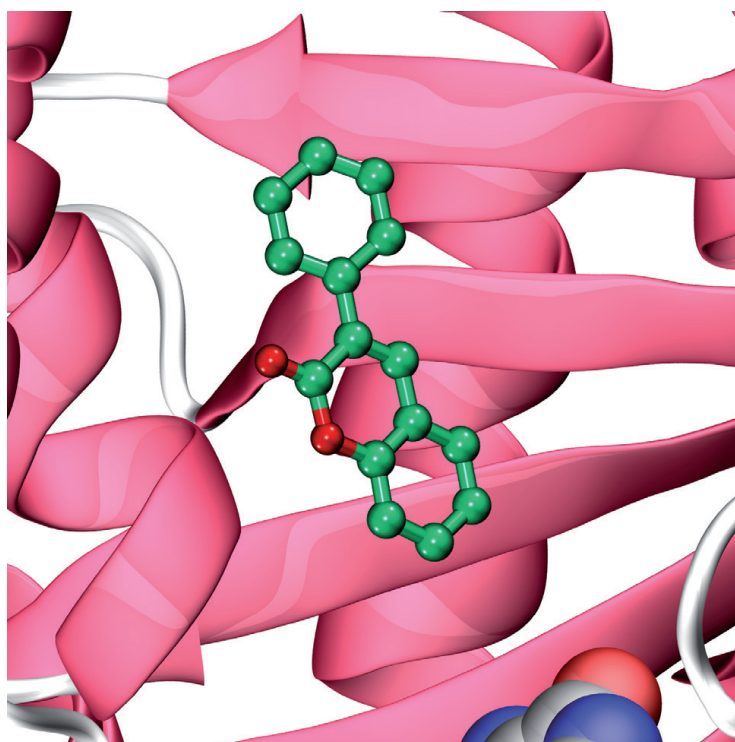Sanna Rauhamäki

Small Molecule Modulators of Amine Oxidation, Nuclear Receptor Signaling and Glucuronidation

3-Phenylcoumarin as a Scaffold of Interest

Sanna Rauhamäki

Small Molecule Modulators of Amine Oxidation, Nuclear Receptor Signaling and Glucuronidation

3-Phenylcoumarin as a Scaffold of Interest

Esitetään Jyväskylän yliopiston matemaattis-luonnontieteellisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston Ambiotica-rakennuksen salissa YAA303
huhtikuun 20. päivänä 2018 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Mathematics and Science of the University of Jyväskylä,
in building Ambiotica, hall YAA303, on April 20, 2018 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2018

Small Molecule Modulators of Amine Oxidation, Nuclear Receptor Signaling and Glucuronidation

3-Phenylcoumarin as a Scaffold of Interest

Sanna Rauhamäki

Small Molecule Modulators of Amine Oxidation, Nuclear Receptor Signaling and Glucuronidation

3-Phenylcoumarin as a Scaffold of Interest

UNIVERSITY OF JYVÄSKYLÄ

Cover picture by Sanna Rauhamäki.

"We need to concentrate on
more than meets the eye"

Twenty Years by Placebo

ABSTRACT

The costs of the drug development process are moderated as computer-aided drug design methods are able to expedite the steps required for lead identification. In fact, computational tools are nowadays virtually indispensable from target identification and validation to preclinical tests due to exponential growth of available information regarding both potential targets and small molecules. One such small molecule with growing number of variations is coumarin. Coumarin scaffold and its various derivatives continue to interest researchers for their vast application potential. Since naturally occurring coumarins are known for example for their antioxidant and anti-inflammatory properties, those molecules are used to guide research endeavors toward similar molecules but with enhanced or newly directed activities. In this doctoral thesis, coumarin derivatives are used to gain novel details regarding monoamine oxidase and nuclear receptor modulation in context relevant for example in neurological conditions and cancer. In order to achieve this, diverse collection of coumarin derivatives is investigated in these targets and corresponding antitargets using both computational and experimental methods. As a result, novel coumarin derivatives with activity in nanomolar range are identified in case of monoamine oxidase B and estrogen receptor α and comparable activity is reached for retinoid-acid-receptor-related orphan receptor γt with novel core. In addition, the usability of coumarin derivatives as assay development tools is put to test by designing selective ligands for glucuronidation. Consequently, the metabolic fate of the coumarins is investigated as they are allocated to metabolizing target using homology models, computational methods and experimental techniques. Subsequently, two coumarin derivatives selective for human uridine 5'-diphospho-glucuronosyltransferase 1A10 are identified.

Keywords: 3-phenylcoumarin, 17β-hydroxysteroid dehydrogenase, cancer, computer-aided drug design, estrogen receptor, monoamine oxidase.

Sanna Rauhamäki, University of Jyväskylä, Department of Biological and Environmental Science, P.O. Box 35, FI-40014 University of Jyväskylä, Finland

Author's address   Sanna Rauhamäki
Department of Biological and Environmental Science
P.O. Box 35
FI-40014 University of Jyväskylä
Finland
sanna.rauhamaki@jyu.fi

Supervisors   Professor Olli Pentikäinen, Ph.D.
Department of Biological and Environmental Science
P.O. Box 35
FI-40014 University of Jyväskylä
Finland

Docent Ulla Pentikäinen, Ph.D.
Department of Biological and Environmental Science
P.O. Box 35
FI-40014 University of Jyväskylä
Finland

Docent Pekka Postila, Ph.D.
Department of Biological and Environmental Science
P.O. Box 35
FI-40014 University of Jyväskylä
Finland

Reviewers   Docent Henri Xhaard, Ph.D.
Division of Pharmaceutical Chemistry and Technology
P.O. Box 56
FI-00014 University of Helsinki
Finland

Docent Lari Lehtiö, Ph.D.
Biocenter Oulu and Faculty of Biochemistry and Molecular
Medicine
P.O. Box 3000
FI-90014 University of Oulu
Finland

Opponent   Docent Maija Lahtela-Kakkonen, Ph.D.
School of Pharmacy
P.O. Box 1627
FI-70211 University of Eastern Finland
Finland

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

The thesis is based on the following original papers, which will be referred to in the text by their Roman numerals I–V.

I   Rauhamäki S., Postila P.A., Niinivehmas S., Kortet S., Schildt E., Pasanen M., Manivannan E., Ahinko M., Koskimies P., Nyberg N., Huuskonen P., Multamäki E., Pasanen M., Juvonen R.O., Raunio H., Huuskonen J. & Pentikäinen O.T. 2018. Structure-Activity Relationship Analysis of 3-phenylcoumarin-Based Monoamine Oxidase B Inhibitors. *Frontiers in Chemistry* 6: 41.

II   Niinivehmas S., Manivannan E., Rauhamäki S., Huuskonen J. & Pentikäinen O.T. 2016. Identification of estrogen receptor α ligands with virtual screening techniques. *Journal of Molecular Graphics and Modeling* 64: 30–39.

III Niinivehmas S., Postila P.A., Rauhamäki S., Manivannan E., Kortet S., Ahinko M., Huuskonen P., Nyberg N., Koskimies P., Lätti S., Multamäki E., Juvonen R.O., Raunio H., Pasanen M., Huuskonen J. & Pentikäinen O.T. 2018. Blocking oestradiol synthesis pathways with potent and selective coumarin derivatives. *Journal of Enzyme Inhibition and Medicinal Chemistry.* In press.

IV Rauhamäki S., Postila P.A., Lätti S., Niinivehmas S., Multamäki E., Liedl K. & Pentikäinen O.T. 2018. Discovery of Retinoic Acid-Related Orphan Receptor γt Inverse Agonists via Docking and Negative Image-Based Screening. Submitted manuscript.

V   Juvonen R.O., Rauhamäki S., Kortet S., Niinivehmas S., Troberg J., Petsalo A., Huuskonen J., Raunio H., Finel M. & Pentikäinen O.T. 2018. Molecular Docking-Based Design and Development of a Highly Selective Probe Substrate for UDP-glucuronosyltransferase 1A10. *Molecular Pharmaceutics* 15: 3: 923–933.

RESPONSIBILITIES OF SANNA RAUHAMÄKI IN THE ORIGINAL PUBLICATIONS OF THE THESIS

The thesis is based on the following original papers, which will be referred to in the text by their Roman numerals I–V.

I   I was responsible for the experimental testing regarding MAO A and B. Elina Multamäki did the MAO A experimental analysis under my advisement. In addition, I performed the docking into MAO B, prepared most of the figures and wrote the manuscript with other coauthors. Pekka Postila was responsible for the SAR analysis. Sami Kortet, Emmi Schildt and Juhani Huuskonen performed the organic synthesis. Mira Ahinko did the PAINS screening. Pasi Koskimies performed the HSD1 measurements. Nina Nyberg, Risto Juvonen and Hannu Raunio did the experimental analysis regarding CYP1A2. Pasi Huuskonen and Markku Pasanen executed the experimental analysis regarding aromatase. Mira Pasanen did preliminary screening for designing MAO B ligands. Sanna Niinivehmas, Elangovan Manivannan, Sami Kortet and Olli Pentikäinen designed the molecules for the selected targets.

II   I was responsible for the fluorescence polarization measurements. Sanna Niinivehmas performed the negative image-based studies and the docking with PLANTS as well as prepared most of the figures. Elangovan Manivannan performed rest of the computational studies. Juhani Huuskonen synthesized the coumarin derivatives. The work was written together with all the co-authors where I was responsible for the sections regarding experimental measurements. Olli Pentikäinen supervised the study. Submitted manuscript of this publication has been previously used in the thesis of Sanna Niinivehmas.

III I was responsible for the MAO A and B experimental analysis and Elina Multamäki executed the MAO A analysis under my advisement. Mira Ahinko did docking-based SAR analysis for CYP1A2. Sami Kortet and Juhani Huuskonen performed the organic synthesis of the compounds. Pasi Huuskonen and Markku Pasanen executed the experimental analysis regarding aromatase. Pasi Koskimies performed the HSD1 measurements. Sakari Lätti performed data analysis. Nina Nyberg, Risto Juvonen and Hannu Raunio did the experimental analysis regarding CYP1A2. Pekka Postila was in charge of the docking-based SAR analysis for HSD1, MAO A, MAO B and aromatase. Pekka Postila also wrote the article with the help of the other authors. Sanna Niinivehmas, Elangovan Manivannan, Sami Kortet and Olli Pentikäinen designed the molecules for the selected targets.

IV I was responsible for the design and execution of the *in vitro* measurements. In addition, I performed data analysis together with Pekka Postila and Olli Pentikäinen. I made most of the figures and participated in the writing of the manuscript with the other coauthors. Elina Multamäki, Sanna Niinivehmas and Olli Pentikäinen performed the virtual screening. Klaus Liedl and Olli Pentikäinen designed the research study.

V   I performed the homology modeling, prepared the figures of them and wrote the relevant sections. Sami Kortet and Juhani Huuskonen did the organic synthesis. Johanna Troberg and Moshe Finel were in control of the protein production. Risto Juvonen and Hannu Raunio were responsible for the experimental analysis. Hannu Raunio, Moshe Finel and Olli Pentikäinen supervised the project.

ABBREVIATIONS

1D	one dimensional
2D	two dimensional
3D	three dimensional
4D	four dimensional
5D	five dimensional
6D	six dimensional
ADMET	absorption, distribution, metabolism, excretion, toxicity
AUC	area under curve
CADD	computer-aided drug design
cryo-EM	cryo-electron microscopy
CYP	cytochrome P450
DUD	Directory of Useful Decoys
ER	estrogen receptor
FAD	flavin adenine dinucleotide
HSD	17β-hydroxysteroid dehydrogenase
MAO	monoamine oxidase
NIB	negative image-based
NMR	nuclear magnetic resonance
PAINS	pan assay interference compounds
PDB	Protein Data Bank
QSAR	quantitative structure-activity relationship
ROC	receiver operator characteristics
ROR	retinoid-acid-receptor-related orphan receptor
SAR	structure-activity relationship
Th17	T helper 17 cells
UGT	uridine 5'-diphospho-glucuronosyltransferase
vdW	van der Waals

1 INTRODUCTION

Proteins are versatile biological macromolecules with a wide range of functions. They maintain the shape of a cell or aid in their movement as structural proteins, carry atoms and small molecules as transport proteins, convey signals for coordinating biological processes as messenger proteins or hormones, catalyze chemical reactions as enzymes and defend the host as antibodies just to mention a few. In order to achieve all this, proteins can assume various shapes and sizes from tiny peptides to molecular machines consisting of multiple domains. More remarkably, regardless of the function, only 20 building blocks, the standard amino acids are typically enough to construct every protein.

The actions of proteins are often guided by small molecules known as ligands which form complexes with the proteins. Depending on the ligand, the consequences of such complex formation may vary greatly. Ligands called agonists are capable of activating their target receptor which induces a biological response. If the protein participates in a chemical reaction involving the bound ligand, the ligand is then known as substrate. Some of these chemical reactions also require additional components called cofactors. However, the biological consequences of ligand binding may also be reversed from the basal level and then the ligand is called inverse agonist. If the function of a protein is inactivated or reduced, the ligand responsible is determined to be antagonist or inhibitor. By forming intricate networks and altering the consequences of protein-ligand interactions, the biological processes are managed on a cellular level.

In drug design the capability of proteins to bind ligands is utilized as pharmaceutical benefits are sought by eliciting desired functions in medically relevant proteins. The ability of a ligand to bind to a target is defined as affinity which is often described as the activity of a ligand. When designing a ligand for pharmaceutical purposes, the goal is to find one exhibiting high affinity for the target in question. On the other hand, promising drug candidates should only display activity against the intended target. In other words, creating selective ligands is an important goal. In order to meet this requirement, detailed interactions which are specific for the protein-ligand complex under

examination, are pursued while also taking into account absorption, distribution, metabolism, excretion and toxicological properties affecting the performance of a drug candidate.

Currently, overwhelming amount of druggable targets are known and new ones are constantly discovered. It has been estimated that the druggable genome contains close to 4500 genes. However, the approved drugs target only approximately 15 % of those genes at present. Thus, finding new pharmaceuticals in cost-effective manner is of the essence. Computer-aided drug design methods aim precisely at abridging the time and money spend between the identification of potential drug candidates and the clinical trials. This doctoral thesis investigates the potential of multiple pharmaceutically significant targets to bind a selection of small molecules that could be modified further. The purpose is to identify the factors determining the selectivity in each individual case. In addition, the metabolism of pharmaceuticals is taken into account as novel tool molecules are designed for improved metabolism analysis.

2 REVIEW OF LITERATURE

2.1 Monoamine oxidases

Monoamine oxidase (EC 1.4.3.4., MAO), discovered in 1928 (Hare 1928), is flavin-containing enzyme which catalyzes the oxidative deamination of neurotransmitters and biogenic amines. The enzyme mainly acts on primary amines but it is also capable of oxidizing some secondary and tertiary amines following the overall reaction:

$$R^1CH_2NR^2R^3 + O_2 + H_2O \rightarrow R^1CHO + NR^2R^3 + H_2O_2$$
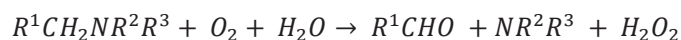
According to the equation, molecular oxygen is used to remove an amine and as a result, a corresponding aldehyde, ammonia (or amine) and hydrogen peroxide are produced. During the oxidation reaction, flavin adenine dinucleotide (FAD) acts as a cofactor and is reduced in the first half-reaction as it accepts a hydride from the substrate. In the second half-reaction the reduced flavin returns to its oxidized state as it is oxidized by molecular oxygen. The exact mechanism for the electron transfer is still debated on and several suggestions that might be substrate specific have been made. As the FAD is capable of accepting more than one electron, the reaction could occur through single electron transfer or through direct hydrogen atom transfer. Alternatively, the reaction might be a variation of a nucleophilic mechanism or occur via direct hydride transfer from the substrate to the flavin (Tipton *et al.* 2004, Gaweska and Fitzpatrick 2011).

Monoamine oxidase has two types in humans, MAO A and MAO B. The presence of two types was originally discovered due to the observation that only MAO A displays sensitivity towards irreversible inhibitor clorgyline (Johnston 1968). MAO A (527 aa, 59.7 kDa) and MAO B (520 aa, 58.8 kDa) share 70 % amino acid sequence identity with the identical five amino acids long peptide SGGCY covalently binding the cofactor FAD. The two types are derived from separate genes which facilitates tissue specific gene regulation and expression (Bach *et al.* 1988). The genes are located tail to tail on X

chromosome and are made of 15 exons with identical exon-intron spacing (Shih *et al.* 1999, Shih and Chen 2004).

Both MAO A and MAO B (Fig. 1A) are located on mitochondrial outer membrane (Schnaitman *et al.* 1967) where the protein is anchored by C-terminal transmembrane region (Binda *et al.* 2002). This helical region is nearly 30 amino acids long and contains the mitochondrial localization signal (Mitoma and Ito 1992). In addition, the protein surfaces have exposed hydrophobic and positively charged amino acids which are thought to be embedded into the membrane or to interact with the phospholipid head group (Binda *et al.* 2002, Son *et al.* 2008). Interestingly, human MAO A crystallizes as a monomer (De Colibus *et al.* 2005) whereas human MAO B obtains dimer conformation during crystallization (Binda *et al.* 2002). The dimer formation of MAO B has been attributed to extensive monomer-monomer interactions encompassing 15 % of the MAO B monomer accessible surface (Binda *et al.* 2002). Although the sequence homology between human MAO A and MAO B is considerable, human MAO A has a unique E151K change (Andrés *et al.* 2004) located next to residues involved in monomer-monomer contacts of the MAO B dimer. The amino acid change destabilizes the dimeric state which is why MAO A crystallizes in monomeric form (De Colibus *et al.* 2005). Consequently, the quaternary assembly observed in the X-ray crystal structures of MAO A and MAO B was at first thought to also comply *in vivo* but it was later shown that both of the proteins obtain the dimeric form on the outer mitochondrial membrane (Upadhyay *et al.* 2008).

The FAD binding domain and the active site region surrounded by two tyrosines right next to the cofactor are conserved in both MAO A and MAO B. It is the only hydrophilic section in the substrate binding cavity and it has vital role in the recognition and guidance of the substrate deamination activity. This is consistent with the concept of both enzymes following the same catalytic mechanism (De Colibus *et al.* 2005, Edmondson *et al.* 2005). As expected, the major differences (Fig. 1C) are located opposite to the bound flavin where the recognition of the substrate takes place inside the hydrophobic cavity.

The structure of human MAO B was obtained first of the two types by Binda *et al.* (2002). It was a 3 Å resolution structure of the enzyme together with inhibitor pargyline which is an analog of clinically used deprenyl. According to the X-ray crystal structure, the substrate binding cavity is divided into two distinctive parts. The substrate is thought to first enter an entrance cavity beneath loop 99–112 (Fig. 1A) as an admission mechanism and then continue towards substrate cavity containing the active site. The loop 99–112 is located close to the membrane binding region, indicating that the substrate entrance takes place from the protein surface facing the mitochondrial membrane. As such, the loop 99–112 might act as initial gating mechanism and the residues F168, L171, I199 and Y326 aligning the section separating the two cavities must then momentarily give in to allow the migration of the substrate into the active site (Binda *et al.* 2002).
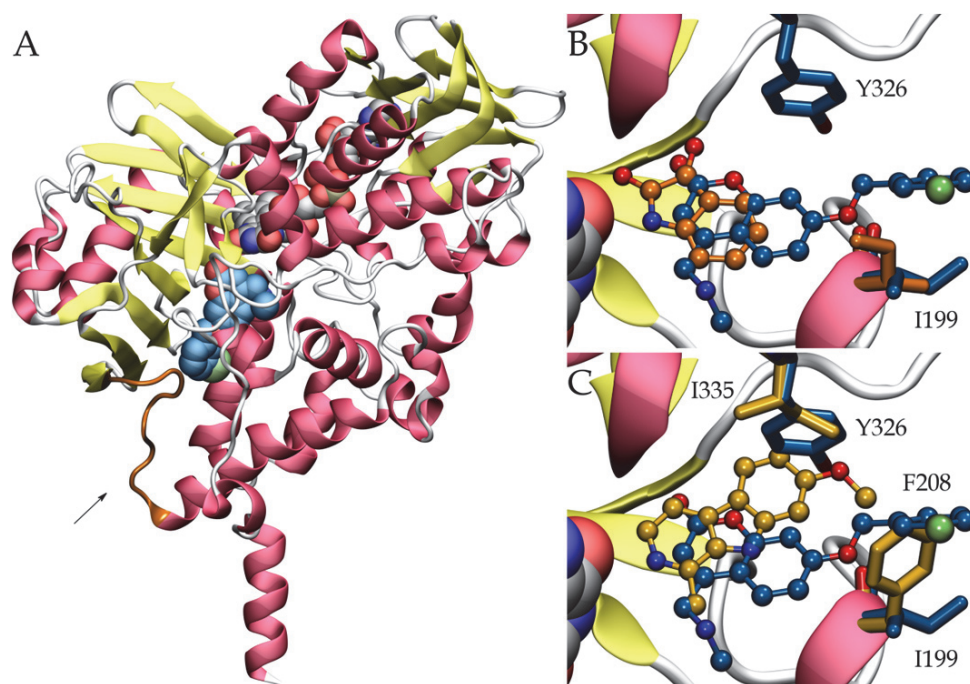
FIGURE 1   A. MAO B (α-helixes in pink, β-sheets in yellow) in complex with 7-(3-chlorobenzyloxy)-4-(methylamino)methyl-coumarin and FAD (as CPK space-filling models with blue or grey carbon atoms) (PDB 2V61 (Binda *et al.* 2007)). On the left, the arrow is pointing at loop 99–112 (in orange) that covers the entrance to the substrate binding cavity and is thus thought to function as an admission mechanism. The loop is located close to the mitochondrial membrane as the α-helix pointing down contains the C-terminal transmembrane region as well as mitochondrial localization signal. B. MAO B is capable of accommodating substrates of varying sizes due to alternative conformation of I199 side chain. While accommodating isatin, I199 acquires a conformation facing the cavity (in orange, PDB 1OJA (Binda *et al.* 2003)) whereas outward facing conformation of I119 enables the binding of larger molecules such as noncovalent inhibitor 7-(3-chlorobenzyloxy)-4-(methylamino)methyl-coumarin (in blue, PDB 2V61 (Binda *et al.* 2007)). C. Comparison of superimposed MAO A and MAO B X-ray crystal structures reveals key features contributing to the substrate selectivity of the enzymes. Based on the superimposition of the structures, harmine in MAO A (in gold, PDB 2Z5X (Son *et al.* 2008)) overlaps with Y326 of MAO B and in contrast, 7-(3-chlorobenzyloxy)-4-(methylamino)methyl-coumarin in MAO B (in blue, PDB 2V61 (Binda *et al.* 2007)) overlaps with F208 of MAO A.

More details of the MAO B binding cavity structure were obtained as a 1.7 Å high-resolution structure of the enzyme complexed with reversible inhibitor isatin was resolved (Binda *et al.* 2003). This structure was compared to a structure resolved with 1,4-diphenyl-2-butene which revealed that I199 side chain is capable of obtaining an alternative conformation depending on the bound substrate. Since 1,4-diphenyl-2-butene is a larger molecule than pargyline or isatin, it occupies both of the previously described cavities. Here I199 is in a key role as its alternative conformation allows the binding of more sizable substrate. Thus, the adjustable nature of the cavity has important role in

the deamination activity directed at varying amines and might make it worthwhile to design more sizable inhibitor candidates (Binda *et al.* 2003, Edmondson *et al.* 2005). The theory was put to test with noncovalent inhibitors safinamine and two coumarin analogs, all of which occupied both of the cavities and obtained the alternative I199 conformation (Fig. 1B) (Binda *et al.* 2007).

Shortly after human MAO B, the 3D structure of human MAO A was also determined (De Colibus *et al.* 2005). The protein was initially crystallized together with its inhibitor clorgyline in the substrate binding cavity. As pargyline in MAO B, clorgyline binds covalently to MAO A flavin N5 (Binda *et al.* 2002, De Colibus *et al.* 2005). In MAO A the substrate binding cavity reaches continuously from the cofactor to loop 210–216 shaping the cavity on the protein surface. On the inside, the cavity is lined with aliphatic and aromatic residues and as with MAO B, is rather hydrophobic (De Colibus *et al.* 2005). Structure with reversible MAO A inhibitor harmine further endorsed the previous observations of the conserved nature of the deamination activity of the two types. The aromatic rings of the reversible inhibitors overlap in the superimposed X-ray crystal structures as they form interactions with the close by amino acids. Comparison of the substrate binding cavities also revealed that I335 in MAO A and Y326 in MAO B are in important role in the substrate selectivity. In the superimposed X-ray crystal structures, harmine overlaps with Y326 in MAO B which could explain its selectivity towards MAO A. Similarly, 1,4-diphenyl-2-butene or 7-(3-chlorobenzyloxy)-4-(methylamino)methyl-coumarin overlaps with F208 in MAO A (Fig. 1C). Therefore the size and shape of the substrate cavity as well as the amino acids I335 and F208 in MAO A, corresponding with Y326 and I199 in MAO B take part in the substrate selectivity of these enzymes (Binda *et al.* 2007, Son *et al.* 2008).

2.1.1 Clinical significance and pharmacology of MAOs

As could be expected from two widely similar enzymes, MAO A and MAO B have distinctive but overlapping substrate specificities. Neurotransmitters containing amino group such as serotonin, norepinephrine and epinephrine are preferentially oxidized by MAO A whereas benzylamine and phenylethylamine are oxidized by MAO B. However, they both share the activity toward dopamine and tyramine (Houslay and Tipton 1974). In addition to substrate preference, human MAO A and MAO B have slight differences in their anatomic location and cellular expression. In general, MAO A and MAO B follow the same distribution pattern in human tissue but certain variations have been observed. In heart, liver, duodenum, blood vessels, lung and spleen the two types can be found in the same cell types but variation in expression levels do exist. For instance, expression of MAO A is particularly strong in duodenum mucosa both in villi and crypts whereas MAO B is strongly present only in the villi. In kidney, pancreas as well as in adrenal and thyroid gland the expression of these enzymes exhibits more variation. Kidney podocytes, Bowman's capsule and collecting ducts, pancreas acinar cells and β cells, adrenal gland capsule

and interstitial cells as well as the cells in thyroid gland are devoid of MAO B but do express MAO A. The opposite has been observed with pancreas α cells where MAO B is present without MAO A (Saura *et al.* 1996b, Rodríguez *et al.* 2000, 2001). Great deal of the MAO expression data has been obtained from tissue samples originating from deceased bodies of the elderly. Ageing alters the MAO expression as it is known that the expression levels vary due to developmental changes. For example in brain, MAO B is expressed more than MAO A from adolescence to senescence but the reverse is true in neonatal brain. In fact, adult ageing is connected to further increase in MAO B but not MAO A levels in brain. The protein levels also vary between different tissue types, MAO B being especially well presented in white matter of temporal cortex and cerebellum (Saura *et al.* 1996a, Tong *et al.* 2013). Since MAOs are responsible for deamination activity directed at neurotransmitters, it is understandable if an imbalance in the function of these enzymes can lead into ailment.

The most concerning side effect related to the use of MAO inhibitors is known as cheese effect. Cheese and other fermented, smoked, pickled or otherwise aged food items contain considerable amounts of tyramine as a result of decarboxylation of tyrosine. Dietary tyramine is metabolized by MAOs present in the gastrointestinal tract where MAO A is the more prevalent type (Hasan *et al.* 1988). Thus, medication inhibiting MAO A, in particular, can cause accumulation of dietary tyramine that reaches the peripheral tissues. There tyramine acts as sympathomimetic amine entering sympathetic neurons and causing the release of noradrenaline. The increase in noradrenaline release mediates vasoconstriction and transient hypertension which can lead into hypertensive crisis and death due to intracranial hemorrhage (Broadley 2010). In order to avoid these detrimental side effects, prescription of MAO A inhibitors requires that the patient complies with detailed dietary restrictions. Such a limiting treatment option might not always be ideal or even feasible, which is why alternatives have been sought from MAO B selective inhibitors.

MAOs as drug target were first discovered in the early 1950s when patients treated with antitubercular agent iproniazid reported psycho energizing effects. The effect was soon pinpointed to MAOs and the MAO inhibitors became the first effective antidepressants. Although irreversible MAO inhibitor iproniazid was eventually removed from market due to hepatotoxicity, other more effective antidepressants irreversibly targeting MAOs such as phenelzine, isocarboxazid and tranylcypromine soon followed (Shulman *et al.* 2013). These early irreversible MAO inhibitors permanently inactivate the enzyme until it is replaced – a process that might not always be desirable. Consequently, reversible MAO inhibitors such as moclobemide, that restore the enzyme activity as soon as the inhibitor dissociates, were created to circumvent this issue. In addition to being reversible or irreversible, MAO inhibitors are categorized by their selectivity. The inhibitors might exhibit selectivity towards MAO A (e.g. moclobemide) or MAO B (e.g. pargyline, selegiline) or be non-selective (e.g. phenelzine, tranylcypromine). The selectivity is often concentration dependent however, and the selectivity might be lost in

high concentrations or doses (Tipton *et al.* 2004, Shulman *et al.* 2013, Finberg and Rabey 2016).

In addition to depression, MAOs have been connected to various neurological disorders and diseases with a wide range of pharmaceutical applications. These diseases include neurodegenerative disorders such as Alzheimer's disease and Parkinson's disease but also mental disorders like schizophrenia (Wayment *et al.* 2001), and generalized anxiety disorder (Tadic *et al.* 2003). In addition, substance abuse (Vanyukov *et al.* 2004), migraines (Filic *et al.* 2005) and diabetes mellitus (Adeghate and Parvez 2004) have all been linked to MAOs (Youdim *et al.* 2006). As a result of gliosis associated with Parkinson's disease, MAO B concentration increases in the brain tissue of the affected people. Because of that, selective MAO B inhibitors have been investigated as potential treatment in order to alleviate the disease symptoms and to slow down the disease progression. MAO B activity is also elevated in the brains of the Alzheimer's disease patients and is thus considered to contribute to the oxidative stress related to this condition. Combining MAO B inhibitors with existing treatments might therefore be beneficial in enhancing cognition in Alzheimer's disease patients (Youdim *et al.* 2006). The levels of monoamines serotonin, norepinephrine and dopamine are lowered in depression because of elevated MAO A expression and, as a result, inhibiting MAO B does not have obvious antidepressant effect in treatment of major depression. Despite of this, selective MAO B inhibitors have been examined as antidepressants with positive outcome. The results have been especially encouraging when the MAO B inhibitor selegiline has been administered transdermally or as a buccally administered solution that permits the use of a lower dose because greater part enters the central nervous system. This way the administered dose avoids the first-pass metabolism and accesses the brain instead of gastrointestinal tract, circumventing also side effects related to the formation of harmful amphetamines as metabolites and the cheese effect caused by MAO A inhibition (Meyer *et al.* 2006, Finberg and Rabey 2016).

Due to neuroprotective capabilities, MAO B is an interesting target for therapeutic applications. Although existing MAO B inhibitors are already in use, the search continues for improved activity and specificity. Since the two MAO types share substrate specificity, creating selective molecules is a challenging but necessary task in order to avoid potentially lethal side effects related to MAO A inhibition.

2.2 Nuclear receptors

By definition, nuclear receptors are transcription factors affecting embryonic development, maintenance of differentiated cell types, metabolism and cell death. Due to their essential nature, dysfunction in the regulation of nuclear receptors has proliferative, reproductive and metabolic consequences such as cancer, infertility and diabetes. Therefore, targeting nuclear receptors with

pharmaceutical agonists or antagonists are among the most used treatment options in medicine (Gronemeyer *et al.* 2004).

In general, nuclear receptors are ligand regulated transcription factors with specific target gene transcription to modulate. They share common mode of action and consist of the same three major domains (Wrange and Gustafsson 1978). The N-terminus is known as transactivation domain and it is the region recognized by coactivators and other transcription factors. In the center is the location of the DNA binding domain that consists of two zinc finger motifs in all but two of the nuclear receptor family members. Lastly, the C-terminus is the location of the ligand binding domain which is fairly conserved as the various family members share common 3D structure (Wurtz *et al.* 1996, Gronemeyer *et al.* 2004). Despite of the common structural features, the existing differences are enough to warrant selective ligand binding. In addition, ligand induced activator function, which has fundamental role in transcriptional coregulator interaction, is included in the ligand binding domain (Gronemeyer *et al.* 2004).

The 3D structures of nuclear receptors with or without a ligand have been an invaluable aid in understanding the underlying mechanisms of the ligand binding and transactivation of this receptor family (Wurtz *et al.* 1996). The ligand binding domain is formed by 11 compactly packed helices. Although the enclosed ligand binding site varies significantly in size among the family members, it is guarded by the 12th helix forming a flexible lid over the cavity (Fig. 2). The orientation of the 12th helix and the related allosteric effects are directly dependent on the chemical composition of the bound ligand. As the residues essential for the ligand induced activator function are also located in this helix, the alternative positioning of the 12th helix has complex regulatory consequences. In other words, the conformational change on the ligand binding domain caused by a bound agonist is actually the reason for nuclear receptor transactivation function. On the contrary, binding of an antagonist would not be able to induce these effects (Wurtz *et al.* 1996, Gronemeyer *et al.* 2004).

2.2.1 Estrogen receptor α and the estradiol synthesis pathway

As is the case with many nuclear receptors, estrogen receptors originate from paralogous genes that were duplicated during the vertebrate evolution. The two receptors, estrogen receptor (ER) α (Green *et al.* 1986) and β (Mosselman *et al.* 1996), are located on different chromosomes, ERα on chromosome 6 and ERβ on chromosome 14 but both of their genes consist of 8 coding exons. Despite of the common evolutionary background and nearly identical DNA binding domains, ERs share only 50 % sequence identity on the ligand binding domain and they differ even more at the N-terminal transactivation domain that has only 30 % sequence identity between the receptors. The differences in sequences are also evident based on the overall sequence length since ERα (595 aa) is over 60 amino acids longer than ERβ (530 aa) (Gronemeyer *et al.* 2004, Lee *et al.* 2012). Moreover, each ER has distinct expression patterns and tissue distributions which are modulated by selective ligands (Zhao *et al.* 2008). However, the

functionalities of ERα and ERβ are interconnected as ERβ often counteracts the effects of ERα and together the two keep the estrogen signaling in balance.
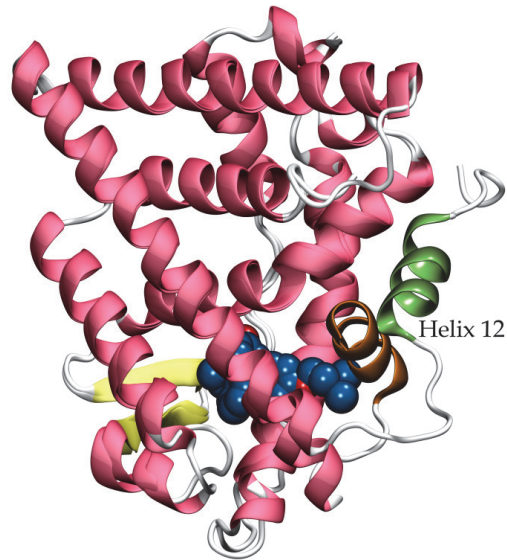
FIGURE 2   The binding of an agonist and an antagonist has a distinct impact on the conformation of the 12th helix of the nuclear receptors. By superimposing two ERα X-ray crystal structures (α-helixes in pink, β-sheets in yellow), the conformational change involving helix 12 becomes apparent. Compared to ERα in complex with the endogenous agonist 17β-estradiol (not shown, PDB 1ERE (Brzozowski *et al.* 1997)) where the helix 12 (in orange) encases the ligand binding cavity, the ligand binding cavity of ERα in complex with an antagonist 4-hydroxytamoxifen (as CPK space-filling model with blue carbon atoms, PDB 3ERT (Shiau *et al.* 1998)) is more exposed as the helix 12 (in green) moves away and obtains an alternative orientation.

Both ERs are activated by 17β-estradiol which is produced either through aromatase or sulfatase pathway (Fig. 3). On aromatase pathway, androgens are converted into 17β-estradiol. Human aromatase also known as CYP19 is present on the final step of the 17β-estradiol synthesis where it transforms androstenedione into estrone and testosterone into 17β-estradiol (Cui *et al.* 2013). In comparison to the aromatase activity, on the sulfatase pathway sulfated estrogens are the precursors of 17β-estradiol production. On the sulfatase pathway, the sulfate group of estrone-3-sulfate is removed by steroid sulfatase producing estrone. Next, 17β-hydroxysteroid dehydrogenase (HSD) 1 and its catalytic counterpart HSD2 participate into the 17β-estradiol biosynthesis. They function on the opposite sides of the 17β-estradiol biosynthesis as HSD1 catalyzes the final reducing step, HSD2 converts 17β-estradiol back into estrone (Simpson 2003, Gronemeyer *et al.* 2004, Vihko *et al.* 2004, Secky *et al.* 2013). Thus, affecting the production of 17β-estradiol provides several methods for pharmaceutical intervention in related ailments.

2.2.2 Clinical significance of the estradiol synthesis pathways

Ovaries are the main source of 17β-estradiol in premenopausal women. Other sources such as mesenchymal cells of the adipose tissue, skin osteoblasts, vascular endothelial cells and many sites in the brain become the primary source postmenopause. In general, the 17β-estradiol synthesis on these other sites is thought to fulfill the 17β-estradiol requirements locally and their contribution to circulating 17β-estradiol is minimal. In males, the main source of circulating 17β-estradiol is the testes and their significance in 17β-estradiol production is considered to remain constant throughout adult life (Simpson *et al.* 2000).

Since ER is activated by 17β-estradiol, it is the direct or indirect target of many treatments related to 17β-estradiol production. Hormone replacement therapies, contraceptives and infertility treatments all target ER (Jordan 2003a, b). However, the most devastating condition related to estradiol synthesis pathway is the endocrine related breast cancer. Selective ER modulators tamoxifen and raloxifen are the best known treatment options in hormone responsive breast cancer. On the other hand, tamoxifen functions as ERα agonists also outside the breast cancer tissue which is why it has been linked to endometrial hyperplasia and cancer. Interestingly, raloxifen has the capability to function as agonist in one tissue and as an antagonist in other. Thus, raloxifen can be used as breast cancer or osteoporosis treatment without endometrial cancer (Simpson *et al.* 2000, Shang and Brown 2002). The success of these ER-targeting breast cancer drugs has generated more research in the field and new alternatives with increased activity and fewer side-effects are constantly developed.

ERα is the main target in ERα-positive breast cancer (Ali and Coombes 2000). Occasionally the cancer may grow resistant of the ERα targeted treatments or it does not exhibit sensitivity to said treatments in the first place. Because of that, alternative methods to affect 17β-estradiol production are sought after in order to treat these types of breast cancer. Obvious alternatives can be found from the estradiol synthesis pathways where aromatase (Simpson and Davis 2001), steroid sulfatase (Woo *et al.* 1998) or HSD1 (Vihko *et al.* 2004, Hanamura *et al.* 2014) have stood out as potential targets. As aromatase and HSD1 expression is increased in ectopic endometrium, they are also potential targets for treating endometriosis (Vihko *et al.* 2004, Dassen *et al.* 2007).
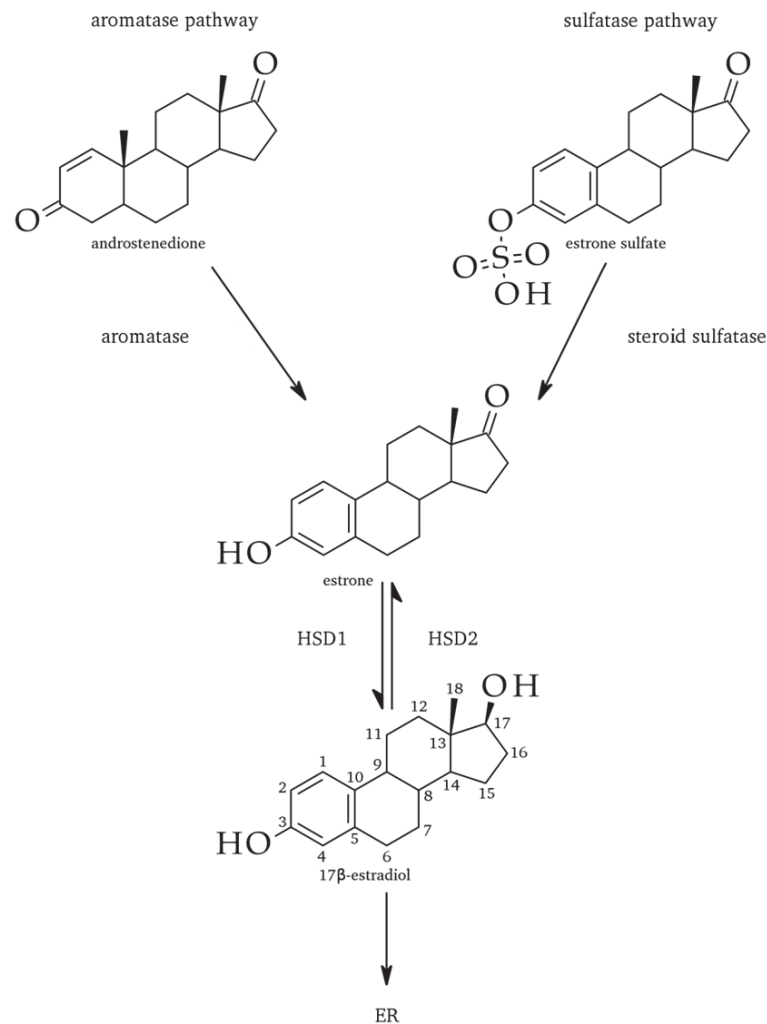
FIGURE 3   Simplified representation of the aromatase and sulfatase pathways of the 17β-estradiol synthesis. Aromatase catalyzes the transformation of androstenedione into estrone (or testosterone into 17β-estradiol, not shown) whereas steroid sulfatase catalyzes the conversion of estrone sulfate into estrone. HSD1 and HSD2 function on the opposite sides of the same reaction equation when HSD1 catalyzes the reduction of estrone into 17β-estradiol and HSD2 converts 17β-estradiol back into estrone. The synthesized 17β-estradiol then activates ER. The 17β-estradiol is numbered according to IUPAC nomenclature of organic chemistry. When referring to the positions of the compound, this numbering is followed and it is additionally applicable to the other steroids presented above where the numbering is omitted for clarity.

2.2.3 Retinoid-acid-receptor-related orphan receptor γt

Retinoid-acid-receptor-related orphan receptor (ROR) γt belongs to the nuclear receptor superfamily. RORγt is an isoform of RORγ expressed in lymphoid tissues. RORγt is located on chromosome 1 and at sequence length of 497 amino acids, it is 21 amino acids shorter than RORγ. The sequences of the two isoforms

are otherwise the same apart from those 21 absent amino acids in the beginning of the RORγt sequence followed by HTS → MRT mutation starting from the first RORγt amino acid. Among the nuclear receptors, RORs are defined as orphan receptors because their endogenous ligands are not known. This does not mean, however, that ligands exhibiting high affinity are nonexistent. For instance, natural products 20α-, 22(*R*)- and 25-hydroxycholesterol have high affinity for RORγ and their binding modes exhibit the characteristics of nuclear receptor agonists, making them also potential endogenous ligands (Solt and Burris 2012, Van Niel *et al.* 2014).

RORγt activity is essential in the proliferation and functionality of T helper 17 (Th17) cells. Th17 cells are abundant at mucosa were they counteract infections caused by bacteria and fungi. As a part of their functionality, Th17 cells produce cytokines such as IL-17 and IL-22 which aid at the Th17 mediated immune response. However, Th17 activity has been connected to inflammatory conditions and autoimmune diseases and being able to limit the Th17 cell functionality has surfaced as a treatment possibility (Littman and Rudensky 2010). Since RORγt is in key role in Th17 cell maturation into functional immune cells, inhibiting RORγt would limit the Th17 cell differentiation.

With this goal in mind, Solt *et al.* (2011) discovered using hydrogen–deuterium exchange mass spectrometry a compound called SR1001 which binds to RORγ ligand binding site inducing a conformational change involving the helix 12. As a result, also the Th17 cell development was inhibited (Solt *et al.* 2011). Encouraging inhibitory effects have also been produced using reversed sulfonamides that act as selective RORγ inverse agonists (Van Niel *et al.* 2014). Currently, both RORγ agonists and inverse agonists are available which contribute to the overall understanding of the nuclear receptor functionality and might aid in alleviating Th17 cell related disorders (Fauber and Magnuson 2014). Additionally, identification of allosteric binding site in RORγt brings forward an alternative approach for inducing RORγt pharmacological antagonism which offers a new outlook into RORγt activity modulation (Scheepstra *et al.* 2015).

