UNIVERSITÀ DEGLI STUDI DI TORINO

Dottorato in
Scienze Farmaceutiche e Biomolecolari

Ph.D. Program in Pharmaceutical and Biomolecular Sciences

2016 - 2019

Computational Strategies for Structure-based Lead Optimization and Analysis of SAR Transfer

PhD Student: Davide Bonanni

Supervisor: Prof. Marco L. Lolli

CHIM/08
2019/2020

Preface

The work reported in this thesis is the result of a research project carried out at the medicinal chemistry research group MEDSynth (Department of Drug Science and Technology, University of Turin, UniTO). The general topic that characterized this project is computer driven lead optimization. The purpose of lead optimisation is to maintain favourable properties in lead compounds, while improving on deficiencies in lead structure.

The first chapter of the thesis describes the computer-aided lead optimization carried out on meds433, a new potent dihydroorotate dehydrogenase (DHODH) inhibitor. Human DHODH is a flavin-dependent mitochondrial enzyme involved in de novo pyrimidine biosynthesis. *h*DHODH overexpression has been associated with acute myelogenous leukemia, a disease for which the standard of intensive care has not changed over decades. Structure-based approaches were used to generate working hypotheses to modify further meds433 and improve its molecular properties. A series of analogues were designed, synthetized and biologically evaluated. The above mentioned medicinal chemistry project is partially reported in the following publication:

- Sainas S, Pippione AC, Lupino E, Giorgis M, Circosta P, Gaidano V, Goyal P, Bonanni D, Rolando B, Cignetti A, Ducime A, Andersson M, Järvå M, Friemann R, Piccinini M, Ramondetti C, Buccinnà B, Al-Karadaghi S, Boschi D, Saglio G, Lolli ML. "Targeting Myeloid Differentiation Using Potent 2-Hydroxypyrazolo[1,5-a]pyridine Scaffold-Based Human Dihydroorotate Dehydrogenase Inhibitors*". J Med Chem.*, **2018**, 61(14): 6034-6055

Two research articles are in preparation and will describe the rest of the results reported in this dissertation.

In the second chapter, the PhD thesis details the developing of new computational strategies aimed to lead optimization. The project was carried out in collaboration with Prof. Jürgen Bajorath at the Department of Life Science Informatics at the b-it Institute of the University of Bonn. The exploration of structure-activity relationships (SARs) is central relevance in drug design. In such situations, one would ideally like to build upon priori knowledge, utilize available SAR information, and evaluate the possibility of an "SAR transfer", i.e., the exploration of an alternative chemical structures that displays similar SAR characteristics and potency progression. So far, few studies have computationally analyzed SAR transfer events.

However, currently lacking are structure-based approaches for the assessment and prediction of SAR transfer. Herein, an innovative computational approach for structure-based analysis of SAR transfer is discussed. To examine the impact of this new approach, the methodology was applied across different targets identified SAR transfer events with high frequency. The above mentioned project is detailed in following publication:

- Bonanni D, Lolli ML, Bajorath J. "Computational Method for Structure-Based Analysis of SAR Transfer". *J. Med. Chem.*, **2020**, 63(3): 1388-1396

Moreover, SAR transfer concept was extent from the singles events to SAR table. Tables were used to identify relevant SARs, which can be transfer among different series of analogues.

Table of contents

1. Lead Optimization of Meds433

1.1 Introduction

1.1.1. The *De Novo* Pyrimidine Biosynthesis Pathway

Pyrimidine nucleosides, including *uridine*, *cytidine* and *thymidine*, play a crucial role in cellular metabolism since they are fundamental precursors for RNA and DNA synthesis. Cells rely on three sources to fulfil their need in pyrimidines:

1. the "*salvage* pathway": salvages pyrimidine nucleosides resulted from DNA and RNA molecules degradation;
2. passive and active transmembrane transporters which enable the uptake of pyrimidines from extracellular body fluid;[1]
3. the "*de novo* biosynthesis pathway" which allows the production of pyrimidines from small metabolites in six different steps.

The relative contribution of the *"salvage"* and *"de novo"* pathways depends on cell type and on their developmental stage. The "salvage pathway" together with the pyrimidines absorption from extracellular fluids are usually enough to satisfy the needs of pyrimidines of non-dividing cells. These mechanisms are not so efficient to provide the right amount of pyrimidines for proliferating cells. Therefore, the *de novo* biosynthesis is essential for proliferating cells to meet their large demand for nucleotide precursors.[2] Pyrimidines biosynthesis is invariably up-regulated in tumours and neoplastic cells and the *de novo* pathway has been linked to the treatment of several diseases including cancers, neurodegenerative disorders, malaria and autoimmune diseases such as rheumatoid arthritis.[3] *De novo* biosynthesis (**Figure 1**) starts when the multifunctional protein CAD catalyses the initial step by physically linking three enzymes: the *carbamoyl-phosphate synthetase* (CPSase), the *aspartate transcarbamylase* (ATCase) and the *dihydroorotase* (DHOase). By doing this, CAD uses two equivalents of ATP, one equivalent of bicarbonate, one of L-glutamine and one of L-aspartate in a three steps reaction to synthetase dihydroorotate. The first enzyme generates carbamoyl phosphate from bicarbonate and the amide nitrogen of glutamine (usually produced from the hydrolysis of the side chain), the second generates carbamoyl aspartate via condensation of carbamoyl phosphate with aspartate. Finally, the third enzyme generates dihydroorotate by a ring closure. The fourth enzymatic step is catalysed by the *human* dihydroorotate dehydrogenase (*h*DHODH) which, from biochemical

and microscopic studies,[4, 5] it is found to be bound to the inner membrane of mitochondria where it converts dihydroorotate (DHO) to orotate (ORO). The electrons are transferred directly to the respiratory chain via ubiquinone (CoQ). At the end, the multifunctional UMP synthase uses orotate to produce UMP, a common precursor of all other pyrimidine nucleosides.[6] UMP synthase is a bifunctional protein that catalyses the last two steps of *de novo* pyrimidines biosynthesis. The mammalian protein consists of a orotate phosphoribosyltransferase (OPRTase) that catalyses the transfer of phosphoribosyl pyrophosphate (PRPP) to orotate forming orotidine-5'-monophosphate (OMP) and a orotidine-5'-phosphate decarboxylase (ODCase) that decarboxylates OMP forming UMP.[2]
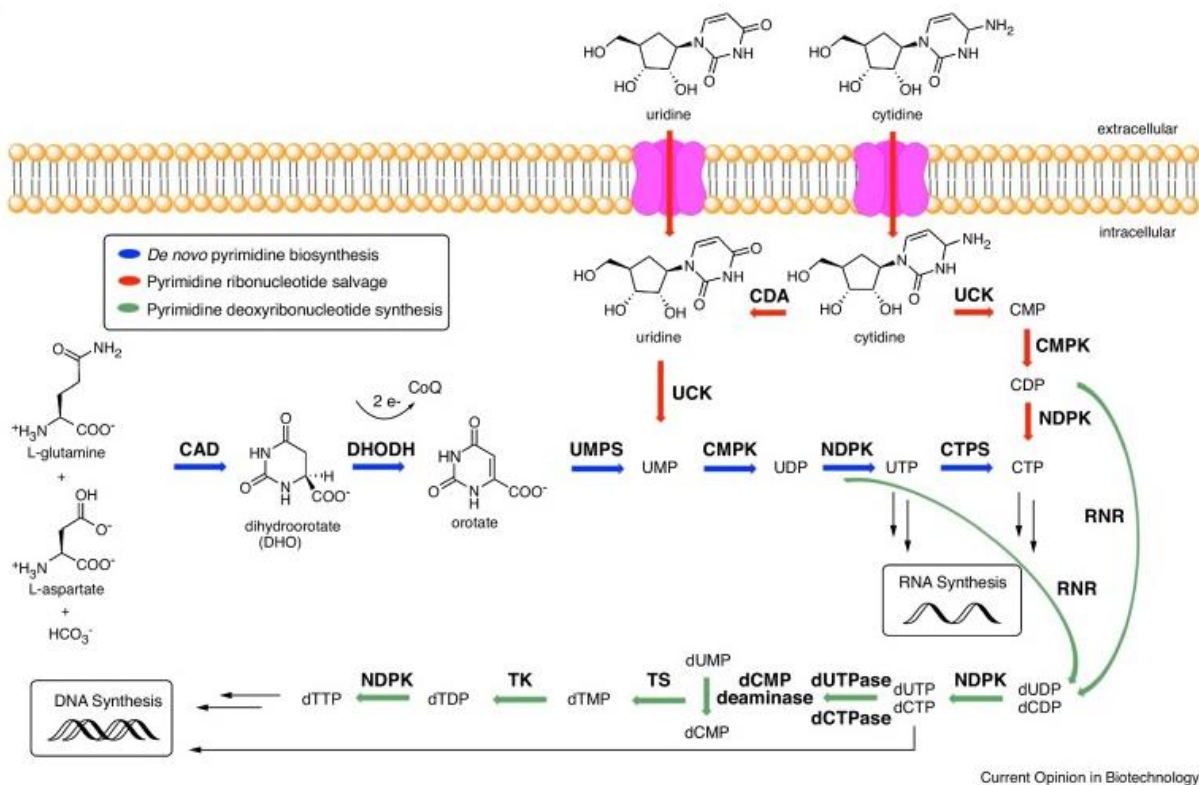
Figure 1. De novo and salvage pyrimidine biosynthesis pathways. OMP: orotidylate, UMP: uridine monophosphate, UDP: uridine diphosphate, , CMP: cytidine monophosphate, CDP: cytidine diphosphate, CTP: cytidine triphosphate, dUDP: deoxy-UDP, dCDP: deoxy-CDP, OPT: orotate phosphorybosyltransferase, UMPS: uridine monophosphate synthetase, CMPK: cytidine monophosphate kinase, NDPK: nucleoside-diphosphate kinase, CTPS: CTP synthetase, CDA: cytidine deaminase, UCK: uridine/cytidine kinase.

1.1.2. Dihydroorotate Dehydrogenase

Human dihydroorotate dehydrogenase belongs to the class 2 DHODH family which has flavin mononucleotide (FMN) as its redox cofactor. *h*DHODH is located in the inner mitochondrial membrane (Fig. 2), where it catalyses the conversion of (S)-dihydroorotate (DHO) to orotate (ORO), in de novo pyrimidine biosynthesis.
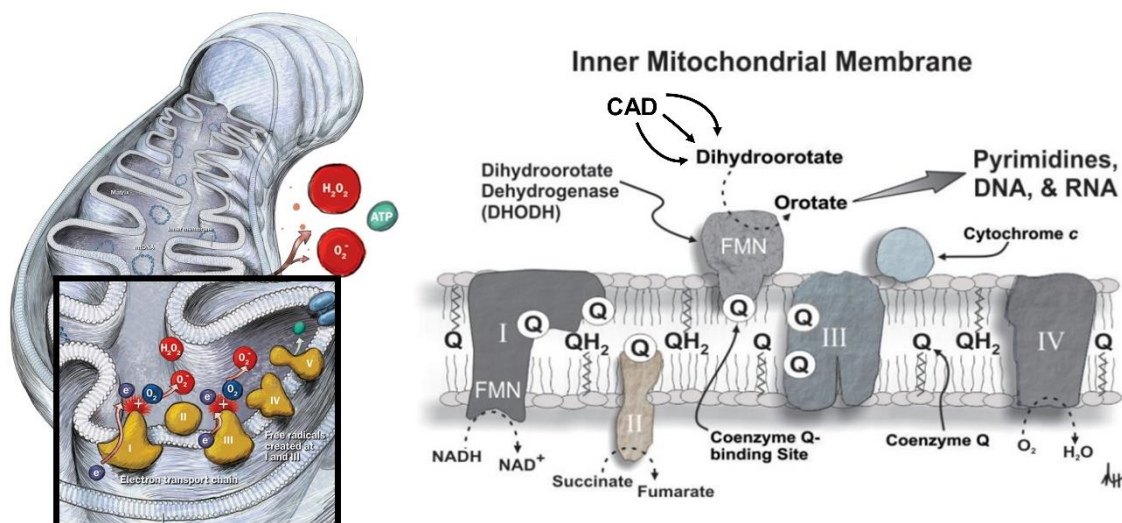

Figure 2. Scheme of respiratory chain inside the mitochondria and enzymes connected to hDHODH. CAD is a multifunctional protein located in the cytosol, which include the other enzymes of the pathway. I, II, III and IV represent enzyme complexes of the respiratory chain.

The kinetic of the oxidation of DHO to orotate ORO takes place through a ping-pong mechanism between two *h*DHODH redox sites (Fig. 3).[7] In the first half of the reaction, electrons are transferred from DHO to the flavin mononucleotide moiety (FMN), which is converted to dihydroflavin mononucleotide (FMNH$_2$), leading to the oxidation to ORO. In the second half of the reaction, ororotate is released before a second cofactor, called coenzyme Q (ubiquinone) binds in a second site and re-oxidizes the FMNH$_2$ to FMN. Electrons are fed into the respiratory chain, with oxygen being the ultimate electron acceptor.[8] Structural studies have revealed two distinct binding sites for DHO/ORO and ubiquinone, respectively. In the de novo pyrimidine biosynthetic pathway, DHODH catalyzes a committed step and thus presents a desirable target for halting pathway flux.
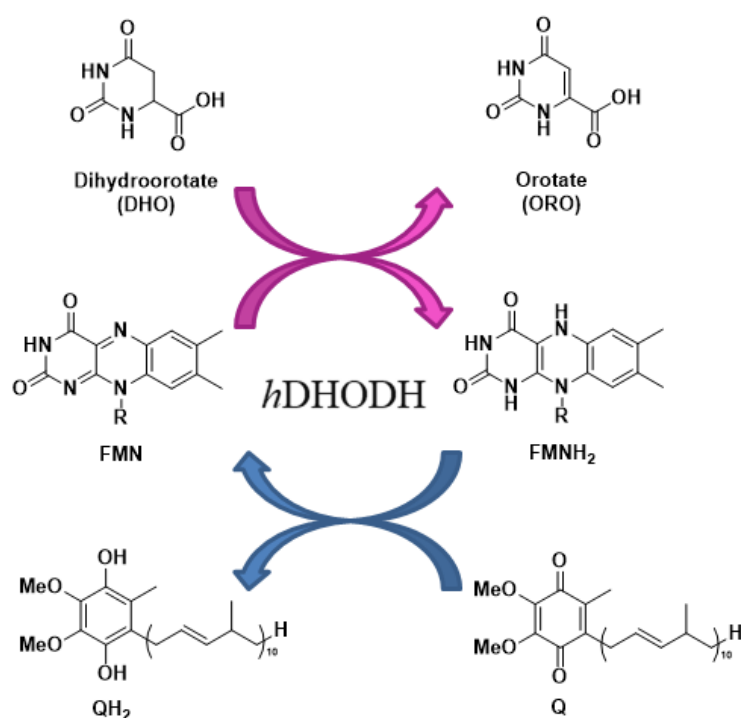
Figure 3. Scheme of DHO (natural substrate) Oxidation to ORO catalysed by hDHODH. FMN and Q are the two co-factors used to complete the cycle.

1.1.3. *h*DHODH Morphology

Human DHODH is composed of two domains: a large C-terminal domain (Met78 - Arg396) consisting of an α/β barrel connected by an extended loop to a smaller N-terminal helical domain (Met30 - Leu68). The cytosolic C-terminal contains the redox site, which is formed by the ORO binding pocket and the cofactor FMN binding site. The small N-terminal domain consists of two α helices, α1 and α2 connected by a short loop. It is used by the enzyme for its connection with the internal mitochondrial membrane and creates a tunnel used by the second cofactor ubiquinone to reach FMN. The presence of this tunnel is specific for class 2 DHODH family (Fig. 4) leading the Q active site as target site of class 2 DHODH inhibitors.[9] In fact, family 1 DHODHs simply have no equivalent tunnel, and the edge of FMN is exposed to solvent. The main differences between class 1 and class 2 DHODH enzymes are:

- The different base in the catalytic site: class 1: Cysteine, class 2: Serine;
- The electron acceptor used: class 1: fumarate / NAD^+, class 2: ubiquinone from respiratory chain;
- N-terminal is more extended in all class 2 than class 1 enzymes.[10]

The reactions catalysed by the enzymes are very different in the two classes because of the nature of the electron acceptors and because of how they interact with the reduced FMN. The variations in the catalytic mechanism of DHODHs from different organisms make them interesting drug targets: i.e. DHODH from one species may be selectively inhibited without affecting the DHODH from another organism. From structural alignment it is known that, despite different residues as active bases, DHODHs from different classes share similarities regarding FMN pocket and pyrimidine binding site. FMN and orotate form a complex and conserved network of hydrogen bonds and hydrophobic interactions with DHODH.
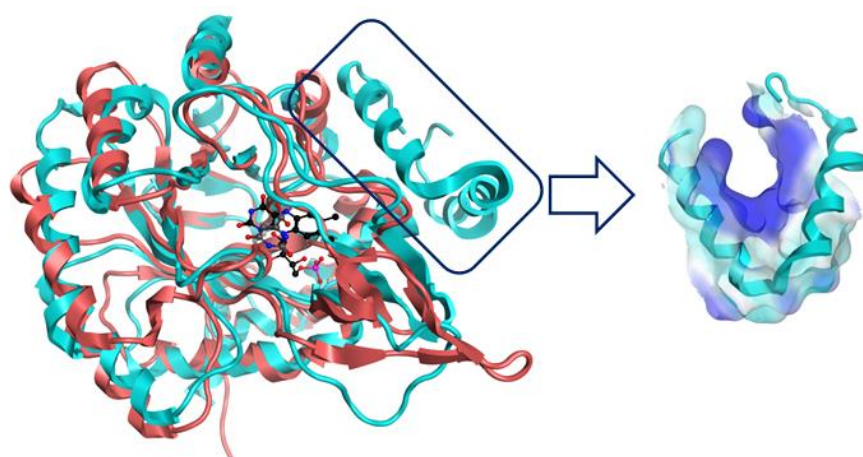
Figure 4. On the left the superimposition of class-1 (red) and class-2 (blue) DHODH structures. Only class-2 DHODHs have an N-terminal domain, enlarged on the right side of the figure.

1.1.4. *h*DHODH Binding Pocket

The N-terminal domain topography and its specificity make this location a proficient target for developing *h*DHODH inhibitors. Baumgartner *et al.* were the first to rationalize the binding pocket topography,[11] dividing it into five sub-sites, according on the intrinsic characteristics of the amino acids involved (Fig. 5). The binding tunnel, named subsite 1, almost exclusively consists of hydrophobic amino acids (residues Leu42, Met43, Leu46, Ala55, Phe58, Leu62, Leu68, Phe98, and Leu359) and this is in agreement with the fact that helices R1 and R2 are involved in membrane association. the narrow end of the tunnel forms a rather polar environment near the redox site, which presents some charged and polar residues (Gln47, Arg136, Tyr356 and Thr360) on both sides (subsites 2 and 3), where the inhibitors can establish salt bridges or hydrogen bonds. The small hydrophobic pocket named subsite 4 is formed from the side chains of Val134 and Val143 and shuts the end of the tunnel. On the opposite side of subsite 4, subsite 5 points the beginning of the lipophilic area.
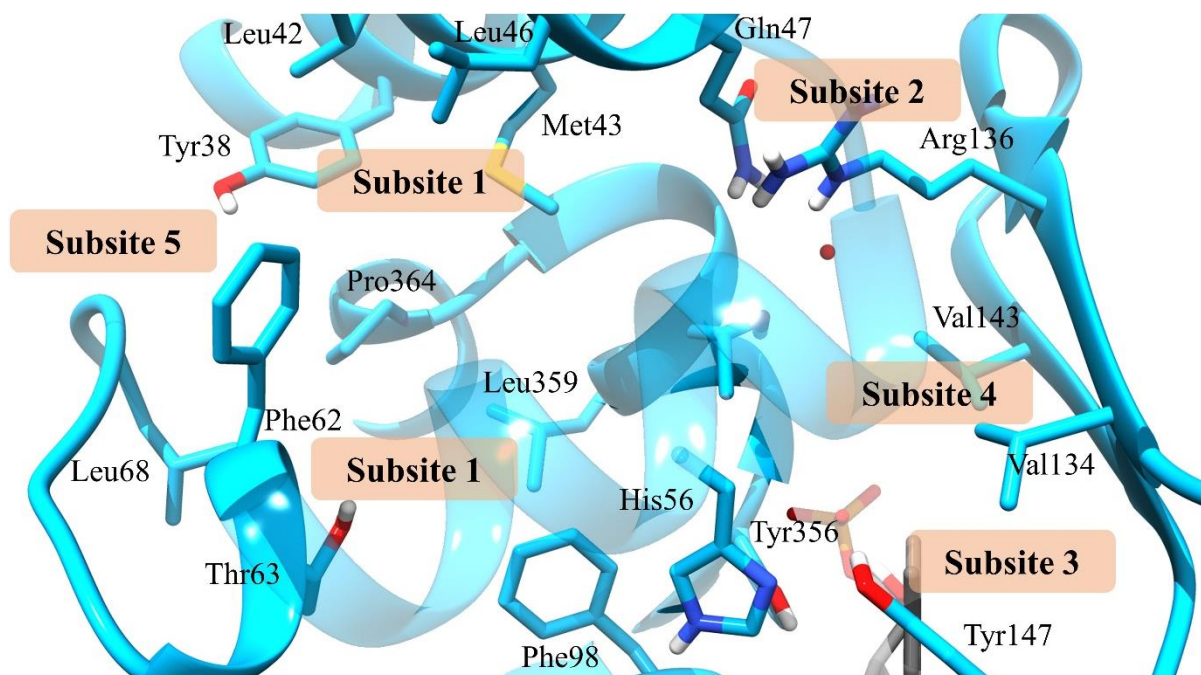
Figure 5. Ubiquinone binding sites of *h*DHODH from X-ray structure 6FMD after the depletion of ligand meds433. Flavin mononucleotide is represented in grey.

1.1.5. Therapeutic Potential of Dihydroorotate Dehydrogenase Inhibitors

Dihydroorotate dehydrogenase is a vital enzyme in the *de novo* pyrimidine biosynthesis pathway, it is crucial for the proliferation of living entities. Pharmacological inhibition of this pathway provides a selective approach to target cells undergoing rapid cellular growth, offering therapeutic opportunities for the treatment of autoimmune diseases, such as rheumatoid arthritis, multiple sclerosis and cancer.[9] The first approved human DHODH inhibitor was Leflunomide,[12] marketed as Arava© by Sanofi-Aventis (Fig. 6), an antirheumatic drug which was approved for the therapy of rheumatoid arthritis and other autoimmune diseases more than 20 years ago. The drug undergoes metabolic isoxazole ring opening to form the active metabolite teriflunomide (A77 1726) (Fig. 6).[13] The latter was developed under the trade name Aubagio® for the treatment of multiple sclerosis.[14] As mentioned before, leflunomide is a disease-modifying anti-rheumatic drug (DMARD) used in the second line treatment of rheumatoid arthritis. Its clinical efficacy and tolerability are approximately comparable to methotrexate and sulfasalazine, but the onset of treatment efficacy occurs more rapidly with leflunomide than with other active agents.[15] *h*DHODH was linked with cancer nearly six decades ago when Smith *et al.* noted elevated DHODH activity in leukemic cells.[16, 17] In the following years several efforts were involved on DHODH studies that culminated in the discovery and clinical evaluation of brequinar, a potent *human* DHODH inhibitor.[18, 19] However, brequinar was unsuccessfully evaluated against several

tumour categories due the conflicting results obtained in its early clinical evaluation. Moreover, brequinar presented a poorly chosen dosing regimens[20], and for this reasons it did not meet the requirements to pass clinical trials. Though, recent studies have uncovered new prospective in this field of cancer treatments, linking *h*DHODH inhibition with the induction of cellular differentiation.[21] In particular, brequinar induced significant myeloid differentiation, delayed disease development and reduced the burden of leukemia-initiating cells in the various AML mouse models, including human cell line xenografts, patient-derived xenografts and syngeneic mouse models. The authors patented these results on the US market in 2017.
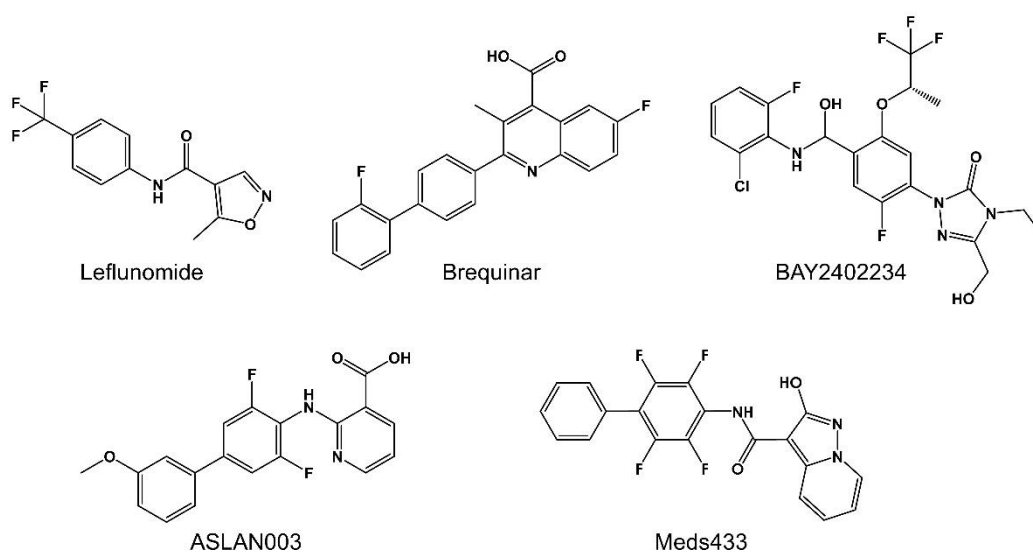

Figure 6. *h*DHODH inhibitors

This novel promising finding sparked the interest on *h*DHODH inhibitors, and several new molecules were designed. Currently, two newly patented *h*DHODH inhibitors are in clinical trials for the treatment of AML (Fig. 6); ASLAN003,[22] Phase II of clinical trials, and BAY2402234,[23] Phase I of clinical trials. In 2018 a novel class of *h*DHODH inhibitors based on a carboxylic group bioisostere 2-hydroxypyrazolo[1,5-a]pyridine (Meds433, Fig. 6), were found to induce myeloid differentiation in AML cell lines, leading to the massive death of leukemic cells.[24] Although important progresses have been made on developing DHODH inhibitors, the mechanism of inhibition-induced differentiation linked to this protein is not yet fully understood. Some studies suggest the differentiation phenomenon may be caused by pyrimidine depletion. In fact, the inhibition of OMP decarboxylase induces differentiation and indicates pyrimidine depletion as a reasonable mechanism.[21] Additionally, the presence of uridine was able to restore the differentiation caused by DHODH inhibition.[24] A potential

connection between pyrimidine depletion and cellular differentiation may be transcriptional elongation. Pyrimidine depletion has been observed to inhibit transcriptional elongation of tumorigenic genes,[25] and in fact, a similar effect is seen inhibiting DHODH in melanoma cells.[26] Despite this, developing DHODH-targeted therapy raising considerable interest among pharmaceutical industry, giving hope for a successful treatment against AML.

