-methoxy-4-oxo-4H-chromene-2-carboxamide (66). HATU (170 mg, 0.43 mmol) and DIPEA (0.23 mL, 1.3 mmol) were sequentially added to a solution of intermediate **65** (60 mg, 0.29 mmol) in a 3:1 mixture DMF/DCM (4 mL) under argon. Reaction mixture stirred at room temperature for 15 minutes, after that 3,4-dimethoxyaniline (44 mg, 0.29 mmol) was added and reaction mixture stirred for other 4 hours until complete consumption of starting material. Water (1 mL) and HCl 2M aq. were added until pH=7, the mixture was extracted with DCM (3 mL x3), collected organic layers were dried over Na_2SO_4 , filtered and concentrated under vacuum. Purification by silica (elution by gradient from 100:0 to 55:45 DCM/EtOAc)

afforded pure title compound **66** (92 mg, 90%). UPLC/MS Rt: 1.66 min (gradient 1), MS (ESI) m/z 356.1, $[M+H]^+$. $[M+H]^+$ Calculated for $C_{19}H_{18}NO_6$: 356.3. 1H NMR (400 MHz, DMSO- d_6) δ 7.78 (dd, $J = 8.4, 8.4$ Hz, 1H), 7.44 (d, $J = 2.4$ Hz, 1H), 7.36 (dd, $J = 8.6, 2.4$ Hz, 1H), 7.34 (d, $J = 8.3$ Hz, 1H), 7.05 (d, $J = 8.3$ Hz, 1H), 6.98 (d, $J = 8.7$ Hz, 1H), 6.76 (s, 1H), 3.88 (s, 3H), 3.77 (s, 3H), 3.76 (s, 3H).

***In vitro* microsomal stability.** 10mM DMSO stock solution of test compound was pre-incubated at 37°C for 15min with mouse liver microsomes added 0.1M Tris-HCl buffer (pH 7.4). The final concentration was 4.6 μ M. After pre-incubation, the co-factors (NADPH, G6P, G6PDH and MgCl₂ pre-dissolved in 0.1M Tris-HCl) were added to the incubation mixture and the incubation was continued at 37°C for 1h. At each time point (0, 5, 15, 30, 60min), 30 μ L of incubation mixture was diluted with 200 μ L cold CH₃CN spiked with 200nM of internal standard, followed by centrifugation at 3500g for 15min. The supernatant was further diluted with H₂O (1:1) for analysis. The concentration of test compound was quantified by LC/MS-MS on a Waters ACQUITY UPLC/MS TQD system consisting of a TQD (Triple Quadrupole Detector) Mass Spectrometer equipped with an Electrospray Ionization interface. The analyses were run on an ACQUITY UPLC BEH C18 (50x2.1mmID, particle size 1.7 μ m) with a VanGuard BEH C18 pre-column (5x2.1mmID, particle size 1.7 μ m) at 40°C, using 0.1% HCOOH in H₂O (A) and 0.1% HCOOH in CH₃CN (B) as mobile phase. Electrospray ionization (ESI) was applied in positive mode. The percentage of test compound remaining at each time point relative to $t=0$ was calculated. The half-lives ($t_{1/2}$) were determined by an one-phase decay equation using a non-linear regression of compound concentration versus time.

***In vitro* Plasma Stability.** 10mM DMSO stock solution of test compound was diluted 50-fold with DMSO-H₂O (1:1) and incubated at 37°C for 2 h with mouse plasma added 5% DMSO (pre-heated at 37°C for 10 min). The final concentration was 2 μ M. At each time point (0, 5, 15, 30, 60, 120min), 50 μ L of incubation mixture was diluted with 200 μ L cold CH₃CN spiked with 200nM of internal standard, followed by centrifugation at 3500g for 20min. The supernatant was further diluted with H₂O (1:1) for analysis. The concentration of test compound was quantified by LC/MS-MS on a Waters ACQUITY UPLC/MS TQD system consisting of a TQD (Triple Quadrupole Detector) Mass Spectrometer equipped with an Electrospray Ionization interface. The analyses were run on an ACQUITY UPLC BEH C18 (50x2.1mmID, particle size 1.7 μ m)

with a VanGuard BEH C18 precolumn (5x2.1mmID, particle size 1.7 μ m) at 40°C, using 0.1% HCOOH in H₂O (A) and 0.1% HCOOH in CH₃CN (B) as mobile phase. Electrospray ionization (ESI) was applied in positive mode. The response factors, calculated on the basis of the internal standard peak area, were plotted over time. When possible, response vs. time profiles were fitted with Prism (GraphPad Software, Inc., USA) to estimate compounds half-life in plasma.

Aqueous kinetic solubility. The aqueous kinetic solubility was determined from a 10mM DMSO stock solution of test compound in Phosphate Buffered Saline (PBS) at pH 7.4. The study was performed by incubation of an aliquot of 10mM DMSO stock solution in PBS (pH 7.4) at a target concentration of 250 μ M resulting in a final concentration of 2.5% DMSO. The incubation was carried out under shaking at 25°C for 24h followed by centrifugation at 21,100g for 30min. The supernatant was analyzed by UPLC/MS for the quantification of dissolved compound by UV at a specific wavelength (215nm). The analyses were performed on a Waters ACQUITY UPLC/MS SQD system consisting of a SQD (Single Quadrupole Detector) Mass Spectrometer equipped with Electrospray Ionization interface. The analyses were run on an ACQUITY UPLC BEH C18 column (50x2.1mmID, particle size 1.7 μ m) with a VanGuard BEH C18 pre-column (5x2.1mmID, particle size 1.7 μ m), using 10mM NH₄OAc in H₂O at pH 5 adjusted with AcOH (A) and 10mM NH₄OAc in MeCN-H₂O (95:5) at pH 5 (B) as mobile phase.

Docking and virtual screening. The structure-based virtual screening was performed using the Schrodinger's Maestro suite.⁴² The D3 internal library (~ 15 000 unique drug-like molecules) was processed with LigPrep⁴³ utility to assign ligands' charges, convert 2D to 3D structures, and generate all tautomeric and ionization states (at pH=7.0 \pm 0.4). These prepared library was used as the input for the docking calculations, which were performed with Glide utility.⁴⁴ Specifically the docking grid of 18 Å was centered on the substrate's (namely, ATP) center-of-mass, and the enhanced planarity for conjugated π groups was considered during the computations. Docking calculations employed a structural ensemble of Pol- η , generated from molecular dynamics (MD) simulations. This ensemble was representative of the transition state (TS) structure for Pol- η catalysis, as obtained from previous classical and quantum simulations performed in our laboratory.³⁴ In detail, we considered three structures with either two, one or none of the two catalytic metal ion (MgA-MgB). When the receptor structure included one or two catalytic metals, the coordination of at least one of them was imposed during the docking to ensure the modeling

of compounds' conformation able to form interaction with the ions. We first performed a standard precision (SP)⁴⁴ ensemble docking using these three receptor models to identify the unique compounds that ranked within the top 1500 entries in all the three independent docking runs. This resulted in the selection of 80 molecules that were subjected to a second ensemble docking performed with the extra-precision (XP) scoring function,⁴⁵ aiming to improve the protein-ligand binding predictions. Thus, such 80 compounds were ranked via their mean score (i.e., averaging the score of each compounds over the three docking runs), and after visual inspection twenty molecules were selected to be tested as Pol- η inhibitors.

The similarity screen was performed with Schrodinger's Canvas.^{46,47} The radial fingerprint was computed for all the compounds of the D3 library, including for ARN18347. This last compound was used as a filter to screen the D3 library by using the Tanimoto similarity score.

Molecular dynamics simulation. Details of the system preparation (including the force fields used) and simulation procedure are described in our previous papers on Pol η .⁴⁸⁻⁵¹ The general AMBER force field (GAFF) was used to describe the bonded and nonbonded interactions of **ARN17212 and compound 21**.⁵² To derive charges for **the ligands**, geometry optimization was performed at the MP2/6-31G* level, followed by electrostatic potential calculation at the HF/6-31G* level using Gaussian 09.⁵³ Subsequently, a restrained electrostatic potential (RESP) fitting was performed using antechamber.⁵⁴ Production MD simulations were performed in the NPT ensemble for 500 ns using GROMACS 2018.8.^{55,56} Geometric parameters and H-bond interactions were analyzed using CPPTRAJ.⁵⁷

Expression and purification. The DNA sequence coding the catalytic core of human Pol η (1-511) was codon optimized for bacterial expression and cloned into a pET28 vector (Genscript, USA) that includes a C-terminal 6x His tag. The translesion synthesis activity of the catalytic core has been shown to equal that of the full-length polymerase.⁵⁸ The protein was overexpressed in *E. Coli* BL21(DE3) cells and purified by chromatography using the ÄKTApurifier system (GE Healthcare Life Sciences) with an anion exchange column (ResourceQ) followed by a HisTrap column. Briefly, 500 ml of cell paste was lysed in 40 ml of 50 mM Tris-HCl pH 7.3, 0.1 M NaCl, 10 % Glycerol (v/v), 5 mM β -mercaptoethanol and 20 mM imidazole containing 0.2 mg/ml lysozyme and a protease inhibitor cocktail (EDTA free; SIGMA). The lysate was further sonicated and then centrifuged at 35000xg, for 30 minutes, at 4°C. The supernatant was filtered through a

0.45 µm filter and then loaded into the ResourceQ column. The flow-through obtained was recovered and directly loaded into a HisTrap column. Finally, the protein was eluted with a gradient of imidazole (60 mM to 500 mM). The fractions containing the eluted protein were combined and stored at -80°C. Protein concentration was determined using the Bradford assay (Bio-Rad) and the purity was assessed by SDS-page and coomassie staining.

Measurement of activity. Activity of Pol η was assessed by end-point measurement of the DNA product synthesized by the enzyme in the presence of all four dNTPs. The DNA template used in the reaction⁵⁹ was created by annealing an IRD700-labelled primer (5'-IRD700-GCAGGTCGACTCCAAAG-3') to a template strand (5'-TCGGTACCGGGTTAGCCTTTGGAGTCGACCTGC-3') as described.⁶⁰ A mixture of DNA template (100 nM final concentration) and Pol η (2 nM final concentration) was set up on ice in assay buffer (50 mM HEPES pH 7.5, 50 mM NaCl, 10 % glycerol, 250 nM each dNTP, 1 mM DTT and 0.05 % Tween 20) and then added into the tubes containing the inhibitors or DMSO (as vehicle; 1 % final assay concentration). The reaction was started by the addition of 5 mM Mg²⁺ (final concentration) and immediately transferred to 37°C. After 60 minutes incubation, the reaction was stopped by the addition of 5X sample loading buffer (90 mM Tris pH 8.3, 90 mM Borate, 2 mM EDTA, 12 % ficoll 400, 7 M Urea and 0.01 % Bromophenol Blue) and incubation for 5 minutes at 70 °C. Pol η extension products were then separated by denaturing electrophoresis (15 % polyacrylamide/TBE/Urea gels; BioRad) and scanned using the ChemiDoc Imaging System (BioRad Laboratories). Scanned images were analyzed using Image Lab software (Version 6.0.0; BioRad Laboratories).

Human Cell Culture. Human cancer cell lines A549 (lung adenocarcinoma, ATCC CCL-185), A375 (malignant melanoma, ATCC CRL-1619), OVCAR3 (ovary adenocarcinoma, ATCC HTB-161) and HEK-293 (epithelial, ATCC CRL-1573) were obtained from ATCC. A549 cells were routinely grown in F-12K medium, supplemented with 2 mM L-glutamine, 10% heat-inactivated FBS and 1% penicillin/streptomycin. A375 and HEK-293 cells were grown in Dulbecco's Modified Eagle Medium (DMEM), supplemented with 2 mM L-glutamine, 10% heat-inactivated FBS and 1% Penicillin/streptomycin. Finally, OVCAR3 cells were grown in RPMI medium, supplemented with 20% heat-inactivated FBS and 2 mM L-glutamine, 0,01 mg/ml insulin solution from bovine pancreas (Merk lifescience), 10 mM HEPES, 1 mM sodium pyruvate and 2,5 gr/L D-

(+)-Glucose solution. All cell lines were grown in a humidified atmosphere of 5% CO₂, at 37 °C. To assess the antiproliferative activity of the compounds, cells were seeded at a density of 3000 cells/well (A375), 5000 cells/well (A549) and 10000 cells/well (OVCAR3, HEK-293) in 96-well plates, and cell viability was measured using the MTT assay. Cells were also co-treated with compound 17, or 21, and cisplatin to determine if the combination treatment sensitizes the cell to the cisplatin drug. Briefly, cells were seeded and after 24 h cisplatin (0-50 µM) or 17 (0-200 µM) or 21 (0-200 µM) was added to each well. The cells were incubated for 48 h. In the co-treatment experiments, cells were treated with cisplatin (0-50 µM) and 10, 20, 30 µM of compound 17; or 35, 50, 100 µM of compound 21.

MTT cell viability assay. Cell viability was measured using the MTT assay. Cells were seeded in 96 well plates. Twenty four hours after seeding, the cells were treated with compounds or vehicle (DMSO, final concentration 0.5%) as control and incubated for 48 hours. Then, MTT solution (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) was added to a final concentration of 0.5 mg/ml and cells were further incubated for 2 hours at 37 °C. After solubilization of the formazan crystals by the addition of a 10% SDS/0.01 N HCl solution, absorbance was measured at 570 nm (reference 690 nm) in a plate reader (Infinite M200, Tecan). Inhibition curves consisted of 9 serial dilutions in triplicate in each case, and results were analyzed as sigmoidal dose-response curves using GraphPad Prism software (version 5.03). Values are reported as the mean±SD of three independent experiments.

Combination index calculation. According to Chou-Talalay method and as already reported in literature, we calculated the combination index (C.I.) of the co-treatment using a fixed concentration of cisplatin and fixed concentration of compound 21 or 17. We calculated the values using the formula reported below:

$$\frac{[\text{Surviving cells treated with the combination}]}{[(\text{Surviving cells treated with Cisplatin}) \times (\text{Surviving cells treated with 17/21})]}$$

Therefore, a ratio ranging from 0.8 to 1.2 denotes an additive effect. Synergism is indicated by a ratio <0.8; antagonism by a ratio >1.2. We reported these results (C.I. values) in the Table 6.

Western blot analysis. A375 and OVCAR3 cells were seeded in 6 well plates (1.5 x 10⁵ and 3 x 10⁵ cells/well respectively) and treated with cisplatin (1 µM and 5 µM for A375; 0.5 µM and 1

μM for OVCAR3) and **17** or **21** ($30 \mu\text{M}$) for 48 h. Then, proteins were extracted by detaching cells with trypsin and after 5 min of centrifugation at 800 rpm, cells pellet were resuspended with Radio Immunoprecipitation Assay Buffer (RIPA) and protein concentration quantified with BCA assay (Euroclone,). Western blot analyses were performed on $20 \mu\text{g}$ of protein extracts, added with 4X of Loading Buffer (0.25 M Tris-HCl pH 6.8, 8% SDS, 0.3 M DTT, 30% Glycerol, 0.4% Bromophenol Blue) and denatured at $95 \text{ }^\circ\text{C}$, for 5 min.

Samples run on precast polyacrylamide gels (NuPAGE 4–12% BisTris Gel, Invitrogen) at 120 V and then transferred onto $0.45 \mu\text{m}$ nitrocellulose membrane (BioRad) for 2 h at 30 V. After blocking in 5% non-fat milk in TBS-Tween 20 0,1% (TBST), membranes were probed overnight at $4 \text{ }^\circ\text{C}$ with anti-GAPDH (1:1000, Invitrogen), and anti-phosphorylated-histone γH2AX (1:1000, Cell signaling).

Next day, membranes were incubated with horseradish peroxidase conjugated secondary antibody (Abcam) at room temperature for 1 h. Proteins were visualized by the chemiluminescent substrate, ECL Star (Euroclone), using the ChemiDoc (BioRad) Chemiluminescence. Densitometry analysis was carried out using ImageJ software.

Immunofluorescence. Cells were seeded (6000 or 10000 cells/well) on Millicell ez slide 8-well (Millipore). Control and treated (with cisplatin $0.5 \mu\text{M}$ or $1 \mu\text{M}$ and with or without **21** at $20 \mu\text{M}$, for 24 h) cells were fixed in formaldehyde 3.7% in PBS for 10 min, and washed twice in PBS. Then triton 0.5% was added for 10 min and wash with PBS for three times. The samples were incubated in 3% Bovine Serum Albumin (BSA) (Sigma- Aldrich, USA) in PBS for 30 min at R.T. and subsequently in the primary antibody, rabbit antibody anti γH2AX (Cell signaling, 1:500 in 1% BSA in PBS), overnight at $4 \text{ }^\circ\text{C}$. After washing, the samples were incubated in anti-rabbit Alexa fluor 488 (Invitrogen) conjugated secondary antibody (1:1000 1% BSA in PBS) for 1 h at R.T., washed with PBS, and added a solution 1:1000 of DAPI (Hoechst 33342, Invitrogen) for 1 min. Finally, after three wash with PBS, the coverslip was mounted. The samples were analyzed with a Leica600 fluorescent microscope.

2.5 References

1. Helleday, T.; Petermann, E.; Lundin, C.; Hodgson, B.; Sharma, R. A., DNA repair pathways as targets for cancer therapy. *Nat Rev Cancer* **2008**, *8* (3), 193-204.
2. Swift, L. H.; Golsteyn, R. M., Genotoxic anti-cancer agents and their relationship to DNA damage, mitosis, and checkpoint adaptation in proliferating cancer cells. *Int J Mol Sci* **2014**, *15* (3), 3403-31.
3. Fu, D.; Calvo, J. A.; Samson, L. D., Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat Rev Cancer* **2012**, *12* (2), 104-20.
4. Rebutti, M.; Michiels, C., Molecular aspects of cancer cell resistance to chemotherapy. *Biochem Pharmacol* **2013**, *85* (9), 1219-26.
5. Vasan, N.; Baselga, J.; Hyman, D. M., A view on drug resistance in cancer. *Nature* **2019**, *575* (7782), 299-309.
6. Chatterjee, N.; Bivona, T. G., Polytherapy and Targeted Cancer Drug Resistance. *Trends Cancer* **2019**, *5* (3), 170-182.
7. Nikolaou, M.; Pavlopoulou, A.; Georgakilas, A. G.; Kyrodimos, E., The challenge of drug resistance in cancer treatment: a current overview. *Clin Exp Metastasis* **2018**, *35* (4), 309-318.
8. Zafar, M. K.; Eoff, R. L., Translesion DNA Synthesis in Cancer: Molecular Mechanisms and Therapeutic Opportunities. *Chem Res Toxicol* **2017**, *30* (11), 1942-1955.
9. Yang, W., An overview of Y-Family DNA polymerases and a case study of human DNA polymerase eta. *Biochemistry* **2014**, *53* (17), 2793-803.
10. Pata, J. D., Structural diversity of the Y-family DNA polymerases. *Biochim Biophys Acta* **2010**, *1804* (5), 1124-35.
11. Yang, W.; Woodgate, R., What a difference a decade makes: insights into translesion DNA synthesis. *Proc Natl Acad Sci U S A* **2007**, *104* (40), 15591-8.
12. Yang, W.; Gao, Y., Translesion and Repair DNA Polymerases: Diverse Structure and Mechanism. *Annu Rev Biochem* **2018**, *87*, 239-261.
13. Wojtaszek, J. L.; Chatterjee, N.; Najeeb, J.; Ramos, A.; Lee, M.; Bian, K.; Xue, J. Y.; Fenton, B. A.; Park, H.; Li, D.; Hemann, M. T.; Hong, J.; Walker, G. C.; Zhou, P., A Small Molecule Targeting Mutagenic Translesion Synthesis Improves Chemotherapy. *Cell* **2019**, *178* (1), 152-159 e11.

14. Patel, S. M.; Dash, R. C.; Hadden, M. K., Translesion synthesis inhibitors as a new class of cancer chemotherapeutics. *Expert Opin Investig Drugs* **2021**, *30* (1), 13-24.
15. Yamanaka, K.; Chatterjee, N.; Hemann, M. T.; Walker, G. C., Inhibition of mutagenic translesion synthesis: A possible strategy for improving chemotherapy? *PLoS Genet* **2017**, *13* (8), e1006842.
16. Sail, V.; Rizzo, A. A.; Chatterjee, N.; Dash, R. C.; Ozen, Z.; Walker, G. C.; Korzhnev, D. M.; Hadden, M. K., Identification of Small Molecule Translesion Synthesis Inhibitors That Target the Rev1-CT/RIR Protein-Protein Interaction. *ACS Chem Biol* **2017**, *12* (7), 1903-1912.
17. Korzhnev, D. M.; Hadden, M. K., Targeting the Translesion Synthesis Pathway for the Development of Anti-Cancer Chemotherapeutics. *J Med Chem* **2016**, *59* (20), 9321-9336.
18. Masutani, C.; Kusumoto, R.; Yamada, A.; Dohmae, N.; Yokoi, M.; Yuasa, M.; Araki, M.; Iwai, S.; Takio, K.; Hanaoka, F., The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta. *Nature* **1999**, *399* (6737), 700-4.
19. Guo, C.; Kosarek-Stancel, J. N.; Tang, T. S.; Friedberg, E. C., Y-family DNA polymerases in mammalian cells. *Cell Mol Life Sci* **2009**, *66* (14), 2363-81.
20. Srivastava, A. K.; Han, C.; Zhao, R.; Cui, T.; Dai, Y.; Mao, C.; Zhao, W.; Zhang, X.; Yu, J.; Wang, Q. E., Enhanced expression of DNA polymerase eta contributes to cisplatin resistance of ovarian cancer stem cells. *Proc Natl Acad Sci U S A* **2015**, *112* (14), 4411-6.
21. Albertella, M. R.; Green, C. M.; Lehmann, A. R.; O'Connor, M. J., A role for polymerase eta in the cellular tolerance to cisplatin-induced damage. *Cancer Res* **2005**, *65* (21), 9799-806.
22. Alt, A.; Lammens, K.; Chiochini, C.; Lammens, A.; Pieck, J. C.; Kuch, D.; Hopfner, K. P.; Carell, T., Bypass of DNA lesions generated during anticancer treatment with cisplatin by DNA polymerase eta. *Science* **2007**, *318* (5852), 967-70.
23. Ceppi, P.; Novello, S.; Cambieri, A.; Longo, M.; Monica, V.; Lo Iacono, M.; Gaj-Levra, M.; Saviozzi, S.; Volante, M.; Papotti, M.; Scagliotti, G., Polymerase eta mRNA expression predicts survival of non-small cell lung cancer patients treated with platinum-based chemotherapy. *Clin Cancer Res* **2009**, *15* (3), 1039-45.
24. Teng, K. Y.; Qiu, M. Z.; Li, Z. H.; Luo, H. Y.; Zeng, Z. L.; Luo, R. Z.; Zhang, H. Z.; Wang, Z. Q.; Li, Y. H.; Xu, R. H., DNA polymerase eta protein expression predicts treatment response and survival of metastatic gastric adenocarcinoma patients treated with oxaliplatin-based chemotherapy. *J Transl Med* **2010**, *8*, 126.

25. Zhou, W.; Chen, Y. W.; Liu, X.; Chu, P.; Loria, S.; Wang, Y.; Yen, Y.; Chou, K. M., Expression of DNA translesion synthesis polymerase eta in head and neck squamous cell cancer predicts resistance to gemcitabine and cisplatin-based chemotherapy. *PLoS One* **2013**, *8* (12), e83978.
26. Saha, P.; Mandal, T.; Talukdar, A. D.; Kumar, D.; Kumar, S.; Tripathi, P. P.; Wang, Q. E.; Srivastava, A. K., DNA polymerase eta: A potential pharmacological target for cancer therapy. *J Cell Physiol* **2021**, *236* (6), 4106-4120.
27. Dorjsuren, D.; Wilson, D. M., 3rd; Beard, W. A.; McDonald, J. P.; Austin, C. P.; Woodgate, R.; Wilson, S. H.; Simeonov, A., A real-time fluorescence method for enzymatic characterization of specialized human DNA polymerases. *Nucleic Acids Res* **2009**, *37* (19), e128.
28. Givens, J. F.; Manly, K. F., Inhibition of RNA-directed DNA polymerase by aurintricarboxylic acid. *Nucleic Acids Res* **1976**, *3* (2), 405-18.
29. Benchokroun, Y.; Couprie, J.; Larsen, A. K., Aurintricarboxylic acid, a putative inhibitor of apoptosis, is a potent inhibitor of DNA topoisomerase II in vitro and in Chinese hamster fibrosarcoma cells. *Biochemical Pharmacology* **1995**, *49* (3), 305-313.
30. Klein, P.; Cirioni, O.; Giacometti, A.; Scalise, G., In vitro and in vivo activity of aurintricarboxylic acid preparations against *Cryptosporidium parvum*. *J Antimicrob Chemother* **2008**, *62* (5), 1101-4.
31. Zafar, M. K.; Maddukuri, L.; Ketkar, A.; Penthala, N. R.; Reed, M. R.; Eddy, S.; Crooks, P. A.; Eoff, R. L., A Small-Molecule Inhibitor of Human DNA Polymerase eta Potentiates the Effects of Cisplatin in Tumor Cells. *Biochemistry* **2018**, *57* (7), 1262-1273.
32. Coggins, G. E.; Maddukuri, L.; Penthala, N. R.; Hartman, J. H.; Eddy, S.; Ketkar, A.; Crooks, P. A.; Eoff, R. L., N-Aroyl indole thiobarbituric acids as inhibitors of DNA repair and replication stress response polymerases. *ACS Chem Biol* **2013**, *8* (8), 1722-9.
33. Wilson, D. M.; Duncton, M. A. J.; Chang, C.; Lee Luo, C.; Georgiadis, T. M.; Pellicena, P.; Deacon, A. M.; Gao, Y.; Das, D., Early Drug Discovery and Development of Novel Cancer Therapeutics Targeting DNA Polymerase Eta (POLH). *Front Oncol* **2021**, *11*, 778925.
34. Genna, V.; Vidossich, P.; Ippoliti, E.; Carloni, P.; De Vivo, M., A Self-Activated Mechanism for Nucleic Acid Polymerization Catalyzed by DNA/RNA Polymerases. *J Am Chem Soc* **2016**, *138* (44), 14592-14598.

35. Nakamura, T.; Zhao, Y.; Yamagata, Y.; Hua, Y. J.; Yang, W., Watching DNA polymerase eta make a phosphodiester bond. *Nature* **2012**, *487* (7406), 196-201.
36. Lopez-Lazaro, M., Distribution and biological activities of the flavonoid luteolin. *Mini Rev Med Chem* **2009**, *9* (1), 31-59.
37. Lin, Y.; Shi, R.; Wang, X.; Shen, H.-M., Luteolin, a Flavonoid with Potential for Cancer Prevention and Therapy. *Current Cancer Drug Targets* **2008**, *8* (7), 634-646.
38. Panche, A. N.; Diwan, A. D.; Chandra, S. R., Flavonoids: an overview. *J Nutr Sci* **2016**, *5*, e47.
39. Dos Santos Ferreira, A. C.; Fernandes, R. A.; Kwee, J. K.; Klumb, C. E., Histone deacetylase inhibitor potentiates chemotherapy-induced apoptosis through Bim upregulation in Burkitt's lymphoma cells. *J Cancer Res Clin Oncol* **2012**, *138* (2), 317-25.
40. Quagliariello, V.; Gennari, A.; Jain, S. A.; Rosso, F.; Iaffaioli, R. V.; Barbarisi, A.; Barbarisi, M.; Tirelli, N., Double-responsive hyaluronic acid-based prodrugs for efficient tumour targeting. *Mater Sci Eng C Mater Biol Appl* **2021**, *131*, 112475.
41. Zhang, J. M., L.; Cui, W., Yang, J.; and Yang, B., Total synthesis of luteolin. *J. Chem. Sci.* **2014**, *38*, 60-61.
42. Schrödinger Release 2015-4: LigPrep; Schrödinger, LLC: New York, 2015.
43. Greenwood, J. R.; Calkins, D.; Sullivan, A. P.; Shelley, J. C., Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *J Comput Aided Mol Des* **2010**, *24* (6-7), 591-604.
44. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S., Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* **2004**, *47* (7), 1739-49.
45. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T., Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J Med Chem* **2006**, *49* (21), 6177-96.
46. Duan, J.; Dixon, S. L.; Lowrie, J. F.; Sherman, W., Analysis and comparison of 2D fingerprints: insights into database screening performance using eight fingerprint methods. *J Mol Graph Model* **2010**, *29* (2), 157-70.

47. Sastry, M.; Lowrie, J. F.; Dixon, S. L.; Sherman, W., Large-scale systematic analysis of 2D fingerprint methods and parameters to improve virtual screening enrichments. *J Chem Inf Model* **2010**, *50* (5), 771-84.
48. <https://academic.oup.com/nar/article/44/6/2827/2499467>.
49. Genna, V.; Gaspari, R.; Dal Peraro, M.; De Vivo, M., Cooperative motion of a key positively charged residue and metal ions for DNA replication catalyzed by human DNA Polymerase- ϵ . *Nucleic Acids Res* **2016**, *44* (6), 2827-36.
50. Genna, V.; Colombo, M.; De Vivo, M.; Marcia, M., Second-Shell Basic Residues Expand the Two-Metal-Ion Architecture of DNA and RNA Processing Enzymes. *Structure* **2018**, *26* (1), 40-50 e2.
51. Genna, V.; Carloni, P.; De Vivo, M., A Strategically Located Arg/Lys Residue Promotes Correct Base Paring During Nucleic Acid Biosynthesis in Polymerases. *J Am Chem Soc* **2018**, *140* (9), 3312-3321.
52. Wang, J., Wolf, R. M., Caldwell J. W., Kollman, P. A., Case, D. A., Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
53. <https://gaussian.com/g09citation/>.
54. Bayly, C. I., Cieplak, P., Cornell, W., and Peter A. Kollman, A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97* (40), 10269–10280.
55. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* **2006**, *25* (2), 247-60.
56. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19-25.
57. Roe, D. R.; Cheatham, T. E., 3rd, PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput* **2013**, *9* (7), 3084-95.
58. Biertumpfel, C.; Zhao, Y.; Kondo, Y.; Ramon-Maiques, S.; Gregory, M.; Lee, J. Y.; Masutani, C.; Lehmann, A. R.; Hanaoka, F.; Yang, W., Structure and mechanism of human DNA polymerase ϵ . *Nature* **2010**, *465* (7301), 1044-8.

59. Beardslee, R. A.; Suarez, S. C.; Toffton, S. M.; McCulloch, S. D., Mutation of the little finger domain in human DNA polymerase eta alters fidelity when copying undamaged DNA. *Environ Mol Mutagen* **2013**, *54* (8), 638-51.
60. Ketkar, A.; Maddukuri, L.; Penthala, N. R.; Reed, M. R.; Zafar, M. K.; Crooks, P. A.; Eoff, R. L., Inhibition of Human DNA Polymerases Eta and Kappa by Indole-Derived Molecules Occurs through Distinct Mechanisms. *ACS Chem Biol* **2019**, *14* (6), 1337-1351.

Chapter 3: Structural studies

During my PhD, I had the opportunity to spend six months abroad. The hosting lab, led by Dr. Marquez at EMBL in Grenoble, has a strong expertise in characterizing challenging macromolecular assemblies via X-rays crystallography. Our aim was to investigate the binding pose of our compounds into the target Pol η /DNA complex. We looked for the best conditions for crystallizing the ternary complex of Pol η /dsDNA with different inhibitors that we have previously identified (Chapter 2). The crystal structure of such complexes would indeed significantly contribute to guiding our structural-based drug design efforts towards more potent Pol η inhibitors.

3.1 The Experimental Approach and Methods

3.1.1. Protein expression. DNA Polymerase η expression started with constructs optimization. The catalytic domain only (aa 1-432) was chosen for the X-rays studies. A longer construct was used for the biochemical assays (aa 1-511). For Pol η (1-511) the protocol was already described in Chapter 2. For the expression of Pol η , one single colony of transformed *E. Coli* BL21(DE3) was picked from the plate and added to a starter culture of 25 mL of LB buffer supplemented with kanamycin 100 $\mu\text{g}/\mu\text{L}$. Bacteria were grown overnight at 37 °C in a shaking incubator (132 rpm). The next day, 10 mL of the saturated overnight culture were diluted into a 1 L of LB culture. The bacterial culture was grown at 37 °C with shaking at 150 rpm for 1 h and 30 minutes. Then the temperature was reduced to 16 °C and further incubated for 1 h- 1 h and 30 min. Once the late-exponential growth phase was reached, when the optical density at 600 nm (OD_{600}) = 0.8 and the culture reached the 16 °C, it was induced using Isopropil- β -D-1-thiogalattopiranoside (IPTG) 0.5 mM. After induction, the bacteria were grown at 16 °C for another night. The expression protocol has been set up testing different growth times, temperatures and IPTG concentrations to increase the yield of protein expression. Bacterial cells were harvested by centrifugation at 6000 xg for 45 minutes at 4 °C. Pellets for both Pol η (1-432) and (1-511) were stored at -80 °C.