2.3   Uridine 5'-diphospho-glucuronosyltransferase 1A10

Uridine 5'-diphospho-glucuronosyltransferase (EC 2.4.1.17, UGT) is an enzyme superfamily which catalyzes a glucuronidation reaction. This conjugation reaction of glucuronic acid is targeted at specific functional groups on lipophilic drugs or other xenobiotics in order to facilitate their secretion through renal pathway. UGTs are the most abundant in liver, however liver is devoid of UGT1A10 which can be found in gastrointestinal tract (Strassburg *et al.* 1999, Rowland *et al.* 2013). The human UGT1A10 sequence is 530 amino acids long and it is formed from UGT1A complex locus on chromosome 2, utilizing alternative promoters and exons (Gong *et al.* 2001).

Currently, the complete mammalian UGT structure has not been solved. This challenges the scientific community to find alternative methods for studying the functionality of this enzyme family. Since UGTs are responsible for clearance and detoxification processes, understanding the determinants of their substrate selectivity are in a vital role for avoiding potentially dangerous drug-drug interactions. This kind of situation may occur if one drug inhibits the metabolism of another drug causing the blood concentration of said drug to increase to toxic level. Since UGT1A10 is expressed the most in small intestine and colon, it has also the potential to affect the bioavailability of orally administered drugs in the system (Rowland *et al.* 2013). Thus, being able to estimate the metabolic consequences of a drug candidate early on in the development process would be essential for avoiding failures related to excretion by UGTs.

2.4 Conventional computer-aided drug design

In drug discovery, computational methods are used to lower the costs and expedite the overall process of transforming a druglike compound into an approved drug. Thanks to a vast volume of biological macromolecule and small molecule information available today, computational tools have become indispensable almost at every step of the drug development process from target identification and validation to preclinical tests.

Computer-aided drug design (CADD) is a term used to describe computational tools and resources involved in the storage, management, analysis and modeling of compounds. The development and use of digital repositories for studying chemical interactions, computer programs for designing compounds and tools for systemic assessment of lead candidates are all part of CADD (Song *et al.* 2009, Ou-Yang *et al.* 2012). Thanks to significant progress over past few decades, CADD is currently a common tool in drug discovery (Zheng *et al.* 2018). Classically, the drug discovery strategies are divided under three main categories depending on the manner of an approach. These categories are structure-based, ligand-based and sequence-based methods. Structure-based drug design includes methods such as molecular docking and *de novo* drug design relying on the structural information of the target macromolecule whereas ligand-based drug design tools like quantitative structure-activity relationship (QSAR) and 2D or 3D similarity assessment, can be applied when structural information is not available. If neither the target structure nor the ligand information is available, sequence-based methods can be applied to identify evolutionally conserved functional sites as potential targets (Jorgensen 2004, Song *et al.* 2009, Ou-Yang *et al.* 2012). However, one single strategy is rarely capable of fulfilling the needs of drug design process which is why combining the strategies in a target-based or phenotype-based manner and incorporating molecular dynamics in order to introduce flexibility is more modern approach (Zheng *et al.* 2018).

Whether compound is considered druglike is determined based on its pharmacokinetic properties. Thus, computational methods are regularly applied to identify undesirable absorption, distribution, metabolism, excretion and toxicological (ADMET) profiles in order to prevent drug safety and efficacy issues as early as possible in the drug development process. One of the most used parameters to ensure the druglikeness and oral availability is Lipinski Rule of Five (Lipinski *et al.* 1997) where druglike compound is determined to have molecular weight under 500, lipophilicity defined as logP under 5, less than 5 hydrogen bond donors and less than 10 hydrogen bond acceptors. Although exceptions to this rule do exist, it provides a good general guideline. In addition, measures should be taken to remove compounds containing toxic or metabolically precarious moieties. For this purpose filters identifying toxicophores or such as Pan Assay Interference Compounds (PAINS) are available. By combining suitable filters, the compounds can be assessed for case specific ADMET property requirements for avoiding undesirable consequences later down the line (Song *et al.* 2009, Lionta *et al.* 2014).

Due to the development of powerful computational technology and increasing amount of success stories, the CADD is nowadays a vital part of creating new pharmaceuticals. As potential druggable targets emerge at accelerating rate (Overington *et al.* 2006) and the druggable genome has been estimated to contain almost 4500 genes (Finan *et al.* 2017) while drugs approved by U.S. Food and Drug Administration (FDA) target only 667 human genome derived proteins (Santos *et al.* 2017), cost efficient and rapid lead discovery methods are in high demand, ensuring growing interest for the CADD. The following chapters describe some of the common tools and conventional approaches used for hit identification and lead optimization over the past decades.

2.4.1 Molecular docking and scoring function

Molecular docking is the most widely applied structure-based drug design technique that has been applied to hit identification and lead optimization in virtual screening since the 1980s (Kuntz *et al.* 1982). Consequently, several molecular docking algorithms have been developed in order to predict the binding mode of a ligand within the ligand binding site. During the process, the binding affinities of the compounds are assessed using scoring functions. The scoring can, in general, be based on force fields which sum the intermolecular interactions, empirical scoring functions which calculate the interactions between binding partners or knowledge-based functions which utilize information gained from known structure-ligand complexes (Lionta *et al.* 2014). While molecular docking can largely predict the correct binding mode of a compound, the ranking based on the scoring functions may not be equally accurate as problems occur with the accuracy of the binding affinity and distinguishing active ligands from inactive molecules (Warren *et al.* 2006, Politi *et al.* 2016, Zheng *et al.* 2018). To circumvent this issue, target-specific scoring

approaches have been established with encouraging outcomes (Knox *et al.* 2007, Wang *et al.* 2017)

2.4.2 Pharmacophore modeling

Pharmacophore concept has been pinpointed to 1898 paper by Paul Ehrlich where pharmacophore was described as characteristic molecular features responsible for bioactivity (Güner and Bowen 2014). Since then the concept has expanded to "a pharmacophore is the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target structure and to trigger (or to block) its biological response" as defined by the International Union of Pure and Applied Chemistry (IUPAC) (Wermuth *et al.* 1998). However, the features are not specific functional groups but abstract models of hydrogen bond donors and acceptors, aromatic rings, hydrophobic groups and charged ions (Zheng *et al.* 2018). Pharmacophore models can be generated using structure-based or ligand-based approach. Structure-based pharmacophore modeling relies on the 3D structure of a target protein as the chemically and spatially relevant features of the active site are used to construct the pharmacophore model. Common issue in structure-based pharmacophore modeling is that too many features might be identified for a specific binding site making practical applications inconvenient (Yang 2010, Zheng *et al.* 2018). Ligand-based approach is applied when the crystal-structure of the target is not available. The common features comprising the pharmacophore model are identified by utilizing 3D structures of known ligands of a specific target. However, the selection of suitable compounds and incorporating ligand flexibility into the models remain as issues (Yang 2010, Zheng *et al.* 2018).

2.4.3 QSAR

One of the most popular computational methods utilized in drug discovery is QSAR modeling. The foundation of QSAR studies is in Hansch-Fujita model reported in 1962 (Hansch *et al.* 1962). In principal, QSAR is an effort to combine structural parameters regarding for example geometrical, thermodynamic, electronic and quantum chemical properties with parameters describing biological activity such as binding affinity and pharmacological properties. In other words, QSAR is a quantitative explanation for structural features of a molecule that contribute to the activity of physiochemical properties (Vaidya *et al.* 2014). Usually QSAR studies are performed in three general steps where first the molecular descriptors are determined, then the model is created and evaluated and finally the model is applied (Vaidya *et al.* 2014, Zheng *et al.* 2018). QSARs can be classified according to the dimensionality of the molecular descriptors where 1D-QSAR models associate activity with properties like pKa or logP and 2D-QSAR begins to connect structural patterns to the biological activity. 3D-QSAR takes into account the three dimensionality of a molecule in explaining the scale and directional inclinations of molecular interactions. 3D-

QSAR is frequently used and several approaches and applications are available (Vaidya *et al.* 2014, Zheng *et al.* 2018). 4D-, 5D- and 6D-QSAR are built on the foundation 3D-QSAR. 4D-QSAR takes into account the flexibility and freedom of alignment by performing ensample averaging with the help of molecular dynamics simulations (Hopfinger *et al.* 1997). 5D-QSAR facilitates the evaluation of multiple induced-fit scenarios and 6D-QSAR examines alternative solvation models (Vedani *et al.* 2005).

2.4.4 Similarity searching

Similarity searching is based on finding molecules similar to the reference structure and it offers a modest yet common method for virtual screening (Johnson *et al.* 1989, Maldonado *et al.* 2006). Accordingly, similarity searching assumes that molecules sharing similar structure also share similar bioactivity. The determination of molecular similarity consists of assigning descriptors to the reference structure and comparing the descriptors to each molecule in a database. The descriptors establish a similarity measure that contains the representation, the similarity coefficient as well as the weighting scheme and the similarity measure can then be used to rank and sort the results of the similarity search (Willett *et al.* 1998, Zheng *et al.* 2018).

Constitutional or 1D descriptors contain information concerning the physical, chemical or biological features of a molecule in addition to its structural properties as atom and bond orders. Configuration or 2D descriptors define the 3D arrangement of the atoms and conformational or 3D descriptors add the atom's thermodynamically stable spatial arrangements. As a whole, thousands of descriptors are known and they range from simple and common ones like molecular codes to complex and specialized that combine multiple descriptors (Maldonado *et al.* 2006). The coefficients transform compatible pairs of molecular representations into numerical values providing a quantitative measure for chemical alikeness. The alikeness can be described as being local or global. Local similarities focus on certain parts such as atoms and functional groups of an object whereas global similarities determine the alikeness between two whole objects. In addition, the coefficients can be grouped into correlation, probability, association and distance by the manner they define the similarities. Similarity coefficients like Tanimoto/Jaccard coefficient (Willett *et al.* 1998) and Tversky coefficient (Tversky 1977) are commonly used. Lastly, weight schemes are used to allocate importance to the components formulating the molecular similarity representation. The components are then combined in order to create various frameworks for automated similarity analysis. The main goal is to be able to keep the computational requirements low enough for efficient comparison without ignoring significant compound properties (Willett *et al.* 1998, Maldonado *et al.* 2006).

2.4.5 Prediction of ligand binding site

The actions of proteins are guided by interactions with other molecules which range from small ligands to other proteins. The locations where these interactions take place are determined as orthosteric sites, allosteric sites or hot spots depending on the type of interaction. Orthosteric sites are ordinary ligand binding sites whereas allosteric sites are indirect, distinctive binding locations. In contrast, hot spots are key amino acid residues involved in protein-protein interactions (Zheng et al. 2018). Consequently, biological functions are routinely modulated through known functional sites. However, in certain cases information concerning functional sites is not available and thus prediction methods can be applied.

The prediction methods of orthosteric sites can be divided into sequence-based and structure-based methods. Sequence-based methods exploit the sequence similarity and evolutionally conserved information of homologous proteins to predict ligand binding sites. Position-specific scoring matrices or graph-based methods can be applied but methods utilizing machine learning benefit from the possibility to directly interpret multiple sequence alignment profiles (Roche et al. 2015). Additionally, multiple features can be combined and information regarding the secondary structure can be included into the prediction method like in TargetS (Yu et al. 2013). Structure-based methods are founded on 3D information of atomic coordinates. Hence, structure-based prediction methods cannot be applied if 3D information is unavailable. In general, structure-based methods can be divided into four categories. Template-based methods take into account homologous proteins and the structural and biological functions they share (Petrey et al. 2015). Geometry-based methods calculate geometric feature values with the goal of identifying curved in impressions and grooves. Energy-based methods calculate interaction energies between chemical probe molecules and the protein surface in order to identify energetically favorable binding regions. Lastly, miscellaneous methods contain approaches that consider for example surface accessibility like LIGSITE[csc] (Huang and Schroeder 2006) and thus cannot be unambiguously placed into the other categories (Roche et al. 2015).

The prediction of allosteric sites consists of various interrelated methods (Collier and Ortiz 2013). In a broad sense, the utilized methods can be classified under feature-based models, structural surveys or molecular simulations. In feature-based models, structural features such as accessible surface area and conserved amino acid residues are incorporated into the prediction. Structural surveys make use of conformational series of protein structures in order to gain insights regarding fluctuations in the protein structure. Fluctuation are also looked into through molecular simulations and approaches utilizing for example bond-bond propensities (Amor et al. 2016) can be used (Collier and Ortiz 2013). Although targeting allosteric sites is considered to have advantages such as higher selectivity and virtually nonexistent chance of overdosage as they do not compete with endogenous ligands, some of these

sites are also targets for endogenous modulators under physiological conditions (Christopoulos 2002).

Hot spots affecting protein-protein interactions can be generally predicted in two ways, either based on unbound protein or bound protein-protein complex structures (Cukuroglu *et al.* 2014). Most of the hot spot prediction methods focus on bound protein-protein complexes where approaches like packing interaction energetics, molecular dynamics simulations and machine learning from known hot spot data have been utilized. Methods based on unbound protein structures are rather scarce but prediction tools using for example amino acid sequences or dynamic fluctuations in high frequency modes have been implemented (Cukuroglu *et al.* 2014).

Predicting orthosteric and allosteric sites as well as hot spots offers rapid ways to discover functional sites suitable for designing small molecule modulators. As the methods for identifying functional sites continue to develop and more functional sites are recognized, the means to modulate biological functions provide an increasing amount of advantageous options.

2.5   Simple coumarins as lead molecules

Coumarin ($C_9H_6O_2$, chromen-2-one, Fig. 4) was first isolated from Tonka beans (*Dipteryx odorata*) in 1820 by A. Vogel. Coumarin and its derivatives are organic molecules which occur naturally in many plant families (Table 1). As a fragrant molecule, coumarin can be found in essential oils and it has been used in food and cosmetic products due to its sweet smell (Borges *et al.* 2005). Simple coumarins which are the focus here, are all build upon the coumarin core (Fig. 4). Along with simple coumarins, there are roughly five other core structures (furanocoumarins, pyranocoumarins, biscoumarins, triscoumarins and coumarinolignans) that can be used to distinguish coumarin derivatives.

In addition to isolating coumarins from numerous plant sources, synthetically created coumarin derivatives expand the amount of currently known coumarin structures (ChEMBL 26.7.2017 996 hits). Coumarins are classically synthetized by Perkin, Pechmann or Knoevenagel reactions but Wittig, Kostanecki–Robinson and Reformatsky reactions are valuable additions to the synthesis of these heterocycles. Each of these reactions have their disadvantages and more efficient, reliable and high-yielding methods continue to be developed (Borges *et al.* 2005). The continuing interest in the coumarin derivatives is fueled by the vast and variable biological activities these compounds seem to have. Consequently, the attention directed to coumarin derivatives aspires to tap into their therapeutic potential and convert the activities into advantageous applications.
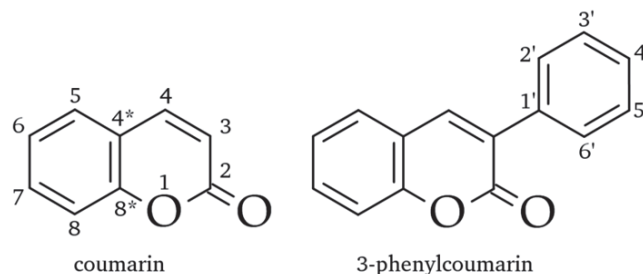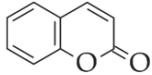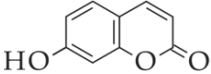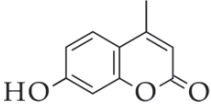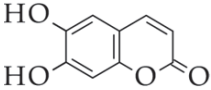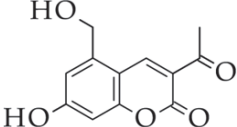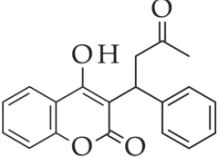
FIGURE 4   The 2D representations of the coumarin scaffold and the 3-phenylcoumarin. The coumarin scaffold has been numbered according to IUPAC nomenclature of organic chemistry. Because the 3-phenyl group on the 3-phenylcoumarin is considered to be a substituent, it follows its own numbering designated with ' character. Thus, the numbered positions and substitutions discussed later in the text follow the numbering seen above, respectively. However, the bond between the coumarin scaffold and the 3-phenyl group is able to rotate and in order to highlight this fact, position 4' is referred to as *para*, positions 3' and 5' as *meta* and positions 2' and 6' as *ortho* in interchangeable manner.

As plant derived secondary metabolites belonging to the flavonoid class of molecules, coumarins are generally considered to have low toxicity (Sandhu *et al.* 2014). Conversely, European Food Safety Authority has set the tolerable daily intake (TDI) of coumarin to 0.1 mg/kg (Aguilar *et al.* 2008) on the grounds of hepatotoxicity. For example, cassia cinnamon has particularly high coumarin content which is why food items rich in cinnamon may create a risk for susceptible individuals. In humans, coumarin is excreted via 7-hydroxylation catalyzed by cytochrome P450 (CYP) 2A6, thus individuals with CYP2A6 polymorphism may be at risk, although the exact mechanism of coumarin hepatotoxicity has not been explicitly determined (Abraham *et al.* 2010) and coumarin intake higher than the TDI does not necessarily equate abnormal liver function (Iwata *et al.* 2016). In contrast to the potential hepatotoxicity, naturally occurring coumarins are known instead to have antiproliferative, antioxidant and anti-inflammatory functionalities which have sparked the interest for designing novel coumarin derivatives for various pharmaceutical applications (Borges *et al.* 2005, Sandhu *et al.* 2014, Stefanachi *et al.* 2018). As a result, some of the coumarin derivatives such as anticoagulants warfarin, acenocoumarol and phenprocoumon (Verhoef *et al.* 2014), antibiotic armillarisin A (Wu *et al.* 2015) as well as furanocoumarins psoralen and methoxsalen used in psoralen+UVA treatment of certain skin conditions (Bethea *et al.* 1999) (Table 1), have also made it as treatment options available to the patients. Persistent efforts amidst coumarins continue to reveal new potential applications in pharmacy for this versatile group of molecules. Thus, coumarins have been implicated for example in treatment of certain types of cancers and Alzheimer's disease (Borges *et al.* 2005, Sandhu *et al.* 2014, Stefanachi *et al.* 2018).

TABLE 1   Overview of commonly used coumarin derivatives. Examples of the typical natural sources of simple coumarins and their commercially relevant application are given.

Compound	Source and application	Reference
 Coumarin	*Dipteryx odorata* Fragrance in cosmetics	(Borges *et al.* 2005)
 Umbelliferone 7-hydroxycoumarin	Apiaceae/Umbelliferae family of celery, carrot and parsley Sunscreen and optical brightener	(Mazimba 2017)
 Hymecromone 4-methylumbelliferone	Synthetic Treatment of biliary spasm	(Nagy *et al.* 2015)
 Esculetin 6,7-dihydroxycoumarin	*Cichorium intybus* L *Fraxinus rhynchophylla* Hance Present in herbal medicine	(Bais and Ravishankar 2001) (Liang *et al.* 2017)
 Armillarisin A 3-acetyl-5-hydroxylmethyl-7-hydroxycoumarin	*Armillaria* genus of honey fungus Treatment of cholecystitis	(Wu *et al.* 2015)
 Warfarin 4-Hydroxy-3-(3-oxo-1-phenyl-butyl)coumarin	Synthetic Anticoagulant	(Link 1959) (Verhoef *et al.* 2014)
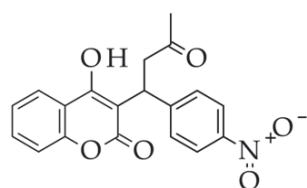
| | Synthetic Anticoagulant | (Verhoef *et al.* 2014) |

Acenocoumarol
4-hydroxy-3-[1-(4-nitrophenyl)-
3-oxobutyl]coumarin
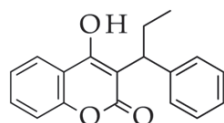
| | Synthetic Anticoagulant | (Verhoef *et al.* 2014) |

Phenprocoumon
3-(1-Phenylpropyl)-
4-hydroxycoumarin
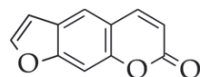
| | *Ficus carica* L | (Chunyan *et al.* 2008) |
| | Psoralen + UVA treatment of skin conditions (vitiligo, psoriasis) | (Bethea *et al.* 1999) |

Psoralen
6,7-Furanocoumarin
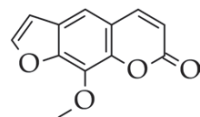
| | *Ammi majus* L | (Schönberg and Sina 1948) |
| | Psoralen + UVA treatment of skin conditions (vitiligo, psoriasis) | (Bethea *et al.* 1999) |

Methoxsalen
8-Methoxyfuranocoumarin

The potential of coumarin and especially 3-phenylcoumarin scaffold (Fig. 4) for building MAO inhibitors, has been explored extensively in the past, for review see (Matos *et al.* 2013b, Patil *et al.* 2013). Catto et al. (2006) studied the 3-, 4- and 7-polysubstituted coumarins and established MAO B selectivity with 7-(m-chlorobenzyloxy)-3-methylcoumarin (Catto *et al.* 2006). Matos *et al.* (2009) explored 6-methyl-3-phenylcoumarins with varying methoxy and halogen substitutions added into the phenyl ring. Inserting two methoxy groups at the positions 3′ and 5′ of the phenyl ring produced both MAO B selectivity and potency (Matos *et al.* 2009a). Furthermore, adding a bromo group in the 8′ position of the phenyl ring of 6-methyl-3-phenylcoumarin improved activity with the *para*-methoxy substituted 3-phenyl ring (Matos *et al.* 2010). The 3-phenylcoumarin derivatives produced MAO B selectivity whereas the 3-benzoylcoumarin, the 6-bromo-8-hydroxy-3-benzoylcoumarin derivatives in particular, exhibited MAO A selectivity (Matos *et al.* 2011b). Matos *et al.* (2011a)

established that the *para* or *meta* moieties at the phenyl ring of 3-phenylcoumarin are beneficial for the MAO B activity (Matos *et al.* 2011a). In addition, Serra et al. (2012) also showed that a chloro group at the *meta* position in the *p*-methoxy-3-phenyl ring on 3-aryl-4-hydroxycoumarin improves MAO B activity (Serra *et al.* 2012).

Anticancer agents from natural sources have attracted attention due to their cancer chemopreventive potential (Kaur *et al.* 2018). Conventional synthetic chemotherapeutic agents often develop resistance and they tend to have severe side effects that are particularly prominent on normal proliferating tissues. Thus, the chemopreventive antioxidative, anti-mutagenic and anti-carcinogenic properties of phytochemicals have led into the use of natural agents as part of cancer management and treatment in addition to the development of new molecules for treatment through molecular hybridization approach (Sandhu *et al.* 2014, Kaur *et al.* 2018). Accordingly, coumarin derivatives have been extensively studied as anticancer agents (Sandhu *et al.* 2014). In breast cancer therapy, the main goal is to prevent the action of estrogens. Estrogens are produced through aromatase or sulfatase pathway and coumarin derivatives have been successfully investigated in both. Sulfatase mediates the cleavage of hormone precursor estrone sulfate into active hormone and, thus, inhibiting sulfatase represents viable therapeutic option in breast cancer. Consequently, coumarin derivatives, especially coumarin sulfamates have been found active in this target (Musa *et al.* 2008). Woo *et al.* (1998) investigated steroidal and nonsteroidal sulfamates as steroid sulfatase inhibitors and found 3,4-dimethylcoumarin 3-*O*-sulfamate especially potent (Woo *et al.* 1998). In addition, Stanway *et al.* (2006) discovered STX64 (667 coumate, 6-Oxo-6,7,8,9,10,11-hexahydrocyclohepta[c]chromen-3-yl sulfamate) as sulfatase inhibitor and have also acquired promising results from Phase I clinical trials (Stanway *et al.* 2006).

Aromatase inhibitors have the potential to affect breast cancer development by reducing the cell proliferation and the estrogen levels. Chen *et al.* (2004) found 4-benzyl-3-(4-chlorophenyl)-7-methoxycoumarin to be a potent competitive aromatase inhibitor and determined that the functional groups 3-(4-chlorophenyl), 4-benzyl, and 7-methoxyl are especially important for the activity of this compound (Chen *et al.* 2004). In addition, coumarin derivatives have been implicated as selective estrogen receptor modulators (Musa *et al.* 2008, Luo *et al.* 2017). This kind of modulators exhibit high affinity to ER but are inactive in other steroid hormone receptors. Selective estrogen receptor modulators are also capable of acting as both ER agonists and antagonists depending on the target tissue. In this regard, substitutions on positions 3 and 7 of 3-substituted coumarins have been found particularly beneficial for ERα activity. Furthermore, coumarin derivatives have been used to create coumarin-estrogen conjugates in order to target ER and simultaneously circumvent non-selectivity and toxicity related issues (Musa *et al.* 2008).

Along with the cases already discussed here, coumarins have exhibited RORγt related activity. Yao *et al.* (2011) have reported a reduction in Th17 cells present in serum and a decrease in expression level of RORγt in collagen-

induced arthritis rat model after administration of plant-based simple coumarin daphnetin (7,8-dihydroxycoumarin) (Yao *et al*. 2011).

3   AIMS OF THE STUDY

In this study, a selection of druggable targets is thoroughly investigated utilizing a readily modifiable core structure, 3-phenylcoumarin. By altering the properties of the ligand candidate, detailed information is sought regarding the pharmacology and binding characteristics. With the intention of examining the requirements for achieving selectivity within the target group consisting of monoamine oxidases, nuclear receptors and uridine 5'-diphospho-glucuronosyltransferases, 3-phenylcoumarins are tailored to correspond with the prevalent binding environment. The study also delves into the metabolic fate and the assay usability of 3-phenylcoumarin derivatives as they are allocated to metabolizing target. Furthermore, computational methods are combined with experimental techniques in order to contribute to the understanding of the interactions between these established targets and 3-phenylcoumarin derivatives.

4 METHODS

Table 2 summarizes the methods used in the original publications comprising the thesis. This chapter briefly discusses the most important methods whereas the detailed explanations are reserved for the original publications.

TABLE 2   Summary of the methods.

Method	Publication
DATABASES	
ChEMBL	II, IV
DUD	II
Specs	II, IV
Easy-to-synthesize	I, II, III, V
PDB	I, II, III, IV, V
LIGAND PREPARATION	
LigPrep	I, II, III, IV, V
ConfGen	III, IV
NEGATIVE IMAGE METHODS	
NIB VOIDOO/FLOOD	II
Panther	I, III, IV, V
SIMILARITY COMPARISON	
ShaEP	II, III, IV
MOLECULAR DOCKING	
GLIDE docking	II, IV
PLANTS docking	I, II, V
MOLECULAR DYNAMICS SIMULATIONS	
NAMD	II
ENRICHMENT METRICS	
ROC AUC	II
MODELING	
Homology modeling	V
EXPERIMENTAL METHODS	
Continuous spectrophotometric assay	I, III
Fluorescence polarization	II, III
Glucuronidation assay, spectrophotometric analysis (by collaborators)	III, V
High-performance liquid chromatography and -mass spectrometry (by collaborators)	III, V

Organic synthesis (by collaborators)	I, II, III, V
Recombinant protein production, mutagenesis and immunodetection (by collaborators)	III, V
Reporter assay system	IV

4.1 Databases

ChEMBL (II, IV) is openly available large-scale bioactivity database which contains mostly manually extracted information from the medicinal chemistry literature (Bento *et al.* 2014). It was used to retrieve molecules with experimentally measured activities. The active molecules are predominantly chosen from a single research article in order to ensure the homogeneity of the experimental factors, although in some cases carefully selected research articles could be included to provide a wider variety of molecules.

A directory of useful decoys, DUD (II), is a collection of active and decoy compounds used to benchmark virtual screening methods (Huang *et al.* 2006). It has been designed to assess docking algorithms by offering decoys with the physical properties of the active compounds but dissimilar topology. There are 2950 ligands for total of 40 targets and each ligand has 36 decoys creating a database of 98 266 compounds. The compounds in DUD are from ZINC (Irwin and Shoichet 2005), which is a database of commercially available compounds meant for virtual screening.

Specs (Specs, Zoetermeer, The Netherlands, www.specs.net) (II, IV) is a commercial supplier of screening compounds. Their database contains chemically diverse, well-characterized molecules with drug-like properties. In addition, the compounds meet the criteria of being biologically active and adhere to ADME (Caldwell *et al.* 1995, Lipinski *et al.* 1997) requirements.

An in-house easy-to-synthesize database (I, II, III, V) consists of 75 coumarin derivatives. The compound collection has been synthesized from affordable starting materials with one-step synthesis, excluding compounds requiring protective groups. The synthesis utilizes relatively fast microwave-assisted organic synthesis making the database an affordable compound collection to use in experimental testing.

Protein Data Bank (PDB) (www.rcsb.org, (Berman *et al.* 2000)) (I, II, III, IV, V) is an archive compiling the structural data of biological macromolecules. It consists of the atomic coordinates of proteins and nucleic acids ranging from small peptides and short DNA strands to complex molecular machines. When structural information is acquired from the PDB to be used in research, resolution, electron density and overall quality of the structure as well as the bound ligand molecules affecting the protein conformation are important factors to consider.

4.2 Ligand preparation

4.2.1 LigPrep (I, II, III, IV, V)

LigPrep is a computational method for generating 3D molecular structures by Schrödinger (Schrödinger LLC, New York, NY, United States of America). The structures LigPrep generates are energetically minimized and the method can be used to create tautomeric and ionization states, ring conformations as well as stereoisomers depending on the needs of the project. In original publications I to V, LigPrep was used to generate 3D structures with protonation states present at pH 7.4.

4.2.2 ConfGen (III, IV)

ConfGen is a conformational search method for generating energetically favorable bioactive 3D conformers by Schrödinger (Schrödinger LLC, New York, NY, United States of America). ConfGen generates the conformers in three steps. First, the variable features are identified, then the conformations are generated and finally the conformers are selected and refined (Watts *et al.* 2010). In both of the original publications III and IV, ConfGen was used as a preparatory step in order to generate biologically active conformers for docking.

4.3 Negative image methods

Two negative image-based (NIB) approaches NIB VOIDOO/FLOOD models and Panther models were used. The intended purpose differs slightly but here they both were applied to cavity analysis and to ligand binding area recognition which can be further refined to produce models for virtual screening.

4.3.1 NIB VOIDOO/FLOOD (II)

In this method, NIB model of the binding site of the protein target is created with NIB VOIDOO/FLOOD (Kleywegt and Jones 1994). The VOIDOO program has been originally designed to detect new cavities or to characterize old ones in macromolecular structures. VOIDOO relies on grid-based cavity search which is able to recognize voids and invaginations. The cavity search is then completed by determining the volume of the found cavities. In addition, FLOOD is small complementary program that attempts to pack as many solvent molecules as possible to the solvent-accessible cavity characterized by VOIDOO. Together these programs can be utilized in model creation for virtual screening.

As a NIB model creation tool, VOIDOO/FLOOD parameters have been adjusted to best describe the environment typical for ligand in a ligand binding cavity of a protein. As the interactions between ligand and protein are dominated by hydrogen bonds and van der Waals (vdW) forces, the probe radius for delineating the cavity was adjusted to 1.25 Å from 1.4 Å default setting. Combined with increased carbon atom vdW radius from 1.85 Å to 2.25 Å, these settings provided 3.50 Å distance between the protein and the negative image. The achieved distance corresponds with the notion that hydrogen bonds can exist when the distances between atoms other than hydrogen are as low as 2.6 Å whereas vdW forces are likely to dominate starting from 3.6 Å. In addition, the charged and polar amino acids were taken into account by decreasing the vdW radius of electronegative atoms oxygen from 1.6 Å to 1.2 Å and nitrogen from 175 Å to 1.2 Å. This allowed the model to be created closer to the charged and polar amino acids amino acids and slightly further away from hydrophobic amino acids (Virtanen and Pentikäinen 2010, Niinivehmas *et al.* 2011). The model was additionally enhanced by assigning electrostatic information for amino acids in ligand binding area using atom-centered MMFF94 charges (Halgren 1996a, b, c, d, Halgren and Nachbar 1996). With the help of the protein charges, the opposite charges were then created for the NIB model data points. This was done by averaging each charge value within 2.7 Å distance from a given model point and assigning the opposite charge to that point. Since proteins are flexible, the alterations in conformation can be taken into consideration by creating NIB models for multiple X-ray crystal structures or for snapshots from MD simulations (Niinivehmas *et al.* 2011). The shape and electrostatic potential of the generated models can then be compared to a dataset with ShaEP (Vainio *et al.* 2009).

4.3.2 Panther (I, III, IV, V)

Panther was developed based on the idea applied to NIB VOIDOO/FLOOD models. Like NIB VOIDOO/FLOOD, Panther can be used to search cavities, analyze cavity volume and fill cavities with molecules. However, in contrast to VOIDOO/FLOOD, Panther algorithm was designed to be suited for virtual screening with accurate recognition and delineation of ligand binding area combined with model creation by filling the cavities in adjustable manner. Panther's asset is that the parameters can be systematically adjusted to accommodate each individual cavity search process with the chance to perfect the model to accurately represent the features of the ligand binding site (Niinivehmas *et al.* 2015).

In Panther protocol, X-ray crystal structure of a protein is utilized in order to create simple atomistic shape-electrostatic model of the ligand binding area. The protocol starts by selecting any Cartesian coordinate (x, y, z) within the cavity of interest. This point is used to identify the cavity, after which the amino acids lining the cavity are pursued. This is accomplished by dividing the protein and possible cofactor atoms found in the PDB of the given structure into spherical segments. From each segment, the atom of an amino acid closest to

the center is selected and this amino acid is, thus, determined to align with the cavity. Once the cavity has been specified, the first model atoms lining the cavity are placed around polar and charged atoms of the residues and the cofactors. These model atoms are placed in a controlled manner which takes into account the angle towards the hydroxyl groups or main chain nitrogen, the geometry and the distance in relation to the target atom. The charges for the model atoms determined this way are either from separate input file or from temperature factor field of the PDB file and they are determined to be the opposite to the charges aligning the cavity. Remainder of the cavity is filled with non-charged atoms starting from the given center point. During the process, atoms overlapping with or too far from the protein are eliminated. Additionally, unconnected atoms which might reside in secondary pockets or in irrelevant grooves are left out of the model. As the final phase, the volume of the cavity is determined with the help of the atom identified to align the cavity. Model atoms outside this volume are removed and the generated charged atoms are confirmed to reside alongside the filler atoms. The newly generated model can then be compared to possible ligands with ShaEP (Vainio *et al.* 2009) in order to look for similarities in shape and electrostatic potential (Niinivehmas *et al.* 2015).

4.4 Similarity comparison

4.4.1 ShaEP (II, III, IV)

ShaEP, derived from the words shape and electrostatic potential, is a tool for superimposing compounds and evaluating their similarity (Vainio *et al.* 2009). In original publications II to IV, ShaEP was used to compare the shape and electrostatic potential of the NIB or Panther models with the ligands in each study.

4.5 Molecular docking

The purpose of molecular docking is to predict the binding mode of a ligand to a protein with *in silico*. Two entwined steps are required in order to accomplish the docking. These steps are the sampling of ligand conformations in the active site of the protein and the ranking of the conformations based on scoring function. Molecular docking software use various sampling algorithms to replicate the experimentally observed binding mode and scoring functions should recognize them by ranking them highest among the generated conformations (Meng *et al.* 2011).

4.5.1 GLIDE docking (II, IV)

GLIDE or grid-based ligand docking with energetics is a docking and scoring method by Schrödinger (Schrödinger LLC, New York, NY, United States of America). As the accuracy of docking while maintaining computational speed required to screen large libraries is the main challenge of any docking program, GLIDE strives to comprehensively examine the ligand binding site in regards to the potential position, orientation and conformation of the ligand. This is obtained by using series of hierarchical filters which are combined with grid presentation of the receptor shape and properties containing a number of fields available to progressively define the score of the position of the ligand. The next step creates initial ligand conformations which are used to probe promising ligand positions. The identified positions are then minimized in the field of the receptor, out of which the lowest-energy positions of ligands continue to Monte Carlo procedure examining nearby torsional minima. Finally, the binding affinity and the ranking of the ligands is predicted using GlideScore which is a GLIDE scoring function, *GScore* described below (Friesner *et al.* 2004, Halgren *et al.* 2004).

The docking using GLIDE started by preparing the protein X-ray crystal structure with the Protein Preparation Wizard in Maestro (Schrödinger LLC, New York, NY, United States of America). The receptor grid as mentioned above was generated by GLIDE to be used in the docking. Both of the modes, GLIDE standard precision (II, IV) and extra precision (II) were used. Depending on the original publication, one of the two scoring functions of GLIDE, *GScore* were applied:

$GScore = 0.065*vdW + 0.130*Coul + Lipo + Hbond + Metal + BuryP + RotB + Site$
(II: GLIDE version 5.7)

$GScore = 0.05*vdW + 0.15*Coul + Lipo + Hbond + Metal + Reward + RotB + Site$
(IV: GLIDE version 6.6)

where *vdW* is van der Waals energy, *Coul* is Coulomb energy, and *Lipo* is lipophilic term which is in the form of lipophilic-lipophilic pair terms in standard precision docking but is derived from hydrophobic grid potential in order to favor suitable hydrophobic interactions in extra precision docking. *Hbond* is hydrogen bonding term, *Metal is* metal binding term, *BuryP* is penalty for buried polar groups and *RotB* is penalty for freezing rotatable bonds. *Site* considers the polar interactions in the active site by rewarding polar atoms which are not hydrogen-bonding in a hydrophobic region. Whereas *BuryP* penalizes buried polar groups in the former equation, *Reward* is used in the latter to reward and penalize various features including the buried polar groups but also hydrophobic enclosures, correlated hydrogen bonds, amide twists and so on, covering all terms which are not explicitly mentioned.

4.5.2 PLANTS docking (I, II, V)

PLANTS or protein-ligand ant system is a docking algorithm based on ant colony optimization. It utilizes virtual ant colony to sample the search area and to find a minimum energy conformation for the ligand in the ligand binding cavity of the target. PLANTS has two scoring functions $PLANTS_{PLP}$ and $PLANTS_{CHEMPLP}$ both of which are a combination of other existing scoring functions. For modeling steric fit between the protein and the ligand, PLANTS relies on piecewise linear potential or PLP scoring function f_{PLP}. Intraligand interactions are taken into account by introducing torsional potential f_{tors} from Tripos force field and a heavy atom clash term f_{clash}. The search is limited to the binding cavity of interest with the term C_{site}. In $PLANTS_{CHEMPLP}$ ChemScore implementation from GOLD is additionally used to include the angle-depended terms f_{hbond} and f_{metal} for hydrogen bonding and metal binding. Thus, the simplified scoring functions can be described as:

$$PLANTS_{PLP} = f_{PLP} + f_{tors} + f_{clash} + C_{site}$$
$$PLANTS_{CHEMPLP} = PLANTS_{PLP} + f_{hbond} + f_{metal}$$

The PLANTS docking and the scoring functions have been described in more detail in (Korb *et al.* 2006, 2009). Here, PLANTS version 1.2 and the scoring function $PLANTS_{CHEMPLP}$ with default parameters were used. Depending on the original publication, the preparatory steps taken before PLANTS docking varied. In original publication I, the docking was performed on the MAO B X-ray crystal structure 2V60 whereas in original publication II, the proteins were prepared using the Protein Preparation Wizard in Maestro (Schrödinger LLC, New York, NY, United States of America) as described in GLIDE docking. For original publication V the shape and electrostatic properties of the ligand binding cavity of UGT1A enzymes were analyzed with Panther.

4.6 Molecular dynamics simulations

4.6.1 NAMD (II)

NAMD Scalable Molecular Dynamics is a code designed to execute high-performance simulations of sizable biomolecular systems. NAMD performs these molecular dynamics in parallel which makes the code vastly scalable not to be run only on personal computers but also on hundreds of processors of a high-end parallel platform or on anything in between. Ease of use has been one of the main goals when developing NAMD, hence it is free to use and compatible with various input files. Other key features include basic simulations such as rigid water molecules and bonds to hydrogen atoms but also advanced simulations like alchemical and conformational free energy calculations (Phillips *et al.* 2005).

In original publication II, NAMD version 2.6 (Phillips *et al.* 2005) was used to run molecular dynamics simulations according to previously published protocol (Postila *et al.* 2011, Ylilauri and Pentikäinen 2012, 2013). The purpose was to introduce flexibility to the protein for the NIB model creation. Amber ff03 force field (Duan *et al.* 2003) parameters were used, adding hydrogen atoms with TLEAP in Antechamber 1.27 (Wang *et al.* 2004). The system was solvated with rectangular box of transferable intermolecular potential three-point, TIP3P water molecules (Åqvist 1990) which extended 13 Å away from the solute. Shortly, the molecular dynamics simulation consisted of (1) energy minimization of the water molecules and amino acid sidechains with conjugate gradient algorithm, keeping the carbon backbone in place with harmonic 5 kcal (mol Å^2)$^{-1}$ force, (2) minimization without constrains, (3) running the simulation at constant 300 K temperature and 1 atm pressure with the same constrains as in first minimization step and (4) repetition of the previous step using Langevin-Piston method. Electrostatic interactions were taken into account by applying cutoff value of 12 Å for vdW interactions and treating long-range electrostatics with Particle Mesh Ewald (Darden *et al.* 1993, Petersen 1995) method. Periodic boundary conditions were used as well.

4.7 Enrichment metrics

4.7.1 ROC AUC (II)

Receiver operating characteristics (ROC) curves are a method for describing the compromises that can be made between the true positive fraction and the false positive fraction by graphing the former against the latter (Metz 1978). In virtual screening ROC curves are used to visualize the efficiency of the selected method to distinguish between active ligands (true positives) and inactive molecules (false positives). ROC is usually presented together with an area under the curve (AUC) which describes the probability to correctly distinguish a randomly selected true positive and a randomly selected false positive from each other (Hanley and McNeil 1982). Consequently, perfect virtual screening method would be able to flawlessly distinguish the two, yielding AUC value of 1.0. On the other hand, if the sampling of the method is completely random, the resulting AUC value is 0.5.

Two methods can be used to determine the enrichment factors *EF*. The first one, $EF_{top\ X\ \%}$, describes the top X % of the results and the second one, $EF_{X\ \%\ FalsePositive}$, the top results until X % of the false positive molecules have been encountered.

$$EF_{top\ X\ \%} = (TruePositive_{X\ \%}\ /\ Molecules_{X\ \%}\ )\ /\ (TruePositive_{all}\ /\ Molecules_{all})$$
$$EF_{X\ \%\ FalsePositive} = (TruePositive_{X\ \%\ FalsePositive}\ /\ TruePositive_{all})\ \cdot 100$$

where $TruePositive_{X \%}$ is the number of active ligands in the top X % of the screened compounds, $Molecules_{X \%}$ is the number of molecules in the top X % of the screened compounds, $TruePositive_{all}$ is the total number of screened ligands and $Molecules_{all}$ is the total number of screened molecules. In the latter equation $TruePositive_{X \% \ FalsePositive}$ is the number of active ligands when X % of the false positive molecules have been found.

4.8 Modeling

In the context of computer-aided drug design, modeling entails all methods used to mimic molecules and their behavior. It can be applied starting from small chemical reactions between few atoms to large molecular machines. Predicting reaction mechanisms, interactions and conformational properties are all essential in order to understand the interactions between biological target and a ligand. Evidently, these are the questions modeling can aid in answering, especially if the information is otherwise unobtainable.

4.8.1 Homology modeling (V)

In homology modelling, the model building usually starts with the identification of suitable templates and continues with the alignment of the sequences. The model coordinates are then generated based on the alignment with the templates after which the created model is optimized and validated (França 2015). In order to model the human UGT1A enzymes, all of their sequences (Q9HAW8 (1A10), O60656 (1A9), Q9HAW9 (1A8), Q9HAW7 (1A7), P19224 (1A6), P35504 (1A5), P22310 (1A4), P35503 (1A3), and P22039 (1A1)) were gathered from the Uniprot Knowledgebase at www.uniprot.org (The UniProt Consortium 2017). The sequences were used to identify protein structures with BLAST (Altschul *et al.* 1990) search against the PDB to act as templates in the model building. Four structures 2O6L (Miley *et al.* 2007), 3HBF (Modolo *et al.* 2009), 3WC4 (Hiromoto *et al.* 2013) and 2C1Z/X (Offen *et al.* 2006) were selected based on the search and these structures served as the foundation of the sequence alignment phase of the model building. The chosen alignment method was protein structure based sequence alignment which was created with Vertaa in BODIL (Lehtonen *et al.* 2004) where 2C1Z was used as the template because it provided the best match with UGT1A10 and had both N- and C-termini. Next the gathered UGT1A sequences were aligned against the structural alignment in BODIL using STRMAT110 matrix (Johnson and Overington 1993) combined with gap penalty of 40. The resulting alignment was examined carefully and adjustments were made to compensate for variations in the sequence length. Models were constructed for each member of the UGT1A family using Modeller version 9.15 (Šali and Blundell 1993).

4.9 Experimental methods

Assay systems were selected based on the protein target under investigation. Thus, three different assay systems were used: Continuous spectrophotometric assay, fluorescence polarization and reporter assay system. Rest of the assays as well as the organic synthesis were performed by collaborators.