1.1.6. Acute Myelogenous Leukemia (AML)

Leukemia is the result of the rapid overproduction of abnormal white blood cells and it is one of the most common type of blood cancer. It occurs when abnormal white blood cells in the bone marrow quickly increase and destroy normal blood cells leaving a person prone to infection. The four most common types of leukemia are:

1. Acute Lymphocytic Leukemia (ALL): the most common in children, it begins in the B or T lymphocytes which are the immature white blood cells;
2. Acute Myeloid Leukemia (AML): the most common in adults;
3. Chronic Lymphocytic Leukemia (CLL): usually affects older adults and it begins in B lymphocytes;
4. Chronic Myeloid Leukemia (CML): rare type of leukemia, it occurs when a genetic change turns the myeloid cells into immature cancer cells.
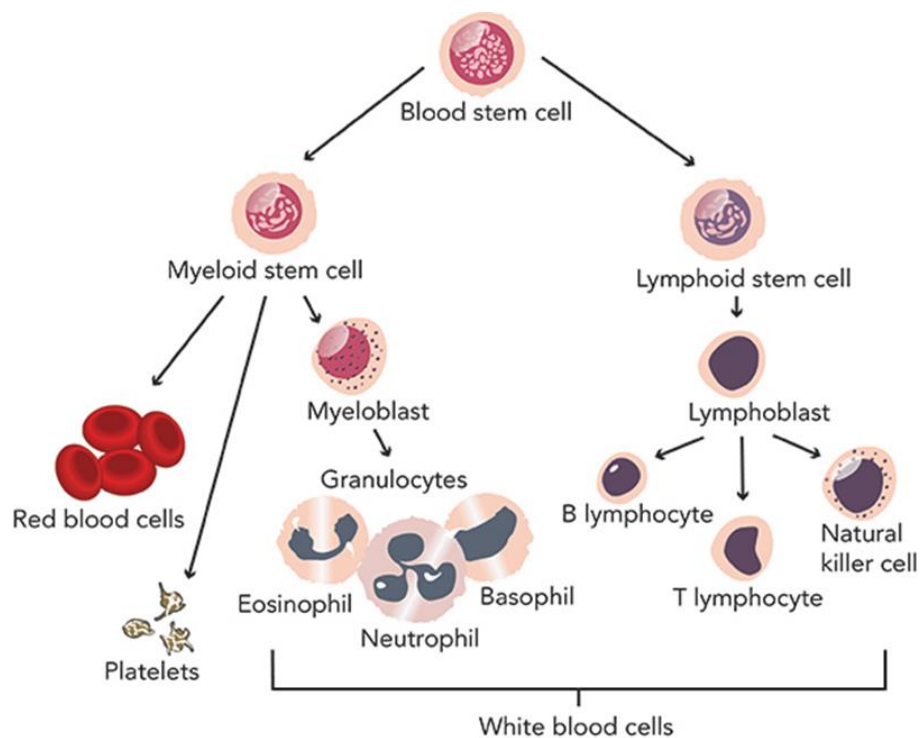
Figure 7. Normal haematopoiesis

14

The chronic form moves more slowly than the acute one and it involves mature cells that might work for a time as they are designed before failing. The acute form moves more rapidly and needs aggressive treatment, it involves immature blood cells that may never mature to work as they are designed. The treatment approaches for this clinically destroying disease depend on the type of white blood cell affected and whether the disease is acute or chronic. All pluripotent cells in the bone marrow proliferate into 2 major cell lineages: the myeloid cells which include erythrocytes, granulocytes and platelets, and the lymphoid cells which include B and T lymphocytes. **Figure 7** shows the normal haematopoiesis pathway: myeloid cells proliferate into their mature cells within the bone marrow, whereas the lymphoid precursors migrate to the lymphoid organs (i.e., lymph nodes, spleen and thymus) to complete maturation. In AML, multiple genetic mutations entail both myeloid stem cells to mature into abnormal myeloblasts and unchecked proliferation. As they multiply, they overwhelm the normal cells in the bone marrow and blood. AML is a clinically devastating heterogeneous disease because mutations can led to abnormality in any step in the cell maturation process (Fig. 8) and even with improvements in diagnosis and supportive care, 5-years survival rate of an adult with AML is only 30%.[21] General treatment involves chemotherapy, targeted therapy drugs, surgery, radiation therapy and stem cells transplant. But its treatment difficulties are mainly due to resistance to drugs, including primary resistance, which exists before the treatment, and secondary resistance, which develops because of repeated use of some induction chemotherapeutic agents.[27] Standard treatment for AML includes intensive induction chemotherapy and consolidation therapy, the composition of induction therapies has remained largely unchanged over more than four decades. For young patients the intensive cytarabine and anthracycline regimen induction therapy is standard of care.[28, 29] Instead, consolidation therapy is used to prevent relapse by eradicating minimal residual disease still present in the bone marrow after induction therapy. There are two main options for consolidation therapy: chemotherapy (including targeted agents) and hematopoietic stem cell transplantation (HSCT).[29] Differentiation therapy is receiving increasing attention due to the remarkable activity of all-trans retinoic acid (ATRA) in combination with arsenic trioxide (Trisenox®) in patients with acute promyelocytic leukemia (APL). The concept of differentiation therapy, in contrast to cytotoxic chemotherapy, aims to restore the normal process of hematopoietic maturation from self-renewing progenitor to terminal differentiated effector cells. Differentiation therapy has revolutionized the treatment of patients with APL.[30] APL is an aggressive type of acute myeloid leukemia in which there are too many immature blood-forming cells (promyelocytes) in the blood and bone marrow.
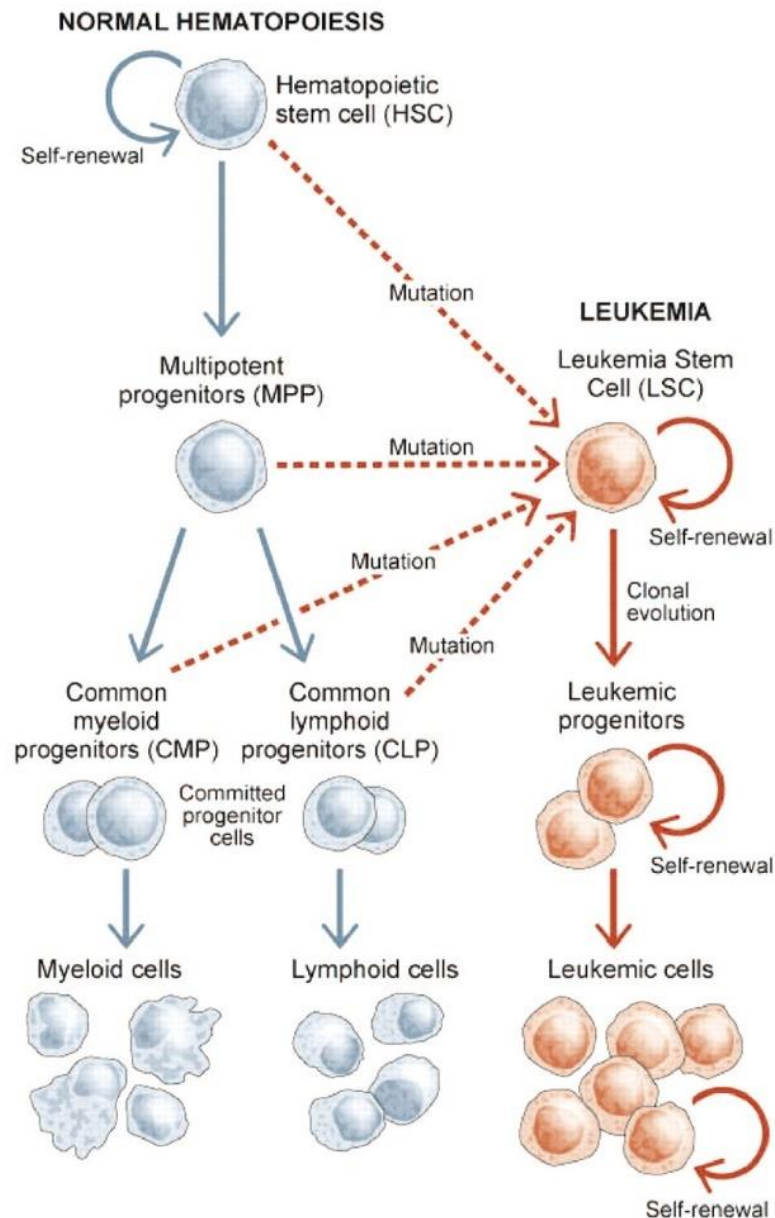
Figure 8. All leukemic stem cells derive from a single abnormal progenitor cell.

APL is caused by a mutation that is acquired over a person's lifetime, usually involving a translocation between chromosomes 15 and 17. This translocation results in fusion oncoproteins involving the retinoic acid receptor. The differentiation treatment may include the use of ATRA which releases the cells from differentiation arrest and permits their normal maturation, bringing the overall survival rates for APL patients in excess of 85%. The results of differentiating therapy in APL patients has encouraged the use of differentiating therapy in patients with AML.[31] At the end of 2016, two publications[21, 30] demonstrated the central role that *h*DHODH plays in AML. Sykes *et al.*,[21] suggested that *h*DHODH can induct myeloid differentiation both *in vitro* and *in vivo* models, delaying the disease development and reducing the burden of leukemia-initiating cells in various AML murine models and *human*

cell line xenografts. One hallmark of AML is that the leukemic blast is arrested at an early stage of differentiation. The recognition that leukemic blasts were frozen at an immature stage of development suggested that new therapies might be directed at promoting differentiation allowing the leukemic blasts to resume their normal maturation to terminally differentiated neutrophils.

1.1.7. Meds433

Meds433 is a potent and selective hDHODH inhibitor with a IC$_{50}$ of 1.2 nM (Fig. 6).[24] Meds433 was designed by scaffold hopping replacement of the quinolone-carboxylate moiety of brequinar with an 2-hydroxypyrazolo[1,5-a]pyridine scaffold. The hydroxypyrazolopyridine motif, connected through an amide bridge to a substituted biphenyl lipophilic moiety, was speculated to play the role of brequinar's carboxylic group by interacting with Arg136 in hDHODH subsite 2 (Fig. 7). The degree of fluorine substitution at the phenyl ring adjacent to the oxadiazole moiety was strongly associated to activity. Moreover, its correlation with stabilization of the compounds bioactive conformations was deeply studied on analogues compounds.[32, 33] Meds433 showed the highest DHODH inhibition among its series and resulted as one of the most potent DHODH inhibitor so far. The X-ray crystallographic structures of meds433 in complex with hDHODH was determined (Fig. 7). The inhibitor was clearly bound in the ubiquinone binding site with a binding mode that can be perfectly superimposed onto the previously reported crystal structure of the complex with brequinar analogue.[24] The tetrafluorobiphenylic moiety of Meds433 is positioned in subsite 1, on the other side of the pocket the hydroxypyrazolo[1,5-a]pyridine substructure is extended over subsites 2 and 3, occupying the innermost part of the cavity. As already mentioned, there is an ion bridge extending to the side chain of Arg136 and a hydrogen bond with the side chain of Gln47. The pyridine motif is able to fit perfectly within the lipophilic subsite 4 (Val134 and Val143). Furthermore, meds433 was extensively investigated on in vitro assay connected to DHODH inhibition. The inhibitor showed a good antiproliferative effect on jurkat T cells, which was reverted by the addition of exogenous uridine, confirming the compound acts as a pyrimidine biosynthesis inhibitor. The immunosuppressive activity was investigated evaluating the effect by the proliferation of phytohemagglutinin-activated peripheral blood mononuclear cells. The result showed a potent inhibition of activated PBMC proliferation reversed by the addition of exogenous uridine. Finally, the effect of Meds433 on AML was evaluated on two tumour cell lines

(U937 and THP1). The lead compound clearly demonstrated to induce myeloid differentiation, leading to the massive death of leukemic cells. Notably, this effect was obtained at a concentration that was 1-log lower than that of the lead brequinar, and was restricted to leukemic cells alone.
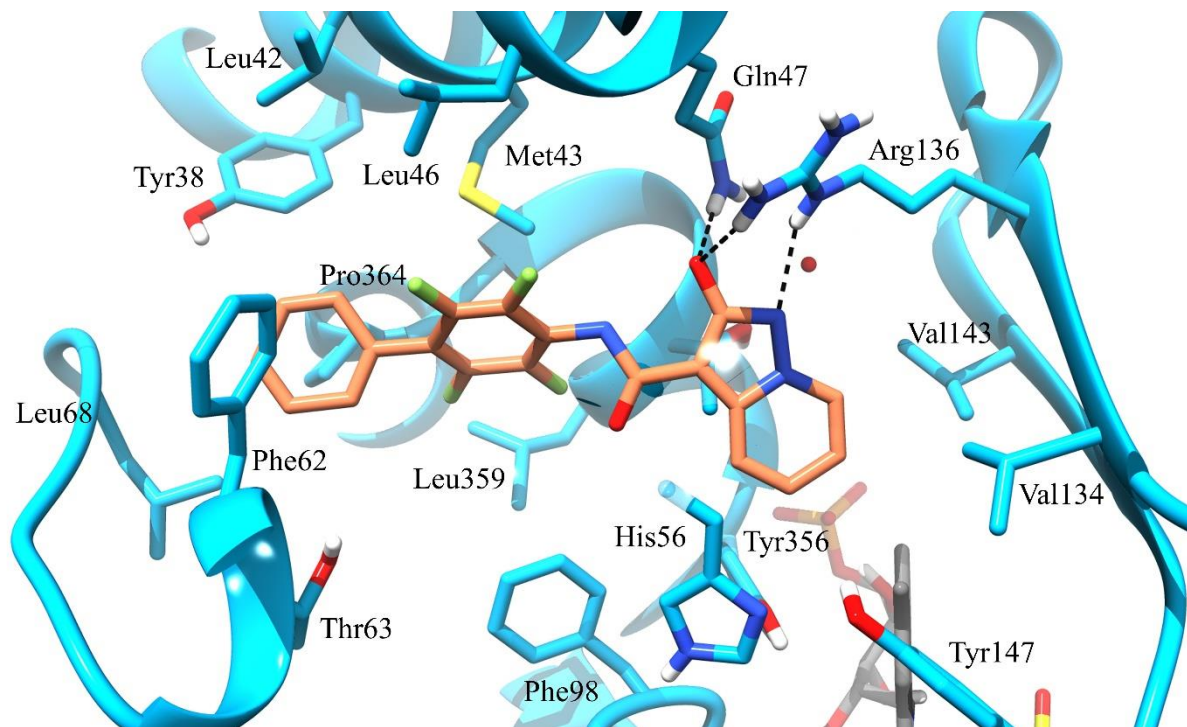
Fig 7. Ubiquinone binding sites of hDHODH cocrystallized with Meds433 (orange) (PDB ID: 6FMD). Hydrogen bonds are shown as dashed lines. Flavin mononucleotide and orotate are represented in grey and yellow, respectively.

1.2. Aim of the Project

Acute myeloid leukemia (AML) is characterized by blasts inability to differentiate into mature cells and to undergo apoptosis.[34] Proapoptotic and differentiation therapies achieve remarkable results in acute promyelocytic leukemia, unfortunately, similarly efficacious differentiation therapies have been lacking for AML so far. The inhibition of *h*DHODH was reported to induce differentiation of diverse AML subtypes. As already discussed, meds433 is a novel DHODH inhibitor and resulted to be effective in the differentiation of AML cell lines, which makes it a promising scaffold for further studies in AML models. This PhD thesis is directed to the computer-driven optimization of meds433, computational methodologies were used to guide the design of new compounds to improve potency and pharmacokinetic properties.

1.3. Results and Discussion

The structure of meds433 can be ideally divided into three substructures (Fig. 8). In the current project, three main studies have been performed on the lead compound in order to optimize its pharmacokinetic and pharmacodynamics properties:

1) Explore the possibility of decorating the A ring of the "pyridine" moiety
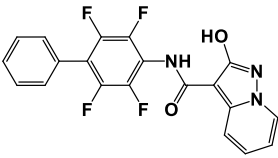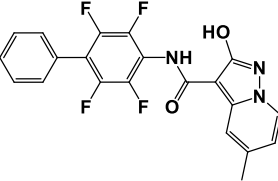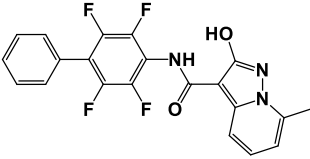2) Replace the tetrafluorobiphenylic scaffold
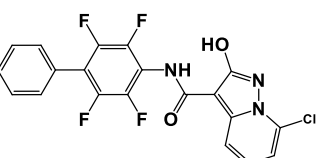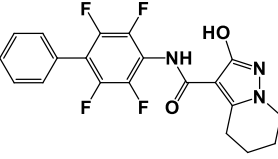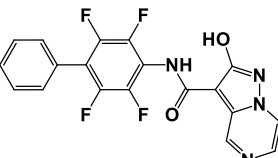3) Investigate modulation on C ring
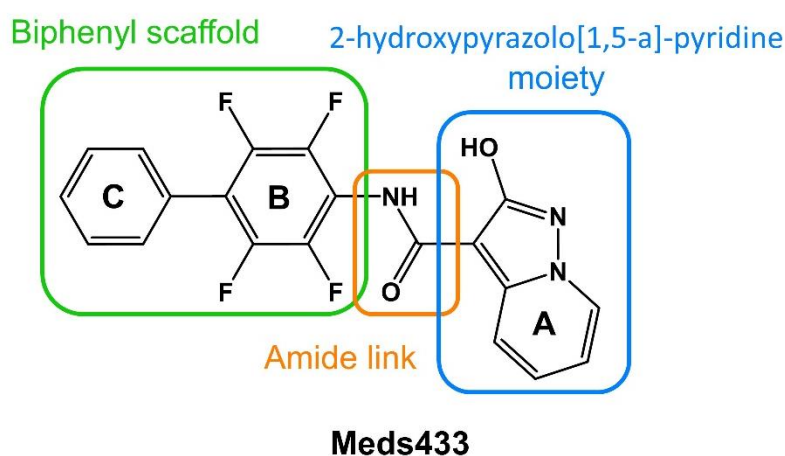
Figure 8. Three main substructure that compose meds433.

1.3.1. Modulation of A Ring

We focused on the possibility of decorating the pyridine motif of meds433 to add an interaction with the small lipophilic pocket created by Val134 and Val143 (subsite 4). Here, the study was performed by means of Molecular Dynamics (MD) free energy perturbation (FEP) (See methods).[35] The four available positions of Meds433, indicated in **Figure 9**, were explored using MD/FEP methyl and chlorine scans to identify the most promising sites for a beneficial hydrogen substitution. The obtained ΔΔG values indicates that chlorine group is generally preferred over methyl in all the positions, since negative values represent an increasing in binding affinity. Among the four sites, position 7 is the most profitable for a substitution as presents the lower energy values, -1.43 kcal/mol for the chlorine and only a marginal effect on the energy for the methyl group (0.15). Replacement of hydrogen in position 6 is the less favorite accord to the higher positive energy values for chlorine (0.87)
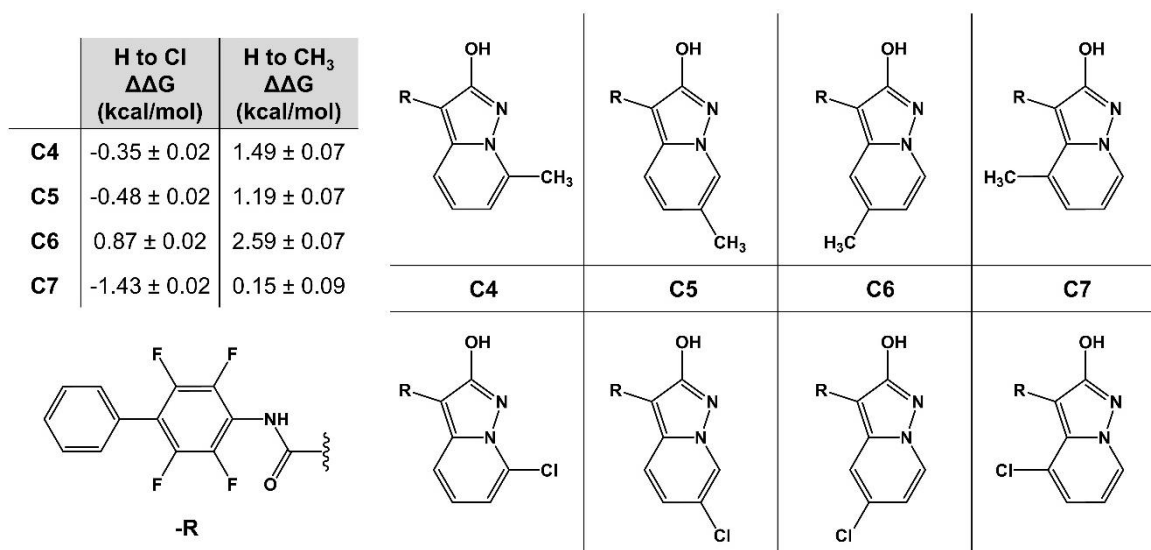
and methyl (2.59). Position 4 and 5 show comparable behavior, substitution in 5 is only slightly better tolerated than in 4 for both groups.

#	Structure	hDHODH[a] IC$_{50}$ ± SE (nM)	log D7.4 ± SDc	Solubility (µM) in PBS
meds433		1.2 ±	2.35 ± 0.02	12
1		35 ± 3	2.47 ± 0.09	2.8
2		4.3 ± 0.5	2.70 ± 0.02	1.4
3		n.d.	2.81 ± 0.13	<LOD
4		5.8 ± 0.4	2.36 ± 0.02	1.4
5		n.d.	n.d.	n.d.

**Table 1.** IC$_{50}$, LogD and solubility of compounds 1 – 5.

Moving to experimental and taking into account MD/FEP results, the two most tolerated positions, 5 and 7, have been considered to be modulated. The derivates with a methyl substituent in positions 5 and 7 were synthetized, leading to compounds **1** and **2** (Tab. 1), respectively. Also, the most promising chlorine derivate compound **3** was synthetized, but its activity was not yet available at the time this manuscript was written. Whereas the substitution of position 5 decreased the activity by 25-fold (**1**, IC$_{50}$ = 35 nM), as compared to meds433, the substitution of position 7 provided a similar profile (**2**, IC$_{50}$ = 4.3 nM). The

reduction of A ring (compound **4**) was studied, and it led to a lower binding affinity (IC$_{50}$ = 5.8 nM) together with a lower solubility. Moving further to catch additional interactions, sub-site 3 is composed by two hydrophilic amino acids, Tyr147 and Tyr356, the side chain of this latter is on the right distance from a possible H-bond acceptor on the pyridine moiety of meds433. Looking at the crystallographic structure of meds433 in complex with *h*DHODH, the oxygen atom in Tyr356 side chain is located 3,43 Å away from C6 of A ring, a suitable distance to form H-bond.

	H to Cl ΔΔG (kcal/mol)	H to CH$_3$ ΔΔG (kcal/mol)
C4	-0.35 ± 0.02	1.49 ± 0.07
C5	-0.48 ± 0.02	1.19 ± 0.07
C6	0.87 ± 0.02	2.59 ± 0.07
C7	-1.43 ± 0.02	0.15 ± 0.09

Figure 9. MD/FEP results of the change in calculated free energy of binding (in kcal/mol), and the computed uncertainty, for the introduction of chlorine and methyl substituents on the 2-hydroxypyrazolo[1,5-a]pyridine motif of Meds433.

A docking study of compound **5**, the analogue of meds433 with a nitrogen atom in position 5 of A ring, clearly showed the additional electrostatic interaction (Fig. 10). Moreover, 50 ns of Molecular Dynamics were performed on the complex compound **5** – DHODH. During the simulation the H-bond interaction between Tyr356(OH) and **5**(N) was maintained in around 80% of frames. In addition, the molecular mechanics energies combined with the generalized Born and surface area continuum solvation (MMGB-SA) method was used to estimate **5** free energy of binding. The result claims the additional electrostatic interaction lead to higher FEB for compound **5** (46.32 kcal/mol) with respect to the meds433 (40.23 kcal/mol). Unfortunately, experimentally data for **5** were unavailable to confirm *in silico* predictions, at the time this manuscript was being drafted.
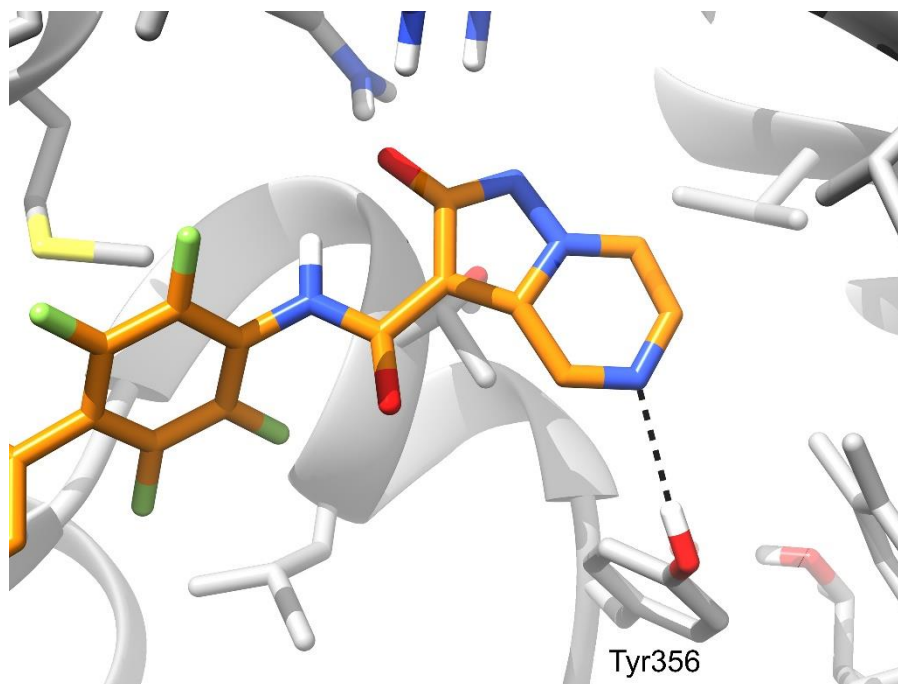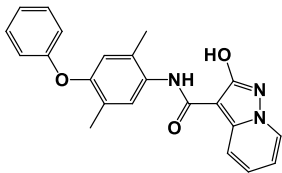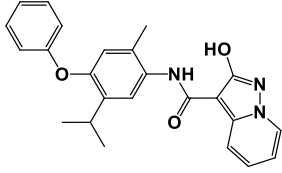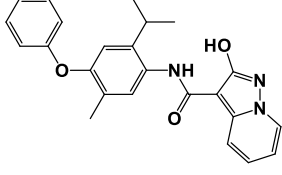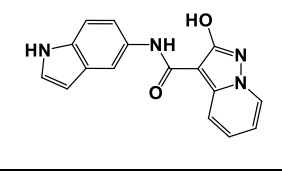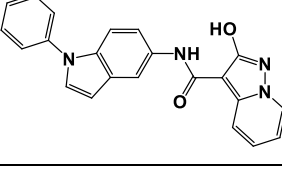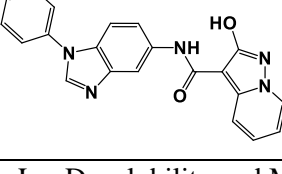
Figure 10. Predicted binding mode of compound **5.** Hydrogen bond is shown as dashed lines.