3.1.2. Protein purification. Cells pellet was resuspended in a 45 mL volume of lysis buffer (50 mM TRIS-HCl pH 7.3, 0.3 M NaCl, 10% glycerol, 1 mM DTT, Protease Inhibitors EDTA-free 0.5x (Roche)), then lysozyme was added (0.2 mg/mL). Cell suspension was then lysed on ice by sonication (5 minutes with rounds of 30 seconds each), 70-75% amplitude, KE 76 Tip, Bandelin Sonoplus HD2070 sonicator). Then, DNase I was added at 100 $\mu\text{g/mL}$ with Mg^{2+} 10 mM for 45 minutes at 4 °C with shaking. After a centrifugation of 30 minutes at 35000 xg, the supernatant was collected, filtered with a 0.45 μm MiniSart Syringe Filter and loaded on a 5 mL HisTrap HP chromatography column (Cytiva) equilibrated with Buffer A (50 mM TRIS-HCl pH 7.3, 0.3 M NaCl, 10 % glycerol (v/v), 1 mM DTT). One washing step was included to remove the excess of DNA, it was performed with 100 % buffer S2 (50 mM TRIS-HCl pH 7.3, 1.5 M NaCl, 10 % glycerol (v/v), 1 mM DTT). The elution was performed with a gradient of buffer B (50 mM TRIS-HCl pH 7.3, 0.5 M NaCl, 10 % glycerol (v/v), 1 mM DTT and 0.5 M Imidazole). Fractions eluted approximately at 40 % buffer B and they were pooled together. The 6xHis-tag was cleaved using PreScission Protease (Sigma, GE27-0843-01) in an overnight reaction at 4 °C. The sample was filtered again and loaded into a 5 mL HiTrap Heparin chromatography column (Cytiva) equilibrated with buffer A1 (50 mM Tris-HCl pH 7.3, 150 mM NaCl, 10 % glycerol (v/v), 1 mM DTT). The elution was obtained with buffer B1 (50 mM Tris-HCl pH 7.3, 1.5 M NaCl, 10 % glycerol (v/v), 1 mM DTT). using a linear gradient up to 100 % buffer B1 in 5 CVs. Eluted protein fractions were collected, pooled together, and quantified by Nanodrop (extinction coefficient of the protein $\epsilon_{280} = 50420$). Purity and degradation were checked through SDS-PAGE gels for each steps. In conclusion, the pool were loaded into size exclusion Superdex 75 pg Increase chromatography 10/300 GL column (Cytiva). Protein storage buffer was 50 mM Hepes pH 7.5, 0.4 M NaCl, 10 % glycerol (v/v), 3 mM DTT. The purified protein was aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

3.1.3. SDS-PAGE was performed using precast polyacrylamide gels (NuPAGE 4-12 % BisTris Gel, Invitrogen). Different concentration of protein samples were mixed with 4X Loading buffer (0.25 M Tris-HCl pH 6.8, 8 % SDS (v/v), 0.3 M DTT, 30 % Glycerol (v/v), 0.4 % Bromphenol Blue (w/v)) prior to denaturation at 95 °C for 5 minutes. After samples loading, precast gels were run in XCell SureLock Mini-Cell Electrophoresis System

(Invitrogen) in MOPS 1 X SDS running buffer (Invitrogen) with a constant voltage of 150 mA for about 90 minutes. Gel were then stained with Coomassie Blue staining buffer (40 % EtOH (v/v), 10 % Acetic Acid (v/v), 0.05% coomassie blue G-250 (w/v)) for 15 – 30 minutes and destained with a destaining Buffer (8 % acetic acid (v/v), 25 % EtOH (v/v)). Protein bands images were visualized and quantified using ChemiDoc Imaging System (BioRad).

3.1.4. Western Blot. SDS-PAGE gels were run as described in the previous section, without Coomassie Staining. Gels were blotted onto a 0.22 µm nitrocellulose membrane (Protran BA83, GE Healthcare) applying a constant voltage of 30 V for 2 hours in NuPage MOPS transfer buffer (Invitrogen) with 10 % methanol (v/v). Protein transfer to the membrane was checked through a fast incubation in a Red Ponceau Solution. The membrane was then incubated for 1 hour in blocking solution (5 % no-fat milk-TBS-T (TBS 1 X (v/v) and 0.1 % Tween20 (v/v)) at room temperature with agitation. The following incubation with primary antibody (anti-HisTag, Thermofisher) diluted 1:1000 in 5 % no-fat milk-TBST (v/v) was performed overnight at 4 °C with agitation. The following day, after three washes (5-10 minutes each) with TBS-T, an incubation with a HRP-conjugated anti-rabbit secondary antibody (Abcam) diluted 1:20000 in 5 % no-fat milk-TBS-T (v/v) was performed for 1 hour at room temperature with agitation. After three additional washes, the membrane was finally incubated with Amersham ECL Star substrate (Euroclone) and exposed to ChemiDoc chemoluminescence for detection (BioRad).

3.1.5. Native Gel Electrophoresis. Native Gel electrophoresis technique was employed using precast native gels (NativePAGE 3-12 % BisTris Gel, Invitrogen). Protein samples were mixed with 4 X Native loading buffer (Invitrogen). After samples loading, precast gels were run in XCell SureLock Mini-Cell Electrophoresis System (Invitrogen) using Cationic and anionic running buffers (Invitrogen), with a constant voltage of 150 mA for about 90 minutes. Gel were then stained with Coomassie Blue staining buffer (40 % EtOH (v/v), 10 % Acetic Acid (v/v), 0.05 % coomassie blue G-250 (w/v)) for 15 – 30 minutes and destained with Destaining Buffer (8 % acetic acid (v/v), 25 % EtOH (v/v)). Protein bands images were visualized and quantified using ChemiDoc Imaging System (BioRad).

3.1.6 Crystallization. The purified protein was used for the crystallization set up. After verifying the purity of the sample via SDS-page, the protein (~ 2 mg/mL) was incubated with the dsDNA for 10 minutes at room temperature with a ratio of 1:1.05 respectively. The sequence of the DNA chosen was undamaged (template: 3' TCGCAGTATTACT 5', primer: 5' TAGCGTCAT 3') and were annealed 1:1 molar ratio in a buffer 10 mM HEPES pH 8.0, 0.1 mM EDTA and 50 mM NaCl by heating for 5 min at 85°C and slow cooling to 25 °C. Then, the MgCl₂ 5 mM was added in the mixture and the sample was ready to be concentrated up to 10 mg/mL using Vivaspin® (500, 10,000 MWCO, PES). Non-hydrolyzable nucleotide was added at last to form ternary complexes. First crystallization attempts were carried out using different kits and an automatic protein crystallization system (crystallization robot), which was also used to screen the optimal crystallization conditions, together with manual methods such as hanging and sitting drop methods. At the same time, following the protocols present in the literature¹⁰, we tried to obtain the protein crystal with dsDNA and with or without the non-hydrolyzable dNTP.

The co-crystallization experiments were performed by incubating the enzyme, in the presence of dsDNA and the normal or non-hydrolyzable nucleotide, with compound 64 or compound 12. We tried different incubations times (10-15-30 minutes) at R.T. or in ice. The two compounds were dissolved in 100 % DMSO to create a stock solution of 200 mM, for compound 64, and 100 mM, for compound 12. The compounds were incubated with the enzyme at 10 mM (compound 64) and 5 mM (compound 12). Soaking was also executed at 20 °C using a Mosquito-LCP robot (TTP Labtech): a drop of compound was mixed with the drop already present in the plate. The final concentrations in the drops were 20 mM (compound 64) or 10 mM (compound 12), adding 20 or 40 nL of compound stock solution. We also tried to dilute the compound 12 at 50 mM mixing 50 % of DMSO (v/v) and 50 % of reservoir solution (v/v) to avoid damaging the protein.

Crystallization experiments were also conducted with a Mosquito-LCP robot (TTP Labtech). The standard drops set up was 100 nL of sample mixed with 100 nL of crystallization condition. First, various commercial screens were used to find the best initial condition for the crystallization, such as Index and Natrix (from Hampton), Midas-plus (from Molecular Dynamics), Nucleix, Classics and PEG_I (from Qiagen). The following crystallization

Automated Crystallography Pipelines

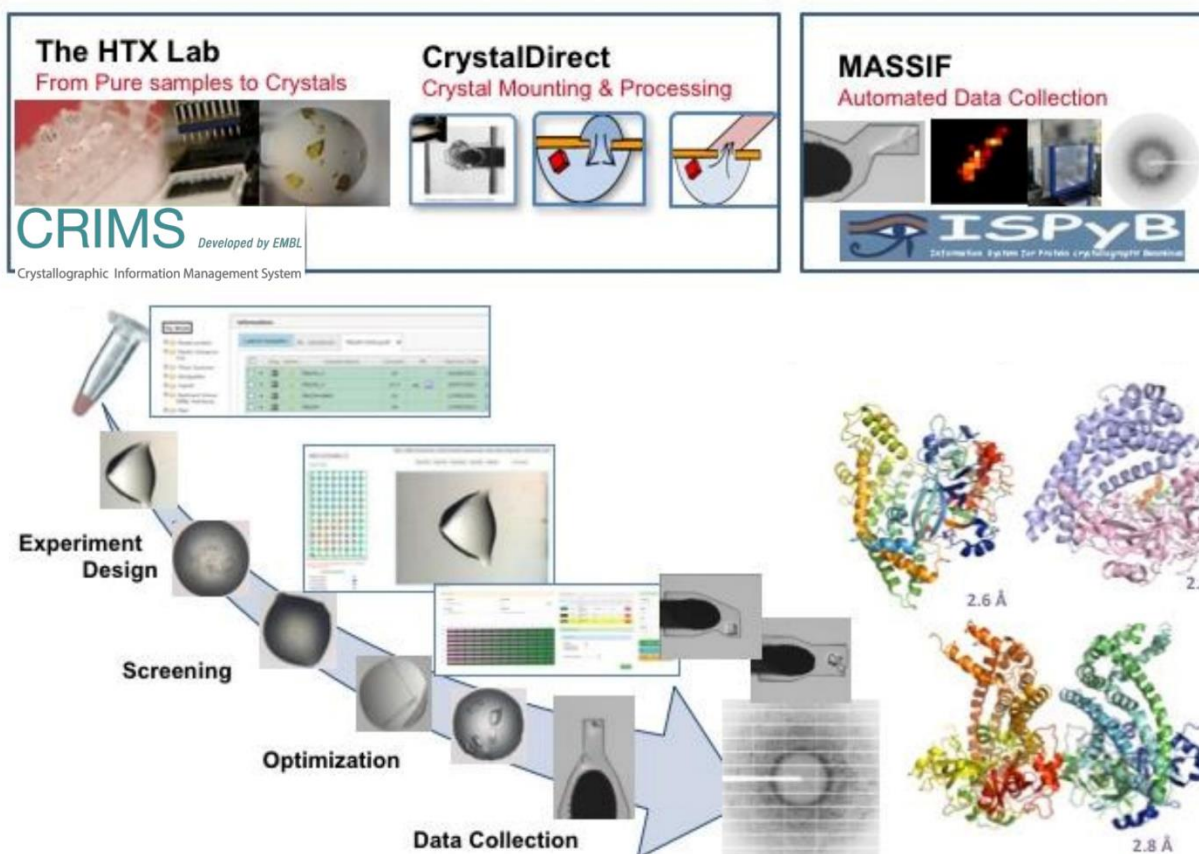


Figure 20 Automated pipeline. Here the entire process is described, which integrate the part of the EMBL HTX Lab process, the CrystalDirect technology and the MASSIF-1 beamline at the ESRF synchrotron.

conditions from literature was tried, namely 0.1 M MES pH 6; 5 mM MgCl₂; 19-21 % PEG 2K MME (v/v). These experiments were conducted using the sitting-drop vapour-diffusion method and CrystalDirect plates (MiTeGen, Ithaca, USA). The CrystalDirect approach for automated crystal harvesting and processing was used. This is a fully-automated process performed at the (HTX Lab) (**Figure 20**).

Both CrystalDirect technology and Mosquito also allows for automated crystal soaking, characterizing the interaction at the small molecule-target complex. Briefly, a small opening is created with the laser on the microplate (indeed a specific microplate is used) and a drop of a solution containing the desired chemicals (i.e. our selected compound) is added on top.

Chemicals can diffuse into the crystallization solution, eventually reaching the crystals. Chemical solutions are formulated in water, DMSO, or other organic solvents. After a certain incubation time the crystals can be collected and analyzed by diffraction.

In an X-ray diffraction experiments, crystals were recovered on pins from cryo-cooled pucks, vitrified and ready for data-collection. At the ESRF we could use MASSIF-1, an automated beamline, so the pins were automatically aligned to the centre of the X-ray beam. The images were then automatically indexed by the ISPyB image processing software pipelines. They offer information on probable spacegroup, unit cell dimensions and, after evaluation of crystal orientation, mosaicity, diffraction spots shape and signal-to-noise ratio. They also suggest data-collection strategies (data collection resolution, rotation range, oscillation width and exposure time) which can be adjusted by the operator as required.

3.1.7 X-ray data processing. The images were integrated and scaled by the ESRF servers with multiple pipelines. Between these, we used autoPROC.

AutoPROC¹ is a system for automatic processing of X-ray diffraction data and it use external programs like XDS/XSCALE, CCP4, POINTLESS and AIMLESS. The typical steps during this process involve: 1) image analysis and spot search; 2) indexing; 3) initial analysis of diffraction quality and detector parameters; 4) refinement of initial unit-cell parameters,

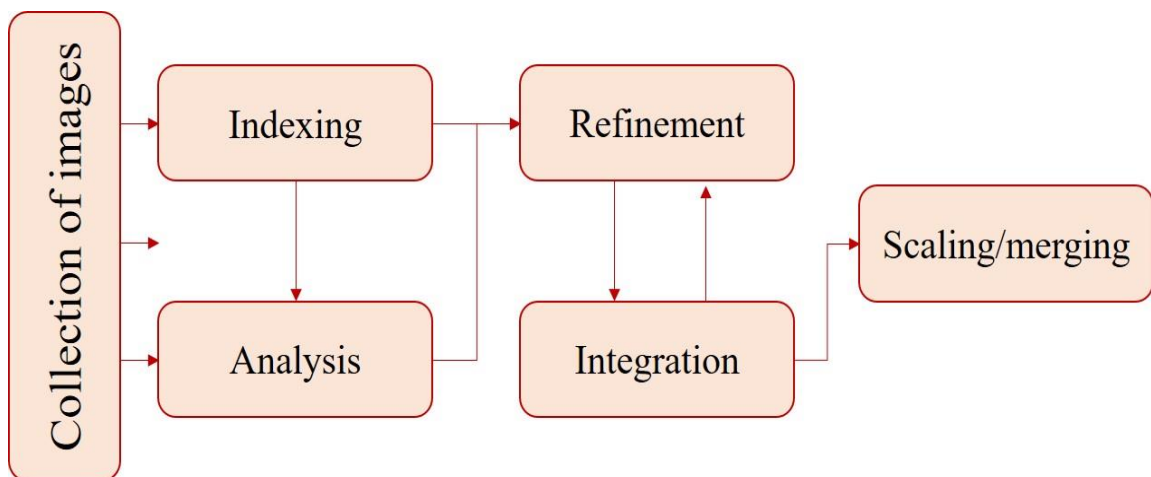


Figure 21 Steps for the initial data processing used by autoPROC.

orientation and mosaicity; 5) determination of the most likely space group; 6) integration of all images and 7) scaling and merging of integrated intensities (**Figure 21**).

Then, to solve the phase problem, molecular replacement (MR) was employed. Specifically, MR² is useful because it is possible to use a known molecular model to solve a unknown crystal structure. It helps to find a solution for the crystallographic phase problem from a previously known structure. Otherwise, it is possible to use two other methods to find the phase: 1) experimental methods (which measure the phase from isomorphous or anomalous differences) or 2) direct methods (which use mathematical relationships between reflection triplets and quartets to bootstrap a phase set for all reflections from phases for a small or random 'seed' set of reflections). The use of MR has become as expected more common, is currently used to solve up to 70% of deposited macromolecular structures, and at its best has the advantages of being fast, cheap and highly automated.

In this thesis, we decide to use Pipedream³. Pipedream is a system to link and automate the data processing with autoPROC, the MR stage with Phaser, the structure refinement with BUSTER and, if needed, automated ligand fitting with RhoFit with a subsequent BUSTER⁴ post-refinement of the top solution. The required input for Pipedream is an input data set, either in the form of unprocessed diffraction images or as a pre-processed .mtz file and an input model.

The refinement, which is a procedure consisting in minimizing the difference between the absolute values of the experimental and the calculated structure factors, was performed either manually or in an automated fashion. For this process WinCOOT⁵ and BUSTER were used. The values in bond lengths, bond angles, dihedral angles, chirality, planarity, Ramachandran phi and psi, rotamers, C-beta deviations, side-chains restraints, clashes, etc. are minimized during refinement by rounds of coordinate remodelling and calculation of an electron density map.

3.1.8 Cryo-EM. Freshly purified protein Pol η (1-432) was used to prepare and then optimize the grids. At the EMBL of Grenoble, we used two different type of grids: Quantifoil Au 300 mesh R 1.2/1.3 and UltrAUfoil 300 mesh 1.2/1.3 (Quantifoil Micro Tools). For the glow discharged we used the Pelco EasiGlow two times, 30 seconds glow at 25 mAmp at 0.3

bar. Instead, to prepare the grids we used the Vitrobot Mark IV (Thermo Fisher Scientific) and, in particular, we tried to use Pol η 1 or 2 mg/mL in the final buffer after the gel filtration, so with 10 % of glycerol (v/v). Moreover, we also tried to reduce the glycerol to 1 or 2 % (v/v). Regarding the Vitrobot, we set up it with 2 or 5 sec blotting force 0, the temperature at 4 °C and humidity 100%. The grids were plunge-frozen in liquid ethane. We added 4 μ L of sample on each side.

3.1.9 Cryo-EM data collection and analysis. Grids were clipped and screened at the Glacios Cryo-TEM equipped with a Falcon 3EC Direct Electron Detector (Thermo Fisher Scientific). The screening was operated at 200 kV. For the data collection a counting mode and a magnification of 150,000x was used, corresponding to a pixel size of 0.94 Å/pixel. Automated data acquisition was performed using EPU, with a defocus value between -0.5 to 2.5 μ m. The movies were recorded with a total dose 40 e/Å² fractionated into 40 frames and the final number of movies was 521. The image processing was performed with cryoSPARC (Structura Biotechnology)⁶. The process started with a motion correction followed by CTF estimation and curation; the particles picking had a box size of 100 Å min and 200 Å max. Then the particles were extracted with a 200 Å box size and a Fourier crop to box size 150 Å. These were used for the 2D classes, we selected the 2D classes with the best shape more similar to the enzyme and then we made an Ab initio reconstruction to generate 3 references for an Heterogeneous refinement followed by Homogeneous refinement and a resolution of 8.3 Å was reached. The template (PDB: 3MR2) were rigid-body fitted into the maps in ChimeraX⁷.

3.1.10 Thermofluor assay - thermal stability quality control. In the facility of the HTX-Lab it was possible to perform the thermofluor of the sample. The protein was diluted to a final protein concentration of 10 μ M in 40 μ L of a solution consisting of 20 mM HEPES pH 7.5, 150 mM NaCl and 5 \times SYPRO Orange (a 5000 \times SYPRO Orange preparation from Molecular Probes, Invitrogen was used as stock solution; catalogue No. S6650). The sample was then exposed to thermal denaturation in a Real Time PCR machine (Stratagene Mx3005P) with a temperature gradient from 20 to 100 °C in steps of 1 °C per minute. Protein unfolding was monitored by the increase in the fluorescence of the SYPRO Orange probe, which was measured every minute using excitation and emission wavelengths of 492 and 516 nm,

respectively. The relative fluorescence emission intensity (R) was plotted as a function of the temperature and the T_m for each individual sample was estimated as the temperature corresponding to the midpoint between the baseline and the point with maximum fluorescence intensity⁸. The same assay was also employed using the Rubic screen (Molecular dimensions), these screens were aimed at helping to determine the protein sample stability according to pH, salt concentration, buffer type and concentration. In particular, 12 μL of protein diluted in its buffer were used. In each well, of the 96-well plate, contains 21 μL of rubic buffer and 4 μL of (6.5 X) protein. For this assay the same program of the thermofluor assay was used.

3.1.11 Differential scanning fluorimetry. Three different samples were prepared: 1) enzyme was incubated with dsDNA (template: 3' TCGCAGTATTACT 5', primer: 5' TAGCGTCAT 3'), the normal nucleotide (dATP) and Ca^{2+} ; 2) the enzyme with dsDNA; 3) the enzyme with dsDNA and Mg^{2+} . All these three samples were then incubated with the compounds ARN24964 (64) 10 mM, ARN17212 (12) 10 mM, ARN25181 (81) 50 mM and with DMSO as control. Thanks to this instrument was possible to screen up to 48 conditions in parallel, using 15 μL of sample per assay with a concentration between 5 $\mu\text{g}/\text{mL}$ to 250 mg/mL (final concentration in our sample was 0.625 mg/mL). A classical set of measure lasts around 1h 30min. Samples were manually loaded into nanoDSF grade standard capillaries (12 μL) in triplicates and transferred to a Prometheus™ NT.48 nanoDSF device. For intrinsic tryptophan fluorescence measurements, the excitation wavelength of 280 nm was used, and the emission of tryptophan fluorescence was measured at 330 nm, 350 nm, and their ratios (350 nm/330 nm).

3.1.12 Isothermal titration calorimetry (ITC).

The work used the platforms of Grenoble Instruct ERIC center.



Figure 22 PEAQ-ITC. On the left, the instrument that was used thanks to the biophysical platform at the EMBL of Grenoble; on the right the schematic representation of the internal part of the instrument. The reference cell (in blue) is always filled in with deionized water; the sample cell (in pink) is filled in manually with the macromolecule (trying to avoid bubbles formation). The syringe has an automatic mechanism for inject the ligand inside.

This technique evaluates the affinity, the enthalpy and the stoichiometry of a binding reaction. It is important is that the buffer of the macromolecule and of the ligand are the same. In the case of this study, the sample for the cell was composed by Pol η (30 μM), dsDNA, Ca^{2+} (5 mM) and in certain experiments also dATP (1 mM) was included. The sample was diluted in a new buffer (50 mM Hepes pH 8, Glycerol 10 % (v/v), 2 mM DTT and 0,15 % DMSO (v/v)). The ligand (the stock is in 100 % DMSO) was diluted to final concentration of 300 μM in the buffer (50 mM Hepes pH 8, 8,5 mM KCl, 10 % glycerol (v/v) and 2 mM DTT). As ligand, we also used MgCl_2 5mM. In particular, 350 μL of sample are needed for filling in the cell and 70 μL of ligand for filling in the syringe. The experiment was set to a temperature of 20 $^\circ\text{C}$, reference power at 5 and 60 sec of delay between the 16 injections (2,5 μL for each injections and 180 sec to be back to the baseline).

3.1.13 Dynamic light scattering (DLS). Pol η (1-432) sample was analysed through DLS to test its tendency to aggregate. In particular, the test was performed in the store buffer (50 mM Hepes pH 7.5, 10 % glycerol (v/v), 0.4 M NaCl and 3 mM DTT) at protein concentrations of 0.8 mg/mL or 1 mg/mL, either immediately after size-exclusion chromatography or after different time points.

The size of the different samples was analysed using Zetasizer Nanoparticles Analyzer Software (Malvern) at 25 °C using the standard operating procedures for size measurements, repeating the measurements scans 13 times for each sample.

3.2 Results

3.2.2 Protocol optimization of Pol η expression and purification

The DNA sequence encoding for the catalytic core of human Pol η (1-432) was codon optimized for bacterial expression and cloned into a pET28 vector (Genscript, USA) that includes a C-terminal 6x His tag. The translesion synthesis activity of the catalytic core has been shown to equal that of the full-length polymerase⁹. The protein was overexpressed in *E. Coli* BL21(DE3) cells and purified by chromatography using an ÄKTApurifier system (GE Healthcare Life Sciences). As reported in the methods section, I optimized two different protocols because at the Italian Institute of Technology (Genova, Italy) all the instrumentations were kept at RT and a sonicator was used to lysate the cells. I reached a good purity of the protein (**Figure 23**) but yield and stability were not optimal, moreover, I had issues to concentrate the sample after the size exclusion. Moreover, I performed DLS to check the aggregation of the protein (data not shown) and I discovered that soon after the size exclusion it starts to aggregate. At the EMBL (Grenoble, France) I had the opportunity to optimize again the purification protocol. In particular, all the akta system were positioned in the cold room, so the temperature was in a range between 4 °- 10 °C. I could use the microfluidizer to lysate the cells. Microfluidizer technology breaks cells gently yet efficiently, resulting in large cell wall fragments. Then, the lysate (with the microfluidizer the volume increase a lot around 150 mL) was loaded into His-trap column, a Nickel column to trap the histidine tag. Through the wash with the high salt buffer, it was possible to remove a big amount of DNA (following the UV 260 nm). The protein was eluted through a gradient of imidazole; usually the protein came out in two different peaks. After this step, it is usually possible to reach 60 mg of protein from 2 L of culture. At this point, the His-tag is removed using PreScission protease in an overnight (16 hours) reaction and the cleaved sample was checked through SDS-page and western blot (**Figure 24E-F**).

Subsequently, the cleaved sample cannot be filtered, because is not stable and tends to aggregate on the membrane of filter, and a Heparin column was performed in the NGC medium-pressure chromatography system. The Hi-Trap heparin column of 5 mL has a 15 mg binding capacity, so the sample was divided and normally three cycle of loading and elution were executed. The eluted product was concentrated to reach 5 mL as final volume and through five injections everything was purified with the gel filtration column. Thanks to these different steps, the enzyme seems to be more stable and was easier to concentrate.

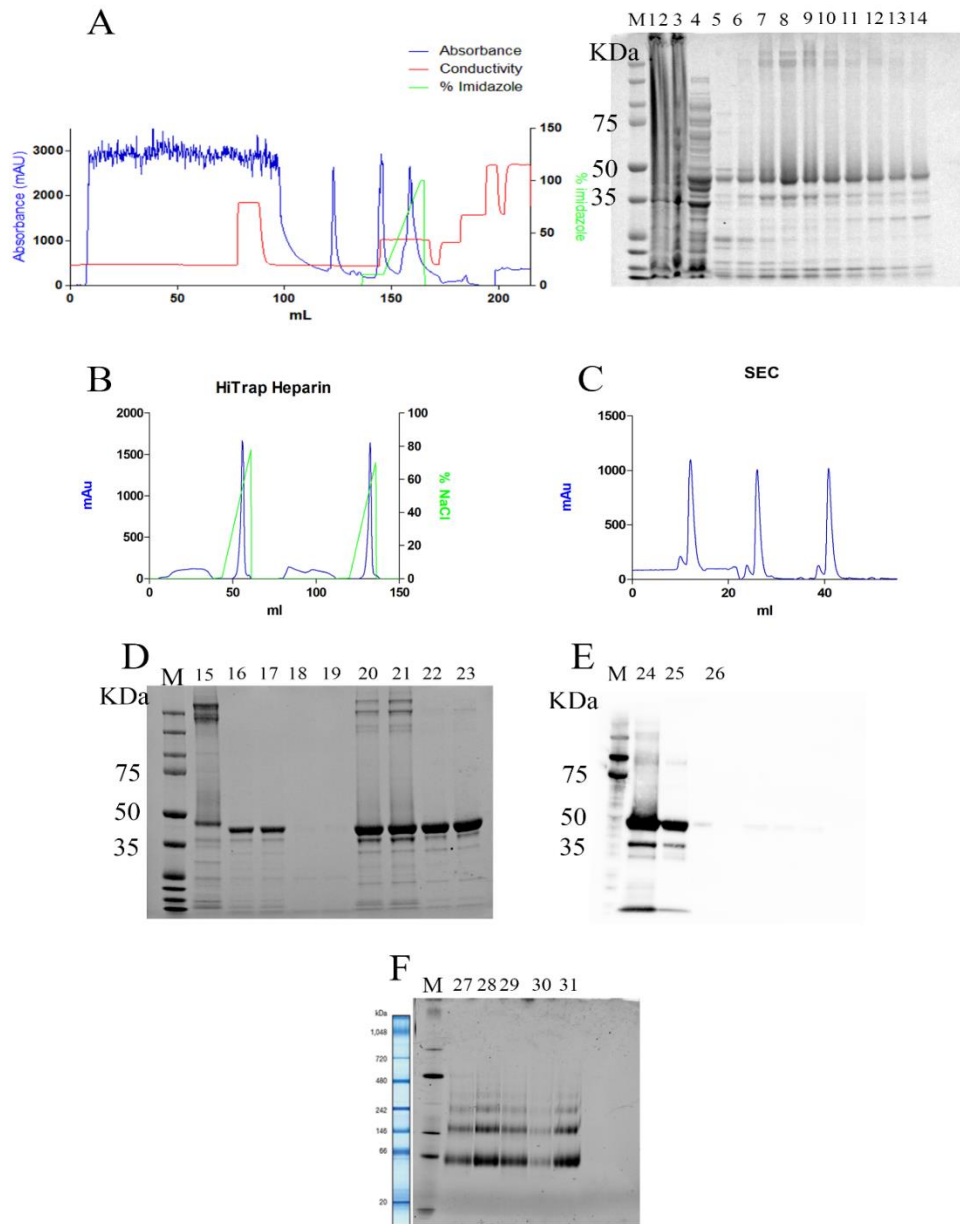


Figure 23 A) Chromatogram of His-trap column, with the loading and the elution of the protein. On the right the SDS-page with the 1) pellet of the expression, 2) the total lysate after sonication, 3) the surnatant; 4) the flow through of the purification and 5-14) the elution of the protein. B) Chromatogram of the Hi-Trap Heparin column. C) Chromatogram of the gel filtration column, the Size exclusion chromatography 75 increase column. D) SDS-page with 15) the elution from the His-trap, 16-17) the sample after the cleavage, 20-21) the elution from the heparin column and 22-23) the protein after SEC. E) the western blot with the antibody against His-tag, the signal is present just in the first two lines (24-25), so the cleavage was completed. F) Native-page gel, with 27-28) the enzyme alone, 29-31) the enzyme with the dsDNA.

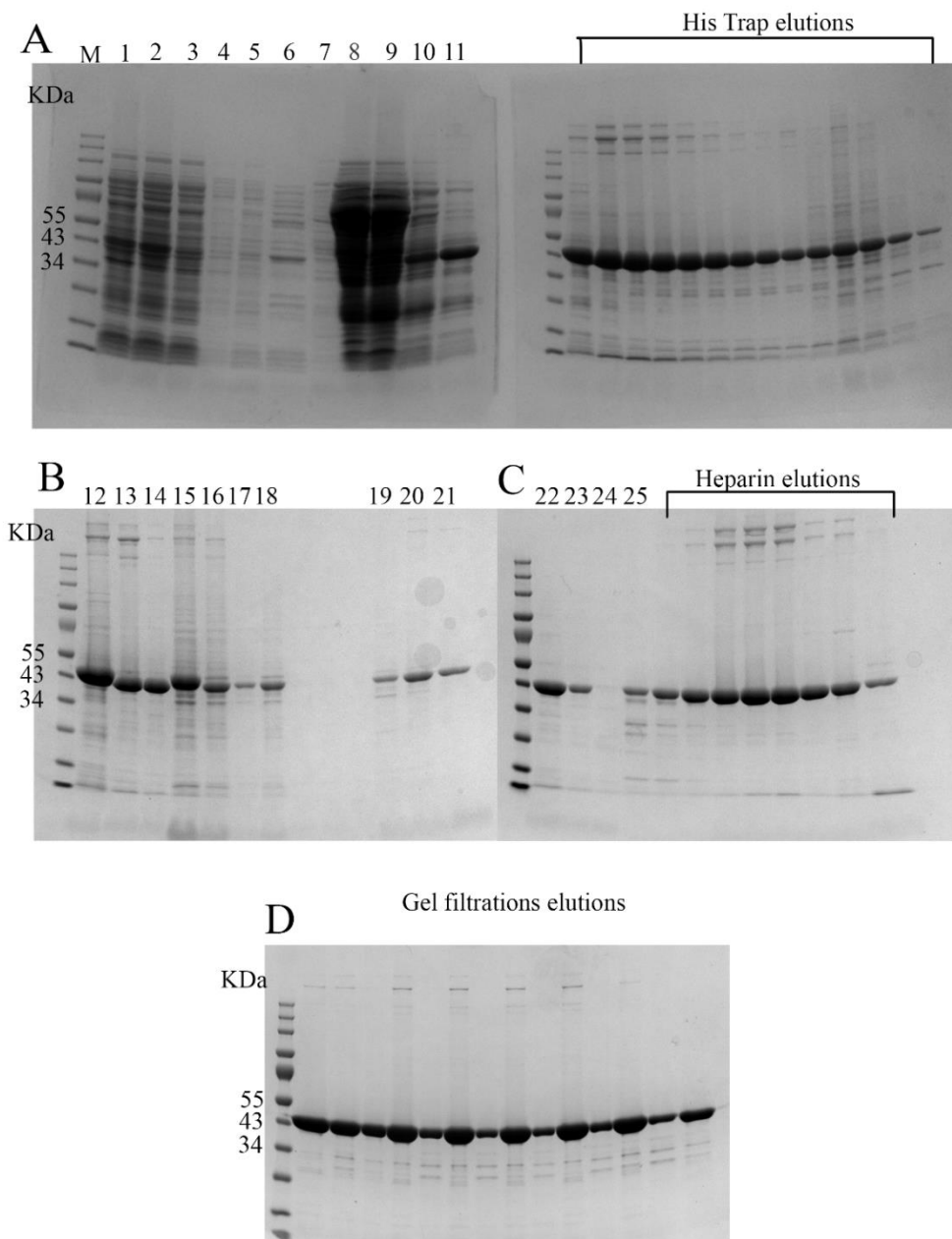


Figure 24 A) SDS-page with 1) total lysate of the protein pellet, 2) surnatant, 3) flow through of the his-trap column, 4-5-6) wash with S2 buffer, 7-8-9-10) 10% of buffer B. The SDS-page on the right is showing the His-trap elution with the gradient of imidazole of the new protocol set up at the EMBL of Grenoble. B) The SDS-page to check the cleavage, 12) pull of the His-trap elution, 13) cleavage, 14) filtered cleavage, 15) pull of the second His-trap elution, 16) cleavage of the second pull after 12 h, 17) cleavage of the second pull after 15 h, 18) filtered cleavage, 19-20-21) second pull cleavage after 18, 19 and 20 h. C) the SDS-page with 22-23) samples loaded in the heparin column, 24-25) Flow through of the heparin, the rest of the samples are the elution of the heparin. D) the SDS-page with all the elution from the injections of the gel filtration column.