4.9.1 Continuous spectrophotometric assay (I)

Peroxide-linked, continuous spectrophotometric assay suitable for measuring monoamine oxidase activity as described by Holt *et al*. (1997) was used to determine the inhibitory potential of the in-house database of easy-to-synthesize coumarin derivatives. The assay is based on the detection of hydrogen peroxide which is a metabolic product formed during deamination of monoamine oxidase substrates. In the assay, the substrate *p*-tyramine is deaminated by the monoamine oxidase. The resulting hydrogen peroxide acts then as an oxidizing agent in the peroxidase reaction catalyzed by horseradish peroxidase. As a result, oxidized 4-aminoantipyrine donates a proton and condenses with vanillic acid forming red quinoneimine dye. The absorption maximum of the quinoneimine dye is at 498 nm which is recorded on a spectrophotometer. With the assay conditions reported, the absorbance change should be approximately 0.35 (Holt *et al.* 1997).

The assay was performed on 96-well plates in 0.2 M potassium phosphate buffer at pH 7.6. The monoamine oxidase and possible inhibitor candidates were first incubated for 30 min at 37 °C in chromogenic solution containing 250 µM vanillic acid, 125 µM 4-aminoantipyrine and 2 U/ml horseradish peroxide calculated in the final volume of 200 µl on the plate. After the incubation, the final volume was reached by adding 20 µl of *p*-tyramine to achieve 0.5 mM concentration on the plate. The change in absorbance was recorded and the monoamine oxidase activity or the inhibitory potential of the tested candidates were calculated accordingly.

4.9.2 Fluorescence polarization (II)

Commercial PolarScreen™ Nuclear Receptor Fluorescence Polarization Competitor Assay Kit, Green (Thermo Fisher Scientific, Waltham, MA, United States of America) was selected for screening potential ERα ligands. The activities of the ligand candidates were first computationally predicted and the most promising hits both purchased from commercial database Specs (Specs, Zoetermeer, The Netherlands) and from the collection of easy-to-synthesize coumarin derivatives were selected to be tested. In this assay, full length ERα and its fluorescent ligand Fluormone™ Tracer form a complex with high polarization value. Compounds capable of displacing the Fluormone™ Tracer cause the fluorescent ligand to tumble rapidly which can be observed as low polarization value. This shift in fluorescence polarization value can then be

used to determine the relative ERα affinity of the ligand candidates as IC_{50} values.

As described in the protocol, a dilution series of the ligands candidates were combined with the ERα and Fluormone™ Tracer complex on black low volume 384-well plates with NBS surface (Corning, Corning, NY, United States of America). The mixed plates were incubated for 2 hours in room temperature before measuring the fluorescence polarization using excitation / emission wavelengths 485 / 535 nm and bandwidths 25 / 20 nm. The 2104 EnVision® Multilabel Plate Reader with EnVision Workstation version 1.7 (PerkinElmer, Waltham, MA, United States of America) was used to record the data.

4.9.3 Reporter assay system (IV)

Human RORγ Reporter Assay System as 96-well Format Assay (INDIGO Biosciences, State College, PA, United States of America) was applied to quantify the activities of computationally predicted ligand candidates against human RORγ. The assay utilizes human reporter cells designed to express high levels of human RORγ hybrids of the both isoforms. The two isoforms are expressed from the RORC gene using tissue-specific promoters. The isoform 1, known as RORγ, *variant 1* mRNA is expressed in various tissues whereas the shorter *variant 2* mRNA of the isoform 2, known as RORγt, is predominantly expressed in immune cells developing in thymus. In the hybrids, the N-terminal DNA binding domain has been replaced with yeast GAL4 DNA binding domain and thus cannot convey any functional differences this domain might have between the native isoforms 1 and 2. In the assay, the high human RORγ expression has been combined with bio-luminescence reporter gene technology where cDNA encoding luciferase originating from North American firefly (*Photinus pyralis*) has been incorporated into the reporter cells. The assay utilizes luciferase detection agent where D-luciferin is oxidized in photon emitting reaction catalyzed by luciferase. Resulting luminescence intensity is then quantified on a luminometer and reported as relative light units (RLU).

Following the assay protocol, appropriate dilution series of the ligand candidates were prepared taking into account the DMSO concentration carried over to the assay reactions. The reporter cells were recovered, dispensed on 96-well collagen coated assay plates and combined with the ligand candidate dilution series. The cells were incubated at 37 ºC, in humidified 5 % CO_2 incubator for 24 hours before placing the luciferase detection reagents on the cells. The assay plates were left to rest for 5 min in room temperature and then the luminescence was quantified using multilabel reader (Victor™ X4, 2030 Multilabel Reader, PerkinElmer, Waltham, MA, USA).

4.10 Illustrations

Protein structures or protein-ligand interactions seen here have been generated using BODIL v. 0.81 (Lehtonen *et al.* 2004), MOLSCRIPT v. 2.1.2 (Kraulis 1991) and Raster3D (Merritt and Bacon 1997) or VMD 1.9.2 (Humphrey *et al.* 1996). The 2D representations of molecules and the reaction pathways seen here have been generated using Dassault Systèmes BIOVIA, BIOVIA Draw 2016, Version 16.1 .NET (32 bit), San Diego: Dassault Systèmes, 2017.

5 RESULTS

5.1 Coumarin scaffold as a lead compound (I, II, III, V)

Simple coumarin derivatives have attracted attention due to their biological activity and therapeutic potential. Essentially, coumarins are known to be effective as antitumor and anti-HIV agents, they are used as anticoagulants and they have shown promising activity as dopaminergic and serotonergic modulators in the central nervous system (Borges *et al.* 2005, Stefanachi *et al.* 2018). The appeal of the coumarin scaffold is further enhanced by the fact that coumarins occur readily in nature. Additionally, modifications of the basic core are easily obtainable and several synthetic pathways, to produce specific substitutions, are available to the organic chemists. In order to explore the potential of the coumarin derivatives as lead molecules for therapeutic applications, easy-to-synthesize database of coumarin derivatives was created. The database focuses mainly on novel 3-phenylcoumarin derivatives for their affordability but also for the interesting properties they entail. With simple substitutions to the coumarin scaffold, the electrostatic properties are altered to provide variation in the database. Furthermore, certain substitutions such as the hydroxylation of the 7-position of the coumarin scaffold produce fluorescent molecules (Rangaswami and Seshadri 1940, Rangaswami *et al.* 1941, Balaiah *et al.* 1942, Sherman and Robins 1968) which, when combined with metabolizing enzymes, provides a possibility to design specific tool molecules to study the enzyme functionality.

5.1.1 Novel 3-phenylcoumarin derivatives as MAO inhibitors (I)

Coumarin derivatives as MAO inhibitors have been studied extensively and promising lead compounds based on coumarin scaffold have been previously identified (Gnerre *et al.* 2000, Catto *et al.* 2006, Matos *et al.* 2009a, b, 2010, 2011b, 2013a, Serra *et al.* 2012, He *et al.* 2014, Mattsson *et al.* 2014). Therefore, database of previously unexamined molecules with related properties offered an interesting research direction. In this study, 52 novel 3-phenylcoumarin

derivatives (I, Figs. 2 and S4) were designed using virtual combinatorial chemistry or rationally *de novo* and their MAO activity was tested using continuous spectrophotometric assay. The activity measurements were complemented with docking-based structure-activity relationship (SAR) analysis, in order to determine the atom level factors contributing to the activity.

The 52 novel 3-phenylcoumarin derivatives were first tested at 10 μM concentration and the molecules exhibiting over 70 % MAO B inhibition (I, Table 1 and S1) were selected for the determination of their IC_{50} values. The initial analysis yielded 24 promising molecules (I, Table 1, Fig. 2) out of which all were determined to have IC_{50} values in nanomolar or low micromolar range as anticipated. More importantly, the results demonstrated that the 3-phenylcoumarin is indeed suitable scaffold for designing MAO B inhibitors. In fact, the best one of the tested compounds, **1** (IC_{50} 56 nM, Fig. 5, Table 3) was measured to have IC_{50} value comparable to known MAO B inhibitor pargyline (IC_{50} 61 nM) which was used as an inhibitor control in the experiments. Overall, more than half of the molecules (**1–6**, **8–10**, **15**, **17**, **20** and **21**) were measured to have IC_{50} values below 500 nM. The rest have IC_{50} values slightly below and above 1 μM but the worst molecule, **14** (IC_{50} 8476 nM) stands out as it was determined to have IC_{50} value just under 10 μM. As MAO A and B are known to have nonspecific inhibitors, examining possible cross reactivity is a routine practice. Due to apparent differences between the binding cavities of MAO A and B (Fig. 1C), the 52 3-phenylcoumarin derivatives were expected to be less suitable as MAO A inhibitors. Thus, the activity of the database in MAO A was determined the same way but the concentration of each tested molecule was 10 times higher at 100 μM. Despite of the increased concentration, the highest inhibition percentages reached nearly 60 % for molecule **43** and molecules **27**, **42** and **45** and **27** were close behind with over 40 % inhibition (I, Table S1). The 24 top MAO B inhibitors showed no or only minimal MAO A inhibition (I, Table 1) which indicates that the molecules are highly selective as MAO B inhibitors.

By scrutinizing the substitutions present in the 3-phenylcoumarin derivatives and combining them with the activity data, it is possible to recognize trends affecting the overall MAO activity. Among the 52 3-phenylcoumarin derivatives, other substitutions to the coumarin scaffold are limited to positions 6, 7 and 8 (I, R1–R3, Fig. 4). On the 3-phenyl ring, substitutions to *ortho-* , *meta-* and *para-*positions (I, R4–R7, Fig. 4) have been introduced. Based on the activity data, having other substitutions to the coumarin scaffold in addition to the 3-phenyl are not necessarily required in order to establish MAO B inhibition (I, see **11** in Table 1, Fig. 2). However, having acetoxy, hydroxyl, methyl, methoxy or a halogen group(s) introduced to the coumarin scaffold positions 6 to 8, may have a positive effect on the MAO B activity. In this regard, methoxy group on position 6 is particularly favorable (I, see **1** in Table 1, Fig. 2). Whether it is a certain position or a certain functional group that determines the MAO B activity ultimately depends on the substitutions on the 3-phenyl ring.

The substitutions on the 3-phenyl ring are actually so important that without them the MAO B activity is lost (I, see **41**, **50** and **51** in Table S1, Fig. S4). Trifluoromethyl at *para*-position of the 3-phenyl ring turned out to be particularly favorable as it can be found on the first and the second best molecules (I, see **1** IC$_{50}$ 56 nM and **2** IC$_{50}$ 138 nM in Table 1) and the third best molecule (I, see **3** in IC$_{50}$ 141 nM, Table 1) possessing structurally similar trifluoromethoxy is essentially as potent. The explanation for this, and other determining factors regarding MAO B activity, were sought after in the docking-based SAR analysis.

Available crystal structures of MAO B complexes with coumarin derived inhibitors (PDB 2V60 and 2V61 (Binda *et al.* 2007)) were selected as the foundation of the SAR analysis for the expected similarities in their occupation of the binding cavity compared to the 3-phenylcoumarin derivatives (I, Fig. 1A–D). One pivotal difference does exist between the previously crystallized coumarin derivatives and the novel 3-phenylcoumarin derivatives, however. Whereas the phenyl ring in the novel molecules is located on position 3, the crystallized molecules contain 3-chlorobenzyloxy substitution on position 7. As a result, when the novel molecules are docked to the X-ray crystal structures, the simulation suggests that coumarin scaffold assumes reverse orientation directing the substitutions on positions 3 to same direction in the binding cavity as the position 7 of the known inhibitors (I, Fig. 1C–D). This reversed orientation leads into interesting outcome regarding the carbonyl oxygen on position 2 according to the docking. Instead of facing the solvent, the carbonyl oxygen is directed toward C172 side chain (Fig. 5). The interaction between the carbonyl and the proton of the thiol group is not as strong as typical hydrogen bond based on electronegativity but this usually weak interaction is likely enhanced in otherwise hydrophobic environment. Although the docking also suggests slight variability in the overall positioning of the 3-phenylcoumarin derivatives due to their substituents (I, Fig. 3D), the most important determining factors for the binding orientation are the hydrophobic interactions elicited by the coumarin scaffold and the 3-phenyl ring which correspond with the surrounding binding cavity (I, Fig. 3A–B). In fact, the positioning of the coumarin scaffold is to some extent analogous even with the MAO B inhibitor isatin in the PDB structure 1OJA (Binda *et al.* 2003), implying that certain positioning of fused ring structures is favored in MAO B. On the other hand, MAO A binding cavity is not as good fit for the 3-phenylcoumarin derivatives due to two residue differences I199→F208 (Fig. 1C), which has been established as a contributor to the enzyme selectivity (Son *et al.* 2008), and L164→F173 (I, Fig. 3C) explaining the observed low aptitude for the enzyme.

According to the general positioning of the 3-phenyl ring, this end of the derivatives is directed at mostly hydrophobic amino acid residues. The tightly fitting enclosure seems to be especially suitable for halogenated substituents, for example for the trifluoromethyl on the *para*-position of the 3-phenyl ring on compound **1** (Fig. 5). Although obviously electronegative halogen substituents might seem like an unlikely fit for the hydrophobic environment, they actually improve the steric packing thanks to resilient vdW interactions while they

maintain the ability to form halogens bonds. Halogen induced increase in lipophilicity perceived as higher logP values (I, Table 1) might additionally support the binding of the halogen-containing derivatives as they proceed to the ligand binding site facing the mitochondrial outer membrane (I, Fig. 1A). Halogens, chlorine and fluorine, are also present in two of the coumarin derivatives in the MAO B X-ray crystal structures PDB 2V60 and 2V61 (Binda *et al.* 2007) where they can be seen to form halogen bonds with the L164 main chain oxygen (I, Fig. 1B). Consequently, the 3-phenylcoumarin derivatives with halogen substitutions on the equivalent *meta*-position of the 3-phenyl ring are similarly capable of substantial MAO B inhibition (I, Fig. 6E–F). The benefits of halogen substituents on 3-phenyl ring are further highlighted by the fact that derivative **15**, with three separate fluorines on *ortho*-, *meta*- and *para*-positions, is at IC_{50} value 292 nM (I, Fig. 6A, Table 1) more active than otherwise similar but only single fluorine on *para*-position containing **21** at IC_{50} value 433 nM (I, Fig. 5D, Table 1).
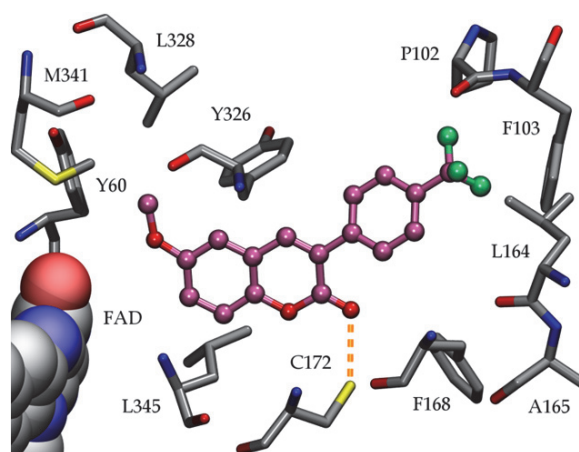
FIGURE 5   The docked position of derivative **1** (I) (in pink) in MAO B (PDB 2V60 (Binda *et al.* 2007)). The rotamer of the methoxy group on position 6 has been adjusted in order to highlight how the methoxy group utilizes the hydrophobic groove aligned with Y60 and L328. Another favorable interaction is the hydrogen bond (in orange) between the carbonyl oxygen on position 2 of the coumarin core and C172. In addition, the trifluoromethoxy on *para*-position of the 3-phenyl ring fits the surrounding hydrophobic environment due to improved steric packing owing to persistent vdW interactions.

Apart from halogen containing derivatives, other promising 3-phenyl ring substituents are methoxy group on *meta*-position and dimethylamine on *para*-position. Looking at derivative **8** which is the most active of the methoxy group on 3-phenyl ring containing derivatives at IC_{50} value 231 nM, the docking placed the methoxy oxygen toward L164 main chain oxygen which might seem like energetically unfavorable position (I, Figs. 2 and 5E, Table 1). However, the undesirable consequences are likely avoided due to intraprotein hydrogen bond within the main chain. In fact, the suitability of methoxy group on this position

is further confirmed by almost as active derivative **9** (I, Figs. 2 and 5F, Table 1) which has IC_{50} value 255 nM and methoxy groups on *meta*-positions on both sides of the 3-phenyl ring. The dimethylamine group on *para*-position of the 3-phenyl ring has similar hydrophobic properties as the trifluoromethyl group on the same position. However, this group is larger in size and lacks the ability to form halogen bonds which is why switching dimethyl amine to the *para*-position of the derivative **1** produces less active inhibitor seen in derivative **18** (I, Figs. and 6D, Table 1) which has IC_{50} value of 617 nM.

The substituents on the coumarin scaffold refine the fit of the coumarin derivatives. In general, the fit of a derivative improves if it contains a substituent capable of both hydrophobic and hydrophilic interactions on positions 6 or 7 of the coumarin scaffold. This kind of benefit becomes evident by comparing derivative **11** which does not have any substitutions on coumarin scaffold to otherwise similar derivative **8**. Derivative **11** is significantly less active with IC_{50} value of 798 nM (I, Figs. 2 and S3F, Table 1) than derivative **8** with IC_{50} value of 231 nM (I, Figs. 2 and 5E, Table 1) only because **8** a has methoxy group on position 6. Having the methoxy group on position 6 of the coumarin scaffold is particularly favorable since the hydrophobic methyl can be packed toward hydrophobic region formed by amino acid residues Y60, Q206, Y326, L328, F343, and M341 while the oxygen makes the derivative as a whole more hydrophilic for the benefit of the fit next to the cofactor and neighboring water molecules. Switching the methoxy group to position 7 makes the alignment of the coumarin derivative more challenging but the overall fit is the combination of the substitutions available on both ends of the derivative.

To conclude, 52 novel 3-phenylcoumarin derivatives were synthetized tested for MAO B activity using spectrophotometry (I, Tables 1 and S1). The most promising derivative **1** had IC_{50} value 56 nM and in total 20 derivatives had IC_{50} values below 1 µM. The activity measurements were complemented with a thorough SAR analysis revealing that precise positioning of particularly favorable substitution may have fundamental benefits for the overall fit of a derivative. Thus, 3-phenylcoumarin is a promising scaffold for building MAO B inhibitors as the properties of the molecule are readily modifiable through various substitutions.

5.1.2 Coumarin derivatives resembling 17β-estradiol as ERα ligands (II)

Coumarin based molecules have been previously used in the lead identification of selective ERα modulators (McKie *et al.* 2004). The properties of 4-hydroxytamoxifen, the metabolite of the first breast cancer drug tamoxifen and raloxifene, a compound with both estrogen agonist and antagonist properties, have contributed to the design. As a result, beneficial manipulations to the coumarin scaffold in order to produce ERα selectivity have been recognized (McKie *et al.* 2004). Since coumarins have been established as ERα modulators, improving the fit of the coumarin scaffold might contribute to the overall usability of this class of compounds.

The structural requirements for a molecule to establish ERα binding are well-known. Basically, the molecule should mimic 17β-estradiol by being capable of similar interactions as the hydroxyl groups on positions 3 and 17 (Fig. 6). In addition, hydrophobicity and ring structures are beneficial features to have on an ER ligand candidate (Fang *et al.* 2001). That in mind, a small in-house database of coumarin derivatives similar in size with 17β-estradiol was created (comparison between estrone and 3-phenylcoumarin on Fig. 7). The molecules were analyzed using virtual screening methods, identifying promising hits with docking and NIB method. Out of these, the molecules predicted to be active were visually inspected and the ones passing the inspection by matching the requirements of an ER ligand were synthesized. This produced 5 molecules with pIC_{50} values between 5.5 and 6.5 (II, Table 2) according to *in vitro* analysis. The top two molecules **3** and **5** can be found on Table 3.
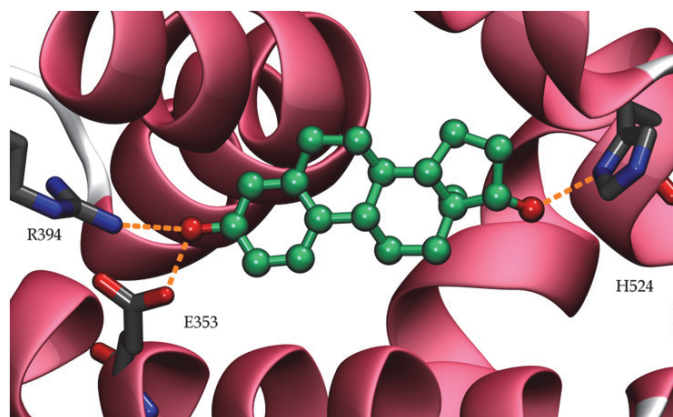
FIGURE 6   Binding of 17β-estradiol (in green) to ERα (PDB 1ERE (Brzozowski *et al.* 1997)). The most relevant interactions between 17β-estradiol and ERα are the hydrogen bonds between the hydroxyl groups on positions 3 and 17 and amino acid residues E353, R394 and H524 (in grey). Thus, compounds capable of similar interactions might be able to demonstrate ERα activity.

Based on the correlation between the predicted and the experimental ERα activity (II, Fig. 5), designing molecules on the grounds of prior understanding of the preferable ER ligand properties proved to be beneficial. Thus, in certain situation thorough understanding of the protein binding cavity and the structural requirements for a ligand may be the most valuable tool for producing novel hits.

5.1.3 Building new inhibitors to block estradiol synthesis pathways using 3-phenylcoumarin derivatives (III)

As with ERα, HSD1 prefers molecules which resemble its natural substrate estrone. Since estrone is the precursor of 17β-estradiol, coumarin derivatives obviously comply with these characteristics as well (Fig. 7). Consequently, 6-

52

and 7-phenylcoumarin derivatives have been previously investigated as HSD1 inhibitors with promising results (Starčević *et al.* 2011). Thus, it could be reasoned that 3-phenylcoumarins with their phenyl group on the opposite end of the coumarin scaffold are capable of corresponding functionality. Furthermore, the 3-phenylcoumarin ring system would likely mimic the hydrophobic packing required for the substrate binding which is also seen in the binding of a known steroid derived inhibitor in PDB structure 3HB5 (Mazumdar *et al.* 2009) (III, Fig. 1D). With these factors in mind, 3-phenylcoumarin scaffold was selected as a starting point for the design of novel non-steroidal HSD1 specific inhibitors.
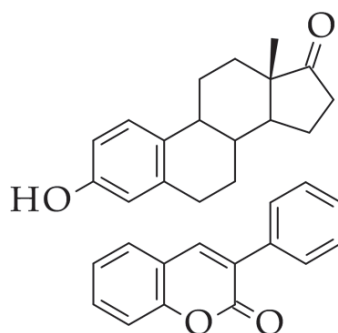
FIGURE 7   Estrone (on top) and 3-phenylcoumarin. The side by side 2D image highlights the fact that 3-phenylcoumarin resembles estrone in size and the similarities of the overall ring systems are also indisputable.

In order to shed light on the most promising substitution options, the HSD1 substrate binding cavity was examined carefully. Upon inspection, the amino acid sidechains of Y219 and S223 and their hydroxyl groups were considered to reside in optimal position for forming hydrogen bonds with the carbonyl oxygen on position 2 of the coumarin scaffold. Alternatively, if facing another direction, the carbonyl oxygen might also be able to form hydrogen bond with R268. However, this would require rotamer adjustments that are not available on the relevant X-ray crystal structure data (PDB 1EQU (Sawicki *et al.* 1999)). To complement this theory, polar moieties were introduced to the positions 6, 7 and 8 of the coumarin scaffold (III, R1–R3, Fig. 2) and/or to the *ortho-* , *meta-* and *para*-positions of the 3-phenyl ring (III, R4–R6, Fig. 2) for obtaining HSD1 inhibition. Eventually, total of nine derivatives were selected for organic synthesis and experimental testing.

According to the activity measurements (III, Table 1), the hypothesis of 3-phenylcoumarin derivatives being able to inhibit HSD1 proved to be correct. The most promising inhibitor **4** produced 47 % inhibition at 100 nM concentration (Table 3). In addition to **4**, **1** and **2** produced over 68 % inhibition at 1 µM and at 5 µM concentration, also **3** and **7** reached at least 62 % inhibition. Out of the tested 3-phenylcoumarin derivatives, **5**, **6**, **8**, **9** and **10** did not inhibit HSD1 or inhibited it only weakly.

If the hydrophobic packing is the only considered factor, the 3-phenylcoumarin derivatives could in theory mimic the steroid position in four different ways. Thus, solely based on this information, it is not possible to decipher the underlying structural basis of the observed HSD1 inhibition. For this reason, the in-house developed NIB Panther protocol (Virtanen and Pentikäinen 2010, Niinivehmas *et al.* 2011, 2015) was utilized in the docking. With Panther, it was predicted how the synthesized 3-phenylcoumarin derivatives accomplish the HSD1 inhibition. As a result, this type of docking-based SAR analysis (III, Fig. 5) was able to explain how the employed substitutions (III, Fig. 2) affect the binding and the consequential inhibitory properties of the 3-phenylcoumarin derivatives (III, Figs. 3 and 4).

According to the docking-based SAR analysis, derivatives **2**, **3**, **4**, **5** and **7** are believed to assume convergent position of the coumarin scaffold (III, Fig. 3). In this position, Y219 and S223 form hydrogen bonds with the carbonyl oxygen as hypothesized. The most potent derivative **4** is additionally firmly coordinated through its substituents. Essentially, the hydroxyl group on *meta*-position on the 3-phenyl of this derivative is able to donate hydrogen bond to E283 while the chlorine on position 6 is capable of forming halogen bonds with S143 and perhaps with Y156 (III, Fig. 3E). In contrast to **2** (III, Fig. 3A) which is able to hydrogen bond with both E283 and H222 through its hydroxyl group on *para*-position, the difference in activity must be explained through the halogen bonding of **4**. The chlorine's negative charge is further accommodated by the protons of V144 and G145 main chain nitrogen which as a whole, turns out to be more favorable substitution than hydroxyl group forming hydrogen bonds only with S143 and Y156. Derivative **5** forms the same interactions on the 3-phenyl substitution as derivative **2** does. The difference between the two molecules lies on the coumarin scaffold where **5** has the hydroxyl group on position 7 instead of position 6 (III, Fig. 3B). Here the hydroxyl group is able to form hydrogen bond only with Y156, drastically weakening the overall activity of the molecule (III, Table 1). Derivative **3** (III, Fig. 3F) is otherwise the same as **4** but it lacks the chlorine on position 6 which explains the decrease in activity (III, Table 1). Derivative **7** (III, Fig. 3C) should be compared to derivative **6** (III, Fig. 3D) since the two have the same substituents present on the 3-phenyl. Hydroxyl group on *para*-position forms the same hydrogen bonds with E283 and H222 as **2** and **5** and fluorine on *meta*-position is additionally able to interact with H222 through a halogen bond. Again, the substituents on the coumarin scaffold are in a key role when it comes to the overall activity of the derivative in question, highlighting the importance of a favorable substitution at this location. Here, derivative **7** (III, Fig. 3C) has a methoxy group on position 6 while the methoxy on derivative **6** (III, Fig. 3D) is located on position 7. On derivative **7** the methoxy group is able to form more coordinated polar interactions with nearby amino acids than on derivative **6** accounting for the activity differences (III, Table 1).

The role of the coumarin scaffold substitutions are further highlighted with the derivative **1** which is expected to assume alternative binding mode for the coumarin scaffold (III, Fig. 4). In this alternative mode, derivative **1** is able to

adopt more favorable position due to its hydroxyl group substitution on position 8. In this position, the carbonyl oxygen of the coumarin scaffold can only form one hydrogen bond but the heterocyclic oxygen compensates for this by partaking in the interactions with Y219. In addition, the hydroxyl on position 8 is capable of accepting a hydrogen bond from H222 and the methoxy on *para*-position of the 3-phenyl is ideally positioned for accepting a hydrogen bond from S143 (III, Fig. 4B).

Following the same principle, the aromatase active site was also identified as a potential target of the hydrophobic packing of 3-substituted coumarin derivatives. However, for the binding to occur, the coumarin derivative should be able to mimic the binding of the natural aromatase substrate androstenedione and accept a hydrogen bond from the neutral side chain of D309 (III, Fig. 6C). The presented 3-phenylcoumarin derivatives cannot comply with this demand but by replacing the 3-phenyl group with imidazole, binding can be achieved as also proven experimentally (III, Fig. 4C, Table 1).

Accordingly, 3-phenylcoumarin is a promising scaffold for building HSD1 inhibitors as it is able to mimic the hydrophobic packing typical for the known active molecules. The same idea can be applied to the design of the aromatase inhibitor where replacing the 3-phenyl with imidazole produces similar results. Thus, both sulfatase pathway and aromatase pathway of the estradiol synthesis can be blocked by 3-substituted coumarin derivatives and the binding can be further adjusted by designing case specific substitutions for the promising candidates presented here.

5.1.4 Selective 7-hydroxycoumarin derivatives as UGT1A10 substrates (V)

The 7-hydroxycoumarin makes an interesting scaffold for designing novel, selective fluorescent substrates for UGTs because the hydroxyl group on position 7 of the coumarin scaffold creates a fluorescent molecule (Rangaswami and Seshadri 1940, Rangaswami *et al.* 1941, Balaiah *et al.* 1942, Sherman and Robins 1968). Other substituents on positions such as 3 or 4 do not abolish the fluorescence but instead modify its intensity. The 7-hydroxyl group of coumarin scaffold is also glucuronidated by many UGTs which has already been utilized in creation of quantitative multi-well plate assay for measuring glucuronidation rate of 7-hydroxy-4-trifluoromethylcoumarin in order to determine UGT activity (Rahikainen *et al.* 2013).

Using a molecular docking-based design method (see chapter 5.3), fluorescent 7-hydroxycoumarin derivatives were created *in silico* for all human UGT enzymes of 1A subfamily. Out of the designed substrates, six (V, Fig. 3) were synthesized and tested for their glucuronidation rates. Initial analysis was performed with pig microsomes containing a selection of UGTs. The observed decrease of fluorescence was linearly dependent on the amount of pig microsomes which indicated that the glucuronidation of 7-hydroxycoumarin derivatives was catalyzed by one or more of the UGTs present in the microsomes (V, Fig. 4). Through these promising results, it was implied that the

glucuronidation rates of the new 7-hydroxycoumarin derivatives could be accurately determined using specific UGT enzymes.

Upon analysis of wide panel of human UGTs (V, 1A1, 1A3, 1A4, 1A6, 1A7, 1A8, 1A9 and 1A10 from 1A family as well as 2A1, 2A2, 2A3, 2B4, 2B7, 2B10, 2B15 and 2B17 from 2A and 2B families, Fig. 5), it was noticed that all of the six substrates were glucuronidated by UGT1A10 the fastest and out of these substrates **2**, **4**, **5** and **6** exhibited also specificity for the enzyme. The substrates **1** and **3** were glucuronidated by UGT1A1 in addition to UGT1A10 but at a lower rate. Furthermore, the newly detected glucuronidation rates were significantly different from the control substrates 7-hydroxycoumarin which was mainly glucuronidated by UGT1A6 and 7-hydroxy-4-trifluoromethylcoumarin which was glucuronidated primarily by UGT1A6 and UGT1A10 but also by other UGTs at a lower rate (V, Fig. 5). Since UGT1A10 is an extrahepatic enzyme with high expression levels in intestine (Sato *et al.* 2014), the glucuronidation rates of the 7-hydroxycoumarin derivatives were additionally analyzed using human liver and intestinal microsomes. Naturally, both of the microsomal preparations contain multiple UGT enzymes, however, human liver microsomes are devoid of functional UGT1A10 (Sato *et al.* 2014). Consequently, the glucuronidation rates of substrates **2**, **4** and **6** in human liver microsomes were very low. Substrates **1**, **3** and **5** performed better but did not reach the glucuronidation rates observed in human intestinal microsomes (V, Fig. 6). Thus, substrates **2**, **4** and **6** would be the most useful for studying UGT1A10 glucuronidation activity in samples containing the enzyme.

As result of modeling based design method, novel UGT1A10 specific 7-hydroxycoumarin derivatives were created. The accomplished specificity was determined to be dependent on the substitutions on the position 3 of the coumarin scaffold. Out of the designed substitutions, 4-dimethylaminophenyl of **4** and triazole of **6** (Table 3) appeared to be especially useful in obtaining UGT1A10 selectivity. These new substrates improve the determination of UGT1A10 activity and offer a starting point for the design of other UGT-specific tool molecules.

5.1.5 Selectivity of the coumarin derivatives (I, II, III, V)

For a compound to be valuable as a potential drug, it should usually be active only in its designed target. Since coumarin derivatives are known to already have therapeutic activities in wide range of proteins (Sandhu *et al.* 2014, Stefanachi *et al.* 2018), it has been of utmost importance to look into the selectivity of the coumarin derivatives presented here also at this stage of the development process.

Monoamine oxidase A and B. The most concerning side effect of MAO inhibition is the accumulation of dietary tyramine which is capable of displacing monoamine deposits causing potentially fatal hypertensive crisis. Since MAO A is the predominant MAO in the gastrointestinal tract (Hasan *et al.* 1988) responsible for the deamination of tyramine, it is important to avoid MAO A inhibition in the design of MAO inhibitors. Notably, most of the tested

3-phenylcoumarin derivatives do not inhibit MAO A at 10 times higher concentration than the concentration used to test MAO B inhibition. Even when they do, the inhibition percent remains below 60 % (I, Table S1). More importantly, the top MAO B inhibitors do not inhibit MAO A (I, Table 1). Based on the MAO B selective 3-phenylcoumarin derivatives such as **1**, the underlying reason for the lack of MAO A activity becomes evident. The ligand binding sites of the two MAOs (I, Fig. 3A and 3C) are critically different in size which is why 3-phenylcoumarin derivatives are able to obtain more favorable interactions in MAO B. Thus, 3-phenylcoumarin derivatives are a promising lead to be used as the foundation of designing MAO B selective inhibitors.
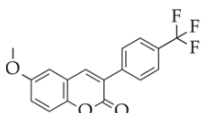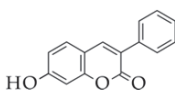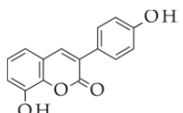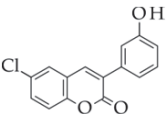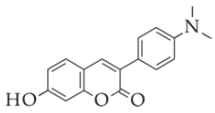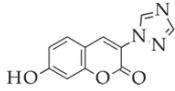
Estrogen receptor a. ERα agonists and antagonists or selective modulators are used for example as infertility treatments, contraception, hormone replacements and ERα positive breast cancer therapies (Jordan 2003a, b). Since ERα is activated by the hormone 17β-estradiol and is responsible for the regulation of vast number of genes, cross reactivity with drugs targeted elsewhere could have detrimental effects. Should a compound function as ERα agonist in addition to its intended use, the compound could potentially end up promoting tumor growth in the breast tissue. Other obvious side effects of agonist or antagonist off-target binding originate from the interference with hormone balance which could at worst range from changes in the secondary sex characteristics to osteoporosis or interfere with already ongoing ER-targeted therapies. As previously established, 3-phenylcoumarin derivatives capable of similar interactions as 17β-estradiol with hydroxyl groups on positions 3 and 17, are potentially active in ERα. Out of the 52 potential MAO B inhibitors (I, Tables 1 and S1) derivatives **12**, **17**, **20**, **22**, **27–30**, **32**, **39–41**, **44**, **47** and **48** would fill this condition and are in fact capable of inhibiting ERα at 10 μM concentration (I, Tables 1 and S1, Figs. 2 and S4). Derivatives **17**, **32** and **47** were unfortunately not included into the analysis due to shortage of materials but based on the obvious trend, they are expected to obtain similar inhibition percentages as the other listed derivatives. Notably, the substitutions suitable for ERα inhibition are not the ones also responsible for the most active MAO B inhibitors (Table 3) and thus, supports the concept that selectivity can be reached by altering the substituents in a case-specific manner.

17β-hydroxysteroid dehydrogenase 1 and 2. HSD1 and its catalytic counterpart HSD2 function on the opposite sides of 17β-estradiol biosynthesis. While HSD1 catalyzes the final reducing step, HSD2 converts 17β-estradiol back into estrone (Fig. 3). The 17β-estradiol producing HSD1 appears to be the predominant type in ectopic endometrium and breast cancer tissue (Vihko *et al.* 2004, Dassen *et al.* 2007, Hanamura *et al.* 2014) which is why inhibiting HSD1 provides an attractive option in treatment of these conditions. Since the point of HSD1 inhibition is to lower the 17β-estradiol production, it is pivotal not to hamper these efforts by simultaneously blocking HSD2 activity as well. Since both HSD1 and ERα are linked to the estradiol synthesis pathway and, in general, prefer similar compounds, they would also be expected to share activity among the 3-phenylcoumarin derivatives. The best example of this among the 52 potential MAO inhibitors is derivative **48** (I, Fig. S4, Table S1; **2** III, Fig. 2, Table

1) which is definitely cross reactive between HSD1 and ERα. The activity is easily explainable due to the ideally positioned hydroxyl groups on *para*-position of the phenyl ring and on position 6 of the coumarin scaffold, mimicking the natural substrates. However, this derivative is not active in HSD2 or either of the MAOs. Essentially, the most active HSD1 inhibitor **4** does not exhibit HSD2 activity (III, Fig. 2, Table 1). Unfortunately activity data regarding this derivative does not exist for ERα or the MAOs. Based on the structure, however, it can be determined that it would lack the features required for optimal binding in all of the three proteins (Table 3). Aromatase could be another potential target for the 3-phenylcoumarin derivatives mimicking steric steroidal positioning and known binding activity at the active site. Due to the particular requirements of the aromatase binding, placing an imidazole on the position 3 of the coumarin core instead of phenyl prove to be able to provide selectivity between aromatase and HSD1 (III, Fig. 2, Table 1).

Through careful design and analysis, 3-phenylcoumarin derivatives have proven to be versatile group of compounds capable of producing selectivity even within highly similar targets sharing binding activities. Accordingly, the presented details required for selectivity in each of the above mentioned targets contribute to the overall knowledge and aid in future efforts of cultivating 3-phenylcoumarin derivatives toward therapeutic applications.

TABLE 3   The most potent 3-phenylcoumarin derivatives from each original publication and their activity in the discussed targets if available.

Compound	MAO B inhibition 10 μM	MAO A inhibition 100 μM	HSD1 inhibition 1μM	ERα binding pIC_{50}	UGT
I **1**	100 %	0 %	-	-	-
II **3**	24 %	10 %	-	6.5	-
II **5**	23 %	22 %	-	6.5	-
III **4**	64 %	-	84 %	-	-
V **4**	90 %	5 %	-	-	1A10
V **6**	-	-	-	-	1A10

5.2 Structure-based virtual screening of RORγt inverse agonists (IV)

As with other nuclear receptors, the ligand binding site of RORγt is very lipophilic which can be also noticed from the high logP values of the active ligands (IV, Fig. 1E). Despite of this, the largely lipophilic ligand binding site contains certain polar residues such as R367 which are in a key role in the binding of the most potent ligands as seen for example in PDB structure 4WLB (IV, Fig. 1A-C, (Van Niel *et al.* 2014)). Analyzing the properties of known inverse agonists aid in recognizing the factors that make a promising lead compound in the virtual screening process.

Two docking protocols were selected for the structure-based virtual screening, GLIDE standard precision and NIB Panther/ShaEP. Active ligands from ChEMBL (Bento *et al.* 2014), 199 in total, were used to validate the selected docking protocols. GLIDE standard precision did not provide high correlation ($R^2 = 0.23$) between the pIC_{50} values of the experimental data and the GLIDE docking energy (IV, Fig. 2A). This can be expected as the experimental data originates from various sources and thus utilizes heterogeneous protocols. Interestingly, many of the ligands with pIC_{50} over 7.5 can be distinguished from the ones with lower activity (IV, Fig. 2A). This can be demonstrated by setting a GLIDE docking energy cut-off value at -11.6 kcal mol^{-1} when 61 % of the these highly active ligands are identified simultaneously leaving out all the molecules with pIC_{50} below 7.5. However, at docking energy -11.4 kcal mol^{-1} enrichment of over 80 is achieved (IV, Fig. 2B). Thus this was decided to be the value used in virtual screening of Specs database in order to identify RORγt high affinity ligands and to include molecules based on their interactions at the binding site. With these configurations, GLIDE docking produced 46 hits out of which 24 also formed at least one hydrogen bond (IV, Table S1).

NIB Panther protocol was used as an alternative screening method since the correlation coefficient achieved by GLIDE was unsatisfactory. Panther was combined with ShaEP scoring. However, the correlation of the pIC_{50} values of the 199 active ligands and the ShaEP similarity of the Panther/ShaEP combination was not much better ($R^2 = 0.27$) (IV, Fig. 2C). Also here, the most active molecules can be visually distinguished based on their docking score defined this time as ShaEP similarity (IV, Fig. 2C) and this was used to guide the decision of the cut-off value. Yet the enrichment was obviously weaker than with GLIDE (IV, Fig. 2D compared to Fig. 2B) and thus the value 0.6 was selected to be used in the virtual screening. NIB Panther/ShaEP screening yielded 5 424 molecules but the number was narrowed down to 123 by introducing a pharmacophore point expecting polar interactions with R367 (Fig. 8). If the molecules overlapped significantly with the protein's main chain, they were excluded, which left 34 molecules remaining. Out of these identified molecules 22 were the same as identified by GLIDE docking (IV, Table S1).

The identified hits were selected for experimental testing where the best molecule **9** (Fig. 8; IV, Fig. 3C) was determined to have IC$_{50}$ value of 587 nM. Overall, 11 molecules were found to have pIC$_{50}$ value 5.0 or above (IV, Table S1, Fig. S5). Out of these 11 molecules, nine were identified by both GLIDE and Panther/ShaEP screening which is encouraging because NIB is remarkably faster than docking. Although the computational models overestimated the tested compounds (IV, Table S1), four new structural cores (IV, Figs. 3 and S2) were discovered to display RORγt inverse agonism.
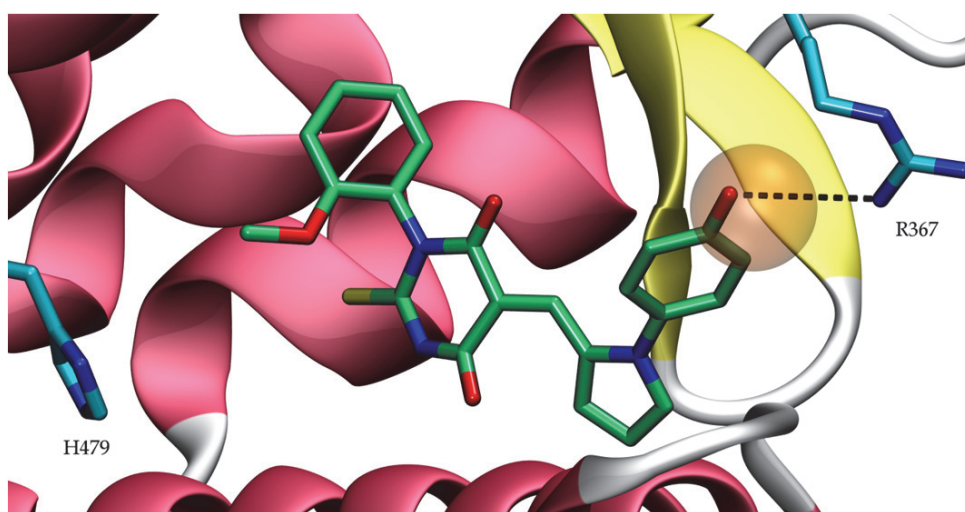
FIGURE 8   The docking placed the most promising molecule **9** (IV) (in green) in the vicinity of R367 instead of H479 (in blue, PDB 4WLB (Van Niel *et al.* 2014)) that has been identified to have critical role in RORγt agonist activity (Kurebayashi *et al.* 2004). This conformation fulfilled the pharmacophore point (in orange) set to expect polar interactions with R367.

5.3 Homology modeling of UGT1A enzymes for designing selective 7-hydroxycoumarin based substrates (V)

Currently, X-ray crystal structure of a mammalian UGT N-terminal domain does not exist. As this domain is highly variable among UGT enzymes, crystal structure of the C-terminal domain does not offer enough details for model building for the substrate binding since the N-terminal domain is heavily involved. Modeling efforts, which include both quantitative structure-activity relationship models and homology models, in order to further understand the functionality of these enzymes, have been previously done with varying success (Dong *et al.* 2012, Raunio *et al.* 2015, Tripathi *et al.* 2016). However, more detailed information is required for an accurate atomistic level analysis of the interactions between the substrate and the enzyme.