1.3.2. Alternatives to Biphenylic Scaffold

In the next series of compounds, the replacement of the biphenylic substructure was investigated in an attempt to improve pharmacokinetic properties and obtain more drug-like compounds. In the previous studies,[24, 33, 36] optimal interaction with lipophilic subsite 1 was only guaranteed with tetrafluoro substitution on the first ring.

#		hDHODH^a IC$_{50}$ ± SE (µM)	log D7.4 ± SDc	Solubility (µM) in PBS	MM-GBSA ± SD (kcal/mol)
Meds433		1.2 ±	2.35 ± 0.02	12	41.06 ± 0.27
6		0.760 ± 0.136	2.30 ± 0.02	47	-35.19 ± 1.30
7		0.480 ± 0.031	2.75 ± 0.01	7.0	-38.52 ± 1.88

		IC50	LogD		MM-GBSA
8		0.043 ± 0.006	2.93 ± 0.09	2.5	-41.52 ± 1.68
9		0.100 ± 0.009	2.91 ± 0.07	n.d.	-50.83 ± 3.58
10		0.0069 ± 0.0009	3.18 ± 0.09	n.d.	-47.35 ± 0.96
11		> 10	0.45 ± 0.02	n.d.	-35.70 ± 0.62
12		2.4 ±	2.89 ± 0.05	n.d.	-43.62 ± 1.24
13		> 5	2.55 ± 0.04	n.d.	-38.77 ± 3.16

Table 2. IC_{50}, LogD, solubility and MM-GBSA results of compounds **6** – **13**.

Conformational analyses[37] underlined the role of incremental fluorine substitution on the first ring in stabilizing the "brequinar-like" binding mode, which has previously been found to be connected with higher inhibitory potency.[38] For example, the removal of two or three fluorine atoms from the biphenylic scaffold of triazole analogue **2** resulted in a dramatic drop in inhibitory activity.[39] However, the presence of the tetrafluorobiphenylic substituent is detrimental for the solubility of our derivatives. We therefore decided to design analogues that lacked a biphenylic scaffold in order to investigate novel possibilities. Inspiration was taken from a study by Das *et al*,[40] in which "brequinar–like" activity was replicated in a brequinar-related compound which included a substituted diphenylether; compound **6** was thus designed. Although a dramatic drop in activity was observed in moving from meds433 (IC_{50} = 1.2 nM), to **6** (IC_{50} = 760 nM), the biphenylic replacement seemed to be required for activity to occur. In order to better understand the role played by substitution on the first

diphenylether ring in influencing activity, a MD study was performed. Starting from the docked conformation of the compound **6** in the target binding site, 50 ns long MD simulation were carried out. The root mean square deviations (RMSD), for ligands and protein α carbons were calculated with respect to the starting structures. The protein structures reached stability after a few ns of simulation and then remained stable with an RMSD of around 0.95 Å. More interesting still is the RMSD changing observed for the compound inside the protein. **Figure 11** shows the average conformations of compounds **6 - 10,** assumed during the MD simulations and the relative RMSD plots. While meds433 shows a stable conformation, with an average RMSD of around 0.6 Å (Fig. 11A), a more highly fluctuating RMSD was observed for compound **6** (**Figure 11B**). This molecule principally assumed two different conformations during the simulation. One is similar to the starting docking pose, identified by an RMSD of around 0.6 Å, and, in the other, the distal phenyl of the diphenylether points towards Tyr38, leading to an RMSD of around 2.0 Å. As far as activity values are concerned, the alternating behavior observed in **6** results in a loss of interaction, which is reflected in reduced activity. Once these considerations taken into account, a series of derivatives substituted on the first ring were designed to avoid the rotation of the distal ring. Compounds **7** and **8** (Fig. 11C and 11D), present a more stable profile, after an initial increase in RMSD, the compounds reached a steady conformation with average RMSDs of around 1.6 and 1.2 Å, respectively. Activity values appear to indicate that the additional degree of freedom presented by diphenylether is strictly correlated with the compounds' binding affinity. Indeed, the introduction of methyl substituents, in compounds **7** and **8**, seems to stabilize the favored binding conformations leading to an optimized interaction with the protein. It is worth noting that the activity of **7** (IC_{50} = 480 nM), which is only slightly higher than that of **6**, clearly indicates how a single substitution is not enough to achieve the rigidity required for optimal binding affinity. On the other hand, the double methyl substitution in compound **8** (IC_{50} = 43 nM) is able to recover the activity, which reaches a nM range, suggesting that molecule rigidity is one of the key attributes for *h*DHODH structure-activity relationship (SAR). Although MD analysis gives important insights into the role of the diphenylether moiety in the binding to *h*DHODH, the free energy of binding (FEB) was still required for the understanding of the structure-activity relationships. MM/GBSA calculations were used to provide a quantitative way to evaluate the different components of interaction energy that contributes to binding of compounds. The MM/GBSA calculations show a good agreement with the experimental activity data.
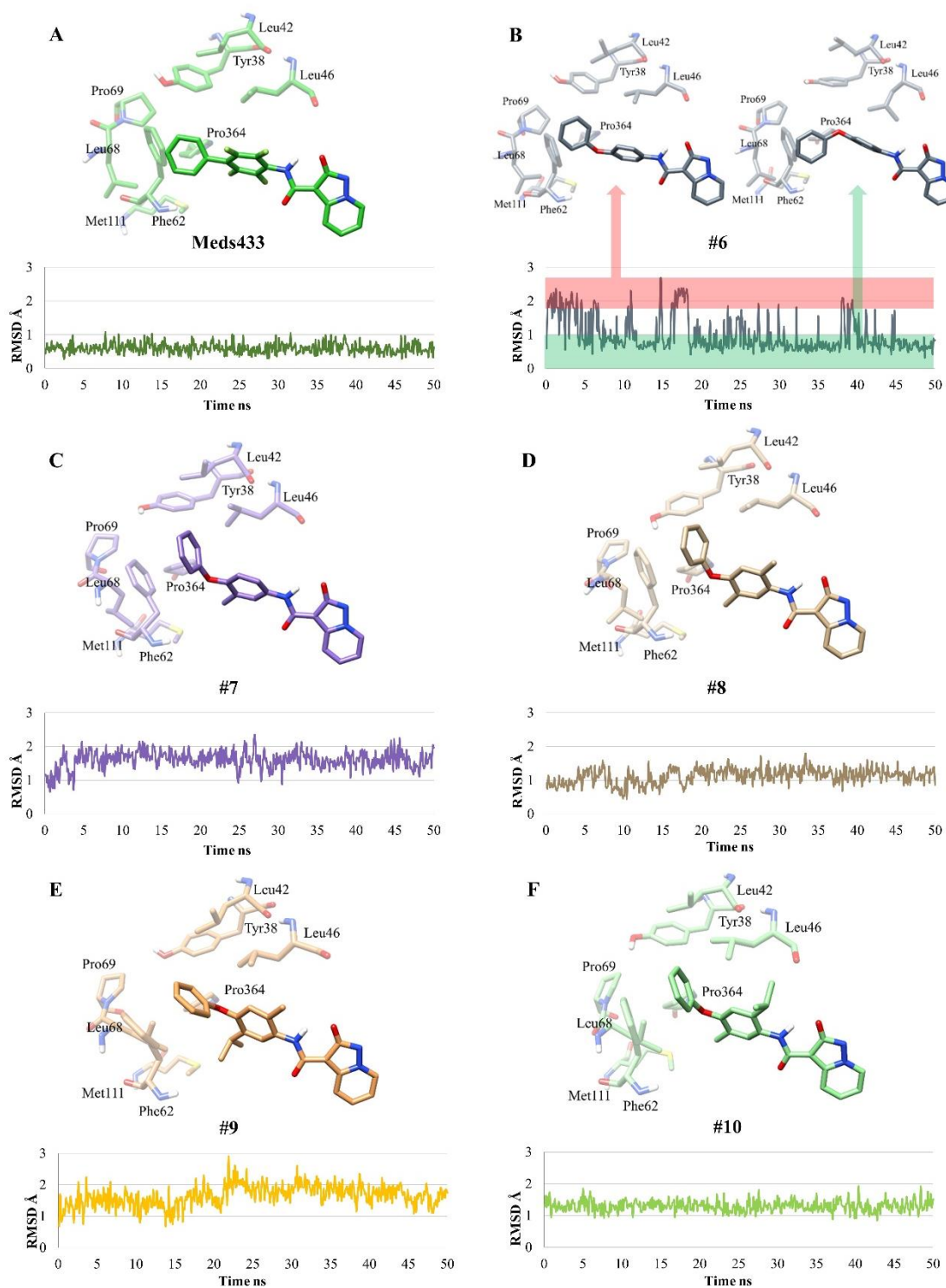
Figure 11. Graphs show the evolution of the RMSD values (calculated for ligand heavy atoms as compared to starting structures) of the *h*DHODH complex analyzed by MD simulation. A, B, C, D, E and F in the figure correspond to compounds Meds433, 6, 7, 8, 9 and 10, respectively. The compounds and surrounding amino acids are in the average conformations assumed during MD simulations.

According to the experiments, the difference in FEB of compound meds433 (ΔTotal = -40.87 kcal/mol) and **6** (ΔTotal = -36.08 kcal/mol) is found to be more than 4 kcal/mol. Moving to

compounds **8** (ΔTotal = -37.18 kcal/mol) and **9** (ΔTotal = -40.32 kcal/mol), obtained by modulation of the first ring of compound **6**, an improvement of the ΔTotal can be observed. While the computed free binding energy for compound **7** is still close to **6**, compound **8** shows a free energy value close to the lead compound. Therefore, two addition derivates of compound **6** were designed to explore the SAR of diphenylether based compounds.
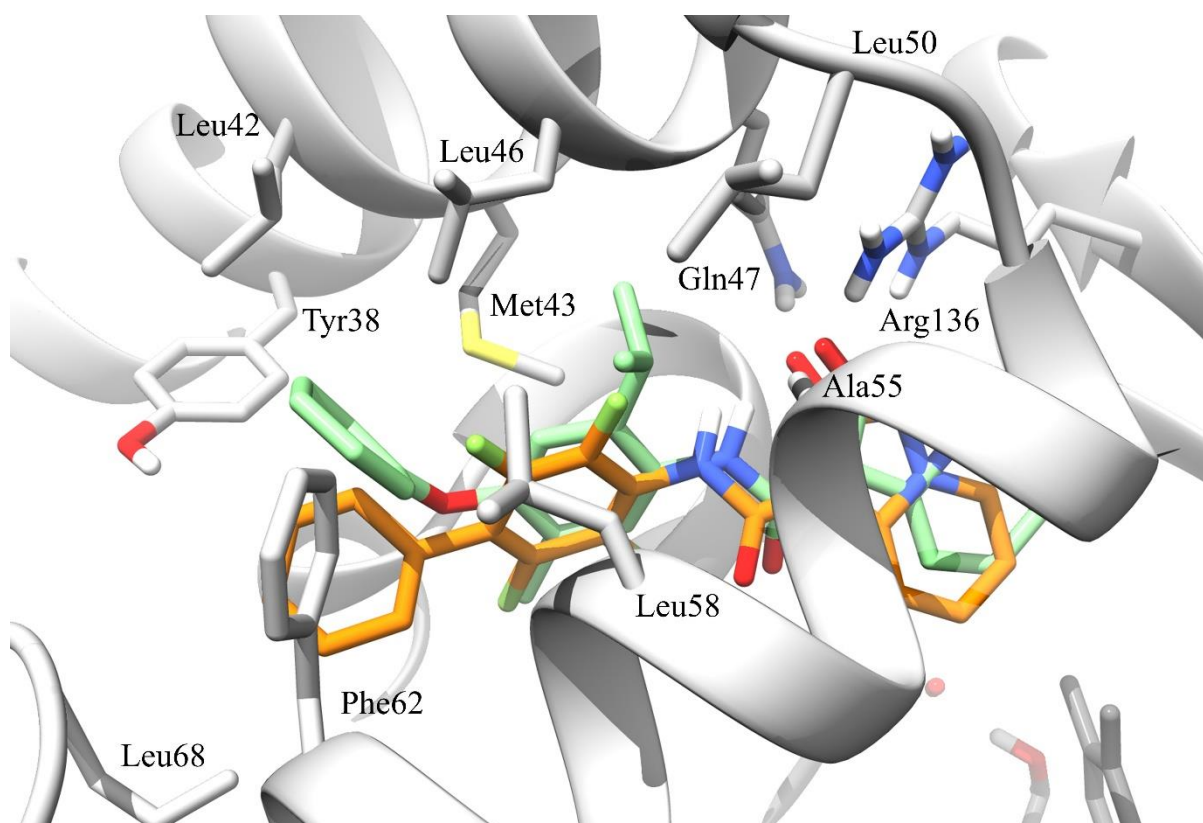
Figure 12. Ubiquinone binding sites of *h*DHODH with bound compound meds433 in orange (PDB ID: 6FMD), superimposed on the average structure of the *h*DHODH-compound **10** complex produced in MD simulation (green). Flavin mononucleotide is represented in gray.

The *o*-methyl and *m*-methyl substituents of **8** were replaced with isopropyl group in compound **9** and **10** (Fig. 11E and 11F), respectively. The predicted FEB increased with respect to compound **8** for both the isopropyl derivates (**9** ΔTotal = -50.83 kcal/mol, **10** ΔTotal = -47.35 kcal/mol,), indicating possible additional lipophilic interaction with sub-site 1. The derivates **9** and **10** were synthetized and tested. Compound **9** showed an IC_{50} of 100 nM, almost 3 fold less than the dimetil derivates **8**. On the other hand, compound **10** instead resulted around 5 times more active than **8** with an IC_{50} of 6.9 nM. The SAR of this series clearly indicate that *meta* position diphenylether requires a small group to avoid the fluctuation of the distal ring, but bulkier groups are unfavorite for the binding. On the contrary, the introduction of larger substituent such as isopropyl on the *orto* position led to a better binding affinity. In fact, the *o*-isopropil fit with a sub-pocket of sub-site 1, formed by

Ala55, Leu50, Leu46, Leu58 (Fig. 12). Moreover, the MD studies performed suggest that there is a preferred binding conformation which depends on the steadiness of the compounds during the MD simulation. **Figure 12** compares the binding modes of Meds433 and **10** in complexes with *h*DHODH. Both molecules show similar behavior, including in their interactions with subsites 2 and 4, where the compounds display the same contact with Arg136 and Gln47. Interestingly, the diphenylether motif in compound **10** is shifted toward the cavity composed by Leu42, Leu46 and Tyr38, with respect to meds433 which places the distal phenyl ring between Phe62 and Pro364. Compared to the tetrafluorinated biphenyl moiety, the diphenylether scaffold could confer to compound **10** and derivates the possibility of reaching a different area of the ubiquinone binding pocket. Inspired by diphenylether compounds, a new series of molecules were designed through a restriction of flexible ether portions into an indole cyclic structure (compounds **11** – **13**). Once the bioactive conformation of diphenylether was identified, it showed that an indole-based scaffold was able to remove one degree of freedom and maintain the position of the phenyl rings inside the binding pocket (Fig. 13). The docking, MD and MM-GBSA calculation were performed with similar modalities as in the previous series of compounds. Despite the good stability during the MD simulation and binding energy comparable to the diphenylether molecules (Tab. 2), the compounds 11, 12, and 13 exhibited IC_{50} values in the µM range, indicating the indole derivates have not affinity for the target. A possible explanation of the weak activity takes into account a high energy barrier that indole series face during the binding process. The rigidity together with the size of the molecules could hinder the entrance of the ligand inside the binding pocket. This explanation justifies the unsuccessful prediction, since MM-GBSA calculations just take into account the bound and the unbound complex.
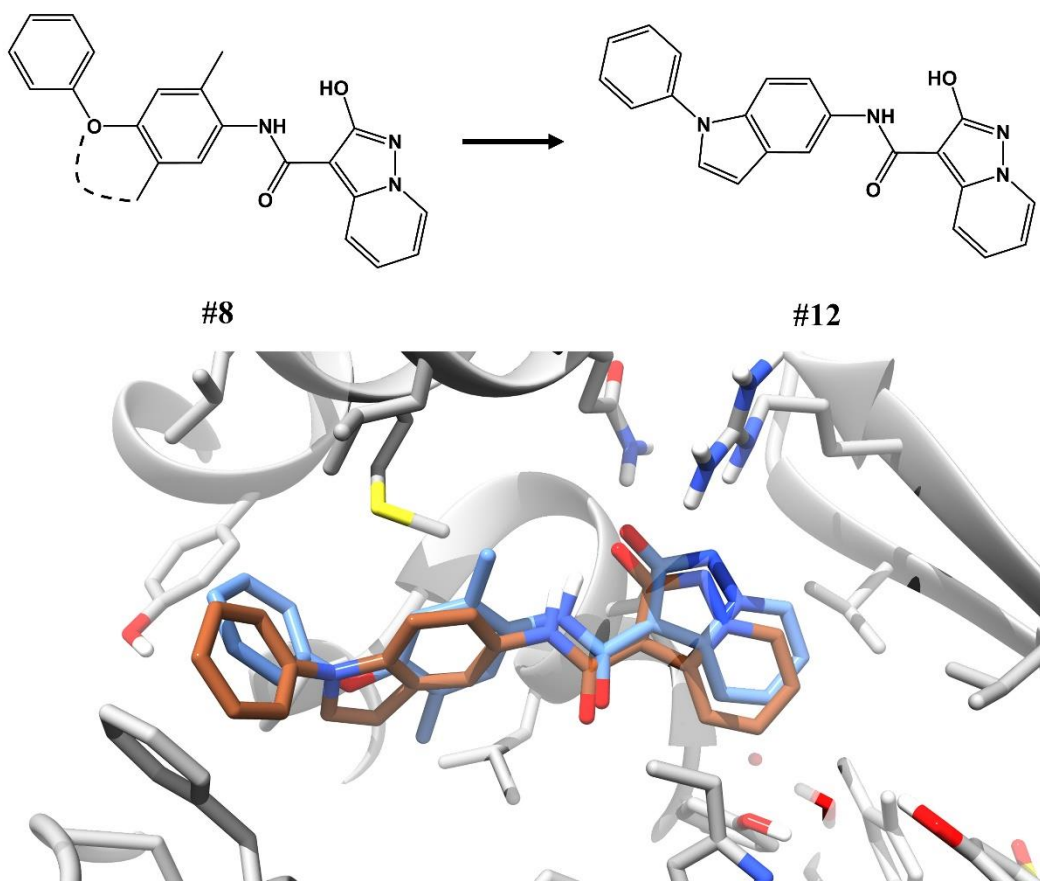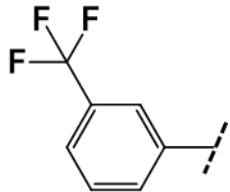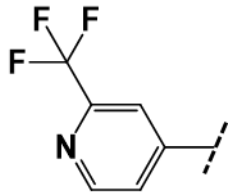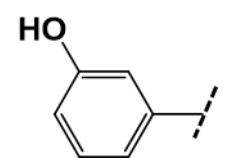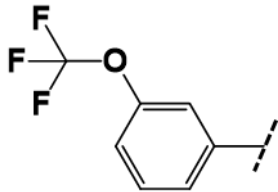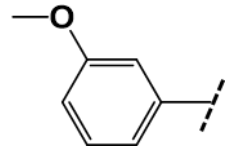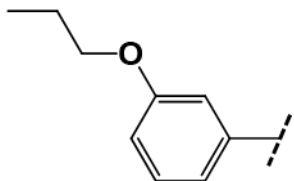
Fig 13. On the top the rational behind the design of compounds **11 – 13.** On the bottom the predicted binding mode of **8** and **12**.

1.3.3. Modulation of C Ring

The optimization of meds433, as well as every drug candidate, it is multivariate problem where physico-chemical and biological space are explored to achieve the optimal solution. This concept was detailed also by Christian *et al.* that underlined how in a series of BAY analogues, clogD lower than 2.5 leads to compounds with insufficient permeability.[23] Lipophilicity plays a central role in the transfer of enzymatic activity into substantial effect on cells, since the ligands must be able to reach the inner mitochondrial leaflet where the DHODH is located. Thus, compounds LogD seems to be well correlated to myeloid differentiation in AML cell lines.
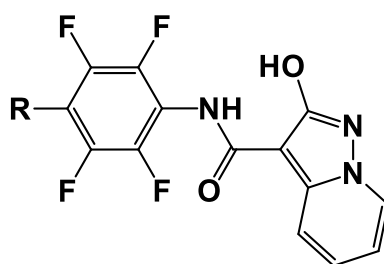
#	-R	$h\text{DHODH}^a$ $IC_{50} \pm SE$ (nM)	log D7.4 ± SDc	Solubility (µM) in PBS
Meds433		1.2 ±	2.35 ± 0.02	12
14		n.d.	0.66 ± 0.08	438
15		n.d.	n.d.	n.d.
16		n.d.	0.98 ± 0.03	47.3
17		n.d.	*Insoluble*	n.d.
18		2.0 ± 0.5	2.09 ± 0.04	n.d.
19		n.d.	*Insoluble*	*Insoluble*

20		2.3 ± 0.8	2.69 ± 0.03	5.6
21		n.d.	1.84 ± 0.06	20.2
22		n.d.	1.82 ± 0.09	55.3
23		n.d.	3.27 ± 0.19	8.1
24		n.d.	*Insoluble*	n.d.
25		n.d.	3.28 ± 0.12	12.9

Table 3. IC_{50}, LogD and solubility of compounds **14 – 25**.

For instance, considering two acid inhibitors with comparable IC_{50} such as meds433 and brequinar, 1.2 nM and 1.0 nM respectively, meds433 with a LogD of 2.35, induces myeloid differentiation at a concentration that was 1-log lower than brequinar with a LogD value of 1.83. Therefore, the further modulation of meds433 were aimed to optimize compound's lipophilicity while retaining high enzymatic inhibition activity. Considering the SAR of our inhibitor, we identified the positions on the C ring (Fig. 8) as suitable substitution points for the purpose. The binding mode of meds433 and derivates places the C ring next to the entrance of the ubuqinone binding pocket (Fig. 7), where the phenyl is mainly involved in hydrophobic interaction with Phe62, Pro364, Leu68 and Tyr38. Moreover, as shown in

Figure 14, substituent in *meta* position are located into an empty area of the binding site, on the border between the pocket and the vacuum. The solvent exposure of these modulations might lead to marginal effect on the inhibitor's activity. In addition, a proper substituent able to extend outside the protein can be exploited to attach ligands to a mitochondria-specific delivery vector. A series of substitutions on C ring were designed and applied to meds433, compounds **14** – **25** were then synthetized. Unfortunately, among the whole series only for compounds **18** and **20** the IC_{50} values were obtained so far.
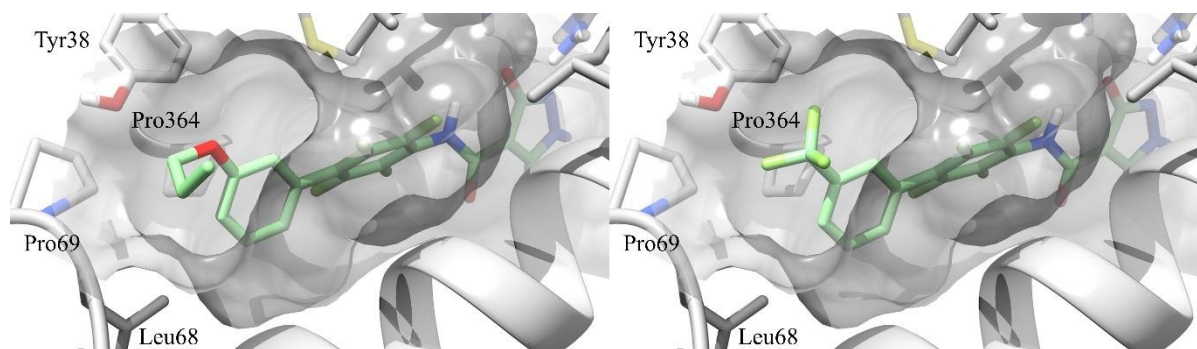

Figure 14. View of entrance of the ubiquinone binding site with predicted binding mode of compound 25, on the left, and compound 20 on the right.