3.2.3 Thermofluor characterization and differential scanning calorimetry

Once we optimized the purification protocol of Pol η (1-432), we evaluated the stability of the enzyme in different buffer using the Rubic screen to understand which the best condition for the crystallization was (Figure 25).

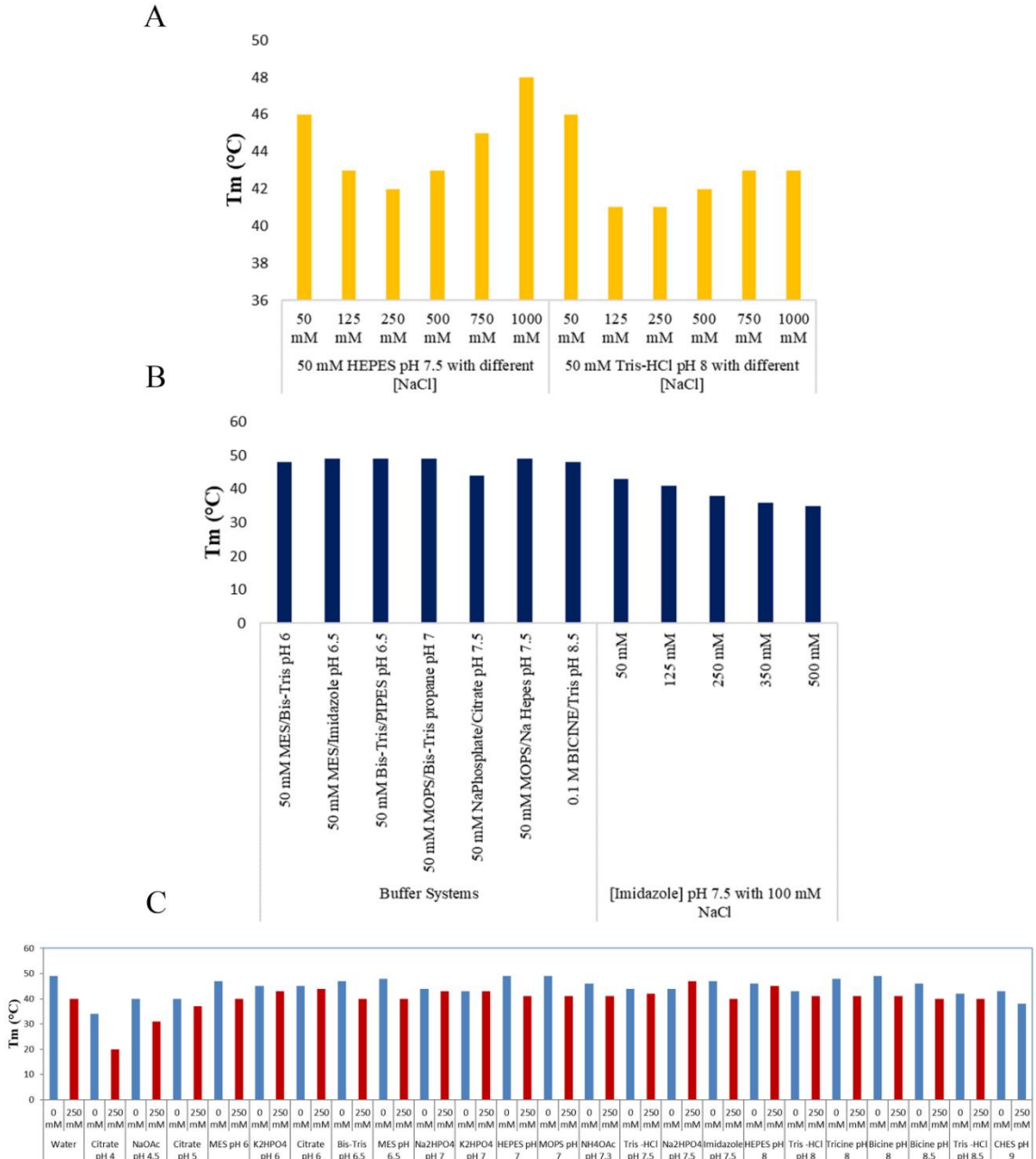


Figure 25 Thermal Shift Assay. A) The T_m related to two buffers at a fixed concentration with different concentration of salt. B) The T_m changing the buffers. C) The T_m with all the possible buffers with different pH. These data were performed one to check the stability of the sample.

This screen from Molecular Dimensions is very useful because allows to screen different pH, salt and buffer.

From this, we understood that the best condition for Pol η (1-432) was pH 5.5/6, that we should decrease the concentration of NaCl that also MES was a good condition for crystallization.

Table 7 Comparison between two different techniques to measure the thermal stability of the enzyme. TSA is the thermal stability assay, DSF is the differential scanning calorimetry performed with Prometheus (as reported in the methods). The '+' sign represents how many degrees higher the Tm of the sample is compared to the control (whose temperature was reported) (1 °C = +). Legend: E = enzyme Pol η ; DNA= dsDNA (sequence reported in the methods); A = nucleotide dATP; 64 = ARN24964 (lead compound); 12= ARN17212 (hit compound).

	TSA	DSF
E	37 °C	/
E+DNA	+++++++	36 °C
E+DNA+A	+++++++	+
E+64	++++	/
E+DNA+64	+++++++	+
E+DNA+A+64	+++++++	++
E+12	+++++++	/
E+DNA+12	+++++++	+
E+DNA+A+12	+++++++	+++

Another useful experiment was performed through the HTX platform: the thermofluor. Each time we started the set up of a crystallization plate, we used a small amount of the sample for a thermofluor to check the stability of the sample and also to check if the sample was temperature sensitive, in order to incubate crystallization plates at 20 ° or 4 °C⁸. We also performed the thermofluor assay in order to realise if the compound could stabilize the enzyme. Thermofluor is not the most reliable technique to have an indirect proof of enzyme: compound binding, therefore we also performed a thermal stability screening through a Prometheus instrument – Nanotemper technologies to make a comparison between the two techniques (Table 7). First, we evaluate the stability of the protein that increases with the binding of the dsDNA and even more with the nucleotide (E+DNA+A), but what is interesting is that the stability seems to improve also with our lead compound 64. Specifically, the Tm increases a lot with compound 64 in the thermofluor assay, while in nanoDSF (Prometheus – Nanotemper Technologies) the Tm seems to increase less. In nanoDSF experiments, compound 12 Tm seems to increase more than for compound 64.

In conclusion, what is clear is that the compounds help the stability of the enzyme. This result is confirmed by both the techniques. To note that the nucleotide used in these experiments was the normal one, not the non-hydrolyzable.

3.2.4 Isothermal titration Calorimetry (ITC)

The characterization of the binding between the enzyme and the compound was performed using the micro PEAQ-ITC (**Figure 22**). For these experiments we used the enzyme Pol η at 30 μM final concentration diluted (after the purification) in the buffer composed by 50 mM Hepes pH 7.5, glycerol 10% (v/v) and 2 mM DTT. The compound 64 was dissolved at 200 mM in DMSO and then diluted to 300 μM . For the dilution we used the buffer formed of 50 mM Hepes pH 7.5, 8.5 mM KCl, 10 % glycerol (v/v) and 2 mM DTT, to be sure that in the cell and in the syringe there was the same buffer, 0,3 % of DMSO was added in the sample with the enzyme. The first experiment was planned to check if the concentration of the protein was fine, then we used as ligand the dATP and in the cell we used the enzyme in complex with the dsDNA and the Mg^{2+} (**Figure 26A**). The K_d , as expected, was in the nanomolar range and the binding reaction came out to be a exothermic reaction. When we performed the second experiment changing the conditions, so the Mg^{2+} in the syringe at 5 mM concentration, and in the cell the enzyme in complex with the dATP and the dsDNA, we noticed that the reaction became endothermic and the K_d resulted to be in the micromolar range (**Figure 26B**). From these results was not possible to confirm the binding of the Mg^{2+} ions, because it could be that the ITC measured the heating of the elongation of the dsDNA reaction. When we added in this system our inhibitor compound 64 (**Figure 26C**), the number of binding sites decrease drastically and there is an order of magnitude increase in enthalpy. This could suggest that conformational changes may occur with the rearrangement of the protein or the polymerisation reaction. The elongation reaction happened very fast, and we cannot distinguish between the binding and the elongation of the dsDNA. However, if we did not add the dsDNA the enzyme was not stable and tends to aggregate. Unfortunately, our enzyme was not suitable for the ITC technique because is a very complex multicomponent system.

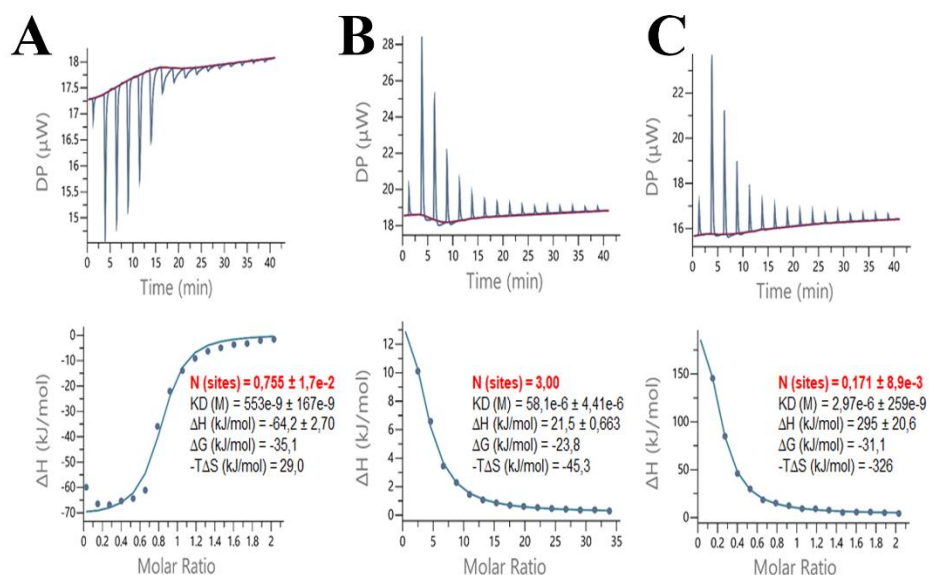


Figure 26 ITC curves. *A)* binding curve between Pol η (in the cell, in complex with the dsDNA and Mg^{2+}) and the dATP (in the syringe); *B)* binding curve between the enzyme (in the cell in complex with the dsDNA and the dATP) and the Mg^{2+} (in the syringe at 5 mM concentration); *C)* binding curve between the enzyme (in the cell in complex with dsDNA, the dATP and the Mg^{2+}) and the compound 64 (in the syringe).

3.2.5 X-ray crystallography

The crystal structure of the catalytic core of Pol η has previously been solved^{10,11, 12,13, 14}. The structure was solved in the presence of dsDNA and with or without the nucleotide. Different types of point mutations of the catalytic core were also studied as well as the mechanism for bypassing different DNA lesions, which was in depth clarified. Both normal and non-hydrolyzable nucleotide were used in these systems.

In this project, together with HTX-Lab, we crystallized Pol η (1-432) in complex with the dsDNA (sequences are reported in the methods) and the dAMPNPP.

We screened numerous conditions and we found at first a crystallization condition with ammonium formate 0.2 M, polyvinylpyrrolidone 10 % w/v and PEG 4000 20 % w/v. The crystals grew in 24 hours but they were small and when we sent them to the MASSIF-1 beamline, we noticed that they were salt crystals. The same happened with crystals grown in di-ammonium phosphate 0.2 M and PEG 3350 20 % w/v. Then, we analysed crystals that were grown in 24 h in 0.1 M MES pH 6.5 and 25% PEG 2000. We let them grow for a minimum of 5 days to a maximum of 15 days in the

crystallization condition before diffraction. We confirmed that this is the best condition for Pol η crystallisation and it is the one that has already been reported by Biertumpfel, et al., 2010. Thanks to these data, we were able to solve two structures with the resolution lower than 2Å. The first one contained dsDNA and Mg^{2+} and the other one contained dsDNA, dAMPNPP and Mg^{2+} (**Figure 27**). It was clear from some structures that the Mg^{2+} ions in the active site were not always two, when the enzyme is lacking of the incoming nucleotide, the Mg^{2+} ion is only one. While the presence of the dAMPNPP or dATP seems to allow two or three ions in the active site^{15,16}. Once we got the right condition and a good resolution, we started soaking with our best compound 64 and with the initial hit compound 12. In the hosting lab, we could not perform the canonical soaking, but we used the mosquito instead and, thank to this technology, we could use very little amount of compound. At the same time, soaking was faster than a manual procedure. We were able to use all the crystals that were grown in each drops, moreover we could better control the incubation time and, thanks to the system that takes picture of the drops we could also check if the crystals were dissolved by the DMSO during the time or not. Another benefit of using mosquito robot to perform soaking was that we could use also the smallest crystal, while manually was not possible because it would have been impossible to catch them. We used 10, 20 or 40 mM of compounds because we previously performed the solubility and stability studies on the two compounds, so we were confident that we could use them at high concentration. Moreover, the crystals seemed to be stable also after an overnight incubation.

In co-crystallisation trials, we instead added the compounds before making the crystallization drops. Specifically, we introduced the compound in the protein solution after the incubation of the enzyme with the dsDNA and the Mg^{2+} and only when the sample had already been concentrated to the final concentration for the crystallization. We also tried conditions with or without the dAMPNPP or dATP. In this case, we used lower concentration of compounds, in particular we tried 1, 2 or 5 mM. We let the compound interact with the enzyme for 10 or 30 minutes, the first time on ice, then at room temperature because on ice the compounds seem to precipitate. Finally, we injected the drops in the plate with crystallization condition 0.1 M MES pH 6.5, 20-28 %w/v PEG 2000 and 5 mM $MgCl_2$. The crystals started to grow overnight, although together with the compounds they became dark, in particular the crystals were dark brown or yellow. Indeed, usually we could take pictures of the drops and we used also the UV to check if the crystals were of proteins or salt, but in the presence of the two compounds the pictures with UV were empty.

Probably these two compounds adsorbed at that wavelength. We obtained a lot of crystals with co-crystal technique, unfortunately a low number of these diffracted when we analysed them through MASSIF-1, and of these in no one we saw the electron density of the compounds. We assumed that the compounds need more Mg^{2+} to bind the active site. Surely, we know from our studies that the compounds somehow could block the enzyme and probably, from the chemical motif, they could chelate the Mg^{2+} ions. Something to consider is that when dAMPNPP is bound in the active site of the enzyme, the nucleotide coordinates two Mg^{2+} ions and blocks the enzyme in a specific conformation. This nucleotide cannot be hydrolysed and may hinder the entrance of the compounds in the active site. At the same time, it was difficult to obtain crystals in the presence of dATP because the polymerisation reaction could go on under those condition. For all these reasons, we tried to perform a double soaking, first with a solution of 5 mM of $MgCl_2$, and then with the compounds. As reported in the literature, it was possible to utilise calcium as an ion instead of magnesium. In this way, the elongation reaction was slowed down and it was possible to obtain crystals with the normal nucleotide. In particular, in the article of Chang et al.¹⁷, they obtained crystals using calcium; they then soaked them in a manganese solution (2 to 10 mM). They had three important results:

- 1) an ion exchange was detected,
- 2) manganese is more able to remain in the active site than magnesium,
- 3) most importantly, in the active site, even with manganese, were present three ions (instead of two).

During our experiments, we tried to replicate their data, so we used the purified enzyme and incubated it with calcium and dsDNA. Once the sample was concentrated, we added dATP and made the drops manually and not using the mosquito. We left the plates at room temperature (instead of 20° C) for about 14 days. Unfortunately, due to lack of time, we were not able to perform the soaking but the crystals were small in size. However, our collaboration with HTX-lab is still ongoing and further experiments are planned to obtain the co-crystal through double soaking. As future experiment, we also plan a high-throughput screening of a fragments library in order to perform a more in depth hit-selection.

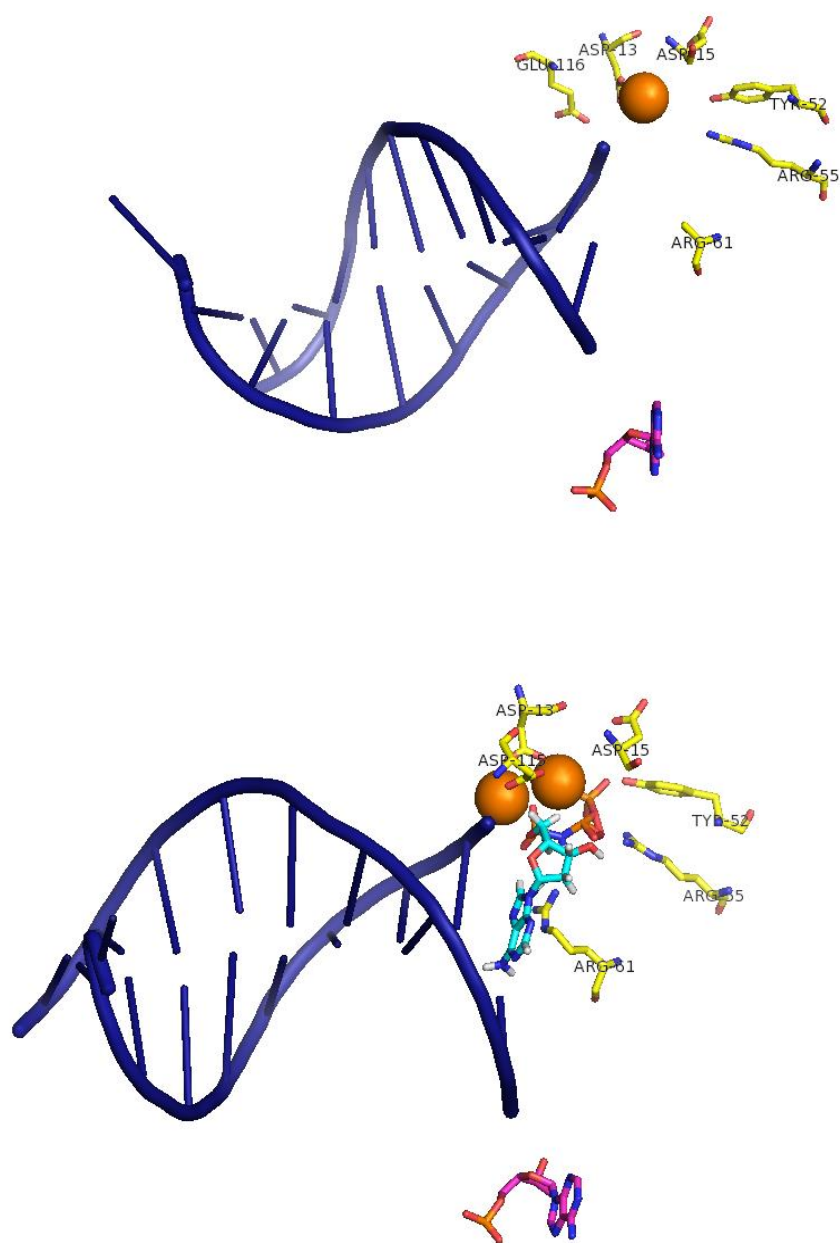


Figure 27 Coordination of the Mg^{2+} ions in the active site. The two structures reported represent the active site in the presence/absence of the nucleotide.

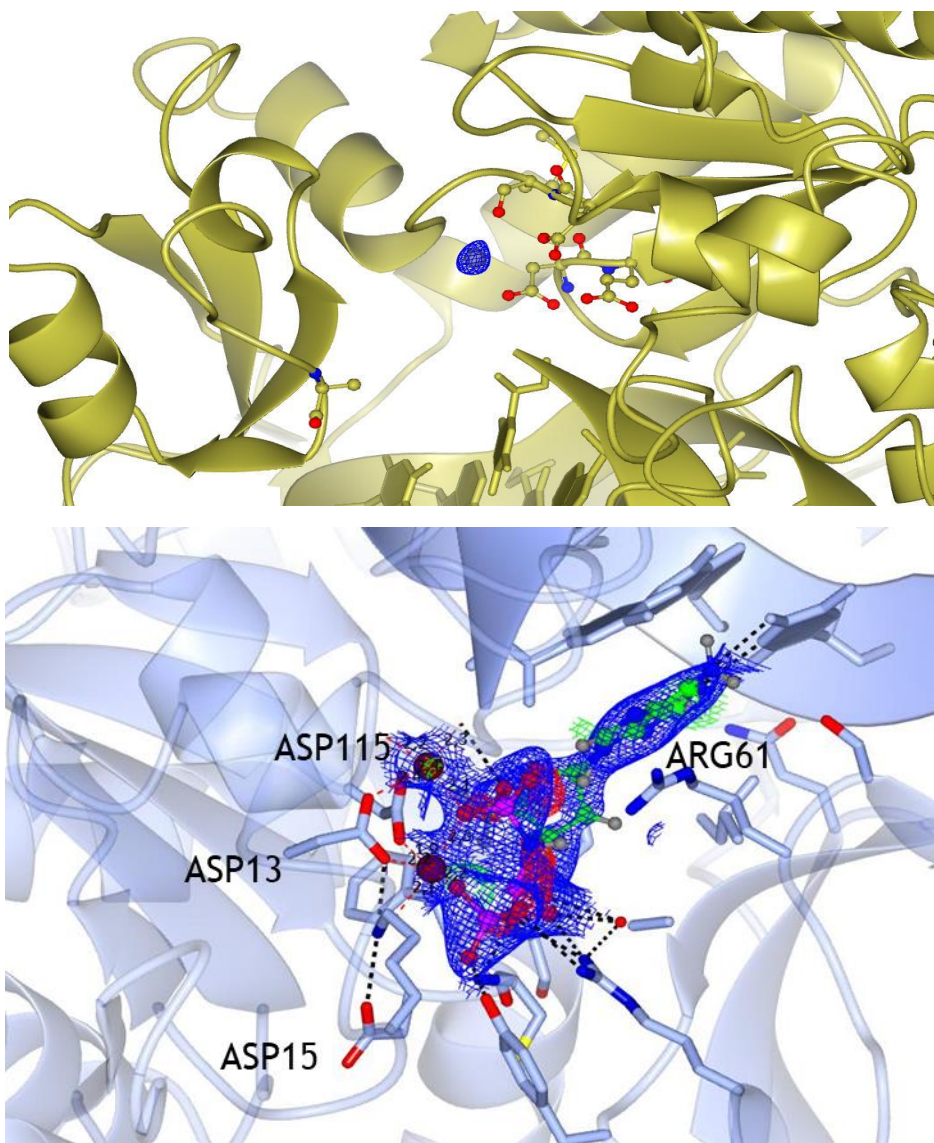


Figure 28 The structure of the active site of the two best structures that we obtain at the EMBL of Grenoble. In blue are reported the density map of the Mg^{2+} ion and of the dAMPNPP (non-hydrolyzable nucleotide).

d

Data reduction statistics	
Space group	P61
Unit cell dimensions (Å\°)	98.315 98.315 80.988 \ 90 90 120
Resolution range (Å)*	18.84-1.897 (2.04-1.897)
mean I/sigma (I)*	9.4 (1.4)
Completeness (%)*	95.3% (56.4%)
CC1/2*	0.996 (0.472)
Rmerge (%)*	0.236 (2.137)
Total reflections*	399751 (19050)
Number of unique reflections*	29671 (1485)
Multiplicity*	13.5 (12.8)
REFINEMENT STATISTICS	
No. Reflections*	29627 (51)
Rwork*	0.200 (0.264)
Rfree*	0.253 (0,307)
Mean B-factor	31.91
RMS bond lengths deviations	0.0115
RMS bond angles deviation	1.44
Molprobability score	1.30
RAMACHANDRAN (%)	
Favored	97.39%
Allowed	1.9%
Outliers	0.47%
*Values in parentheses are for the highest-resolution shell	

Figure 29 Statistics data of the structure with the two Mg²⁺ ions and the nucleotide. The statistic of the other structure cannot be presented because refinement is not complete.

3.2.6 Cryo-EM

At the EMBL of Grenoble, in the HTX-Lab I had also the opportunity to know and learn the basics of the Cryo-EM technique. The technique has improved enormously in the last few years and it is now comparable in throughput to X-rays, but like all techniques, it has its advantages and disadvantages. Among the advantages is that the sample does not need to be very concentrated, therefore if it is not stable or if it is very difficult to obtain in large amount, it is easier to test with cryo-EM than with crystallography. On the other hand, a disadvantage is that this technique is not yet able to solve the structure of proteins with a molecular weight below 100 kDa with good

resolution. We were aware that the active site of our enzyme did not meet the cut-off criteria, as the molecular weight was around 60 KDa including dsDNA. However, given the difficulty in obtaining the purified protein at 10 mg/mL, and especially the difficulty in obtaining the co-crystals, we decided to try the technique to achieve our goal.

As a first step, we optimised the grids. There are no studies in the literature on using this technique to solve Pol η 's structure, so we started by setting everything from the beginning: we tried using both UltrAufoils and Quantifoils. As a concentration, we tried 2 mg/mL and 1 mg/mL of the sample that already contained the protein, dsDNA, Mg^{2+} and also dAMPNPP. We knew that the buffer in which the protein was stored was no good because there was 10 % glycerol (v/v) in it, so we diluted the protein and tried to lower the glycerol concentration as well. Finally, we tested the samples with: 10 % glycerol (v/v) and 2 mg/mL protein; with 2% glycerol (v/v) and 1 mg/mL protein; and lastly with 2% glycerol (v/v) and 2 mg/mL protein. We used approximately 4 μ L of sample on each side of the grid. From the first screen of the grids, we found that the best condition was the one with 10 % glycerol (v/v) and 2 mg/mL protein. We also tried grids with no glycerol, 1 mg/mL protein and 0.01 % triton (v/v). We screened 4 grids, of which only one, the Quantifoil Au 300 1.2/1.3, was acceptable for an initial data collection (**Figure 30**). The data collection was set up overnight and 521 movies were collected. Using cryoSPARC, we analysed the data. We were able to choose 2D classes, selected 40 000 particles and obtain a 3D model. Unfortunately the resolution was very low at 8.26 Å, so the fitting made with the structure in the literature (PDB: 3MR2) was not the best (**Figure 31**). In particular, we noticed that part of the structure seemed to be missing, probably because it was one of the very flexible parts of the enzyme. Or, we could assume an opening of the active site caused by the entry of the non-hydrolysable nucleotide. This could also be related to the result obtained with ITC where we had seen an increase in enthalpy, which could also be due to the conformational change of the enzyme. Here too, unfortunately, it was not possible to carry out any further experiments due to lack of time with Glacios. However, given the size of the enzyme and that our compound is a small molecule, it would not have been the best technique to be able to resolve the structure of the complex and analyse the binding interactions between the protein and the inhibitor at the atomic level.

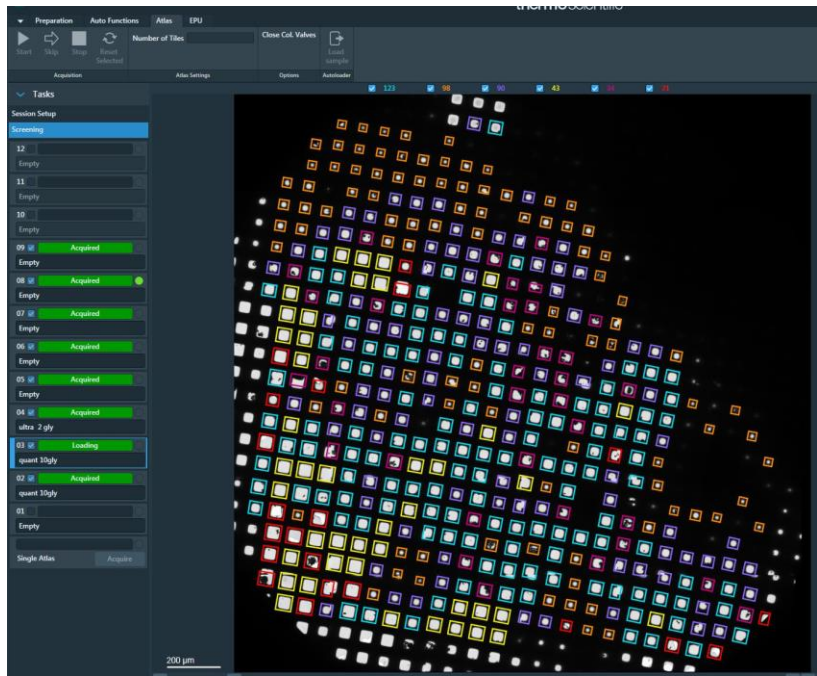


Figure 30 The grid chosen during the screening and the 2D classes. Selected for the 3D reconstruction.

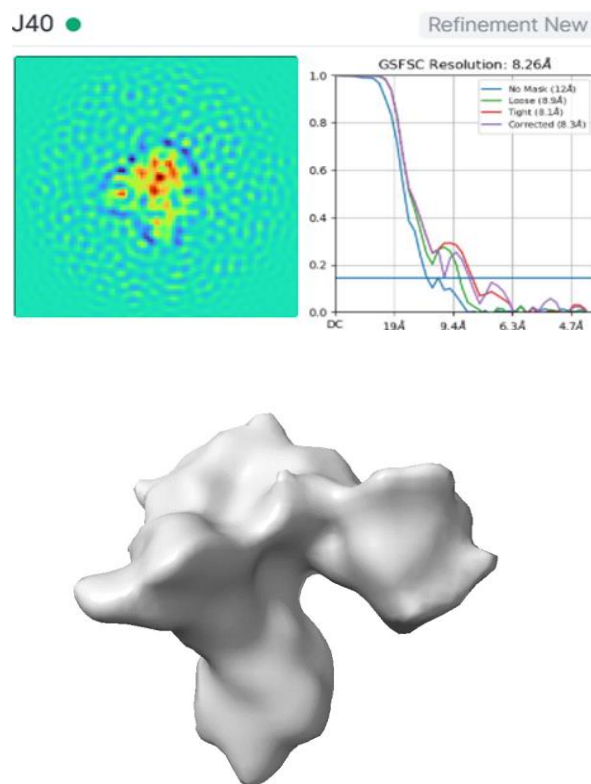


Figure 31 Refinement. The graph of the final refinement indicates that the resolution reached is at 8.26 Å. The resolution was very low and it was impossible to fit the structure in this model.

3.3 Discussion

In summary, we confirmed the interaction between the compounds and the active site of the enzyme through the gel-based elongation assay. Indeed, thanks to this assay we measured the inhibitory activity of each compound, among 35 analogs we were able to select 10 best compounds with an IC_{50} lower than 20 μ M. The most active ones were also evaluated on cancer cell lines, in particular we used three different cell lines: A375, A549 and OVCAR3. Unfortunately, we did not reach a clear result with ITC because it is a very complex multi-component system in which is difficult to measure just the energy of the binding. We also tried the MicroScale Thermophoresis (MST) (data not shown), although the compounds are fluorescent and, for this reason, the compound could affect the binding curve.

We saw that compound 64 could help to stabilize the protein through TSA and DSC, but, at the same time, we were not able to obtain the co-crystal structure. These compounds at higher concentration are probably not stable for long time. We need to further investigate the stability of the compounds over time. The crystals that we obtain from the co-crystallization became darker, this could mean that the protein interact with the inhibitors, but they did not diffract.

Until now, there are not structure of the co-crystal with the same construct. For this reason, we are planning to continue the crystallization experiments, also we would like to try to use the full length sequence of the protein to have another chance with the Cryo-EM technique. Regarding Cryo-EM, until now there is only one structure of Pol η at 20 Å of resolution, we improved the resolution with these few experiments. In the future, we could keep improving the resolution by trying to reduce the glycerol content while using the full length protein construct.