BLAST (Altschul *et al.* 1990) search against PDB using the sequences of the human UGT1A family (see section 4.8.1) produced four templates with moderate sequence similarity to be used in the model building. Out of these, 2O6L (Miley *et al.* 2007) is the only mammalian UGT crystal structure and it contains the C-terminal domain. As the UDP-glucuronic acid binds to this domain, it has highly conserved sequence among other UGTs. Rest of the templates 3HBF (Modolo *et al.* 2009), 3WC4 (Hiromoto *et al.* 2013) and 2C1Z/X (Offen *et al.* 2006) contained both the sugar-nucleotide cofactor and a small aglycone molecule which were considered to be beneficial for the model and thus fulfilled the structural requirements. The selected structures were used to build models for the human UGT1A family using structure-based sequence alignment method where 2C1Z was used as the template for its complete C- and N-termini containing structure and sequence similarity with UGT1A10. The sequences of the human UGT1A family were then aligned with the structural alignment of the selected structures. Since the main focus of the model building was on the glucuronidation of 7-hydroxycoumarin derivatives, conserved lipophilic region that spans the endoplasmic reticulum and ends in so called cytoplasmic tail was excluded from the homology model. In addition, structures such as loops and amino acid side chains outside the catalytic site were not optimized.

The created models were used to design selective 7-hydroxycoumarin substrates with decrease in fluorescence as a result of glucuronidation. At first, a collection of 7-hydroxycoumarin scaffolds from the easy-to-synthesize database were docked into the UGT models, providing several possible conformations. Out of these conformations, one placed the 7-hydroxy group toward the catalytic site, in the vicinity of the UDP-glucuronic acid (V, Fig. 2A) and found an optimal cavity nearby for the coumarin scaffold (V, Fig. 2B). At this position, the coumarin scaffold could be stabilized (Fig. 9) by accepting a hydrogen bond from Q101 in UGT1A8, UGT1A9 and UGT1A10 on the carbonyl oxygen on position 2 (V, Fig. 2A), according to the created models. With R102 in UGT1A6, this position could form even stronger stabilizing interactions. On the contrary, similar stabilizing interactions are not possible with D103 in UGT1A1, E104 in UGT1A3 and UGT1A4, or P101 in UGT1A7, making the coumarin scaffold an unlikely substrate for these enzymes. In addition, Q103 in UGT1A10 might donate a hydrogen bond to the oxygen on position 1 of the coumarin scaffold while corresponding R103 in UGT1A7, UGT1A8 and UGT1A9 blocks the binding to some extent due to its size (V, Fig. 2A).
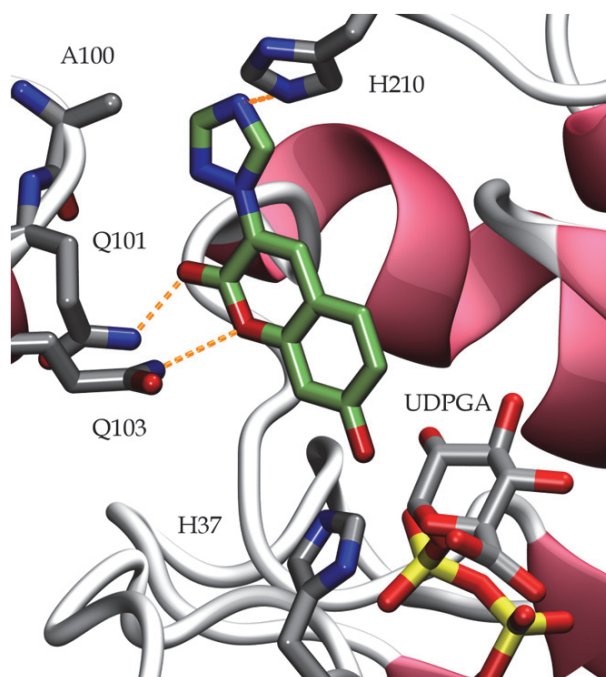
FIGURE 9   Docked position of 7-hydroxycoumarin derivative **6** with triazole (V) on position 3 (in green) in UGT1A10 model. Based on the model, interactions (in orange) stabilizing the substrate could be formed through amino acid residues Q101, Q103 and H210 (in grey), supporting the glucuronidation of the position 7 facing H37 and UDPGA.

According to the homology models, the 7-hydroxycoumarin docking site analysis revealed an additional space next to the position 3 of the coumarin scaffold. In UGT1A1 and UGT1A10 this space was thought to be large enough for five- or six-membered ring substituents (V, Fig. 2B) but in UGT1A6, F212 and E101 visibly obstruct the binding of 3-substituted 7-hydroxycoumarins. Other differences surrounding this additional space included H210 and A100 in UGT1A10 (Fig. 9) whereas UGT1A1 has hydrophobic M213 and polar N102 at the corresponding locations. Thus, this could indicate that UGT1A1 encounters problems while trying to accommodate large substituents at the 3-position of the coumarin scaffold. The modeled cavity was explored by docking the 7-hydroxycoumarin derivatives to the UGT1A10 model with the restriction of 7-hydroxyl group facing the catalytic site. Based on the docking, the most promising 7-hydroxycoumarin derivatives contained these substituents on position 3: **1** 4-fluorophenyl, **2** 4-hydroxyphenyl, **3** 4-methoxyphenyl, **4** 4-dimethylaminophenyl, **5** 4-methylphenyl or **6** triazole. From these 6 molecules, 7-hydroxy-3-triazolecoumarin is especially interesting since it has the possibility to stabilize through H210 in UGT1A10 but M213 in UGT1A1 does not have such capability. The role of these amino acids was tested by creating a new UGT1A10 H210M mutant. However, the results were not as straight forward as expected as the outcome of the glucuronidation kinetics (V, Table 1)

turned out to be capricious. The glucuronidation rate decreased in all cases (V, Fig. 7) but the V_{max} increased in case on **2** and **6**.

Despite of the lack of full mammalian UGT X-ray crystal structure, the building of homology models of the UGT1A family was successful in the design of selective substrates. The created models were utilized in the design of selective 7-hydroxycoumarin derivatives which could serve as fluorescent tool molecules in UGT activity studies. As a result, two 7-hydroxycoumarin derivatives with **4** 4-dimethylaminophenyl and **6** triazole substitutions on position 3 with UGT1A10 selectivity were created.

6 DISCUSSION

As naturally occurring compounds with known therapeutic applications (Table 1), coumarin derivatives are uniquely fascinating as lead molecules. Here an easy-to-synthetize database of coumarin derivatives has been explored as a potential source of novel coumarin-based lead molecules. The novel compounds have been targeted at versatile group of known druggable proteins with pharmaceutical applications ranging from neurological disorders to cancer and autoimmune conditions. The metabolism of drugs is also touched upon as coumarin derivatives have been investigated as tool molecules in order to further understanding of the renal elimination of xenobiotics.

6.1 Pharmacological potential of 3-phenylcoumarin derivatives

Originally, 3-phenylcoumarin derivatives were created as MAO B inhibitors because of the biological activity of *trans*-resveratrol (3,5,4'-trihydroxy-*trans*-stilbene) for example as a vasodilator (Vilar *et al.* 2006). Since then, coumarin derivatives and especially 3-phenylcoumarin derivatives have attracted a great deal of attention and various analogs have been found active in MAO B (Matos *et al.* 2009a, b, 2010, 2011a, b, 2013a, b, Santana *et al.* 2010, Viña *et al.* 2012b, Serra *et al.* 2012, Viña *et al.* 2012a) and some also in MAO A (Mattsson *et al.* 2014). Thus, 3-phenylcoumarin scaffold offers a unique opportunity to thoroughly explore the ligand binding cavities of MAOs. Despite of the extensive work already contributed to coumarin derivatives as MAO B inhibitors, the coumarin derivatives in the easy-to-synthesize database are novel in this regard. In general, the novel coumarin derivatives utilize similar substitution as the previously examined compounds but in alternative combinations. Review by Matos *et al.* (2013b) lists the substitutions used in the previously published compounds (Matos *et al.* 2013b). Because of that, it has been possible to approximate the minimum requirements in order to obtain MAO B inhibition with 3-phenylcoumarin derivatives while still maintaining selectivity.

The most active MAO B inhibitor (I) **1** (6-methoxy-3-(4-(trifluoromethyl)phenyl)coumarin) exhibited high potency at IC_{50} value 56 nM. Furthermore, the compound was found selective for the MAO B (Table 3) with trifluoromethyl and methoxy substitutions on the opposite ends optimally complementing the hydrophobic binding cavity (I, Fig. 3, Table 1). The effects of 3-phenylcoumarin substitutions have been reviewed by Carradori and Silvestri (2015). The most potent MAO B inhibitor according to the review is 3-phenylcoumarin from the work of Matos *et al.* (2011) with methyl substitutions at *para*-position of the 3-phenyl ring and position 6 of the coumarin scaffold. In addition, docking has placed the oxygen of the coumarin scaffold facing C172, making the formation of hydrogen bond possible. In the same work, having a halogen, a bromine atom, on *meta*-position on the 3-phenyl ring combined with a methoxy substitution on *para*-position was also found beneficial (Matos *et al.* 2011a, Carradori and Silvestri 2015).

The results presented here correspond well with these previous findings. Trifluoromethyl is able to fill the binding cavity and obtain favorable hydrophobic interactions in a comparable manner as smaller methyl group does. Similar impact can be observed with the conversion of methyl into methoxy group on the position 6 of the coumarin scaffold. The importance of these modifications is highlighted by taking a look at **5** (II, Fig. 2, Table 1) which has the same trifluoromethyl substitution on 3-phenyl ring as **1** and the methyl substitution on the position 6 of the coumarin scaffold of the previously reported compound. According to the experimental data (I, Table 1), the compound loses activity when the position 6 of the coumarin scaffold houses methyl (**5**) instead of methoxy (**1**). The experimentally obtained IC_{50} values are not directly comparable to the results obtained elsewhere because the assay methods or experimental conditions rarely match. However, by taking a look at the results gained regarding the activity of known MAO B inhibitor pargyline, some conclusions may be drawn. The IC_{50} value of pargyline was 61 nM (I, Table 1, Fig. S1) which is on comparable level with the compound **1**. Thus, the novel MAO B inhibitor is capable of inhibiting MAO B on a similar level as pargyline. Previously pargyline has been reported to have an IC_{50} value of 8 nM (Fisar *et al.* 2010) which might indicate that the results obtained here underestimate the inhibition activity of these compounds. In terms of lipophilicity and potency, expressed as logP and pIC_{50} (I, Fig. S6, Table S2) values, the 24 novel MAO B inhibitors fall in the middle of previously identified 3-phenylcoumarin derivatives. To conclude, the novel reported 3-phenylcoumarin derivatives exhibit similar characteristics regarding lipophilicity and potency as previously reported compounds and as such, contribute to the overall understanding of the factors affecting the MAO B activity of 3-phenylcoumarin derivatives.

Since ERα (II) has been extensively studied as a pharmaceutical target, the structural requirements for establishing ERα binding are well-known. Typically, the better the compound is able to mimic 17β-estradiol and its interactions generated by the hydroxyl groups on positions 3 and 17, the stronger it will bind. Additionally, hydrophobicity and ring structures are also

known to be beneficial for ER ligand candidates (Fang *et al.* 2001). Accordingly, 3-phenylcoumarin derivatives which are easily modifiable to comply with these requirements have been previously used in lead identification of selective ERα modulators. McKie *et al.* (2004) aspired to combine the properties of 4-hydroxytamoxifen and raloxifene in order to create a compound with both estrogen agonist and antagonist characteristics using a 3-phenyl-7-hydroxycoumarin as the core structure (McKie *et al.* 2004). As expected based on the previous work and the knowledge of the natural ERα activator 17β-estradiol, compounds **3** and **5** (II, Table 2) were found to be the most active ERα ligands. With hydroxyl groups both on *para*-position of the 3-phenyl ring and position 8 of the coumarin scaffold, the compound **5** was found to be as active as compound **3** (II, Table 2) which is equal to the 3-phenyl-7-hydroxycoumarin previously utilized as a core structure (McKie *et al.* 2004). Consequently, the results obtained here imply that by altering the substitutions on *para*-position of the 3-phenyl ring and positions 7 and 8 of the coumarin scaffold, the ERα activity of 3-phenylcoumarin can be altered.

Similarly to ERα, HSD1 (III) prefers molecules resembling its natural substrate estrone. As the precursor of 17β-estradiol, estrone has hydroxyl group on position 3 but a ketone group on position 17. Accordingly, coumarin derivatives would be expected to comply with these characteristics and in fact, 6- and 7-phenylcoumarin derivatives have been previously investigated as HSD1 inhibitors with promising results (Starčević *et al.* 2011). Because of that, 3-phenylcoumarins with their phenyl group on the opposite end of the coumarin scaffold could be expected to demonstrate parallel functionality. Besides, as a hydrophobic ring system, 3-phenylcoumarin is likely to mimic the hydrophobic packing seen in the substrate binding of the inhibitor in the PDB structure 3HB5 (Mazumdar *et al.* 2009) (III, Fig. 1D). As anticipated, 3-phenylcoumarin was found to be a promising scaffold for building HSD1 inhibitors as it obtains the hydrophobic packing typical for the known active molecules. By applying the same idea with the design of aromatase inhibitors, active compound was produced by replacing the 3-phenyl with imidazole. Therefore, it is possible to block both sulfatase and aromatase pathway of the estradiol synthesis using 3-substituted coumarin derivatives. Additionally, as the coumarin derivatives are readily substitutable, the binding can be further adjusted by designing case-specific substitutions for the promising candidates identified here.

The fact that hydroxyl group on position 7 of the coumarin scaffold (V) creates a fluorescent molecule has been common knowledge for a long time (Rangaswami and Seshadri 1940, Rangaswami *et al.* 1941, Balaiah *et al.* 1942, Sherman and Robins 1968). Other substituents modify its intensity but do not abolish the fluorescence completely. In addition, the 7-hydroxyl group of coumarin scaffold is known to be the target of glucuronidation by many UGTs. Due to the resulting loss of fluorescence, this useful property has been utilized in creation of quantitative multi-well plate assay for measuring glucuronidation rate of 7-hydroxy-4-trifluoromethylcoumarin in order to determine UGT activity (Rahikainen *et al.* 2013). Therefore, 7-hydroxycoumarin is considered as an attractive option for designing novel fluorescent substrates for UGT1A

family. Since complete mammalian UGT crystal structures are not currently available, modeling based design method was applied in order to obtain novel UGT1A10 specific 7-hydroxycoumarin derivatives. Subsequently, specificity was accomplished as a result of the substitutions on the position 3 of the coumarin scaffold. In this regard, 4-dimethylaminophenyl of derivative **4** and triazole of derivative **6** were found to be especially suitable for achieving UGT1A10 selectivity as the glucuronidation of these derivatives by the other tested human UGTs was minimal (V, Fig. 5). With the help of the newly designed substrates the determination of UGT1A10 activity in tissues can be improved and additionally, they serve as a foundation for the design of other UGT specific tool molecules.

Although it might be assumed that simple substitutions on a common and simple scaffold cannot provide novel functionalities as they might already be tested, the coumarin derivatives presented here have proven particularly fascinating. Seemingly similar compounds have been successfully used to establish selective MAO B inhibition, to block 17β-estradiol synthesis and subsequent ERα activity by targeting three different macromolecules and to create fluorescent tool compounds for metabolism analysis (Table 3). However, as a privileged scaffold (Welsch *et al.* 2010), coumarin is known to be capable of acting as a ligand in various targets but this kind of promiscuity is also typical for PAINS (Baell and Holloway 2010, Baell and Walters 2014). Due to the fluorescence properties and the various activities presented here and elsewhere, the 3-phenylcoumarin derivatives have the potential to be PAINS. Thus, that possibility was taken into account as the 52 novel 3-phenylcoumarin derivatives were filtered with PAINS filter (I). Only one of the derivatives, derivative **50** (I, Fig. S4, Table S1), was indicated as potential PAINS. This compound, in particular, differs from the rest of the derivatives by having a 1-methylpiperidine substituent, which might contribute to the PAINS properties, on position 8 of the coumarin scaffold. However, filters do not always accurately predict PAINS-like behavior or they might mistake a compound as PAINS. Thus, in order to truly confirm that the 3-phenylcoumarin derivatives are not PAINS, additional experiments should be run to determine for example lack of incubation effect for recognizing irreversible inhibition and resistance to mild detergents for colloidal aggregation (Aldrich *et al.* 2017). Despite of the simple substitutions used to produce the in-house database of coumarin derivatives, the available variations were enough to generate selectivity across this small but comprehensive selection of targets. Therefore, there is a reason to believe that selectivity could also be reached in a wider assay of targets.

By investigating the functionalities of these compounds with variety of targets, novel understanding of the target has been gained. Still, there is a chance that the ligands assume alternative binding modes to the ones presented here. In fact, multiple binding modes are relatively common (Mobley and Dill 2009) and they might be a factor here, especially, since the 3-phenylcoumarin derivatives can be considered to have symmetrical features. Thus, the presented binding modes might not explore the complete landscape available but, instead, provide a thorough examination of one likely candidate.

6.2   Virtual screening

RORγt (IV) is a transcription factor present in the development of Th17 cells. Regulating Th17 cells has become increasingly important as these cells have been linked to detrimental inflammation promoting effects in cancer and in autoimmune diseases such as arthritis and psoriasis. Accordingly, several RORγt modulators have already been identified (Kumar *et al.*, 2012, Solt *et al.* 2011, 2012, 2015, Fauber *et al.* 2013, 2014, 2015, Huh *et al.* 2013, Khan *et al.* 2013, Fauber and Magnuson 2014, Van Niel *et al.* 2014, Gege *et al.* 2014, Nishiyama *et al.* 2014, René *et al.* 2015, Toyama *et al.* 2015).

Here two structure-based virtual screening methods were applied in hopes of identifying RORγt inverse agonists. Regrettably, both GLIDE standard precision (R^2 = 0.23) and NIB Panther/ShaEP screening (R^2 = 0.27) did not provide high correlation between the scoring and the experimental data of known ligands. However, this can be anticipated as the collected experimental information was obtained from various sources with heterogeneous protocols. In both cases the virtual screening cut-off was set slightly low for identifying highly active compounds which would be able to fulfill the binding requirements. Consequently, the NIB Panther/ShaEP screening method required further modification before reasonable amount of hits was reached but ultimately both of the methods identified similar number of hits from the Specs database. Accordingly, GLIDE docking produced 46 hits with 24 of them forming at least one hydrogen bond (IV, Table 1) and NIB Panther/ShaEP screening identified 34 hits after the additional modifications. In the end, GLIDE docking and Panther/ShaEP screening identified 22 same molecules (IV, Table 1). The identified hits were tested experimentally and the best molecule was determined to have IC_{50} value of 587 nM. In total, 11 molecules had pIC_{50} value 5.0 or above (IV, Table 1) and 9 of these were identified by both GLIDE and Panther/ShaEP docking. Based on the correlation between the scoring and the experimental results, it is fair to conclude that the computational models overestimated the activity of the tested compounds. On the other hand, the assay system in question was cell based and that could lead into certain problems in the assay performance. For example, various factors from the laboratory environment to the cell culture practice of the researcher could make the experimental results inconsistent, not to mention the observed effect might not correlate with the protein-ligand interaction. Repeating the assay would be beneficial for the overall reliability of the activity data regarding the identified hits.

6.3   Homology modeling

Typically, if the 3D structure of a potential drug target is not known, X-ray crystallography or nuclear magnetic resonance (NMR) (França 2015) and

nowadays in increasing amounts cryo-electron microscopy (cryo-EM) (Murata and Wolf 2018) can be used to acquire the structure. However, all these techniques have their limitations. In addition to being rather expensive and time consuming, X-ray crystallography can only be applied to proteins which crystallize, often excluding transmembrane proteins, and NMR cannot solve structures beyond 100 kDa. NMR also requires the proteins to be highly purified which can be the limiting step. As the most important drug targets are often large in size or transmembrane proteins or both, the experimental requirements significantly hinder the rate at which 3D structures are acquired for drug design purposes (França 2015). Since cryo-EM does not require 3D crystals and macromolecules can be even observed in their natural environment, cryo-EM answers to many of the traditional challenges in solving biomolecular structures. However, cryo-EM has lower molecular weight limitation estimated at 38 kDa at present and the data analysis can be laborious (Murata and Wolf 2018). Regardless, acquiring the 3D structure of a target protein of interest is not always feasible. In that case, homology modeling may bridge the gap between the available 3D structures and the demand for new drug design targets (França 2015).

At present, X-ray crystal structure of a mammalian UGT (V) does not exist. Therefore, the functionality of these enzymes has been examined through modeling efforts which include both quantitative structure-activity relationship models and homology models (Laakkonen and Finel 2010, Dong *et al.* 2012, Ghemtio *et al.* 2014, Raunio *et al.* 2015, Tripathi *et al.* 2016). That being said, other efforts to model UGTs of the 1A family do exist, however they may not be reasonably comparable with the models obtained here. Due to low sequence identity, especially on the N-terminal, the created model is highly dependable on the selected templates and the alignment. Thus, the created model reflects the applied modeling effort which is why models of the same protein may be remarkably different from one model to another. For instance, the models created here were built in order to design selective substrates from 7-hydroxycoumarin which lose fluorescence as a result of glucuronidation. Accordingly, irrelevant sections were left out and the homology model was not optimized for loop positions and amino acid side chains outside the binding cavity. As the models worked for their intended use and were used to successfully design selective fluorescent tool molecules for UGT1A10, it would be counterproductive to scrutinize the model features in contrast to models created by others for other purposes. However, as a general homology modeling strategy for improved model precision, comparing existing models could aid in recognizing regions that are accurate to model.

7 CONCLUSIONS

In the light of the results obtained during the study, following conclusions regarding the individual studies can be drawn:

I   Selective coumarin-based inhibitors were identified for MAO B and the trifluoromethyl group on *para*-position of the most promising 3-phenylcoumarin derivative presents a unique substituent to be considered for further development.

II   Based on the understanding of the ligands preferred by ERα, 3-phenylcoumarin derivatives with substituents capable of mimicking the 17β-estradiol binding were found active in this target.

III   Coumarin scaffold can be modified to selectively inhibit both the aromatase and sulfatase pathways of the 17β-estradiol synthesis.

IV   Both NIB screening based on Panther/ShaEP and GLIDE docking identified four new structural cores as RORγt inverse agonists.

V   Designed by homology modeling and confirmed by experimental testing, selective 7-hydroxycoumarin-based tool molecules were identified for UGT1A10.

Acknowledgements

First of all I would like to thank my supervisors. Professor Olli Pentikäinen, I truly admire your vast knowledge and deep understanding of science. It has been a privilege to immerse myself into our field in your guidance and to be included in your team. I am grateful for the opportunities you have provided and for the progress I have made as a researcher. Docent Ulla Pentikäinen, thank you for your kind advisement and support during the course of my Ph.D. studies. Docent Pekka Postila, your insight has been invaluable in the final stretch of my thesis work. Thank you for teaching the secrets to an efficient publication process and for preparing me for the things to come.

Next, I would like to express my gratitude to my opponent, Docent Maija-Lahtela-Kakkonen who kindly accepted the invitation. I would also like to thank the preliminary examiners of my dissertation manuscript Docent Lari Lehtiö and Docent Henri Xhaard for their thorough and insightful feedback during the preliminary examination.

Then, I acknowledge the members of my thesis follow-up group Docent Maija Vihinen-Ranta and Docent Juhani Huuskonen. Both of you have given me plenty of valuable advises at different stages of my Ph.D. studies.

My fellow CBL-group members, thank you for your input and companionship during the time we have spent working together. Dr. Sanna Niinivehmas, I am grateful for having your support during these four years. In addition, thank you for being Sanna with all it entails and for showing me the importance of arginine. I appreciate how both you and Sakari welcomed me also in your private life and made me feel included from the very first day onward. Mira Ahinko, I value your enthusiasm and it has been a pleasure to work alongside such a young and bright mind. Sakari Lätti, thank you for the lunch discussions and the endless ideas for alternative career plans should this path prove futile.

I thank especially Arja Mansikkaviita but also Laura Pitkänen, Alli Liukkonen and Tatu Haataja for the help and advice I have received in the lab. Additionally, my family taught me to always aspire to understand how things work which ignited my passion for science. My parents and sisters, thank you for your constant support.

Finally, my other half Erkki, thank you for your patience during this seemingly endless process. You are my rock, the one I can count on. I appreciate everything you have done and continue to do for me and I hope I can return the favor. Special thanks belongs also to my fur babies Ilo and Fera, you have made sure I get a daily dose of fresh air and offered much needed paw therapy.

YHTEENVETO (RÉSUMÉ IN FINNISH)

Pienmolekyylit amiinien oksidaation, tumareseptorien signaloinnin ja glukuronidaation muovaajina – 3-fenyylikumariini tutkimuksen kohteena

Proteiinit ovat monimuotoisia makromolekyylejä, jotka vastaavat suuresta joukosta tehtäviä. Ne voivat esimerkiksi olla rakenteellisia proteiineja, jotka huolehtivat solujen muodosta ja liikkeestä. Välittäjäproteiinit tai hormonit puolestaan koordinoivat biologisia prosesseja kuljettamalla viestejä ja entsyymit katalysoivat kemiallisia reaktioita. Vasta-aineet taasen auttavat puolustautumaan sairaudenaiheuttajia vastaan. Jotta proteiinit voivat aikaansaada kaiken tämän, niiden koko ja muoto vaihtelee pienistä peptideistä suuriin, useita yksiköitä eli domeeneja sisältäviin komplekseihin. Riippumatta koosta tai rakenteesta, jokainen proteiini koostuu 20 erilaisesta rakennuspalikasta eli aminohaposta – vain niiden järjestys vaihtelee.

Proteiinien toimintaa ohjaavat pienmolekyylit, joita kutsutaan ligandeiksi. Ligandit muodostavat proteiinien kanssa komplekseja, aikaansaaden biologisen vastineen. Ligandista riippuen tämä vastine saattaa olla proteiinin aktivoituminen, jolloin ligandia kutsutaan agonistiksi. Jos taas proteiini osallistuu kemialliseen reaktioon, jossa sitoutunut ligandi on keskeisessä osassa, kutsutaan tällaista ligandia substraatiksi. Toisaalta ligandin sitoutumisen biologiset seuraukset saattavat olla täysin päinvastaiset, jolloin ligandia kutsutaan käänteisagonistiksi. Jos proteiinin toiminta puolestaan hidastuu tai lakkaa, on kyseessä antagonisti tai inhibiittori. Muodostamalla yksityiskohtaisia verkostoja ja vaikuttamalla proteiini-ligandi vuorovaikutuksien aikaansaamiin seurauksiin, biologisia vuorovaikutuksia hallinnoidaan solutasolla.

Lääkeainesuunnittelussa proteiinien kykyä sitoa ligandeja hyödynnetään kun lääkinnällisesti hyödyllisiä vaikutuksia yritetään aikaansaada merkittävissä kohdeproteiineissa. Ligandin sitoutumisvoimakkuutta proteiiniin kuvaillaan affiniteettina, joka myös kuvailee ligandin aktiivisuutta – mitä parempi affiniteetti, sitä aktiivisempi ligandi. Suunnitellessa ligandia käytettäväksi lääkeaineena, tavoitteena on löytää sellainen ligandi, jolla on korkea affiniteetti juuri toivottua kohdetta kohtaan. Samalla lupaavien lääkekandidaattien tulee välttää aktiivisuutta muita kohteita kohtaan. Toisin sanoen, tavoitteena on luoda selektiivisiä ligandeja. Tämän tavoitteen saavuttamiseksi pyritään kehittämään sellaisia vuorovaikutuksia, jotka ovat ominaisia kulloinkin tarkasteltavalle proteiini-ligandi komplekseille.

Tällä hetkellä tunnetaan valtava määrä lääkinnällisesti merkittäviä kohteita ja uusia kohteita löydetään jatkuvasti lisää. Siksi uusien lääkeaineiden kustannustehokas kehittäminen on erityisen tärkeää. Tietokoneavusteinen lääkeainesuunnittelu pyrkii nimenomaan kaventamaan lääkeaineiden tunnistamisen ja kliinisten tutkimusten välistä kuilua sekä ajankäytön että taloudellisten vaatimusten suhteen. Tässä väitöskirjassa tarkastellaan amiinioksidaasien, tumareseptoreiden ja uridiini-5′-difosfoglukuronyylitransferaasien kykyä sitoa 3-fenyylikumariineja, jotka ovat helposti räätälöitävissä kunkin kohteen tarpei-

siin. Tavoitteena on tunnistaa ne tekijät, jotka määrittelevät selektiivisyyden kussakin tapauksessa. Tämän lisäksi lääkeaineiden aineenvaihdunta otetaan huomioon luomalla uusia työkalumolekyylejä tarkempaa aineenvaihdunnan analysointia varten.

Ensimmäisessä osatyössä (I) pyrittiin tunnistamaan selektiivisiä mono-amiinioksidaasi B:n (MAO B) inhibiittoreita. Testatusta joukosta 3-fenyylikumariinijohdannaisia, lupaavin molekyyli **1** saavutti IC_{50}-arvon 56 nM (I, Kuva 3, Taulukko 1) eli nanomolaarisen pitoisuuden, jossa puolet entsyymin toiminnasta estyy. Lisäksi kyseinen molekyyli todettiin selektiiviseksi MAO B:tä kohtaan, jolloin se ei estä muiden testattujen proteiinien toimintaa. Mole-kyylin **1** 3-fenyylikumariinin trifluorimetyyli- ja metoksisubstituutiot molekyy-lin vastakkaisissa päissä täydentävät optimaalisesti hydrofobisen sitoutumis-alueen. Nämä saavutetut tulokset vastaavat hyvin muiden aiemmin saavutta-mia tuloksia, vaikka kokeelliset tulokset eivät olekaan täysin verrattavissa toi-siinsa poikkeavien menetelmien takia. Voidaan kuitenkin todeta, että paras mo-lekyyli **1** on yhtä hyvä inhibiittori kuin tunnettu inhibiittori pargyliini (I, Kuva S1), sillä näiden molekyylien IC_{50}-arvot vastaavat toisiaan. Koska tämä joukko uusia 3-fenyylikumariinijohdannaisia pitää sisällään samankaltaisia ominai-suuksia kuin aikaisemmin löydetyt MAO B:tä inhiboivat johdannaiset, tämä osatyö osaltaan täydentää ymmärrystä niitä ominaisuuksia kohtaan, jotka ai-kaansaavat MAO B:tä kohtaan aktiivisen 3-fenyylikumariinijohdannaisen.

Toisessa osatyössä (II) tarkasteltiin estrogeenireseptoria α (ERα), joka on erittäin tutkittu lääkekehityksen kohde. Tästä syystä ne rakenteelliset ominai-suudet, jotka aikaansaavat molekyylin sitoutumisen ERα:aan, tunnetaan hyvin. Tyypillisesti hyvin sitoutuva molekyyli kykenee jäljittelemään ERα:n luonnol-lista ligandia 17β-estradiolia ja sen sitoutumiseen vaadittavia vuorovaikutuksia, kuten 3- ja 17-hydroksyylien aikaansaamia vetysidoksia ja steroidille ominaisen rengasrakenteen hydrofobisia vuorovaikutuksia. Tässä mielessä 3-fenyyli-kumariinijohdannaiset ovat muokattavuutensa vuoksi hyviä ehdokkaita sovel-lettaviksi ERα-kohdennettuina lääkeaineina. Perustuen aiempiin havaintoihin ja ERα:n rakenteeseen liittyvään tuntemukseen, molekyylit **3** ja **5** (II, Taulukko 2) todettiin kaikkein aktiivisimmaksi ERα ligandeiksi. Molekyylillä **3** on hydrok-syyliryhmä kumariinin 7-asemassa ja molekyylillä 5 sekä kumariinin 8-asemassa että 3-fenyylin *para*-asemassa. Tästä voidaan päätellä, että vaihtele-malla näissä asemissa sijaitsevia substituutioita, voidaan vaikuttaa 3-fenyyli-kumariinin aktiivisuuteen ERα ligandina.

Kuten ERα, kolmannen osatyön (III) 17β-hydroksisteroidi dehydrogenaasi 1 (HSD1) suosii molekyylejä, jotka muistuttavat sen luonnollista substraattia estronia. Estroni on 17β-estradiolin esiaste, joten sillä on 17β-estradiolin tapaan 3-asemassa hydroksyyliryhmä, mutta 17-asemassa on karbonyyliryhmä. Vas-taavasti kumariinijohdannaisten voidaan olettaa mukautuvan myös näihin vaa-timuksiin ja 6- sekä 7-fenyylikumariineilla onkin havaittu olevan HSD1:tä inhi-boivia ominaisuuksia. Koska 3-fenyylikumariinin rengasrakenne muistuttaa läheisesti muita aktiiviseksi havaittuja fenyylikumaiineja, voidaan 3-fenyyli-kumariinijohdannaisilla olettaa olevan rinnakkaista toiminnallisuutta. Odote-

tusti 3-fenyylikumariini todettiin lupaavaksi lähtökohdaksi lähteä rakentamaan HSD1 inhibiittoreita, sillä se tavoittaa tunnettujen aktiivisten molekyylien kantaisen hydrofobisen pakkautumisen sitoutumisalueella. Samaa ideaa hyödynnettiin myös suunniteltaessa inhibiittoria aromataasille, jolloin aikaansaatiin aktiivinen molekyyli korvaamalla 3-fenyyli imidatsolilla. Käyttämällä 3-aseman suhteen substituoituja kumariinijohdannaisia on siis mahdollista estää 17β-estradiolin steroidigeneesi sekä sulfataasiin että aromataasiin perustuvilla reiteillä. Koska kumariinijohdannaiset ovat helposti muokattavissa, voidaan osatyössä tunnistettujen lupaavien molekyylien sitoutumista parantaa kuhunkin tarkoitukseen soveltuvilla substituutioilla.

Neljännessä osatyössä (IV) käytettiin kahta rakenteeseen perustuvaa virtuaaliseulontamenetelmää, joiden avulla pyrittiin tunnistamaan retinoiinihapporeseptorin sukuisen orporeseptori γt:n (RORγt) käänteisagonisteja. Käytettyjen menetelmien GLIDE ($R^2 = 0{,}23$) ja Panther/ShaEP ($R^2 = 0{,}27$) pisteytysfunktioiden antamat tulokset eivät kuitenkaan kumpikaan korreloineet kovin hyvin kokeellisten tulosten kanssa. Tämä oli odotettavissa, sillä kokeelliset tulokset ovat peräisin useista lähteistä ja niiden tuottamisessa on käytetty vaihtelevia menetelmiä. Tästä syystä molemmissa menetelmissä pyrittiin tunnistamaan kaikkein aktiivisimmat molekyylit, jotka täyttävät sitoutumiseen liittyvät vaatimukset, asettamalla aktiivisten ja inaktiivisten erotteluun käytetty raja hieman alhaiseksi. Molempien menetelmien avulla tunnistettiin lopulta samankaltainen määrä aktiivisia molekyylejä, GLIDE tunnisti 46 molekyyliä, joista 24 muodostivat ainakin yhden vetysidoksen ja Panther/ShaEP tunnisti 34 molekyyliä (IV, Taulukko 1). Näistä molekyyleistä 22 olivat samoja. Tunnistetut aktiiviset molekyylit testattiin kokeellisesti ja kaikkein aktiivisimman molekyylin IC_{50}-arvoksi saatiin 587 nM. Kaiken kaikkiaan 11 molekyylillä oli pIC_{50}-arvo 5,0 tai korkeampi (IV, Taulukko 1) ja näistä molekyyleistä 9 tunnistettiin molempien menetelmien avulla. Tulosten korrelaation perusteella kumpikin menetelmä arvioi testatut molekyylit todellista aktiivisemmiksi, mutta toisaalta paras molekyyli ennustettiin kaikkein tarkimmin.

Viidennessä osatyössä (V) tehtiin homologiamalleja uridiini-5'-difosfoglukuronyylitransferaasi 1A (UGT1A) perheelle, sillä niille ei tunneta röntgenkristallografista rakennetta. Homologiamallin perusteella pyrittiin suunnittelemaan 7-hydroksikumariiniin perustuvia selektiivisiä substraatteja UGT1A10 alatyypille, sillä UGT-entsyymien aikaansaaman 7-hydroksiryhmään kohdistuvan glukuronidaatioreaktion seurauksena 7-hydroksikumariinin fluoresenssi katoaa. Seuraamalla muutosta fluoresenssissa, pystytään tehokkaammin tarkastelemaan entsyymin toimintaa, joten tällaiset työkalumolekyylit ovat hyödyksi UGT1A10-entsyymin tutkimuksessa. Homologiamallien perusteella tunnistettiin UGT1A10-selektiivisiä 7-hydroksikumariinijohdannaisia. Etenkin molekyylien **4** 4-dimetyyliaminofenyyli- ja **6** triatsolisubstituutio 3-asemassa todettiin hyödylliseksi UGT1A10 selektiivisen glukuronidaation saavuttamiseksi. Tällä tavoin tunnistettuja molekyylejä voidaan kehittää eteenpäin sekä hyödyntää pyrittäessä luomaan selektiivisiä fluoresoivia työkalumolekyylejä myös muille UGT1A-perheen entsyymeille.

REFERENCES

Abraham K., Wöhrlin F., Lindtner O., Heinemeyer G. & Lampen A. 2010. Toxicology and risk assessment of coumarin: Focus on human data. *Molecular Nutrition and Food Research* 54: 228–239.

Adeghate E. & Parvez H. 2004. The Effect of Diabetes Mellitus on the Morphology and Physiology of Monoamine Oxidase in the Pancreas. *NeuroToxicology* 25: 167–173.

Aguilar F., Autrup H.N., Barlow S., Castle L., Crebelli R., Dekant W., Engel K., Gontard N., Gott D.M., Grilli S., Gürtler R., Larsen J.C., Leclercq C., Leblanc J., Xavier F., Mennes W., Milana M.R., Pratt I., Magdalena I., Maria C., Tobback P.P. & Toldrá F. 2008. Coumarin in flavourings and other food ingredients with flavouring properties – Scientific Opinion of the Panel on Food Additives, Flavourings, Processing Aids and Materials in Contact with Food (AFC). *The EFSA Journal* 793: 1–15.

Aldrich C., Bertozzi C., Georg G.I., Kiessling L., Lindsley C., Liotta D., Merz K.M., Schepartz A. & Wang S. 2017. The Ecstasy and Agony of Assay Interference Compounds. *Journal of Chemical Information and Modeling* 57: 387–390.

Ali S. & Coombes R.C. 2000. Estrogen Receptor Alpha in Human Breast Cancer: Occurrence and Significance. *Journal of Mammary Gland Biology and Neoplasia* 5: 271–281.

Altschul S.F., Gish W., Miller W., Myers E.W. & Lipman D.J. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403–410.

Amor B.R.C., Schaub M.T., Yaliraki S.N. & Barahona M. 2016. Prediction of allosteric sites and mediating interactions through bond-to-bond propensities. *Nature Communications* 7: 1–13.

Andrés A.M., Soldevila M., Navarro A., Kidd K.K., Oliva B. & Bertranpetit J. 2004. Positive selection in MAOA gene is human exclusive: Determination of the putative amino acid change selected in the human lineage. *Human Genetics* 115: 377–386.

Bach A.W., Lan N.C., Johnson D.L., Abell C.W., Bembenek M.E., Kwan S.W., Seeburg P.H. & Shih J.C. 1988. cDNA cloning of human liver monoamine oxidase A and B: molecular basis of differences in enzymatic properties. *Proceedings of the National Academy of Sciences of the United States of America* 85: 4934–4938.

Baell J.B. & Holloway G.A. 2010. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *Journal of Medicinal Chemistry* 53: 2719–2740.

Baell J. & Walters M.A. 2014. COMMENT: Chemical con artists foil drug discovery. *Nature* 513: 481–483.

Bais H.P. & Ravishankar G.A. 2001. *Cichorium intybus* L - Cultivation, processing, utility, value addition and biotechnology, with an emphasis on current status and future prospects. *Journal of the Science of Food and Agriculture* 81: 467–484.

76

Balaiah V., Seshadri T.R. & Venkateswarlu V. 1942. Visible Fluorescence and Chemical Constitution of Compounds of the Benzopyrone Group Part III. Further Study of Structural Influences in Coumarins. *Proceedings of the Indian Academy of Sciences - Section A* 16A: 68–82.

Bento A.P., Gaulton A., Hersey A., Bellis L.J., Chambers J., Davies M., Krüger F.A., Light Y., Mak L., McGlinchey S., Nowotka M., Papadatos G., Santos R. & Overington J.P. 2014. The ChEMBL bioactivity database: An update. *Nucleic Acids Research* 42: 1083–1090.

Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N. & Bourne P.E. 2000. The Protein Data Bank. *Nucleic Acids Research* 28: 235–242.

Bethea D., Fullmer B., Syed S., Seltzer G., Tiano J., Rischko C., Gillespie L., Brown D. & Gasparro F.P. 1999. Psoralen photobiology and photochemotherapy: 50 years of science and medicine. *Journal of Dermatological Science* 19: 78–88.

Binda C., Li M., Hubalek F., Restelli N., Edmondson D.E. & Mattevi A. 2003. Insights into the mode of inhibition of human mitochondrial monoamine oxidase B from high-resolution crystal structures. *Proceedings of the National Academy of Sciences of the United States of America* 100: 9750–9755.

Binda C., Newton-Vinson P., Hubálek F., Edmondson D.E. & Mattevi A. 2002. Structure of human monoamine oxidase B, a drug target for the treatment of neurological disorders. *Nature Structural Biology* 9: 22–26.

Binda C., Wang J., Pisani L., Caccia C., Carotti A., Salvati P., Edmondson D.E. & Mattevi A. 2007. Structures of human monoamine oxidase B complexes with selective noncovalent inhibitors: Safinamide and coumarin analogs. *Journal of Medicinal Chemistry* 50: 5848–5852.

Borges F., Roleira F., Milhazes N., Santana L. & Uriarte E. 2005. Simple Coumarins and Analogues in Medicinal Chemistry: Occurrence, Synthesis and Biological Activity. *Current Medicinal Chemistry* 12: 887–916.

Broadley K.J. 2010. The vascular effects of trace amines and amphetamines. *Pharmacology and Therapeutics* 125: 363–375.

Brzozowski A.M., Pike A.C.W., Dauter Z., Hubbard R.E., Bonn T., Engström O., Öhman L., Greene G.L., Gustafsson J.Å. & Carlquist M. 1997. Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* 389: 753–758.

Caldwell J., Gardner I. & Swales N. 1995. An Introduction to Drug Disposition: The Basic Principles of Absorption , Distribution , Metabolism , and Excretion. *Toxicologic Pathology* 23: 102–114.

Carradori S. & Silvestri R. 2015. New Frontiers in Selective Human MAO‐B Inhibitors Miniperspective. *Journal of Medicinal Chemistry* 58: 6717–6732.

Catto M., Nicolotti O., Leonetti F., Carotti A., Favia A.D., Soto-Otero R., Méndez-Álvarez E. & Carotti A. 2006. Structural Insights into Monoamine Oxidase Inhibitory Potency and Selectivity of 7-Substituted Coumarins from Ligand- and Target-Based Approaches. *Journal of Medicinal Chemistry* 49: 4912–4925.

Chen S., Cho M., Karlsberg K., Zhou D. & Yuan Y.C. 2004. Biochemical and Biological Characterization of a Novel Anti-aromatase Coumarin Derivative. *Journal of Biological Chemistry* 279: 48071–48078.

Christopoulos A. 2002. Allosteric Binding Sites on Cell-Surface Receptors: Novel Targets for Drug Discovery. *Nature Reviews Drug Discovery* 1: 198–210.

Chunyan C., Bo S., Ping L., Jingmei L. & Ito Y. 2008. Isolation and Purification of Psoralen and Bergapten from *Ficus carica* L. Leaves by High-Speed Countercurrent Chromatography. *Journal of Liquid Chromatography & Related Technologies* 32: 136–143.

Colibus L. De, Li M., Binda C., Lustig A., Edmondson D.E. & Mattevi A. 2005. Three-dimensional structure of human monoamine oxidase A (MAO A): relation to the structures of rat MAO A and human MAO B. *Proceedings of the National Academy of Sciences of the United States of America* 102: 12684–12689.

Collier G. & Ortiz V. 2013. Emerging computational approaches for the study of protein allostery. *Archives of Biochemistry and Biophysics* 538: 6–15.

Cui J., Shen Y. & Li R. 2013. Estrogen synthesis and signaling pathways during ageing: from periphery to brain. *Trends in Molecular Medicine* 19: 197–209.

Cukuroglu E., Engin H.B., Gursoy A. & Keskin O. 2014. Hot spots in protein-protein interfaces: Towards drug discovery. *Progress in Biophysics and Molecular Biology* 116: 165–173.

Darden T., York D. & Pedersen L. 1993. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* 98: 10089–10092.

Dassen H., Punyadeera C., Kamps R., Delvoux B., Langendonckt A. Van, Donnez J., Husen B., Thole H., Dunselman G. & Groothuis P. 2007. Estrogen metabolizing enzymes in endometrium and endometriosis. *Human Reproduction* 22: 3148–3158.

Dong D., Ako R., Hu M. & Wu B. 2012. Understanding substrate selectivity of human UDP-glucuronosyltransferases through QSAR modeling and analysis of homologous enzymes. *Xenobiotica* 42: 808–820.

Duan Y., Wu C., Chowdhury S., Lee M.C., Xiong G., Zhang W., Yang R., Cieplak P., Luo R., Lee T., Caldwell J., Wang J. & Kollman P. 2003. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *Journal of Computational Chemistry* 24: 1999–2012.