Both compounds showed, as one might expect, a value of enzymatic inhibition comparable to lead compound meds433 (Tab. 3). The substitution of hydrogen atom with fluorine is a classical bioisoteric replacement, which in this case has a marginal effect on the potency of the molecule. Furthermore, also the introduction of a small lipophilic group such as CF_3 in the same position does not affect the activity. This, validate the predicted binding mode of **20,** where trifluoromethyl is placed in empty area of the receptor. Compounds solubility was evaluated at pH 7.4 in phosphate buffered saline (PBS), at 25 °C. The solubility of most compounds in table 3 show values that largely underline the reference limit of 20 μM, thus they are considered insoluble. As expected, in the target compounds **17 - 20**, the modulation of the C ring with -F atom or -CF3 group result in a further decrease of solubility. A different situation is observed for compound **14**, **16**, **21**, and **22**. In detail, compound **14** and **22**, which are characterize by an extra oxygen atom able to form hydrogen bond with water, showed a better solubility with respect to meds433. For compounds **16** and **21** the presence of a pyridine ring yield to a good solubility. Finally, compounds **23** and **25** showed a solubility comparable with the lead compound. Considering the Log D parameters, normally the optimal range, necessary for the absorption of the drug through phenomena of passive permeability or by diffusion, turns out to be value between 1 and 3.[41] For the all series Log D was measured using the shake flask method at pH 7.4 (See methods). Compounds **14** and **16**

fall out of the optimal range of log D due their high solubility, on the other hand compounds **17**, **19** and **24** are completely insoluble as a result of their increased lipophilicity. Compounds **20**, **23** and **25** present a higher logD compared to medss433, in particular, the introduction of propoxy group in **25** lead to an optimization of lipophilicity maintaining the same solubility. Noteworthy, **16** and **21** show a very low lipophilicity due to the presence of the pyridine in the molecular scaffold, and in vitro data has shown that this reflects in the inability of the molecules to pass the membranes.

1.3.4. Cell-Based Assay

The effect of *h*DHODH inhibitors discussed above were evaluated on AML cell line THP1. The differentiation process was tracked by analyzing CD14 expression, as this antigen is typically present in mature myeloid cells, whereas the apoptotic process is marked searching the positive response to the Annexin V^+ (See methods). In order to limit the number of experiments, the tests were conducted using a compound concentration of 1 µM, then the most promising molecules were selected to repeat the cell assay at the lower concentration of 0.1 µM.
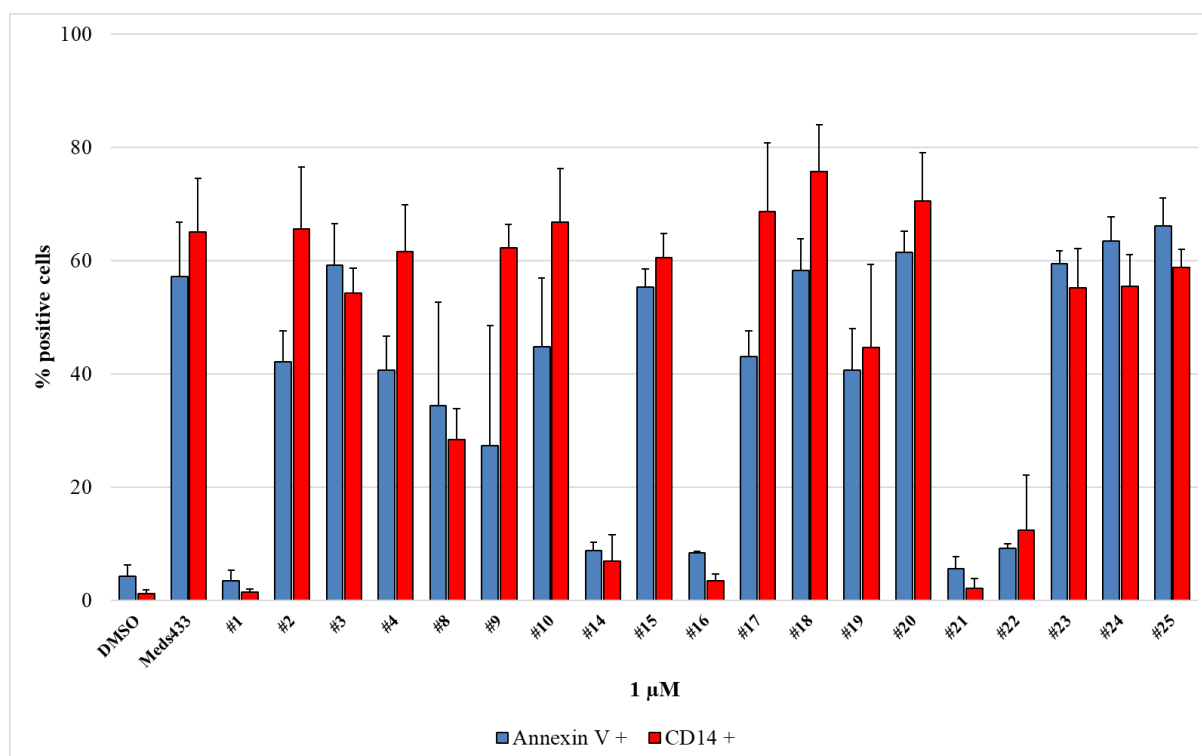
Figure 15. Differentiation and apoptosis induction with 1 µM of inhibitor expressed as the proportion of CD14 and Annexin V positive cells (THP-1). The compounds are reported as comparisons while DMSO indicates cells treated with dimethyl sulfoxide only.

As shown in the graphic in **Figure 14**, at 1 µM of concentration, most of the compounds induce the phenomena of differentiation of myeloid cells and the apoptotic, except for **1**, **14**, **16**, **21**, **22**. As discussed above, the ineffectiveness of these compounds is due to the weak activity or the low lipophilicity that hinders the passage through the cell membrane. Then, the experiment was repeated with only the most promising compounds at 0.1 µM concentration of inhibitors. The results shown in **Figure 15** indicate that **2**, **4**, **8**, **9**, **17**, **19** are not able to induce differentiation or apoptosis in AML cells at this concentration. Among the dyphenilether series, only compound **10** result effective at 0.1 µM, although it does not reach the lead compound meds433. The introduction of chlorine atom in compound **3** decrease the effect *in vitro*.
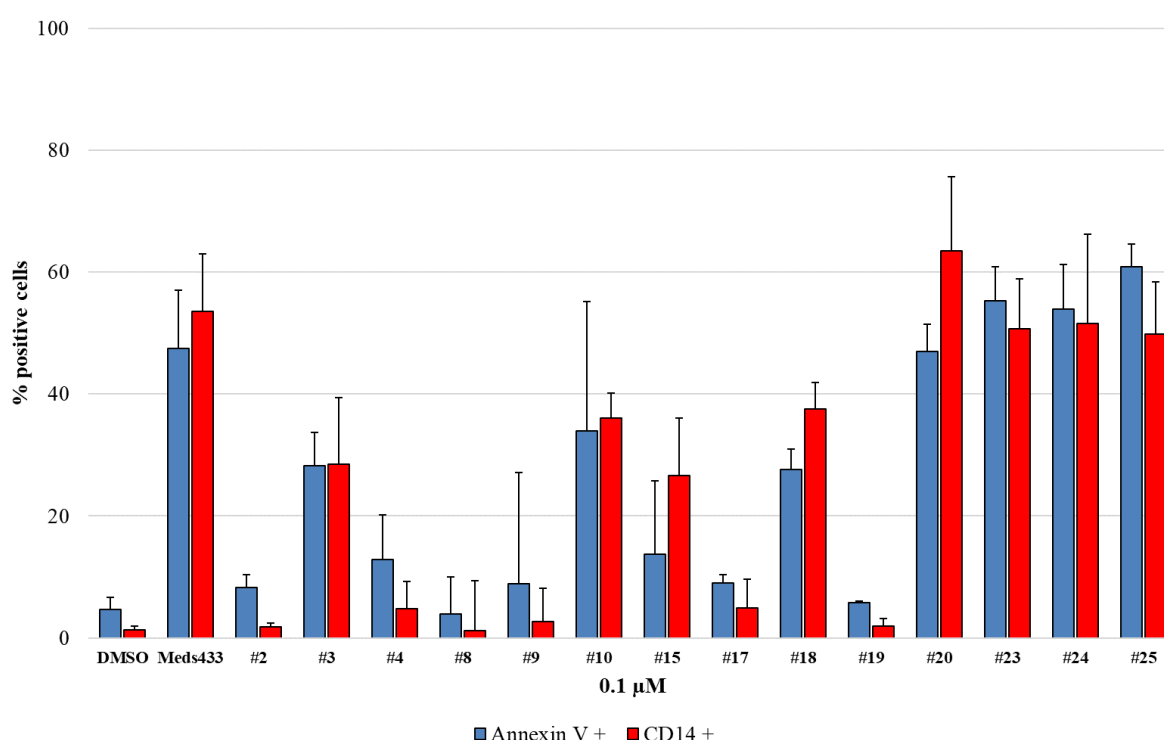
Figure 15. Differentiation and apoptosis induction with 0.1 µM of inhibitor expressed as the proportion of CD14 and Annexin V positive cells (THP-1). The compounds are reported as comparisons while DMSO indicates cells treated with dimethyl sulfoxide only.

The modulation applied on **15** and **18** lead to lower differentiation and apoptotic effect, on the other hand substitution on *meta* position of C ring yield comparable or better results with respect to meds433 for compounds **20**, **23**, **24** and **25**. This confirm substituents in *meta* position are favourite compared to the ones in *para*. To further demonstrate the connection between differentiation and *h*DHODH inhibition, the differentiation and apoptosis experiments of the best six compounds of the series (**10**, **18**, **20**, **23**, **24** and **25**) were repeated in the presence of uridine, and the complete rescue of the phenomenon was observed. Finally,

in **Table 4** reported the EC_{50} of the most promising compounds of the series compared with brequinar and BAY-2402234 are reported. We can note that, compounds **10** and **18** had a higher EC_{50} with respect to the lead compound. For **20, 23** and **24** we can observe instead a similar potency to meds433 in terms of induction of apoptotic process and differentiation. Noteworthy, compound **25** shows a better profile respect meds433, but around 3-fold less efficient compared to phase-1 *h*DHODH inhibitor BAY-2402234.

	EC_{50} (µM)	
	Apoptosis	**Differentiation**
Brequinar	missing	0.27
BAY-2402234	0.003334	0.002391
Meds433	missing	0.0297
#10	0.1094	0.06922
#18	0.2222	0.05887
#20	0.04289	0.03204
#23	0.03260	0.02795
#24	0.02961	0.0269
#25	0.018	0.01454

Table 4. Cellular EC50 (µM) for apoptosis and differentiation responsivity expressed as the proportion of Annexin V and CD14 positive cells (THP-1) respectively.

1.4. Conclusions

Meds433 is a recently identified *h*DHODH inhibitor based on hydroxypyrazolo[1,5-a]pyridine scaffold. Starting from compound meds433, we have explored a series of derivates using different computer aided techniques, with the aim of investigating and optimizing the pharmacokinetic and pharmacodynamic profile of our lead compound. The most suitable position for a beneficial replacement of the hydrogen on the pyridine motif of meds433 were explored with MD/FEP method, the best results were synthetized and tested. To increase compound solubility, two alternative scaffolds to tetraflurinatebyphenil moiety were studied through MD and MM-GBSA calculations. On diphenylether derivates the introduction of aliphatic groups resulted fundamental to stabilize the bioactive conformation and lead to obtain additional derivates in the low nM range. The second series based on indole scaffold resulted to be weakly active. Furthermore, distal phenyl ring (C ring) of meds433 was

subjected to several modulation, and the *meta* position was identified as able to place substituent next to the entrance of the binding pocket. This behaviour allowed to modulate pharmacokinetic properties while leaving the compound potency unchanged. Meds433 already proved its efficacy to induce myeloid differentiation in AML cell lines, leading to the massive death of leukemic cells. The derivates analyzed here were tested on AML cells THP1 and their differentiation and apoptotic profile were evaluated. Derivates substituted on *meta* position of C ring showed satisfactory values of EC_{50} in differentiation and induction of apoptotic process on cells. In particular, compound **25** was 2-fold more efficient in both assay in comparison to meds433. We can conclude that several hydroxypyrazolo[1,5-a]pyridine derivates demonstrated a potential effect on treatment of AML in vitro, making these molecules candidates for further in vivo studies in AML models.

1.5. Methods

1.5.1. *In Silico* Protein Preparation

All analyses were carried out on the *h*DHODH protein conformation that was extracted from the X-ray data using PDB ID: 6FMD. The missing protein loops were built and the crystal structure of the protein underwent an optimization process using the Protein Preparation Wizard tool, implemented in Maestro[TM] GUI.[42] Missing hydrogen atoms were added and bond orders were assigned. The prediction of protonation states for the protein was accomplished using PROPKA[TM], with the pH set at 7.4.

1.5.2. Docking

Docking studies on the compounds shown herein were performed using Glide/Inducet Fit Docking Protocol (IFD). [43] Docking was performed using the bound crystallographic ligand as the centroid of the box. The standard IFD protocol was used. Protein preparation constrained refinement and Glide XP redocking were set, while the other parameters were kept in their default states.

1.5.3. Molecular Dynamics

Molecular dynamic simulations were performed using GROMACS (version 5.0.5).[44] The parameter files for the complexes were prepared using the tLeap module of Ambertools and the amber ff14SB force field.[45, 46] Ligand and cofactor (Flavin mononucleotide and Orotate), parameters were obtained using the Antechamber module[47] and AM1-BCC charge methods.

[48] The starting structures were immersed in a pre-equilibrated cubic box of around 25000 TIP3P water molecules and chloride ions were added to maintain the electrical neutrality of the simulated systems. The systems were minimized over 6000 steps of the steepest descent algorithm before MD simulations were performed. The minimized structures were used as a starting point for the MD simulations. During the equilibration steps, the protein α carbons were kept fixed with a constraint of 1000 kcal/mol. In the first step, a constant volume simulation (NVT), was performed, during which the system was heated from 0 to 300 K over three 100-ps steps; 100, 200 and 300 K. The second, isothermal- and isobaric ensemble (NPT), was performed using the Parrinello-Rahman algorithm for 1 ns of dynamic simulation. Finally, 50 ns MD production trajectories were run without restraint, collecting frames at 100 ps intervals and using a 2 fs time step. Particle mesh Ewald (PME),[49] was used to treat the long-range electrostatic interactions in MM minimization and MD simulations. All bonds were constrained using the LINCS algorithm.[50]

1.5.4. Free Binding Energy Calculations

Molecular mechanics/generalized born surface area (MM/GBSA) is a widely used technique to calculate the binding free energy between receptor and ligand.[51-53] Here, the MM/GBSA method was employed to compute the binding free energy of DHODH in complex with compounds **4**, **8**, **9**, **10**. The free energies were calculated on basis of the last 40 ns of MD trajectories. The ΔG_{bind} of protein-ligand complexes were computed with the following equation:

$$\Delta G_{bind} = \Delta H - T \, \Delta S \approx \Delta E_{MM} + \Delta G_{sol} - T \, \Delta S$$

$$\Delta E_{MM} = \Delta E_{internal} + \Delta E_{ele} + \Delta E_{vdw}$$

$$\Delta G_{so} = \Delta G_{GB} + \Delta G_{SA}$$

where ΔE_{MM}, $T\Delta S$ and ΔG_{sol} represent the gas phase MM energy, conformational entropy, and solvation free energy, respectively. ΔE_{MM} consists of van der Waals energy ΔE_{vdw}, electrostatic ΔE_{ele} and $\Delta E_{internal}$ of the bond, angle, and dihedral energies. The Generalized Born (GB) model was used to compute the polar solvation free energies ΔG_{GB}.[54] And, the LCPO method wa employed to calculate the nonpolar solvation contribution energy ΔG_{SA}.[55] To compute the entropy calculation, 400 snapshots were extracted from the simulated trajectories every 100 ps. All binding free energy calculations were carried out by

AmberTools14 and AMBER14.[46] MMPBSA.py program were used to decompose the contribution energies of individual residues.[56]

1.5.5. FEP Analysis

All simulations were performed in GROMACS (version 5.0.5), more details are provided in the supplementary material section. Two different sets of calculations were performed for each alchemical transformation; one on the ligand-protein solvated complex and the other on the ligand into the solvent. FESetup[57] was used to prepare the input. Free energies were obtained via the implementation of multiple Bennet acceptance ratios (MBAR), which were provided by python package pymbar (https://github.com/choderalab/pymbar),[58] using the Alchemical analysis tool (https://github.com/MobleyLab/alchemical-analysis).[59]

1.5.6. Solubility Assay at pH 7.4

Solubility was assayed in phosphate buffered saline (PBS: 12 mM with NaCl 137 mM and KCl 2.7 mM, pH 7.4) Each solid compound (1 mg) was added to 1 mL of PBS. The samples were shaken in an orbital shaker at 25 °C for 24 h. These suspensions were filtered through a PTFE 0.45 μm filter (VWR), and the solutions were chromatographically analyzed. Quantitative analysis was performed on a HPLC-UV system (MERK -HITACHI), equipped with an auto sampler of 60 μL injection volume (MERK- HITACHI AS-2000A), a binary HPLC pump (MERK-HITACHI L- 6200 IP), and a diode array detector (MERK-HITACHI L-4250). LC analysis was performed using an Agilent Zorbax SB-Phenyl column 4.6 × 250, 5 μm). Analyses were carried out at a flow rate of 1 mL/ min using gradient elution with eluent A being trifluoroacetic acid (TFA), 0.1% in water, and B TFA 0.1% in MeOH for brequinar and compounds 4−10. The analyses started with 40% of eluent B, and the following gradient profile was used: (time min, % B) 18.0, 100%; 26.0, 100%; 28.0, 40%. For compound 5, eluent A was TFA 0.1% in water and eluent B acetonitrile. The gradient profile was as follows: (time, % B) 0, 50%; 7.5, 50%; 22.4, 100%; 32.4, 100%. Single compound quantification was made using the relative calibration curve, which was obtained by analyzing standard solutions in MeOH. Solubility is expressed as micromolar concentration of the saturated solution.

1.5.7. Log D (pH 7.4)

The partition coefficients between n-octanol and PBS at pH 7.4 ($\log D^{7.4}$) were obtained using the shake-flask technique at room temperature. In the shake-flask experiments, 50 mM of phosphate buffered saline pH 7.4 was used as the aqueous phase. The organic (n- octanol) and aqueous phases were mutually saturated by shaking for 4 h. The compounds were solubilized in the buffered aqueous phase at the highest concentration compatible with solubility, and appropriate amounts of n-octanol were added. The two phases were shaken for about 20 min, by which time the partitioning equilibrium of solutes had been reached, and then centrifuged (10 000 rpm, 10 min). The concentration of the solutes was measured in the aqueous phase by UV spectrophotometer (Varian Cary 50BIO); absorbance values (recorded for each compound at the wavelength of maximum absorption) were interpolated in calibration curves obtained using standard solutions of the compounds ($r^2 > 0.99$). Each log D value is an average of at least six measurements.

1.5.8. *h*DHODH Inhibition Assay

Inhibitory activity was assessed by monitoring the reduction of 2,6-dichloroindophenol (DCIP), which is associated with the oxidation of dihydroorotate as catalyzed by the DHODH enzyme. The enzyme was preincubated for 5 min at 37 °C in Tris-buffer solution (pH 8.0), with coenzyme Q10 (100 μM), with the compounds to be tested used at different concentrations (final DMSO concentration 0.1% v/v), with DCIP (50 μM). The reaction was initiated by the addition of DHO (500 μM), and the reduction was monitored at $\lambda = 650$ nm. The initial rate was measured in the first 5 min ($\varepsilon = 10\ 400\ \mathrm{M^{-1}\ cm^{-1}}$), and an IC_{50} value was calculated, using GraphPad Prism 7 software. Values are means ± SE of three independent experiments.

1.5.9. Proliferation Assay

The proliferation of AML cell lines THP1 was evaluated using a flow cytometer. Cell lines were labeled with CFSE dye according to the protocol described above. After labeling, cell lines were plated (1 × 104) and cultured with hDHODH inhibitor molecules (0.01−10 μM) for 3 days. At the end of the cultures, cells were harvested, and 1 μg/mL of propidium iodide was added to exclude dead cells before acquisition. The proliferation of cell lines was quantified on viable cells as % of PICSFE cells.

1.5.10. Differentiation Assay

The 1×104 cells (THP1) were plated in 96-well round-bottom plates, and hDHODH inhibitors were added, from 0.1 to 10 μM, to a volume of 200 μL of medium. The differentiation kinetics was monitored from day 1 to day 4 for U937, and to day 5 for THP1. Cells were washed and stained with CD11b and CD14. The differentiation assay was also performed in the presence of uridine 100 μM and analyzed on day 3.

2. Computational Method for Structure-Based Analysis and Application of SAR Transfer

2.1. Introduction

Hit-to-lead and lead optimization are the basic steps in drug discovery. They are based on the generation of analogues to establish dose-response behavior, collect preliminary structure-activity relationship (SAR) information, improve compound potency, further evolve SARs, and ultimately balance pharmacokinetic and pharmacodynamics molecular properties.[60] Compound optimization efforts are largely guided by the expertise of medicinal chemists, which are supported by the state-of-art of computational methodologies. The process is far from being routine and roadblocks are frequently encountered. During lead optimization, compounds with promising SARs might fail due undesirable characteristics such as low bioavailability, metabolic liability or unwanted side effects, which then prevent their further development. In these cases, experts seek to replace such compounds with analogue compounds able to present a better profile, expecting them to also maintain activity and show similar SAR characteristics. This concept is also referred to as SAR transfer from one compound class or series to another.[60, 61] SAR transfer notion is a common task in practical medicinal chemistry taking part in the lead optimization process. Empirical SAR transfer attempts can be supported by computational search calculations for compound series in which pairs of analogues with corresponding substitutions display comparable SAR progression for the same target.[61, 62] In addition to circumventing roadblocks in multi-property optimization, SAR transfer might also be considered when investigating parallel series for their development potential. Nowadays, the constantly increasing volume of crystallographic data becoming available allows, in most of the cases, SARs investigations with the aid of target structures. This is typically done by iteratively solving X-ray structures of complexes between a target protein and analogues. Then, SAR progression is monitored on the basis of crystallographic ligand-target interactions.[63] Alternatively, series of analogues can also be docked or superimposed onto complex structures with representative compounds.[64, 65] The different interactions can be studied with technique such as molecular dynamics and free energy calculations to relate the activity to chemical modifications.[66-68] Furthermore, fragment-based design has emerged as an alternative approach in drug discovery, complex structures with weakly active fragments are determined as seeds for incremental or combinatorial generation of increasingly drug-like and potent compounds.[69] The latter

methodology has emerged as an alternative approach in drug discovery, growing considerably in the past years, being used routinely in both industry and academia. On the basis of these approaches, SARs of evolving series can be investigated in a structure-based manner. For a systematic exploration and exploitation of SAR transfer events, going beyond case-by-case analysis, the application of computational methods is essential. So far, SAR transfer has only been investigated computationally on the basis of compounds and analogue series. By contrast, there currently is no methodology available for structure-based assessment of SAR transfer, although insights into ligand-target interactions obtained from structures of complexes add another dimension to SAR analysis.

2.2. Aim of the Project

SAR transfer is a well-known technique used in medicinal chemistry. Although few computational methodologies are available to analyze SAR transfer, none of them relies on the three-dimensional structure of ligand-target complexes. Moreover, literature lacks studies on the statistical assessment of SAR transfer events, which become relevant for practical application during lead optimization. The objective of this project is the development of a generally applicable computational methodology for the identification of SAR transfer and other SAR events that combines the comparison of three-dimensional ligand binding modes with the search for other active compounds having varying core structures and substituents. Then, proof of concept was established using the new approach to identify SAR transfer events for selected targets, providing opportunities for follow-up investigations of SAR transfer and compound optimization. Once a SAR transfer event has been detected for an X-ray ligand or compound series of interest, newly derived core structures from analogue ensembles can be used as starting points for the generation of transfer series. Multiple analogue series linked to each other by the principles of SAR transfer define a SAR table, which extend the view form single SAR events to an expanded scenario that includes related SAR transfer data from various X-ray structures. Moreover, SAR tables offer the opportunity to evaluate the contribution of multiple substructures (fragments) obtaining relevant and interpretable SARs to guide practical optimization efforts. In the following chapters, the approach for structure-based SAR transfer analysis is introduced and exemplary applications for different target proteins and crystallographic ligands are presented.

2.3. Result and Discussion

2.3.1. Structure-based SAR Transfer Concept

Principles of the methodology introduced herein are illustrated in **Figure 1**. Comparisons of X-ray structures of protein-ligand complexes are carried out to identify pairs of chemically distinct ligands with similar binding modes for a given target. For ligands of a pair, core structures and shared substituents are identified on the basis of molecular fragmentation and three-dimensional fragment comparisons and a search is carried out for structural analogues of each ligand. Obtaining two pairs of analogues with distinct core structures and the same substitution is the desired outcome, as illustrated in **Figure 1A**. For each pair of analogues, the potency difference (ΔP) accompanying the corresponding chemical modification is determined, which provides the basis for calculating the $\Delta\Delta P$ value as a measure of SAR transfer. If $\Delta\Delta P$ is small, an SAR transfer event is identified. Hence, interpretation of the SAR for the final analogue example is straightforward from a medicinal chemistry perspective and the results are easy to rationalize. However, a key aspect of the methodology is that:

(i) Core structures and substituents of ligands are not predefined but determined on the basis of systematic molecular fragmentation, comparison of compound fragments in three dimensions, and recombination of fragments into final cores and substituents.

This methodological aspect is "hidden" in the example shown in **Figure 1A** and its notion is not essential for interpreting the results. However, superposition-based molecular organization requires the development of a systematic fragmentation scheme to identify shared fragments in compared ligands and derive molecular cores, which provide the basis for the identification of analogues, as schematically illustrated in **Figure 1B**. The fragmentation and analogue search method is detailed in the Experimental Section (see "Systematic Compound Comparison"). In addition, there are two other characteristic features of the approach that enable three-dimensional SAR transfer assessment including:

(ii) Identification of shared fragments representing substituents and core structures is strictly based on three-dimensional similarity.

(iii) For a pair of X-ray ligands, multiple cores and substituent fragments can be obtained.