3.4 References

- (1) Vonrhein, C.; Flensburg, C.; Keller, P.; Sharff, A.; Smart, O.; Paciorek, W.; Womack, T.; Bricogne, G. Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr D Biol Crystallogr* **2011**, *67* (Pt 4), 293-302. DOI: 10.1107/S0907444911007773
- (2) Evans, P.; McCoy, A. An introduction to molecular replacement. *Acta Crystallogr D Biol Crystallogr* **2008**, *64* (Pt 1), 1-10. DOI: 10.1107/S0907444907051554
- (3) A., S.; P., K.; C., V.; O., S.; T., W.; C., F.; W., P.; I., T.; R., F.; M., W.; et al. *Pipedream, version 1.3.1*. Global Phasing Ltd, Cambridge, United Kingdom, 2021. (accessed).
- (4) Bricogne, G.; Blanc, E.; Brandl, M.; Flensburg, C.; Keller, P.; Paciorek, W.; Roversi, P.; Sharff, A.; Smart, O.; Vonrhein, C. BUSTER Version XYZ Global Phasing Ltd. *Cambridge, United Kingdom* **2011**.
- (5) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010**, *66* (Pt 4), 486-501. DOI: 10.1107/S0907444910007493
- (6) Punjani, A.; Rubinstein, J. L.; Fleet, D. J.; Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **2017**, *14* (3), 290-296. DOI: 10.1038/nmeth.4169
- (7) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Meng, E. C.; Couch, G. S.; Croll, T. I.; Morris, J. H.; Ferrin, T. E. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* **2021**, *30* (1), 70-82. DOI: 10.1002/pro.3943
- (8) Dupeux, F.; Rower, M.; Seroul, G.; Blot, D.; Marquez, J. A. A thermal stability assay can help to estimate the crystallization likelihood of biological samples. *Acta Crystallogr D Biol Crystallogr* **2011**, *67* (Pt 11), 915-919. DOI: 10.1107/S0907444911036225
- (9) Beardslee, R. A.; Suarez, S. C.; Toffton, S. M.; McCulloch, S. D. Mutation of the little finger domain in human DNA polymerase eta alters fidelity when copying undamaged DNA. *Environ Mol Mutagen* **2013**, *54* (8), 638-651. DOI: 10.1002/em.21807
- (10) Biertumpfel, C.; Zhao, Y.; Kondo, Y.; Ramon-Maiques, S.; Gregory, M.; Lee, J. Y.; Masutani, C.; Lehmann, A. R.; Hanaoka, F.; Yang, W. Structure and mechanism of human DNA polymerase eta. *Nature* **2010**, *465* (7301), 1044-1048. DOI: 10.1038/nature09196

- (11) Zhao, Y.; Biertumpfel, C.; Gregory, M. T.; Hua, Y. J.; Hanaoka, F.; Yang, W. Structural basis of human DNA polymerase η -mediated chemoresistance to cisplatin. *Proc Natl Acad Sci U S A* **2012**, *109* (19), 7269-7274. DOI: 10.1073/pnas.1202681109
- (12) Ummat, A.; Silverstein, T. D.; Jain, R.; Buku, A.; Johnson, R. E.; Prakash, L.; Prakash, S.; Aggarwal, A. K. Human DNA polymerase η is pre-aligned for dNTP binding and catalysis. *J Mol Biol* **2012**, *415* (4), 627-634. DOI: 10.1016/j.jmb.2011.11.038
- (13) Patra, A.; Su, Y.; Zhang, Q.; Johnson, K. M.; Guengerich, F. P.; Egli, M. Structural and Kinetic Analysis of Miscoding Opposite the DNA Adduct 1,N6-Ethenodeoxyadenosine by Human Translesion DNA Polymerase η . *J Biol Chem* **2016**, *291* (27), 14134-14145. DOI: 10.1074/jbc.M116.732487
- (14) Weng, P. J.; Gao, Y.; Gregory, M. T.; Wang, P.; Wang, Y.; Yang, W. Bypassing a 8,5'-cyclo-2'-deoxyadenosine lesion by human DNA polymerase η at atomic resolution. *Proc Natl Acad Sci U S A* **2018**, *115* (42), 10660-10665. DOI: 10.1073/pnas.1812856115
- (15) Wang, J.; Smithline, Z. B. Crystallographic evidence for two-metal-ion catalysis in human pol η . *Protein Sci* **2019**, *28* (2), 439-447. DOI: 10.1002/pro.3541
- (16) Yoon, H.; Warshel, A. Simulating the fidelity and the three Mg mechanism of pol η and clarifying the validity of transition state theory in enzyme catalysis. *Proteins* **2017**, *85* (8), 1446-1453. DOI: 10.1002/prot.25305
- (17) Chang, C.; Lee Luo, C.; Gao, Y. In crystallo observation of three metal ion promoted DNA polymerase misincorporation. *Nat Commun* **2022**, *13* (1), 2346. DOI: 10.1038/s41467-022-30005-

3

Chapter 4: Hyaluronic acid-based prodrugs for efficient bioavailability

Hyaluronic acid (HA) is a well-studied natural polymer of the glycosaminoglycan heteropolysaccharides (GAGs) group¹. HA and all the GAGs group are the main component of the extracellular matrix. The chemical structure is simple, it is composed by repetition of disaccharide units of D-glucuronic acid and N-acetyl-D-glucosamine with alternating β -1,3 and β -1,4 glycosidic bonds. Usually, the HA scaffold can vary in molecular weight depending on the number of repetitions included. Two mechanisms of HA degradation are known: one is mediated by specific enzymes called hyaluronidases (HYALs), the other is mediated by oxidative damage due to reactive oxygen species^{2,3}.

To date, interest in hyaluronic acid has grown significantly, mainly because it has been reported to be an excellent carrier for drugs and small molecules. The polymer indeed has good biocompatibility and it is capable to target CD44⁴ receptors that are actually involved in HA internalization in cells. Interestingly, CD44 receptors are known to be overexpressed in several solid cancers and in inflammatory pathologies. On these premises, we started a collaboration with Dr. Tirelli, principal investigator of the Polymers and Biomaterials lab at the IIT (Genova). In one of their works, they improved the bioavailability of quercetin solubilizing it in a HA prodrug scaffold. Specifically, Tirelli's group has exploited a boronic ester as linker between HA and quercetin. The group is stable at neutral or mildly basic pH, although it can be cleaved when in acidic pH or through the exposure to hydrogen peroxide⁵ leading to the release of the drug.

Quercetin is a drug that belongs to the flavonoids⁶, a class of natural polyphenols known for their role as oxidant scavenger. They are used to treat pathologies such as cancer but also in chemoprevention. However, quercetin has the disadvantage of being quickly eliminated through the bile, thus explaining the need of improving its bioavailability⁷ (**Figure 32**). Our inhibitors belong to the flavonoid class. By studying their pharmacokinetic properties, we realised that they have excellent kinetic solubility and good metabolic properties, but they have a half-life that is not optimal for *in vivo* studies. Therefore, we started this collaboration for the synthesis of prodrugs with hyaluronic acid to see whether this formulation could improve the bioavailability of our small molecules while maintaining the inhibitory and synergistic activity with cisplatin. Here, we report

our preliminary results using the prodrug on our three cancer cell lines. This is a promising starting point to improve the drug likeness of our compound 64.

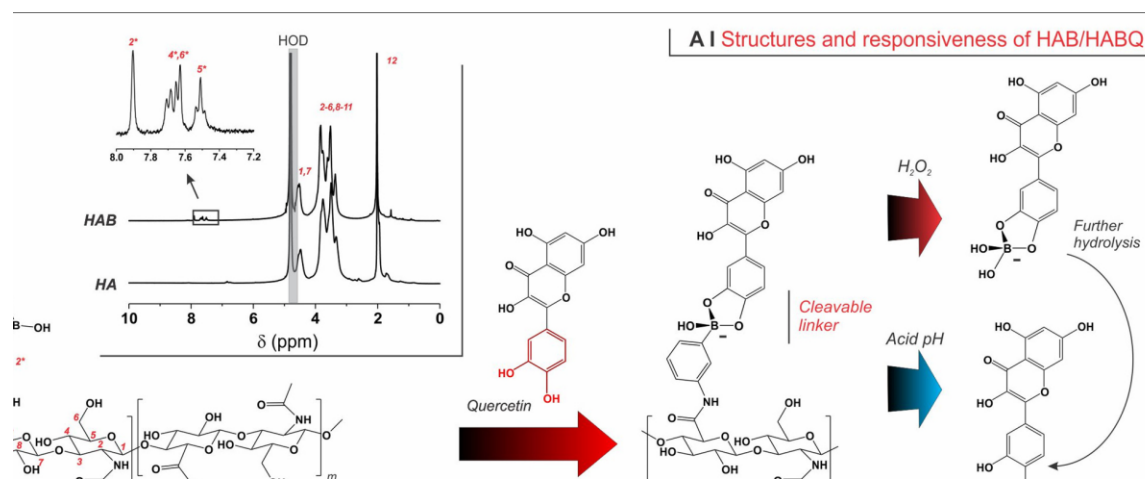


Figure 32 HAB (left) boronic acids are introduced via amidation of HA. Boronates bind catechols of quercetin and reversibly release them at low pH (transition of boronic esters from quaternarized to trigonal form followed by hydrolysis), or in the presence of oxidants such as H_2O_2 . (Image taken from Quagliarello et al., 2021)

4.1 Methods

4.1.2 Human cell culture. We used the same human cancer cell lines as reported in the Experimental section of the Chapter 2. Briefly, we used the A375, from malignant melanoma; A549, from non-small cell lung cancer; OVCAR3, from ovarian cancer. The MTT assay was also performed using the conditions previously reported in the Experimental section of the Chapter 2.

4.1.3 Cell proliferation assay. The CellTiter 96® AQueous One Solution Cell Proliferation Assay was used (Promega). It is a colorimetric assay to measure the metabolic activity of the cells. In particular, tetrazolium compound [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulphophenyl)-2H-tetrazolium, inner salt; MTS(a)] and an electron coupling reagent (phenazine ethosulfate; PES) were used. PES has enhanced chemical stability, which allows it to be combined with MTS to form a stable solution. The MTS tetrazolium compound (Owen's reagent) is bio-reduced by cells into a coloured formazan product that is soluble in tissue culture medium. This conversion is presumably accomplished by NADPH or

NADH produced by dehydrogenase enzymes in metabolically active cells. Assays are performed by adding a small amount (20 μ L) of the reagent directly into wells, incubating for 1 hour and then recording the absorbance at 490nm with the Spark Microplate multimodal reader instrument (Tecan). The advantage of this assay is that the cells remain alive, so after the reading it is possible to use those cells to calculate the amount of the proteins using a BCA kit.

4.1.4 Internalization of the prodrug. Cells were seeded in 24-well plates at a density of 5000 (A375), 10000 (A549 and OVCAR3) cells/well and allowed to adhere overnight. Cells were incubated with fluorescently-labeled HA-PBA-QC nanoparticles at 10 and 50 μ M (0.5 mL) for specific time points (0, 6, 24, 48 and 72 h). Untreated cells were used as a control. At the end of the incubation period, the HA-bio-containing medium was removed. Cells were washed two times with PBS (0.4 mL) and detached using a trypsin-EDTA solution (0.2 mL) in order to eliminate any residual membrane-bound nanoparticle. This step was included to allow the detection of internalized nanoparticles exclusively. 0.75 mL of PBS was added, cell were transferred to an Eppendorf and then centrifuged (120g for 5 min). After that, supernatant was discarded, cells were resuspended in 0.75 mL of PBS and recentrifuged. Finally, cells were resuspended in 100 μ L RIPA and stocked overnight at -80 $^{\circ}$ C. The total internalized materials, was measured from the fluorescence intensity of cell lysates by using a calibration of HA-PBA-QC-Cy3 in cell lysate. The total protein content per well was then measured using the bicinchoninic acid kit for protein determination (BCA assay form Euroclone) according to the manufacturer instructions.

4.1.5 Immunofluorescence. Cells were seeded (1000 (A375) or 5000 (A549 and OVCAR3) cells/well) on 96-well plate (CellCarrierTM-96 ultra, black, 96-well, Clear bottom, with lid, TC treated, sterile). Control and treated (with 50 μ M of HA-PBA-QC-Cy3 for 24, 48 and 72 h) cells were fixed in formaldehyde 3.7% in PBS for 10 min and washed twice in PBS. Then, triton 0.5 % (v/v) was added for 10 min and wash with PBS for three times. Phalloidin (Abcam, ab176753) was added diluted 1:1000 in PBS 1X and incubated for 10 min, then removed and washed three times. Lastly, the 1:1000 of DAPI (Hoechst 33342, Invitrogen) was added for 2 min, then washed three times with PBS 1X. Using a 96-well plate was possible to add 100 μ L of PBS1X and the samples were ready to be analysed with the Leica600 fluorescent microscope.

4.2 Results and discussions

4.2.1 Comparison between the prodrug and drug effects

As reported in the previous chapters, the best compounds were 64 and 61. However, compound 64 was found to be the better of the two as it was not cytotoxic when used alone and, more importantly, it showed a synergistic effect with cisplatin.

Firstly, we have synthesised the prodrugs. They will be called from now on as follows:

64	ARN24964
61	ARN24961
QC	QUERCETIN
HA-PBA-64	Hyaluronic Acid- PhenylBoronic Acid- ARN24964
HA-PBA-61	Hyaluronic Acid- PhenylBoronic Acid- ARN24961
HA-PBA-QC	Hyaluronic Acid- PhenylBoronic Acid- Quercitin

Once we got the prodrugs, we wanted to make sure that when used, they would have the same effect as the compounds alone, therefore we treated all three cell lines for 48h with the prodrugs and we evaluated their effect through the MTT assay.

Table 8 LD₅₀ values of the treatment for 48 and 72 hours of the drug and the prodrug alone, using MTT assay. These are preliminary data from two duplicates \pm SD.

	LD ₅₀ [μ M] 48H					
	64	HA-PBA-64	61	HA-PBA-61	QC	HA-PBA-QC
A375	>100	>100	29 \pm 13	>100	35.7 \pm 20.2	>100
A549	>100	>100	60 \pm 5	>100	>100	>100
OVCAR3	>100	>100	5.2 \pm 1.3	>100	>100	>100
	LD ₅₀ [μ M] 72H					
	64	HA-PBA-64	61	HA-PBA-61	QC	HA-PBA-QC
A375	>100	>100	17.9	>100	40	>100
A549	>100	>100	15	>100	17.3	>100
OVCAR3	>100	>100	5 \pm 2.5	>100	>100	>100

These experiments confirmed that the prodrugs used alone did not alter the effect of the compound. On the contrary, the synthesis of the prodrugs appears to have decreased cytotoxicity of compound 61, which was very cytotoxic on its own (Table 8).

Subsequently, we investigated the effect of prodrugs co-treatment with cisplatin using the same concentrations for both the molecules. Indeed, in 48h, the effect of HA-PBA-64 was not synergistic with cisplatin, but neither did it diminish its efficacy. Instead, HA-PBA-61 and HA-PBA-QC appeared to have an antagonistic effect with cisplatin. This effect can be explained by the fact that, as anticipated, quercetin and all flavonoids have a role as ROS scavengers, so instead of helping the tumour cell to go into apoptosis after cisplatin-induced damage, they have the opposite effect.

When using prodrugs, it has to be considered that the compound does not reach the cell directly and does not enter as quickly as when it is not modified. On the contrary, in this case it is necessary to wait for the hyaluronic acid to recognise the CD44 receptor, for it to be internalised via lysosomes (although the mechanism is still not entirely clear) and for the compound to be released in order to act. Therefore, we did internalisation experiments to see if the compound needed more time to enter and to be as active as the compound on its own.

To follow the internalisation of the prodrug, we used a further modified compound. In particular, a fluorophore called Cy3 was also added. The three cell lines, A375, A549 and

OVCAR3 were treated with two different concentrations of prodrug: 10 and 50 μM at different time points (0, 6, 24, 48 and 72 h) (**Figure 33**).

Specifically, when we speak of μM concentration of prodrug, we mean the concentration of compound that was attached to the hyaluronic acid polymer by a reaction as reported in the

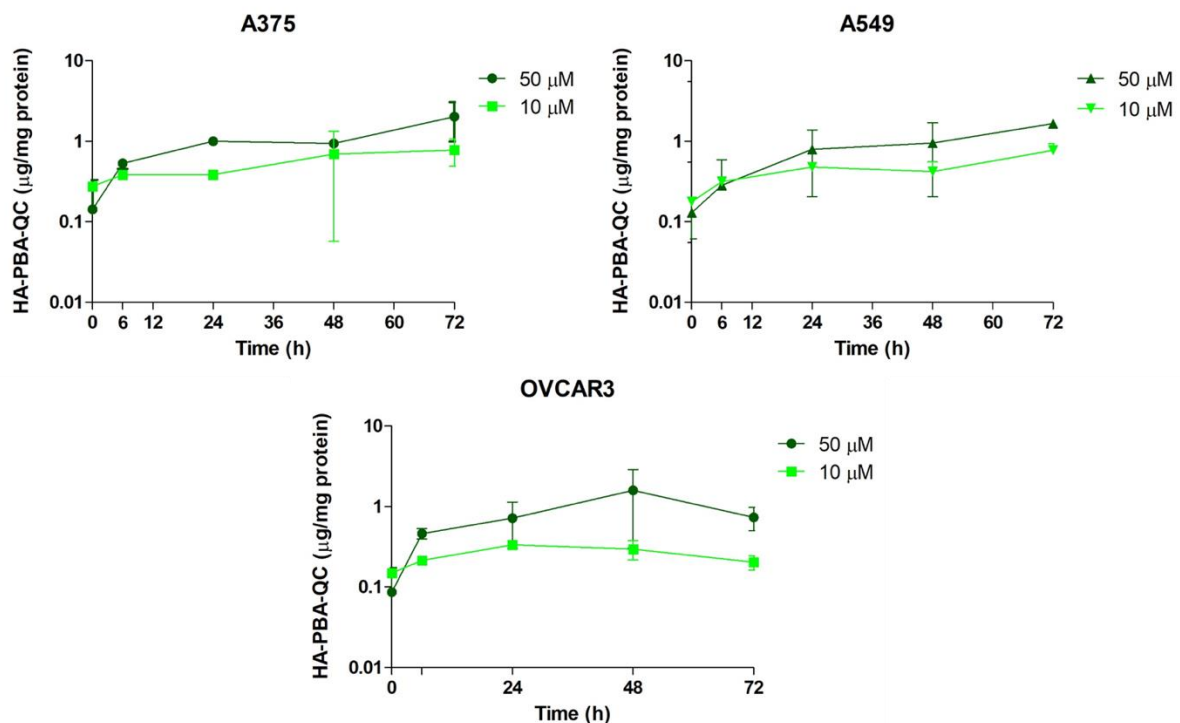


Figure 33 Internalization of the prodrug of the quercetin. Two different concentrations were used: dark green the 50 μM concentration of the drug bound to the HA; light green the 10 μM concentration of the drug bound to the HA. The reported results are from three independent experiments (mean \pm SD). No statistical significance from two-way ANOVA test.

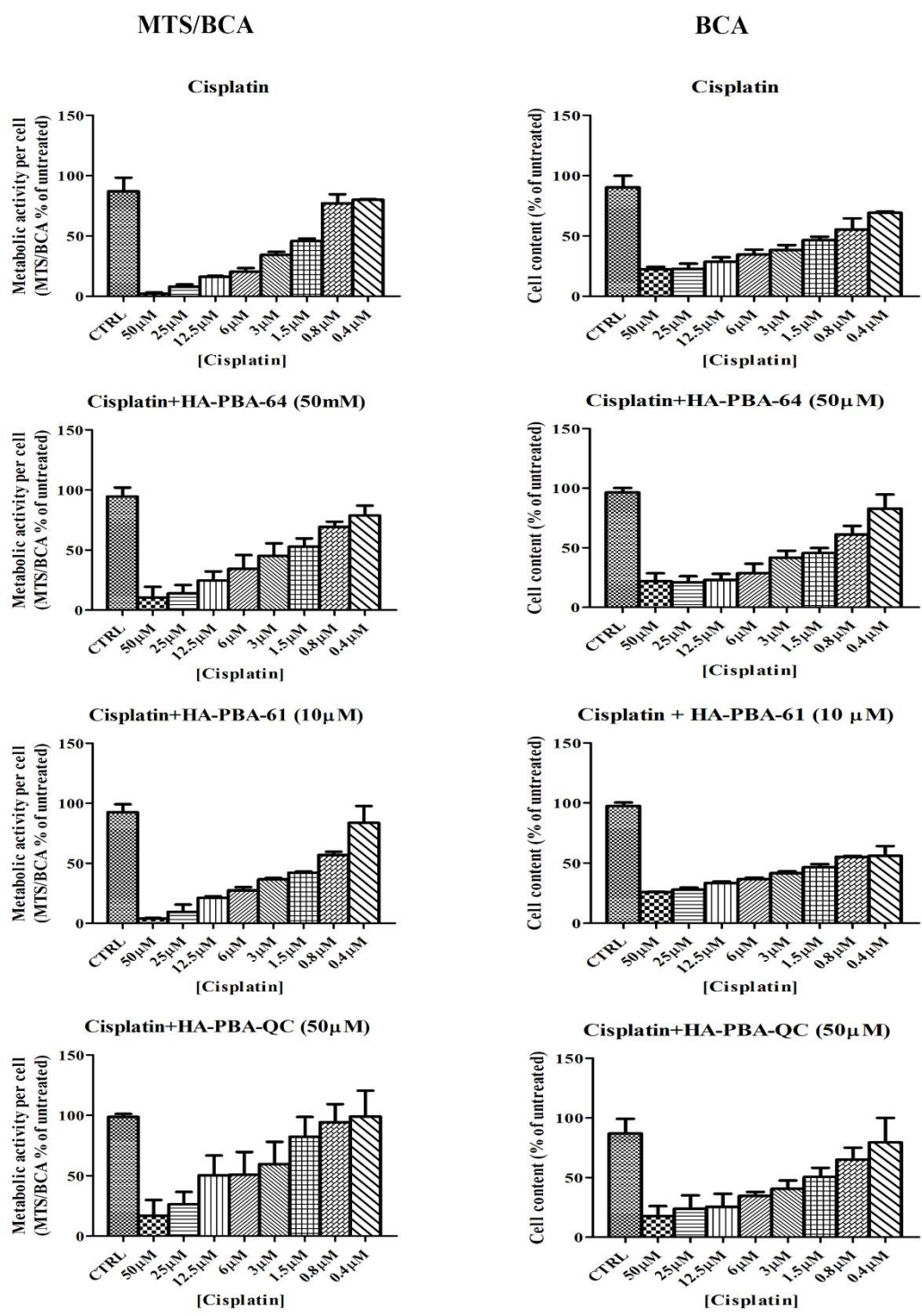


Figure 34 A375. Co-treatment using cisplatin (serial dilution, concentrations are reported in the X axis), and a fixed concentration (reported in the brackets) of the prodrugs. (Two independent experiments \pm SD).

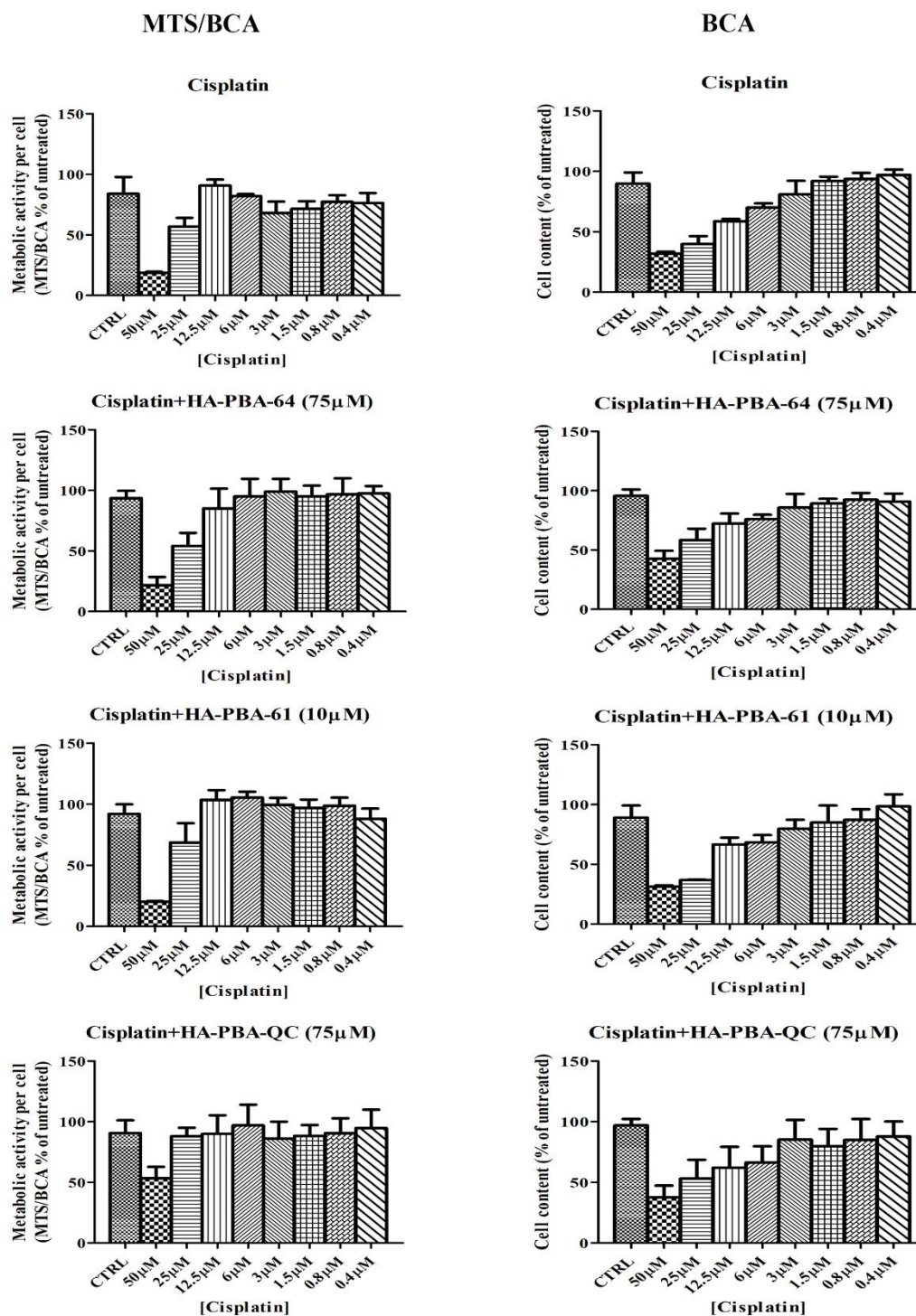


Figure 35 A549. Co-treatment using cisplatin (serial dilution, concentrations are reported in the X axis), and a fixed concentration (reported in the brackets) of the prodrugs. (Two independent experiments \pm SD).

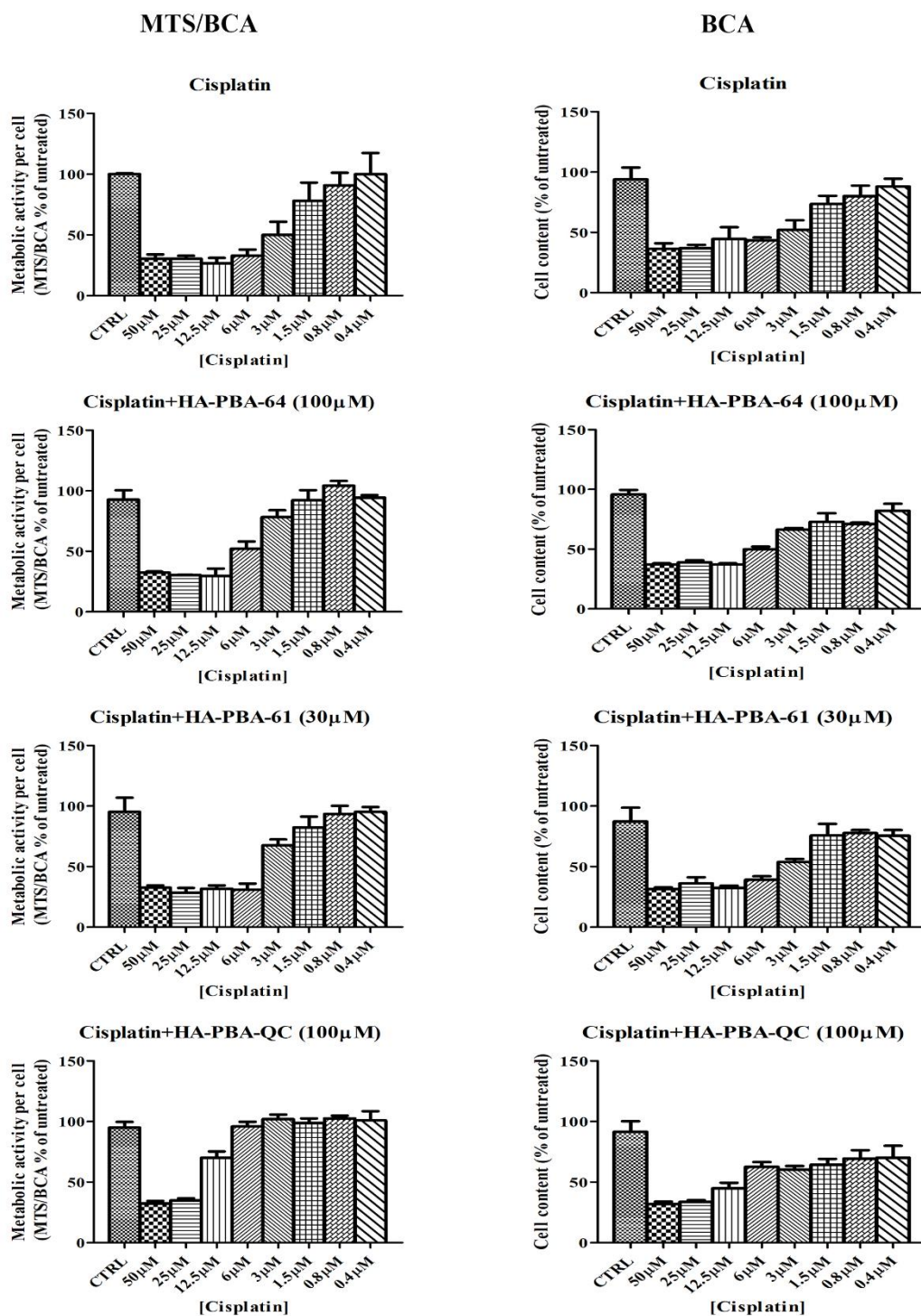


Figure 36 OVCAR3. Co-treatment using cisplatin (serial dilution, concentrations are reported in the X axis), and a fixed concentration (reported in the brackets) of the prodrugs. (Two independent experiments \pm SD).

methods of Quagliarello et al. The exact amount of compound attached via the boronic ester is measured by performing experiments using HPLC.

Indeed, it does not appear that internalisation changes much according to concentration of XXX. Especially after 6 h in lines A375 and A549, the two concentrations are perfectly overlapping as can be seen in **Figure 33**. Whereas after 72h, a plateau would appear to be reached with the highest prodrug concentration. In the OVCAR3 there would seem to be a peak of internalisation after 48h, which drops slightly after 72h, this behaviour will be better investigated. We then repeated all treatments and co-treatments with cisplatin for 72 h instead of 48 h. The results obtained from these experiments with MTT were again in line with those reported in chapter 2 and thus that compound 64 has a synergistic effect with cisplatin, while compound 61 and quercetin appear to have an antagonistic effect. We also used another cell viability assay: MTS. It is still used to assess the metabolic activity of cells, but has the advantage that it is faster, because you do not have to dissolve the formazan crystals, but just add the reagent directly into the culture medium and read at 490nm after one hour. Moreover, the cells will not die with the MTS, for this reason we could wash the well with PBS and add the RIPA buffer to lysate the cells to measure the amount of the protein with BCA kit. Thanks to this, we could measure the metabolic activity per number of cells, and these data are more accurate than MTT (**Figure 34**). What comes out from these experiments is that in all the three cell lines, the metabolic activity after 72h of co-treatment with cisplatin and fixed concentration of the prodrug seems to be completely absent. At the same time they are in line with the BCA results.