Edmondson D.E., Mattevi A., Binda C., Li M. & Hubálek F. 2005. Structure and Mechanism of Monoamine Oxidase. *Burger's Medicinal Chemistry and Drug Discovery* 11: 1983–1993.

Fang H., Tong W., Shi L.M., Blair R., Perkins R., Branham W., Hass B.S., Xie Q., Dial S.L., Moland C.L. & Sheehan D.M. 2001. Structure - Activity relationships for a large diverse set of natural, synthetic, and environmental estrogens. *Chemical Research in Toxicology* 14: 280–294.

Fauber B.P., Gobbi A., Robarge K., Zhou A., Barnard A., Cao J., Deng Y., Eidenschenk C., Everett C., Ganguli A., Hawkins J., Johnson A.R., La H.,

78

Norman M., Salmon G., Summerhill S., Ouyang W., Tang W. & Wong H. 2015. Discovery of imidazo[1,5-a]pyridines and -pyrimidines as potent and selective RORc inverse agonists. *Bioorganic and Medicinal Chemistry Letters* 25: 2907–2912.

Fauber B.P., Leon Boenig G. De, Burton B., Eidenschenk C., Everett C., Gobbi A., Hymowitz S.G., Johnson A.R., Liimatta M., Lockey P., Norman M., Ouyang W., René O. & Wong H. 2013. Structure-based design of substituted hexafluoroisopropanol- arylsulfonamides as modulators of RORc. *Bioorganic and Medicinal Chemistry Letters* 23: 6604–6609.

Fauber B.P. & Magnuson S. 2014. Modulators of the Nuclear Receptor Retinoic Acid Receptor-Related Orphan Receptor-γ (RORγ or RORc). *Journal of Medicinal Chemistry* 57: 5871–5892.

Fauber B.P., René O., Burton B., Everett C., Gobbi A., Hawkins J., Johnson A.R., Liimatta M., Lockey P., Norman M. & Wong H. 2014. Identification of tertiary sulfonamides as RORc inverse agonists. *Bioorganic and Medicinal Chemistry Letters* 24: 2182–2187.

Filic V., Vladic A., Stefulj J., Cicin-Sain L., Balija M., Sucic Z. & Jernej B. 2005. Monoamine oxidases A and B gene polymorphisms in migraine patients. *Journal of the Neurological Sciences* 228: 149–153.

Finan C., Gaulton A., Kruger F.A., Lumbers R.T., Shah T., Engmann J., Galver L., Kelley R., Karlsson A., Santos R., Overington J.P., Hingorani A.D. & Casas J.P. 2017. The druggable genome and support for target identification and validation in drug development. *Science Translational Medicine* 9: eaag1166.

Finberg J.P.M. & Rabey J.M. 2016. Inhibitors of MAO-A and MAO-B in psychiatry and neurology. *Frontiers in Pharmacology* 7: 1–15.

Fisar Z., Hroudová J. & Raboch J. 2010. Inhibition of monoamine oxidase activity by antidepressants and mood stabilizers. *Neuroendocrinology Letters* 31: 645–656.

França T.C.C. 2015. Homology modeling: An important tool for the drug discovery. *Journal of Biomolecular Structure and Dynamics* 33: 1780–1793.

Friesner R.A., Banks J.L., Murphy R.B., Halgren T.A., Klicic J.J., Mainz D.T., Repasky M.P., Knoll E.H., Shelley M., Perry J.K., Shaw D.E., Francis P. & Shenkin P.S. 2004. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry* 47: 1739–1749.

Gaweska H. & Fitzpatrick P.F. 2011. Structures and Mechanism of the Monoamine Oxidase Family. *Biomolecular Concepts* 2: 365–377.

Gege C., Schlüter T. & Hoffmann T. 2014. Identification of the first inverse agonist of retinoid-related orphan receptor (ROR) with dual selectivity for RORβ and RORγt. *Bioorganic and Medicinal Chemistry Letters* 24: 5265–5267.

Ghemtio L., Soikkeli A., Yliperttula M., Hirvonen J., Finel M. & Xhaard H. 2014. SVM classification and CoMSIA modeling of UGT1A6 interacting molecules. *Journal of Chemical Information and Modeling* 54: 1011–1026.

Gnerre C., Catto M., Leonetti F., Weber P., Carrupt P.A., Altomare C., Carotti A. & Testa B. 2000. Inhibition of Monoamine Oxidases by Functionalized

Coumarin Derivatives: Biological Activities, QSARs, and 3D-QSARs. *Journal of Medicinal Chemistry* 43: 4747–4758.

Gong Q.H., Cho J.W., Huang T., Potter C., Gholami N., Basu N.K., Kubota S., Carvalho S., Pennington M.W., Owens I.S. & Popescu N.C. 2001. Thirteen UDPglucuronosyltransferase genes are encoded at the human UGT1 gene complex locus. *Pharmacogenetics* 11: 357–368.

Green S., Walter P., Kumar V., Krust A., Bornert J.-M., Argos P. & Chambon P. 1986. Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A. *Nature* 320: 134–139.

Gronemeyer H., Gustafsson J.-Å. & Laudet V. 2004. Principles for Fodulation of the Nuclear Receptor Superfamily. *Nature Reviews Drug Discovery* 3: 950–964.

Güner O.F. & Bowen J.P. 2014. Setting the Record Straight: The Origin of the Pharmacophore Concept. *Journal of Chemical Information and Modeling* 54: 1269–1283.

Halgren T.A. 1996a. Merck Molecular Force Field. I. Basis, Form, Scope, Parametrization, and Performance of MMFF94. *Journal of Computational Chemistry* 17: 490–519.

Halgren T.A. 1996b. Merck Molecular Force Field. II. MMFF94 van del Waals and Electrostatic Parameters for Intermolecular Interactions. *Journal of Computational Chemistry* 17: 520–552.

Halgren T.A. 1996c. Merck Molecular Force Field. III. Molecular Geometries and Vibrational Frequencies for MMFF94. *Journal of Computational Chemistry* 17: 553–586.

Halgren T.A. 1996d. Merck Molecular Force Field. V. Extension of MMFF94 Using Experimental Data , Additional Computational Data , and Empirical Rules. *Journal of Computational Chemistry* 17: 616–641.

Halgren T.A., Murphy R.B., Friesner R.A., Beard H.S., Frye L.L., Pollard W.T. & Banks J.L. 2004. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *Journal of Medicinal Chemistry* 47: 1750–1759.

Halgren T.A. & Nachbar R.B. 1996. Merck Molecular Force Field. IV. Conformational Energies and Geometries for MMFF94. *Journal of Computational Chemistry* 17: 587–615.

Hanamura T., Niwa T., Gohno T., Kurosumi M., Takei H., Yamaguchi Y., Ito K.I. & Hayashi S.I. 2014. Possible role of the aromatase-independent steroid metabolism pathways in hormone responsive primary breast cancers. *Breast Cancer Research and Treatment* 143: 69–80.

Hanley A.J. & McNeil J.B. 1982. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology* 143: 29–36.

Hansch C., Maloney P.P., Fujita T. & Muir R.M. 1962. Correlation of Biological Activity of Phenoxyacetic Acids with Hammett Substituent Constants and Partition Coefficients. *Nature* 194: 178–180.

Hare M.L. 1928. Tyramine oxidase: A new enzyme system in liver. *The Biochemical journal* 22: 968–979.

Hasan F., McCrodden J.M., Kennedy N.P. & Tipton K.F. 1988. The involvement of intestinal monoamine oxidase in the transport and metabolism of tyramine. *Journal of Neural Transmission, Supplement* 26: 1–9.

He X., Chen Y.Y., Shi J.B., Tang W.J., Pan Z.X., Dong Z.Q., Song B.A., Li J. & Liu X.H. 2014. New coumarin derivatives: Design, synthesis and use as inhibitors of hMAO. *Bioorganic and Medicinal Chemistry* 22: 3732–3738.

Hiromoto T., Honjo E., Tamada T., Noda N., Kazuma K., Suzuki M. & Kuroki R. 2013. Crystal structure of UDP-glucose:anthocyanidin 3-*O*-glucosyltransferase from *Clitoria ternatea*. *Journal of Synchrotron Radiation* 20: 894–898.

Holt A., Sharman D.F., Baker G.B. & Palcic M.M. 1997. A Continuous Spectrophotometric Assay for Monoamine Oxidase and Related Enzymes in Tissue Homogenates. *Analytical Biochemistry* 244: 384–392.

Hopfinger A.J., Wang S., Tokarski J.S., Jin B., Albuquerque M., Madhav P.J. & Duraiswami C. 1997. Construction of 3D-QSAR Models Using the 4D-QSAR Analysis Formalism. *Journal of the American Chemical Society* 119: 10509–10524.

Houslay M.D. & Tipton K.F. 1974. A Kinetic Evaluation of Monoamine Oxidase Activity in Rat Liver Mitochondrial Outer Membranes. *The Biochemical Journal* 139: 645–652.

Huang B. & Schroeder M. 2006. LIGSITEcsc: Predicting ligand binding sites using the Connolly surface and degree of conservation. *BMC Structural Biology* 6: 1–11.

Huang N., Shoichet B.K. & Irwin J.J. 2006. Benchmarking Sets for Molecular Docking Benchmarking Sets for Molecular Docking. *Journal of Medicinal Chemistry* 49: 6789–6801.

Huh J.R., Englund E.E., Wang H., Huang R., Huang P., Rastinejad F., Inglese J., Austin C.P., Johnson R.L., Huang W. & Littman D.R. 2013. Identification of Potent and Selective Diphenylpropanamide RORγ Inhibitors. *ACS Medicinal Chemistry Letters* 4: 79–84.

Humphrey W., Dalke A. & Schulten K. 1996. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* 14: 33–38.

Irwin J.J. & Shoichet B.K. 2005. ZINC - A free database of commercially available compounds for virtual screening. *Journal of Chemical Information and Modeling* 45: 177–182.

Iwata N., Kainuma M., Kobayashi D., Kubota T., Sugawara N., Uchida A., Ozono S., Yamamuro Y., Furusyo N., Ueda K., Tahara E. & Shimazoe T. 2016. The relation between Hepatotoxicity and the Total Coumarin Intake from Traditional Japanese Medicines Containing Cinnamon Bark. *Frontiers in Pharmacology* 7: 1–9.

Johnson M., Lajiness M. & Maggiora G. 1989. Molecular similarity: a basis for designing drug screening programs. *Progress in clinical and biological research* 291: 167–171.

Johnson M.S. & Overington J.P. 1993. A structural basis for sequence comparisons. An evaluation of scoring methodologies. *Journal of molecular biology* 233: 716–738.

Johnston J.P. 1968. Some observations upon a new inhibitor of monoamine oxidase in brain tissue. *Biochemical Pharmacology* 17: 1285–1297.

Jordan V.C. 2003a. Antiestrogens and Selective Estrogen Receptor Modulators as Multifunctional Medicines. 1. Receptor Interactions. *Journal of Medicinal Chemistry* 46: 883–908.

Jordan V.C. 2003b. Antiestrogens and Selective Estrogen Receptor Modulators as Multifunctional Medicines. 2. Clinical Considerations and New Agents. *Journal of Medicinal Chemistry* 46: 1081–1111.

Jorgensen W.L. 2004. The Many Roles of Computation in Drug Discovery. *Science* 303: 1813–1818.

Kaur V., Kumar M., Kumar A., Kaur K., Dhillon V.S. & Kaur S. 2018. Pharmacotherapeutic potential of phytochemicals: Implications in cancer chemoprevention and future perspectives. *Biomedicine and Pharmacotherapy* 97: 564–586.

Khan P.M., El-Gendy B.E.D.M., Kumar N., Garcia-Ordonez R., Lin L., Ruiz C.H., Cameron M.D., Griffin P.R. & Kamenecka T.M. 2013. Small molecule amides as potent ROR-γ selective modulators. *Bioorganic and Medicinal Chemistry Letters* 23: 532–536.

Kleywegt G.J. & Jones T.A. 1994. Detection, Delineation, Measurement and Display of Cavities in Macromolecular Structures. *Acta Crystallographica Section D: Biological Crystallography* 50: 178–185.

Knox A.J.S., Meegan M.J., Sobolev V., Frost D., Zisterer D.M., Williams D.C. & Lloyd D.G. 2007. Target Specific Virtual Screening: Optimization of an Estrogen Receptor Screening Platform. *Journal of Medicinal Chemistry* 50: 5301–5310.

Korb O., Stützle T. & Exner T.E. 2006. PLANTS: Application of Ant Colony Optimization to Structure-Based Drug Design. *Lecture Notes in Computer Science* 4150: 247–258.

Korb O., Stützle T. & Exner T.E. 2009. Empirical Scoring Functions for Advanced Protein-Ligand Docking with PLANTS. *Journal of Chemical Information and Modeling* 49: 84–96.

Kraulis P.J. 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *Journal of Applied Crystallography* 24: 946–950.

Kumar N., Kamenecka T., Lyda B., Khan P., Chang M.R., Garcia-Ordonez R.D., Cameron M., Ferguson J., Mercer B.A., Hodder P., Rosen H. & Griffin P.R. Identification of a novel selective inverse agonist probe and analogs for the Retinoic acid receptor-related Orphan Receptor Gamma (RORγ) 2012 Apr 16 [Updated 2013 Mar 14]. *Probe Reports from the NIH Molecular Libraries Program [Internet]. Bethesda (MD): National Center for Biotechnology Information (US).*

Kumar N., Lyda B., Chang M.R., Lauer J.L., Solt L.A., Burris T.P., Kamenecka T.M. & Griffin P.R. 2012. Identification of SR2211: A Potent Synthetic RORγ-Selective Modulator. *ACS Chemical Biology* 7: 672–677.

82

Kuntz I.D., Blaney J.M., Oatley S.J., Langridge R. & Ferrin T.E. 1982. A geometric approach to macromolecule-ligand interactions. *Journal of Molecular Biology* 161: 269–288.

Kurebayashi S., Nakajima T., Kim S.C., Chang C.Y., McDonnell D.P., Renaud J.P. & Jetten A.M. 2004. Selective LXXLL peptides antagonize transcriptional activation by the retinoid-related orphan receptor RORγ. *Biochemical and Biophysical Research Communications* 315: 919–927.

Laakkonen L. & Finel M. 2010. A Molecular Model of the Human UDP-Glucuronosyltransferase 1A1, Its Membrane Orientation, and the Interactions between Different Parts of the Enzyme. *Molecular Pharmacology* 77: 931–939.

Lee H.-R., Kim T.-H. & Choi K.-C. 2012. Functions and physiological roles of two types of estrogen receptors, ERα and ERβ, identified by estrogen receptor knockout mouse. *Laboratory Animal Research* 28: 71–76.

Lehtonen J. V, Still D.-J., Rantanen V.-V., Ekholm J., Björklund D., Iftikhar Z., Huhtala M., Repo S., Jussila A., Jaakkola J., Pentikäinen O., Nyrönen T., Salminen T., Gyllenberg M. & Johnson M.S. 2004. BODIL: a molecular modeling environment for structure-function analysis and drug design. *Journal of Computer-Aided Molecular Design* 18: 401–419.

Liang C., Ju W., Pei S., Tang Y. & Xiao Y. 2017. Pharmacological activities and synthesis of esculetin and its derivatives: A mini-review. *Molecules* 22: 387.

Link K.P. 1959. The Discovery of Dicoumarol and Its Sequels. *Circulation* 19: 97–107.

Lionta E., Spyrou G., Vassilatis D.K. & Cournia Z. 2014. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Current Topics in Medicinal Chemistry* 14: 1923–1938.

Lipinski C.A., Lombardo F., Dominy B.W. & Feeney P.J. 1997. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Develop ment Settings. *Advanced Drug Delivery Reviews* 23: 3–25.

Littman D.R. & Rudensky A.Y. 2010. Th17 and Regulatory T Cells in Mediating and Restraining Inflammation. *Cell* 140: 845–858.

Luo G., Li X., Zhang G., Wu C., Tang Z. & Liu L. 2017. Novel SERMs based on 3-aryl-4-aryloxy-2H-chromen-2-one skeleton - A possible way to dual ER a / VEGFR-2 ligands for treatment of breast cancer. *European Journal of Medicinal Chemistry* 140: 252–273.

Maldonado A.G., Doucet J.P., Petitjean M. & Fan B.T. 2006. Molecular similarity and diversity in chemoinformatics: From theory to applications. *Molecular Diversity* 10: 39–79.

Matos M.J., Terán C., Pérez-Castillo Y., Uriarte E., Santana L. & Viña D. 2011a. Synthesis and Study of a Series of 3-Arylcoumarins as Potent and Selective Monoamine Oxidase B Inhibitors. *Journal of Medicinal Chemistry* 54: 7127–7137.

Matos M.J., Vazquez-Rodriguez S., Uriarte E., Santana L. & Viña D. 2011b. MAO inhibitory activity modulation: 3-Phenylcoumarins versus 3-benzoylcoumarins. *Bioorganic and Medicinal Chemistry Letters* 21: 4224–4227.

Matos M.J., Vilar S., Gonzalez-Franco R.M., Uriarte E., Santana L., Friedman C., Tatonetti N.P., Viña D. & Fontenla J.A. 2013a. Novel (coumarin-3-yl)carbamates as selective MAO-B inhibitors: Synthesis, in vitro and in vivo assays, theoretical evaluation of ADME properties and docking study. *European Journal of Medicinal Chemistry* 63: 151–161.

Matos M.J., Viña D., Janeiro P., Borges F., Santana L. & Uriarte E. 2010. New halogenated 3-phenylcoumarins as potent and selective MAO-B inhibitors. *Bioorganic and Medicinal Chemistry Letters* 20: 5157–5160.

Matos M.J., Viña D., Picciau C., Orallo F., Santana L. & Uriarte E. 2009a. Synthesis and evaluation of 6-methyl-3-phenylcoumarins as potent and selective MAO-B inhibitors. *Bioorganic and Medicinal Chemistry Letters* 19: 5053–5055.

Matos M.J., Viña D., Quezada E., Picciau C., Delogu G., Orallo F., Santana L. & Uriarte E. 2009b. A new series of 3-phenylcoumarins as potent and selective MAO-B inhibitors. *Bioorganic and Medicinal Chemistry Letters* 19: 3268–3270.

Matos M.J., Viña D., Vazquez-Rodriguez S., Uriarte E. & Santana L. 2013b. Focusing on New Monoamine Oxidase Inhibitors: Differently Substituted Coumarins As An Interesting Scaffold. *Current Topics in Medicinal Chemistry* 12: 2210–2239.

Mattsson C., Svensson P. & Sonesson C. 2014. A novel series of 6-substituted 3-(pyrrolidin-1-ylmethyl)chromen-2-ones as selective monoamine oxidase (MAO) A inhibitors. *European Journal of Medicinal Chemistry* 73: 177–186.

Mazimba O. 2017. Umbelliferone: Sources, chemistry and bioactivities review. *Bulletin of Faculty of Pharmacy, Cairo University* 55: 223–232.

Mazumdar M., Fournier D., Zhu D.-W., Cadot C., Poirier D. & Lin S.-X. 2009. Binary and ternary crystal structure analyses of a novel inhibitor with 17β-HSD type 1: a lead compound for breast cancer therapy. *The Biochemical Journal* 424: 357–366.

McKie J.A., Bhagwat S.S., Brady H., Doubleday M., Gayo L., Hickman M., Jalluri R.K., Khammungkhune S., Kois A., Mortensen D., Richard N., Sapienza J., Shevlin G., Stein B. & Sutherland M. 2004. Lead identification of a potent benzopyranone selective estrogen receptor modulator. *Bioorganic and Medicinal Chemistry Letters* 14: 3407–3410.

Meng X.-Y., Zhang H.-X., Mezei M. & Cui M. 2011. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Current Computer-Aided Drug Design* 7: 146–157.

Merritt E.A. & Bacon D.J. 1997. Raster3D: Photorealistic Molecular Graphics. *Methods in Enzymology* 277: 505–524.

Metz C.E. 1978. Basic principles of ROC analysis. *Seminars in Nuclear Medicine* 8: 283–298.

Meyer J.H., Ginovart N., Boovariwala A., Sagrati S., Hussey D., Garcia A., Young T., Praschak-Rieder N., Wilson A.A. & Houle S. 2006. Elevated Monoamine Oxidase A Levels in the Brain - An Explanation for the Monoamine Imbalance of Major Depression. *Archives of General Psychiatry* 63: 1209–1216.

84

Miley M.J., Zielinska A.K., Keenan J.E., Bratton S.M., Radominska-Pandya A. & Redinbo M.R. 2007. Crystal Structure of the Cofactor-Binding Domain of the Human Phase II Drug-Metabolism Enzyme UDP-Glucuronosyltransferase 2B7. *Journal of Molecular Biology* 369: 498–511.

Mitoma J. & Ito A. 1992. Mitochondrial Targeting Located at Its Carboxy Signal Terminus of Rat Liver Monoamine Oxidase B Is Targeting Signal of Monoamine Oxidase B. *Journal of Biochemistry* 24: 20–24.

Mobley D.L. & Dill K.A. 2009. Binding of Small-Molecule Ligands to Proteins: 'What You See' Is Not Always 'What You Get'. *Structure* 17: 489–498.

Modolo L. V., Li L., Pan H., Blount J.W., Dixon R.A. & Wang X. 2009. Crystal Structures of Glycosyltransferase UGT78G1 Reveal the Molecular Basis for Glycosylation and Deglycosylation of (Iso)flavonoids. *Journal of Molecular Biology* 392: 1292–1302.

Mosselman S., Polman J. & Dijkema R. 1996. ERβ: identification and characterization of a novel human estrogen receptor. *FEBS Letters* 392: 49–53.

Murata K. & Wolf M. 2018. Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. *BBA - General Subjects* 1862: 324–334.

Musa M., Cooperwood J. & Khan M.O. 2008. A Review of Coumarin Derivatives in Pharmacotherapy of Breast Cancer. *Current Medicinal Chemistry* 15: 2664–2679.

Nagy N., Kuipers H.F., Frymoyer A.R., Ishak H.D., Bollyky J.B., Wight T.N. & Bollyky P.L. 2015. 4-Methylumbelliferone treatment and hyaluronan inhibition as a therapeutic strategy in inflammation, autoimmunity, and cancer. *Frontiers in Immunology* 6: 1–11.

Niel M.B. Van, Fauber B.P., Cartwright M., Gaines S., Killen J.C., René O., Ward S.I., Leon Boenig G. De, Deng Y., Eidenschenk C., Everett C., Gancia E., Ganguli A., Gobbi A., Hawkins J., Johnson A.R., Kiefer J.R., La H., Lockey P., Norman M., Ouyang W., Qin A., Wakes N., Waszkowycz B. & Wong H. 2014. A reversed sulfonamide series of selective RORc inverse agonists. *Bioorganic and Medicinal Chemistry Letters* 24: 5769–5776.

Niinivehmas S.P., Salokas K., Lätti S., Raunio H. & Pentikäinen O.T. 2015. Ultrafast protein structure-based virtual screening with Panther. *Journal of Computer-Aided Molecular Design* 29: 989–1006.

Niinivehmas S.P., Virtanen S.I., Lehtonen J. V., Postila P.A. & Pentikäinen O.T. 2011. Comparison of virtual high-throughput screening methods for the identification of phosphodiesterase-5 inhibitors. *Journal of Chemical Information and Modeling* 51: 1353–1363.

Nishiyama Y., Nakamura M., Misawa T., Nakagomi M., Makishima M., Ishikawa M. & Hashimoto Y. 2014. Structure-activity relationship-guided development of retinoic acid receptor-related orphan receptor gamma (RORγ)-selective inverse agonists with a phenanthridin-6(5H)-one skeleton from a liver X receptor ligand. *Bioorganic and Medicinal Chemistry* 22: 2799–2808.

Offen W., Martinez-Fleites C., Yang M., Kiat-Lim E., Davis B.G., Tarling C.A., Ford C.M., Bowles D.J. & Davies G.J. 2006. Structure of a flavonoid

glucosyltransferase reveals the basis for plant natural product modification. *The EMBO Journal* 25: 1396–1405.

Ou-Yang S., Lu J., Kong X., Liang Z., Luo C. & Jiang H. 2012. Computational drug discovery. *Acta Pharmacologica Sinica* 33: 1131–1140.

Overington J.P., Al-Lazikani B. & Hopkins A.L. 2006. How many drug targets are there? *Nature Reviews Drug Discovery* 5: 993–996.

Patil P.O., Bari S.B., Firke S.D., Deshmukh P.K., Donda S.T. & Patil D.A. 2013. A comprehensive review on synthesis and designing aspects of coumarin derivatives as monoamine oxidase inhibitors for depression and Alzheimer's disease. *Bioorganic and Medicinal Chemistry* 21: 2434–2450.

Petersen H.G. 1995. Accuracy and efficiency of the particle mesh Ewald method. *The Journal of Chemical Physics* 103: 3668–3679.

Petrey D., Chen T.S., Deng L., Garzon J.I., Hwang H., Lasso G., Lee H., Silkov A. & Honig B. 2015. Template-based prediction of protein function. *Current Opinion in Structural Biology* 32: 33–38.

Phillips J.C., Braun R., Wang W., Gumbart J., Tajkhorshid E., Villa E., Chipot C., Skeel R.D., Kalé L. & Schulten K. 2005. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* 26: 1781–1802.

Politi R., Convertino M., Popov K., Dokholyan N. V & Tropsha A. 2016. Docking and Scoring with Target-Specific Pose Classifier Succeeds in Native-Like Pose Identification But Not Binding Affinity Prediction in the CSAR 2014 Benchmark Exercise. *Journal of Chemical Information and Modeling* 56: 1032–1041.

Postila P.A., Ylilauri M. & Pentikäinen O.T. 2011. Full and Partial Agonism of Ionotropic Glutamate Receptors Indicated by Molecular Dynamics Simulations. *Journal of Chemical Information and Modeling* 51: 1037–1047.

Rahikainen T., Häkkinen M.R., Finel M., Pasanen M. & Juvonen R.O. 2013. A high throughput assay for the glucuronidation of 7-hydroxy-4-trifluoromethylcoumarin by recombinant human UDP-glucuronosyltransferases and liver microsomes. *Xenobiotica* 43: 853–861.

Rangaswami S. & Seshadri T.R. 1940. A Note On Certain Constitutional Factors Controlling Visible Fluorescence in Compounds of the Benzo-pyrone Group. *Proceedings of the Indian Academy of Sciences - Section A* 12: 375–380.

Rangaswami S., Seshadri T.R. & Venkateswarlu V. 1941. The Remarkable Fluorescence of Certain Coumarin Derivatives. *Proceedings of the Indian Academy of Sciences - Section A* 13: 316–321.

Raunio H., Kuusisto M., Juvonen R.O. & Pentikäinen O.T. 2015. Modeling of interactions between xenobiotics and cytochrome P450 (CYP) enzymes. *Frontiers in Pharmacology* 6: 123.

René O., Fauber B.P., Boenig G.D.L., Burton B., Eidenschenk C., Everett C., Gobbi A., Hymowitz S.G., Johnson A.R., Kiefer J.R., Liimatta M., Lockey P., Norman M., Ouyang W., Wallweber H.A. & Wong H. 2015. Minor Structural Change to Tertiary Sulfonamide RORC Ligands Led to Opposite Mechanisms of Action. *ACS Medicinal Chemistry Letters* 6: 276–281.

Roche D.B., Brackenridge D.A. & McGuffin L.J. 2015. Proteins and Their Interacting Partners: An Introduction to Protein-Ligand Binding Site

86

Prediction Methods. *International Journal of Molecular Sciences* 16: 29829–29842.

Rodríguez M.J., Saura J., Billett E.E., Finch C.C. & Mahy N. 2001. Cellular localization of monoamine oxidase A and B in human tissues outside of the central nervous system. *Cell and Tissue Research* 304: 215–220.

Rodríguez M.J., Saura J., Finch C.C., Mahy N. & Billett E.E. 2000. Localization of Monoamine Oxidase A and B in Human Pancreas, Thyroid, and Adrenal Glands. *Journal of Histochemistry and Cytochemistry* 48: 147–151.

Rowland A., Miners J.O. & Mackenzie P.I. 2013. The UDP-glucuronosyltransferases: Their role in drug metabolism and detoxification. *International Journal of Biochemistry and Cell Biology* 45: 1121–1132.

Šali A. & Blundell T.L. 1993. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *Journal of Molecular Biology* 234: 779–815.

Sandhu S., Bansal Y., Silakari O. & Bansal G. 2014. Coumarin hybrids as novel therapeutic agents. *Bioorganic and Medicinal Chemistry* 22: 3806–3814.

Santana L., Orallo F., Viña D., Matos Joao Correia P.C., Quezada E., Yañes J., Vilar V. & Uriarte E. 2010. Use of derivates of 6-substituted 3-phenylcoumarins and preparation of new derivates.

Santos R., Ursu O., Gaulton A., Bento A.P., Donadi R.S., Bologa C.G., Karlsson A., Al-Lazikani B., Hersey A., Oprea T.I. & Overington J.P. 2017. A comprehensive map of molecular drug targets. *Nature Reviews Drug Discovery* 16: 19–34.

Sato Y., Nagata M., Tetsuka K., Tamura K., Miyashita A., Kawamura A. & Usui T. 2014. Optimized Methods for Targeted Peptide-Based Quantification of Human Uridine 5'-Diphosphate-Glucuronosyltransferases in Biological Specimens Using Liquid Chromatography-Tandem Mass Spectrometry. *Drug Metabolism and Disposition* 42: 885–889.

Saura J., Bleuel Z., Ulrich J., Mendelowitsch A., Chen K., Shih J.C., Malherbe P., Prada M. Da & Richards J.G. 1996a. Molecular neuroanatomy of human monoamine oxidases A and B revealed by quantitative enzyme radioautography and in situ hybridization histochemistry. *Neuroscience* 70: 755–774.

Saura J., Nadal E., Berg B. Van Den, Vila M., Bombí J.A. & Mahy N. 1996b. Localization of monoamine oxidases in human peripheral tissues. *Life Sciences* 59: 1341–1349.

Sawicki M.W., Erman M., Puranen T., Vihko P. & Ghosh D. 1999. Structure of the ternary complex of human 17β-hydroxysteroid dehydrogenase type 1 with 3-hydroxyestra-1,3,5,7-tetraen-17-one (equilin) and NADP+. *Proceedings of the National Academy of Sciences of the United States of America* 96: 840–845.

Scheepstra M., Leysen S., Almen G.C. van, Miller J.R., Piesvaux J., Kutilek V., Eenennaam H. van, Zhang H., Barr K., Nagpal S., Soisson S.M., Kornienko M., Wiley K., Elsen N., Sharma S., Correll C.C., Trotter B.W., Stelt M. van der, Oubrie A., Ottmann C., Parthasarathy G. & Brunsveld L. 2015. Identification of an allosteric binding site for RORγt inhibition. *Nature Communications* 6: 8833.

Schnaitman C., Erwin V.G. & Greenawalt J.W. 1967. The submitochondrial localization of monoamine oxidase. *The Journal of Cell Biology* 32: 719 LP-735.

Schönberg A. & Sina A. 1948. Xanthotoxin from the Fruits of *Ammi majus* L. *Nature* 161: 481–482.

Secky L., Svoboda M., Klameth L., Bajna E., Hamilton G., Zeillinger R., Jäger W. & Thalhammer T. 2013. The Sulfatase Pathway for Estrogen Formation: Targets for the Treatment and Diagnosis of Hormone-Associated Tumors. *Journal of Drug Delivery* 2013: 1–13.

Serra S., Ferino G., Matos M.J., Vázquez-Rodríguez S., Delogu G., Viña D., Cadoni E., Santana L. & Uriarte E. 2012. Hydroxycoumarins as selective MAO-B inhibitors. *Bioorganic and Medicinal Chemistry Letters* 22: 258–261.

Shang Y. & Brown M. 2002. Molecular Determinants for the Tissue Specificity of SERMs. *Science* 295: 2465–2468.

Sherman W.R. & Robins E. 1968. Fluorescence of substituted 7-hydroxycoumarins. *Analytical Chemistry* 40: 803–805.

Shiau A.K., Barstad D., Loria P.M., Cheng L., Kushner P.J., Agard D.A. & Greene G.L. 1998. The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen. *Cell* 95: 927–937.

Shih J.C. & Chen K. 2004. Regulation of MAO-A and MAO-B Gene Expression. *Current Medicinal Chemistry* 11: 1995–2005.

Shih J.C., Chen K. & Ridd M.J. 1999. Monoamine Oxidase: From Genes to Behavior. *Annual Review of Neuroscience* 22: 197–217.

Shulman K.I., Herrmann N. & Walker S.E. 2013. Current Place of Monoamine Oxidase Inhibitors in the Treatment of Depression. *CNS Drugs* 27: 789–797.

Simpson E.R. 2003. Sources of estrogen and their importance. *Journal of Steroid Biochemistry and Molecular Biology* 86: 225–230.

Simpson E.R. & Davis S.R. 2001. Minireview: Aromatase and the Regulation of Estrogen Biosynthesis–Some New Perspectives. *Endocrinology* 142: 4589–4594.

Simpson E., Rubin G., Clyne C., Robertson K., O'Donnell L., Jones M. & Davis S. 2000. The Role of Local Estrogen Biosynthesis in Males and Females. *Trends in Endocrinology and Metabolism* 11: 184–188.

Solt L.A., Banerjee S., Campbell S., Kamenecka T.M. & Burris T.P. 2015. ROR Inverse Agonist Suppresses Insulitis and Prevents Hyperglycemia in a Mouse Model of Type 1 Diabetes. *Endocrinology* 156: 869–881.

Solt L.A. & Burris T.P. 2012. Action of RORs and Their Ligands in (Patho)physiology. *Trends in Endocrinology and Metabolism* 23: 619–627.

Solt L.A., Kumar N., He Y., Kamenecka T.M., Griffin P.R. & Burris T.P. 2012. Identification of a Selective RORγ Ligand That Suppresses TH17 Cells and Stimulates T Regulatory Cells. *ACS Chemical Biology* 7: 1515–1519.

Solt L.A., Kumar N., Nuhant P., Wang Y., Lauer J.L., Liu J., Istrate M.A., Kamenecka T.M., Roush W.R., Vidović D., Schürer S.C., Xu J., Wagoner G., Drew P.D., Griffin P.R. & Burris T.P. 2011. Suppression of TH17

differentiation and autoimmunity by a synthetic ROR ligand. *Nature* 472: 491–494.

Son S.-Y., Ma J., Kondou Y., Yoshimura M., Yamashita E. & Tsukihara T. 2008. Structure of human monoamine oxidase A at 2.2-A resolution: the control of opening the entry for substrates/inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* 105: 5739–5744.

Song C.M., Lim S.J. & Tong J.C. 2009. Recent advances in computer-aided drug design. *Briefings in Bioinformatics* 10: 579–591.

Stanway S.J., Purohit A., Woo L.W.L., Sufi S., Vigushin D., Ward R., Wilson R.H., Stanczyk F.Z., Dobbs N., Kulinskaya E., Elliott M., Potter B.V.L., Reed M.J. & Coombes R.C. 2006. Phase I Study of STX 64 (667 Coumate) in Breast Cancer Patients: The First Study of a Steroid Sulfatase Inhibitor. *Clinical Cancer Research* 12: 1585–1592.

Starčević Š., Brožič P., Turk S., Cesar J., Lanišnik Rižner T. & Gobec S. 2011. Synthesis and Biological Bvaluation of (6- and 7-Phenyl) Coumarin Derivatives as Selective Nonsteroidal Inhibitors of 17β-hydroxysteroid Dehydrogenase Type 1. *Journal of Medicinal Chemistry* 54: 248–261.

Stefanachi A., Leonetti F., Pisani L., Farmaco F., Moro A., Orabona E. & Bari I.-. 2018. Coumarin: A Natural, Privileged and Versatile Scaffold for Bioactive Compounds. *Molecules* 23: 250.

Strassburg C.P., Strassburg a, Nguyen N., Li Q., Manns M.P. & Tukey R.H. 1999. Regulation and function of family 1 and family 2 UDP-glucuronosyltransferase genes (UGT1A, UGT2B) in human oesophagus. *The Biochemical journal* 338 (Pt 2: 489–498.

Tadic A., Rujescu D., Szegedi A., Giegling I., Singer P., Möller H.-J. & Dahmen N. 2003. Association of a MAOA Gene Variant with Generalized Anxiety Disorder, but not With Panic Disorder or Major Depression. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* 117B: 1–6.

The UniProt Consortium. 2017. UniProt: The universal protein knowledgebase. *Nucleic Acids Research* 45: D158–D169.

Tipton K.F., Boyce S., O'Sullivan J., Davey G.P. & Healy J. 2004. Monoamine Oxidases: Certainties and Uncertainties. *Current Medicinal Chemistry* 11: 1965–1982.

Tong J., Meyer J.H., Furukawa Y., Boileau I., Chang L.-J., Wilson A.A. & Houle S. 2013. Distribution of Monoamine Oxidase Proteins in Human Brain: Implications for Brain Imaging Studies. *Journal of Cerebral Blood Flow and Metabolism* 33: 863–871.

Toyama H., Nakamura M., Hashimoto Y. & Fujii S. 2015. Design and synthesis of novel ROR inverse agonists with a dibenzosilole scaffold as a hydrophobic core structure. *Bioorganic and Medicinal Chemistry* 23: 2982–2988.

Tripathi S.P., Prajapati R., Verma N. & Sangamwar A.T. 2016. Predicting substrate selectivity between UGT1A9 and UGT1A10 using molecular modelling and molecular dynamics approach. *Molecular Simulation* 42: 270–288.

Tversky A. 1977. Features of similarity. *Psychological Review* 84: 327–352.

Upadhyay A.K., Borbat P.P., Wang J., Freed J.H. & Edmondson D.E. 2008. Determination of the Oligomeric States of Human and Rat Monoamine Oxidases in the Outer Mitochondrial Membrane and Octyl β-D-Glucopyranoside Micelles Using Pulsed Dipolar Electron Spin Resonance Spectroscopy. *Biochemistry* 47: 1554–1566.

Vaidya A., Jain S., Jain S., Jain A.K. & Agrawal R.K. 2014. Quantitative Structure-Activity Relationships: A Novel Approach of Drug Design and Discovery. *Journal of Pharmaceutical Sciences and Pharmacology* 1: 219–232.

Vainio M.J., Puranen J.S. & Johnson M.S. 2009. ShaEP : Molecular Overlay Based on Shape and Electrostatic Potential. *Journal of Chemical Information and Modeling* 49: 492–502.

Vanyukov M.M., Maher B.S., Devlin B., Tarter R.E., Kirillova G.P., Yu L.M. & Ferrell R.E. 2004. Haplotypes of the Monoamine Oxidase Genes and the Risk for Substance Use Disorders. *American Journal of Medicinal Genetics Part B: Neuropsychiatric Genetics* 125B: 120–125.

Vedani A., Dobler M. & Lill M.A. 2005. Combining Protein Modeling and 6D-QSAR. Simulating the Binding of Structurally Diverse Ligands to the Estrogen Receptor. *Journal of Medicinal Chemistry* 48: 3700–3703.

Verhoef T.I., Redekop W.K., Daly A.K., Schie R.M.F. Van, Boer A. De & Maitland-Van Der Zee A.H. 2014. Pharmacogenetic-guided dosing of coumarin anticoagulants: Algorithms for warfarin, acenocoumarol and phenprocoumon. *British Journal of Clinical Pharmacology* 77: 626–641.

Vihko P., Härkönen P., Soronen P., Törn S., Herrala A., Kurkela R., Pulkka A., Oduwole O. & Isomaa V. 2004. 17β-Hydroxysteroid dehydrogenases - Their role in pathophysiology. *Molecular and Cellular Endocrinology* 215: 83–88.

Vilar S., Quezada E., Santana L., Uriarte E., Yánez M., Fraiz N., Alcaide C., Cano E. & Orallo F. 2006. Design, synthesis, and vasorelaxant and platelet antiaggregatory activities of coumarin-resveratrol hybrids. *Bioorganic and Medicinal Chemistry Letters* 16: 257–261.

Viña D., Matos M.J., Ferino G., Cadoni E., Laguna R., Borges F., Uriarte E. & Santana L. 2012a. 8-Substituted 3-Arylcoumarins as Potent and Selective MAO-B Inhibitors: Synthesis, Pharmacological Evaluation, and Docking Studies. *ChemMedChem* 7: 464–470.

Viña D., Matos M.J., Yáñez M., Santana L. & Uriarte E. 2012b. 3-Substituted coumarins as dual inhibitors of AChE and MAO for the treatment of Alzheimer's disease. *Medicinal Chemistry Communications* 3: 213–218.

Virtanen S.I. & Pentikäinen O.T. 2010. Efficient Virtual Screening Using Multiple Protein Conformations Described as Negative Images of the Ligand-Binding Site. *Journal of Chemical Information and Modeling* 50: 1005–1011.

Wang Y., Li L., Zhang B., Xing J., Chen S., Wan W., Song Y., Jiang H., Jiang H., Luo C. & Zheng M. 2017. Discovery of Novel Disruptor of Silencing Telomeric 1-Like (DOT1L) Inhibitors using a Target-Specific Scoring Function for the (S)-Adenosyl-l-methionine (SAM)-Dependent Methyltransferase Family. *Journal of Medicinal Chemistry* 60: 2026–2036.

Wang J.M., Wolf R.M., Caldwell J.W., Kollman P. a & Case D. a. 2004. Development and Testing of a General Amber Force Field. *Journal of Computational Chemistry* 25: 1157–1174.

Warren G.L., Andrews C.W., Capelli A.M., Clarke B., LaLonde J., Lambert M.H., Lindvall M., Nevins N., Semus S.F., Senger S., Tedesco G., Wall I.D., Woolven J.M., Peishoff C.E. & Head M.S. 2006. A Critical Assessment of Docking Programs and Scoring Functions. *Journal of Medicinal Chemistry* 49: 5912–5931.

Watts K.S., Dalal P., Murphy R.B., Sherman W., Friesner R.A. & Shelley J.C. 2010. ConfGen: A Conformational Search Method for Efficient Generation of Bioactive Conformers. *Journal of Chemical Information and Modeling* 50: 534–546.

Wayment H.K., Schenk J.O. & Sorg B.A. 2001. Characterization of Extracellular Dopamine Clearance in the Medial Prefrontal Cortex: Role of Monoamine Uptake and Monoamine Oxidase Inhibition. *Journal of Neuroscience* 21: 35–44.

Welsch M.E., Snyder S.A. & Stockwell B.R. 2010. Privileged scaffolds for library design and drug discovery. *Current Opinion in Chemical Biology* 14: 347–361.

Wermuth C.G., Ganellin C.R., Lindberg P. & Mitscher L.A. 1998. Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure and Applied Chemistry* 70: 1129–1143.

Willett P., Barnard J.M. & Downs G.M. 1998. Chemical Similarity Searching. *Journal of Chemical Information and Computer Sciences* 38: 983–996.

Woo L.W.L., Howarth N.M., Purohit A., Hejaz H.A.M., Reed M.J. & Potter B.V.L. 1998. Steroidal and Nonsteroidal Sulfamates as Potent Inhibitors of Steroid Sulfatase. *Journal of Medicinal Chemistry* 41: 1068–1083.

Wrange O. & Gustafsson J.A. 1978. Separation of the Hormone- and DNA-binding Sites of the Hepatic Glucocorticoid Receptor by Means of Proteolysis. *Journal of Biological Chemistry* 253: 856–865.

Wu P., Guo Y., Jia F. & Wang X. 2015. The Effects of Armillarisin A on Serum IL-1β and IL-4 and in Treating Ulcerative Colitis. *Cell Biochemistry and Biophysics* 72: 103–106.

Wurtz J.-M., Bourguet W., Renaud J.-P., Vivat V., Chambon P., Moras D. & Gronemeyer H. 1996. A canonical structure for the ligand-binding domain of nuclear receptors. *Nature Structural Biology* 3: 87–94.

Yang S.Y. 2010. Pharmacophore modeling and applications in drug discovery: Challenges and recent advances. *Drug Discovery Today* 15: 444–450.

Yao R., Fu Y., Li S., Tu L., Zeng X. & Kuang N. 2011. Regulatory effect of daphnetin, a coumarin extracted from *Daphne odora*, on the balance of Treg and Th17 in collagen-induced arthritis. *European Journal of Pharmacology* 670: 286–294.

Ylilauri M. & Pentikäinen O.T. 2012. Structural Mechanism of N-Methyl-D-Aspartate Receptor Type 1 Partial Agonism. *PLoS ONE* 7: e47604.

Ylilauri M. & Pentikäinen O.T. 2013. MMGBSA as a Tool to Understand the Binding Affinities of Filamin-Peptide Interactions. *Journal of Chemical Information and Modeling* 53: 2626–2633.

Youdim M.B.H., Edmondson D. & Tipton K.F. 2006. The therapeutic potential of monoamine oxidase inhibitors. *Nature Reviews Neuroscience* 7: 295–309.