These characteristics set the fragmentation scheme introduced herein apart from others used underlying, for example, the generation of scaffold trees,[70] scaffold keys,[71] matched molecular pairs,[72, 73] or SAR matrices.[74] The newly introduced three-dimensional fragmentation approach was specifically designed for structured-based SAR transfer assessment, the aim of our approach.
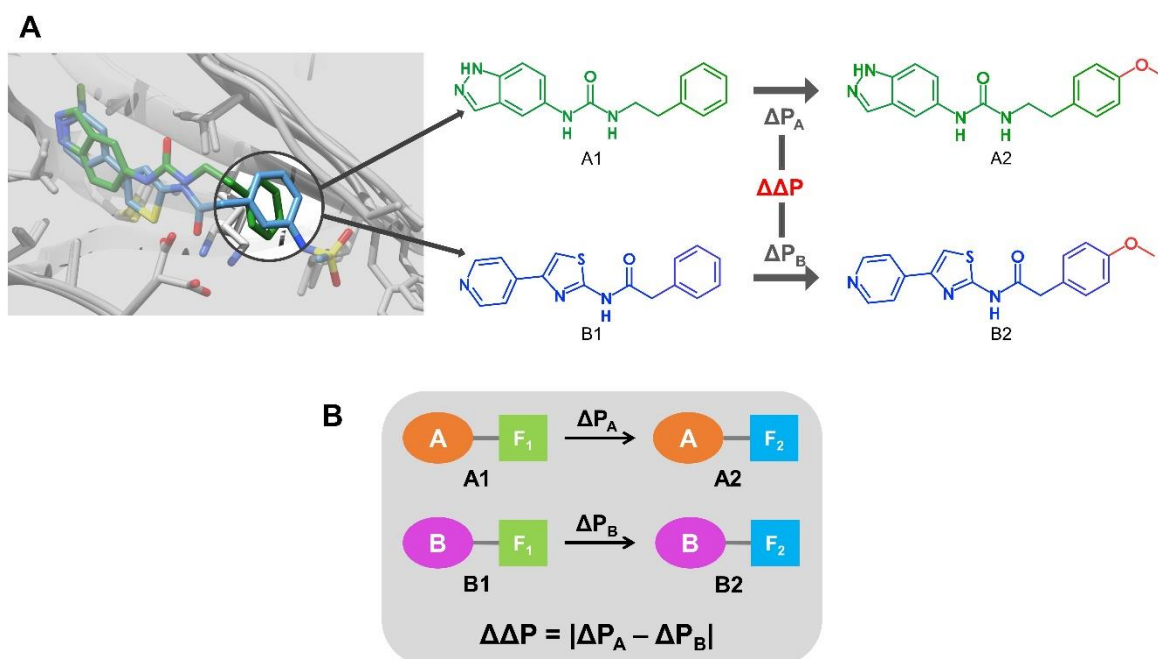
Figure 1. Methodological concept underlying structure-based SAR transfer analysis. (A) The representation illustrates basic principles of computational structure-based SAR transfer analysis. On the left, complex X-ray structures of Rho kinase with two inhibitors are superimposed (PDB ID: 3V8S and 5BML). The black circle highlights a shared phenyl ring. On the right, the two inhibitors are displayed together with corresponding active analogues that are distinguished by a methoxy substitution (red) at the shared phenyl substructure (fragment). From the ΔP (potency difference) value of each X-ray inhibitor – analogue pair, $\Delta\Delta P$ is determined. (B) The analysis scheme is generalized. A1 and B1 are different crystallographic ligands with corresponding binding modes and a shared fragment F1 (green) but distinct core structures (orange and magenta, respectively). A2 and B2 represent analogues of A1 and B2, respectively, which have corresponding cores and share a fragment F2 that is distinct from F1. Thus, each analogue pair (A1, A2) and (B1, B2) is distinguished by the same substitution. For the two potency changes ΔPA and ΔPB accompanying the substitution in different analogue pairs, hence providing an alternative structural context, the $\Delta\Delta P$ value is calculated as an indicator of potential SAR transfer events.

2.3.2. Different Types of SAR Events

On the basis of $\Delta\Delta P$ values according to **Figure 1**, different categories of SAR events can be defined. For the purpose of our analysis, three different types of events were distinguished. First and foremost, a small $\Delta\Delta P$ value, ideally going against zero, identifies an SAR transfer event. In addition, if potency changes associated with a given substitution in corresponding analogue pairs are of moderately varying magnitude but have the same direction, i.e., an

increase (+) or decrease (-), we regard the event as a (consistent) SAR trend. By contrast, if potency changes are opposing (+/-) the event is categorized as an SAR inversion. These SAR events are further differentiated using different $\Delta\Delta P$ threshold values (see methods).

2.3.3. Systematic Analysis

For the structure-based investigation of SAR transfer, structures of different targets and their ligands were investigated. The general workflow is outlined in **Figure 2**. On the basis of multiple structural alignments of complexes with distinct ligands, spatially aligned pairs of ligands were obtained for each target and subjected to three-dimensional fragmentation in order to identify shared fragments and core structures. Considering substituent fragments and cores, a search for corresponding analogues was carried out to obtain analogue ensembles for $\Delta\Delta P$ calculations.
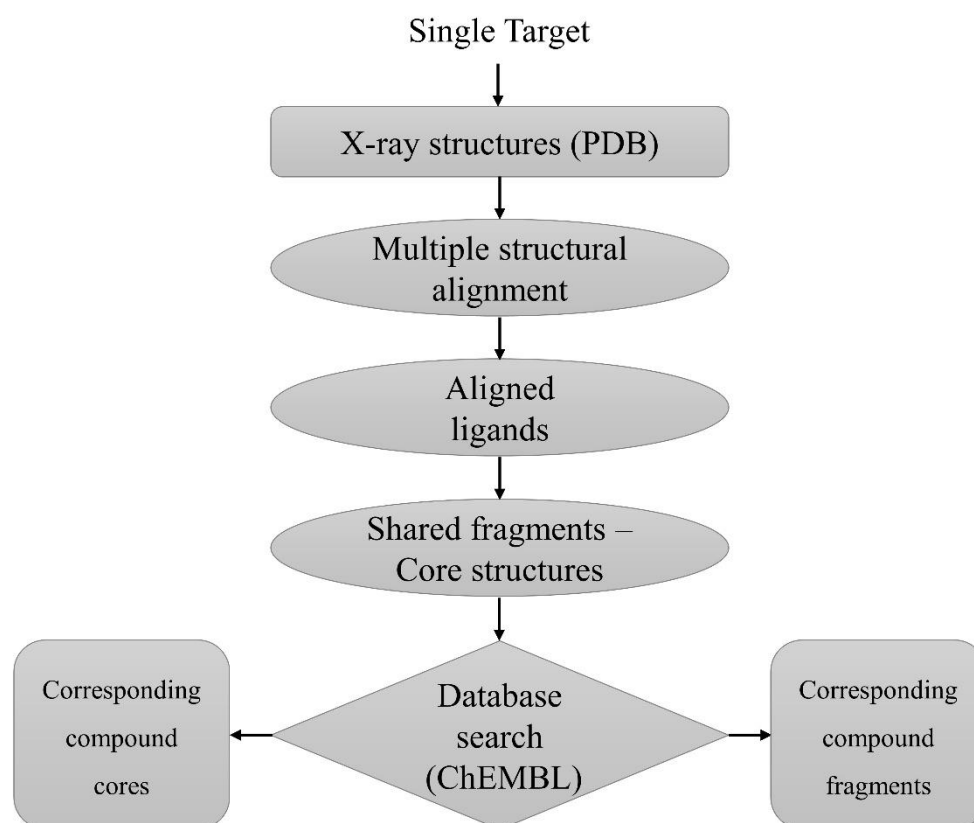
Figure 2. Analysis workflow. The flowchart outlines the protocol for the generation of analogue ensembles and $\Delta\Delta P$ analysis. For a given target, available complex X-ray structures are retrieved and superposed. Spatially aligned ligands with similar binding modes are selected and examined for shared fragments and cores. Then, a database search is performed for active analogues with

corresponding cores and fragments, respectively, to identify analogue ensembles for ΔΔP calculations according to Figure 1.

Table 1 reports the targets and number of complex X-ray structures used in our analysis. In addition, compound and ΔΔP statistics are provided. In each case, 20 different analogue ensembles were generated for ΔΔP calculations. The number of selected X-ray structures was typically small. For eleven targets, only two to three complex structures were taken for which well-defined ligand activity data were available. Nonetheless, in each case, 20 different analogue ensembles were obtained. This was a consequence of generating multiple shared fragments and cores for a pair of superimposed ligands, representing a specific feature of our methodology.

ChEMBL TID	Target protein	# X-ray	# ΔΔP	Mean ΔΔP	ΔΔP σ	# CPD	Pot. range
2360	Carbonic anhydrase I	2	20	0.71	0.60	43	3.89
4618	Dihydrofolate reductase	4	20	0.54	0.37	38	3.48
202	Adenosine kinase	2	20	0.52	0.43	26	1.58
261	Thyroid hormone receptor alpha	3	20	1.45	1.00	46	4.38
4681	Methionine aminopeptidase 2	4	20	0.49	0.53	28	3.50
4282	Beta-secretase 1	5	20	0.47	0.43	15	2.10
3587	Purine nucleoside phosphorylase	11	20	0.38	0.30	33	2.20
3305	Hypoxanthine-guanine phosphoribosyltransferase	6	20	1.07	0.92	28	3.44
1908389	Dual specificity mitogen-activated protein kinase kinase 1	2	20	0.50	0.40	42	3.77
1926	Serine/threonine-protein kinase AKT	3	20	0.80	0.59	65	4.67
3589	Aldo-keto-reductase family 1 member C3	5	20	1.19	1.22	25	3.79
205	Leukotriene A4 hydrolase	8	20	0.91	0.80	62	3.67
1937	Rho-associated protein kinase 1	3	20	0.64	0.63	47	4.86
4338	Testis-specific androgen-binding protein	5	20	0.88	0.78	27	7.65
1860	Histone deacetylase 2	4	20	0.76	0.54	30	4.92
4822	Egl nine homolog 1	2	20	0.24	0.20	45	4.05

3922	Carbonic anhydrase II	3	20	0.84	0.51	18	5.44
3231	Dihydrofolate reductase	3	20	0.85	0.55	25	2.22
5697	Thyroid hormone receptor beta-1	3	20	0.65	0.50	25	3.18
1947	Mitogen-activated protein kinase kinase kinase 12	2	20	0.63	0.54	32	4.93
Total		**80**	**400**	**0.73**	**0.68**	**680**	

Table 1. Targets, compounds, and activity data. The table lists crystallographic target proteins with their ChEMBL target ID (TID) used for generating the $\Delta\Delta P$ data set. For each target, the number of complex X-ray structures that were used for the analysis (# X-ray), mean of 20 $|\Delta\Delta P|$ calculations, standard deviation (σ), total number of compounds (# CPD) including X-ray ligands and database analogues, and their logarithmic potency (Pot.) range are reported.

2.3.4. SAR Transfer: Proof of Concept

The results in **Table 1** reveal that mean $\Delta\Delta P$ values were generally small although compound potency ranges mostly covered several orders of magnitude. With only three exceptions, target-based mean $\Delta\Delta P$ values fell well within one order of magnitude. **Figure 3** shows the distribution $|\Delta\Delta P|$ values in a boxplot format, displaying a narrow interquartile range and a number of statistical outliers, yielding a low $|\Delta\Delta P|$ median value of 0.53.
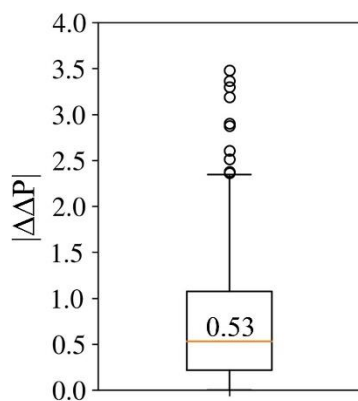
Figure 3. Distribution of $|\Delta\Delta P|$ values. The boxplot reports the global distribution of $|\Delta\Delta P|$ values across all targets. Boxplots report the smallest value (bottom line), first quartile (lower boundary of the box), median value (orange line), third quartile (upper boundary of the box), largest value (top line), and statistical outliers (points below the smallest or above the largest value). The median of the $|\Delta\Delta P|$ distribution (0.53) is reported.

Taken together, these findings already indicated that SAR transfer events were frequently detected, despite the presence of in part very large compound potency variations. In addition, high $\Delta\Delta P$ values were also observed. This is illustrated in **Figure 4**, which shows examples of SAR events with increasing $\Delta\Delta P$ values.
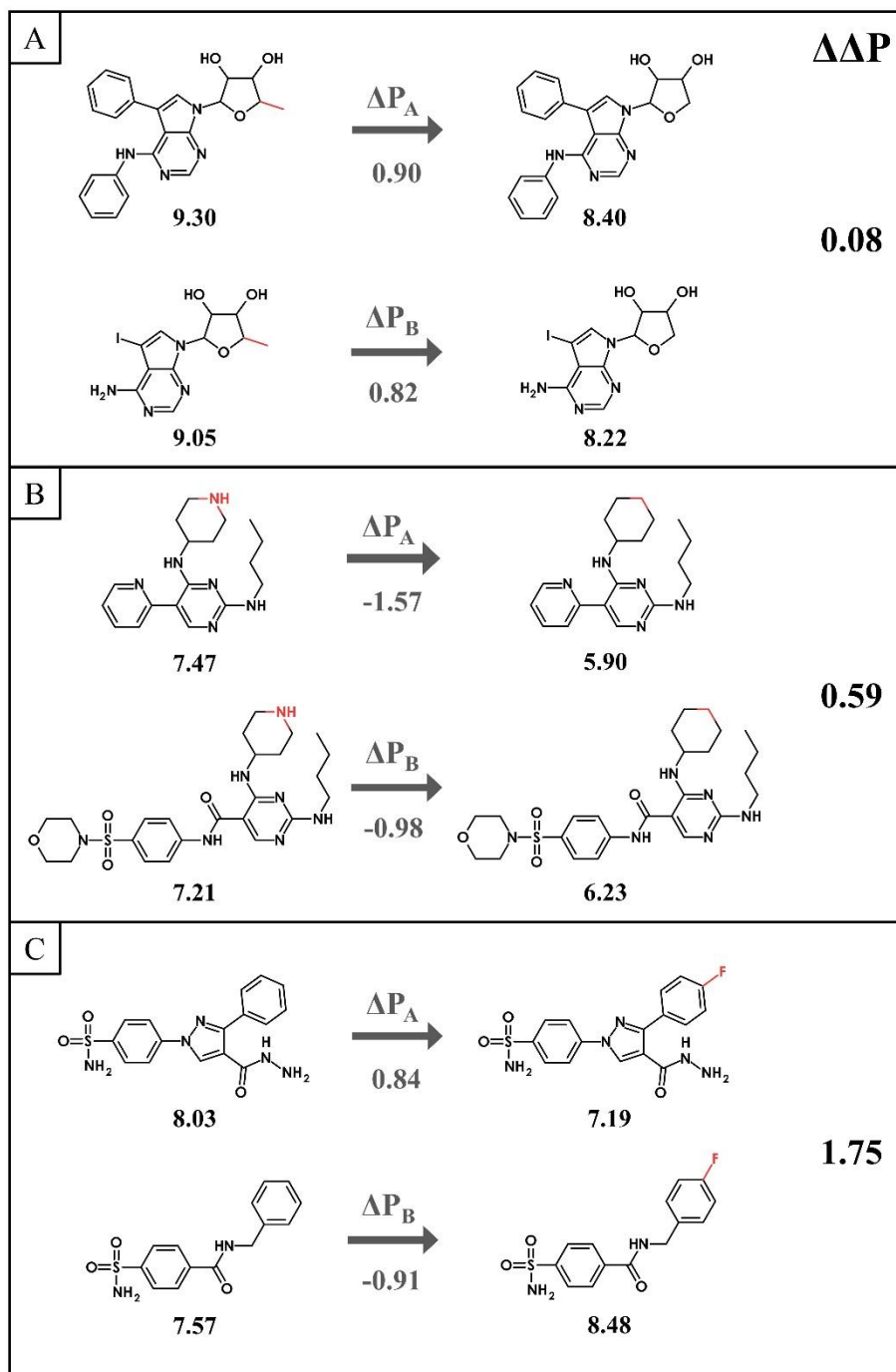
Figure 4. Exemplary analog ensembles. Shown are analogue ensembles with increasing $\Delta\Delta P$ values including inhibitors of (A) adenosine kinase, (B) proto-oncogene tyrosine-protein kinase MER, and (C) carbonic anhydrase II. Structural modifications that distinguish different fragments are colored red.

The analogue ensemble in **Figure 4A** represents a typical SAR transfer event. The examples shown in **Figure 4** illustrate the ability of the approach to identify analogue pairs with corresponding substitutions (shared fragments) having rather different core structures, which supports the relevance of the systematic three-dimensional compound fragmentation and recombination approach. **Figure 5** reports the results of our systematic analysis of all detected SAR events for different |ΔΔP| threshold values. SAR events belonging to all three categories were consistently detected. However, SAR transfer instances dominated the distributions for all three thresholds. At the lowest |ΔΔP| threshold, corresponding to an only 3-fold difference in relative potency between analogue pairs, 53% of all events corresponded to SAR transfer. At the intermediate (5-fold relative difference) and highest (10-fold) threshold, 65% and 74% of the detected events corresponded to SAR transfer, respectively. Compared to SAR transfer events, SAR trends accounting for larger differences having the same directionality were underrepresented (9% - 18% for decreasing thresholds). On the other hand, SAR inversions were more frequently observed than SAR trends, with 16% - 29% for decreasing thresholds. Taken together, the findings revealed that SAR transfer was detected with high frequency using the structure-based computational approach. **Figure 6** shows examples of different SAR events.
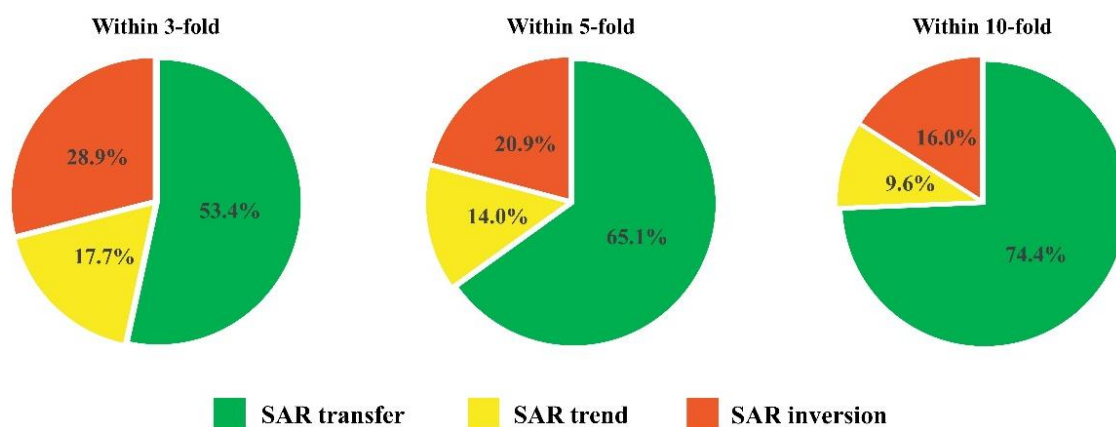
Figure 5. Distribution of different SAR events. For |ΔΔP| thresholds corresponding to at most 3-, 5-, or 10-fold differences, pie charts report the proportion of SAR transfer events, SAR trends, and SAR inversions across all targets derived from differences between compound potency values in analogue ensembles.

In **Figure 6A**, a representative SAR transfer event involving MAP3K12 kinase inhibitors is shown. Analogue pairs are distinguished by an iodine to isopropyl substitution of related yet distinct core structures leading to an equivalent potency increase producing nanomolar inhibitors. **Figure 6B** shows an ensemble of Ser/Thr-protein kinase AKT inhibitors with an

ortho,meta-dichloro to para-chloro modification of closely related small scaffolds, yielding a decrease in potency in the same direction.
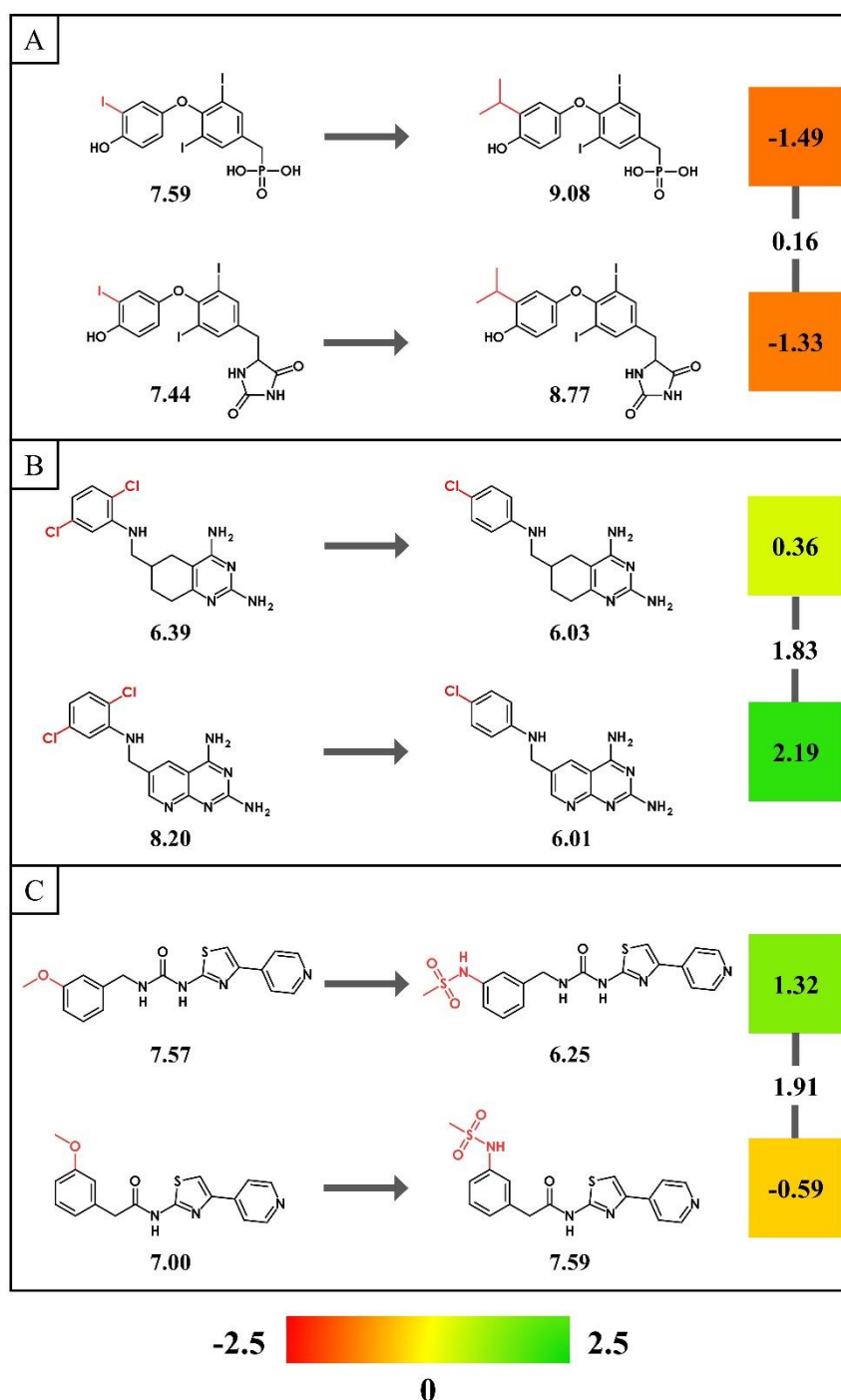
Figure 6. SAR categories. Shown are examples of different SAR categories introduced herein including (A) SAR transfer, (B) consistent SAR trend, and (C) SAR inversion. The representation is according to Figure 4. In addition, ΔP values are given in boxes color-coded using a continuous spectrum as indicated at the bottom. Accordingly, red/green ΔP combinations indicate SAR inversion events. ΔΔP values are reported between colored boxes.

However, the decrease for the analogue pair at the bottom exceeds two orders of magnitude whereas the potency decrease for the analogue pair at the top is well within an order of

magnitude. These potency effects would be regarded as a weak SAR trend. In **Figure 6C**, an SAR inversion is depicted for pairs of dihydrofolate reductase inhibitors related by a methoxy to sulfonamide substitution at different positions in a terminal phenyl ring causing a decrease (analogue pair at the top) and increase in potency (bottom), respectively. Taken together, these examples also illustrate the variety of structural substituent-core relationships detected through spatial fragmentation, recombination, and analogue searching, which yield interpretable SARs.