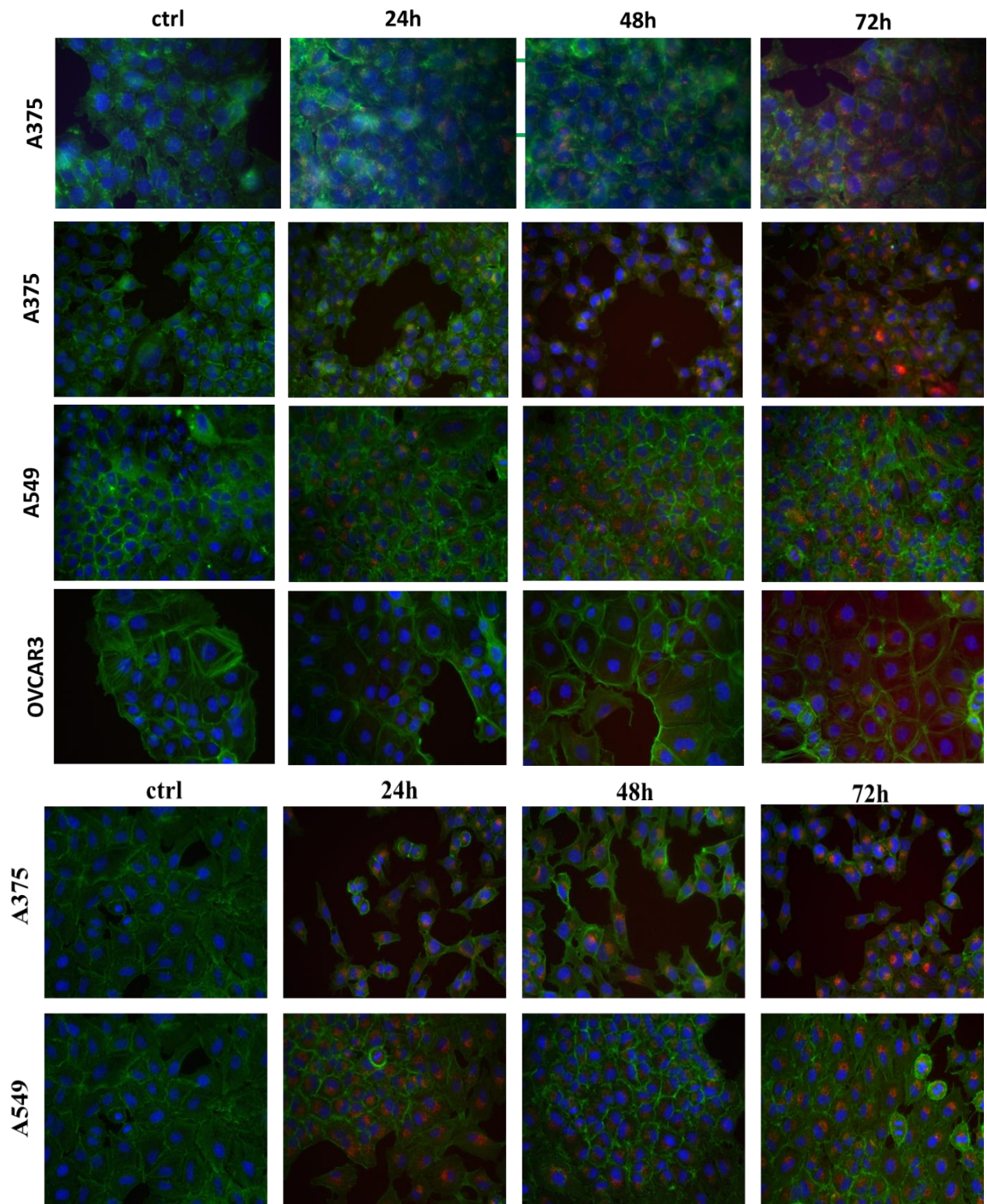


Figure 37 Immunofluorescence to follow the internalization of HA-PBA-Cy3 [20 μ M] (only A375) and HA-PBA-QC-Cy3 [50 μ M] (red) at 3 different time points. Below HA-PBA-64-Cy3 [50 μ M] (red) at 3 different time points. Nuclei were stained with Hoechst 33342 (blue). Actin was stained with the phalloidin (green). Nuclei were stained with Hoechst 33342 (blue). Actin was stained with the phalloidin (green).

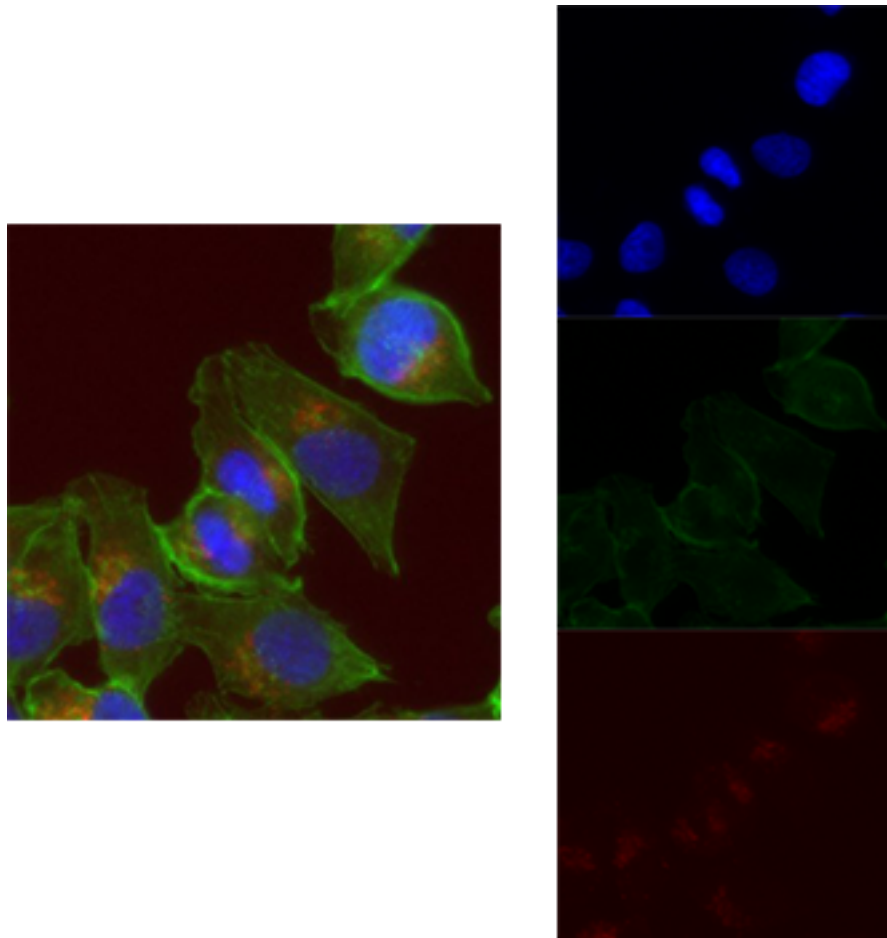


Figure 38 A375. Zoom to check the localization on the treatment with HA-PBA-64-Cy3 [50 μ M].

To further confirm the internalization data from a qualitative point of view, we performed the immunofluorescence on A375 cell line first with only the HA labelled with the Cy3 (**Figure 35**). At 24, 48 and 72 h.

Then, we also performed immunofluorescence experiments on all the three cell lines with the HA-PBA-QC-Cy3 (**Figure 37**) at the same time points to be sure that the presence of the compound would not change the behaviour of the HA alone. Finally, we checked the internalization of the HA-PBA-64-Cy3 at the same time points but only in A375 and A549 because these two cell lines seem to respond more to the internalization of the HA. We used for all these experiments always the same concentration of 50 μ M. As we can see from the images 7 and 8, we confirmed that after 72h the prodrug is completely inside the cells and that the A375 cell lines are the most sensitive cells to the internalization and to the co-treatment. Probably,

because the overexpressed the CD44 receptors. Regarding the internalization of our prodrug, HA-PBA-64-Cy3, we noticed (**Figure 38**) that the localization seems to be perinuclear. It has to be specified that the fluorescence that we could see derived from the hyaluronic acid, not from the compound. For this reason, it is possible that the hyaluronic acid can be accumulated around the nucleus, while the released compound, thanks to the acid environment, can be internalized inside the nucleus.

In conclusion, thanks to the prodrug investigated here, we have improved the bioavailability of our inhibitors and we also improved their solubility. At the same time, we were able to preserve the inhibitory activity of compound 64. Further studies are needed to better characterize this prodrug; the next steps will be a further investigation through additional *in vivo* studies.

4.3 References

- (1) Marinho, A.; Nunes, C.; Reis, S. Hyaluronic Acid: A Key Ingredient in the Therapy of Inflammation. *Biomolecules* **2021**, *11* (10). DOI: 10.3390/biom11101518
- (2) Bastow, E. R.; Byers, S.; Golub, S. B.; Clarkin, C. E.; Pitsillides, A. A.; Fosang, A. J. Hyaluronan synthesis and degradation in cartilage and bone. *Cell Mol Life Sci* **2008**, *65* (3), 395-413. DOI: 10.1007/s00018-007-7360-z
- (3) Fallacara, A.; Baldini, E.; Manfredini, S.; Vertuani, S. Hyaluronic Acid in the Third Millennium. *Polymers (Basel)* **2018**, *10* (7). DOI: 10.3390/polym10070701
- (4) Rios de la Rosa, J. M.; Tirella, A.; Gennari, A.; Stratford, I. J.; Tirelli, N. The CD44-Mediated Uptake of Hyaluronic Acid-Based Carriers in Macrophages. *Adv Healthc Mater* **2017**, *6* (4). DOI: 10.1002/adhm.201601012
- (5) Gennari, A.; Gujral, C.; Hohn, E.; Lallana, E.; Cellesi, F.; Tirelli, N. Revisiting Boronate/Diol Complexation as a Double Stimulus-Responsive Bioconjugation. *Bioconjug Chem* **2017**, *28* (5), 1391-1402. DOI: 10.1021/acs.bioconjchem.7b00080
- (6) Quagliariello, V.; Gennari, A.; Jain, S. A.; Rosso, F.; Iaffaioli, R. V.; Barbarisi, A.; Barbarisi, M.; Tirelli, N. Double-responsive hyaluronic acid-based prodrugs for efficient tumour targeting. *Mater Sci Eng C Mater Biol Appl* **2021**, *131*, 112475. DOI: 10.1016/j.msec.2021.112475
- (7) Orrego-Lagaron, N.; Martinez-Huelamo, M.; Quifer-Rada, P.; Lamuela-Raventos, R. M.; Escribano-Ferrer, E. Absorption and disposition of naringenin and quercetin after simultaneous administration via intestinal perfusion in mice. *Food Funct* **2016**, *7* (9), 3880-3889. DOI: 10.1039/c6fo00633g

Chapter 5: Design, synthesis, dynamic docking,
biochemical characterization, and in vivo
pharmacokinetics studies of novel topoisomerase II
poisons with promising antiproliferative activity

Jose M. Arencibia,#¹ Nicoletta Brindani,#¹ Sebastian Franco-Ulloa,¹ **Michela Nigro**,¹ Jissy Akkarapattiakal Kuriappan,¹ Giuliana Ottonello,² Sine Mandrup Bertozzi,² Maria Summa,² Stefania Giroto,¹ Rosalia Bertorelli,² Andrea Armirotti,² Marco De Vivo*.¹

#Equally contributed

This manuscript has been published on the Journal of Medicinal Chemistry. 2020 Apr 9;63(7):3508-3521. doi: 10.1021/acs.jmedchem.9b01760. Epub 2020 Mar 30. PMID: 32196342; PMCID: PMC7997578.

Supporting information:

https://pubs.acs.org/doi/suppl/10.1021/acs.jmedchem.9b01760/suppl_file/jm9b01760_si_001.pdf

5.1 Introduction

Human topoisomerase II (topoII) enzymes are a validated target to treat cancer because of their role in modifying the topology of entangled DNA strands during vital cellular processes like replication and transcription.¹⁻⁴ Several topoII anticancer inhibitors are clinically available. One example is etoposide, which is used to treat a variety of cancers, including leukemia and ovarian cancer.⁵⁻⁹ However, drug resistance and the possibility of severe side effects of topoII-targeting drugs means that researchers continue to seek novel safer topoII inhibitors.^{6, 10-17}

Small molecules targeting topoII are classified as either topoII poisons or topoII catalytic inhibitors.¹⁸⁻²⁰ These two classes of topoII blockers differ in their mode of action. TopoII poisons act by trapping the covalent topoII/DNA cleavage complex, which is formed during the catalytic cycle required for DNA topology modification. A covalent and stable topoII/DNA cleavage complex eventually leads to the accumulation of double-strand breaks, causing cell death.^{1, 12-13, 18, 21-23} The chemotherapy drug etoposide (**Figure 39**) acts via this mechanism, although its pharmacological action can lead to severe side-effects.²⁴⁻²⁶ Additional examples of anticancer drugs²⁷⁻³⁰ that act as a topoII poison are doxorubicin, mitoxantrone, salvicine, and teniposide. These drugs are frontline therapies for a wide range of solid and hematological malignancies.³¹⁻³³

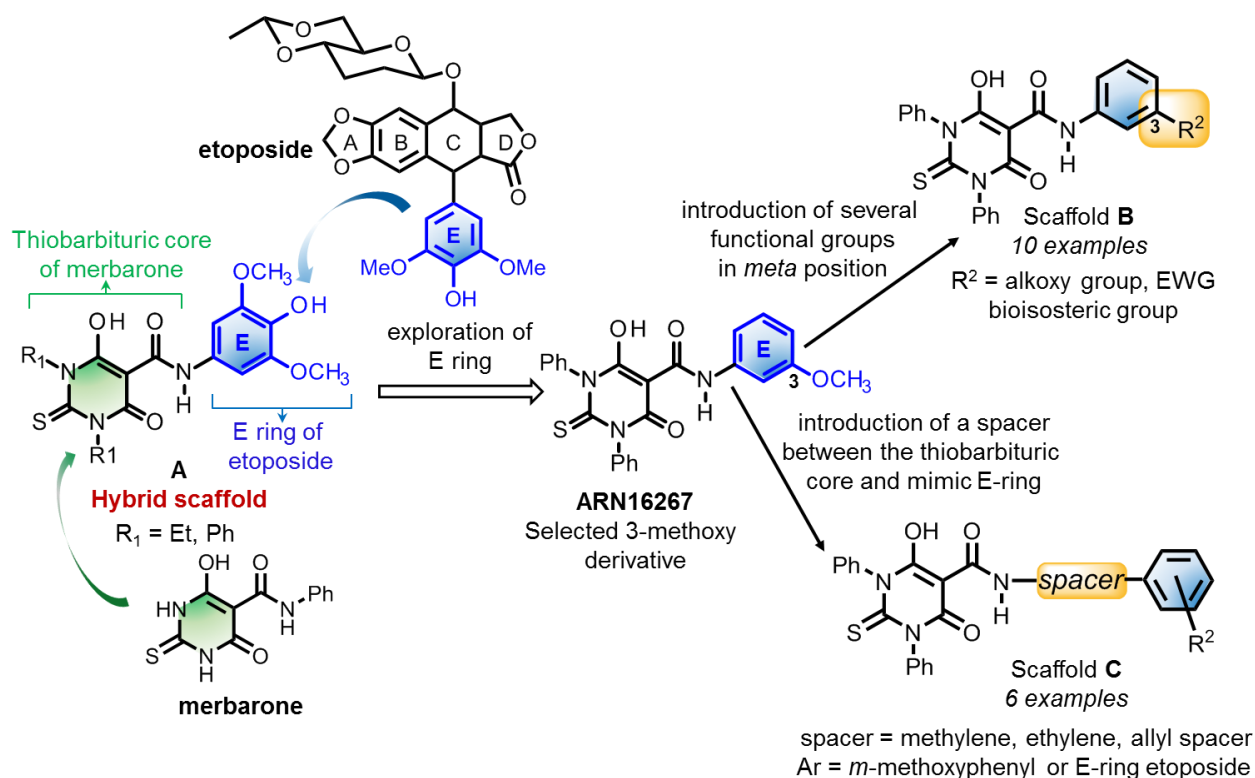


Figure 39 Hybrid topoisomerase II poison with the scaffold A (left) was explored to discover ARN16267 as a potent topoisomerase II blocker.³⁹ Here, the hybrid scaffold A was expanded to generate structures of type B (right above), with several functional groups introduced in the meta position. Structures of type C (right below) were generated by introducing a spacer between the thiobarbituric core and the mimic E-ring.

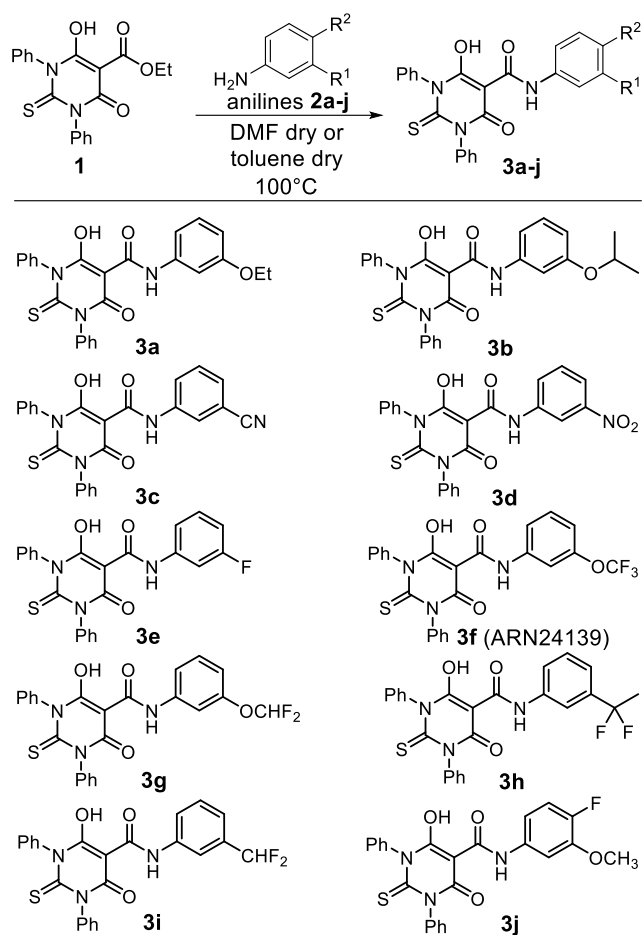
Topoisomerase II catalytic inhibitors act differently than poisons and do not generate an accumulation of topoisomerase II/DNA cleavage complex. Instead, topoisomerase II catalytic inhibitors act, for example, by inhibiting DNA binding and/or by blocking the ATP-binding site in topoisomerase II, thus preventing ATP hydrolysis, which is needed for topoisomerase II function.¹⁹ One notable example is merbarone, which was one of the first and most promising topoisomerase II inhibitors (**Figure 39**).³⁴⁻³⁷ Merbarone is a thiobarbituric derivative (6-hydroxy-4-oxo-*N*-phenyl-2-thioxo-1H-pyrimidine-5-carboxamide) that blocks topoisomerase II catalysis and inhibits proliferation of several cancer cell lines.³⁵ Notably, merbarone underwent clinical trials as a chemotherapy drug.^{34, 36} These trials failed because merbarone displayed nephrotoxicity issues and did not generate the expected efficacy.³⁸

Recently, we used a pharmacophore hybridization strategy to realise a first set of new topoisomerase II poisons.³⁹ They were rationally designed by combining key pharmacophoric elements of etoposide and merbarone to generate a new etoposide-merbarone hybrid active scaffold.³⁹⁻⁴⁰ In particular, we

designed, synthesized, and characterized a first set of compounds that feature the thiobarbituric core of merbarone linked via an amide bond to the E-ring of etoposide (type A structure, Figure 1). This design generated new *N,N'*-diphenyl derivatives that potently block human topoII.³⁹ In addition, our SAR studies clarified the effect of ethyl and phenyl substitutions at each nitrogen of the thiourea moiety, as well as the influence of the number and/or position of hydroxyl and methoxy substituents on the mimic E-ring.³⁹ Importantly, we identified compound ARN16267 (IC₅₀ = 30 ± 6 μM, structure in Figure 36, which was originally named compound **3** in Ref³⁹), which is a more potent topoII blocker than the template compounds, i.e. etoposide (IC₅₀ = 120 ± 10 μM) and merbarone (IC₅₀ = 120 ± 12 μM).³⁹ Intriguingly, we found that ARN16267 was the most efficient of this new chemical class in generating accumulation of topoII/DNA cleavage complex. This suggests that ARN16267 may act as a topoII poison, although this mechanism was less marked than that of etoposide.³⁹

These results prompted us to investigate the SAR of these new hybrid topoII blockers. Here we present an additional sixteen derivatives that expand the initial panel of merbarone-etoposide hybrid molecules.³⁹ As described in Figure 36, we used ARN16267 as our best starting point for further derivatization of its core scaffold, generating scaffolds of type B and C (**Figure 36**). We thus identified a new hybrid derivative (**3f**, ARN24139 – see Scheme 1) with improved human topoII inhibitory activity (7.3 ± 1.5 μM). In addition, our results confirm that this new class of hybrid compounds acts as topoII poisons, generating accumulation of topoII/DNA cleavage complex. Our dynamic docking simulations support binding of **3f** at the cleavage complex. Additionally, **3f** showed high kinetic solubility and metabolic stability, as well as a promising antiproliferative activity in the low μM range against DU145, HeLa, and A549 cancer cell lines. Finally, we found **3f** to have a good pharmacokinetic profile in vivo. Thus, **3f** can be added to the pipeline of compounds that are active against topoII with promising anticancer activity.⁴¹⁻⁴³

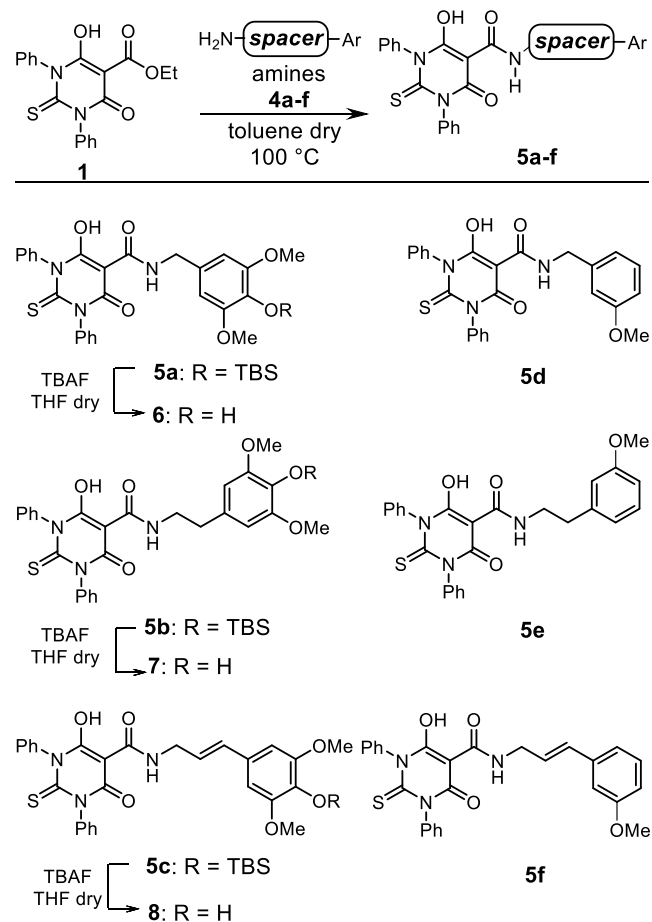
5.2 Results and discussion



Scheme 1. Synthesis of compounds **3a-j**. Our lead compound **3f** (ARN24139) is in the right column.

Exploring the structure of the new hybrid scaffold. For our new set of hybrid analogues and based on our previous work and results,³⁹ we initially expanded our SAR studies by exploring the effect of diverse functional groups in the *meta* position on the mimic E-ring (**3a-j**, Scheme 1). First, we synthesized two new compounds with an ethoxy (**3a**) or isopropoxy (**3b**) group in the *meta* position on the mimic E-ring. We previously found that replacing the methoxy in ARN16267 with a hydroxyl group significantly decreased topoII inhibitory activity.³⁹ We thus substituted the original methoxy in ARN16267 with a cyano (**3c**), nitro (**3d**), or fluoro (**3e**) to modulate the electron density of the mimic E-ring. Additionally, we investigated the bioisosteric replacement of the methoxy group of ARN16267, and generated three additional new hybrid compounds, each

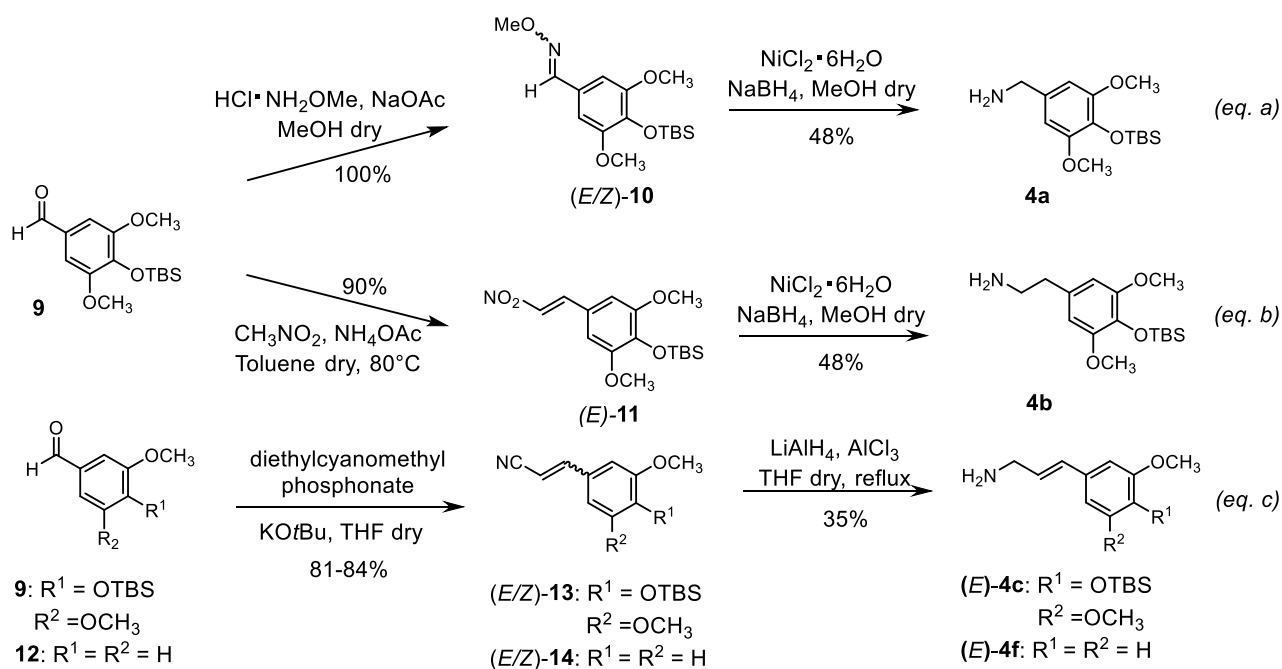
bearing a trifluoromethoxy (**3f**), difluoromethoxy (**3g**), or difluoroethoxy (**3h**) group in the *meta* position on the mimic E-ring. Similarly, we synthesized a derivative with the difluoromethyl group in the *meta* position on the mimic E-aromatic ring (**3h**). Finally, given that a fluorine proximal to a methoxy can influence the overall electronic behavior of the aromatic ring,⁴⁴ we inserted a fluoro in the *para* position of ARN16267 to obtain **3j** (Scheme 1).



Scheme 2. Synthesis of compounds **6-8**, **5d-f**.

In the crystal structure of the ternary topoII/DNA/Etoposide complex, the drug molecule is stabilized by interactions with Asp463 and Arg487.²⁵ To favour the formation of these interactions for our hybrid compounds, we introduced a flexible spacer between the thiobarbituric core and the mimic E-ring (compounds **6-8**, and **5d-f**, Scheme 2).^{25, 45-47} To test this hypothesis, we generated an additional set of six compounds with a link-mediated increased flexibility (**6-8**, **5d-f**).

Chemistry. The sixteen new derivatives were synthesized through regular amidation of ester **1**³⁹ with amines **2a-j** and **4a-f** using either DMF or toluene dry as a solvent, with yields that ranged from 34 to 72% (Schemes 1 and 2). As previously described,⁴⁸ in the presence of amine **2j** substituted with a fluorine in *para* and a methoxy group in *meta* position, the formylation side reaction performed by DMF was preferred over the alternative and desired reaction with ester **1**.⁴⁸⁻⁴⁹ To circumvent this problem, the reaction was conducted in toluene dry at 100 °C, obtaining **3j** with good yield (72%). The same strategy was used to synthesize **6-8** and **5d-f** (Scheme 2), where different hydrocarbon chains were introduced between the amide and the aromatic ring. In particular, amines **4a-c** and **4f** (used to prepare **6-8** and **5f**) were synthesized in two steps starting from silyl protected syringaldehyde **9** and *m*-anisaldehyde **12**, respectively (Scheme 3). Compound **4a** is a benzylamine featuring an aromatic ring with the same functionalization of the E ring of etoposide. This was obtained through the quantitative conversion of aldehyde **9** in the related *O*-methyl oxime **10**, which was reduced into the desired amine **4a** with NiCl₂·6H₂O and NaBH₄ with 48% yield (Scheme 3, eq a).



Scheme 3. Synthesis of amines **4a-c** and **4f**.

The Henry reaction between protected syringaldehyde **9**⁵⁰ and nitromethane in the presence of ammonium acetate gave (*E*)-nitrostyrene **11**, with an excellent 90% yield. Compound **11** was then

completely reduced with NiCl₂ and NaBH₄ into phenethenamine **4b** with 48% yield (Scheme 3, eq b). The Horner-Wadsworth-Emmons (HWE) reaction between aldehyde **9**⁵⁰ and diethylcyanomethylphosphonate in the presence of potassium *tert*-butoxide in THF gave a 1:0.12 mixture of (*E*)- and (*Z*)- acrylonitrile **13**, in 84% yield. The chemoselective nitrile reduction with LiAlH₄ and aluminium trichloride afforded (*E*)-allylamine **4c** after chromatography purification, with an acceptable 35% yield (Scheme 3, eq c). The same sequential transformations of olefination and reduction allowed the introduction of an allylic spacer for amine **4f**, starting from *m*-anisaldehyde **12**, with a 22% overall yield (Scheme 3, eq c). In this case, too, the (*E*)-isomer was obtained pure after chromatography purification. Thus, the key amidation reaction between silylated amines **4a-c** with ester **1** in toluene dry at 100 °C generated amides **5a-c**, which were deprotected using TBAF. This yielded our final targets **6-8** with a free hydroxylic group in para position, with 22-38% yield after two steps (Scheme 2). Similarly, *m*-methoxy amine counterparts **5d-f** were obtained with 38-62% yield using related amines **4d-f** (Scheme 2).

TopoII inhibitory activity of the novel hybrid compounds. We measured the inhibitor activity of our new compounds against human topoII, using a topoII decatenation assay (Table 1). Notably, in our assay, etoposide returned an IC₅₀ of 47.5 ± 2.2 μM, which agrees with the activity reported by the manufacturer of the decatenation assay kit.⁵¹ Additionally, merbarone showed an IC₅₀ of 26.0 ± 4.7 μM, which is in line with that reported previously (IC₅₀ = 40 μM) using a plasmid relaxation assay.³⁵

Interestingly, greater bulkiness of the alkyl chain on the oxygen in the *meta* position, as in **3a** and **3b**, improved the potency of these compounds, as compared to ARN16267, which has an IC₅₀ of 16.1 ± 2.4 μM, as measured in the decatenation assay used in this study. In fact, **3b**, with the bulkier substituent, had an IC₅₀ of 9.7 ± 2.6 μM, while the ethoxy substitution in **3a** returned an IC₅₀ of 12.4 ± 3.7 μM (Table 1). We then found that the electron-withdrawing nitro group, in the *meta* position on the mimic E-ring in **3d**, returned an IC₅₀ of 10.2 ± 4.6 μM. Other electron-withdrawing groups such as the cyano (**3c**) and the fluoro (**3e**) returned a comparable activity to that of ARN16267 (Table1). This is in line with our previous demonstration of the unfavourability of an electron-donating group, such as a hydroxyl substitution, at this position.³⁹ Similarly, introducing a fluorine proximal to the methoxy group (**3j**) was detrimental for activity, with an IC₅₀ of 22.5 ± 5.8 μM.

We then investigated the bioisosteric replacement of the methoxy group in *meta* position on the mimic E-ring. Interestingly, all the bioisosteric analogues **3f-h** displayed a better activity than ARN16267: introducing a trifluoromethoxy group in **3f** returned a 2-fold increased activity with an IC₅₀ of 7.3 ± 1.5 μM, while difluoroethyl in **3h** returned an IC₅₀ of 9.2 ± 0.2 μM, and the difluoromethoxy group in **3g** returned an IC₅₀ = 11.4 ± 2.4 μM. Notably, the difluoromethyl analogue **3i** also had this improved activity, which confirms that the additional interactions provided by fluorinated groups can compensate the loss of the oxy-moiety in *meta* position of the aromatic (mimic) E-ring.

Table 1. Inhibitory and antiproliferative activities of sixteen new derivatives **3a-j**, **6-9**, **5d-f**. Antiproliferative activity in cancer cells of etoposide and merbarone was measured in Ref³⁹.