Yu D.J., Hu J., Yang J., Shen H. Bin, Tang J. & Yang J.Y. 2013. Designing Template-Free Predictor for Targeting Protein-Ligand Binding Sites with Classifier Ensemble and Spatial Clustering. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 10: 994–1008.

Zhao C., Dahlman-Wright K. & Gustafsson J.-Å. 2008. Estrogen receptor β: an overview and update. *Nuclear Receptor Signaling* 6: 1–10.

Zheng M., Zhao J., Cui C., Fu Z., Li X., Liu X., Ding X., Tan X., Li F., Luo X., Chen K. & Jiang H. 2018. Computational chemical biology and drug design: Facilitating protein structure, function, and modulation studies. *Medicinal Research Reviews*: 1–37.

Åqvist J. 1990. Ion-water interaction potentials derived from free energy perturbation simulations. *The Journal of Physical Chemistry* 94: 8021–8024.

ORIGINAL PAPERS

I

STRUCTURE-ACTIVITY RELATIONSHIP ANALYSIS OF 3-PHENYLCOUMARIN-BASED MONOAMINE OXIDASE B INHIBITORS

by

Sanna Rauhamäki, Pekka A. Postila, Sanna Niinivehmas, Sami Kortet, Emmi Schildt, Mira Pasanen, Elangovan Manivannan, Mira Ahinko, Pasi Koskimies, Niina Nyberg, Pasi Huuskonen, Elina Multamäki, Markku Pasanen, Risto O. Juvonen, Hannu Raunio, Juhani Huuskonen & Olli T. Pentikäinen 2018

Frontiers in Chemistry 6: 41.

Check for
updates

Structure-Activity Relationship Analysis of 3-Phenylcoumarin-Based Monoamine Oxidase B Inhibitors

Sanna Rauhamäki[1], Pekka A. Postila[1], Sanna Niinivehmas[1], Sami Kortet[1,2],
Emmi Schildt[1,2], Mira Pasanen[1], Elangovan Manivannan[1,3], Mira Ahinko[1],
Pasi Koskimies[4], Niina Nyberg[5], Pasi Huuskonen[5], Elina Multamäki[1], Markku Pasanen[5],
Risto O. Juvonen[5], Hannu Raunio[5], Juhani Huuskonen[2]* and Olli T. Pentikäinen[1,6]*

[1] Computational Bioscience Laboratory, Department of Biological and Environmental Science & Nanoscience Center, University of Jyväskylä, Jyväskylä, Finland, [2] Department of Chemistry & Nanoscience Center, University of Jyväskylä, Jyväskylä, Finland, [3] School of Pharmacy, Devi Ahilya University, Madhya Pradesh, India, [4] Forendo Pharma Ltd., Turku, Finland, [5] School of Pharmacy, University of Eastern Finland, Kuopio, Finland, [6] MedChem.fi, Institute of Biomedicine, University of Turku, Turku, Finland

Monoamine oxidase B (MAO-B) catalyzes deamination of monoamines such as neurotransmitters dopamine and norepinephrine. Accordingly, small-molecule MAO-B inhibitors potentially alleviate the symptoms of dopamine-linked neuropathologies such as depression or Parkinson's disease. Coumarin with a functionalized 3-phenyl ring system is a promising scaffold for building potent MAO-B inhibitors. Here, a vast set of 3-phenylcoumarin derivatives was designed using virtual combinatorial chemistry or rationally *de novo* and synthesized using microwave chemistry. The derivatives inhibited the MAO-B at 100 nM—1 μM. The IC$_{50}$ value of the most potent derivative **1** was 56 nM. A docking-based structure-activity relationship analysis summarizes the atom-level determinants of the MAO-B inhibition by the derivatives. Finally, the cross-reactivity of the derivatives was tested against monoamine oxidase A and a specific subset of enzymes linked to estradiol metabolism, known to have coumarin-based inhibitors. Overall, the results indicate that the 3-phenylcoumarins, especially derivative **1**, present unique pharmacological features worth considering in future drug development.

Keywords: 3-phenylcoumarin, monoamine oxidase B (MAO-B), structure-activity relationship (SAR), virtual drug design, Parkinson's disease

INTRODUCTION

During neuronal signaling, neurotransmitters are released from the presynaptic cell into the synaptic cleft, from where they bind into their specific receptors embedded on the postsynaptic membrane. The membrane lipid bilayer, especially its anionic phospholipid constituents, has been suggested to play a role in the small-molecule entry processes with the receptors (Orłowski et al., 2012; Postila et al., 2016; Mokkila et al., 2017). Moreover, to assure that the neurotransmission remains transient, the neurotransmitters are removed quickly from the synaptic cleft via enzymatic degradation and cellular uptake.

When inside the neuron, monoamine neurotransmitters such as norepinephrine and dopamine are either recycled or destined for deactivation through oxidative deamination (RCH_2NHR' + H_2O + O_2 = $RCHO$ + $R'NH_2$ + H_2O_2) by monoamine oxidases A (MAO-A; E.C. 1.4.3.4) and B (MAO-B; E.C. 1.4.3.4). These enzymes are integral monotopic proteins that anchor themselves as dimers onto the mitochondrial outer membrane surface by protruding their α-helical C-termini into the lipid bilayer (**Figure 1A**). Moreover, both subtypes A and B deaminate preferentially their respective substrates to aldehydes: MAO-A catalyzes serotonin, norepinephrine, and to some extent dopamine; and MAO-B catalyzes dopamine, phenethylamine, benzylamine and to a lesser extent norepinephrine (Shih et al., 1999; Edmondson et al., 2005; Gaweska and Fitzpatrick, 2011).

The MAO-B, which is the target of this study, is connected to neurodegenerative disorders such as Alzheimer's disease but also mental disorders such as schizophrenia, anorexia nervosa, depression and attention deficit disorder. In all of these conditions, the involvement of MAO-B in the metabolism of dopamine and other amines is in a key role (Youdim et al., 2006; Carradori and Silvestri, 2015). For instance, due to gliosis associated with Parkinson's disease, increased levels of MAO-B speed up degradation of dopamine in the motor neurons. MAO-B inhibitors decrease the degradation and boost dopamine concentration in the synapse. Thus, instead of introducing more dopamine, the neurotransmitter levels are elevated by inhibiting MAO-B. As a result, MAO-B inhibitors such as selegiline are used in treatment of Parkinson's disease, moreover, their neuroprotective effects can benefit Alzheimer's disease patients (Youdim et al., 2006). Due to these hepatotoxic effects of irreversibly binding MAO inhibitors, reversible inhibitors such as moclobemide were developed (Youdim et al., 2006; Finberg and Rabey, 2016). The MAO inhibitors can exhibit selectivity toward MAO-A (moclobemide) or MAO-B (pargyline, selegiline) or be non-selective (phenelzine, tranylcypromine). The selectivity, which can be lost in high dosages, is important for avoiding MAO-A inhibition related cheese effect (Youdim et al., 2006; Finberg and Rabey, 2016).

A vast amount of different types of MAO inhibitors are described in the literature and for example the ChEMBL database lists inhibition data for thousands of compounds. The specific problem in the development of MAO-specific ligands is that the promising compounds have potential to become active on other amine oxidases such as vascular adhesion protein 1 (Nurminen et al., 2010, 2011). Here, the aim was to probe the MAO-B activity and selectivity effects of different substitutions on the coumarin core by focusing, especially, on the 3-phenylcoumarin (or 3-arylcoumarin). Notably, there exist two X-ray crystal structures with structurally related coumarin analogs in which 3-chlorobenzyloxy groups are attached at the C7-position (**Figures 1B–D**). The studied set

of 3-phenylcoumarin derivatives with different R1-R7 groups (**Figure 1E**) introduced in this study make an important addition to the earlier studies in which the potential of coumarin core, including 61 3-phenylcoumarin derivatives (Matos et al., 2009b, 2010, 2011a,b; Santana et al., 2010; Serra et al., 2012; Viña et al., 2012a,b), to block MAO-A and MAO-B has been explored (Borges et al., 2005; Catto et al., 2006; Matos et al., 2009a, 2010, 2011a; Serra et al., 2012; Ferino et al., 2013; Joao Matos et al., 2013; Patil et al., 2013). The compounds were designed using virtual combinatorial chemistry or rationally *de novo* and binding were probed via molecular docking prior to synthesis or *in vitro* testing.

Initially, 52 derivatives of the 3-phenylcoumarin core were synthesized and tested here for the first time for MAO-B inhibition using a specifically tailored spectrophotometric assay (Supplementary Table S1) (Holt et al., 1997). Next, 24 of the derivatives (**Figure 2**, **Table 1**), producing >70% inhibition at $10\,\mu M$, were selected for further analysis. These derivatives inhibited MAO-B at a ~100 nM to ~1 μM range, while the most potent derivative **1** produces ~50–60 nM inhibition (**Table 1**, **Figure 2**). Finally, the potency of the derivatives for inhibiting estrogen receptor (ER), 17-β-hydroxysteroid dehydrogenase 1 (HSD1), aromatase (CYP19A1), and cytochrome P450 1A2 (CYP1A2), the topics of both our prior (Niinivehmas et al., 2016) and ongoing studies, was also considered. A docking-based structure-activity relationship (SAR) analysis (**Figure 2**) was performed with all of the synthetized 3-phenylcoumarins focusing mainly on the 24 most potent compounds.
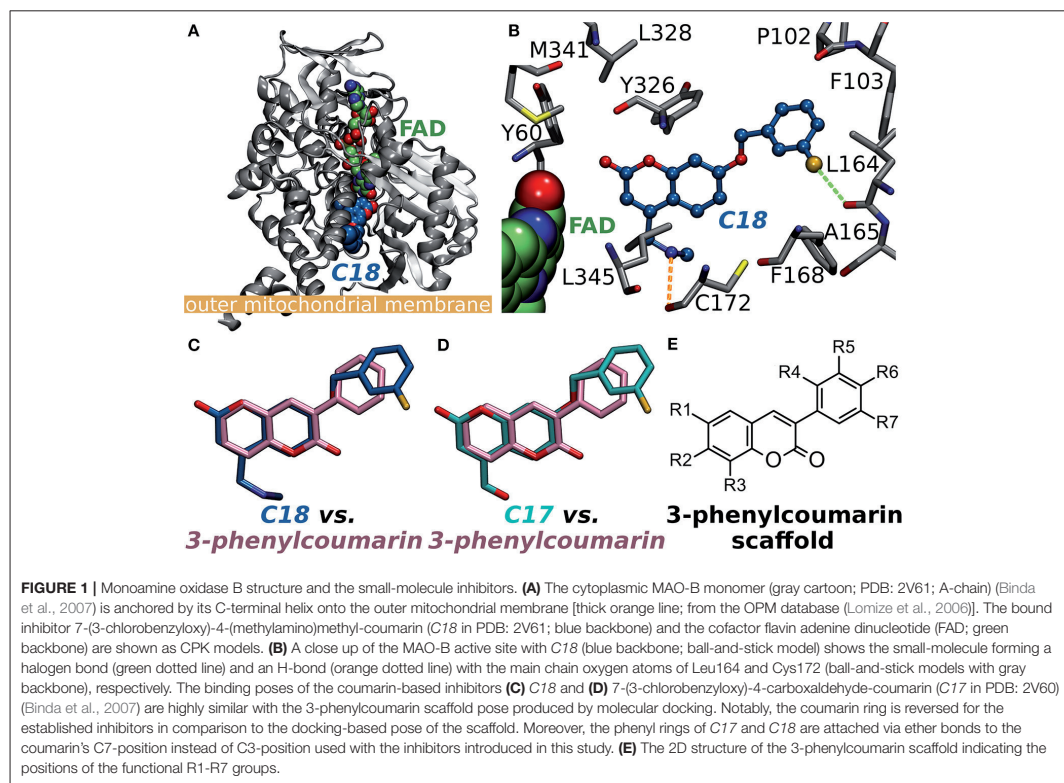
In short, this study explores thoroughly the pharmacological potential of 3-phenylcoumarin (**Figure 1E**) for blocking the MAO-B activity (**Table 1**, Supplementary Table S1) and, furthermore, explains the basis of the inhibitory effect on the atom level.

MATERIALS AND METHODS

Virtual Combinatorial Chemistry

The 3-phenylcoumarin was chosen as the scaffold of interest for building new MAO-B-specific inhibitors (see section The Alignment of the 3-Phenylcoumarin Scaffold at the Active Site). The analogs were designed using virtual combinatorial chemistry or virtual synthesis. In the initial stages, methoxy group was included at the R1 or R2 position (**Figure 1E**) in the coumarin core due to its predicted favorability at the active site. The R4-R7 substituents of the 3-phenyl ring (**Figure 1E**) were designed by combining phenylacetic acid with either 6-methoxycoumarin or 7-methoxycoumarin. The preliminary combinatorial compound library was generated using MAESTRO version 9.3 CombiGlide (CombiGlide, version 2.8, Schrödinger, LLC, New York, NY, USA) and Combinatorial Screening module. The compounds were docked with GLIDE and scored using GlideScore. Some of these derivatives with promising potency and selectivity profile in this study (**8**, **10**, **25**, **37**) were eventually synthesized, albeit using different chemistry (see section Chemical Procedure), and tested *in vitro*. Majority of the final derivatives were designed *de novo* after performing the initial docking simulations with the virtual synthesis products.

FIGURE 1 | Monoamine oxidase B structure and the small-molecule inhibitors. **(A)** The cytoplasmic MAO-B monomer (gray cartoon; PDB: 2V61; A-chain) (Binda et al., 2007) is anchored by its C-terminal helix onto the outer mitochondrial membrane [thick orange line; from the OPM database (Lomize et al., 2006)]. The bound inhibitor 7-(3-chlorobenzyloxy)-4-(methylamino)methyl-coumarin (*C18* in PDB: 2V61; blue backbone) and the cofactor flavin adenine dinucleotide (FAD; green backbone) are shown as CPK models. **(B)** A close up of the MAO-B active site with *C18* (blue backbone; ball-and-stick model) shows the small-molecule forming a halogen bond (green dotted line) and an H-bond (orange dotted line) with the main chain oxygen atoms of Leu164 and Cys172 (ball-and-stick models with gray backbone), respectively. The binding poses of the coumarin-based inhibitors **(C)** *C18* and **(D)** 7-(3-chlorobenzyloxy)-4-carboxaldehyde-coumarin (*C17* in PDB: 2V60) (Binda et al., 2007) are highly similar with the 3-phenylcoumarin scaffold pose produced by molecular docking. Notably, the coumarin ring is reversed for the established inhibitors in comparison to the docking-based pose of the scaffold. Moreover, the phenyl rings of *C17* and *C18* are attached via ether bonds to the coumarin's C7-position instead of C3-position used with the inhibitors introduced in this study. **(E)** The 2D structure of the 3-phenylcoumarin scaffold indicating the positions of the functional R1-R7 groups.
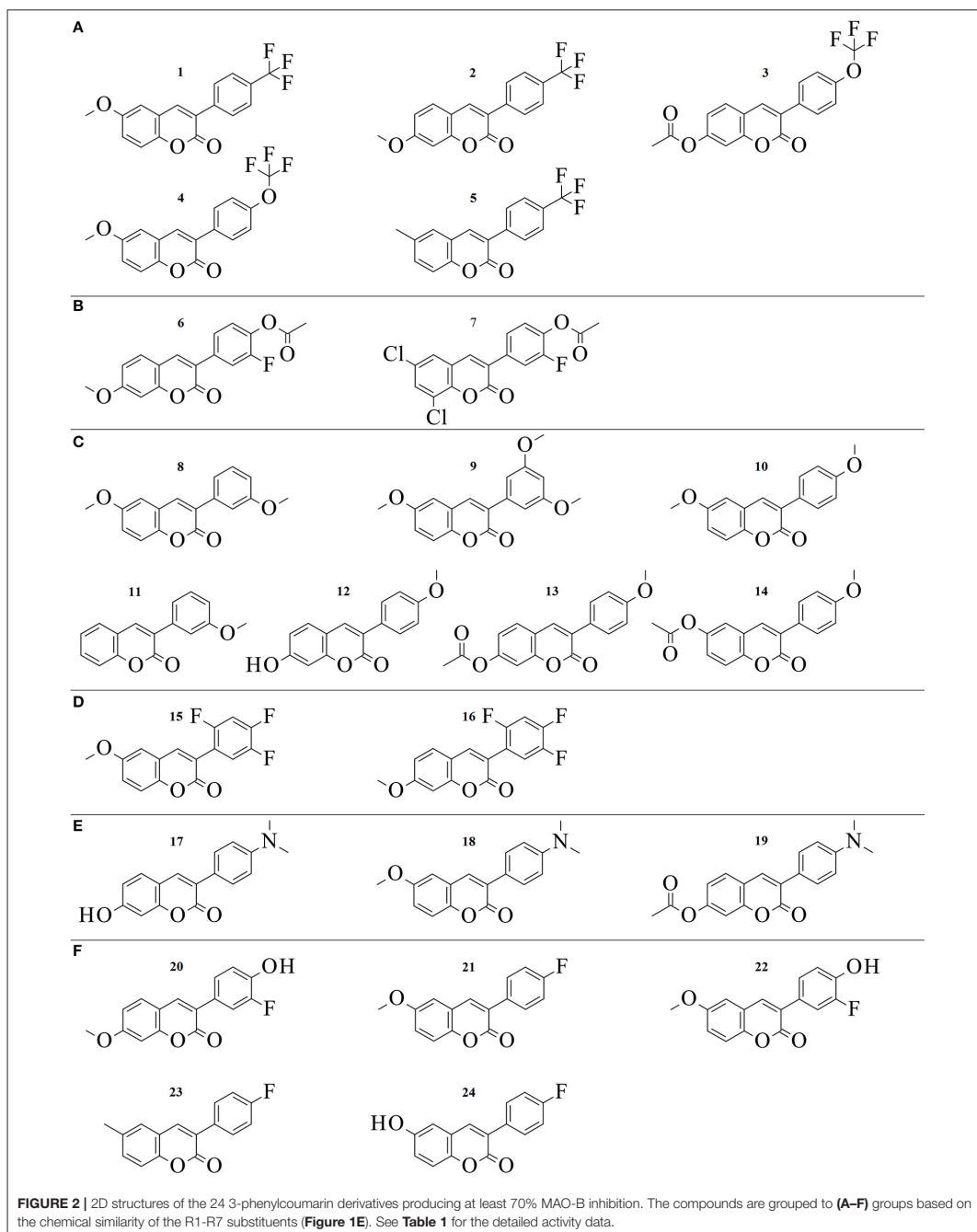
Chemical Procedure

All reactions were carried out using commercial materials (Sigma-Aldrich, Mannheim, Germany) and reagents without further purification unless otherwise noted. Reaction mixtures were heated by the CEM Discover microwave apparatus. All reactions were monitored by thin layer chromatography (TLC) on silica gel plates. ^{1}H NMR and ^{13}C NMR data was recorded on a Bruker Avance 400 MHz spectrometer or Bruker Avance III 300 MHz spectrometer. Chemical shifts are expressed in parts per million values (ppm) and are designated as s (singlet), br s (broad singlet), d (doublet), dd (double doublet), and t (triplet). Coupling constants (*J*) are expressed as values in hertz (Hz). The mass spectra were recorded using Micromass LCT ESI-TOF equipment. Elemental analyses were done with Elementar Vario EL III elemental analyzer. The coumarin derivatives were synthesized using Perkin-Oglialor condensation reaction. The method was developed from the earlier published procedures and transferred to microwave reactor and it was published earlier by authors (Niinivehmas et al., 2016).

A typical procedure: A mixture of salicylaldehyde derivative (2 mmol) and phenyl acetic acid derivative (2.1 mmol), acetic acid anhydride (0.6 ml), and triethylamine (0.36 ml) were placed in a microwave reactor tube and this mixture was heated at 100–170°C with microwave apparatus (100–200 W) for 10–20 min. After cooling, 2 ml of 10% NaHCO$_3$ solution was added and the precipitate was filtered, dried and recrystallized from EtOH/H$_2$O or acetone/H$_2$O mixture. The acetyl group(s) were removed by treating the compound with 2 M MeOH/NaOH(aq) (1:1) solution for 30–60 min at r.t. The solution was acidified with 2 M HCl(aq,) and the precipitate was filtered and recrystallized if needed.

Based on the elemental analysis and/or ^{1}H-NMR the purity of compounds was >95%.

6-methoxy-3-(4-(trifluoromethyl)phenyl)-2H-chromen-2-one (1). Yield: 76%; ^{1}H-NMR (400 MHz, CDCl$_3$) δ: 3.86 (s, 3H, CH$_3$O-), 6.99 (s, 1H, H-5), 7.14 (d, 1H, J^3 = 7.7 Hz, H-7), 7.29 (d, J^3 = 8.9 Hz, H-8), 7.69 (d, 2H, J^3 = 7.9 Hz, H-2', H-6'), 7.58 (m, 3H, H-4, H-3', H-5'); ^{13}C-NMR (100.6 MHz, CDCl$_3$) δ: 55.99, 110.24, 117.73, 119.78, 120.02, 125.51 (q, J^{C-F} = 4 Hz), 127.37, 129.05, 130.85 (q, J^{C-F} = 32 Hz), 138.41, 140.88, 148.33, 156.44, 160.42. HRMS(ESI): calc. for C$_{17}$H$_{11}$F$_3$O$_3$Na$_1$ 343.0558, found 343.0574; elemental anal. for C$_{17}$H$_{11}$F$_3$O$_3$, calc. C% 63.76, H% 3.46, found C% 63.25, H% 3.51.

FIGURE 2 | 2D structures of the 24 3-phenylcoumarin derivatives producing at least 70% MAO-B inhibition. The compounds are grouped to **(A–F)** groups based on the chemical similarity of the R1-R7 substituents (**Figure 1E**). See **Table 1** for the detailed activity data.

TABLE 1 | The activity data on the 24 most potent 3-phenylcoumarin derivatives.

Group	ID	MAO-B inhibition IC$_{50}$ nM	QPlogPo/w	MAO-B inhibition % (10 μM)	MAO-A inhibition % (100 μM)	ER inhibition % (10 μM)	HSD1 inhibition % (1 μM)	CYP1A2 inhibition IC$_{50}$ μM
Control	c	61[1]	2.43[1]	102.00[1]	100.89[2]	106.60[3]	N/A	N/A
A	01	56	4.08	99.53	0.00	N/A	0	124.00
	02	138	4.11	99.58	0.00	N/A	1	N/A
	03	141	3.33	100.44	22.03	N/A	0	280.00
	04	317	4.22	101.96	0.00	N/A	0	7.00
	05	343	4.35	105.33	0.00	1.08	0	171.00
B	06	189	2.47	99.92	0.00	N/A	21	N/A
	07	888	3.36	91.01	0.00	N/A	0	46.00
C	08	231	3.11	111.93	0.00	0	0	2.30
	09	255	3.21	80.21	0.00	N/A	0	84.00
	10	400	3.15	97.57	10.14	N/A	0	15.00
	11	798	3.06	90.33	0.00	0.29	4	1.60
	12	955	2.49	85.89	24.57	91.34	3	170.00
	13	1946	2.41	85.89	2.48	N/A	0	570.00
	14	8476	2.34	75.75	N/A	N/A	1	87.51
D	15	292	3.73	87.16	0.00	0	12	3.00
	16	1433	3.71	77.63	N/A	8.80	33	4.50
E	17	384	2.80	90.14	4.74	N/A	5	35.00
	18	617	3.49	93.86	0.00	0	1	17.00
	19	866	2.79	85.41	0.00	N/A	15	370.00
F	20	391	2.71	100.82	0.00	86.10	46	30.00
	21	433	3.32	88.77	0.00	0	0	3.00
	22	831	2.73	94.86	0.00	55.38	54	1.50
	23	902	3.58	83.49	0.00	0	11	3.00
	24	1058	2.61	89.10	14.18	0	20	3.00

*N/A = not available. Controls: [1]pargyline, [2]clorgyline, [3]kit control. The compounds are grouped (A–F) based on the chemical similarity of the R1–R7 substituents (**Figure 1E**).*

SCHEME 1 | The synthesis of 3-phenylcoumarin analogs.
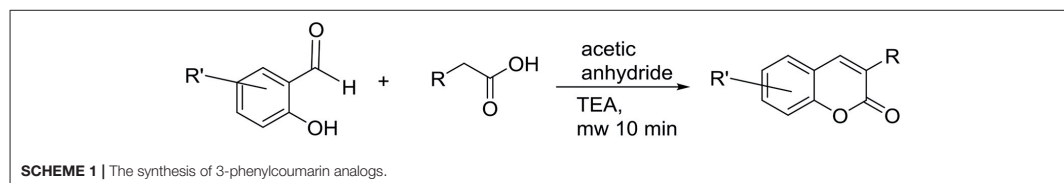
6-methoxy-3-(4-(trifluoromethyl)phenyl)-2H-chromen-2-one (2). Yield: 80%; ^{1}H-NMR (300 MHz, d^6-DSMO) δ: 3.88 (s, 3H, CH$_3$O-), 6.99 (s, 1H, J^3 = 8.7 Hz, J^4 = 2.4 Hz, H-6), 7.03 (d, 1H, J^4 = 2.4 Hz, H-7), 7.71 (d, 1H, J^3 = 8.6 Hz, H-8), 7.79 (d, 2H, J^3 = 8.3 Hz, H-2', H-6'), 7.93 (d, 2H, H-3', H-5'), 8.32 (s, 1H, H-4); ^{13}C-NMR (75.5 MHz, d^6-DSMO) δ: 55.97, 100.25, 112.80, 121.57, 122.38, 120.02, 124.97 (q, J^{C-F} = 4 Hz), 128.29 (q, J^{C-F} = 32 Hz), 128.97, 129.99, 139.01, 142.10, 155.09, 159.62, 162.82. HRMS(ESI) calc for C$_{17}$H$_{11}$F$_3$O$_3$Na$_1$ [M + Na]$^+$: 343.05525, found 343.05610.

2-oxo-3-(4-(trifluoromethoxy)phenyl)-2H-chromen-7-yl acetate (3). (Dobelmann-Mara et al., 2017) Yield: 54%; %; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 2.27 (s, 3H, CH$_3$C(O)O-), 7.20 (dd, 1H, J^3 = Hz, J^4 = Hz, H-6), 7.33 (d, 1H, J^4 = Hz, H-8), 7.47 (d, 2H, J^3 = Hz, H-3', H-5'), 7.81 (d, 1H, J^3 = 8.4 Hz, H-5), 8.32 (s, 1H, H-4); ^{13}C-NMR (100 MHz, d^6-DMSO) δ: 20.86 109.74, 117.23, 118.88, 120.75, 124.84, 129.42, 129.60, 130.52, 133.85, 140.73, 148.35, 152.90, 153.55, 159.40, 168.78; HRMS(ESI) calc. for C$_{18}$H$_{11}$F$_3$O$_5$Na$_1$ [M + Na]$^+$ 387.0457, found 387.0481.

6-methoxy-3-(4-(trifluoromethoxy)phenyl)-2H-chromen-2-one (4). Yield: 52%; ^{1}H-NMR (400 MHz, CDCl$_3$) δ: 3.86 (s, 3H, CH$_3$O-), 6.98 (d, 1H, J^4 = 3 Hz, H-5), 7.12 (dd, J^3 = 9.1 Hz, J^4 = 3 Hz, H-7), 7.27-7.30 (m, 3H, H-8, H-3', H-5'), 7.74 (d, 2H, J^3 = 8.9 Hz, H-2', H-6'); 7.77 (s, 1H, H-4); ^{13}C-NMR (100 MHz, CDCl$_3$) δ: 55.99, 110.16, 117.70, 119.68, 119.92, 120.97, 127.41, 130.26, 133.51, 140.20, 148.21, 149.67, 156.41, 160.65. HRMS(ESI) calc for C$_{17}$H$_{11}$F$_3$O$_4$Na$_1$ [M + Na]$^+$: 359.05071, found 359.05260. elemental anal. for C$_{17}$H$_{11}$F$_3$O$_4$·0.5H$_2$O, calc. C% 59.14, H% 3.50, found C% 58.99, H% 3.25.

6-methyl-3-(4-(trifluoromethyl)phenyl)-2H-chromen-2-one (5). Yield: 54%; ^{1}H-NMR (400 MHz, CDCl$_3$) δ: 7.27 (d, 1H, J^4 = 2.2 Hz, H-5), 7.35-7.38 (m, 2H, H-7, H-8), 7.70 (d, J^3 = 8.2 Hz, H-2', H-6'), 7.82 (m, 3H, H-4, H-3', H-5'); ^{13}C-NMR (100 MHz, CDCl$_3$) δ: 20.92, 116.46, 119.22, 122.80, 125.53 (q, J^{C-F} = 4 Hz), 126.98, 128.07, 129.05, 130.80 (q, J^{C-F} = 33 Hz), 133.29, 134.62, 138.50, 141.08, 152.04 160.53; HRMS(ESI) calc for C$_{17}$H$_{12}$Cl$_2$O$_4$Na$_1$ [M + Na]$^+$: 373.0005, found 372.9998.

2-fluoro-4-(7-methoxy-2-oxo-2H-chromen-3-yl)phenyl acetate (6). Yield 75%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 2.35 (s, 3H, CH$_3$C(O)O-Ph), 3.88 (s, 3H, CH$_3$O-Ph), 6.99 (dd, 1H, J^3 = 8.6 Hz, J^4 = 2.4 Hz, H-6), 7.05 (d, 1H, J^4 = 2.4 Hz, H-8), 7.37 (dd, J^3 = 9.3 Hz, J^{H-F} = 8.3 Hz, H-5'), 7.62 (ddd, 1H, J^3 = 8.5 Hz, J^4 = 2.1 Hz, J^{H-F} = 0.8 Hz, H-6'), 7.68 (d, J = 8.6 Hz, 1H, H-5), 7.74 (dd, J^{H-F} = 12.1 Hz, J^4 = 2.0 Hz, H-3'), 8.31 (s, 1H, H-4); ^{13}C-NMR (100 MHz, d^6-DMSO) δ: 20.19, 55.97, 100.25, 112.79, 116.35 (d, J^{C-F} = 20.3 Hz), 121.02, 121.03, 123.83, 124.79 (d, J^{C-F} = 3.2 Hz), 129.86, 134.24 (d, J^{C-F} = 7.7 Hz), 137.20 (d, J^{C-F} = 13.1 Hz), 141.55, 153.00 (J^{C-F} = 246.1 Hz), 154.92, 159.65, 162.69, 168.19. HRMS(ESI) calc for C$_{18}$H$_{13}$F$_1$O$_5$Na$_1$ [M + Na]$^+$: 351.06447, found 351.06240; elemental anal. for C$_{18}$H$_{13}$F$_1$O$_5$ C% 65.85, H% 3.99, found C% 65.28 H% 4.02.

4-(6,8-dichloro-2-oxo-2H-chromen-3-yl)-2-fluorophenyl acetate (7). Yield 58%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 2.36 (s, 3H, CH$_3$C(O)O-), 7.43 (dd, J^3 = 9.3 Hz, J^{H-F} = 8.3 Hz, H-5'), 7.67 (ddd, 1H, J^3 = 8.4 Hz, J^4 = 2.1 Hz, J^{H-F} = 0.8 Hz, H-6'), 7.74 (dd, J^{H-F} = 11.8 Hz, J^4 = 2.0 Hz, H-3'), 7.84 (d, 1H, J^4 = 2.4 Hz, H-7), 7.97 (d, 1H, J^4 = 2.4 Hz, H-5), 8.32 (s, 1H, H-4); ^{13}C-NMR (100 MHz, d^6-DMSO) δ: 20.72, 117.23 (d, J^{C-F} = 21 Hz), 121.13, 122.17, 124.65, 125.74 (d, J^{C-F} = 3.3 Hz), 127.29, 128.80, 131.47, 133.70 (d, J^{C-F} = 7.7 Hz), 138.50 (d, J^{C-F} = 12.9 Hz), 140.12, 147.94, 152.30, 154.75. 158.73; HRMS(ESI): calc. for C$_{17}$H$_9$Cl$_2$F$_1$O$_4$Na$_1$ [M + Na]$^+$: 388.9760, found 388. 9762.

6-methoxy-3-(3-methoxyphenyl)-2H-chromen-2-one (8). Yield 78%; ^{1}H-NMR (400 MHz, CDCl$_3$) δ: 3.85 (s, 3H, CH$_3$O-Ph), 3.86 (s, 3H, CH$_3$O-Ph), 6.93-6.97 (m, 2H, H-4', H-5), 7.10 (dd, 1H, J^3 = 9.0, Hz, J^4 = 1.9 Hz, H-7), 7.25-7.29 (m, 3H, H-8, H-2', H-6'), 7.35 (t, 1H, J^3 = 8.2 Hz, H-5'), 7.76 (s, 1H, H-4); ^{13}C-NMR (100 MHz, CDCl$_3$) δ: 55.69, 56.18, 110.28, 114.57, 114.88, 117.78, 119.55, 120.28, 121.26, 128.80, 129.79, 136.43, 140.13, 148.34, 156.47, 159.88, 160.91; HRMS(ESI): calc. for C$_{17}$H$_{14}$O$_4$Na$_1$ [M + Na]$^+$: 305.07898, found 305.07950; elemental anal. for C$_{14}$H$_{14}$O$_4$ calc. C% 72.33, H% 5.00, found C% 72.41, H% 4.88.

3-(3,5-dimethoxyphenyl)-6-methoxy-2H-chromen-2-one (9). (Vilar et al., 2006) Yield 59%; ^{1}H-NMR (400 MHz,

d^6-DMSO) δ: 3.79 (s, 6H, CH$_3$O-Ph), 3.82 (s, 3H, CH$_3$O-Ph), 6.56 (t, 1H, J^4 = 2.3 Hz, H-4'), 6.89 (d, 2H, J^4 = 2.3 Hz, H-2', H-6'), 7.20 (dd, 1H, J^3 = 9.0 Hz, J^4 = 3.0 Hz, H-7), 7.31 (d, 1H, J^4 = 3.0 Hz, H-5), 7.36 (d, 1H, J^3 = 9.0 Hz, H-8), 8.23 (s, 1H, H-4); ^{13}C-NMR (100 MHz, d^6-DMSO) δ: 55.30, 55.66, 100.48, 106.71, 110.69, 116.90, 119.36, 119.78, 126.75, 136.44, 140.66, 147.33, 155.62, 159.53, 160.16; HRMS(ESI): calc. for C$_{18}$H$_{16}$O$_5$Na$_1$ [M + Na]$^+$: 335.08954, found 305.09010; elemental anal. for C$_{14}$H$_{14}$O$_4$ calc. C% 69.22, H% 5.16, found C% 68.80, H% 5.14.

6-methoxy-3-(4-methoxyphenyl)-2H-chromen-2-one (10). (Prendergast, 2001; Ferino et al., 2013) Yield 79%; ^{1}H-NMR (400 MHz, CDCl$_3$) δ: 3.847 (s, 3H, CH$_3$O-Ph), 3.852 (s, 3H, CH$_3$O-Ph), 6.95-6.98 (m, 3H, H-5, H-3', H-5'), 7.07 (dd, 1H, J^3 = 9.0 Hz, J^4 = 2.9 Hz, H-7), 7.27 (d, 1H, J^4 = 8.8 Hz, H-5), 7.66 (d, 2H, J^3 = 8.9 Hz, H-2', H-6'), 7.70 (s, 1H, H-4); ^{13}C-NMR (100 MHz, CDCl$_3$) δ:55.69, 56.16, 110.11, 114.23, 117.69, 119.05, 120.51, 127.47, 128.49, 130.18, 138.63, 148.11, 156.44, 160.49, 161.24; HRMS(ESI): calc. for C$_{17}$H$_{14}$O$_4$Na$_1$ [M + Na]$^+$: 305.07898, found 305.07910; elemental anal. for C$_{17}$H$_{14}$O$_4$ C% 72.33, H% 5.00, found C% 72.34, H% 4.86.

3-(3-methoxyphenyl)-2H-chromen-2-one (11). (Kirkiacharian et al., 1999) Yield 81%; ^{1}H-NMR (400 MHz, CDCl$_3$) δ: 3.86 (s, 3H, CH$_3$O-Ph), 6.95 (ddd, 1H, J^3 = 8.2 Hz, J^4 = 2.3 Hz, J^4 = 2.5 Hz, H-4'), 7.26-7.37 (m, 5H, H-6, H-8, H-2', H-5', H-6'), 7.51-7.53 (m, 2H, H-5, H-7), 7.81 (s, 1H, H-4); ^{13}C-NMR (100 MHz, CDCl$_3$) δ: 55.69, 114.56, 114.86, 116.76, 119.94, 121.24, 124.81, 128.26, 128.49, 129.80, 131.76, 136.35, 140.28, 153.85, 159.88, 160.76; HRMS(ESI): calc. for C$_{16}$H$_{12}$O$_3$Na$_1$ [M + Na]$^+$: 275.06841, found 275.06540; elemental anal. for C$_{16}$H$_{12}$O$_3$ calc. C% 76.18, H% 4.79, found C% 75.94, H% 4.67.

7-hydroxy-3-(4-methoxyphenyl)-2H-chromen-2-one (12). (Prendergast, 2001) Yield 81%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 3.79 (s, 3H, CH$_3$O-Ph), 6.74 (s, 1H, H-8), 6.81 (d, 1H, J^3 = 8.5 Hz, H-6), 6.99 (d, 2H, J^3 = 8.3 Hz, H-3', H-5'), 7.57 (d, 1H, J^3 = 8.4 Hz, H-5), 7.65 (d, 2H, J^3 = 8.3 Hz, H-2, H-6'), 8.08 (s, 1H, H-4), 10.54 (s, 1H, HO-Ph); ^{13}C-NMR (100 MHz, d^6-DMSO) δ: 55.18, 101.66, 112.10, 113.29, 113.61, 121.84, 127.30, 129.48, 129.70, 139.73, 154.63, 159.14, 160.20, 160.88; HRMS(ESI): calc. for C$_{16}$H$_{12}$O$_4$Na$_1$ [M + Na]$^+$: 291.06333, found 291.06160.

3-(4-methoxyphenyl)-2-oxo-2H-chromen-7-yl acetate (13). (Bhandri et al., 1949) Yield 67%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 7.02 (d, 2H, J^3 = 7.8 Hz, H-3', H-5'), 7.17 (d, 1H, J^3 = 8.3 Hz, H-6), 7.29 (d, 1H, H-8), 7.69 (d, 2H, J^3 = 7.8 Hz, H-2', H-6'), 7.79 (d, 1H, J^3 = 8.2 Hz, H-5), 8.19 (s, 1H, H-4); ^{13}C-NMR (100 MHz, d^6-DMSO) δ: 20.85, 55.22, 109.57, 113.68, 117.49, 118.68, 125.69, 126.72, 129.19, 129.75, 138.62, 152.36, 153.15, 159.59, 168.81; HRMS (ESI): Calc for C$_{18}$H$_{14}$O$_5$Na$_1$ [M + Na]$^+$: 333.07389, found 333.07220. Elemental analysis for C$_{18}$H$_{14}$O$_5$ calc C% 69.67 H% 4.55, found C% 69.58 H% 4.52.

3-(4-methoxyphenyl)-2-oxo-2H-chromen-6-yl acetate (14). Yield 34%; 1H-NMR (300 MHz, d6-DMSO) δ: 2.31 (s, 3H, CH3C(O)O-), 3.81 (s, 3H, CH3O-), 7.03 (d, 2H, J3 = 8.7 Hz, H-3', H-5'), 7.37 (dd, 1H, J3 = 8.9 Hz, J4 = 2.5 Hz, H-7), 7.47 (d, 1H, J3 = 8.9 Hz, H-8), 7.54 (d, 1H, J4 = 2.5 Hz, H-5), 7.70 (d, 2H, J3 = 8.7 Hz, H-2', H-6'), 8.15 (s, 1H, H-5); 13C-NMR (75 MHz,

d6-DMSO) δ: 20.73, 55.21, 113.68, 116.83, 120.07, 120.53, 125.00, 126.61, 127.03, 129.84, 138.30, 146.39, 150.18, 159.64, 159.71, 169.22. HRMS (ESI): Calc for C18H14O5 [M + H]+: 311.0914, found 311.0908.

6-methoxy-3-(2,4,5-trifluorophenyl)-2H-chromen-2-one (15). Yield 80%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 3.81 (s, 3H, CH$_3$O-Ph), 7.26 (dd, 1H, J^3 = 9.0 Hz, J^4 = 3.0 Hz, H-7), 7.31 (d, 1H, J^4 = 3.0 Hz, H-5), 7.41 (d, 1H, J^3 = 9.0 Hz, H-8), 7.64-7.77 (m, 2H, H-2', H-6'), 8.18 (s, 1H, H-4); ^{13}C-NMR (100 MHz, d^6-DMSO) δ: 55.73, 106.31 (dd, J^{C-F} = 21 Hz, J^{C-F} = 22 Hz), 110.90, 117.25, 119.12, 119.39, 119.55, 120.07, 120.91, 143.74, 145.70 (d, J^{C-F} = 242 Hz), 147.64, 149.34 (J^{C-F} = 252 Hz), 155.13 (J^{C-F} = 248 Hz), 155.79, 158.78. HRMS (ESI): Calc for C$_{16}$H$_9$F$_3$O$_3$Na$_1$ [M + Na]$^+$: 329.04015, found 329.04090. Elemental analysis for C$_{16}$H$_9$F$_3$O$_3$: calc C% 62.75 H% 2.96, found C% 62.62 H% 3.15.

7-methoxy-3-(2,4,5-trifluorophenyl)-2H-chromen-2-one (16). Yield 85 %; ^{1}H-NMR (300 MHz, d^6-DMSO) δ: 3.88 (s, 3H, CH$_3$O-Ph), 7.00 (dd, 1H, J^3 = 8.6 Hz, J^4 = 2.4 Hz, H-6), 7.06 (d, 1H, J^4 = 2.3 Hz, H-8), 7.61-7.6 (m, 3H, H-5, H-2', H-6'), 8.17 (s, 1H, H-4); ^{13}C-NMR (75.5 MHz, d^6-DMSO) δ: 56.02, 100.49, 106.21 (dd, J^{C-F} = 21 Hz, J^{C-F} = 21 Hz), 112.24, 112.85, 116.85, 119.30, 119.57, 129.95, 144.06, 145.67 (d, J^{C-F} = 242 Hz), 148.93 (d, J^{C-F} = 250) Hz, 155.10 (d, J^{C-F} = 245 Hz), 155.22, 158.89, 162.98; HRMS (ESI): Calc for C$_{16}$H$_9$F$_3$O$_3$Na$_1$ [M + Na]$^+$: 329.04015, found 329.03980.

3-(4-(dimethylamino)phenyl)-7-hydroxy-2H-chromen-2-one (17). (Kirkiacharian et al., 2003) In the first step 7-acetoxy-3-(4-(dimethylamino)phenyl)-2H-chromen-2one was obtained. Yield: 70%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 2.31 (s, 3H, CH$_3$C(O)O-Ph), 2.95 (s, 6H, (CH$_3$)$_2$N-Ph), 6.77 (d, J^3 = 9.0 Hz, 2H, H-2', H-6'), 7.14 (dd, J^3 = 8.4 Hz, J^4 = 2.2 Hz, 1H, H-5), 7.26 (d, J^4 = 2.2 Hz, 1H, H-8), 7.63 (d, J^3 = 9.0 Hz, 2H, H-3', H-5') 7.76 (d, J^3 = 8.5 Hz, 1H, H-5), 8.11 (s, 1H, H-4); ^{13}C-NMR (100.6 MHz, d^6-DMSO) δ: 20.85, 39.84, 109.44, 111.58, 117.76, 118.57, 121.57, 126.00, 128.82, 129.11, 136.46, 150.45, 151.90, 152.77, 159.74, 168.85. In the second step 7-hydroxy-3-(4-(dimethylamino)phenyl)-2H-chromen-2one was obtained. Yield: 85% yellow solid; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 2.94 (s, 6H, (CH$_3$)$_2$N-), 6.72 (d, J^4 = 2.3 Hz, 1H, H-8), 6.75 (d, J^3 = 9.0 Hz, 2H, H-2', H-6'), 6.79 (dd, J^3 = 8.4 Hz, J^4 = 2.3 Hz, 1H, H-5),), 7.55 (d, J^3 = 8.5 Hz, 1H, H-5), 7.58 (d, J^3 = 9.0 Hz, 2H, H-3', H-5'), 7.99 (s, 1H, H-4); ^{13}C-NMR (100.6 MHz, d^6-DMSO) δ: 39.92, 101.59, 112.33, 113.16, 122.30, 122.32, 129.34, 137.83, 150.07, 154.27, 160.30, 160.41; HRMS (ESI): Calc for C$_{17}$H$_{15}$N$_1$O$_3$Na$_1$ [M + Na]$^+$: 304.09496, found 304.09480; elemental anal. for C$_{17}$H$_{15}$N$_1$O$_3$, calc. C% 72.58, H% 5.37, N% 4.98, found C% 72.45, H% 5.40, N% 5.15.