2.3.5. From SAR Transfer Events to SAR Tables

Transfer of evolving SARs from one compound class to another is a topical issue in medicinal chemistry. For example, applying the SAR transfer concept provides possible alternatives when encountering roadblocks during lead optimization. Accordingly, one would use the SAR gain from SAR transfer events for practical optimization of known or new compounds. Although the analysis of SAR transfer described so far provides valuable information, the very large number of generated data points makes it difficult to rationalize them for practical studies. Specifically, every SAR event contains four different compounds formed by one pair of core structures and one pair of fragments, providing multiple readouts and varying SAR information. Since the methodology relies on combination of molecules, the growth of results is exponential, and thousands of SAR events can be detected for a single target. In absence of an exact query, retrieving SARs from single events can be time-consuming even for an expert. For this reason, it is necessary to assemble and simplify SAR transfer data into an easy-to-read format. By definition, a SAR event refers only to a single chemical modification on two alternative core structures, which represents a small part of the complete SAR picture where hundreds of molecules belonging to different series are usually analyzed. Once an SAR transfer event has been detected for an X-ray ligand or compound series of interest, newly derived core structures from analogue ensembles can be used as starting points for the generation of transfer series. Thus, transfer series can be organized into so-called SAR tables. The analogy between SAR transfer scheme and the SAR table format is illustrated in **Figure 7**. Tables are based on the same rules as SAR event but extend the three-dimensional fragment comparisons (and analogous search) to complete ligand structures. The outcome is a table of compounds organized according transfer series and shared substituents (fragments). Cores and fragments can be easily inspected in SAR tables. Molecules in rows share the same core structure and differ in their fragments. Thus, each row contains a series

of analogues (transfer series). On the other hand, columns contain compounds with different cores and a common substituent. Therefore, SAR tables reveal structure-activity relationships from two points of view: the first one is along the horizontal axis, where fragments can be inspected evaluating their contribution. The second is along the vertical axis to investigate varying cores associated with the same fragments.
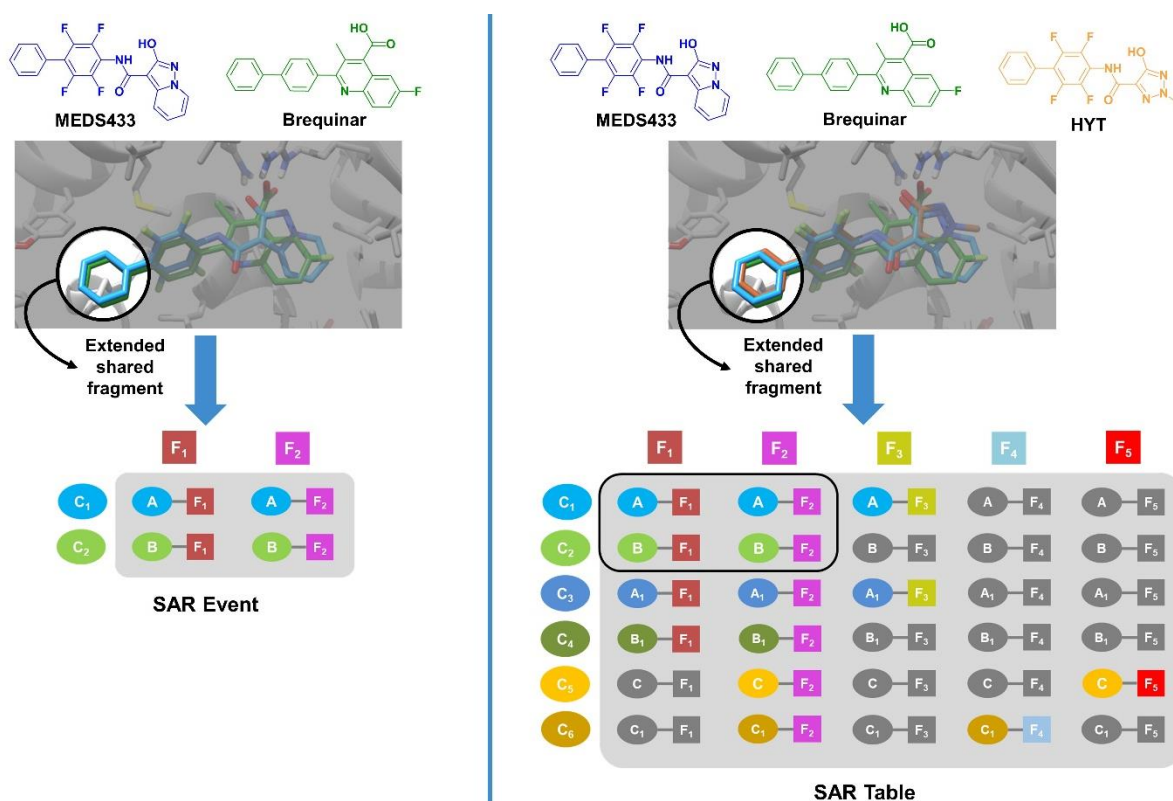
Figure 7. Comparison between SAR transfer event and SAR Table. On the left, methodological concept accounting for an SAR transfer event is schematically summarized. Complex X-ray structures of *h*DHODH with two inhibitors are superimposed (MEDS433, PDB ID: 6FMD; brequinar, PDB ID: 4D3G). On the right, the organization and composition of the SAR Table are illustrated. In this example, the complex of hDHODH - HYT (PDB ID: 5MUT) is superimposed on corresponding complexes with MEDS433 and brequinar. The black square in the table highlights the SAR transfer event. Combinations of core and fragment colored in grey represent virtual analogues (compounds lacking activity data).

Moreover, the SAR Tables might contain virtual molecules, compounds derived by the combination of cores and fragments missing activity data. Taking into account that fragments are the result of a specific fragmentation scheme for each compound, the combination of fragment and core structure originating from different series might generate virtual compounds that are synthetically inaccessible or chemically meaningless. Furthermore, a quartet of four different compounds formed by the combination of two alternative core structures and two fragments yield an SAR transfer event inside the SAR Table (Fig. 1). Focusing the study on accessible ligands, the contribution of single fragments in transfer
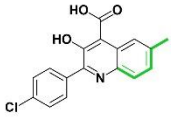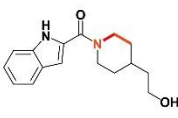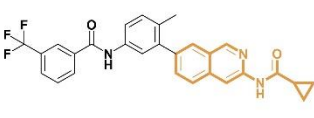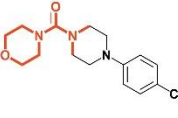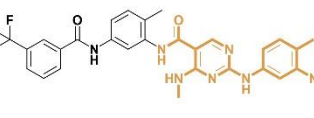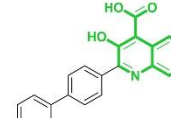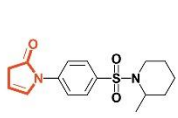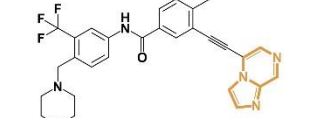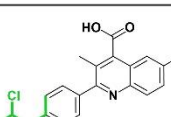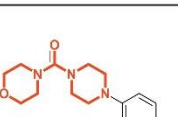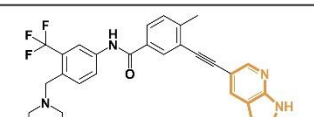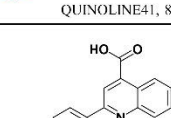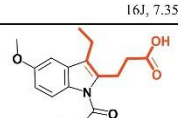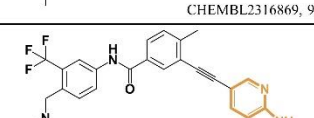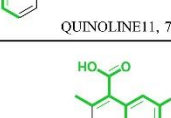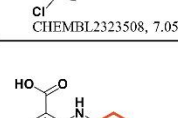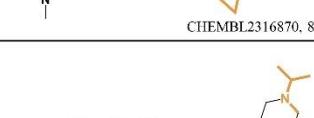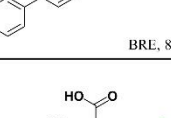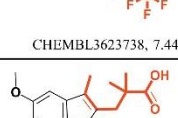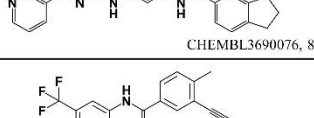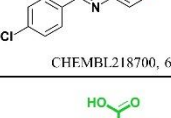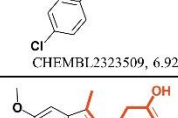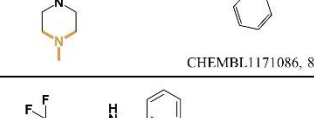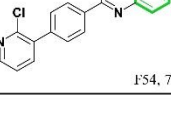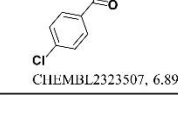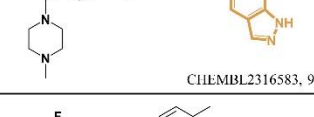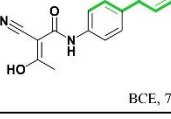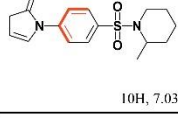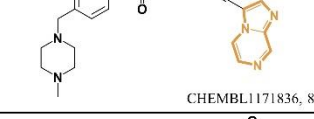
series is easy to rationalize through ΔP, in analogy with the SAR transfer concept. The mean potency of a series is used as a reference and subtracted from single potency values to obtain ΔP (Fig. 8). Compounds that present a positive ΔP value are associated with fragments making favourable contribution to binding. On the contrary, molecules with activity values below the average present an unsuitable substructure for the target. Therefore, ΔPs from different series, if present, are combined in each SAR Table to yield a score for evaluating and ranking fragments. For fragments presenting multiple ΔP values, a weighted average was considered (See Methods). The top-ranked fragments indicate the substructure having the most positive impact on the interaction with the target. Moreover, the transfer series associated with each fragment can be analysed to obtain an insight into the local SAR information. Notably, for fragment analysis, it is important to inspect the substructures in their molecular context. Although fragments are represented by a well-defined substructures, considering them as isolated structures leads to information loss concerning ligand-target interactions. Moreover, the overall methodology is based on available active compounds and X-ray ligands, thus known or significant SAR information might be missing in the overall analysis of fragments, due to data incompleteness or lacking crystallographic structures. Therefore, interpretation of results captures in an SAR Table from a medicinal chemistry perspective supports successful application of the approach.
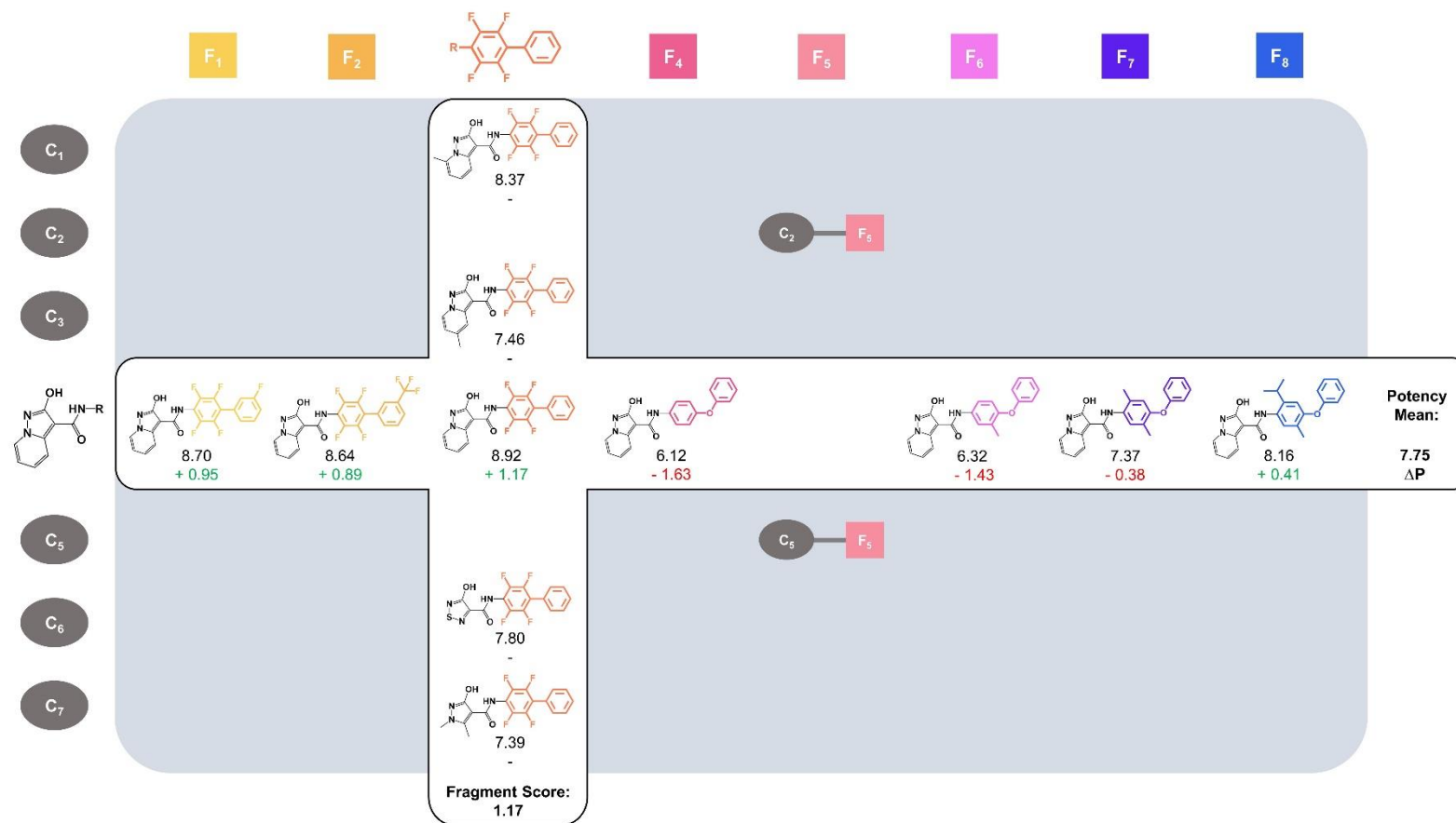
Figure 8. Extract of SAR table relative to *h*DHODH. Logarithmic potency and ΔP are shown under the structures.

2.3.6. SAR Table: Proof of Concept

In the previous chapter, we have shown how SAR Tables enable the evaluation of substructure contribution to binding affinity for series of analogues. In order to investigate the correlation of top-ranked fragments and ligand-target interactions, we have systematically generated and analysed SAR Tables of aldo-keto reductase 3 (AKR1C3, 353 compounds, 42 X-ray), tyrosine-protein kinase ABL (996 compounds, 20 X-ray) and *h*DHODH (643 compounds, 64 X-ray). For each target, shared fragments together with their transfer series were extracted and ranked according their fragment scores. The top10 results provided by SAR Tables are reported in **Table 2**. The fragment structures range from few atoms to larger scaffolds, which is the result of the fragmentation scheme and analogous search method used for the analysis. It is interesting to notice that some of the most potent compounds available for the targets are present in the top 10 fragment lists. This finding indicates that the fragment score successfully accounts for the relevance of shared substructures, since top-ranked solutions highlight most optimized fragments among the explored structures. Moreover, fragments can be inspected on the basis of transfer series to gain insights into SAR progression. **Figure 9** shows transfer series associated with fragment #1 of AKR1C3. The series is composed of three compounds, the best fragment belongs to the crystallographic ligand WDX (PDB ID: 4WDX) and consists of the piperidine substructure. On the basis of the WDX crystallographic structure, (Fig. 10) fragment #1 is not directly involved in the interaction within the binding site However, compared to the other shared fragment its shape if relevant for orienting the hydroxyl group for the interaction with Glu222. The analogues CHEMBL2413856 and CHEMBL2413858 were more than 100 times less active compared to WDX, although they contained similar scaffolds. This potency difference suggests the importance of the WDX H-bonding interaction with Glu222 for activity compared to the other analogues. Another series for AKR1C3 is shown in **Figure 11**. In this case, fragment #6 is shared by two transfer series associated with 2 and 3-aminobezoicacid as core structures, respectively. The combination of p-pentafluorosulfanephenyl on both scaffolds leads to an increase in potency compared to the other derivatives, except for CHEMBL121626. Noteworthy, top-ranked fragments are not always substructures of most potent compounds, but their contribution is consistent among different series.

Target

Rank #	*h*DHODH	AKR1C3	Tyrosine-protein kinase ABL
1	CHEMBL217918, 6.67	CHEMBL2413860, 8.35	CHEMBL3685175, 8.70
2	DUH, 8.92	16J, 7.35	CHEMBL3780219, 7.16
3	CHEMBL218062, 8.22	10H, 7.03	CHEMBL2324924, 9.31
4	QUINOLINE41, 8.01	16J, 7.35	CHEMBL2316869, 9.11
5	QUINOLINE11, 7.30	CHEMBL2323508, 7.05	CHEMBL2316870, 8.88
6	BRE, 8.16	CHEMBL3623738, 7.44	CHEMBL3690076, 8.27
7	CHEMBL218700, 6.68	CHEMBL2323509, 6.92	CHEMBL1171086, 8.57
8	F54, 7.58	CHEMBL2323507, 6.89	CHEMBL2316583, 9.30
9	BCE, 7.50	10H, 7.03	CHEMBL1171836, 8.64
10	QUINOLINE11, 7.29	CHEMBL3623735, 7.24	CHEMBL3894450, 8.16

Table 2. Top 10 ranked fragments in the SAR Table for *h*DHODH, AKR1C3 and tyrosine-protein

kinase ABL. For fragments present in multiple series, only the first structure is shown. Compound names and logarithmic potency are reported. Shared fragments are depicted in bold.

Figure 12 reports transfer series of fragment #6 for tyrosine-protein kinase ABL. Here, the analogues revealed an increase in activity for the substitution of the piperazine with small lipophilic groups, such in CHEBL3690076 and CHEBL3690075.
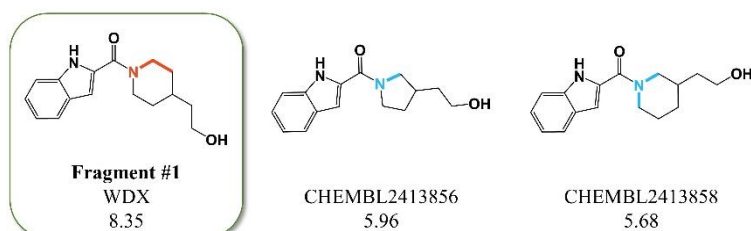
Figure 9. The first fragment ranked in the SAR Table of AKR1C3 together with the relative transfer series. Shared fragments are depicted in bold. Numbers report the logarithmic potency.

On the other hand, the substitution of piperazine ring or the replacement of the hydrogen with carboxyl group led to low binding affinity. Moving to *h*DHODH analysis, the transfer series associated with fragment #2 is detailed in **Figure 13**. This series consists of analogues of meds433, which represents the most potent compound.
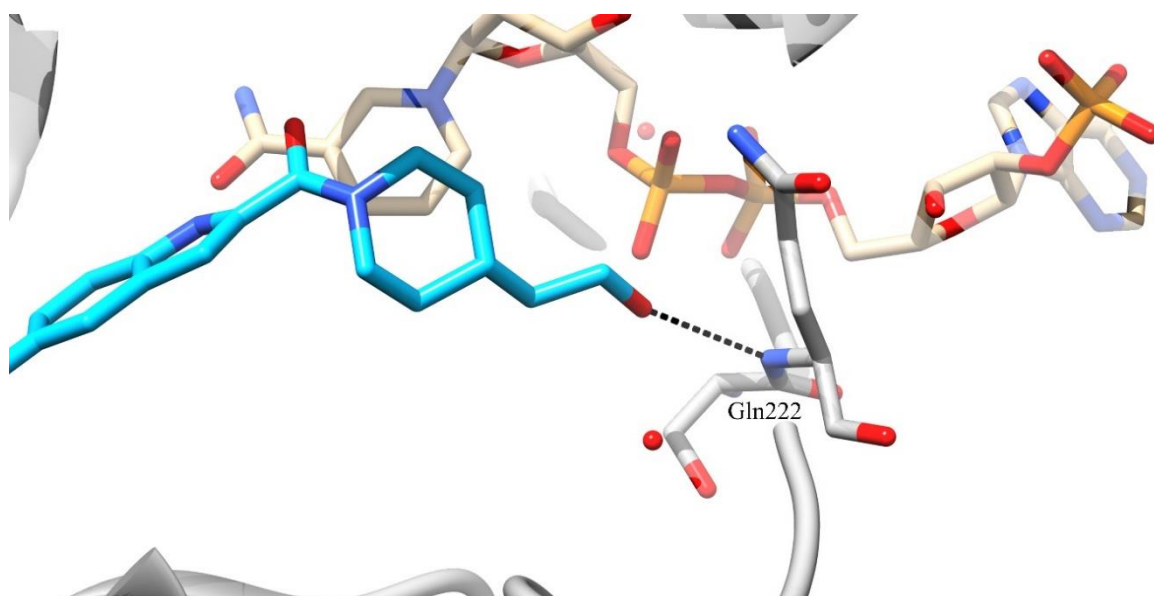
Figure 10. Cocrystal structure of compound WDX (blue) in AKR1C3 binding sites. NADP is represented in beige. Hydrogen bond is shown as dashed lines.

Overall, meds433 core is modulated with three scaffolds: biphenylic, diphenylether, and indole, the last two of which have been analysed and discussed in chapter 1.3.2. Since all the available derivatives are identified in this series, the trend of SAR progression can be easily recognized. **Figure 14** shows the two series captured for fragment #7. For both core structures, the additional bromine atom is beneficial for potency.
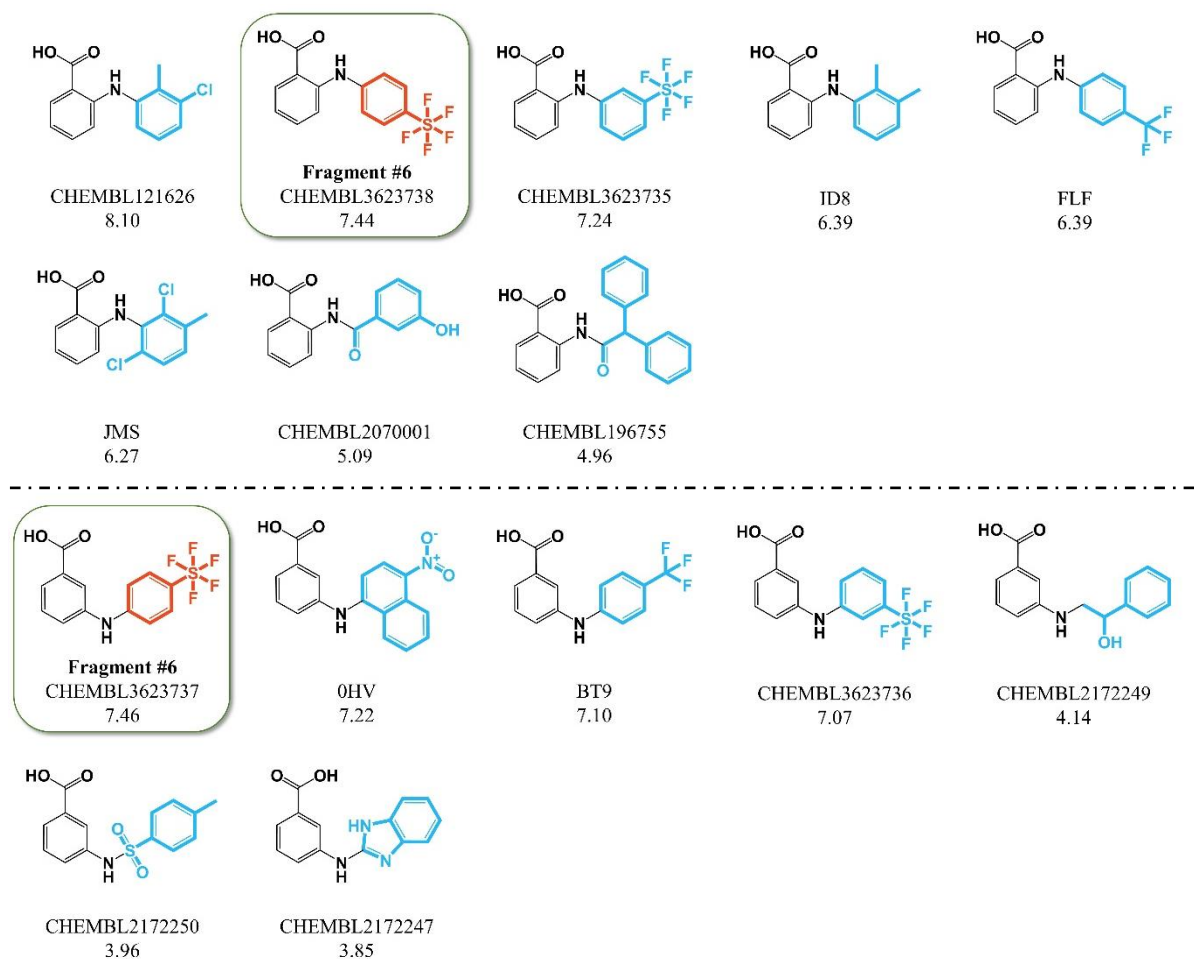
Figure 11. The fragment at rank at position 6 in the SAR Table for AKR1C3 is shown with the related transfer series. Shared fragments are depicted in bold. Numbers report the logarithmic potency.