Entry	IC50 (μM)	DU145 (μM)	HeLa (μM)	A549 (μM)
etoposide	47.5 ± 2.2	1.0 ± 0.8	2.4 ± 0.9	1.3 ± 0.1
merbarone	26.0 ± 4.7	18.9 ± 2.0	62.3 ± 6.4	40.0 ± 2.7
ARN16267	16.1 ± 2.4	7.6 ± 0.8	5.5 ± 1.3	4.7 ± 0.3
3a	12.4 ± 3.7	5.7 ± 1.4	5.2 ± 0.9	3.6 ± 0.4
3b	9.7 ± 2.6	5.5 ± 0.1	4.2 ± 0.3	3.0 ± 0.3
3c	14.4 ± 3.8	14.4 ± 4.1	12.4 ± 1.3	16.9 ± 0.3
3d	10.2 ± 4.6	6.9 ± 0.3	11.9 ± 2.6	18.5 ± 3.1
3e	15.1 ± 4.2	5.6 ± 0.6	5.6 ± 0.5	4.1 ± 0.2
3f	7.3 ± 1.5	4.7 ± 0.1	3.8 ± 0.3	3.1 ± 0.1
3g	10.2 ± 1.9	6.5 ± 0.9	6.4 ± 0.5	4.8 ± 0.4
3h	9.2 ± 0.2	2.7 ± 2.5	2.5 ± 0.4	3.0 ± 1.1
3i	11.4 ± 2.4	7.8 ± 0.2	4.9 ± 1.4	4.4 ± 0.2
3j	22.5 ± 5.8	7.8 ± 0.6	5.3 ± 1.0	4.6 ± 0.1
5d	22.5 ± 7.2	7.7 ± 0.1	6.1 ± 1.1	4.8 ± 0.4
5e	15.8 ± 3.4	3.3 ± 2.5	3.4 ± 0.1	3.2 ± 1.2
5f	8.3 ± 2.3	5.0 ± 2.9	5.5 ± 1.1	4.4 ± 0.6
6	107.8 ± 10.1	8.8 ± 0.3	19.1 ± 4.8	14.7 ± 0.3
7	74.4 ± 13.6	13.9 ± 7.6	13.3 ± 1.7	14.8 ± 1.4
8	28.0 ± 4.4	9.8 ± 2.1	9.3 ± 0.2	9.5 ± 1.5

After evaluating the inhibitory activity of this first subset of derivatives, we assessed the activity of the hybrid molecules with a flexible spacer connecting the thiobarbituric core and the mimic E-ring (**5d-f** and **6-8**, Scheme 2). We started evaluating the activity of the compound **6**, where we inserted a methylene substitution and that contains the exact E-ring of etoposide. This first modification reduced the potency ($IC_{50} = 107.8 \pm 10.1 \mu\text{M}$), with a 7-fold drop in activity compared to ARN16267 (Table 1). Conversely, inserting the same substitution in **5d** only decreased 1.4-fold the inhibitory activity ($IC_{50} = 22.5 \pm 7.2 \mu\text{M}$) compared to the parent compound ARN16267 (Table 1). This result confirms that a methoxy group, alone, in the mimic E-ring increases the potency of this hybrid scaffold, as also observed previously.³⁹ Increasing the spacer length was also beneficial, improving the IC_{50} from over 100 μM for **6** to 74.4 ± 13.6 and $28.0 \pm 4.4 \mu\text{M}$ for **7** and **8**, respectively. This positive trend in potency could be due to a more balanced structure where the flexibility introduced by having the hydrocarbon chain (i.e., the spacer) is compensated by the introduced rigidity of the C=C double bond embedded in the allylic system in **8**. These compounds were active when the 3-methoxy was retained as the only substituent, with IC_{50} values of $22.5 \pm 7.2 \mu\text{M}$ for **5d**, $15.8 \pm 3.4 \mu\text{M}$ for **5e**, and $8.3 \pm 2.3 \mu\text{M}$ for **5f**. Notably, **5f** showed a 2-fold increase in IC_{50} compared to its template ARN16267 (which has no spacer). These data thus suggest that a flexible substituent connecting the thiobarbituric core and the mimic E-ring in our hybrid scaffold may facilitate the orientation of our molecules inside the active site of topoII, increasing their inhibitory potency (see docking results, below).

Antiproliferative activity against cultured human cancer cells, metabolic stability, chemical solubility, and topoII poisoning. The antiproliferative activity of all compounds was evaluated in i) DU145, an androgen-independent prostate cancer cell line; ii) HeLa, a cervical cancer cell line; and iii) A549, a lung adenocarcinoma cell line (Table 1). Notably, all new compounds showed good antiproliferative activity with IC_{50} values in the low μM range. Among the most active compounds in inhibiting topoII activity, **3f**, **3h**, and **5e** showed cytotoxicity with IC_{50} values lower than 5 μM (Table 1).

After this initial evaluation of the new set of hybrid compounds for their inhibitory activity against topoII *in vitro* and for their biological cytotoxicity, we selected **3b**, **3f-i**, and **5e** for further evaluation. We assessed their metabolic stability using mouse serum and mouse liver microsomes,

and we assessed their kinetic solubility in neutral buffer. These compounds had excellent plasma and microsomal stability with half-time values greater than 120 and 60 min, respectively. Additionally, **3f**, **3g**, and **3i** displayed excellent solubility in aqueous buffer (pH 7.4), with values greater than 200 μM (Table 2).

Table 9 Kinetic solubility of **3b**, **3f-h**, **5e**.

Entry	Kinetic solubility (μM)
ARN16267	236
3b	34
3f	224
3g	208
3h	6
3i	238
5e	122

In view of these results, **3f**, **3g**, and **3i** were tested in a cleavage complex formation assay to further ascertain their mode of action as topoII poisons. As shown in Figure 37, all these hybrid molecules were confirmed to be poisons and thus able to generate an accumulation of topoII/DNA cleavage complex. In particular, **3f** had the greatest poison efficacy, being about 1.6-fold better than the template ARN16267, at 200 μM concentration. Given **3f**'s promising *in vitro* activity as a topoII poison and its overall drug-like profile, we examined its binding mode with topoII α its *in vivo* pharmacokinetic profile in mice.

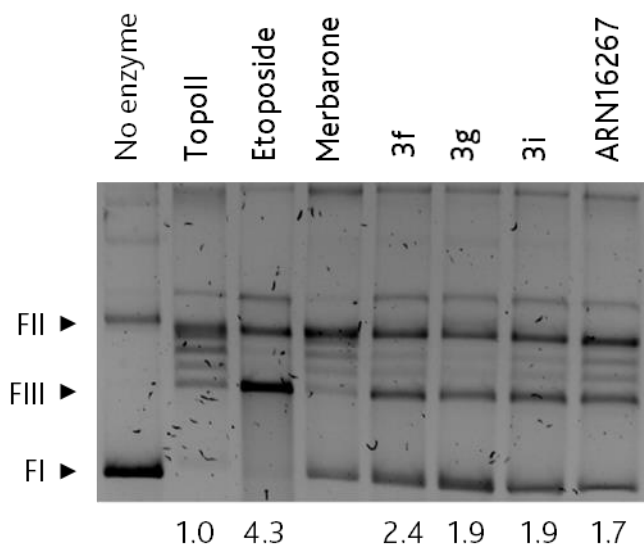


Figure 40 Poison activity of the hybrid compounds. Agarose gel electrophoresis of plasmid DNA incubated in the absence (no enzyme) or presence of 1 U topoII containing either 1% DMSO as control vehicle (no enzyme and topoII lanes) or 200 μ M of the compound. Labels are shown above each lane. Numbers in the bottom correspond to the normalized intensity of the linear form (FIII). Plasmid forms are indicated by the arrow-points on the right: supercoiled (FI), nicked (FII), and linear (FIII).

Docking and molecular simulations of 3f bound to the target We used docking and atomistic force-field-based molecular dynamics (MD) simulations to model **3f** bound to the topoII/DNA cleavage site.⁵²⁻⁵⁵ First, the crystal structure of the topoII α isoform (PDBID: 5GWK) was used for the docking studies.^{25, 56} As seen in Figure S2, when the compound was first docked into the cleavage site, the mimic E-ring slightly shifted relative to the position of the E-ring of etoposide in the crystal.²⁵ Our calculations revealed several key contacts between the ligand and vicinal residues that confer the system a stable, inhibited conformation, thus endorsing the compound's action as a topoII poison. The protein aids the ligand's anchoring within the pocket by a cation- π link formed between Arg487 and the E-ring. The neighbouring DNA bases also contribute to the stabilization of the complex. Specifically, the G₊₅ and C₋₁ bases display π -stacking interactions with the heterocycle inherited from merbarone (**Figure 41**). Similarly, the T₊₁ and C₋₁ bases align to the phenyl substituents at the thiobarbituric core in a perpendicular fashion resembling a T-shaped π -interaction. Finally, an H-bond between T₊₁ and the ligand's N-H moiety was also identified.

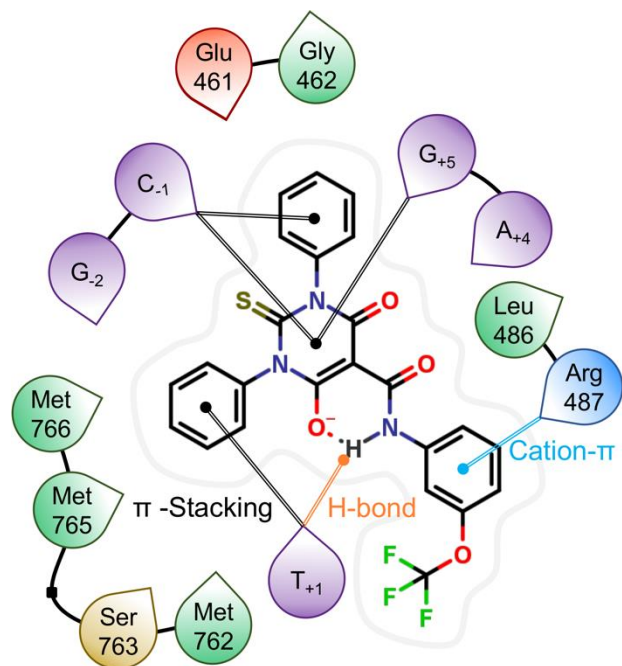


Figure 41 Interaction diagram of the lead compound **3f** bound to the DNA/topoII complex as derived from docking calculations. The computational study flags a cation- π interaction with Arg487, π -stacking with various DNA bases, and an H-bond with T_{+1} , as the main drivers for ligand binding.

Once the general interaction pattern was established, we proceeded to perform equilibrium molecular dynamics (MD) simulations (~200 ns) to analyze the evolution and stability of the ternary topoII/DNA/**3f** model system.^{14, 57-58} For this, distances representative of the interactions described above were tracked (**Figure 39A**). The simulations corroborate the role of Arg487 in stabilizing the drug at the cleaved site. Indeed, we monitored the distance between the carbon atom of the guanidinium group and the centroid of the E-ring and found that it remains under 6 Å for 99% of the simulation (**Figure 39B**).

The H-bonds between the ligand and the cleaved complex were also examined. It is worth to note that the -NH linker present in compound **3f** adds to the rigidity of the molecule. In fact, it enables the formation of a six-membered intramolecular pseudo-cycle. The cycled configuration was present for ~79% of the overall simulation time. Nonetheless, the -NH group was found to intermittently invert in order to form an H-bond with the oxygen from the deoxyribose ring of G_{+5} . The latter bond persisted for a total of ~17% of the production run, thus making a complementary interaction for the stabilization of the cleaved complex holo-form.

Similarly, we examined the staggered π -stacking formed between the thiobarbituric cycle and the G_{+5} and C_{-1} bases (**Figure 39B**). Here, the distance to the G_{+5} base remains consistently smaller at a stable value of 3.6 ± 0.3 Å, whereas the distance to C_{-1} extends to 4.4 ± 0.4 Å. A similar pattern is observed between the phenyl groups and the vicinal base pairs. In summary, the binding mode most frequently visited is stable in our MD simulations too, with key cation- π , H-bonds, and stacking interactions formed with the surrounding DNA/protein complex.

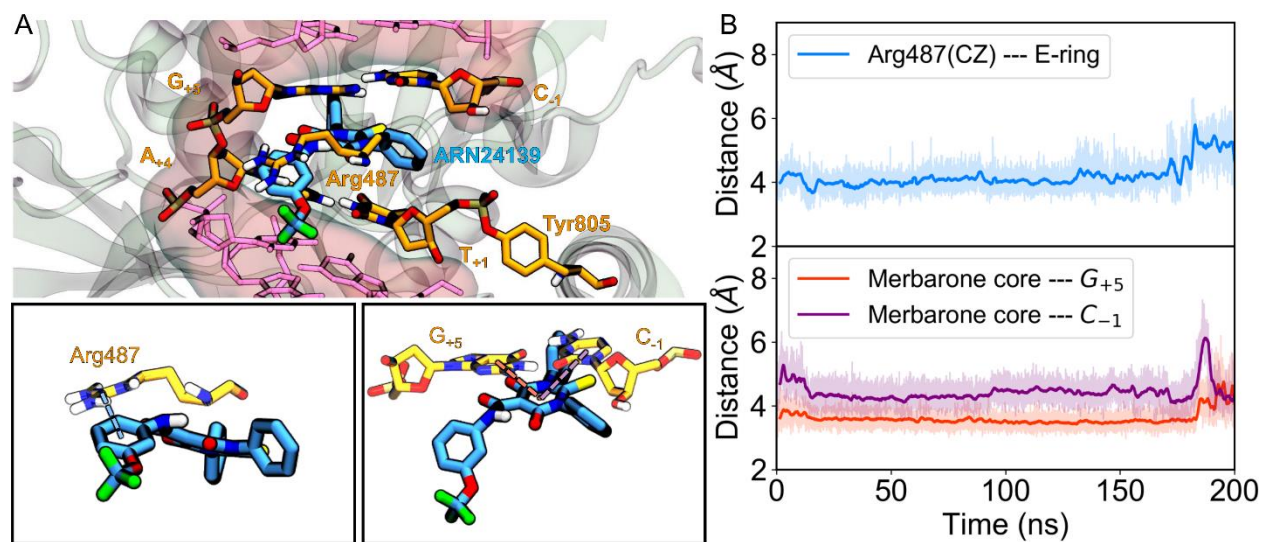


Figure 42 Dynamic description for the binding of **3f** to the in the DNA cleavage/religation active site of topoiII α (PDBID: 5GWK). *A*) The lead compound in its binding mode. The residues directly interacting with the drug are shown in orange carbons, the rest of the DNA in pink, and the protein in white. The insets zoom into the distances of interest, particularly, those of Arg487 with the E-ring, and the merbarone core with the G_{+5} and C_{-1} bases. *B*) Evolution of the distances of interest over time.

***In vivo* pharmacokinetics.** Finally, based on the overall results and drug-like profile, **3f** was selected as our lead for *in vivo* pharmacokinetics studies, as a preparatory characterization for future analyses of *in vivo* drug efficacy in animal models of cancer. We tested two different routes of administration: i) intravenous (I.V.) injection at a concentration of 3 mg/Kg (n=3 for each time point) and oral (P.O.) treatment at a dose of 10 mg/Kg (n=3 for each time point). Despite the relatively low thermodynamic solubility of **3f** (30 μ M in PBS), the compound reached the target concentration in the formulation used for the *in vivo* experiments. The mean plasma concentration versus time profile of **3f** is shown in Figure 40, and the corresponding pharmacokinetic parameters

are summarized in the inserted table. During the PK studies, via either I.V. or P.O. administration, **3f** was well tolerated by all the animals, and no treatment-related clinical signs were observed.

The peak plasma **3f** concentration for I.V. was observed at the earliest time point (5min after administration), and the concentration of **3f** in plasma was above the lower limit of quantification throughout the sampling period. The I.V. profile of **3f** presents a very fast distribution phase with a C_{max} of 7366 ng/mL, followed by a slower exposure phase. The compound was still detectable after 2 h at a concentration of 551 ng/mL, with Cl_p value of 0.004 L/min/Kg. After oral administration (10 mg/Kg), plasma concentration of about 400 ng/mL was reached relatively fast (1 h) and it was stably maintained for at least 6 h. The maximum concentration was achieved at approximately 2 h post oral administration, 428 ng/mL.

These data indicate that the compound **3f** is well tolerated. Indeed, the animal behavior and the obtained PK profiles indicate that the dose of **3f** can be increased in view of repeated experiments. This would be extremely important, given the observed very high protein binding of this compound (>99% in both mouse and human plasma) that plays against the possibility to hit the target at the active concentration measured *in vitro*.

Mouse PK Profile ARN24139

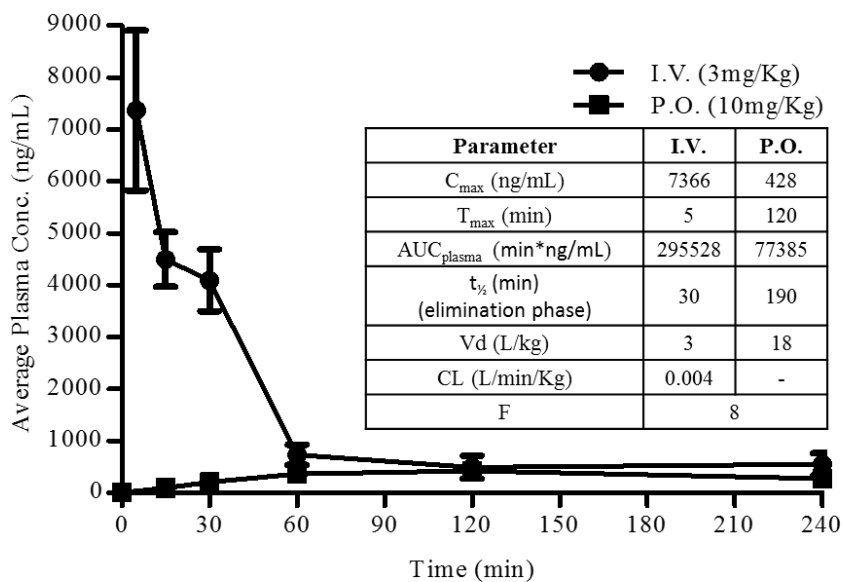


Figure 43 Mouse PK profiles of **3f** following intravenous (I.V.) and oral administration (P.O.) at 3 and 10mg/Kg, respectively. The observed and calculated PK parameters following intravenous (I.V.) and oral administration (P.O.). The bioavailability F was calculated to 8% based on the AUC (area under curve) from $t=0$ to 240 min.

5.3 Conclusions

Based on our previous results on a novel hybrid scaffold with structural elements of merbarone and etoposide,³⁹ we have here reported the design, synthesis, and extensive experimental-computational characterization of new hybrid molecules that act as topoII poisons. The resulting SAR elucidated the key structural features that enhanced the potency and antiproliferative activity of our new etoposide-merbarone hybrid compounds. These new compounds were often equipotent and sometimes more potent relative to the reference compounds merbarone and etoposide. Inhibitory activity was improved by introducing a bulkier group in *meta* position of the mimic E-ring (**3a** and **3b**, Scheme 1). Incorporating electron-withdrawing groups preserved or slightly improved the inhibitory activity (**3c-e**, Scheme 1), while the bioisosteric substitution with fluorine-embedding groups (**3f-h**, Scheme 1) was highly favorable. Furthermore, in the structural design of these new bioactive hybrid molecules, the combined functionalization of both the aromatic E-ring and the hydrocarbon spacer was essential to fine-tuning the drug-target interactions, as proven by the activity of more flexible hybrid topoII poisons (**6-8**, and **5d-f**, Scheme 2). Taken together, the inhibitory activity and extensive analyses of the drug-likeness profile indicate the novel derivative **3f** (ARN24139) as the most drug-like topoII poison of this novel chemical series. This lead compound has also shown promising antiproliferative activity against cancer cell lines and a favorable pharmacokinetic profile, which are optimal features for future *in vivo* efficacy studies.

5.4 Experimental sections

Chemistry

General considerations. All the commercially available reagents and solvents were used as purchased from vendors without further purification. Dry solvents were purchased from Sigma-Aldrich. Automated column chromatography purifications were done using a Teledyne ISCO apparatus (CombiFlash Rf) with pre-packed silica gel columns of different sizes (from 4 g up to 24 g) and mixtures of increasing polarity of cyclohexane and ethyl acetate (EtOAc) or dichloromethane (DCM) and methanol (MeOH). NMR data were collected on 400 MHz or 600 MHz (^1H) and 100 MHz or 150 MHz (^{13}C). Spectra were acquired at 300 K, using deuterated dimethylsulfoxide ($\text{DMSO}-d_6$) or deuterated chloroform (CDCl_3) as solvents. For ^1H -NMR, data are reported as follows: chemical shift, multiplicity (s= singlet, d= doublet, dd= double of doublets, t= triplet, q= quartet, m= multiplet), coupling constants (Hz), and integration. UPLC/MS analyses were run on a Waters ACQUITY UPLC/MS instrument consisting of an SQD (Single Quadrupole Detector). The analyses were performed on an ACQUITY UPLC BEH C_{18} column (50x2.1mmID, particle size 1.7 μm) with a VanGuard BEH C_{18} pre-column (5x2.1mmID, particle size 1.7 μm) (LogD>1). The mobile phase was 10mM NH_4OAc in H_2O at pH 5 adjusted with AcOH (A) and 10mM NH_4OAc in MeCN- H_2O (95:5) at pH 5 (B). Electrospray ionization in positive and negative mode was applied in the mass scan range 100-500Da. Depending on the analysis method used, a different gradient increasing the proportion of mobile phase B was applied. For analysis method A, the mobile-phase B proportion increased from 5% to 95% in 3 min. For analysis method B, the mobile-phase B proportion increased from 50% to 100% in 3 min. High-resolution mass spectrometry (HRMS) was carried out on a Waters Synapt G2 Quadrupole-Tof Instrument equipped with an ESI ion source. The analyses were run on an ACQUITY UPLC BEH C_{18} column (50x2.1mmID, particle size 1.7 μm), using H_2O + 0.1% formic acid (A) and MeCN + 0.1% formic acid as mobile phase. All final compounds displayed $\geq 95\%$ purity as determined by NMR and UPLC/MS analysis.

3-(1,1-Difluoroethyl)aniline (2h). A solution of 3-nitroacetophenone (165 mg, 1 mmol) in CH_2Cl_2 dry (4 ml) was treated with [Bis(2-methoxyethyl)amino]sulphur trifluoride 50wt% solution in toluene (2.5 mmol) at room temperature under argon. EtOH (24 μL , 0.4 mmol) was

added and the reaction mixture was stirred at room temperature for 48h. After which time, the solution was poured into NaHCO₃ saturated solution, extracted with CH₂Cl₂ (2x5mL). Combined organic layers were dried with Na₂SO₄, filtered and concentrated under vacuum. Purification by silica gel flash chromatography (elution by gradient from 100 to 95/5 Cyclohexane/EtOAc) afforded pure 1-(1,1-Difluoroethyl)-3-nitrobenzene (73 mg, 39% yield). ¹H NMR (400MHz, CDCl₃): δ 8.38 (bs, 1H), 8.31 (d, *J* = 7.7 Hz, 1H), 7.85 (d, *J* = 7.7 Hz, 1H), 7.64 (dd, *J* = 7.9, 7.9 Hz, 1H), 1.98 (t, ³*J*_{H-F} = 18.2 Hz, 3H). Then tin chloride dihydrate (440 mg, 1.95 mmol) was added to a solution of compound 1-(1,1-Difluoroethyl)-3-nitrobenzene (73 mg, 0.39 mmol) in 3 mL of ethanol. The reaction mixture was refluxed for one hour. The mixture was slowly poured on cooled water. The pH was adjusted to 7 by addition of an aqueous 5N solution of sodium hydroxide, then adjusted to 9 by addition of an aqueous NaHCO₃ saturated solution. The product was extracted with EtOAc (3x5mL). The organic phases were combined, dried over Na₂SO₄, filtered and concentrated under vacuum. Purification by silica gel flash chromatography (elution by gradient from 100 to 80/20 Cyclohexane/EtOAc) afforded pure compound **3h** (58 mg, 95% yield). UPLC/MS: Rt = 1.75 min (method A), MS (ESI) m/z: 158.0 [M+H]⁺, C₈H₁₀F₂N [M+H]⁺ calculated: 158.1. ¹H NMR (400MHz, CDCl₃): δ 7.19 (dd, *J* = 7.9, 7.9 Hz, 1H), 6.89 (d, *J* = 7.8 Hz, 1H), 6.82 (bs, 1H), 6.73 (d, *J* = 7.5 Hz, 1H), bs (2H), 1.89 (t, ³*J*_{H-F} = 18.2 Hz, 3H).

General procedure 1: method A for amide formation

A solution 0.5M of ethyl 6-hydroxy-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-tetrahydropyrimidine-5-carboxylate **1** (1 equiv) and an appropriate aniline (1 equiv) in DMF dry was stirred at 100 °C for 30 min, then cooled to room temperature, and the product was precipitated with water, filtered, and rinsed with MeOH, yielding pure compound.

General procedure 2: method B for amide formation

A solution 0.5M of ethyl 6-hydroxy-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-tetrahydropyrimidine-5-carboxylate **1** (1 equiv) and an appropriate amine (1.2 equiv) in toluene dry was stirred at 100 °C for 2 h. After completion of reaction, the solvent was removed under vacuum. The product was purified by flash chromatography and/or by trituration with MeOH, yielding pure compound.

***N*-(3-ethoxyphenyl)-6-hydroxy-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide (3a)**. Title compound was prepared according to general procedure 1 using: 3-ethoxyaniline **2a** (31 mg, 0.23 mmol), ester **1** (86 mg, 0.23 mmol) in anhydrous DMF (0.46 mL).

Then, water (3 mL) was added, the resulting precipitate was filtered and rinsed with water (2 mL) and MeOH (2 mL), yielding **3a** (70 mg, 66% yield). UPLC/MS: Rt = 1.99 min (method A), MS (ESI) m/z: .460.4 [M+H]⁺, C₂₅H₂₂N₃O₄S [M+H]⁺ calculated: 460.5. HRMS (AP-ESI) m/z calc for C₂₅H₂₂N₃O₄S [M+H]⁺ 460.1331, found 460.1331. ¹H NMR (400MHz, DMSO-*d*₆): δ 11.63 (s, 1H, NH), 7.51-7.33 (m, 10H, Ph), 7.29 (dd, *J* = 8.1, 8.1 Hz, 1H, Ar), 7.14 (dd, *J* = 2.2, 2.2 Hz, 1H, Ar), 7.10 (dd, *J* = 8.0, 1.9 Hz, 1H, Ar), 6.80 (dd, *J* = 8.2, 2.3 Hz, 1H, Ar), 4.02 (q, *J* = 7.0 Hz, 2H), 1.30 (t, *J* = 6.9 Hz, 3H). ¹³C (100MHz, DMSO-*d*₆): δ 178.3 (CS), 168.7 (Cq), 159.0 (Cq), 139.0 (Cq), 136.4 (Cq), 130.1 (CH), 129.1 (CH), 129.0 (CH), 128.4 (CH), 114.0 (CH), 112.1 (CH), 108.1 (CH), 84.0 (Cq), 63.3 (CH₂), 14.6 (CH₃).

6-hydroxy-*N*-(3-isopropoxyphenyl)-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-

tetrahydropyrimidine-5-carboxamide (3b). Title compound was prepared according to general procedure 1 using: 3-isopropoxyaniline **2b** (45 mg, 0.30 mmol), ester **1** (110 mg, 0.30 mmol) in anhydrous DMF (0.6 mL). Then, water (3 mL) was added, the resulting precipitate was filtered and rinsed with water (2 mL) and MeOH (2 mL), yielding **3b** (50 mg, 35% yield). UPLC/MS: Rt = 2.10 min (method A), MS (ESI) m/z: 472.4 [M-H]⁻, C₂₆H₂₂N₃O₄S [M-H]⁻ calculated: 472.5. HRMS (AP-ESI) m/z calc for C₂₆H₂₄N₃O₄S [M+H]⁺ 474.1488, found 474.1494. ¹H NMR (400MHz, CDCl₃): δ 11.85 (s, 1H, NH), 7.50-7.60 (m, 6H, Ph), 7.36-7.33 (m, 4H, Ph), 7.27 (dd, *J* = 8.0, 5.0 Hz, 1H, Ar), 7.08-7.04 (m, 2H, Ar), 6.79 (dd, *J* = 8.3, 2.4 Hz, 1H, Ar), 4.54 (quint, *J* = 6.0 Hz, 1H), 1.36 (d, *J* = 6.0 Hz, 6H). ¹³C (100MHz, CDCl₃): δ 178.6 (CS), 169.3 (Cq), 167.8 (Cq), 162.3 (Cq), 158.7 (Cq), 139.2 (Cq), 138.1 (Cq), 136.2 (Cq), 130.2 (CH), 129.8 (CH), 129.7 (CH), 129.3 (CH), 129.1 (CH), 128.8 (CH), 128.6 (CH), 114.2 (CH), 114.0 (CH), 109.5 (CH), 83.6 (Cq), 70.3 (CH), 22.1 (CH₃).

***N*-(3-cyanophenyl)-6-hydroxy-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-tetrahydropyrimidine-5-**

carboxamide (3c). Title compound was prepared according to general procedure 1 using: 3-cianoaniline **2c** (48 mg, 0.41 mmol), ester **1** (150 mg, 0.41 mmol) in anhydrous DMF (0.82 mL). Then, water (4 mL) was added, the resulting precipitate was filtered and rinsed with water (4 mL) and MeOH (4 mL), yielding **3c** (122 mg, 68% yield) as light pink amorphous solid. UPLC/MS: Rt = 1.75 min (method A), MS (ESI) m/z: 441.4 [M+H]⁺, C₂₄H₁₇N₄O₃S [M+H]⁺ calculated: 441.4. HRMS (AP-ESI) m/z calc for C₂₄H₁₇N₄O₃S [M+H]⁺ 441.1021, found 441.1025. ¹H NMR (400MHz, DMSO-*d*₆): δ 11.72 (s, 1H, NH), 8.10 (bs, 1H, Ar), 7.86 (d, *J* = 8.2 Hz, 1H, Ar), 7.66 (d, *J* = 7.7 Hz, Ar), 7.58 (dd, *J* = 8.0, 8.0 Hz, 1H, Ar), 7.51-7.33 (m, 10H, Ph). ¹³C (100MHz,

DMSO-*d*₆): δ 178.2 (CS), 168.7 (Cq), 164.1 (Cq), 139.1 (Cq), 138.8 (Cq), 130.5 (CH), 129.1 (CH), 129.0 (CH), 128.9 (CH), 126.8 (CH), 125.1 (CH), 118.3 (Cq), 111.9 (Cq), 84.7 (Cq).

6-hydroxy-*N*-(3-nitrophenyl)-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide (3d). Title compound was prepared according to general procedure 1 using: 3-nitroaniline **2d** (38 mg, 0.27 mmol), ester **1** (100 mg, 0.27 mmol) in anhydrous DMF (0.54 mL). Then, water (4 mL) was added, the resulting precipitate was filtered and rinsed with water (4 mL) and MeOH (4 mL), yielding **3d** (70 mg, 57% yield) as pale yellow amorphous solid. UPLC/MS: Rt = 1.86 min (method A), MS (ESI) m/z: 461.5 [M+H]⁺, C₂₃H₁₇N₄O₅S [M+H]⁺ calculated: 461.5. HRMS (AP-ESI) m/z calc for C₂₃H₁₇N₄O₅S [M+H]⁺ 461.0920, found 461.0924. ¹H NMR (400MHz, DMSO-*d*₆): δ 11.84 (s, 1H, NH), 8.59 (dd, *J* = 2.2, 2.2 Hz, 1H, Ar), 8.01 (ddd, *J* = 8.2, 2.3, 0.8 Hz, 1H, Ar), 7.89 (ddd, *J* = 8.2, 2.1, 0.8 Hz, 1H, Ar), 7.64 (dd, *J* = 8.2, 8.2 Hz, 1H, Ar), 7.50-7.31 (m, 10H, Ph). ¹³C NMR (100MHz, CDCl₃): δ 178.2 (CS), 169.9 (Cq), 168.1 (Cq), 162.2 (Cq), 148.9 (Cq), 138.9 (Cq), 137.6 (Cq), 136.8 (Cq), 130.3 (CH), 129.9(CH), 129.8 (CH), 129.6 (CH), 129.3 (CH), 128.6 (CH), 128.5 (CH), 127.0 (CH), 120.6 (CH), 116.6 (CH), 84.1 (Cq).

***N*-(3-fluorophenyl)-6-hydroxy-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide (3e)**. Title compound was prepared according to general procedure 1 using: 3-fluoroaniline **2e** (48 μ L, 0.50 mmol), ester **1** (184 mg, 0.50 mmol) in anhydrous DMF (1 mL). Then, water (5 mL) was added, the resulting precipitate was filtered and rinsed with water (5 mL) and MeOH (5 mL), yielding **3e** (82mg, 38% yield) as white amorphous solid. UPLC/MS: Rt = 1.86 min (method A), MS (ESI) m/z: 434.3 [M+H]⁺, C₂₃H₁₇FN₃O₃S [M+H]⁺ calculated: 434.5. HRMS (AP-ESI) m/z calc for C₂₃H₁₇FN₃O₃S [M+H]⁺ 434.0975, found 434.0974. ¹H NMR (600MHz, CDCl₃): δ 11.91 (s, 1H, NH), 7.59-7.56 (m, 4H, Ar), 7.53-7.49 (m, 2H, Ar), 7.41 (ddd, *J* = 10.4, 2.3, 2.3 Hz, 1H, Ar), 7.35 (m, 5H), 7.18 (dd, *J* = 8.8, 2.0 Hz, 1H, Ar), 6.94 (ddd, *J* = 8.2, 8.2, 2.5 Hz, 1H, Ar). ¹³C NMR (150 MHz, CDCl₃): δ 178.4 (CS), 169.5 (Cq), 167.9 (Cq), 163.8 (Cq), 163.0 (d, ¹*J*_{CF} = 246.3 Hz, Cq), 139.1 (Cq), 137.9 (Cq), 136.8 (d, ³*J*_{CF} = 10.8 Hz, Cq), 130.6 (d, ³*J*_{CF} = 9.1 Hz, CH), 129.9 (CH), 129.8 (CH), 129.4 (CH), 129.2 (CH), 128.7 (CH), 128.5 (CH), 117.2 (d, ⁴*J*_{CF} = 3.0 Hz, CH), 113.2 (²*J*_{CF} = 21.5 Hz, CH), 109.4 (CH, ²*J*_{CF} = 21.4 Hz, CH), 83.8 (Cq). ¹⁹F NMR (565 MHz): δ -110.3 (s).