3-(4-(dimethylamino)phenyl)-6-methoxy-2H-chromen-2-one (18). Yield 55%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 2.96 (s, 6H, (CH$_3$)$_2$N-Ph), 3.81 (s, 3H, CH$_3$O-Ph), 6.77 (d, 2H, J^3 = Hz, H-3', H-5'), 7.14 (dd, 1H, J^3 = 3.0 Hz, J^4 = 9.0 Hz, H-7), 7.28 (d, 1H, J^4 = 3.0 Hz, H-5), 7.33 (d, 1H, J^3 = 9.0 Hz, H-8), 7.63 (d, 2H, J^3 = 9.0 Hz, H-2', H-6'), 8.06 (s, 1H, H-4); ^{13}C-NMR (100.6 MHz, d^6-DMSO) δ: 39.93, 110.27, 111.59, 116.68, 118.16, 120.35, 121.73, 126.96, 129.15, 136.79, 146.79, 150.46, 155.59, 160.06; HRMS (ESI): Calc for C$_{18}$H$_{17}$N$_1$O$_3$Na$_1$ [M + Na]$^+$: 318.11061, found 318.11050; elemental anal. for C$_{18}$H$_{17}$N$_1$O$_3$, calc. C% 73.20, H% 5.80, N% 4.74, found C% 72.75, H% 5.83, N% 4.45.

3-(4-(dimethylamino)phenyl)-2-oxo-2H-chromen-7-yl acetate (19). Yield 70%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 2.31 (s, 3H, CH$_3$C(O)O-Ph), 2.95 (s, 6H, (CH$_3$)$_2$N-Ph), 6.77 (d, 2H, J^3 = 9.0 Hz, H-3', H-5'), 7.14 (dd, 1H, J^3 = 8.4 Hz, J^4 = 2.2 Hz, H-6), 7.26 (d, 1H, J^4 = 2.2 Hz, H-8), 7.63 (d, 2H, J^3 = 9.0 Hz, H-2', H-6'), 7.76 (d, 1H, J^3 = 8.5 Hz, H-5), 8.11 (s, 1H, H-4; ^{13}C-NMR (100 MHz, d^6-DMSO) δ: 20.86, 39.84, 109.44, 111.58, 117.76, 118.57, 121.58, 126.00, 128.82, 129.11 136.46, 150.45, 151.90, 152.77, 159.74, 168.85; HRMS (ESI): Calc for C$_{19}$H$_{17}$N$_1$O$_4$Na$_1$ [M + Na]$^+$: 346.10553, found 346.10640.

3-(3-fluoro-4-hydroxyphenyl)-7-methoxy-2H-chromen-2-one (20). In the first step 2-fluoro-4-(7-methoxy-2-oxo-2H-chromen-3-yl)phenyl acetate was obtained. Yield 75%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 2.35 (s, 3H, CH$_3$C(O)O-Ph), 3.88 (s, 3H, CH$_3$O-Ph), 6.99 (dd, 1H, J^3 = 8.6 Hz, J^4 = 2.4 Hz, H-6), 7.05 (d, 1H, J^4 = 2.4 Hz, H-8), 7.37 (t, 1H, J = 8.3 Hz, H-6'), 7.62 (d, 1H, J^3 = 8.5 Hz, H-5'), 7.68 (d, J = 8.6 Hz, 1H, H-5), 7.74 (dd, J^{H-F} = 12.1 Hz, J^4 = 2.0 Hz, H-3'), 8.31 (s, 1H, H-4); ^{13}C-NMR (100 MHz, d^6-DMSO) δ: 20.19, 55.97, 100.25, 112.79, 116.35 (d, J^{C-F} = 20.3 Hz), 121.02, 121.03, 123.83, 124.79 (d, J^{C-F} = 3.2 Hz), 129.86, 134.24 (d, J^{C-F} = 7.7 Hz), 137.20 (d, J^{C-F} = 13.1 Hz), 141.55, 153.00 (J^{C-F} = 246 Hz), 154.92, 159.65, 162.69, 168.19. In the second step 3-(3-fluoro-4-hydroxyphenyl)-7-methoxy-2H-chromen-2-one was obtained. Yield 70%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 3.87 (s, 3H, CH$_3$O-Ph), 6.96-7.03 (m, 3H, H-6, H-8, H-5'), 7.41 (d, 1H, J^3 = 8.4, H-6'), 7.57 (dd, 1H, J^{H-F} = 13.1 Hz, J^4 = 2.2 Hz (H-H), 1H, H-2'), 7.66 (d, 1H, J^3 = 8.4, H-5), 8.18 (s, 1H, H-4), 10.09 (s, 1H, Ph-OH). ^{13}C-NMR (75.5 MHz, d^6-DMSO) δ: 55.91, 100.16, 112.61, 113.04, 115.95 (d, J^{C-F} = 20 Hz), 117.37 (d, J^{C-F} = 3.3 Hz), 121.78 (J^{C-F} = 2.0 Hz), 124.54 (d, J^{C-F} = 3.0 Hz), 126.08 (d, J^{C-F} = 7.0 Hz), 129.49, 139.62, 145.0 (J^{C-F} = 13 Hz), 150.46 (d, J^{C-F} = 240 Hz), 154.52, 159.87, 162.19; HRMS (ESI): Calc for C$_{16}$H$_{11}$F$_1$O$_4$Na$_1$ [M + Na]$^+$: 309.0539, found 309.0553.

3-(4-fluorophenyl)-6-methoxy-2H-chromen-2-one (21). Yield 58%; ^{1}H-NMR (400 MHz, d^6-acetone) δ: 3.87 (s, 3H, CH$_3$O-Ph), 7.19-7.33 (m, 5H, H-5, H-7, H-8, H-3', H-5'), 7.83 (dd, 2H, J^{HF} = 5.4 Hz, J^{H-H} = 9.0 Hz, H-2', H-6'), 8.12 (s, 1H, H-4); ^{13}C-NMR (100 MHz, d^6-acetone) δ: 56.17, 111.34, 115.84 (d, J^{C-F} = 22 Hz), 117.85, 120.04, 121.04, 127.79, 131.62 (d, J^{C-F} = 8 Hz), 132.41 (d, J^{C-F} = 3 Hz), 140.82, 148.82, 157.13, 160.64, 163.72 (d, J^{C-F} = 247 Hz); HRMS (ESI): Calc for C$_{16}$H$_{11}$F$_1$O$_3$Na$_1$ [M + Na]$^+$: 293.05899, found 293.05850; elemental anal. for C$_{16}$H$_{11}$F$_1$O$_3$, calc C% 71.11, H% 4.10, found C% 71.10, H% 4.10.

3-(3-fluoro-4-hydroxyphenyl)-6-methoxy-2H-chromen-2-one (22). In the first step 2-fluoro-4-(6-methoxy-2-oxo-2H-chromen-3-yl)phenyl acetate was obtained. Yield 66%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 2.33 (s, 3H, CH$_3$C(O)O-Ph), 3.82 (s, 3H, (CH$_3$O-Ph), 7.23 (dd, 1H, J^3 = 9.0 Hz, J^4 = 3.0 Hz, H-7), 7.30 (d, 1H, J^4 = 3.0 Hz, H-5), 7.35 (d, 1H, J^3 = 9.2 Hz, H-8), 7.61 (d, 1H, J^3 = 8.5 Hz, H-5'), 7.75 (dd, 1H, J^{H-F} = 12.0 Hz, J^4 = 1.7 Hz (H-H), 1H, H-3'), 8.30 (s, 1H, H-4); ^{13}C-NMR (100.6 MHz, d^6-DMSO) δ: 20.22, 55.69, 110.83, 116.67, 117.02, 119.66, 123.96,

125.10, 135.96, 141.18, 147.44, 151.78, 154.23, 155.70, 159.53, 168.21. In the second step 3-(3-fluoro-4-hydroxyphenyl)-6-methoxy-2H-chromen-2-one was obtained. Yield 71%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 3.81 (s, 3H, (CH$_3$O-Ph), 7.02 (dd, 1H, J^3 = 9.2 Hz, H-6'), 7.18 (dd, 1H, J^3 = 9.0 Hz, J^4 = 3.0 Hz, H-7), 7.28 (d, 1H, J^4 = 2.9 Hz, H-5), 7.42 (d, 1H, J^3 = 8.4 Hz, H-5'), 7.57 (dd, 1H, J^{H-F} = 13.0 Hz, J^4 = 2.2 Hz (H-H), 1H, H-2'), 8.17 (s, 1H, H-4), 10.19 (s, 1H, Ph-OH); ^{13}C-NMR (100.6 MHz, d^6-DMSO) δ: 55.66, 110.59, 116.67, 117.02, 119.66, 123.96, 125.10, 135.96, 141.18, 147.44, 151.78, 154.23, 155.70, 159.53, 168.21. HRMS (ESI): Calc for C$_{16}$H$_{11}$F$_1$O$_4$Na$_1$ [M + Na]$^+$: 309.0539, found 309.0521.

3-(4-fluorophenyl)-6-methyl-2H-chromen-2-one   (23). (Chauhan et al., 2016) Yield 74%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 2.38 (s, 3H, CH$_3$-Ph), 7.27-7.35 (m, 3H, H-3', H-5', H-8), 7.43 (dd, 1H, J_3 = 8.5 Hz, J_4 = 2.1 Hz, H-7), 7.55 (d, 1H, J_4 = 1.4 Hz, H-5), 7.77 (dd, 2H, J^{HF} = 5.7 Hz, J^{H-H} = 9.0 Hz, H-2', H6'), 8.18 (s, 1H, H-4); ^{13}C-NMR (100.6 MHz, d^6-DMSO) δ: 20.26, 115.11 (d, J^{H-F} = 21.5 Hz), 115.64, 119.16, 125.76, 128.20, 130.70 (d, J^{H-F} = 8.4 Hz), 131.10 (d, J^{H-F} = 3.2 Hz), 132,61, 133.80, 140.48, 151.10, 159.82, 162.17 (d, J^{H-F} = 245 Hz); HRMS (ESI): Calc for C$_{16}$H$_{11}$F$_1$O$_2$Na$_1$ [M + Na]$^+$: 277.06408, found 277.06390; Elemental anal. for C$_{16}$H$_{11}$F$_1$O$_2$, calc C% 75.58, H% 4.36, found C% 75.42, H% 4.33.

3-(4-fluorophenyl)-6-hydroxy-2H-chromen-2-one (24). In the first step 3-(4-fluorophenyl)-2-oxo-2H-chromen-6-yl acetate was obtained and used as such for the next step. In the second step 3-(4-fluorophenyl)-6-hydroxy-2H-chromen-2-one was obtained. Yield 65%; ^{1}H-NMR (300 MHz, d^6-DMSO) δ: 7.04 (dd, 1H, J^3 = 8.8 Hz, J^4 = 2.9 Hz, H-7), 7.09 (d, 1H, J^4 = 2.8 Hz, H-5), 7.24-7.29 (m, 3H, H-3', H-5', H-8), 7.75 (dd, 2H, J^{HF} = 5.6 Hz, J^{H-H} = 8.9 Hz, H-2', H6'), 8.13 (s, 1H, H-4), 9.72 (s, 1H, HO-Ph); ^{13}C-NMR (75.5 MHz, d^6-DMSO) δ: 112.561, 115.03 (d, J^{H-F} = 21.5 Hz), 116.71, 119.78, 119.93, 125.80, 130.70 (d, J^{H-F} = 8.2 Hz), 131.18 (d, J^{H-F} = 3.2 Hz), 140.50, 146.35, 159.92, 162.17 (d, J^{H-F} = 246 Hz); HRMS (ESI): Calc for C$_{15}$H$_9$F$_1$O$_3$Na$_1$ [M + Na]$^+$: 279.04334, found 279.0444.

Monoamine Oxidase A and B

Both monoamine oxidase A (MAO-A) and B (MAO-B) protein and the reagents for the chromogenic solution of vanillic acid (4-hydroxy-3-methoxylbenzoic acid, 97% purity), 4-aminoantipyrine (reagent grade), horseradish peroxidase and the substrate tyramine hydrochloride (minimum 99% purity) as well as the potassium phosphate buffer, which was prepared using potassium phosphate dibasic trihydrate (≥99% ReagentPlus^{TM}) and potassium phosphate monobasic (minimum 98% purity, molecular biology tested), were purchased from Sigma-Aldrich (St. Louis, MO, USA) for the spectrophotometric assay.

The protocol for continuous spectrophotometric assay (Holt et al., 1997) was followed in the activity measurements. The assay was performed in 0.2 M potassium phosphate buffer pH 7.6 on 96-well plates (Nunc^{TM} 96F microwell plate without a lid, Nunc A/S, Roskilde, DK) in 200 μl total volume. The chromogenic solution containing 1 mM vanillic acid, 500 μM 4-aminoantipyrine and 8 U/ml horseradish peroxidase in 0.2 M potassium phosphate buffer pH 7.6 was mixed anew for each

measurement. 5 mM tyramine solution was used as the substrate. In order to determine the activity of both MAO-B and MAO-A, concentration series as duplicates were prepared. The protein was combined with the chromogenic solution and incubated 30 min at 37°C. The background signal was measured using multilabel reader (Victor^{TM} X4, 2030 Multilabel Reader, PerkinElmer, Waltham, MA, USA) at A$_{490}$ before reaching the total 200 μl volume by adding 20 μl of tyramine to final concentration of 0.5 mM on the plate. As a result, the final concentration of the chromogenic solution on the plate was 250 μM vanillic acid, 125 μM 4-aminoantipyrine and 2 U/ml horseradish peroxidase. After adding the substrate, the plates were measured 300 times every 15 s using 1 s exposure time. The device was set to 37°C for the duration of the experiment.

Based on the activity measurement, suitable concentrations were chosen for both MAO-B and MAO-A to be used in the inhibition studies (Supplementary Figures S1, S2, and S5, **Table 1**, Supplementary Table S1). The experiment conditions should produce absorbance change of ∼0.35 (Holt et al., 1997). With MAO-B, this was reached using 10 μl (equals 50 μg of protein with enzymatic activity 3.2 units per well) of the protein and running the experiment for 2 h (Supplementary Figures S1, S2, and S5, **Table 1**, Supplementary Table S1). MAO-A was significantly more active, providing absorbance change of >0.5 with 5 μl (equals 25 μg of protein with enzymatic activity 1.05 units per well) of protein and, consequently, the reaction maximum was reached already in 30 min (Supplementary Figure S5, **Table 1**, Supplementary Table S1). Thus, a wide panel of coumarin derivatives was analyzed at 10 μM (**Table 1,** Supplementary Table S1) and those 3-phenylcoumarin derivatives producing >70% inhibition were selected for further analysis (**Table 1**, **Figure 2**). The selected 24 candidates were measured as duplicates on a dilution series ranging from 50 μM to 1 nM, and based on the normalized measurement results, IC$_{50}$ values were calculated (**Table 1**). The same wide panel of coumarin derivatives was additionally used to analyze the MAO-A inhibition at 100 μM (**Table 1,** Supplementary Table S1).

GRAPHPAD PRISM 5.03 (GraphPad Software Inc., CA, USA) was used to normalize the spectrophotometric assay data where the maximal signal was reached at the lowest concentration of 10^{-8} or 10^{-9} depending on the sample and the starting concentration of $5·10^{-5}$ acted as the lowest point of signal. The measured data was then fitted on a curve using non-linear regression with the equation for log[inhibitor] vs. response. The IC50 values were therefore determined based on the curve fit. The fitted curves are shown on −log scale in Supplementary Figures S1, S2.

17-β-Hydroxysteroid Dehydrogenase 1

Inhibition of the 17-β-hydroxysteroid dehydrogenase 1 (HSD1) was determined by HPLC using recombinant human HSD1 proteins, produced in Sf9-insect cells, as described earlier (Messinger et al., 2009). The assay was performed in a final volume of 0.2 ml buffer (20mM KH2PO4, 1mM EDTA, pH 7.4) containing 0.1 mg/ml protein, 1mM cofactor NADPH, 30 nM substrate estrone or estradiol, 800,000 cpm/ml of tritium

labeled estrone ([3H]-E1) or estradiol ([3H]-E2) and inhibitor concentrations in the range of 0.1–5 mM. Triplicate samples were incubated for 25 min at RT. The reaction was stopped by addition of 20 ml 10% trichloroacetic acid per sample. After incubation the substrate and the product of enzymatic conversion [3H]-E1 and [3H]-E2, were separated and quantified by HPLC (Alliance 2790, Waters) connected to an online - counter (Packard Flow Scintillation Analyzer). The ratio of [3H]-E1 converted to [3H]-E2, or *vice versa*, determines the sample conversion percentage. Inhibition efficiencies were calculated by comparing the conversion percentages of the samples including inhibitors with those of conversion controls (without inhibitors).

Aromatase

Aromatase (CYP19A1) activity was measured as described previously (Pasanen, 1985) by using human placental microsomes and 50 nM [3H]-androstenedione as a substrate and inhibitor concentrations in the range of 60–1,000 nM. Aromatase activities were measured as released [3H]-H_2O in Optiphase Hisafe 2 scintillation liquid (Perkin Elmer, USA) with a Wallac 1450 MicroBeta Trilux scintillation counter (Perkin Elmer, USA). As a positive control for aromatase inhibition, 1 μM finrozole (generous gift from Olavi Pelkonen, University of Oulu, Finland) was used.

Cytochrome P450 1A2

Inhibition of CYP1A2 activity was determined with commercial heterologously expressed human CYP1A2 enzyme (Corning Inc., Corning, NY, USA) as described earlier (Korhonen et al., 2005). The metabolic activity was not in the scope of this particular study. The assay was adapted to the 96-well plate format. In each well, a 150 μL incubation volume contained 100 mM Tris-HCl buffer (pH 7.4), 4.2 mM $MgCl_2$, 1 μM 7-ethoxyresorufin, 0.5 pmol of cDNA expressed CYP1A2, 0-40 mM inhibitor, and a NADPH-generating system. All inhibitors were dissolved in ethanol, and the final concentration of ethanol was 2% in all incubations. The reaction was initiated by adding the NADPH-regenerating system after a 10 min preincubation at 37°C, and after a 20 min incubation, the reaction was terminated by the addition of 110 μL of 80% acetonitrile/20% 0.5 M Tris base. The formed fluorescence was measured with a Victor2 plate counter (Perkin-Elmer Life Sciences Wallac, Turku, Finland) at 570 nm excitation and 616 nm emission.

Estrogen Receptor

The pIC50 values for the derivatives (**Table 1**, Supplementary Table S1) were measured with green PolarScreen™ ER Alpha Competitor Assay (Life Technologies, CA, The United States of America) kit, following the manufacturer protocol as previously described (Niinivehmas et al., 2016). The final concentration of the compounds ranged from 0.0007 to 10 000 nM in the dilution series which were performed as duplicates. The molecules were combined with 25 nM ERα and 4.5 nM fluormone in the assay buffer and placed on black low volume 384-well assay plate with NBS surface (Corning, NY, The United States of America). After mixing the assay plate, it was incubated for 2 h in RT. The fluorescence polarization was measured using excitation wave

length 485 and emission wave length 535 with bandwidths of 25/20 nm on a 2104 EnVision® Multilabel Plate Reader which had EnVision Workstation version 1.7 (PerkinElmer, MA, The United States of America).

Computational Methods

The small-molecule ligand structures were drawn in 3D and their tautomeric states at pH 7.4 were built using LIGPREP module in MAESTRO 2016-3 (Schrödinger, LLC, New York, NY, USA, 2016). The derivatives were docked to the X-ray crystal structure of MAO-B (PDB: 2V60) (Binda et al., 2007) with PLANTS 1.2 (Korb et al., 2009) using 10 Å radius and the C8 atom of inhibitor *C18* (PDB: 2V60) was used as the center. The R1-methoxy group rotamers of compounds **1**, **8**, **9**, **21**, **15**, **18**, and **22** were manually adjusted to indicate how the groups exploit the small hydrophobic niche in the cavity (green sector in **Figures 3A,B**). The 2D structures of the 3-phenylcoumarin scaffold and the 24 most potent inhibitor derivatives shown in **Figures 1E**, **2** were drawn with BIOVIA DRAW 2016 (Dassault Systèmes, San Diego, CA, USA, 2016). **Figures 1A–D**, **3–5** were prepared using BODIL (Lehtonen et al., 2004) and VMD 1.9.2 (Humphrey et al., 1996). The negative images of the MAO-B and MAO-A binding cavities shown in **Figure 3A** and C were outlined using PANTHER (Niinivehmas et al., 2011, 2015) and visualized with BODIL, MOLSCRIPT (Kraulis, 1991), and RASTER3D (Merritt and Murphy, 1994).

RESULTS AND DISCUSSION

Spectrophotometric Activity Measurements for Monoamine Oxidase B

All of the 52 derivatives were docked, synthetized and tested experimentally. Those 24 compounds that provided IC$_{50}$ values below 10 μM were tested more thoroughly (**Table 1**). The fact that 24 of the synthesized derivatives with a wide variety of different R1-R7 groups (**Figure 2**) passed the 70% threshold indicates that the 3-phenylcoumarin is indeed a highly suitable scaffold for building MAO-B inhibitors. Notably, eight of these tested derivatives (**3**, **9–13**, **17,** and **23** in **Figure 2**) had been synthesized previously (Bhandri et al., 1949; Kirkiacharian et al., 1999, 2003; Prendergast, 2001; Vilar et al., 2006; Ferino et al., 2013; Chauhan et al., 2016; Dobelmann-Mara et al., 2017), however, this is the first time they are tested for MAO-B activity. The novel derivative **1** is the most potent inhibitor of the analog set with the IC$_{50}$ value of 56 nM (**Figure 2**, **Table 1**); meanwhile, the rest of the tested derivatives are evenly distributed within a range of 0.1–10 μM (**Figure 2**, **Table 1**).

By focusing solely on the R1-R7 constituents of the derivatives (**Figures 1E**, **2**) and the activity data (**Table 1**) it is possible to outline trends that determine which functional groups, positions or their combinations establish and weaken or improve the MAO-B inhibition.

Although the R1 and R2 groups in the coumarin ring are not necessarily required for establishing MAO-B inhibition (see **11**; **Figure 2**, Supplementary Figure S3F; **Table 1**), the activity measurements indicate that adding a methoxy, hydroxyl, acetoxy, methyl or even halogen group(s) into the ring can facilitate strong

FIGURE 3 | The active site of monoamine oxidase B with docked 3-phenylcoumarin derivatives. **(A)** A negative image of the MAO-B active site shown as a transparent surface indicates the space available for inhibitor binding with docked derivative **1** (ball-and-stick model; **Figure 2**). **(B)** A cross section, showing half of the active site, displays the contours (opaque surface) that roughly match the inhibitor shape and conformation. The colored sectors highlight specific sections of the cavity dedicated to different aspects of the 3-phenylcoumarin derivative binding: 3-phenyl ring (orange), the R4-R7 groups of the 3-phenyl ring (red), coumarin ring (yellow), the hydrophobic niche occupied by the R1/R2-groups of the coumarin ring (green). **(C)** A negative image of the MAO-A active site shows that only two residue changes (Ile199→ Phe208; Leu164→ Phe173) are enough to prevent 3-phenylcoumarin analog binding. **(D)** The docked poses of the 23 most potent 3-phenylcoumarin derivatives show what space is collectively occupied by the new inhibitors. See **Figure 1** for details.

inhibition (**Table 1**). As a rule of thumb, introducing R1-methoxy group produces strong MAO-B inhibition (e.g., **1**; **Figure 2**; **Table 1**). In contrast, inserting for example a bulky R3 substituent such as acetoxy group weakens the inhibition considerably (26, 35, 47; Supplementary Figure S4; Supplementary Table S1). Whether the R1 or R2 position or any specific functional group in particular is favored depends on the composition of the 3-phenyl ring's R4-R7 constituents.

In fact, the activity data indicates that the R4-R7 substituents are vital for assuring strong MAO-B inhibition and without any 3-phenyl substituents, the activity is lost (41, 50, 52; Supplementary Figure S4, Supplementary Table S1). The most potent inhibitors were **1** (IC_{50} of ∼56 nM; **Table 1**) and **2** (IC_{50} of ∼138 nM; **Table 1**) housing R6-trifluoromethyl, but **3** (IC_{50} of ∼141 nM; **Table 1**) with structurally similar R6-trifluoromethoxy group is almost equally potent. The combination of the R6-acetoxy and R7-fluorine groups in **6** (IC_{50} of ∼189 nM) produces relatively strong inhibition. Furthermore, housing just one methoxy group at the R7 position (**8**; IC_{50} of ∼230 nM) or two methoxy groups at both R5 and R7 positions (**9**; IC_{50} of ∼255 nM) assures < 300 nM inhibition.

The effects of the R4-R7 groups of the 3-phenyl ring and the R1-R3 groups of coumarin ring (**Figure 2**) for the derivative binding and inhibition are detailed below in a docking-based structure-activity relationship (SAR) analysis.

The Alignment of the 3-Phenylcoumarin Scaffold at the Active Site

The 3-phenylcoumarin derivative binding at the MAO-B active site is based on the premise that the coumarin and phenyl ring systems occupy roughly the same 3D space as the equivalent ring systems of the coumarin-based inhibitors co-crystallized with the enzyme (PDB: 2V60, 2V61; **Figures 1A–D**) (Binda et al., 2007). The fundamental difference between the 3-phenylcoumarin derivatives and those coumarin inhibitors with validated binding poses is that the coumarin alignment is reversed and the phenyl ring is attached to the C3-position instead of the C7-position (**Figures 1C,D**).

What is more, the "canonical" coumarin ring positioning inside the pocket is somewhat analogous to even simpler double ring constructs such as the indole of inhibitor isatin (PDB: 1OJA) (Binda et al., 2003). In fact, the hydrophobicity of the aromatic coumarin (yellow sector in **Figures 3A,B**) and 3-phenyl (orange sector in **Figures 3A,B**) rings is vital for establishing the MAO-B binding and it outweighs all other favorable interactions such as hydrogen or halogen bonding (*via* sigma hole) in importance (**Figure 4**). Thus, although the docking suggests variability in the coumarin and 3-phenyl ring positioning for the 3-phenylcoumarin derivatives due to different R1-R7 substituents, the hydrophobic interactions of the aromatic rings are highly similar between them (**Figure 3D**).

It is also noteworthy that the coumarin's C2-carbonyl is not facing the solvent based on the molecular docking simulations

FIGURE 4 | Structure-activity relationship (SAR) analysis of the 3-phenylcoumarin derivatives.
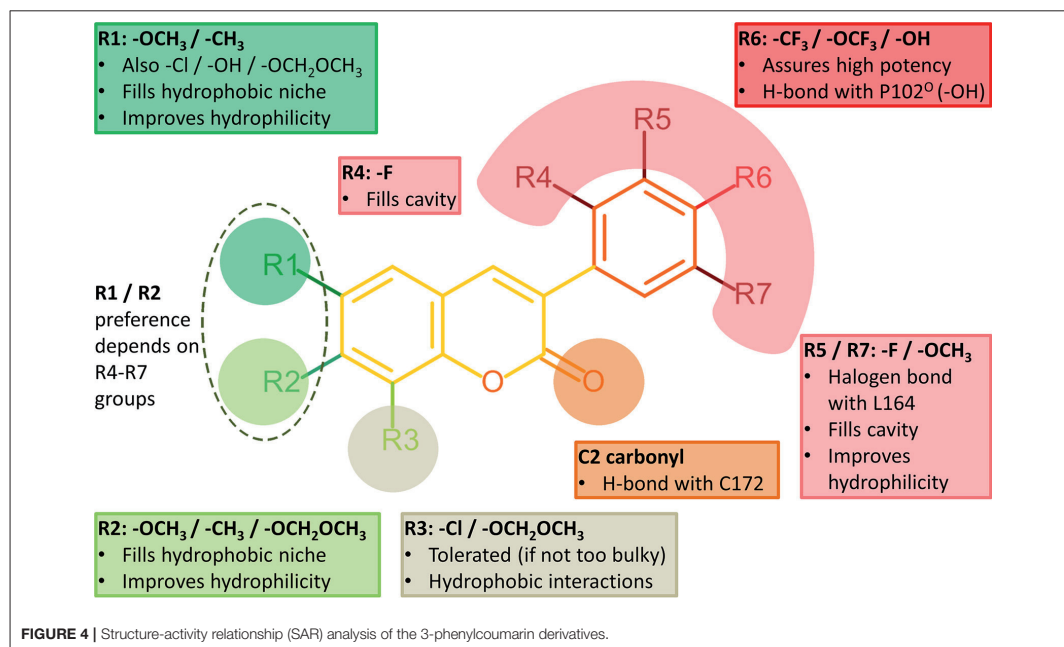
(**Figure 3D**). Paradoxically, this does not matter, because the carbonyl group finds an atypical interaction partner from the thiol group of Cys172 side chain (**Figure 4**). Although the C2-carbonyl cannot form a full-fledged H-bond with the proton of the thiol group, the hydrophobic environment of the cavity likely enhances this ordinarily weak interaction between the two groups.

R6-Trifluoromethyl Packing Produces the Strongest Inhibition

Halogen substituents in the 3-phenyl ring ensure strong MAO-B inhibition (**Figure 4**). This makes sense with MAO-B, because despite their apparent electronegativity the halogen substituents actually improve the steric packing of small-molecules via persistent van der Waals interactions while also retaining the ability to act as a halogen bond donor. Both of these properties should assist inhibitor binding into the active site that is mostly hydrophobic (**Figures 3A,B**). Besides, the increased lipophilicity conveyed by the halogen substituents (logP values in **Table 1**) should assist the 3-phenylcoumarin derivatives in aggregating on the outer mitochondrial membrane on route to the MAO-B active site (**Figure 1A**).
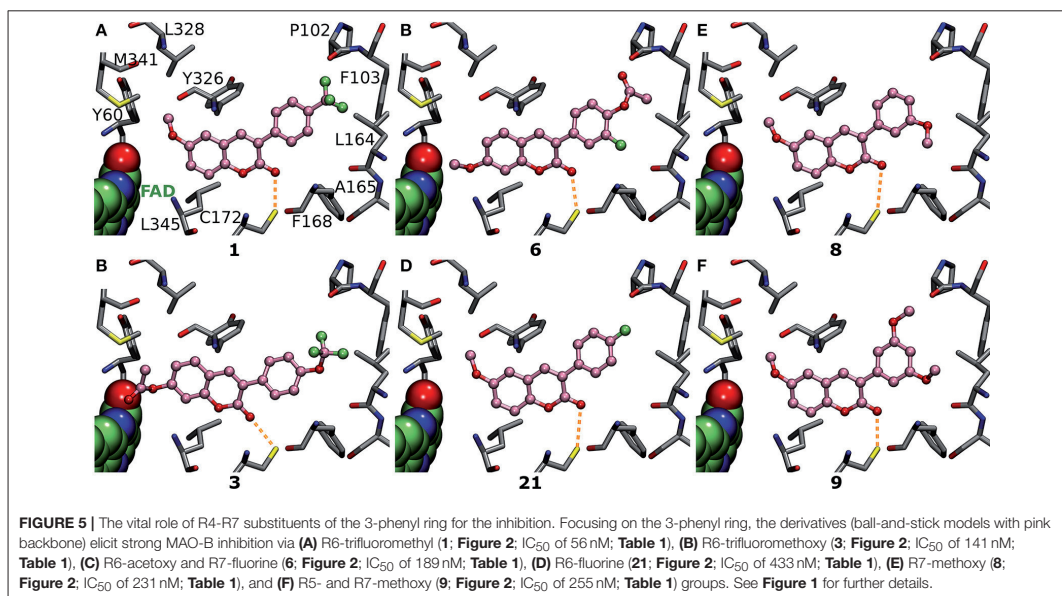
The most potent derivative **1** (**Figure 2**, **Table 1**) has trifluoromethyl group at the R6 position in the 3-phenyl ring. The derivative is relatively flat when bound at the active site and the proximal R6-group cannot flex out of this plane (**Figure 5A**). The trifluoromethyl of **1** fits very snugly into the hydrophobic end of the cavity (red sector in **Figures 3A,B**). The

high shape complementarity of this cavity part and the R6-trifluoromethyl of **1** is typical for this bulky moiety in drug compounds. Thus, the R6-group alignment of **1** is mostly relying on the collective potency of individually weak van der Waals interactions (**Figures 3A,B, 5A**).

Replacing the R6-trifluoromethyl of derivative **1** with a trifluoromethoxy in **4** (**Figure 2**) produces six times lower MAO-B inhibition (**Table 1**, Supplementary Figure S3B). This happens because the trifluoromethoxy already fills the available space almost optimally (**Figures 3A,B, 5A**) and elongating the substituent with an ether bond does not improve the fit (Supplementary Figure S3B). In fact, there is no extra wiggle room to fit the trifluoromethoxy (**Figures 3A,B**), if the 3-phenylcoumarin scaffold would be kept at the "canonical" position (**Figures 1C,D**). Hence, the coumarin ring of **4** pushes slightly closer to the cofactor. Although the binding site residues can adjust slightly in response to this shift, the realignment or rather misalignment of the scaffold (Supplementary Figure S3B) imposes an energetic cost that is reflected in the MAO-B inhibition (**Table 1**). In addition, depending on the rotamer pose of the R6-trifluorometoxy, a hydrogen bond could be bridged between a fluorine atom and the Pro102^O by a water molecule (not shown).

The Effects of Halogenation on the 3-Phenyl Ring Alignment

The chlorine and fluorine substituents of prior coumarin-based inhibitors form halogen bond with the Leu164^O based on

FIGURE 5 | The vital role of R4-R7 substituents of the 3-phenyl ring for the inhibition. Focusing on the 3-phenyl ring, the derivatives (ball-and-stick models with pink backbone) elicit strong MAO-B inhibition via **(A)** R6-trifluoromethyl (**1**; **Figure 2**; IC_{50} of 56 nM; **Table 1**), **(B)** R6-trifluoromethoxy (**3**; **Figure 2**; IC_{50} of 141 nM; **Table 1**), **(C)** R6-acetoxy and R7-fluorine (**6**; **Figure 2**; IC_{50} of 189 nM; **Table 1**), **(D)** R6-fluorine (**21**; **Figure 2**; IC_{50} of 433 nM; **Table 1**), **(E)** R7-methoxy (**8**; **Figure 2**; IC_{50} of 231 nM; **Table 1**), and **(F)** R5- and R7-methoxy (**9**; **Figure 2**; IC_{50} of 255 nM; **Table 1**) groups. See **Figure 1** for further details.
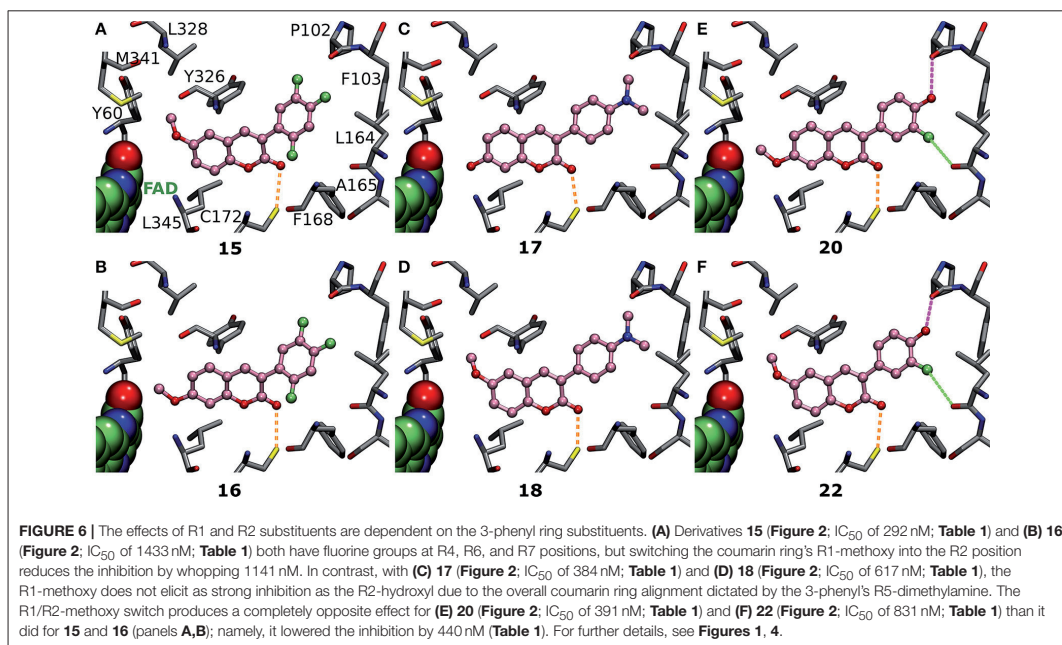
X-ray crystallography (PDB: 2V60, 2V61; **Figures 1A–D**; Binda et al., 2007). Accordingly, it is not surprising that those 3-phenylcoumarin derivatives with single halogen substituent at their 3-phenyl rings are also capable of blocking the MAO-B activity (**Figure 4**, **Table 1**).

Although it is known that fluorine is the poorest halogen bond donor (Cavallo et al., 2016), the R7-fluorine groups of **20** and **22** (**Figure 2**) could form halogen bond with the Leu164° (**Figures 6E,F**) similarly to the halogens of previously published inhibitors with validated binding modes (**Figures 1B–D**; Binda et al., 2007). In fact, the R7-halogen groups of **20** and **22** are inserted into the exact same position as the halogen groups of the established inhibitors (**Figure 1B** vs. **Figures 6E,F**). The MAO-B inhibition (**Table 1**) is reinforced further by the R6-hydroxyl group H-bonding with the Pro102^O (magenta dotted lines in **Figures 6E,F**). Because both **20** and **22** are bonding simultaneously with the Leu164^O and the Pro102^O, they elicit equivalent or stronger inhibition than derivatives **21** (**Figure 5D**), **23** (Supplementary Figure S3K), and **24** (Supplementary Figure S3L) that do not retain either one of these two interactions. Docking suggests that replacing the R6-hydroxyl with an acetoxy group prevents **6** (**Figure 2**) from forming direct halogen or hydrogen bonds (**Figure 5C**), but the R6-acetoxy and R7-fluorine could potentially connect via a water bridge with the Pro102^O (not shown). Despite this, the hydrophobic packing of the R6-acetoxy in **6** against the hydrophobic residues, mainly Phe103 (**Figure 5C**), is likely the reason behind doubling the inhibition in comparison to **20** (IC_{50} value of 391 vs. 189 nM; **Table 1**, **Figure 6E**).

Introducing fluorine to the R6 position of the 3-phenyl ring in derivatives **21**, **23**, and **24** (**Figure 2**) produces MAO-B inhibition ranging from 433 to 1,060 nM (**Table 1**). Due to the overall planarity of the 3-phenylcoumarin scaffold (**Figures 1C,D**), the R6-fluorine (**Figure 5D**, Supplementary Figures S3K,L), cannot take on the equivalent site occupied by the halogens of validated coumarin-based inhibitors that form halogen bond with the Leu164° (**Figure 1B**; Binda et al., 2007). In addition, the R6-fluorine is too limited in size to fill the end of the binding cavity as completely as for example the trifluoromethyl of **1** does (**Figures 3A,B**, **5A**). In addition, the R6-fluorine groups of derivatives **21**, **23**, and **24** (**Figure 5D**, Supplementary Figure S3K,L) reside within a suitable distance to form a halogen bond with the Pro102^O (3.6 Å), however, the available angles seem to rule out actual bonding.

Derivatives **15** and **16** (**Figure 2**) house three fluorine atoms at their 3-phenyl groups' R4, R6, and R7 positions (**Figures 6A,B**). In the case of **15** (**Figure 6A**), these halogen substituents assure an IC_{50} value that is almost 150 nM stronger than what is seen with derivatives housing only a single fluorine moiety at the R6 or R7 position (**21**, **23,** and **24**; **Figure 5D**, Figure S3K-L, **Table 1**). This is achieved by filling the hydrophobic cavity end (orange and red sectors in **Figure 3**) efficiently with the 3-phenyl ring and its fluorine moieties (**Figures 6A,B**). The fit is better for a 3-phenyl ring with the R5-trifluoromethyl than what is seen with the ring housing three separate fluorine substituents (**Figure 5A** vs. **Figure 6A**) and; accordingly, derivative **15** is not as potent MAO-B inhibitor as **1** (IC_{50} 292 vs. 56 nM; **Table 1**). In addition, depending on the 3-phenyl ring pose, the R4 or R7 fluorine

FIGURE 6 | The effects of R1 and R2 substituents are dependent on the 3-phenyl ring substituents. **(A)** Derivatives **15** (**Figure 2**; IC$_{50}$ of 292 nM; **Table 1**) and **(B) 16** (**Figure 2**; IC$_{50}$ of 1433 nM; **Table 1**) both have fluorine groups at R4, R6, and R7 positions, but switching the coumarin ring's R1-methoxy into the R2 position reduces the inhibition by whopping 1141 nM. In contrast, with **(C) 17** (**Figure 2**; IC$_{50}$ of 384 nM; **Table 1**) and **(D) 18** (**Figure 2**; IC$_{50}$ of 617 nM; **Table 1**), the R1-methoxy does not elicit as strong inhibition as the R2-hydroxyl due to the overall coumarin ring alignment dictated by the 3-phenyl's R5-dimethylamine. The R1/R2-methoxy switch produces a completely opposite effect for **(E) 20** (**Figure 2**; IC$_{50}$ of 391 nM; **Table 1**) and **(F) 22** (**Figure 2**; IC$_{50}$ of 831 nM; **Table 1**) than it did for **15** and **16** (panels **A,B**); namely, it lowered the inhibition by 440 nM (**Table 1**). For further details, see **Figures 1**, **4**.

groups could again potentially act as weak halogen bond donors to the Phe168^O or the Leu164^O, respectively (not shown).

The Effects of the Methoxy and Dimethylamine Groups for the 3-Phenyl Alignment

Derivatives with proximal methoxy groups (**Figure 2**), especially at the R7 position, assure relatively strong MAO-B inhibition (**Figure 4**) and produce at best 230 nM inhibition (e.g., **8** in **Figure 2**, **Table 1**).

Based on the docking, derivatives **8** and **11** (**Figure 2**) flip their R7-methoxy groups toward the Leu164° (**Figure 5E**, Supplementary Figure S3F), which is shielded from a clash with the methoxy group by forming intra-protein H-bond with the Phe168^N (not shown). Inserting an extra R5-methoxy into the 3-phenyl of **8** to produce otherwise identical derivative **9** (**Figure 2**) weakens the inhibition slightly (IC$_{50}$ difference of 23 nM; **Table 1**), because the added methoxy group is unable to form particularly favorable interactions with the nearby Pro102° (**Figure 5F**). With derivatives **10** or **13** (**Figure 2**), the methoxy group is added to the phenyl ring's para position, and due to the planarity of the 3-phenylcoumarin scaffold, there is an energetic penalty for pushing the group toward either side of the cavity end (red sector in **Figures 3A,B**). Accordingly, to avoid a scaffold misalignment, the R6-methoxy group of **10** (and **13**) points directly toward the side chains of Phe103, Pro104, Trp119, and Ile199 (Supplementary Figures S3E,F), which, in turn, produces roughly 170 nM difference in the IC$_{50}$ values with otherwise

identical **8** (**Figure 5E**, **Table 1**) in favor of the R7-methoxy position.

A dimethylamine group at the 3-phenyl ring's para position (a.k.a. dimethylaniline; **Figure 2**) produces moderately strong MAO-B inhibition (**Table 1**) for derivatives **17** (**Figure 2**; IC$_{50}$ value of 400 nM), **18** (**Figure 2**; IC$_{50}$ value of 798 nM), and **19** (**Figure 2**; IC$_{50}$ value of 955 nM). This is due to the ability of the R6-dimethylamine to fill the cavity end (red sector in **Figures 3A,B**) similarly to the R6-trifluoromethyl of **1** (**Figures 5A,B** vs. **Figures 6C,D**, Supplementary Figure S3J). The downside is that the bulkier R6-substituent cannot form halogen or hydrogen bonds with water or residues nor push against either side of the cavity and, most importantly, it causes unfavorable coumarin alignment. Accordingly, the R6-dimethylamine of derivatives **17–19** packs directly against the side chains of Phe103, Pro104, Trp119, Leu164, and Ile316 (**Figures 4C,D**, Supplementary Figure S3J).

Refining the Alignment via the R1–R3 Substituents of the Coumarin Ring

Inserting a functional group such as methoxy to the R1/R2 position of the coumarin ring (**Figure 2**), capable of forming both hydrophobic and hydrophilic interactions, generally improves the MAO-B inhibition (**Figure 4**, **Table 1**).

The benefits of this sort of dual-purpose group are evident when comparing the activity of otherwise identical derivatives with and without the proximal group; i.e., **11**, that lacks only the R1-methoxy of **8** (Supplementary Figure S3F vs. **Figure 5E**),

produces significantly lower inhibition (IC_{50} value of 798 vs. 231 nM; **Table 1**). On one hand, the methyl of the R1-methoxy group of **8** (**Figure 5E**) packs into a hydrophobic niche formed by the side chains of Tyr60, Gln206, Tyr326, Leu328, Phe343, and Met341 (green sector in **Figures 3A,B**). On the other hand, the methoxy's oxygen increases the 3-phenyl ring's hydrophilicity and softens the clash of the coumarin ring with the solvent shielding the cofactor (**Figure 5E**).