Taken together, SAR progression if the series clearly indicates that a small substituent in position 6 of quinoline moiety is favoured over bulky group as in CHEMBL387288 and CHEMBL384983, which are only weakly potent. Moreover, the additional substituents in position 5 of quinoline are not well tolerated and leaving the scaffold unsubstituted results in a decrease in potency.
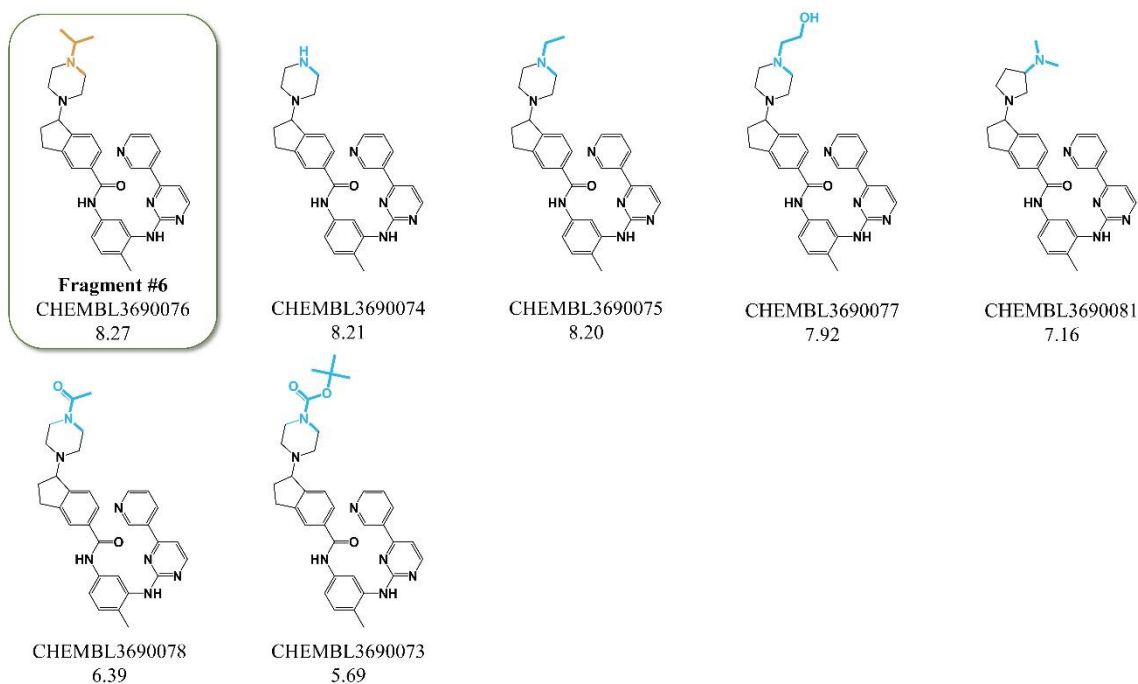
Figure 12. The sixth fragment ranked in the SAR table analysis on tyrosine-protein kinase ABL together with the relative transfer series. Shared fragments are highlighted in bold. Numbers report the logarithmic potency. (see Figure 11)
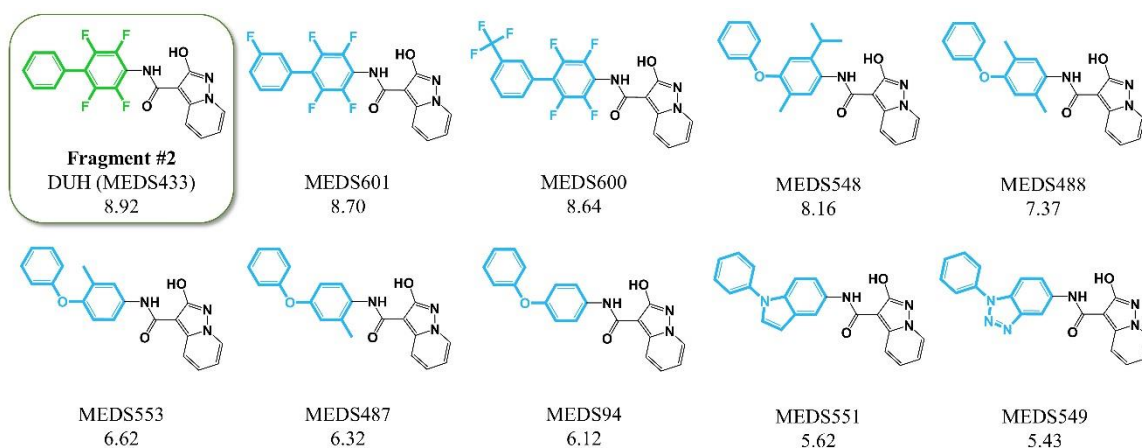
Figure 13. The second fragment ranked in the SAR table analysis on *h*DHODH together with the relative transfer series. Shared fragments are highlighted in bold. Numbers report the logarithmic potency. (see Figure 11)
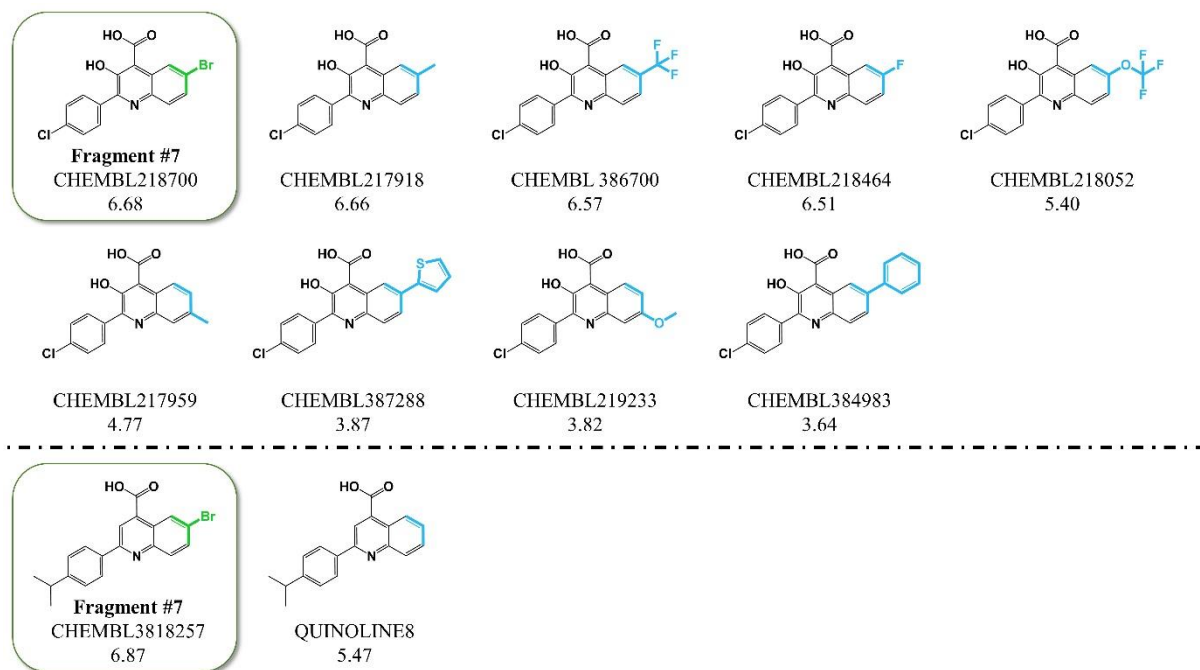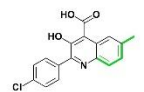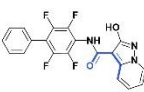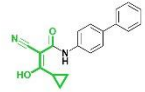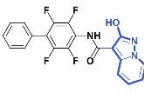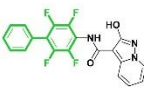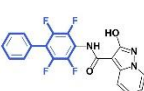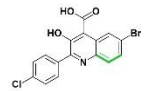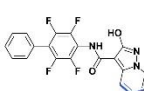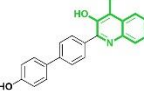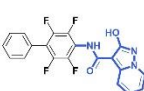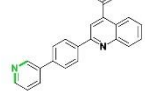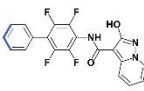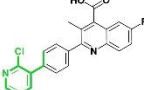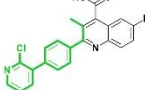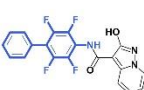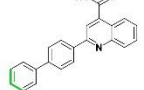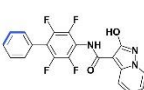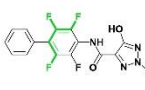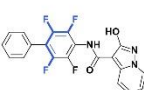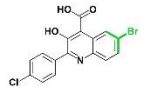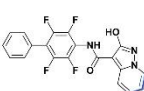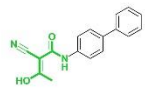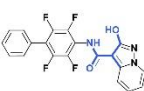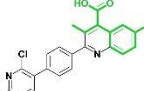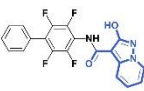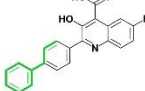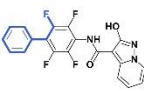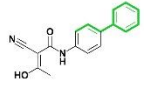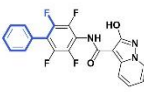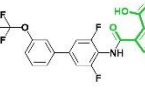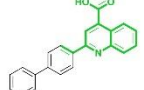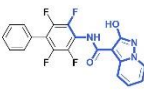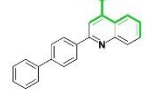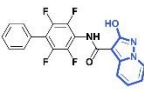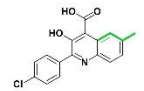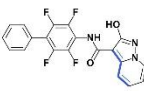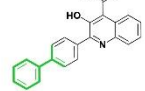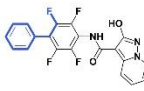
Figure 14. The seventh fragment ranked in the SAR table analysis on *h*DHODH together with the relative transfer series. Shared fragments are highlighted in bold. Numbers report the logarithmic potency. (see Figure 11)

2.3.7. SAR Transfer Applied to Meds433

Lead optimization process involves iterative rounds of synthesis and characterisation of a potential drug candidate to gain insights as to how chemical structure and activity are related in light of interactions with given targets. As illustrated above, scoring and ranking of fragments present in SAR Tables make it possible to characterize SARs and identify fragment structures that significantly contribute to ligand-target interactions. Accordingly, one would exploit such information to guide the optimization process of new ligands. To these end, SAR Tables can be used to transfer selected fragment structures between series or to specific compounds. Herein, we explored the transferability of hDHODH's top-ranked fragment to the inhibitor meds433. The best 25 fragments from SAR Tables were coupled to meds433 shared fragments in **Table 3**. The analysis allows associating each fragment with a specific substructure of meds433. Although the structure-based method utilized to asses shared fragment is designed to identify substructures with similar shape and interactions, practically, the transferability of fragments between different series might be difficult for synthetic reasons. Therefore, in some cases, fragment structures of meds433 cannot be directly replaced, requiring additional analysis to achieve similar SAR progression for the target compound. It should be considered that different shared fragments in **Table 3** represent the chemical modification applied in Chapter 1. For instance, fragment #4 is defined by a 2-

chloropyridine structure and is found in the brequinar analogue with a logarithmic potency value of 8.01. For meds433, the shared fragment is represented by the distal phenyl ring alignment of which with brequinar analogues is well established in different crystallographic structures (Fig. 7). Then, shared fragment of meds433 might be replaced with fragment #4 to design new analogues for the series developed in Chapter 1.3.3. In addition, fragment #7 yields another possible analogue of meds433. This substructure is defined by a bromine substitution at position 6 of quinoline-4-carboxylic acid derivative. According to the SAR transfer principle, we can apply the same modulation on meds433, replacing a hydrogen with a bromine atom at the position 6 of the hydroxypyrazolo[1,5-a]pyridine moiety. Shared fragments are expected to form similar interactions within the binding site, although this cannot be predicted with certainty given the limited data on which transfer analysis is usually based.

	Ranked fragments	Meds433 shared fragment		Ranked fragments	Meds433 shared fragment
#1	CHEMBL217918, 6.67		#13	MD7, 7.72	
#2	DUH, 8.92		#14	CHEMBL218700, 6.68	
#3	CHEMBL218062, 8.22		#16	QUINOLINE14, 6.40	
#4	QUINOLINE41, 8.01		#17	QUINOLINE41, 8.01	
#5	QUINOLINE11, 7.30		#18	HYT, 7.35	
#7	CHEMBL218700, 6.68		#19	BCE, 7.50	
#8	F54, 7.58		#21	CHEMBL41719, 8.30	
#9	BCE, 7.50		#22	ILF, 8.70	
#10	QUINOLINE11, 7.29		#23	QUINOLINE11, 7.29	
#11	CHEMBL217918, 6.67		#24	CHEMBL219648, 8.00	
#12	QUINOLINE11, 7.30		#25	MD7, 7.72	

Table 3. Top 25 ranked fragments associated with mdes433's shared substructures. Compound names

and logarithmic potency are reported. Shared fragments are depicted in bold.

The four available substituent positions in pyridine substructure of meds433 were investigated and the most promising analogues were synthetized. Predicted binding affinity favoured the addition of a halogen atom such as chlorine proximal to position 6. Assessment of this prediction awaits the availability of the corresponding experimental data. Finally, the fragment #25 in **Table 3** is present in the brequinar analogue F51 with logarithmic potency of 7.55. The related shared fragment of meds433 is the 2-hydroxypyrazolo[1,5-a]pyridine motif, which was originally designed to mimic carboxylic acid function of brequinar.[24] Consequently, the two shared fragments show a good superimposition inside the binding pocket.
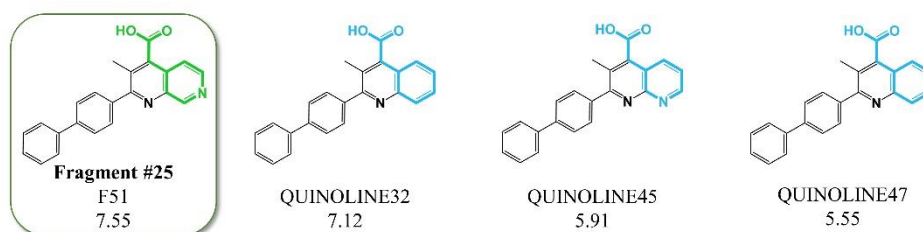
Figure 15. The twenty-fifth fragment ranked in the SAR table analysis on *h*DHODH together with the relative transfer series. Shared fragments are highlighted in bold. Numbers report the logarithmic potency. (see Figure 11)

The interesting feature of fragment #25 is the additional nitrogen in position 7 of the quinoline heterocycle. Inspecting the series in light of this substructure (Fig. 15) reveals that analogies only differ by the position or the presence of the nitrogen atom in the quinoline moiety. Therefore, the SAR transfer series clearly indicates that the introduction of this nitrogen atom yields an additional interaction with the target which improves activity. Such an interaction is confirmed by X-ray structure of compound F51 in complex with *h*DHODH.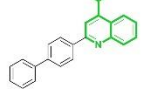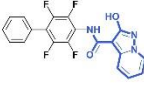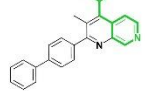 The nitrogen incorporated into the 1,7-naphthyridine ring is within 3.3 Å of the Tyr356 hydroxyl group, forming an additional H-bonding interaction within the binding site (Fig. 16).[75] Following the SAR transfer concept, one would introduce a similar modification in meds433. An analogue of the query compound with an additional nitrogen added to the hydroxypyrazolo[1,5-a]pyridine moiety was already proposed in Chapter 1.3.1. Moreover, *in silico* studies carried out on this compound supported the idea of hydrogen bond transfer of the naphthyridine derivative to meds433, which also waits experimental confirmation.
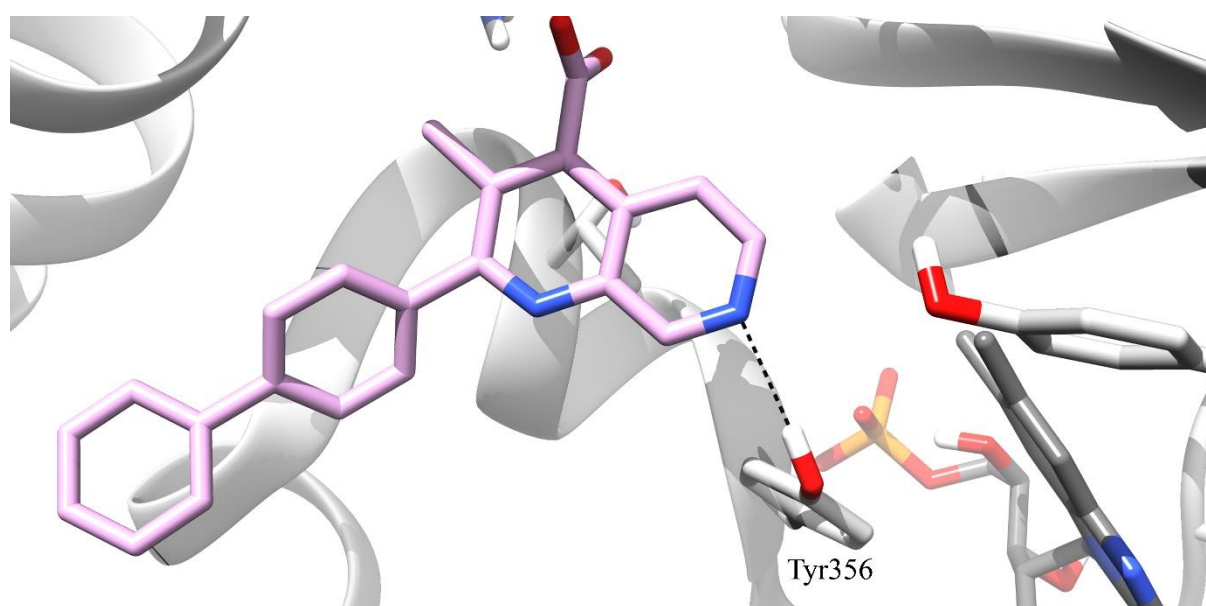
Figure 16. Cocrystal structure of compound F51 (pink) with *h*DHODH b. FAD is represented in grey. The hydrogen bond is shown as a dashed line.

2.4. Conclusion

Lead optimization is the one of the most important steps in drug development. The aim is to explore key interactions between the receptor and the ligand and utilize these insights in the design of new analogues with improved potency, selectivity or other pharmacological properties. SAR transfer analysis offers opportunities to apply knowledge obtained from one compound series to another. The application of the SAR transfer concept also provides possible alternatives when encountering roadblocks during lead optimization process. In addition, analysis of SAR transfer events is of interest from a basic science perspective to better understand how different compound classes might yield similar SAR progression for a given target. This ultimately depends on the nature of ligand-target interactions. Accordingly, one would also like to take structural features of targets and ligand-target interactions into account when studying SAR transfer. However, so far SAR transfer has only been investigated at the level of compounds and analog series. In this work, we have introduced a new computational methodology for structure-based identification and analysis of SAR transfer events. The approach involves a unique three-dimensional molecular fragmentation and recombination scheme for defining substituent fragments and core structures in pairs of superimposed crystallographic ligands and for identifying matching analogues. Importantly, for a given ligand pair, multiple shared fragments and cores are usually obtained, which are used to identify different analogues and associated SAR events. Thus, by design, the approach makes it possible to map multiple analogues with different core-substituent

combinations, which represents a unique feature of the approach. In our proof-of-concept investigation, we have deliberately considered analogue ensembles consisting of two (crystallographic/database) pairs to assess the suitability of the approach. Extending the comparisons by adding further analogues from compound database search is straightforward and relevant for practical applications. We have compared crystallographic ligands bound to a variety of targets and obtained multiple analogue ensembles in each case, often with very large differences in potency. On the basis of $\Delta\Delta P$ calculations, SAR transfer events were identified with high frequency, confirming the utility of the approach and providing opportunities for follow-up investigations of SAR transfer and compound optimization. To examine further the impact of SAR transfer for practical applications, single SAR transfer events were combined on the basis of aligned ligands. This resulted in newly introduced SAR Tables, which organized ensemble of analogues according their core structure and shared fragments. With the aim of evaluating shared substructures, we designed a score to assess the contribution of fragments for each series of analogues (transfer series). The approach was tested on three different targets in order to assess the reliability of the method. Several top-ranked fragments were associated with substructures of the most potent compounds, indicating the ability of the approach to detect relevant SARs. Moreover, the analysis of transfer series highlights the presence of specific ligand-target interactions, providing insight into SAR progression. Finally, we investigated the transferability of shared fragment among different series. Shared fragments of meds433 were coupled to the best substructure detected for *h*DHODH, providing specific suggestions for new meds433 analogues. Some of these suggestions have been already analyzed in previous lead optimization studies, supporting the utility of SAR transfer analysis for practical optimization efforts.

2.5. Methods

2.5.1. Target-ligand X-ray Structures

For each of 20 targets (**Table 1**) as well as the targets used for the generation of SAR Tables , X-ray structures of complexes with different ligands were retrieved from the RCSB Protein Data Bank (PDB).[76] Targets were required to originate from different families.[77] In addition, multiple complex X-ray structures and other active compounds had to be available. X-ray structures were only selected if they were determined at a resolution higher than 3.0 Å. For the 20 targets, a total of 80 qualifying complex structures were chosen. For all crystallographic ligands, potency (K_i, K_d, and/or IC_{50}) values were extracted from BindingDB,[78] PDBind,[79] and ChEMBL.[80] Preference was given to assay-independent equilibrium or constants whenever available. Multiple measurements of the same type had to fall within the same order of magnitude; otherwise, the ligand was discarded. If the measurements fell within one order of magnitude, the mean was calculated as the final potency value. In order to obtain $\Delta\Delta P$ values for pairs of X-ray ligands and structural analogues, it was in some instances required to numerically compare different types of potency measurements because equilibrium constants were not available for all crystallographic ligands. This approximation made it possible to obtain 20 $\Delta\Delta P$ values for each target. Structures of the same target with different ligands were superimposed using the multiple alignment function of the Molecular Operating Environment.[81] The superposition yielded the spatial alignment of bound ligands for further analysis.

2.5.2. Candidate Compounds

As a source for identifying structural analogues of crystallographic ligands, candidate compounds with activity against the targets discussed here were selected from ChEMBL (release 24). Only active compounds with reported direct target interactions (ChEMBL target relationship type "D"), highest assay confidence (confidence score 9), and numerically defined equilibrium constants were considered. Approximate potency annotations such as ">", "<", and "~" were discarded. If multiple potency measurements were available for a compound and target, the mean was calculated as the final potency value if all values fell within the same order of magnitude. ChEMBL dataset was augmented with additional activity values for hDHODH and AKR1C3 collected from literature.[24, 75, 82, 83] Compounds named "QUINOLINE" were obtained from Madak *et al.*,[75] and "MEDS" from Sainas *et al.*[24] or generated in-house (unpublished data) .

2.5.3. Systematic Compound Comparison

For systematically comparing crystallographic ligands and other active compounds and identifying alternative cores and fragments, a new compound fragmentation and recombination scheme was developed. **Figure 7** below illustrates the approach detailed in the following step-by-step using a practical example.

Identification of shared basic fragments

From aligned crystallographic ligands, three-dimensional basic scaffolds (BSCs) were extracted by only retaining C, O, N, S, P, and B atoms, replacing these atoms with carbons, and converting double and triple bonds to single bonds. BSCs including ring structures they contained were then systematically fragmented through bond deletions to generate all possible substructures consisting of three to six heavy atoms. These substructures (fragments) were stored with their atomic coordinates. The fragmentation procedure was carried out using in-house Python code. For each ensemble of ligands, pairwise fragment comparisons were carried out. For example, if three crystallographic ligands A, B, and C were available, fragments of the same size (number of atoms) were compared for pairs A/B, B/C, and A/C. In each case, pairs of fragments with root mean square deviation (RMSD) values of at most 1.0 Å were retained as shared basic fragments and mapped back to the original ligand structure.

Assignment of core structures and shared extended fragments

To consistently define core structures and shared substituent fragments for pairs of crystallographic ligands, a structural organization scheme was applied following two design principles specific to our analysis:

(i) Shared basic fragments identified through systematic BSC comparison were recombined with available R-groups, yielding shared extended fragments (which are simply referred to as shared fragments in the Results section).

(ii) For pairs of X-ray ligands, multiple cores and shared extended fragment were constructed, if possible.
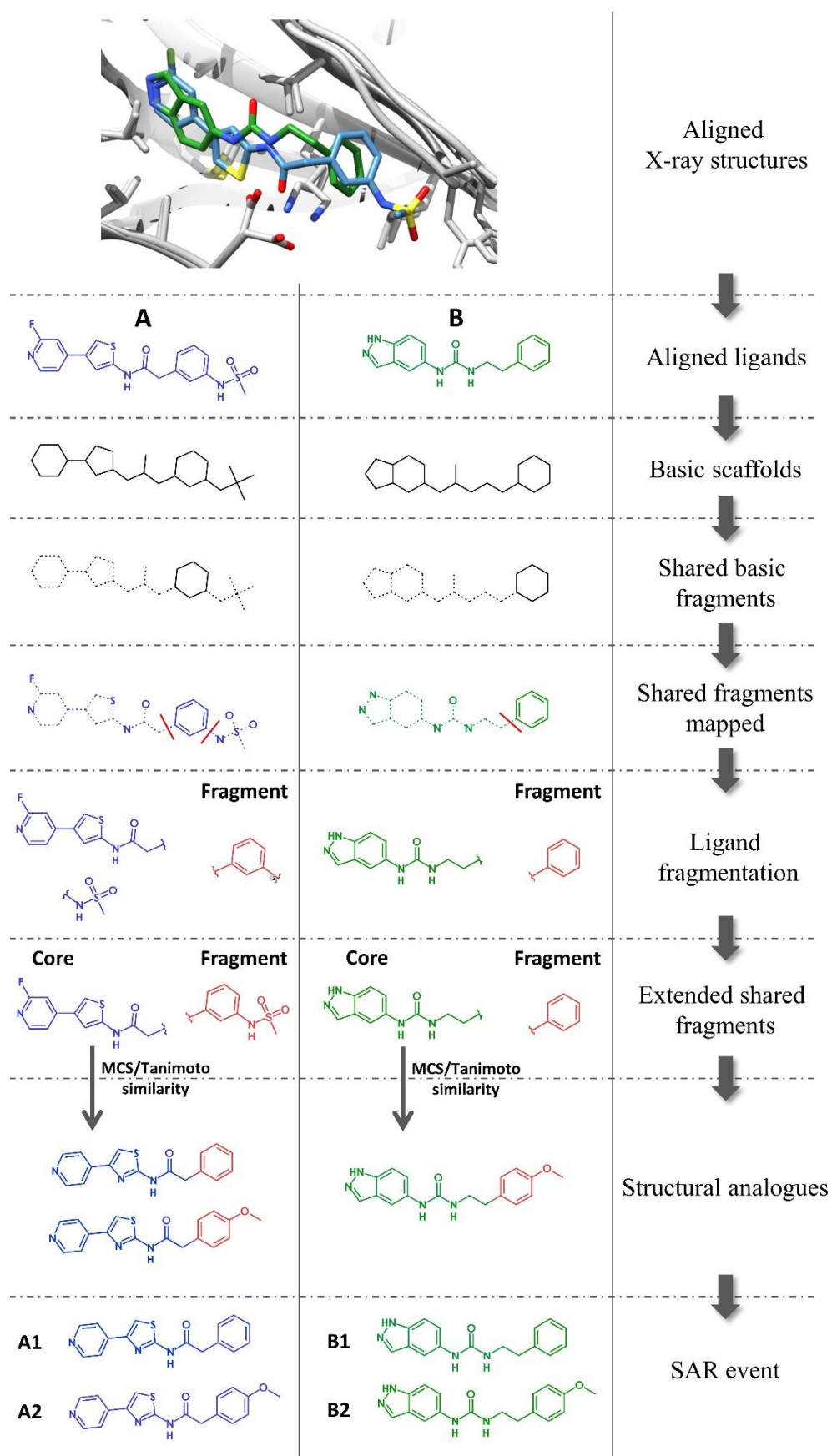
Figure 7. From ligand comparison to the detection of SAR events. The representation illustrates individual steps of the methodology detailed in the text including ligand comparison, fragmentation,

recombination, and identification of analogues, ultimately leading to the generation of analogue ensembles for the assessment of SAR events.

To these ends, the structural context of mapped shared basic fragments was systematically analyzed by detecting all bound substructures, regardless of bond orders. Substructures with at most five heavy atoms were assigned to the fragment, while substructures with more than five atoms were assigned to the core. Atoms forming core-fragment bonds were indexed.

Identification of structural analogues

A search for structural analogues of X-ray ligands with shared fragments was carried out using candidate compounds from ChEMBL. Therefore, the core(s) of each ligand were used for maximum common substructure (MCS)19 calculations and compounds containing the core and matching all core atoms bound to the fragment were selected MCS calculations were carried out in Python with the aid of the OpenEye chemistry toolkit.[20] Tanimoto similarity[21] using tree fingerprints, which systematically generate all possible unique trees from a molecular graph,[19] was calculated to compare each core and all MCSs that were identified. MCSs with at least 50% similarity were retained and ranked in the order of decreasing similarity. Ranked MCSs represented alternative cores and the residual structures of the corresponding compounds represented alternative shared fragments.

2.5.4. SAR Transfer Criteria and Different SAR Events

Logarithmic ΔP values were calculated for pairs of analogues sharing the same core. In addition, $\Delta\Delta P$ values were calculated as the difference between ΔP values for pairs of compounds containing the same fragments. For SAR transfer assessment, logarithmic $\Delta\Delta P$ values were assigned to three different categories describing different SAR events:

(i) $|\Delta\Delta P| \leq x$: SAR transfer.

(ii) $x < |\Delta\Delta P| \leq 2x$, both ΔPs values have the same (+/-): SAR trend.

(iii) $x < |\Delta\Delta P|$, ΔPs values have opposite signs: SAR inversion.