6-hydroxy-4-oxo-1,3-diphenyl-2-thioxo-*N*-(3-(trifluoromethoxy)phenyl)-1,2,3,4-tetrahydropyrimidine-5-carboxamide (3f). Title compound was prepared according to general procedure 1 using: 3-trifluoromethoxyaniline **2f** (54 μ L, 0.41 mmol), ester **1** (150 mg, 0.41 mmol)

in anhydrous DMF (0.82 mL). Then, water (4 mL) was added, the resulting precipitate was filtered and rinsed with water (4 mL) and MeOH (4 mL), yielding **3f** (79 mg, 39% yield) as white amorphous solid. UPLC/MS: Rt = 2.05 min (method A), MS (ESI) m/z: 500.3 [M+H]⁺, C₂₄H₁₇F₃N₃O₄S [M+H]⁺ calculated: 500.5. HRMS (AP-ESI) m/z calc for C₂₄H₁₇F₃N₃O₄S [M+H]⁺ 500.0892, found 500.0883. ¹H NMR (400MHz, DMSO-*d*₆): δ 11.70 (s, 1H, NH), 7.71 (s, 1H, Ar), 7.54-7.31 (m, 12H, Ar), 7.19 (d, *J* = 7.7 Hz, 1H, Ar). ¹³C (150MHz, CDCl₃): δ 178.4 (CS), 169.6 (Cq), 168.0 (Cq), 162.2 (Cq), 149.7 (Cq), 139.0 (Cq), 137.8 (Cq), 136.8 (Cq), 130.6 (CH), 129.9 (CH), 129.8 (CH), 129.5 (CH), 129.2 (CH), 128.7 (CH), 128.5 (CH), 120.5 (q, ¹J_{CF} = 258 Hz, Cq), 119.9 (CH), 118.4 (CH), 114.6 (CH), 83.8 (Cq). ¹⁹F NMR (565 MHz): δ -57.8 (s).

***N*-(3-(difluoromethoxy)phenyl)-6-hydroxy-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-**

tetrahydropyrimidine-5-carboxamide (3g). Title compound was prepared according to general procedure 1 using: 3-(difluoromethoxy)aniline **2g** (51 μL, 0.41 mmol), ester **1** (150 mg, 0.41 mmol) in anhydrous DMF (0.82 mL). Then, water (4 mL) was added, the resulting precipitate was filtered and rinsed with water (4 mL) and MeOH (4 mL), yielding **3g** (67 mg, 34% yield) as white amorphous solid. UPLC/MS: Rt = 1.95 min (method A), MS (ESI) m/z: 482.4 [M+H]⁺, C₂₄H₁₈F₂N₃O₃S [M+H]⁺ calculated: 482.5. HRMS (AP-ESI) m/z calc for C₂₄H₁₈F₂N₃O₃S [M+H]⁺ 482.0986, found 482.0984. ¹H NMR (600MHz, CDCl₃): δ 11.9 (s, 1H, NH), 7.58-7.50 (m, 6H, Ar), 7.38-7.31 (m, 7H), 7.00 (d, *J* = 8.1 Hz, 1H, Ar), 6.52 (t, *J* = 73.4 Hz, 1H, CHF₂). ¹³C (150MHz, CDCl₃): δ 178.4 (CS), 169.6 (Cq), 167.9 (Cq), 162.2 (Cq), 151.6 (Cq), 139.1 (Cq), 137.9 (Cq), 136.7 (Cq), 130.6 (CH), 129.9 (CH), 129.8 (CH), 129.5 (CH), 129.2 (CH), 128.7 (CH), 128.5 (CH), 118.6 (CH), 117.1 (CH), 115.7 (t, ¹J_{CF} = 260.5 Hz, CH), 113.3 (CH), 83.8 (Cq). ¹⁹F NMR (565 MHz): δ -81.2 (s).

***N*-(3-(1,1-difluoroethyl)phenyl)-6-hydroxy-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-**

tetrahydropyrimidine-5-carboxamide (3h). Title compound was prepared according to general procedure 1 using: 3-(1,1-difluoroethyl)aniline **2h** (50 mg, 0.32 mmol), ester **1** (117 mg, 0.32 mmol) in anhydrous DMF (0.64 mL). Then, water (3 mL) was added, the resulting precipitate was filtered and rinsed with water (3 mL) and MeOH (3 mL), yielding **3h** (104 mg, 68% yield) as white amorphous solid. UPLC/MS: Rt = 2.08 min (method B), MS (ESI) m/z: 480.1 [M+H]⁺, C₂₅H₂₀F₂N₃O₃S [M+H]⁺ calculated: 480.5. HRMS (AP-ESI) m/z calc for C₂₅H₂₀F₂N₃O₃S [M+H]⁺ 480.1193, found 480.1194. ¹H NMR (400MHz, CDCl₃): δ 11.95 (s, 1H, NH), 7.63-7.47 (m, 8H, Ar), 7.44 (dd, *J* = 7.9, 7.9 Hz, 1H, Ar), 7.37 (d, *J* = 7.9 Hz, 1H, Ar), 7.33-7.29 (m, 4H), 1.90 (t, *J*

= 18.2 Hz, 3H). ^{13}C (100MHz, CDCl_3): δ 178.5 (CS), 169.5 (Cq), 167.9 (Cq), 162.3 (Cq), 139.7 (t, $^2J_{\text{CF}} = 27.3$ Hz, Cq), 139.1 (Cq), 137.9 (Cq), 135.6 (Cq), 129.9 (CH), 129.8 (CH), 129.5 (CH), 129.2 (CH), 128.7 (CH), 122.9 (CH), 122.5 (t, $^3J_{\text{CF}} = 5.9$ Hz, CH), 121.3 (t, $^1J_{\text{CF}} = 239.5$ Hz, CF_2), 118.2 (t, $^3J_{\text{CF}} = 6.4$ Hz, CH), 83.7 (Cq), 26.0 (t, $^2J_{\text{CF}} = 29.5$ Hz, CH_3). ^{19}F NMR (565 MHz): δ -87.1 (s).

***N*-(3-(difluoromethyl)phenyl)-6-hydroxy-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-**

tetrahydropyrimidine-5-carboxamide (3i). Title compound was prepared according to general procedure 1 using: 3-(difluoromethyl)aniline **2i** (43 mg, 0.30 mmol), ester **1** (110 mg, 0.30 mmol) in anhydrous DMF (0.60 mL). Then, water (3 mL) was added, the resulting precipitate was filtered and rinsed with water (3 mL) and MeOH (3 mL), yielding **3i** (45 mg, 32% yield) as white amorphous solid. UPLC/MS: $R_t = 1.94$ min (method A), MS (ESI) m/z : 466.4 $[\text{M}+\text{H}]^+$, $\text{C}_{24}\text{H}_{18}\text{F}_2\text{N}_3\text{O}_3\text{S}$ $[\text{M}+\text{H}]^+$ calculated: 466.5. HRMS (AP-ESI) m/z calc for $\text{C}_{24}\text{H}_{18}\text{F}_2\text{N}_3\text{O}_3\text{S}$ $[\text{M}+\text{H}]^+$ 466.1037, found 466.1041. ^1H NMR (600MHz, CDCl_3): δ 11.97 (s, 1H, NH), 7.68 (s, 1H), 7.61-7.47 (m, 8H, Ar), 7.38-7.31 (m, 5H, Ar), 6.63 (t, $J = 56.5$ Hz, 1H, CHF_2). ^{13}C (150MHz, CDCl_3): δ 178.5 (CS), 169.6 (Cq), 167.9 (Cq), 162.2 (Cq), 139.1 (Cq), 137.9 (Cq), 135.9 (Cq), 135.8 (t, $^2J_{\text{CF}} = 22.6$ Hz, Cq), 130.0 (CH), 129.9 (CH), 129.8 (CH), 129.5 (CH), 129.2 (CH), 128.7 (CH), 128.5 (CH), 123.8 (CH), 123.3 (t, $^3J_{\text{CF}} = 6.1$ Hz, CH), 119.0 (t, $^3J_{\text{CF}} = 6.1$ Hz, CH), 119.0 (t, $^3J_{\text{CF}} = 6.2$ Hz, CH), 114.0 (t, $^1J_{\text{CF}} = 240.8$ Hz, CH), 83.8 (Cq). ^{19}F NMR (565 MHz): δ -111.5 (s).

***N*-(4-fluoro-3-methylphenyl)-6-hydroxy-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-**

tetrahydropyrimidine-5-carboxamide (3j). Title compound was prepared according to general procedure 2 using: 4-fluoro-3-methoxy-aniline **2j** (61 mg, 0.43 mmol), ester **1** (132 mg, 0.36 mmol) in anhydrous toluene (0.72 mL). Then, the solvent was removed under vacuum, the residue was treated with water (3 mL), the resulting precipitate was filtered and rinsed with MeOH (3 mL), yielding **3j** (126 mg, 72% yield) as white amorphous solid. UPLC/MS: $R_t = 1.94$ min (method A), MS (ESI) m/z : 464.4 $[\text{M}+\text{H}]^+$, $\text{C}_{24}\text{H}_{19}\text{FN}_3\text{O}_4\text{S}$ $[\text{M}+\text{H}]^+$ calculated: 464.5. HRMS (AP-ESI) m/z calc for $\text{C}_{24}\text{H}_{19}\text{FN}_3\text{O}_4\text{S}$ $[\text{M}+\text{H}]^+$ 464.1080, found 464.1084. ^1H NMR (600MHz, CDCl_3): δ 11.84 (s, 1H, NH), 7.59-7.49 (m, 6H, Ar), 7.34-7.31 (m, 4H), 7.09-7.06 (m, 3H). ^{13}C (150MHz, CDCl_3): δ 178.5 (CS), 169.1 (Cq), 167.8 (Cq), 162.3 (Cq), 150.4 ($^1J_{\text{CF}} = 245.4$ Hz, Cq), 148.1 ($^2J_{\text{CF}} = 11.3$ Hz, Cq), 139.1 (Cq), 138.0 (Cq), 131.5 (d, $^4J_{\text{CF}} = 3.3$, Cq), 129.9 (CH), 129.8 (CH), 129.4 (CH), 129.1 (CH), 128.7 (CH), 128.5 (CH), 116.5 (d, $^2J_{\text{CF}} = 19.6$ Hz, CH), 114.1 (d, $^3J_{\text{CF}} = 6.8$ Hz, CH), 107.6 (CH), 83.6 (Cq), 56.4 (OCH_3). ^{19}F NMR (565 MHz): δ -136.7.

[4-[tert-butyl(dimethyl)silyl]oxy-3,5-dimethoxy-phenyl]methanamine (4a). A suspension of compound **10** (250 mg, 0.77 mmol), in 7 mL of MeOH dry, was cooled to 0 °C and treated with NiCl₂·6H₂O (734.0 mg, 3.09 mmol). The resulting mixture was stirred at the same temperature for 5 min before the addition of NaBH₄ (290 mg, 7.66 mmol). After 30 min, the reaction mixture was quenched with saturated aqueous NH₄Cl (10 mL) solution and extracted with EtOAc (3×15 mL). The combined extracts were dried over Na₂SO₄ and concentrated under vacuum. Flash chromatographic purification (elution by gradient from 100 to 80/20 DCM/MeOH·NH₃ 1N) afforded compound **4a** (110 mg, 48% yield) as a viscous oil. UPLC/MS: Rt = 0.94 min (Method B), MS (ESI) m/z: 281 of main fragment. ¹H NMR (400MHz, CDCl₃): δ 6.51 (s, 2H), 3.80 (bs, 2H), 3.79 (s, 6H), 1.00 (s, 9H, *t*Bu TBS), 0.12 (s, 6H, CH₃ TBS).

2-[4-[tert-butyl(dimethyl)silyl]oxy-3,5-dimethoxy-phenyl]ethanamine (4b). A suspension of compound **11** (250 mg, 0.74 mmol), in 7 mL of MeOH dry, was cooled to 0 °C and treated with NiCl₂·6H₂O (703.6 mg, 2.96 mmol). The resulting mixture was stirred at the same temperature for 5 min before the addition of NaBH₄ (279.9 mg, 7.40 mmol). After 30 min, the reaction mixture was quenched with saturated aqueous NH₄Cl (10 mL) solution and extracted with EtOAc (3×15 mL). The combined extracts were dried over Na₂SO₄ and concentrated under vacuum. Flash chromatographic purification (elution by gradient from 100 to 80/20 DCM/MeOH·NH₃ 1N) afforded compound **4b** (110 mg, 48% yield) as a viscous oil. UPLC/MS: Rt = 1.16 min (method B), MS (ESI) m/z: 312.2 [M+H]⁺, C₁₆H₃₀NO₃Si [M+H]⁺ calculated: 312.2. ¹H NMR (400MHz, DMSO-*d*₆): δ 6.45 (s, 2H), 3.75 (s, 6H), 2.75 (m, 2H), 2.55 (m, 2H), 1.00 (s, 9H, *t*Bu TBS), 0.12 (s, 6H, CH₃ TBS).

(E)-3-[4-[tert-butyl(dimethyl)silyl]oxy-3,5-dimethoxy-phenyl]prop-2-en-1-amine (4c).

A 2M solution of LiAlH₄ in THF (1.37 mL, 2.73 mmol) was added to a suspension of AlCl₃ (363 mg, 2.73 mmol) in THF dry (6 mL) at 0°C under argon. After 10 min, a solution of intermediate **13** (774 mg, 0.78 mmol) in 5 mL of THF dry was added dropwise. The mixture was stirred for 30 min at 50°C, and then cooled at 0°C, quenched with ice water (5 mL). The pH was adjusted to 9-10 with NaOH 2M solution. The mixture was extracted with EtOAc (3x20 mL). The combined organic extracts were dried over Na₂SO₄ and concentrated under vacuum. The ¹H-NMR of the crude of the reaction showed the presence of *E/Z* isomers in ratio 1:0.12. Flash chromatographic purification (elution by gradient from 100 to 80/20 DCM/MeOH·NH₃ 1N) afforded compound **(E)-4c** (90 mg, 35% yield) as a viscous oil. UPLC/MS: Rt = 1.26 min (method B), MS (ESI) m/z:

307.2 main fragment. ^1H NMR (400MHz, CDCl_3): δ 6.58 (s, 2H), 6.41 (dd, $J = 16.0, 1.5$ Hz, 1H), 6.20 (ddd, $J = 15.8, 6.0, 6.0$ Hz, 1H), 3.80 (s, 6H), 3.47 (dd, $J = 6.0, 1.4$ Hz, 2H), 1.00 (s, 9H, *t*Bu TBS), 0.12 (s, 6H, CH_3 TBS).

(*E*)-2-(3-methoxyphenyl)ethenamine (4f). A 2M solution of LiAlH_4 in THF (1.75 mL, 3.5 mmol) was added to a suspension of AlCl_3 (467 mg, 3.5 mmol) in anhydrous THF (8 mL) at 0°C under argon. After 10 min, a solution of intermediate 14 (159 mg, 1.0 mmol) in 6 mL of anhydrous THF was added dropwise. The mixture was stirred for 30 min at 50°C , and then cooled at 0°C , quenched with ice water (7 mL). The pH was adjusted to 9-10 with NaOH 2M solution. The mixture was extracted with EtOAc (3x20 mL). The combined organic extracts were dried over Na_2SO_4 and concentrated under vacuum. The ^1H -NMR of the crude of the reaction showed the presence of *E/Z* isomers in ratio 1:0.10. Flash chromatographic purification (elution by gradient from 100 to 80/20 DCM/MeOH- NH_3 1N) yielded title compound (*E*)-**4f** (45 mg, 27% yield) as yellow viscous oil. UPLC/MS: $R_t = 1.15$ min (method A), MS (ESI) m/z : 147.0 main fragment. ^1H NMR (400MHz, CDCl_3): δ 7.22 (dd, $J = 7.9, 7.9$ Hz, 1H), 6.97 (d, $J = 7.8$ Hz, 1H), 6.91 (dd, $J = 2.0, 2.0$ Hz, 1H), 6.78 (dd, $J = 8.1, 2.6$ Hz, 1H), 6.48 (ddd, $J = 15.9, 1.7, 1.7$ Hz, 1H), 6.32 (ddd, $J = 15.8, 5.8, 5.8$ Hz, 1H), 3.81 (s, 3H), 3.49 (dd, $J = 5.8, 1.3$ Hz, 2H).

***N*-(4-((*tert*-butyldimethylsilyloxy)-3,5-dimethoxybenzyl)-6-hydroxy-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide (5a)**. Title compound was prepared according to general procedure 2 using: amine **4a** (80 mg, 0.27 mmol), ester **1** (83 mg, 0.22 mmol) in anhydrous toluene (0.44 mL). Then, the solvent was removed under vacuum. Flash chromatographic purification (elution by gradient from 100 to 85/15 Cyclohexane/EtOAc) afforded **5a** (95 mg, 70% yield) as a viscous oil. UPLC/MS: $R_t = 2.08$ min (method B), MS (ESI) m/z : 620.3 $[\text{M}+\text{H}]^+$, $\text{C}_{32}\text{H}_{38}\text{N}_3\text{O}_6\text{SSi}$ $[\text{M}+\text{H}]^+$ calculated: 620.8. ^1H NMR (400MHz, CDCl_3): δ 10.13 (dd, $J = 5.8, 5.8$ Hz, 1H, NH), 7.55-7.42 (m, 6H, Ph), 7.30-7.23 (m, 4H, Ph), 6.44 (s, 2H, Ar), 4.48 (d, $J = 5.9$ Hz, 2H), 3.76 (s, 6H), 1.00 (s, 9H), 0.12 (s, 6H). ^{13}C (100MHz, CDCl_3): δ 179.0 (CS), 170.4 (CONH), 167.5 (Cq), 162.2 (Cq), 152.0 (Cq), 128.7 (CH), 127.8 (CH), 105.8 (CH), 83.1 (Cq), 56.0 (OCH₃), 45.0 (CH₂), 25.9 (CH₃, TBS), 18.8 (Cq, TBS), -4.5 (CH₃, TBS).

***N*-(4-((*tert*-butyldimethylsilyloxy)-3,5-dimethoxyphenethyl)-6-hydroxy-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide (5b)**. Title compound was prepared according to general procedure 2 using: amine **4b** (40 mg, 0.12 mmol), ester **1** (37 mg, 0.10 mmol) in anhydrous toluene (0.50 mL). Then, the solvent was removed under vacuum. Flash

chromatographic purification (elution by gradient from 100 to 75/25 Cyclohexane/EtOAc) afforded **5b** (27 mg, 42% yield) as a viscous oil. UPLC/MS: Rt = 2.38 min (method B), MS (ESI) m/z: 634.2 [M+H]⁺, C₃₃H₄₀N₃O₆SSi [M+H]⁺ calculated: 634.8. ¹H NMR (400MHz, CDCl₃): δ 10.04 (dd, *J* = 5.7, 5.7 Hz, 1H, NH), 7.54-7.43 (m, 6H, Ph), 7.29-7.22 (m, 4H, Ph), 6.35 (s, 2H, Ar), 3.71 (s, 6H), 3.65 (ddd, *J* = 6.8, 6.8, 6.8 Hz, 2H), 2.80 (dd, *J* = 6.8, 6.8 Hz, 2H), 1.00 (s, 9H), 0.10 (s, 6H).

(E)-N-(3-(4-((tert-butyldimethylsilyl)oxy)-3,5-dimethoxyphenyl)allyl)-6-hydroxy-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide (5c). Title compound was prepared according to general procedure 2 using: amine **4c** (65 mg, 0.20 mmol), ester **1** (63 mg, 0.17 mmol) in anhydrous toluene (0.34 mL). Then, the solvent was removed under vacuum. Flash chromatographic purification (elution by gradient from 100 to 75/25 Cyclohexane/EtOAc) afforded **5c** (43 mg, 39% yield) as a viscous oil. UPLC/MS: Rt = 2.30 min (method B), MS (ESI) m/z: 646.3 [M+H]⁺, C₃₄H₃₉N₃O₆SSi [M+H]⁺ calculated: 646.8. ¹H NMR (400MHz, CDCl₃): δ 10.06 (dd, *J* = 5.3, 5.3 Hz, 1H, NH), 7.55-7.42 (m, 6H, Ph), 7.30-7.26 (m, 4H, Ph), 6.53 (s, 2H, Ar), 6.47 (ddd, *J* = 15.7, 1.4, 1.4 Hz, 1H), 6.02 (ddd, *J* = 15.7, 6.5, 6.5 Hz, 1H), 4.17 (ddd, *J* = 6.5, 6.5, 1.4 Hz, 2H), 3.79 (s, 6H), 1.00 (s, 9H), 0.12 (s, 6H). ¹³C (100MHz, CDCl₃): δ 179.0 (CS), 170.6 (CONH), 167.4 (Cq), 162.3 (Cq), 151.8 (Cq), 139.4 (Cq), 138.4 (Cq), 134.7 (CH), 129.7 (CH), 129.2 (CH), 128.9 (CH), 128.6 (CH), 103.9 (CH, 2C), 83.1 (Cq), 55.9 (OCH₃), 42.6 (CH₂), 25.9 (CH₃, TBS), 18.9 (Cq, TBS), -4.5 (CH₃, TBS)

6-hydroxy-N-(3-methoxybenzyl)-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide (5d). Title compound was prepared according to general procedure 2 using: 3-methoxybenzylamine **4d** (21 μL, 0.16 mmol), ester **1** (50 mg, 0.14 mmol) in anhydrous toluene (0.28 mL). Then, the solvent was removed under vacuum. Flash chromatographic purification (elution by gradient from 100 to 75/25 Cyclohexane/EtOAc) afforded **5d** (39 mg, 62% yield) as an amorphous white solid. UPLC/MS: Rt = 2.26 min (method A), MS (ESI) m/z: 460.2 [M+H]⁺, C₂₅H₂₂N₃O₄S [M+H]⁺ calculated: 460.5. HRMS (AP-ESI) m/z calc for C₂₅H₂₂N₃O₄S [M+H]⁺ 460.1331, found 460.1325. ¹H NMR (400MHz, DMSO-*d*₆): δ 10.26 (dd, *J* = 6.0, 6.0 Hz, 1H, NH), 7.47-7.35 (m, 6H, Ar), 7.30-7.25 (m, 5H, Ar), 6.91-6.85 (m, 3H, Ar), 4.55 (d, *J* = 6.2 Hz, 2H), 3.73 (s, 3H). ¹³C (100MHz, DMSO-*d*₆): δ 178.6 (CS), 169.7 (Cq), 159.4 (Cq), 139.3 (Cq), 138.8 (Cq), 129.8 (CH), 129.0 (CH), 128.2 (CH), 119.7 (CH), 113.6 (CH), 112.7 (CH), 82.8 (Cq), 55.1 (OCH₃), 43.7 (CH₂).

6-hydroxy-*N*-(3-methoxyphenethyl)-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-

tetrahydropyrimidine-5-carboxamide (5e). Title compound was prepared according to general procedure 2 using: 3-methoxyphenethylamine **4e** (38 μ L, 0.26 mmol), ester **1** (80 mg, 0.22 mmol) in anhydrous toluene (0.44 mL). Then, the solvent was removed under vacuum. Flash chromatographic purification (elution by gradient from 100 to 75/25 Cyclohexane/EtOAc) afforded **5e** (43 mg, 41% yield) as an amorphous white solid. UPLC/MS: Rt = 2.37 min (method A), MS (ESI) m/z: 474.1 [M+H]⁺, C₂₆H₂₄N₃O₄S [M+H]⁺ calculated: 474.1. HRMS (AP-ESI) m/z calc for C₂₆H₂₄N₃O₄S [M+H]⁺ 474.1488, found 474.1489. ¹H NMR (400MHz, CDCl₃): δ 10.03 (dd, J = 6.0, 6.0 Hz, 1H, NH), 7.54-7.43 (m, 6H, Ar), 7.29-7.20 (m, 5H, Ar), 6.79-6.72 (m, 3H, Ar), 3.76 (s, 3H), 3.67 (ddd, J = 6.9, 6.9, 6.9 Hz, 2H), 2.87 (dd, J = 7.2 Hz, 2H). ¹³C (150MHz, CDCl₃): δ 179.0 (CS), 170.8 (Cq), 167.3 (Cq), 162.2 (Cq), 159.9 (Cq), 139.4 (Cq), 139.1 (Cq), 138.4 (Cq), 129.9 (CH), 129.6 (CH), 129.1 (CH), 128.9 (CH), 128.8 (CH), 128.7 (CH), 121.0 (CH), 114.4 (CH), 112.5 (CH), 83.0 (Cq), 55.3 (OCH₃), 41.9 (CH₂), 35.5 (CH₂).

(*E*)-6-hydroxy-*N*-(3-(3-methoxyphenyl)allyl)-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-

tetrahydropyrimidine-5-carboxamide (5f). Title compound was prepared according to general procedure 2 using: amine **4f** (45 mg, 0.28 mmol), ester **1** (85 mg, 0.23 mmol) in anhydrous toluene (0.46 mL). Then, the solvent was removed under vacuum, flash chromatographic purification (elution by gradient from 100 to 70/30 Cyclohexane/EtOAc) afforded **5f** (42 mg, 38% yield) as an amorphous white solid. UPLC/MS: Rt = 2.33 min (method A), MS (ESI) m/z: .486.1 [M+H]⁺, C₂₇H₂₄N₃O₄S [M+H]⁺ calculated: 486.5. ¹H NMR (600MHz, CDCl₃): δ 10.11 (dd, J = 5.5, 5.5 Hz, 1H, NH), 7.56-7.48 (m, 7H, Ar), 7.33-7.30 (m, 3H, Ar), 7.25 (dd, J = 8.0, 8.0 Hz, 1H, Ar), 6.95 (d, J = 7.7 Hz, 1H, Ar), 6.89 (dd, J = 2.0, 2.0 Hz, 1H, Ar), 6.84 (dd, J = 8.0, 2.2 Hz, 1H), 6.56 (d, J = 15.8 Hz, 1H), 6.18 (ddd, J = 15.8, 6.4, 6.4 Hz, 1H), 4.22 (ddd, J = 6.1, 6.1, 1.4 Hz, 2H), 3.83 (s, 3H). ¹³C NMR (150MHz, CDCl₃): δ 179.0 (CS), 170.7 (CONH), 167.4 (Cq), 162.2 (Cq), 159.9 (Cq), 139.3 (Cq), 138.4 (Cq), 137.4 (Cq), 134.1 (CH), 129.8 (CH), 129.7 (CH), 129.2 (CH), 128.9 (CH), 128.8 (CH), 128.6 (CH, 2C), 122.8 (CH), 119.3 (CH), 114.0 (CH), 111.9 (CH), 83.1 (Cq), 55.4 (OCH₃), 42.5 (CH₂).

General procedure 3: TBS deprotection. A solution 0.5M of silylated precursor (1 equiv) was treated with TBAF 1M solution in THF (1.5 equiv). The reaction mixture stirred for 3 h. Then, the mixture was diluted with EtOAc, washed with water, and concentrated under vacuum. The crude material was purified by flash chromatography.

6-hydroxy-*N*-(4-hydroxy-3,5-dimethoxybenzyl)-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide (6). Title compound was prepared according to general procedure 3 using: intermediate **5a** (90 mg, 0.14 mmol), TBAF 1M in THF (220 μ L, 0.27 mmol) in anhydrous THF (0.28 mL). The crude was purified by silica gel flash chromatography (elution by gradient from 100 to 60/40 Cyclohexane/EtOAc) to yield **6** (42 mg, 54%) as pale yellow amorphous solid. UPLC/MS: Rt = 2.08 min (method A), MS (ESI) m/z: 504.2 [M-H]⁻, C₂₆H₂₂N₃O₆S [M-H]⁻ calculated: 504.5. HRMS (AP-ESI) m/z calc for C₂₆H₂₄N₃O₆S [M+H]⁺ 506.1386, found 506.1373. ¹H NMR (400MHz, DMSO-*d*₆): δ 10.18 (dd, *J* = 6.2, 6.2 Hz, 1H, NH), 8.37 (s, 1H, OH), 7.48-7.35 (m, 6H, Ph), 7.30-7.23 (m, 4H, Ph), 6.65 (s, 2H, Ar), 4.46 (d, *J* = 6.1 Hz, 2H), 3.73 (s, 6H). ¹³C (100MHz, DMSO-*d*₆): δ 178.6 (CS), 169.4 (Cq), 150.0 (Cq), 139.2 (Cq), 135.3 (Cq), 128.9 (CH), 128.1 (CH), 126.7 (CH), 106.1 (CH), 82.8 (Cq), 56.1 (OCH₃), 43.7 (CH₂).

6-hydroxy-*N*-(4-hydroxy-3,5-dimethoxyphenethyl)-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide (7). Title compound was prepared according to general procedure 3 using: intermediate **5b** (26 mg, 0.04 mmol), TBAF 1M in THF (60 μ L, 0.06 mmol) in anhydrous THF (0.1 mL). The crude was purified by silica gel flash chromatography (elution by gradient from 100 to 60/40 Cyclohexane/EtOAc) to yield **7** (11 mg, 53%) as pale yellow amorphous solid. UPLC/MS: Rt = 2.11 min (method A), MS (ESI) m/z: 518.1 [M-H]⁻, C₂₇H₂₄N₃O₆S [M-H]⁻ calculated: 518.6. HRMS (AP-ESI) m/z calc for C₂₇H₂₆N₃O₆S [M+H]⁺ 520.1542. ¹H NMR (400MHz, CDCl₃): δ 10.07 (dd, *J* = 5.8, 5.8 Hz, 1H, NH), 7.55-7.42 (m, 6H, Ph), 7.29-7.22 (m, 4H, Ph), 6.40 (s, 2H, Ar), 5.41 (s, 1H, OH), 3.8 (s, 6H), 3.65 (ddd, *J* = 6.1, 6.1, 6.1 Hz, 2H), 2.81 (dd, *J* = 6.8 Hz, 6.8 Hz, 2H). ¹³C (150MHz, CDCl₃): δ (CS), 179.0 (CS), 170.6 (CONH), 167.3 (Cq), 162.2 (Cq), 147.3 (Cq), 139.4 (Cq), 138.4 (Cq), 133.8 (Cq), 129.7 (CH), 129.6 (CH), 129.2 (CH), 128.9 (CH), 128.8 (CH), 128.6 (CH), 105.5 (CH), 83.0 (Cq), 56.4 (OCH₃), 42.2 (CH₂), 35.5 (CH₂).

(*E*)-6-hydroxy-*N*-(3-(4-hydroxy-3,5-dimethoxyphenyl)allyl)-4-oxo-1,3-diphenyl-2-thioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide (8). Title compound was prepared according to general procedure 3 using: intermediate **5c** (40 mg, 0.06 mmol), TBAF 1M in THF (90 μ L, 0.09 mmol) in anhydrous THF (0.12 mL). The crude was purified by silica gel flash chromatography (elution by gradient from 100 to 60/40 Cyclohexane/EtOAc) to yield **8** (29 mg, 88%) as pale yellow amorphous solid. UPLC/MS: Rt = 2.11 min (method A), MS (ESI) m/z: 530.3 [M-H]⁻, C₂₈H₂₄N₃O₆S [M-H]⁻ calculated: 530.6. HRMS (AP-ESI) m/z calc for C₂₈H₂₆N₃O₆S [M+H]⁺

532.1542, found 532.1524. ¹H NMR (400MHz, DMSO-*d*₆): δ 10.01 (dd, *J* = 6.2, 6.2 Hz, 1H, NH), 8.44 (s, 1H, OH), 7.47-7.28 (m, 11H, Ph, OH), 6.68 (m, 2H, Ar), 6.46 (d, *J* = 15.8 Hz, 1H), 6.17 (ddd, *J* = 15.8, 6.2, 6.2 Hz, 1H), 4.13 (dd, *J* = 6.0, 6.0, 1.4 Hz, 2H), 3.75 (s, 6H). ¹³C (150 MHz, DMSO-*d*₆): δ 178.7 (CS), 169.6 (CONH), 148.1 (Cq), 139.4 (Cq), 135.7 (Cq), 132.7 (Cq), 129.1 (CH), 128.3 (CH), 126.7 (Cq), 121.7 (CH), 104.0 (CH), 83.2 (Cq), 58.0 (OCH₃), 41.9 (CH₂).

1-[4-[tert-butyl(dimethyl)silyl]oxy-3,5-dimethoxy-phenyl]-*N*-methoxy-methanimine (10).

Sodium acetate (138 mg, 1.69 mmol) and *N*-Methylhydroxylamine hydrochloride (141 mg, 1.69 mmol) were added to a solution of compound **9**⁵⁰ (250 mg, 0.844 mmol) in MeOH dry (5 ml) under argon. The reaction mixture was stirred at 50°C for 5h until completion of reaction. The solvent was removed under vacuum, and the residue was suspended in water (5 mL) and the product was extracted with EtOAc (5x3 mL). Collected organic layers were dried with Na₂SO₄, filtered and concentrated under vacuum affording desired product **10** as a mixture of E/Z isomers (260 mg, 95% yield). The product was used as such without further purification. UPLC/MS: Rt = 2.42 min (method B), MS (ESI) m/z: 326.2 [M+H]⁺, C₁₆H₂₇NO₃Si [M+H]⁺ calculated: 326.2. ¹H NMR (400MHz, CDCl₃) of *major isomer*: δ 7.96 (s, 1H), 6.78 (s, 2H), 3.95 (s, 3H), 3.82 (s, 6H), 1.00 (s, 9H, *t*Bu TBS), 0.13 (s, 6H, CH₃ TBS).

1-[(1,1-Dimethylethyl)dimethylsilyl]oxy]-2,6-dimethoxy-4-[(1*E*)-2-nitroethenyl]benzene (11).

Nitromethane (2.7 ml, 50.5 mmol) was carefully added to a mixture of aldehyde **9**⁵⁰ (300 mg, 1.01 mmol) and ammonium acetate (77.1 mg, 1.01 mmol) in toluene dry (10 mL) under argon. The reaction mixture was stirred for 20 hours at reflux under argon. Then the reaction mixture was cooled at room temperature, quenched with water (10 mL), and extracted with EtOAc (2x10 mL). Collected organic layers were dried over Na₂SO₄, filtered and concentrated under vacuum. Flash chromatographic purification (elution by gradient from 100 to 95/5 Cyclohexane/EtOAc) afforded compound (*E*)-**11** (308 mg, 90% yield) as amorphous yellow solid. UPLC/MS: Rt = 2.41 min (method B), MS (ESI) m/z: 340.2 [M+H]⁺, C₁₆H₂₆NO₅Si [M+H]⁺ calculated: 340.1. ¹H NMR (400MHz, CDCl₃): δ 7.93 (d, *J* = 13.5 Hz, 1H), 7.52 (d, *J* = 13.5 Hz, 1H), 6.73 (s, 2H), 3.84 (s, 6H), 1.01 (s, 9H, *t*Bu TBS), 0.15 (s, 6H, CH₃ TBS).

3-[4-[tert-butyl(dimethyl)silyl]oxy-3,5-dimethoxy-phenyl]prop-2-enenitrile 13. To a solution of diethylcyanomethyl phosphonate (180 μL, 1.1 mmol) in THF (8 mL) was added *t*-BuOK (125 mg, 1.1 mmol) at ice-water bath temperature with stirring for 30 min. After that, aldehyde **9**⁵⁰ (300 mg, 1.0 mmol) in THF (3 mL) was added dropwise into the above mixture at room temperature

and was stirred overnight. The reaction mixture was quenched with water and extracted with EtOAc, washed with brine, dried over anhydrous Na₂SO₄, filtered and concentrated under vacuum. Flash chromatographic purification (elution by gradient from 100 to 85/15 Cyclohexane/EtOAc) afforded title compound **13** (270 mg, 84% yield) as an *E/Z* mixture in ratio 1:0.12. UPLC/MS: Rt = 2.28 min (method B), MS (ESI) m/z: 320.2 [M+H]⁺, C₁₇H₂₆NO₃Si [M+H]⁺ calculated: 320.2. ¹H NMR (400MHz, CDCl₃) for major isomer: δ 7.29 (d, *J* = 16.5 Hz, 1H), 6.63 (s, 2H), 5.71 (d, *J* = 16.5 Hz, 1H), 3.82 (s, 6H), 1.00 (s, 9H, *t*Bu TBS), 0.12 (s, 6H, CH₃ TBS).

3-(3-methoxyphenyl)prop-2-enitrile (14). To a solution of diethylcyanomethyl phosphonate (523 μL, 3.2 mmol) in anhydrous THF (20 mL) was added *t*-BuOK (391 mg, 3.2 mmol) at ice-water bath temperature with stirring for 30 min. After that, to this mixture *m*-anisaldehyde **12** (400 mg, 2.94 mmol) in anhydrous THF (8 mL) was added dropwise at room temperature and was stirred overnight. The reaction mixture was quenched with water and extracted with EtOAc, washed with brine, dried over anhydrous Na₂SO₄, filtered and concentrated under vacuum. Flash chromatographic purification (elution by gradient from 100 to 85/15 Cyclohexane/EtOAc) afforded title compound **14** (412 mg, 81% yield) as an *E/Z* mixture in ratio 1:0.18. UPLC/MS: Rt = 1.96 min (method A), MS (ESI) m/z: 160.0 [M+H]⁺, C₁₀H₁₀NO [M+H]⁺ calculated: 160.1. ¹H NMR (400MHz, CDCl₃) for major isomer: δ 7.37 (d, *J* = 16.7 Hz, 1H), 7.31 (d, *J* = 7.9 Hz, 1H), 7.04 (d, *J* = 7.7 Hz, 1H), 6.98 (dd, *J* = 8.2, 2.6 Hz, 1H), 6.95 (m, 1H), 5.87 (d, *J* = 16.6 Hz, 1H), 3.83 (s, 3H).

Biology

Cell Viability Assay. Human cancer cell lines A549 (lung adenocarcinoma, ATCC CCL-185), DU-145 (androgen-independent prostate cancer, ATCC HTB-81) and HeLa (cervical carcinoma, ATCC CCL-2) were obtained from ATCC. Cells were routinely grown in Minimal Essential Medium containing Eagle's salts and L-Glutamine supplemented with 10% heat-inactivated FBS in a humidified atmosphere of 5% CO₂ at 37 °C. To assess the antiproliferative activity of the compounds, cells were seeded at a density of 2500 cells/well (HeLa) or 5000 cells/well (A549, DU-145) in 96-well plates, and cell viability was measured using the MTT assay as described previously.³⁹ Values are reported as the mean ± SD of two independent experiments.

Topoisomerase II activity assay. The activity of topoII was measured using a decatenation assay (Inspiralis) following the manufacturer's instructions. Compounds were dissolved in DMSO and

used at a concentration ranging from 200 to 1 μ M. Final DMSO concentration in the assay was \leq 1%. Reaction mixtures were incubated for 30 min at 37 °C and terminated with STEB buffer (40% (w/v) sucrose, 100 mM Tris-HCl pH8, 1 mM EDTA, 0.5 mg/ml Bromophenol Blue). Reaction products were resolved by electrophoresis in 1% agarose gels containing SYBR Safe DNA stain (Invitrogen), scanned and quantified using the ChemiDoc system (BioRad). IC₅₀ values were obtained with GraphPad Prism software (version 5.03) using the band intensities of the dose–response gels. Values are reported as the mean \pm SD of two independent experiments.

Topoisomerase II cleavage assay. Poison activity of the compounds was evaluated using a cleavage complex assay (Inspiralis) as described previously.⁶⁰ Compounds were tested at a fixed concentration of 200 μ M in the presence of 1 U of topoisomerase II and 500 ng of pBR233 plasmid at 37 °C for 6 min. Final DMSO concentration in the assay was 1%. Reaction products were subjected to electrophoresis in a 1% agarose gel, stained with SYBR Safe DNA stain and DNA bands visualized and quantified as described above.

Computational studies

Structural model. The crystal structure of the alpha isoform of human topoII, co-crystallized with etoposide, was downloaded from the RCSB PDB repository, namely PDBid 5GWK (alpha). ARN24319 was considered for the docking and classical molecular dynamic (MD) studies. The protein structure was processed with the Protein Preparation Wizard in the Schrödinger 2017 suite.⁶¹ The ligands' structure was generated and prepared with Ligprep for molecular docking, using the OPLS2005 force field and charges. All possible protonation and ionization states were generated at a pH of 7.4. Stereoisomers were generated with a limit of 32 stereoisomers per ligand.

Docking calculations. The receptor grid for each target was prepared using the OPLS2005 force field. We specified the area surrounding the co-crystallized ligand (i.e. etoposide) as the receptor-binding pocket. The grid center was set to be the centroid of the bound etoposide. The cubic grid had a side length of 20 Å. For the receptor, we included aromatic hydrogen atoms as potential H-bond donors and halogens as potential acceptors. After grid preparation, ligands were first docked into the generated receptor grids using the extra precision (XP) scoring function. Flexible ligand sampling was considered in the docking procedure. All poses were subjected to post-docking minimization. The conformational degrees of freedom of the ligands were extensively explored by allowing nitrogen inversions as well as multiple ring conformations.

Classical MD Simulations: The most prevalent binding mode obtained from the docking studies was used for MD simulations with GROMACS version 5.1. All bonds were constrained using the P-LINCS algorithm, with an integration time step of 2 fs. The Verlet cutoff scheme was used with a minimum cutoff of 1.2 nm for short-range Lennard-Jones interactions and the real-space contribution to the fourth-ordered Ewald algorithm, which was used to compute long-range electrostatic interactions. Dispersion correction was applied to energy and pressure terms. Periodic boundary conditions were applied in all three dimensions. Each system was equilibrated in two phases, during which restraints were placed on protein and DNA heavy atoms. The first equilibration was done under an NVT ensemble for 500 ps using the v-rescale thermostat ($\tau_T = 0.1$ ps), to heat the systems until a temperature of 310 K. The NVT thermalization was followed by a 500 ps-long NPT pressurization using the same thermostat and the Parrinello-Rahman barostat ($\tau_P = 2.0$ ps and $\kappa = 4.5 \times 10^{-4}$ bar $^{-1}$) to equilibrate the pressure at 1 bar. Production simulations were carried out under an NPT ensemble in the absence of any restraints. A 200 ns production run was conducted for the complex. The analysis was carried out using programs within the GROMACS package and Python-based in-house scripts.

5.5 References

1. Pommier, Y. Drugging Topoisomerases: Lessons and Challenges. *ACS Chem. Biol.* **2013**, *8*, 82-95.
2. Deweese, J. E.; Osheroff, N. The DNA Cleavage Reaction of Topoisomerase II: Wolf in Sheep's Clothing. *Nucleic Acids Res.* **2009**, *37*, 738-748.
3. Deweese, J. E.; Osheroff, M. A.; Osheroff, N. DNA Topology and Topoisomerases: Teaching a "Knotty" Subject. *Biochem. Mol. Biol. Educ.* **2008**, *37*, 2-10.
4. Iacopetta, D.; Rosano, C.; Puoci, F.; Parisi, O. I.; Saturnino, C.; Caruso, A.; Longo, P.; Ceramella, J.; Malzert-Freon, A.; Dallemagne, P.; Rault, S.; Sinicropi, M. S. Multifaceted Properties of 1,4-Dimethylcarbazoles: Focus on Trimethoxybenzamide and Trimethoxyphenylurea Derivatives as Novel Human Topoisomerase II Inhibitors. *Eur. J. Pharm. Sci.* **2017**, *96*, 263-272.
5. Rosenblum, D.; Joshi, N.; Tao, W.; Karp, J. M.; Peer, D. Progress and Challenges Towards Targeted Delivery of Cancer Therapeutics. *Nat Commun* **2018**, *9*, 1410.
6. Hu, W.; Huang, X. S.; Wu, J. F.; Yang, L.; Zheng, Y. T.; Shen, Y. M.; Li, Z. Y.; Li, X. Discovery of Novel Topoisomerase II Inhibitors by Medicinal Chemistry Approaches. *J. Med. Chem.* **2018**, *61*, 8947-8980.
7. Pommier, Y.; Sun, Y.; Huang, S. N.; Nitiss, J. L. Roles of Eukaryotic Topoisomerases in Transcription, Replication and Genomic Stability. *Nat. Rev. Mol. Cell Biol.* **2016**, *17*, 703-721.
8. Nitiss, J. L. DNA Topoisomerase II and Its Growing Repertoire of Biological Functions. *Nat. Rev. Cancer* **2009**, *9*, 327-337.
9. Lindsey, R. H., Jr.; Pendleton, M.; Ashley, R. E.; Mercer, S. L.; Deweese, J. E.; Osheroff, N. Catalytic Core of Human Topoisomerase II α : Insights into Enzyme-DNA Interactions and Drug Mechanism. *Biochemistry* **2014**, *53*, 6595-6602.
10. Liang, X.; Wu, Q.; Luan, S.; Yin, Z.; He, C.; Yin, L.; Zou, Y.; Yuan, Z.; Li, L.; Song, X.; He, M.; Lv, C.; Zhang, W. A Comprehensive Review of Topoisomerase Inhibitors as Anticancer Agents in the Past Decade. *Eur. J. Med. Chem.* **2019**, *171*, 129-168.
11. Wang, W.; Tse-Dinh, Y. C. Recent Advances in Use of Topoisomerase Inhibitors in Combination Cancer Therapy. *Curr. Top. Med. Chem.* **2019**, *19*, 730-740.

12. Hande, K. R. Topoisomerase II Inhibitors. *Cancer Chemother. Biol. Response Modif.* **2003**, *21*, 103-125.
13. Baldwin, E. L.; Osheroff, N. Etoposide, Topoisomerase II and Cancer. *Curr. Med. Chem. Anticancer Agents* **2005**, *5*, 363-372.
14. Palermo, G.; Minniti, E.; Greco, M. L.; Riccardi, L.; Simoni, E.; Convertino, M.; Marchetti, C.; Rosini, M.; Sissi, C.; Minarini, A.; De Vivo, M. An Optimized Polyamine Moiety Boosts the Potency of Human Type II Topoisomerase Poisons as Quantified by Comparative Analysis Centered on the Clinical Candidate F14512. *Chem. Commun. (Camb.)* **2015**, *51*, 14310-14313.
15. Baviskar, A. T.; Amrutkar, S. M.; Trivedi, N.; Chaudhary, V.; Nayak, A.; Guchhait, S. K.; Banerjee, U. C.; Bharatam, P. V.; Kundu, C. N. Switch in Site of Inhibition: A Strategy for Structure-Based Discovery of Human Topoisomerase IIalpha Catalytic Inhibitors. *ACS Med. Chem. Lett.* **2015**, *6*, 481-485.
16. Kadayat, T. M.; Park, S.; Shrestha, A.; Jo, H.; Hwang, S. Y.; Katila, P.; Shrestha, R.; Nepal, M. R.; Noh, K.; Kim, S. K.; Koh, W. S.; Kim, K. S.; Jeon, Y. H.; Jeong, T. C.; Kwon, Y.; Lee, E. S. Discovery and Biological Evaluations of Halogenated 2,4-Diphenyl Indeno[1,2-B]Pyridinol Derivatives as Potent Topoisomerase IIalpha-Targeted Chemotherapeutic Agents for Breast Cancer. *J. Med. Chem.* **2019**, *62*, 8194-8234.
17. Gouveia, R. G.; Ribeiro, A. G.; Segundo, M.; de Oliveira, J. F.; de Lima, M.; de Lima Souza, T. R. C.; de Almeida, S. M. V.; de Moura, R. O. Synthesis, DNA and Protein Interactions and Human Topoisomerase Inhibition of Novel Spiroacridine Derivatives. *Bioorg. Med. Chem.* **2018**, *26*, 5911-5921.
18. Pogorelcnik, B.; Perdih, A.; Solmajer, T. Recent Developments of DNA Poisons--Human DNA Topoisomerase IIalpha Inhibitors--as Anticancer Agents. *Curr. Pharm. Des.* **2013**, *19*, 2474-2488.
19. Pogorelcnik, B.; Perdih, A.; Solmajer, T. Recent Advances in the Development of Catalytic Inhibitors of Human DNA Topoisomerase IIalpha as Novel Anticancer Agents. *Curr. Med. Chem.* **2013**, *20*, 694-709.
20. Riddell, I. A.; Agama, K.; Park, G. Y.; Pommier, Y.; Lippard, S. J. Phenanthriplatin Acts as a Covalent Poison of Topoisomerase II Cleavage Complexes. *ACS Chem. Biol.* **2016**, *11*, 2996-3001.

21. Delgado, J. L.; Hsieh, C. M.; Chan, N. L.; Hiasa, H. Topoisomerases as Anticancer Targets. *Biochem J.* **2018**, *475*, 373-398.
22. Froelich-Ammon, S. J.; Osheroff, N. Topoisomerase Poisons: Harnessing the Dark Side of Enzyme Mechanism. *J. Biol. Chem.* **1995**, *270*, 21429-21432.
23. Gibson, E. G.; King, M. M.; Mercer, S. L.; Deweese, J. E. Two-Mechanism Model for the Interaction of Etoposide Quinone with Topoisomerase IIalpha. *Chem. Res. Toxicol.* **2016**, *29*, 1541-1548.
24. Bailly, C. Contemporary Challenges in the Design of Topoisomerase II Inhibitors for Cancer Chemotherapy. *Chem. Rev.* **2012**, *112*, 3611-3640.
25. Wu, C. C.; Li, T. K.; Farh, L.; Lin, L. Y.; Lin, T. S.; Yu, Y. J.; Yen, T. J.; Chiang, C. W.; Chan, N. L. Structural Basis of Type II Topoisomerase Inhibition by the Anticancer Drug Etoposide. *Science* **2011**, *333*, 459-462.
26. Wendorff, T. J.; Schmidt, B. H.; Heslop, P.; Austin, C. A.; Berger, J. M. The Structure of DNA-Bound Human Topoisomerase II Alpha: Conformational Mechanisms for Coordinating Inter-Subunit Interactions with DNA Cleavage. *J. Mol. Biol.* **2012**, *424*, 109-124.
27. Cowell, I. G.; Sondka, Z.; Smith, K.; Lee, K. C.; Manville, C. M.; Sidorchuk-Lesthuruge, M.; Rance, H. A.; Padget, K.; Jackson, G. H.; Adachi, N.; Austin, C. A. Model for MLL Translocations in Therapy-Related Leukemia Involving Topoisomerase IIbeta-Mediated DNA Strand Breaks and Gene Proximity. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 8989-8994.
28. Pendleton, M.; Lindsey, R. H., Jr.; Felix, C. A.; Grimwade, D.; Osheroff, N. Topoisomerase II and Leukemia. *Ann. N. Y. Acad. Sci.* **2014**, *1310*, 98-110.
29. Azarova, A. M.; Lyu, Y. L.; Lin, C. P.; Tsai, Y. C.; Lau, J. Y.; Wang, J. C.; Liu, L. F. Roles of DNA Topoisomerase II Isozymes in Chemotherapy and Secondary Malignancies. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 11014-11019.
30. Cowell, I. G.; Austin, C. A. Mechanism of Generation of Therapy Related Leukemia in Response to Anti-Topoisomerase II Agents. *Int. J. Environ. Res. Public Health* **2012**, *9*, 2075-2091.
31. Pommier, Y.; Leo, E.; Zhang, H.; Marchand, C. DNA Topoisomerases and Their Poisoning by Anticancer and Antibacterial Drugs. *Chem. Biol.* **2010**, *17*, 421-433.

32. Hu, C. X.; Zuo, Z. L.; Xiong, B.; Ma, J. G.; Geng, M. Y.; Lin, L. P.; Jiang, H. L.; Ding, J. Salvicine Functions as Novel Topoisomerase II Poison by Binding to Atp Pocket. *Mol. Pharmacol.* **2006**, *70*, 1593-1601.
33. Walker, J. V.; Nitiss, J. L. DNA Topoisomerase II as a Target for Cancer Chemotherapy. *Cancer Invest.* **2002**, *20*, 570-589.
34. Dimaggio, J. J.; Warrell, R. P., Jr.; Muindi, J.; Stevens, Y. W.; Lee, S. J.; Lowenthal, D. A.; Haines, I.; Walsh, T. D.; Baltzer, L.; Yaldae, S.; Young, C. V. Phase I Clinical and Pharmacological Study of Merbarone. *Cancer Res.* **1990**, *50*, 1151-1155.
35. Fortune, J. M.; Osheroff, N. Merbarone Inhibits the Catalytic Activity of Human Topoisomerase IIalpha by Blocking DNA Cleavage. *J. Biol. Chem.* **1998**, *273*, 17643-17650.
36. Malik, U. R.; Dutcher, J. P.; Caliendo, G.; Lasala, P.; Mitnick, R.; Wiernik, P. H. Phase II Trial of Merbarone in Patients with Malignant Brain Tumors. *Med. Oncol.* **1997**, *14*, 159-162.
37. Spallarossa, A.; Rotolo, C.; Sissi, C.; Marson, G.; Greco, M. L.; Ranise, A.; La Colla, P.; Busonera, B.; Loddo, R. Further SAR Studies on Bicyclic Basic Merbarone Analogues as Potent Antiproliferative Agents. *Bioorg. Med. Chem.* **2013**, *21*, 6328-6336.
38. Larsen, A. K.; Escargueil, A. E.; Skladanowski, A. Catalytic Topoisomerase II Inhibitors in Cancer Therapy. *Pharmacol. Ther.* **2003**, *99*, 167-181.
39. Ortega, J. A.; Riccardi, L.; Minniti, E.; Borgogno, M.; Arencibia, J. M.; Greco, M. L.; Minarini, A.; Sissi, C.; De Vivo, M. Pharmacophore Hybridization to Discover Novel Topoisomerase II Poisons with Promising Antiproliferative Activity. *J. Med. Chem.* **2018**, *61*, 1375-1379.
40. Ranise, A.; Spallarossa, A.; Schenone, S.; Bruno, O.; Bondavalli, F.; Pani, A.; Marongiu, M. E.; Mascia, V.; La Colla, P.; Loddo, R. Synthesis and Antiproliferative Activity of Basic Thioanalogues of Merbarone. *Bioorg. Med. Chem.* **2003**, *11*, 2575-2589.
41. Minniti, E.; Byl, J. A. W.; Riccardi, L.; Sissi, C.; Rosini, M.; De Vivo, M.; Minarini, A.; Osheroff, N. Novel Xanthone-Polyamine Conjugates as Catalytic Inhibitors of Human Topoisomerase IIalpha. *Bioorg. Med. Chem. Lett.* **2017**, *27*, 4687-4693.
42. Oviatt, A. A.; Kuriappan, J. A.; Minniti, E.; Vann, K. R.; Onuorah, P.; Minarini, A.; De Vivo, M.; Osheroff, N. Polyamine-Containing Etoposide Derivatives as Poisons of Human Type II Topoisomerases: Differential Effects on Topoisomerase IIalpha and IIbeta. *Bioorg. Med. Chem. Lett.* **2018**, *28*, 2961-2968.

43. Riccardi, L. G., V.; De Vivo, M. . Metal–Ligand Interactions in Drug Design. *Nat. Rev. Chem.* **2018**, *2*, 100-112.
44. Meanwell, N. A. Synopsis of Some Recent Tactical Application of Bioisosteres in Drug Design. *J. Med. Chem.* **2011**, *54*, 2529-2591.
45. Tanaka, R. H., N. Structure of Etoposide. *Anal. Sci.* **2007**, *23*, x29-x30.
46. Wilstermann, A. M.; Bender, R. P.; Godfrey, M.; Choi, S.; Anklin, C.; Berkowitz, D. B.; Osheroff, N.; Graves, D. E. Topoisomerase II - Drug Interaction Domains: Identification of Substituents on Etoposide That Interact with the Enzyme. *Biochemistry* **2007**, *46*, 8217-8225.
47. Bender, R. P.; Jablonksy, M. J.; Shadid, M.; Romaine, I.; Dunlap, N.; Anklin, C.; Graves, D. E.; Osheroff, N. Substituents on Etoposide That Interact with Human Topoisomerase IIalpha in the Binary Enzyme-Drug Complex: Contributions to Etoposide Binding and Activity. *Biochemistry* **2008**, *47*, 4501-4509.
48. Ding, S.; Jiao, N. N,N-Dimethylformamide: A Multipurpose Building Block. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 9226-9237.
49. Yang Heung, B. J. Convenient N-Formylation of Amines in Dimethylformamide with Methyl Benzoate under Microwave Irradiation. *Bull. Korean Chem. Soc.* **2010**, *31*, 1424-1426.
50. Pettit, G. R.; Searcy, J. D.; Tan, R.; Cragg, G. M.; Melody, N.; Knight, J. C.; Chapuis, J. C. Antineoplastic Agents. 585. Isolation of Bridelia Ferruginea Anticancer Podophyllotoxins and Synthesis of 4-Aza-Podophyllotoxin Structural Modifications. *J. Nat. Prod.* **2016**, *79*, 507-518.
51. Inspiralis Limited. <https://www.Inspiralis.Com/Assets/Technicaldocuments/Human-Topo-II-Alpha-Decatenation-Assay-Protocol.Pdf> (Accessed Oct 9, 2019).
52. De Vivo, M.; Cavalli, A. Recent Advances in Dynamic Docking for Drug Discovery. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7*, e1320.
53. Franco-Ulloa, S.; La Sala, G.; Miscione, G. P.; De Vivo, M. Novel Bacterial Topoisomerase Inhibitors Exploit Asp83 and the Intrinsic Flexibility of the DNA Gyrase Binding Site. *Int. J. Mol. Sci.* **2018**, *19*, 453.
54. Kuriappan, J. A.; Osheroff, N.; De Vivo, M. Smoothed Potential MD Simulations for Dissociation Kinetics of Etoposide to Unravel Isoform Specificity in Targeting Human Topoisomerase II. *J. Chem. Inf. Model.* **2019**, *59*, 4007-4017.
55. De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* **2016**, *59*, 4035-4061.

56. Wang, Y. R.; Chen, S. F.; Wu, C. C.; Liao, Y. W.; Lin, T. S.; Liu, K. T.; Chen, Y. S.; Li, T. K.; Chien, T. C.; Chan, N. L. Producing Irreversible Topoisomerase II-Mediated DNA Breaks by Site-Specific Pt(II)-Methionine Coordination Chemistry. *Nucleic Acids Res.* **2017**, *45*, 10861-10871.
57. Palermo, G.; Stenta, M.; Cavalli, A.; Dal Peraro, M.; De Vivo, M. Molecular Simulations Highlight the Role of Metals in Catalysis and Inhibition of Type II Topoisomerase. *J. Chem. Theory Comput.* **2013**, *9*, 857-862.
58. Palermo, G.; Cavalli, A.; Klein, M. L.; Alfonso-Prieto, M.; Dal Peraro, M.; De Vivo, M. Catalytic Metal Ions and Enzymatic Processing of DNA and RNA. *Acc. Chem. Res.* **2015**, *48*, 220-228.
59. Jacob, D. A.; Mercer, S. L.; Osheroff, N.; Dewese, J. E. Etoposide Quinone Is a Redox-Dependent Topoisomerase II Poison. *Biochemistry* **2011**, *50*, 5660-5667.
60. Bandele, O. J.; Osheroff, N. Cleavage of Plasmid DNA by Eukaryotic Topoisomerase II. *Methods Mol. Biol.* **2009**, *582*, 39-47.
61. Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. *J. Comput. Aided Mol. Des.* **2013**, *27*, 221-234.

Chapter 6: Conclusions and future perspectives

The characterization and the inhibition of DNA polymerase η is a project that started long time ago in the lab of my supervisor Dr. De Vivo at IIT. This lab has a strong expertise in the field of two metal ions enzymes, specifically, in studying polymerases and topoisomerases. The DNA polymerase η belongs to the Y-family of polymerases and is involved in the repair of the dsDNA through the TLS mechanism. It is not essential for the cell survival, although is the most efficient to bypass the UV-lesions as the pyrimidine dimers. It lacks in fidelity, for this reason it is also able to bypass, when overexpressed, damages induced by external agents as chemotherapeutic drugs like cisplatin. This often brings to drug resistance. In particular, it is well studied that cancer cells can develop this mechanism of defense, over time. In my PhD, I contributed to a project aimed at discovering a new class of small molecules capable to inhibit Pol η . To achieve our goal, the project was built into two main areas:

- 1) Set up the protocol for the purification of the enzyme, and the inhibition experiments also in cancer cell lines to understand, from a molecular point of view, the mechanism of action of the compounds;
- 2) Understand, from a structural point of view, the mechanism of action of the compounds and in particular the atomic interactions of the compounds with the enzyme to aid the structural based drug design efforts.

Thanks to our multidisciplinary team, we found a hit compound through a virtual screening campaign of the IIT-D3 library (~15000 small molecules readily usable) and the medicinal chemists of the group synthesized 35 analogs. I evaluated these 35 analogs first with a gel-based elongation assay. I used an available purification protocol of the Pol η (1-511) and then set up the gel-based assay to screen all the compounds in a fast, cheap and reliable way. I calculated all the IC_{50} values and then we decided to select the 10 best compounds with an IC_{50} lower/around than 20 μ M. We studied the DMPK properties of these 10 analogs and we found two compounds with a good kinetic solubility, plasma and microsomal stability. Consequently, I decided to continue our studies with compound 64 and the compound 61. They were evaluated on three different cancer cell lines (A375, A549 and OVCAR3) and also on a normal cell line (HEK293). I co-treated the cells with different concentrations of cisplatin and the compounds, trying to achieve the best

combination of drugs using the lowest possible concentration of our inhibitors. In this way I discovered that compound 64 was less potent than compound 61, but seems to be more suitable for our purpose because was not cytotoxic when we used it alone, mainly compound 64 could, in combination with cisplatin, decreased the LD₅₀ value. Moreover, it acts with a synergistic effect with cisplatin, in contrast to the compound 61, which has not. In conclusion, we reached our goal of finding a new potent inhibitor of Pol η to overcome drug resistance; this result is presented in a manuscript published in the European Journal of Medicinal Chemistry (Munafò F, Nigro M, Brindani N, et al., 2023) (Chapter 2). The compound belongs to the flavonoids family, which is known to contain very easily metabolized compounds. Thus, in order to improve its biodistribution and bioavailability, we started a new collaboration with Dr. Tirelli from IIT, to develop a new hyaluronic acid-based prodrug to improve the targeting and the pharmacokinetics properties of the inhibitor 64.

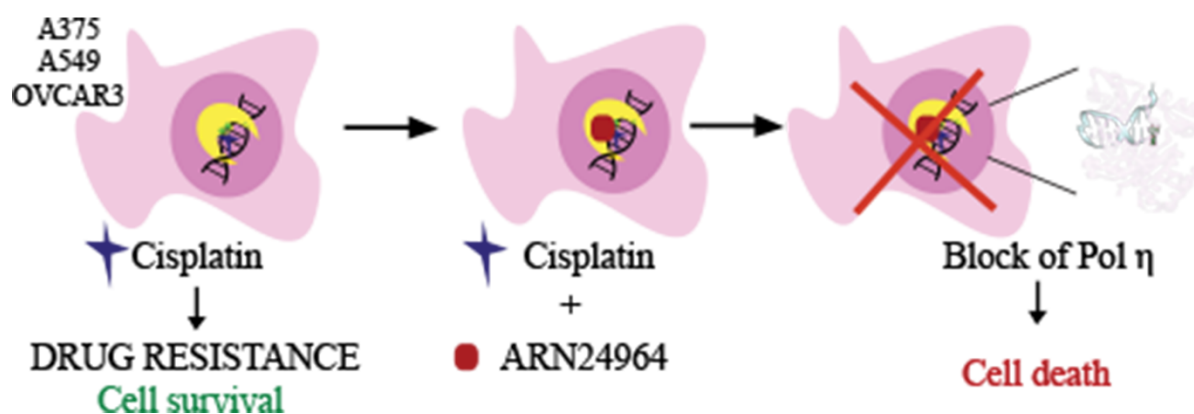


Figure 44 Summary of our results. We screened 34 compounds, and we tested only two compounds in three cancer cell lines. Co-treating the cells with compound ARN24964 and cisplatin, we found out that they have a synergistic effect leading to cancer cell death.

Our inhibitor could be a promising starting point to develop a new and more potent class of Pol η inhibitors. We are aware that it is less potent than the PNR-7-02 and that we did not characterize its selectivity, yet. Although, this compound has a promising DMPK profile.

Additionally, I spent 6 months at the EMBL of Grenoble in the group of Dr Marquez, head of the crystallization facility. He built an automated pipeline to crystallize protein and to co-crystallize protein with inhibitors. We obtained crystals with an excellent resolution, and confirmed the data already present in literature. In particular, we managed to visualize the catalytic site of the enzyme with the dsDNA and the Mg^{2+} ions both in presence and absence of the nucleotide. Unfortunately, we did not obtain crystals with our compounds, although we solve two structures that will be deposited in the PDB databank.

Based on all these results, we are planning future work. We are keenly aware that this inhibitor belongs to a class of compounds known to be poorly selective and easily metabolized. Therefore, we will further characterize the use of the prodrugs in combination with cisplatin to improve the PK properties. We are planning to determine the PK *in vivo*. Moreover, we would like to use 3D cell culture such as organoids, to better understand the mechanism of internalization of these prodrugs. Furthermore, we already have a second class of compounds ready to be tested in cancer cell lines. This is a completely new class of analogs, with promising expectations because these compounds do not belong to the flavonoids group and we believe they could be readily go *in vivo*.

At the same time, our collaboration with the EMBL of Grenoble will continue. We have already started the high-throughput screening of a fragment library of the HTX-Lab in order to find a novel pocket or an allosteric binding site in which compounds will not compete with nucleotides for binding.

As a side project, I worked also on human topoisomerase II α inhibitors. This enzyme is another two metal ions-based catalyst, involved in crucial pathways for cell survival. Due to its ability of acting on DNA topology, it is involved in several types of cancer. To date there are some chemotherapeutic drugs already available to treat cancer, but there is often the problem induced by drug resistance. For this reason, our aim here again was to discover a novel class of inhibitors that help in escaping resistance. Here, we found a novel class of topoisomerase II poisons with excellent DMPK properties and potency. I evaluated the compounds through a decatenation assay (Inspiralis), and a cleavage assay to better elucidate their mechanism of action. Afterwards, I tested

the best compounds on three different cancer cell lines (HeLa, DU145 and A549). Compound 3f showed a good cytotoxicity in all the three cancer cell lines. We now have a new aim: we started a new FEP (free energy perturbation) guided campaign to enhance this class of inhibitors. This effort led us to synthesized other 14 compounds, all with a good potency and cytotoxicity in all the three cancer cell lines. These preliminary data need to be further explored in order to find a potential novel inhibitor more potent than the 3f already published.