So-defined SAR events were calculated for three logarithmic $|\Delta\Delta P|$ thresholds (x values) of 0.5, 0.7 and 1.0, which approximately correspond to 3-, 5-, and 10-fold relative changes in potency, respectively.

2.5.5. SAR Table Composition

For systematically generating SAR Tables, a similar scheme was used to compare candidate compounds for SAR transfer.

Identification of shared basic fragments

From aligned crystallographic ligands, three-dimensional basic scaffolds (BSCs) were extracted by only retaining C, O, N, S, P, and B atoms, replacing these atoms with carbons, and converting double and triple bonds to single bonds. BSCs including ring structures they contained were then systematically fragmented through bond deletions to generate seven possible substructures consisting of one the fallowing: propane, butane, pentane, hexane, isobutane, isopentane, 2-methylpentane. According to the data retrieved for the previous SAR transfer analysis, more than 98% of analogues found by analogues search were accountable to the seven selected shared basic fragments. Therefore, all the possible substructures consisting of three to six heavy atoms used in SAR transfer method were reduced to seven to simplify further analysis and avoid redundancy. These substructures (fragments) were stored with their atomic coordinates. The fragmentation procedure was carried out using in-house Python code. For each ensemble of ligands, pairwise fragment comparisons were carried out. In each case, pairs of fragments with root mean square deviation (RMSD) values of at most 1.0 Å were retained as shared basic fragments and mapped back to the original ligand structure.

Shared basic fragment classification

For each of the seven fragment types, all basic shared fragment retrieved from pairwise ligand comparison were recorded in an all vs. all matrix. A value of "1" was set for each combination of fragments with RMSD value of at most 1.0 Å, "0" was used otherwise. A total of seven matrixes was generated for different targets.

Identification of structural analogues and SAR table organization

For each matrix, basic shared fragment with binary value of "1" were retained and mapped back to the original ligand structure. Then, to define core structures and shared substituent fragments for the ensemble of crystallographic ligands, the same structural organization scheme as for SAR transfer analysis was applied. Structural analogues of X-ray ligands with shared fragments were identified as described for retrieval of SAR transfer events., resulting in an ensemble of X-ray ligands and analogues. Each compound was represented by a core

structures and a shared fragment. From the pool of compounds, SAR Table was built as follows:

(i)   Unique core structures were used as row indices

(ii)   Unique shared fragments were used to define columns

Tables are filled with the compounds representing given core-fragment combinations. At least one compound had to be present for each row and column and potency values were recorded.

Fragment score calculation

Logarithmic ΔP values were calculated in SAR Tables for each compound belonging to a transfer series with at least two compounds. For each SAR Table, average row ΔP values were determined and subtracted from individual potency values. Fragment scores were obtained for columns with at least one ΔP value. In case of a single ΔP fragment, the score was equal to this value. For multiple ΔP values, the fragment score was calculated as a weighted average according to the following formula:

$$Fragment\ score\ =\ \frac{\sum_{i=1}^{n} \Delta P_i \cdot \omega_i}{\sum_{i=1}^{n} \omega_i}$$

ΔP_i = Series ΔP

ω_i = number of compounds in the series

2.5.6. Meds433 Shared Fragments

Fragment scores were calculated and sorted and SAR tables for the top 25 fragments were generated. Only SARTables containing meds433 were considered. The transfer series associated with selected fragments and fragments shared with meds433 were iteratively retrieved from SAR Tables using in-house Python code.

3. Bibliography

1. Pastor-Anglada, M.; Perez-Torras, S., Nucleoside transporter proteins as biomarkers of drug responsiveness and drug targets. *Front. Pharmacol.* **2015,** *6*, 13.

2. Evans, D. R.; Guy, H. I., Mammalian pyrimidine biosynthesis: fresh insights into an ancient pathway. *J. Biol. Chem.* **2004,** *279* (32), 33035-8.

3. Dorasamy, M. S.; Choudhary, B.; Nellore, K.; Subramanya, H.; Wong, P. F., Dihydroorotate dehydrogenase Inhibitors Target c-Myc and Arrest Melanoma, Myeloma and Lymphoma cells at S-phase. *J. Cancer* **2017,** *8* (15), 3086-3098.

4. Jones, M. E., Pyrimidine nucleotide biosynthesis in animals: genes, enzymes, and regulation of UMP biosynthesis. *Annu Rev Biochem* **1980,** *49*, 253-79.

5. Carrey, E. A.; Dietz, C.; Glubb, D. M.; Loffler, M.; Lucocq, J. M.; Watson, P. F., Detection and location of the enzymes of de novo pyrimidine biosynthesis in mammalian spermatozoa. *Reproduction* **2002,** *123* (6), 757-68.

6. Lucas-Hourani, M.; Dauzonne, D.; Munier-Lehmann, H.; Khiar, S.; Nisole, S.; Dairou, J.; Helynck, O.; Afonso, P. V.; Tangy, F.; Vidalain, P. O., Original Chemical Series of Pyrimidine Biosynthesis Inhibitors That Boost the Antiviral Interferon Response. *Antimicrob Agents Chemother* **2017,** *61* (10).

7. Knecht, W.; Henseling, J.; Loffler, M., Kinetics of inhibition of human and rat dihydroorotate dehydrogenase by atovaquone, lawsone derivatives, brequinar sodium and polyporic acid. *Chem. Biol. Interact.* **2000,** *124* (1), 61-76.

8. Liu, S.; Neidhardt, E. A.; Grossman, T. H.; Ocain, T.; Clardy, J., Structures of human dihydroorotate dehydrogenase in complex with antiproliferative agents. *Structure* **2000,** *8* (1), 25-33.

9. Reis, R. A. G.; Calil, F. A.; Feliciano, P. R.; Pinheiro, M. P.; Nonato, M. C., The dihydroorotate dehydrogenases: Past and present. *Arch Biochem Biophys* **2017,** *632*, 175-191.

10. Singh, A.; Maqbool, M.; Mobashir, M.; Hoda, N., Dihydroorotate dehydrogenase: A drug target for the development of antimalarials. *Eur J Med Chem* **2017,** *125*, 640-651.

11. Baumgartner, R.; Walloschek, M.; Kralik, M.; Gotschlich, A.; Tasler, S.; Mies, J.; Leban, J., Dual binding mode of a novel series of DHODH inhibitors. *J. Med. Chem.* **2006,** *49* (4), 1239-47.

12. Sanders, S.; Harisdangkul, V., Leflunomide for the treatment of rheumatoid arthritis and autoimmunity. *Am. J. Med. Sci.* **2002,** *323* (4), 190-3.

13. Kalgutkar, A. S.; Nguyen, H. T.; Vaz, A. D.; Doan, A.; Dalvie, D. K.; McLeod, D. G.; Murray, J. C., In vitro metabolism studies on the isoxazole ring scission in the anti-inflammatory agent leflunomide to its active alpha-cyanoenol metabolite A771726: mechanistic similarities with the cytochrome P450-catalyzed dehydration of aldoximes. *Drug Metab. Dispos.* **2003,** *31* (10), 1240-50.

14. Teriflunomide (AUBAGIO). Multiple sclerosis: just a metabolite of leflunomide. *Prescrire Int.* **2015,** *24* (158), 61-4.

15. Li, E. K.; Tam, L. S.; Tomlinson, B., Leflunomide in the treatment of rheumatoid arthritis. *Clin. Ther.* **2004,** *26* (4), 447-59.

16. Smith, L. H., Jr.; Baker, F. A., Pyrimidine metabolism in man. I. The biosynthesis of orotic acid. *J. Clin. Invest.* **1959,** *38* (5), 798-809.

17. Smith, L. H., Jr.; Baker, F. A.; Sullivan, M., Pyrimidine metabolism in man. II. Studies of leukemic cells. *Blood* **1960,** *15*, 360-9.

18. Dexter, D. L.; Hesson, D. P.; Ardecky, R. J.; Rao, G. V.; Tippett, D. L.; Dusak, B. A.; Paull, K. D.; Plowman, J.; DeLarco, B. M.; Narayanan, V. L.; et al., Activity of a novel 4-quinolinecarboxylic acid, NSC 368390 [6-fluoro-2-(2'-fluoro-1,1'-biphenyl-4-yl)-3-methyl-4-quinolinecarb oxylic acid sodium salt], against experimental tumors. *Cancer Res.* **1985,** *45* (11 Pt 1), 5563-8.

19. Schwartsmann, G.; Peters, G. J.; Laurensse, E.; de Waal, F. C.; Loonen, A. H.; Leyva, A.; Pinedo, H. M., DUP 785 (NSC 368390): schedule-dependency of growth-inhibitory and antipyrimidine effects. *Biochem. Pharmacol.* **1988,** *37* (17), 3257-66.

20. Cramer, D. V., Brequinar sodium. *Pediatr. Nephrol.* **1995,** *9 Suppl*, S52-5.

21. Sykes, D. B.; Kfoury, Y. S.; Mercier, F. E.; Wawer, M. J.; Law, J. M.; Haynes, M. K.; Lewis, T. A.; Schajnovitz, A.; Jain, E.; Lee, D.; Meyer, H.; Pierce, K. A.; Tolliday, N. J.; Waller, A.; Ferrara, S. J.; Eheim, A. L.; Stoeckigt, D.; Maxcy, K. L.; Cobert, J. M.; Bachand, J.; Szekely, B. A.; Mukherjee, S.; Sklar, L. A.; Kotz, J. D.; Clish, C. B.; Sadreyev, R. I.; Clemons, P. A.; Janzer, A.; Schreiber, S. L.; Scadden, D. T., Inhibition of Dihydroorotate Dehydrogenase Overcomes Differentiation Blockade in Acute Myeloid Leukemia. *Cell* **2016,** *167* (1), 171-186 e15.

22. Zhou, J.; Quah, J. Y.; Ng, Y.; Chooi, J. Y.; Toh, S. H.; Lin, B.; Tan, T. Z.; Hosoi, H.; Osato, M.; Seet, Q.; Ooi, A. G. L.; Lindmark, B.; McHale, M.; Chng, W. J., ASLAN003, a potent dihydroorotate dehydrogenase inhibitor for differentiation of acute myeloid leukemia. *Haematologica* **2019**.

23. Christian, S.; Merz, C.; Evans, L.; Gradl, S.; Seidel, H.; Friberg, A.; Eheim, A.; Lejeune, P.; Brzezinka, K.; Zimmermann, K.; Ferrara, S.; Meyer, H.; Lesche, R.; Stoeckigt, D.; Bauser, M.; Haegebarth, A.; Sykes, D. B.; Scadden, D. T.; Losman, J. A.; Janzer, A., The novel dihydroorotate dehydrogenase (DHODH) inhibitor BAY 2402234 triggers differentiation and is effective in the treatment of myeloid malignancies. *Leukemia* **2019,** *33* (10), 2403-2415.

24. Sainas, S.; Pippione, A. C.; Lupino, E.; Giorgis, M.; Circosta, P.; Gaidano, V.; Goyal, P.; Bonanni, D.; Rolando, B.; Cignetti, A.; Ducime, A.; Andersson, M.; Jarva, M.; Friemann, R.; Piccinini, M.; Ramondetti, C.; Buccinna, B.; Al-Karadaghi, S.; Boschi, D.; Saglio, G.; Lolli, M. L., Targeting Myeloid Differentiation Using Potent 2-Hydroxypyrazolo[1,5- a]pyridine Scaffold-Based Human Dihydroorotate Dehydrogenase Inhibitors. *J. Med. Chem.* **2018,** *61* (14), 6034-6055.

25. Tan, J. L.; Fogley, R. D.; Flynn, R. A.; Ablain, J.; Yang, S.; Saint-Andre, V.; Fan, Z. P.; Do, B. T.; Laga, A. C.; Fujinaga, K.; Santoriello, C.; Greer, C. B.; Kim, Y. J.; Clohessy, J. G.; Bothmer, A.; Pandell, N.; Avagyan, S.; Brogie, J. E.; van Rooijen, E.; Hagedorn, E. J.; Shyh-Chang, N.; White, R. M.; Price, D. H.; Pandolfi, P. P.; Peterlin, B. M.; Zhou, Y.; Kim, T. H.; Asara, J. M.; Chang, H. Y.; Young, R. A.; Zon, L. I., Stress from Nucleotide Depletion Activates the Transcriptional Regulator HEXIM1 to Suppress Melanoma. *Mol. Cell* **2016,** *62* (1), 34-46.

26. White, R. M.; Cech, J.; Ratanasirintrawoot, S.; Lin, C. Y.; Rahl, P. B.; Burke, C. J.; Langdon, E.; Tomlinson, M. L.; Mosher, J.; Kaufman, C.; Chen, F.; Long, H. K.; Kramer, M.; Datta, S.; Neuberg, D.; Granter, S.; Young, R. A.; Morrison, S.; Wheeler, G. N.; Zon, L. I., DHODH modulates transcriptional elongation in the neural crest and melanoma. *Nature* **2011,** *471* (7339), 518-22.

27. Lin, M.; Chen, B., Advances in the drug therapies of acute myeloid leukemia (except acute wpromyelocytic leukemia). *Drug Des. Devel. Ther.* **2018,** *12*, 1009-1017.

28. Erba, H. P., Finding the optimal combination therapy for the treatment of newly diagnosed AML in older patients unfit for intensive therapy. *Leuk. Res.* **2015,** *39* (2), 183-91.

29. Medinger, M.; Lengerke, C.; Passweg, J., Novel therapeutic options in Acute Myeloid Leukemia. *Leuk Res Rep* **2016,** *6*, 39-49.

30. Lewis, T. A.; Sykes, D. B.; Law, J. M.; Munoz, B.; Rustiguel, J. K.; Nonato, M. C.; Scadden, D. T.; Schreiber, S. L., Development of ML390: A Human DHODH Inhibitor That Induces Differentiation in Acute Myeloid Leukemia. *ACS Med. Chem. Lett.* **2016,** *7* (12), 1112-1117.

31. Xin, J.; You, D.; Breslin, P.; Li, J.; Zhang, J.; Wei, W.; Cannova, J.; Volk, A.; Gutierrez, R.; Xiao, Y.; Ni, A.; Ng, G.; Schmidt, R.; Xia, Z.; Pan, J.; Chen, H.; Patel, M. M.; Kuo, P. C.; Nand, S.; Kini, A. R.; Zhang, J.; Chen, J.; Zhu, J.; Zhang, J., Sensitizing acute myeloid leukemia cells to induced differentiation by inhibiting the RIP1/RIP3 pathway. *Leukemia* **2017,** *31* (5), 1154-1165.

32. Bonomo, S.; Tosco, P.; Giorgis, M.; Lolli, M.; Fruttero, R., The role of fluorine in stabilizing the bioactive conformation of dihydroorotate dehydrogenase inhibitors. *J. Mol. Model.* **2013,** *19* (3), 1099-107.

33. Lolli, M. L.; Giorgis, M.; Tosco, P.; Foti, A.; Fruttero, R.; Gasco, A., New inhibitors of dihydroorotate dehydrogenase (DHODH) based on the 4-hydroxy-1,2,5-oxadiazol-3-yl (hydroxyfurazanyl) scaffold. *Eur J Med Chem* **2012,** *49*, 102-9.

34. Kavianpour, M.; Ahmadzadeh, A.; Shahrabi, S.; Saki, N., Significance of oncogenes and tumor suppressor genes in AML prognosis. *Tumour Biol.* **2016,** *37* (8), 10041-52.

35. Williams-Noonan, B. J.; Yuriev, E.; Chalmers, D. K., Free Energy Methods in Drug Design: Prospects of "Alchemical Perturbation" in Medicinal Chemistry. *Journal of Medicinal Chemistry* **2017**.

36. Sainas, S.; Pippione, A. C.; Giorgis, M.; Lupino, E.; Goyal, P.; Ramondetti, C.; Buccina, B.; Piccinini, M.; Braga, R. C.; Andrade, C. H.; Andersson, M.; Moritzer, A. C.; Friemann, R.; Mensa, S.; Al-Kadaraghi, S.; Boschi, D.; Lolli, M. L., Design, synthesis, biological evaluation and X-ray structural studies of potent human dihydroorotate dehydrogenase inhibitors based on hydroxylated azole scaffolds. *Eur J Med Chem* **2017,** *129*, 287-302.

37. Bonomo, S.; Tosco, P.; Giorgis, M.; Lolli, M.; Fruttero, R., The role of fluorine in stabilizing the bioactive conformation of dihydroorotate dehydrogenase inhibitors. *J. Mol. Model.* **2013,** *19* (3), 1099-1107.

38. Baumgartner, R.; Walloschek, M.; Kralik, M.; Gotschlich, A.; Tasler, S.; Mies, J.; Leban, J., Dual Binding Mode of a Novel Series of DHODH Inhibitors. *J. Med. Chem.* **2006,** *49* (4), 1239-1247.

39. Sainas, S.; Pippione, A. C.; Giorgis, M.; Lupino, E.; Goyal, P.; Ramondetti, C.; Buccina, B.; Piccinini, M.; Braga, R. C.; Andrade, C. H.; Andersson, M.; Moritzer, A.-C.; Friemann, R.; Mensa, S.; Al-Kadaraghi, S.; Boschi, D.; Lolli, M. L., Design, synthesis, biological evaluation and X-ray structural studies of potent human dihydroorotate dehydrogenase inhibitors based on hydroxylated azole scaffolds. *Eur. J. Med. Chem.* **2017,** *129*, 287-302.

40. Das, P.; Deng, X.; Zhang, L.; Roth, M. G.; Fontoura, B. M. A.; Phillips, M. A.; De Brabander, J. K., SAR-Based Optimization of a 4-Quinoline Carboxylic Acid Analogue with Potent Antiviral Activity. *ACS Med. Chem. Lett.* **2013,** *4* (6), 517-521.

41. Waring, M. J., Lipophilicity in drug discovery. *Expert Opin Drug Discov* **2010,** *5* (3), 235-48.

42. **Schrödinger Release 2017-3:** Maestro, Schrödinger, LLC, New York, NY, 2017.

43. Kujawski, J.; Bernard, M. K.; Jodlowska, E.; Czaja, K.; Drabinska, B., On the Interactions of Leflunomide and Teriflunomide Within Receptor Cavity-Nmr Studies and Energy Calculations. *J Mol Model* **2015,** *21* (5), 105 - 117.

44. 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.

45. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **2015,** *11* (8), 3696-713.

46. D.A. Case; D.S. Cerutti; T.E. Cheatham, I.; T.A. Darden; R.E. Duke; T.J. Giese; H. Gohlke; A.W. Goetz; D. Greene; N. Homeyer; S. Izadi; A. Kovalenko; T.S. Lee; S. LeGrand; P. Li, C. L.; J. Liu, T. L.; R. Luo; D. Mermelstein; K.M. Merz; G. Monard; H. Nguyen; I. Omelyan; A. Onufriev; F. Pan; R. Qi; D.R. Roe; A. Roitberg; C. Sagui; C.L. Simmerling; W.M. Botello-Smith; J. Swails; R.C. Walker; J. Wang; R.M. Wolf; X. Wu; L. Xiao; York, D. M.; and P.A. Kollman, AMBER 2017. *University of California, San Francisco* **(2017)**.

47. 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), 247-260.

48. Jakalian, A.; Jack, D. B.; Bayly, C. I., Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. *J Comput Chem* **2002,** *23* (16), 1623-1641.

49. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A Smooth Particle Mesh Ewald Method. *J Chem Phys* **1995,** *103* (19), 8577-8593.

50. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M., LINCS: A Linear Constraint Solver for Molecular Simulations. *J Comput Chem* **1997,** *18* (12), 1463-1472.

51. Hou, T.; Wang, J.; Li, Y.; Wang, W., Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J Chem Inf Model* **2011,** *51* (1), 69-82.

52. Xu, L.; Sun, H.; Li, Y.; Wang, J.; Hou, T., Assessing the Performance of MM/PBSA and MM/GBSA Methods. 3. The Impact of Force Fields and Ligand Charge Models. *J Phys Chem B* **2013,** *117* (28), 8408-21.

53. Genheden, S.; Ryde, U., The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. *Expert Opin Drug Discov* **2015,** *10* (5), 449-61.

54. Bashford, D.; Case, D. A., Generalized Born Models of Macromolecular Solvation Effects. *Annu Rev Phys Chem* **2000,** *51* (1), 129-52.

55. Weiser, J.; Shenkin, P. S.; Still, W. C., Approximate Atomic Surfaces from Linear Combinations of Pairwise Overlaps (LCPO). *J Comput Chem* **1999,** *20* (2), 217-230.

56. Miller, B. R., 3rd; McGee, T. D., Jr.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E., MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J Chem Theory Comput* **2012,** *8* (9), 3314-3321.

57. Loeffler, H. H.; Michel, J.; Woods, C., FESetup: Automating Setup for Alchemical Free Energy Simulations. *J Chem Inf Model* **2015,** *55* (12), 2485-90.

58. Shirts, M. R.; Chodera, J. D., Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J Chem Phys* **2008,** *129* (12), 124105-10.

59. Klimovich, P. V.; Shirts, M. R.; Mobley, D. L., Guidelines for the Analysis of Free Energy Calculations. *J Comput Aided Mol Des* **2015,** *29* (5), 397-411.

60. Wermuth, C. G., The Practice of Medicinal Chemistry. *Academic Press-Elsevier: Burlington, San Diego, USA, London, U.K.* **2008**.

61. Wassermann, A. M.; Bajorath, J., A data mining method to facilitate SAR transfer. *J. Chem. Inf. Model.* **2011,** *51* (8), 1857-66.

62. Gupta-Ostermann, D.; Wawer, M.; Wassermann, A. M.; Bajorath, J., Graph mining for SAR transfer series. *J. Chem. Inf. Model.* **2012,** *52* (4), 935-42.

63. Blundell, T. L.; Jhoti, H.; Abell, C., High-throughput crystallography for lead discovery in drug design. *Nat. Rev. Drug Discov.* **2002,** *1* (1), 45-54.

64. de Ruyck, J.; Brysbaert, G.; Blossey, R.; Lensink, M. F., Molecular docking as a popular tool in drug design, an in silico travel. *Adv. Appl. Bioinform. Chem.* **2016,** *9*, 1-11.

65. Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J., Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* **2004,** *3* (11), 935-49.

66. Liu, X.; Shi, D.; Zhou, S.; Liu, H.; Liu, H.; Yao, X., Molecular dynamics simulations and novel drug discovery. *Expert Opin Drug Discov* **2018,** *13* (1), 23-37.

67. Godden J.W., B. J., Computational molecular dynamics: Challenges, methods, ideas (book review). *Theor Chem Acc* **2002,** *107*, 250-251.

68. Cournia, Z.; Allen, B.; Sherman, W., Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J. Chem. Inf. Model.* **2017,** *57* (12), 2911-2937.

69. Erlanson, D. A.; McDowell, R. S.; O'Brien, T., Fragment-based drug discovery. *J. Med. Chem.* **2004,** *47* (14), 3463-82.

70. Schuffenhauer, A.; Ertl, P.; Roggo, S.; Wetzel, S.; Koch, M. A.; Waldmann, H., The scaffold tree--visualization of the scaffold universe by hierarchical scaffold classification. *J. Chem. Inf. Model.* **2007,** *47* (1), 47-58.

71. Ertl, P., Intuitive ordering of scaffolds and scaffold similarity searching using scaffold keys. *J. Chem. Inf. Model.* **2014,** *54* (6), 1617-22.

72. Weber, J.; Achenbach, J.; Moser, D.; Proschak, E., VAMMPIRE: a matched molecular pairs database for structure-based drug design and optimization. *J. Med. Chem.* **2013,** *56* (12), 5203-7.

73. Hussain, J.; Rea, C., Computationally efficient algorithm to identify matched molecular pairs (MMPs) in large data sets. *J. Chem. Inf. Model.* **2010,** *50* (3), 339-48.

74. Wassermann, A. M.; Haebel, P.; Weskamp, N.; Bajorath, J., SAR matrices: automated extraction of information-rich SAR tables from large compound data sets. *J. Chem. Inf. Model.* **2012,** *52* (7), 1769-76.

75. Madak, J. T.; Cuthbertson, C. R.; Miyata, Y.; Tamura, S.; Petrunak, E. M.; Stuckey, J. A.; Han, Y.; He, M.; Sun, D.; Showalter, H. D.; Neamati, N., Design, Synthesis, and Biological Evaluation of 4-Quinoline Carboxylic Acids as Inhibitors of Dihydroorotate Dehydrogenase. *J. Med. Chem.* **2018,** *61* (12), 5162-5186.

76. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *Nucleic Acids Res.* **2000,** *28* (1), 235-42.

77. UniProt Consortium, T., UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **2018,** *46* (5), 2699.

78. Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K., BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res.* **2007,** *35* (Database issue), D198-201.

79. Wang, R.; Fang, X.; Lu, Y.; Wang, S., The PDBbind database: collection of binding affinities for protein-ligand complexes with known three-dimensional structures. *J. Med. Chem.* **2004,** *47* (12), 2977-80.

80. Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A. P.; Chambers, J.; Mendez, D.; Mutowo, P.; Atkinson, F.; Bellis, L. J.; Cibrian-Uhalte, E.; Davies, M.; Dedman, N.; Karlsson, A.; Magarinos, M.

P.; Overington, J. P.; Papadatos, G.; Smit, I.; Leach, A. R., The ChEMBL database in 2017. *Nucleic Acids Res.* **2017,** *45* (D1), D945-D954.

81. Chemical Computing Group ULC, S. S. W., Suite #910, Montreal, QC, Canada, H3A 2R7, , Molecular Operating Environment (MOE), 2018.01. 2018.

82. Pippione, A. C.; Giraudo, A.; Bonanni, D.; Carnovale, I. M.; Marini, E.; Cena, C.; Costale, A.; Zonari, D.; Pors, K.; Sadiq, M.; Boschi, D.; Oliaro-Bosso, S.; Lolli, M. L., Hydroxytriazole derivatives as potent and selective aldo-keto reductase 1C3 (AKR1C3) inhibitors discovered by bioisosteric scaffold hopping approach. *Eur J Med Chem* **2017,** *139*, 936-946.

83. Pippione, A. C.; Carnovale, I. M.; Bonanni, D.; Sini, M.; Goyal, P.; Marini, E.; Pors, K.; Adinolfi, S.; Zonari, D.; Festuccia, C.; Wahlgren, W. Y.; Friemann, R.; Bagnati, R.; Boschi, D.; Oliaro-Bosso, S.; Lolli, M. L., Potent and selective aldo-keto reductase 1C3 (AKR1C3) inhibitors based on the benzoisoxazole moiety: application of a bioisosteric scaffold hopping approach to flufenamic acid. *Eur J Med Chem* **2018,** *150*, 930-945.