Switching the R1-methoxy of **1** into the R2 position in **2** (**Figure 2**) makes the alignment of the coumarin ring more challenging, because the R2-methoxy is unable to occupy the same hydrophobic niche (green sector in **Figures 3A,B**) as the R1-methoxy (**Figure 4A** vs. Supplementary Figure S3A). Although the R1/R2 methoxy switch, by all means, does not prevent binding, it leads to ~80 nM reduction in the IC_{50} value (**Table 1**). Paradoxically, the opposite and considerably larger difference in inhibition is produced by the R1/R2 switch, when comparing the activity of derivatives **20** and **22** (**Figure 2**; **Table 1**). Accordingly, **20** with the R2-methoxy of (IC_{50} value of 391 nM; **Table 1**) provides twice as strong inhibition as **22** with the R1-methoxy (IC_{50} value of 831 nM; **Table 1**). The vast difference is caused by the coordinated R6/R7 interactions of the 3-phenyl ring, which pushes the coumarin ring closer to the Tyr326 side chain—a critical shift that is stunted by the R1-methoxy of **22** (**Figure 5E** vs. **Figure 5F**).

Replacing the R2-acetoxy of **3** (**Figure 2**) with the R1-methoxy in **4** (**Figure 2**) weakens the inhibition ~180 nM (**Table 1**). The coumarin ring of **4** is pushed closer to the cofactor due to the addition of the R6-trifluoromethoxy into the 3-phenyl ring (**Figure 5B** vs. Supplementary Figure S3B) and, in this new pose, the methyl of the R2-acetoxy is able to occupy the small hydrophobic niche (green sector in **Figures 3A,B**), meanwhile, exposing the acetoxy's oxygen atoms to the solvent (**Figure 3B**). However, substituting the R1-methoxy of **18** with the R2-acetoxy in **19** (**Figure 2**) does not improve the inhibition; instead, the IC_{50} value is reduced by ~250 nM (**Table 1**). This happens, because the R6-dimethylamine of **19** (Supplementary Figure S3J) is not forcing the scaffold to align close to the cofactor the same way as the R6-trifluoromethoxy does (**Figure 5B** vs. **Figures 6C,D**). In contrast, replacing the R1-methoxy of **18** with the R2-hydroxyl in **17** improves the inhibition (IC_{50} improvement of 234 nM; **Table 1**) by promoting water solubility near the cofactor (**Figure 6C** vs. **Figure 6D**).

The R6 and R7 interactions of **7** (**Figure 2**) are expected to remind closely those of **6** (Supplementary Figure S3D vs. **Figure 5C**), but its coumarin ring's R1- and R3-chlorine groups weaken the inhibition ~700 nM (**Table 1**). The R2-methoxy of **6** is able to play into the hydrophobic/hydrophilic dual nature of the cavity end facing the cofactor (**Figure 5C**) without occupying the small hydrophobic niche (green sector in **Figures 3A,B**). In this respect, the R1-chlorine is too bulky to occupy this specific niche although a methoxy group at the same position should be able to occupy the available space (e.g., **1** in **Figure 5A**).

Selectivity of the 3-Phenylcoumarin Derivatives

Determining the specificity and subtype selectivity of the 3-phenylcoumarin derivatives for MAO-B is needed to evaluate their true pharmacological potential. Unintended off-target effects with other proteins can render even the most promising drug candidates useless, ambiguous or even toxic. Here, the focus is put on MAO-A which has shared activity with MAO-B in deamination of dopamine and dietary amines tyramine and tryptamine. In addition, the effects of the derivatives are tested with a specific subset of enzymes, including HSD1, aromatase, CYP1A2, and ER, whose function is linked to different stages of estradiol action and metabolism. These particular enzymes were looked at with the derivatives, because they are known to have structurally similar ligands or even coumarin-based inhibitors based on prior studies and our upcoming study (Mattsson et al., 2014; Niinivehmas et al., 2016; Niinivehmas et al., unpublished results).

Monoamine oxidase A (MAO-A) is more prevalent than the subtype B in the gastrointestinal tract and, accordingly, the MAO-A inhibition can cause accumulation of tyramine from dietary sources. Because tyramine can displace neurotransmitters leading to potentially fatal hypertensive crisis, it is highly desirable to design MAO-B-specific inhibitors lacking MAO-A activity. The vast majority of the novel derivatives do not produce MAO-A inhibition at 100 μM despite the fact that it is ten times the concentration used in this study to determine MAO-B inhibition percentage (**Table 1**, Supplementary Table S1). Furthermore, only in those few cases where inhibition was detected, especially with the most potent MAO-B derivatives, it remains at moderate or close to non-existent level (**Table 1**). The strongest MAO-A inhibition was elicited by derivatives **42** and **43** (48.86 and 56.76%), but derivatives **27** and **45** (43.83 and 43.36%) are close runner-ups and next analogs down the list are already much weaker (Supplementary Figure S4, Supplementary Table S1). Notably, **1**, which is the most potent MAO-B inhibitor of the derivative set with the IC_{50} value of 56 nM, does not produce MAO-A inhibition at 100 μM (**Table 1**). The molecular basis for the lack of MAO-A activity is evident, when comparing the shape and size of the active sites of the two enzyme subtypes in the context of 3-phenylcoumarin binding (**Figure 3A** vs. **Figure 3B**).

17-β-hydroxysteroid dehydrogenase 1 (HSD1), which functions as the catalyst of the final reducing step in the estradiol biosynthesis, is often overexpressed in breast cancer and endometriotic tissue (Vihko et al., 2004; Dassen et al., 2007; Hanamura et al., 2014). Thus, specific inhibition of HSD1 has potential to reduce effective estradiol levels in the treatments. Although the synthesized 3-phenylcoumarin set contains several molecules that exhibit activity toward HSD1, the inhibition was generally very weak and the active compounds are not among the most potent MAO-B inhibitors. Of the 24 most potent MAO-B inhibitors, the strongest HSD1 inhibition could be recorded for **20** and **22** (46 and 54%; **Figure 2**, **Table 1**); however, considerably higher activity (48.20–83.90%) was seen with derivatives **30**, **31**, **33**, **38**, and **48** (Supplementary

Figure S4, Supplementary Table S1). Modest HSD1 inhibition (12–33%) was also elicited by **6**, **15**, **16**, **23**, **24** (**Figure 2**, **Table 1**) and **51** (Supplementary Figure S4, Supplementary Table S1). Importantly, derivative **1**, which is the most potent MAO-B inhibitor of the derivative set, does not inhibit HSD1.

Aromatase (CYP19A1) inhibition, which is important for blocking local estradiol synthesis for example in breast cancer treatment (Pasqualini et al., 1996), was not detected with the derivatives (**Table 1**, Supplementary Table S1). Although 3-phenylcoumarin should be able to sterically mimic the steroidal positioning at the active site (not shown), it would have to house a clear-cut H-bond acceptor at the R5/R7-position in the 3-phenyl to facilitate aromatase binding. This is, because X-ray crystallography shows that the Asp309 side chain is in neutral state at pH 7.4 and donating a proton to the carbonyl group of inhibitor androstenedione (PDB: 3EQM) (Ghosh et al., 2009). Inserting a hydroxyl group to the R5/R7 position could put an H-bond acceptor to this same location with the 3-phenylcoumarins (see **31**, **38**, **40**, **42**, **43**; Supplementary Figure S4, Supplementary Table S1). However, because the hydroxyl always has a dual role as an H-bond donor as well, any aromatase binding by the derivatives remains theoretical as it is prevented by a proton donor clash. The issue is described more thoroughly in our upcoming study (Niinivehmas et al., unpublished results).

Estrogen receptor (ER) agonists/antagonists or selective modulators are developed for infertility, contraception, hormone replacement, and ER positive breast cancer therapies. If the MAO-B inhibitors would function also as ER agonists, they could promote tumorigenesis in the breast tissue as a side effect. Unintended ER inhibition could also disturb natural estrogen levels or interrupt ER-targeted therapies. The measurements indicate that the 3-phenylcoumarin derivatives either are a hit or miss when considering ER inhibition. Although the ER activity could not be measured for all of the analogs due to running out of the synthesis products, the acquired results overwhelmingly support our prior findings stating that the R2-hydroxyl or the R6-hydroxyl/halogen is needed to prompt ER activity (Niinivehmas et al., 2016). This ER-specific effect is prominent with **12**, **20**, **22**, **27**, **28**, **29**, **30**, **39**, **40**, **41**, **44**, and **48** (**Table 1**, Supplementary Table S1, **Figure 2**, Supplementary Figure S4) and, moreover, ER activity is predicted for **17** and likely for **32** and **47** based on the well-established trend.

Cytochrome P450 1A2 (CYP1A2) catalyzes the oxidation of xenobiotics, especially polyaromatic hydrocarbons and steroid hormone-sized compounds such as 3-phenylcoumarins, into more soluble form for excretion (Zhou et al., 2010). Accordingly, it was prudent to get a rough estimate of the CYP1A2 inhibition levels for the novel 3-phenylcoumarin derivatives as well. In general, all of the derivatives inhibited CYP1A2 at some level (**Table 1**, Supplementary Table S1); however, typically the most potent CYP1A2 inhibitors such as **21**–**24** were less potent MAO-B inhibitors (**Table 1**). Similar to MAO-A, HSD1, and aromatase, the most potent MAO-B derivative **1** displayed only low CYP1A2 activity (IC$_{50}$ value of 124 μM; **Table 1**).

Overall Assessment on the Druglikeness

As a whole, the selectivity analysis indicates that the cross-reactivity of 3-phenylcoumarins can be managed or even avoided via specific functional group substitutions without taking away the MAO-B activity. Coumarins in general do not belong to the PAINS (pan assay interference compounds) category as it is a privileged scaffold structure. Only derivative **50**, which is not a potent MAO-B inhibitor (Supplementary Table S1, Supplementary Figure S4), was recognized as a potential PAINS ligand by PAINS3 filter (or A filter) in CANVAS module in MAESTRO (Baell and Holloway, 2010). In the ChEMBL database, ~14,200 coumarin derivatives are included (observed online in 8.2.2018), which indicates that the scaffold can be tailored to target multitude of proteins. Despite this, the literature does not raise widespread concerns that the coumarin-based compounds in particular would cause harmful cross-reactivity or selectivity issues. The 24 active derivatives presented in this study (**Table 1**, **Figure 2**) have lower potency than some of the prior 3-phenylcoumarin compounds (Supplementary Figure S6, Supplementary Table S2) (Matos et al., 2009b, 2011a,b; Santana et al., 2010; Viña et al., 2012a); however, one has to be aware of fact that these results originate from different laboratories and activity assays and are, therefore, not fully comparable. To a degree this is the case even for the positive control pargyline (Fišar et al., 2010). Importantly, the new compounds follow closely the Lipinski rule of five regarding the logP value (logP < 5) and remain in the logP range of 2–4. Moreover, the ligand-lipophilicity efficiency (LiPE) values of the new analogs suggest reasonable druglikeness (Freeman-Cook et al., 2013). What is more, derivative **1** clearly has the most promising selectivity profile of the derivatives for future consideration, because it is not only the most potent MAO-B inhibitor of the set but it is also selective against the other tested enzymes.

CONCLUSION

A broad set of 3-phenylcoumarin derivatives was designed using virtual combinatorial chemistry or rationally *de novo*, synthesized and tested for MAO-B inhibition potency using spectrophotometry (Supplementary Table S1). The results further validate prior studies suggesting that the 3-phenylcoumarin is a suitable scaffold for building potent small-molecule MAO-B inhibitors by functionalizing its ring systems. A moderate MAO-B inhibition could be achieved by inserting a wide variety of functional groups into the coumarin (R1–R3; **Figure 4**) or 3-phenyl (R4–R7; **Figure 4**) rings (Supplementary Table S1). Twenty-four of the derivatives (**Figures 2**, **3D**) were found to elicit >70% inhibition (**Table 1**, Supplementary Figures S1, S2). These promising derivatives inhibit the MAO-B at a ~100 nM to ~1 μM range (**Table 1**), while the most potent derivative **1** produces ~56 nM MAO-B inhibition. A molecular docking-based (**Figures 5**, **6**, Supplementary Figure S3) SAR analysis (**Figure 4**) describe the determinants of the MAO-B binding and inhibition at the atomistic level. Firstly, without any kind of the 3-phenyl substituents, no inhibition was detected. Although both hydrogen and halogen bonding can assist the 3-phenyl alignment and facilitate inhibition (**Figures 6E,F**, **Table 1**), the

ability of the functionalized ring to fill the hydrophobic end of the binding cavity (red sector in **Figures 3A,B**) is the most important property for ensuring strong MAO-B inhibition (e.g., R6-trifluoromethyl of **1**; **Figure 5A**). Secondly, the SAR analysis reveals that a spot-on placement and composition of the coumarin ring's substituents can further enhance the MAO-B inhibition (**Figure 2**, **Table 1**), however, these effects are ultimately dependent on the scaffold alignment, which, in turn, depends on the 3-phenyl ring substituents (**Figure 4**). The cross-reactivity analysis focusing on MAO-A and a subset of estradiol metabolism-linked HSD1, aromatase, CYP1A2 and ER highlighted the potential of the 3-phenylcourmains, especially the most potent MAO-B derivative **1**, for producing selective MAO-B inhibition. Finally, the most potent 3-phenylcoumarin analogs presented in this study are estimated to operate at close to optimal ligand-lipophilicity efficiency—a feature highlighting their overall druglikeness.

AUTHOR CONTRIBUTIONS

SR: was responsible for the experimental testing regarding MAO-A and MAO-B; EMu: performed the MAO-A experimental analysis; SR: did the docking into MAO-B and prepared most of the figures; PAP: was responsible for the final SAR analysis; SK, ES, and JH: performed the organic synthesis; MA: did the PAINS screening; PK: performed the HSD1 measurements; NN, RJ, and HR: did the experimental analysis regarding CYP1A2; PH and MaP: executed the experimental analysis regarding aromatase; MiP: did preliminary screening for designing MAO-B ligands; SN, EMa, SK, and OTP: designed the molecules for the selected targets; SN, EMa, and OTP: designed the study. All the coauthors were involved in the manuscript preparation and approved the final version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2018.00041/full#supplementary-material

REFERENCES

Baell, J. B., and Holloway, G. A. (2010). New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* 53, 2719–2740. doi: 10.1021/jm901137j

Bhandri, P. R., Bose, J. L., and Siddiqui, S. (1949). 3-Phenylcoumarin series. I. Synthesis of some new members of the series. *J. Sci. Ind. Res.* 8B, 189–192.

Binda, C., Li, M., Hubalek, F., Restelli, N., Edmondson, D. E., and Mattevi, A. (2003). Insights into the mode of inhibition of human mitochondrial monoamine oxidase B from high-resolution crystal structures. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9750–9755. doi: 10.1073/pnas.1633804100

Binda, C., Wang, J., Pisani, L., Caccia, C., Carotti, A., Salvati, P., et al. (2007). Structures of human monoamine oxidase B complexes with selective noncovalent inhibitors: safinamide and coumarin analogs. *J. Med. Chem.* 50, 5848–5852. doi: 10.1021/jm070677y

Borges, F., Roleira, F., Milhazes, N., Santana, L., and Uriarte, E. (2005). Simple coumarins and analogues in medicinal chemistry: occurrence, synthesis and biological activity. *Curr. Med. Chem.* 12, 887–916. doi: 10.2174/0929867053507315

Carradori, S., and Silvestri, R. (2015). New frontiers in selective human MAO-B inhibitors miniperspective. *J. Med. Chem.* 58, 6717–6732. doi: 10.1021/jm501690r

Catto, M., Nicolotti, O., Leonetti, F., Carotti, A., Favia, A. D., Soto-Otero, R., et al. (2006). Structural insights into monoamine oxidase inhibitory potency and selectivity of 7-substituted coumarins from ligand- and target-based approaches. *J. Med. Chem.* 49, 4912–4925. doi: 10.1021/jm060183l

Cavallo, G., Metrangolo, P., Milani, R., Pilati, T., Priimagi, A., Resnati, G., et al. (2016). The halogen bond. *Chem. Rev.* 116, 2478–2601. doi: 10.1021/acs.chemrev.5b00484

Chauhan, P., Ravi, M., Singh, S., Prajapati, P., and Yadav, P. P. (2016). Regioselective [small alpha]-arylation of coumarins and 2-pyridones with phenylhydrazines under transition-metal-free conditions. *RSC Adv.* 6, 109–118. doi: 10.1039/C5RA20954D

Dassen, H., Punyadeera, C., Kamps, R., Delvoux, B., Van Langendonckt, A., Donnez, J., et al. (2007). Estrogen metabolizing enzymes in endometrium and endometriosis. *Hum. Reprod.* 22, 3148–3158. doi: 10.1093/humrep/dem310

Dobelmann-Mara, L., Riedmueller, S., and Schraub, M. (2017). *Compounds for Optically Active Devices.* PCT Int. Appl. A1 20170302.

Edmondson, D. E., Mattevi, A., Binda, C., Li, M., and Hubálek, F. (2005). Structure and mechanism of monoamine oxidase. *Burger's Med. Chem. Drug Discov.* 11, 1983–1993. doi: 10.1002/0471266949.bmc111

Ferino, G., Cadoni, E., Matos, M. J., Quezada, E., Uriarte, E., Santana, L., et al. (2013). MAO Inhibitory activity of 2-arylbenzofurans versus 3-arylcoumarins: synthesis, *in vitro* study, and docking calculations. *ChemMedChem* 8, 956–966. doi: 10.1002/cmdc.201300048

Finberg, J. P. M., and Rabey, J. M. (2016). Inhibitors of MAO-A and MAO-B in psychiatry and neurology. *Front. Pharmacol.* 7:340. doi: 10.3389/fphar.2016.00340

Fisar, Z., Hroudová, J., and Raboch, J. (2010). Inhibition of monoamine oxidase activity by antidepressants and mood stabilizers. *Neuroendocrinol. Lett.* 31, 645–656.

Freeman-Cook, K. D., Hoffman, R. L., and Johnson, T. W. (2013). Lipophilic efficiency: the most important efficiency metric in medicinal chemistry. *Future Med. Chem.* 5, 113–115. doi: 10.4155/fmc.12.208

Gaweska, H., and Fitzpatrick, P. F. (2011). Structures and mechanism of the monoamine oxidase family. *Biomol. Concepts* 2, 365–377. doi: 10.1515/BMC.2011.030

Ghosh, D., Griswold, J., Erman, M., and Pangborn, W. (2009). Structural basis for androgen specificity and oestrogen synthesis in human aromatase. *Nature* 457, 219–223. doi: 10.1038/nature07614

Hanamura, T., Niwa, T., Gohno, T., Kurosumi, M., Takei, H., Yamaguchi, Y., et al. (2014). Possible role of the aromatase-independent steroid metabolism pathways in hormone responsive primary breast cancers. *Breast Cancer Res. Treat.* 143, 69–80. doi: 10.1007/s10549-013-2788-3

Holt, A., Sharman, D. F., Baker, G. B., and Palcic, M. M. (1997). A continuous spectrophotometric assay for monoamine oxidase and related enzymes in

tissue homogenates. *Anal. Biochem.* 244, 384–392. doi: 10.1006/abio.1996.9911

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. *J. Mol. Graph.* 14, 33–8, 27–8. doi: 10.1016/0263-7855(96)00018-5

Joao Matos, M., Vina, D., Vazquez-Rodriguez, S., Uriarte, E., and Santana, L. (2013). Focusing on new monoamine oxidase inhibitors: differently substituted coumarins as an interesting scaffold. *Curr. Top. Med. Chem.* 12, 2210–2239. doi: 10.2174/1568026611212200008

Kirkiacharian, S., Chidiack, H., Philibert, D., Van De Velde, P., and Bouchoux, F. (1999). Binding affinity to steroid hormone receptors and antiproliferative action on MCF-7 cells of coumarinic derivatives and isoflavonoids. *Ann. Pharm. Fr.* 57, 332–339.

Kirkiacharian, S., Lormier, A. T., Resche-Rigon, M., Bouchoux, F., and Cerede, E. (2003). Synthesis and binding affinity of 3-aryl-7-hydroxycoumarins to human α and β estrogen receptors. *Ann. Pharm. Fr.* 61, 51–56.

Korb, O., Stützle, T., and Exner, T. E. (2009). Empirical scoring functions for advanced protein-ligand docking with PLANTS. *J. Chem. Inf. Model.* 49, 84–96. doi: 10.1021/ci800298z

Korhonen, L. E., Rahnasto, M., Mähönen, N. J., Wittekindt, C., Poso, A., Juvonen, R. O., et al. (2005). Predictive three-dimensional quantitative structure-activity relationship of cytochrome P450 1A2 inhibitors. *J. Med. Chem.* 48, 3808–3815. doi: 10.1021/jm0489713

Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24, 946–950. doi: 10.1107/S0021889891004399

Lehtonen, J. V., Still, D.-J., Rantanen, V.-V., Ekholm, J., Björklund, D., Iftikhar, Z., et al. (2004). BODIL: a molecular modeling environment for structure-function analysis and drug design. *J. Comput. Aided. Mol. Des.* 18, 401–419. doi: 10.1007/s10822-004-3752-4

Lomize, M. A., Lomize, A. L., Pogozheva, I. D., and Mosberg, H. I. (2006). OPM: orientations of proteins in membranes database. *Bioinformatics* 22, 623–625. doi: 10.1093/bioinformatics/btk023

Matos, M. J., Terán, C., Pérez-Castillo, Y., Uriarte, E., Santana, L., and Viña, D. (2011a). Synthesis and study of a series of 3-arylcoumarins as potent and selective monoamine oxidase B inhibitors. *J. Med. Chem.* 54, 7127–7137. doi: 10.1021/jm200716y

Matos, M. J., Vazquez-Rodriguez, S., Uriarte, E., Santana, L., and Viña, D. (2011b). MAO inhibitory activity modulation: 3-Phenylcoumarins versus 3-benzoylcoumarins. *Bioorganic Med. Chem. Lett.* 21, 4224–4227. doi: 10.1016/j.bmcl.2011.05.074

Matos, M. J., Viña, D., Janeiro, P., Borges, F., Santana, L., and Uriarte, E. (2010). New halogenated 3-phenylcoumarins as potent and selective MAO-B inhibitors. *Bioorganic Med. Chem. Lett.* 20, 5157–5160. doi: 10.1016/j.bmcl.2010.07.013

Matos, M. J., Viña, D., Picciau, C., Orallo, F., Santana, L., and Uriarte, E. (2009a). Synthesis and evaluation of 6-methyl-3-phenylcoumarins as potent and selective MAO-B inhibitors. *Bioorganic Med. Chem. Lett.* 19, 5053–5055. doi: 10.1016/j.bmcl.2009.07.039

Matos, M. J., Viña, D., Quezada, E., Picciau, C., Delogu, G., Orallo, F., et al. (2009b). A new series of 3-phenylcoumarins as potent and selective MAO-B inhibitors. *Bioorg. Med. Chem. Lett.* 19, 3268–3270. doi: 10.1016/j.bmcl.2009.04.085

Mattsson, C., Svensson, P., and Sonesson, C. (2014). A novel series of 6-substituted 3-(pyrrolidin-1-ylmethyl)chromen-2-ones as selective monoamine oxidase (MAO) A inhibitors. *Eur. J. Med. Chem.* 73, 177–186. doi: 10.1016/j.ejmech.2013.11.035

Merritt, E. A., and Murphy, M. E. P. (1994). Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 50, 869–873. doi: 10.1107/S0907444994006396

Messinger, J., Husen, B., Koskimies, P., Hirvelä, L., Kallio, L., Saarenketo, P., et al. (2009). Estrone C15 derivatives–a new class of 17beta-hydroxysteroid dehydrogenase type 1 inhibitors. *Mol. Cell. Endocrinol.* 301, 216–224. doi: 10.1016/j.mce.2008.10.022

Mokkila, S., Postila, P. A., Rissanen, S., Juhola, H., Vattulainen, I., and Róg, T. (2017). Calcium assists dopamine release by preventing aggregation on the inner leaflet of presynaptic vesicles. *ACS Chem. Neurosci. Acschemneuro.* 8, 1242–1250. doi: 10.1021/acschemneuro.6b00395

Niinivehmas, S. P., Manivannan, E., Rauhamäki, S., Huuskonen, J., and Pentikäinen, O. T. (2016). Identification of estrogen receptor α ligands with virtual screening techniques. *J. Mol. Graph. Model.* 64, 30–39. doi: 10.1016/j.jmgm.2015.12.006

Niinivehmas, S. P., Salokas, K., Lätti, S., Raunio, H., and Pentikäinen, O. T. (2015). Ultrafast protein structure-based virtual screening with Panther. *J. Comput. Aided Mol. Des.* 29, 989–1006. doi: 10.1007/s10822-015-9870-3

Niinivehmas, S. P., Virtanen, S. I., Lehtonen, J. V., Postila, P. A., and Pentikäinen, O. T. (2011). Comparison of virtual high-throughput screening methods for the identification of phosphodiesterase-5 inhibitors. *J. Chem. Inf. Model.* 51, 1353–1363. doi: 10.1021/ci1004527

Nurminen, E. M., Pihlavisto, M., Lázár, L., Pentikäinen, U., Fülöp, F., and Pentikäinen, O. T. (2011). Novel hydrazine molecules as tools to understand the flexibility of vascular adhesion protein-1 ligand-binding site: toward more selective inhibitors. *J. Med. Chem.* 54, 2143–2154. doi: 10.1021/jm200059p

Nurminen, E. M., Pihlavisto, M., Lázár, L., Szakonyi, Z., Pentikäinen, U., Fülöp, F., et al. (2010). Synthesis, *in vitro* activity, and three-dimensional quantitative structure–activity relationship of novel hydrazine inhibitors of human vascular adhesion protein-1. *J. Med. Chem.* 53, 6301–6315. doi: 10.1021/jm100337z

Orłowski, A., Grzybek, M., Bunker, A., Pasenkiewicz-Gierula, M., Vattulainen, I., Männistö, P. T., et al. (2012). Strong preferences of dopamine and l -dopa towards lipid head group: importance of lipid composition and implication for neurotransmitter metabolism. *J. Neurochem.* 122, 681–690. doi: 10.1111/j.1471-4159.2012.07813.x

Pasanen, M. (1985). Human placental aromatase activity: use of a C18 reversed-phase cartridge for separation of tritiated water or steroid metabolites in placentas from both smoking and non-smoking mothers *in vitro. Biol. Res. Pregnancy Perinatol.* 6, 94–99.

Pasqualini, J., Chetrite, G., Blacker, M., Feinstein, M.-C., Delalonde, L., Talbi, M., et al. (1996). Concentrations of estrone, estradiol, and estrone sulfate and evaluation of sulfatase and aromatase activities in pre- and postmenopausal breast cancer patients. *J. Clin. Endocrinol. Metab.* 81, 1460–1464.

Patil, P. O., Bari, S. B., Firke, S. D., Deshmukh, P. K., Donda, S. T., and Patil, D. A. (2013). A comprehensive review on synthesis and designing aspects of coumarin derivatives as monoamine oxidase inhibitors for depression and Alzheimer's disease. *Bioorganic Med. Chem.* 21, 2434–2450. doi: 10.1016/j.bmc.2013.02.017

Postila, P. A., Vattulainen, I., and Róg, T. (2016). Selective effect of cell membrane on synaptic neurotransmission. *Sci. Rep.* 6:19345. doi: 10.1038/srep19345

Prendergast, P. T. (2001). *Use of Flavones, Coumarins and Related Compounds to Treat Infections.* PCT Int. Appl., A2 20010118.

Santana, L., Orallo, F., Viña, D., Matos Joao Correia, P. C., Quezada, E., Yañes, J., et al. (2010). *Use of Derivates of 6-Substituted 3-Phenylcoumarins and Preparation of New Derivates.* Available online at: http://www.google.com.pg/patents/WO2010086484A1?cl=en

Serra, S., Ferino, G., Matos, M. J., Vázquez-Rodríguez, S., Delogu, G., Viña, D., et al. (2012). Hydroxycoumarins as selective MAO-B inhibitors. *Bioorganic Med. Chem. Lett.* 22, 258–261. doi: 10.1016/j.bmcl.2011.11.020

Shih, J. C., Chen, K., and Ridd, M. J. (1999). Monoamine oxidase: from genes to behavior. *Annu. Rev. Neurosci.* 22, 197–217. doi: 10.1146/annurev.neuro.22.1.197

Vihko, P., Härkönen, P., Soronen, P., Törn, S., Herrala, A., Kurkela, R., et al. (2004). 17β-Hydroxysteroid dehydrogenases - their role in pathophysiology. *Mol. Cell. Endocrinol.* 215, 83–88. doi: 10.1016/j.mce.2003.11.021

Vilar, S., Quezada, E., Santana, L., Uriarte, E., Yánez, M., Fraiz, N., et al. (2006). Design, synthesis, and vasorelaxant and platelet antiaggregatory activities of coumarin–resveratrol hybrids. *Bioorg. Med. Chem. Lett.* 16, 257–261. doi: 10.1016/j.bmcl.2005.10.013

Viña, D., Matos, M. J., Ferino, G., Cadoni, E., Laguna, R., Borges, F., et al. (2012a). 8-substituted 3-arylcoumarins as potent and selective MAO-B inhibitors:

synthesis, pharmacological evaluation, and docking studies. *ChemMedChem* 7, 464–470. doi: 10.1002/cmdc.201100538

Viña, D., Matos, M. J., Yáñez, M., Santana, L., and Uriarte, E. (2012b). 3-Substituted coumarins as dual inhibitors of AChE and MAO for the treatment of Alzheimer's disease. *Med. Chem. Commun.* 3, 213–218. doi: 10.1039/C1MD00221J

Youdim, M. B. H., Edmondson, D., and Tipton, K. F. (2006). The therapeutic potential of monoamine oxidase inhibitors. *Nat. Rev. Neurosci.* 7, 295–309. doi: 10.1038/nrn1883

Zhou, S.-F., Wang, B., Yang, L.-P., and Liu, J.-P. (2010). Structure, function, regulation and polymorphism and the clinical significance of human cytochrome P450 1A2. *Drug Metab. Rev.* 42, 268–354. doi: 10.3109/036025 30903286476

Supplementary Material

Structure-Activity Relationship Analysis of 3-Phenylcoumarin-Based Monoamine Oxidase B Inhibitors

Sanna Rauhamäki, Pekka A. Postila, Sanna Niinivehmas, Sami Kortet, Emmi Schildt, Mira Pasanen, Elangovan Manivannan, Mira Ahinko, Pasi Koskimies, Niina Nyberg, Pasi Huuskonen, Elina Multamäki, Markku Pasanen, Risto O. Juvonen, Hannu Raunio, Juhani Huuskonen*, Olli T. Pentikäinen*

* Correspondence: Juhani Huuskonen: juhani.s-p.huuskonen@jyu.fi (synthesis),
Olli T. Pentikäinen: olli.pentikainen@utu.fi (modeling, *in vitro*)

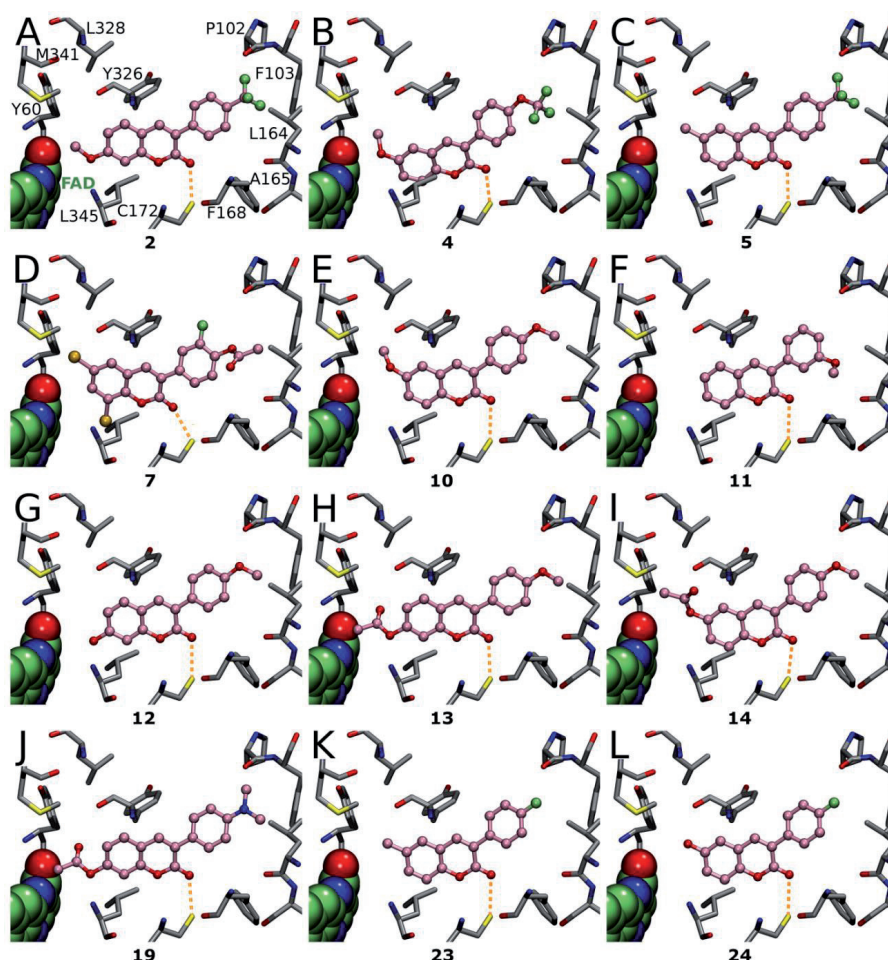## 1   Supplementary Figures and Tables

1.1   Supplementary Figures
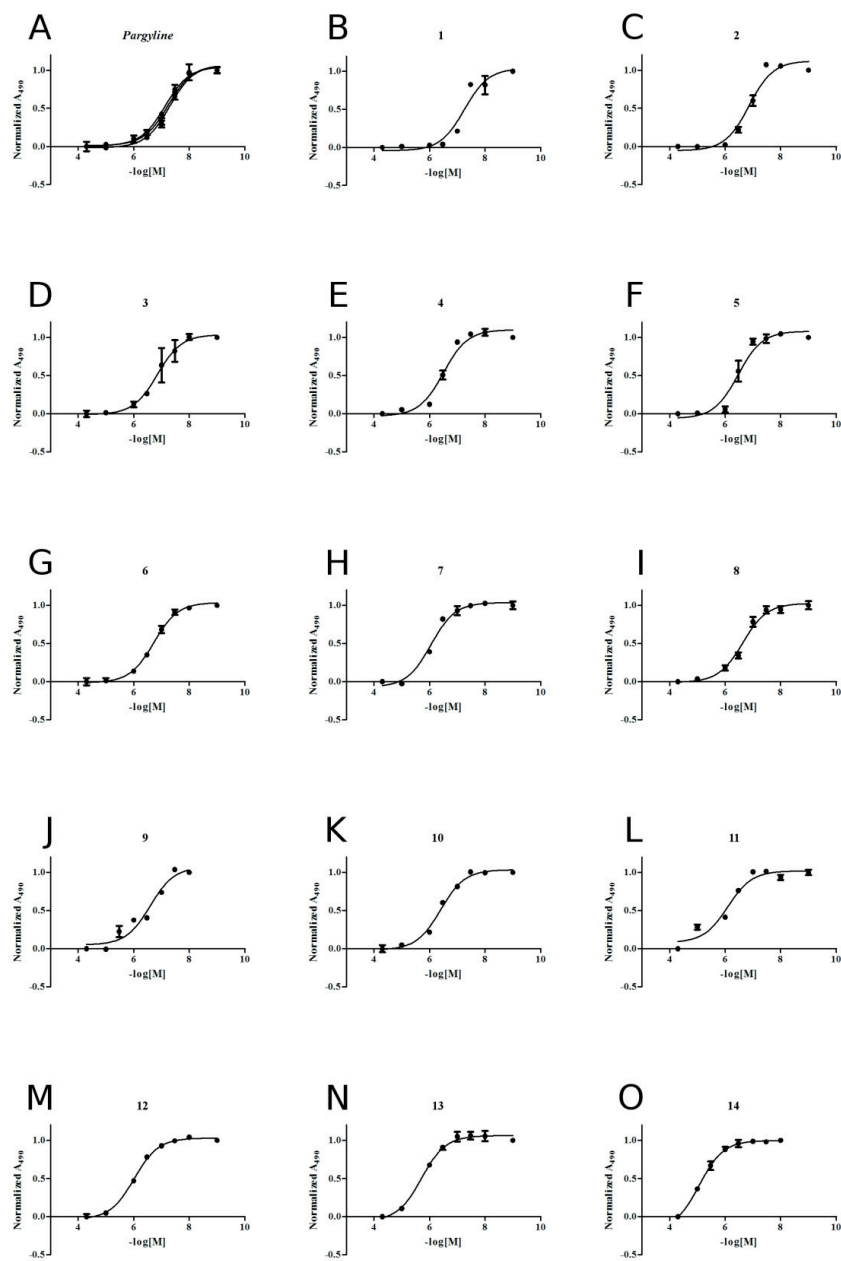
**Supplementary Figure 1.** MAO-B IC$_{50}$ graphs for pargyline and 3-phenylcoumarin derivatives 1-14.
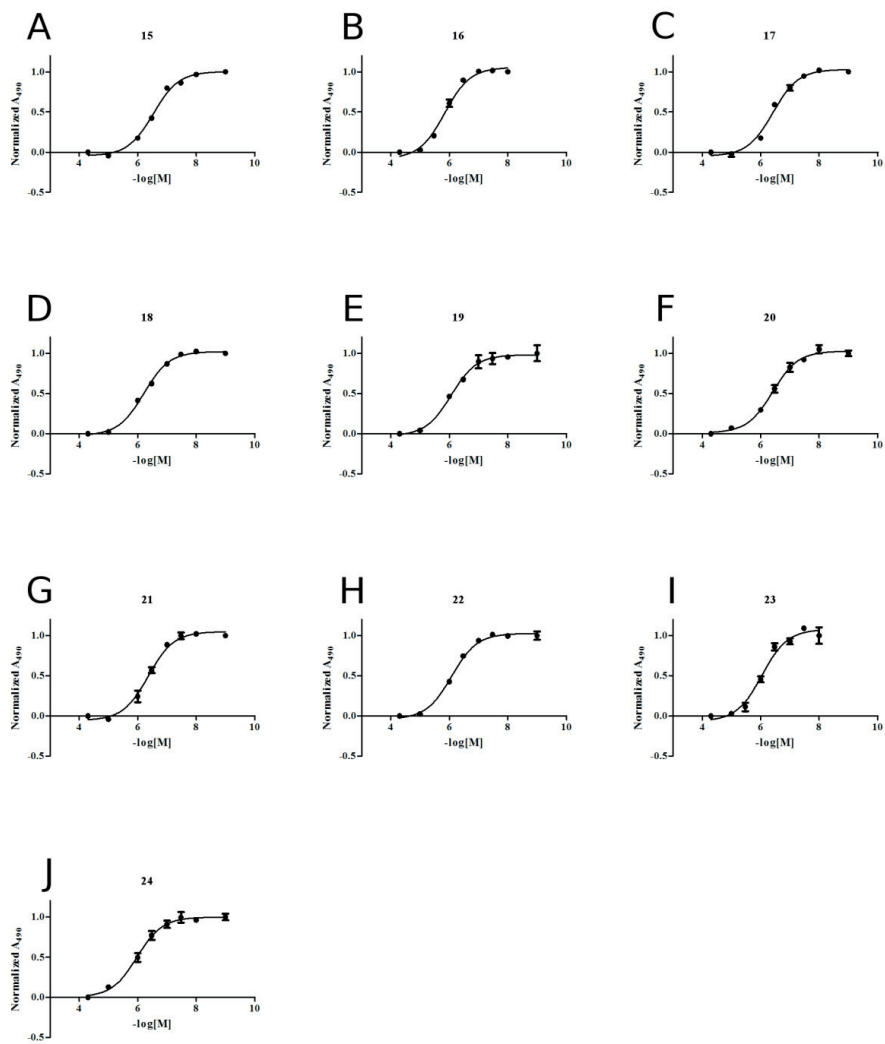
**Supplementary Figure 2.** MAO-B IC_{50} graphs for 3-phenylcoumarin derivatives 15-24.

Supplementary Figure 3. The binding modes of selected 3-phenylcoumarin inhibitors. (A) Derivatives **2** (IC_{50} 138 nM), (B) **4** (IC_{50} 317 nM) and (C) **5** (IC_{50} 343 nM) have three fluorine atoms at R6 but the R2-methoxy of **2** ensures stronger MAO-B inhibition than the R1-methoxy of **4** or the R1-methyl of **5**. (D) Derivative **7** (IC_{50} 888 nM) has halogens on R1, R3 and R6 but it is hardly more active than (K) **23** (IC_{50} 902 nM). (E) Derivatives **10** (IC_{50} 400 nM), (G) **12** (IC_{50} 955 nM), (H) **13** (IC_{50} 1946 nM) and (I) **14** (IC_{50} 8476 nM) have methoxy on R6 and (F) **11** (IC_{50} 798 nM) has it on R7. Paired with methoxy on 3-phenyl, smaller substitutes like methoxy on R1 of **10** or hydroxyl on R2 of **12** are more suitable than acetoxy on R2 of **13** or on R1 of **14**. (J) Same applies on (J) **19** (IC_{50} 866 nM) which is less active than molecules **17** and **18** (Fig. 6) with the same R6. However, the smaller substitutes on (K) **23** (IC_{50} 902 nM) and (L) **24** (IC_{50} 1058 nM) do not compare to the suitable interactions of R1 methoxy of **21** (Fig. 5) or the combination of hydroxyl on R6 and halogen on R7 of **20** and **22** (Fig. 6). See Fig. 1, Fig. 2 and Table 1 for further details.
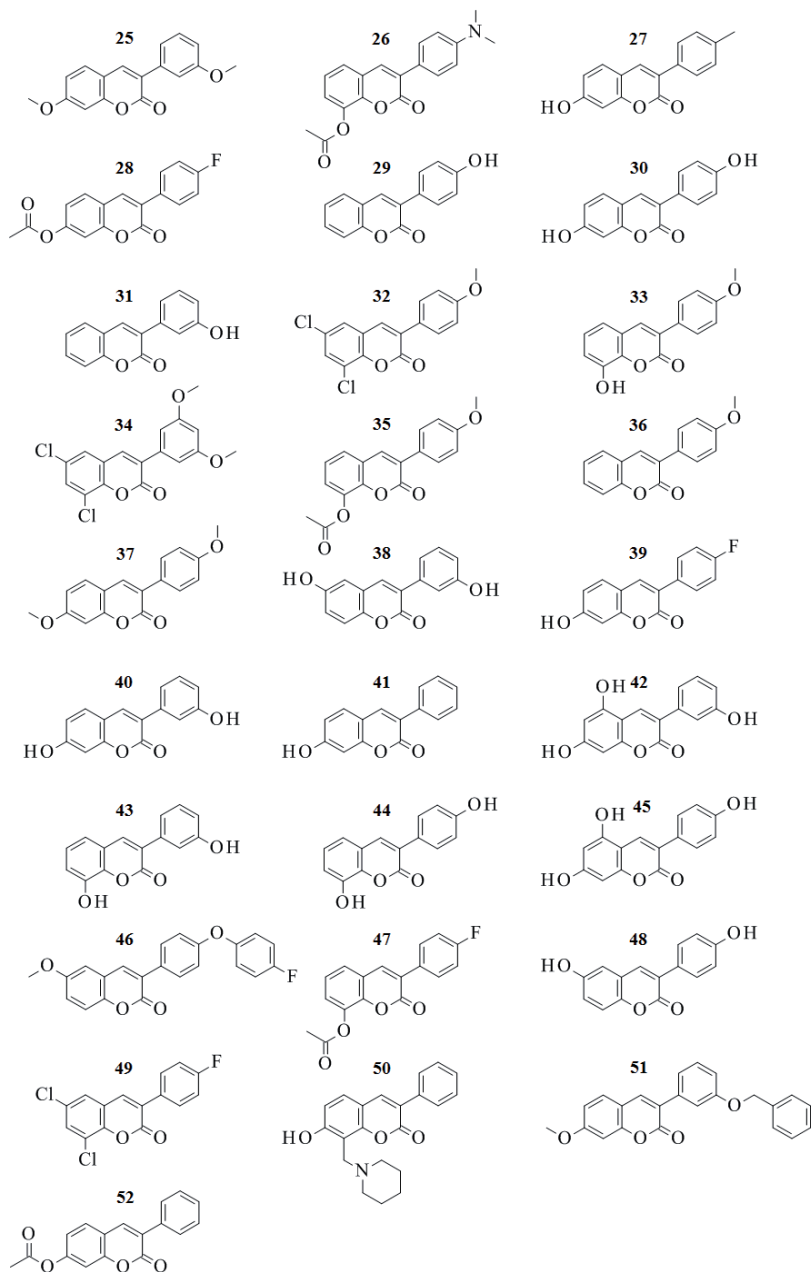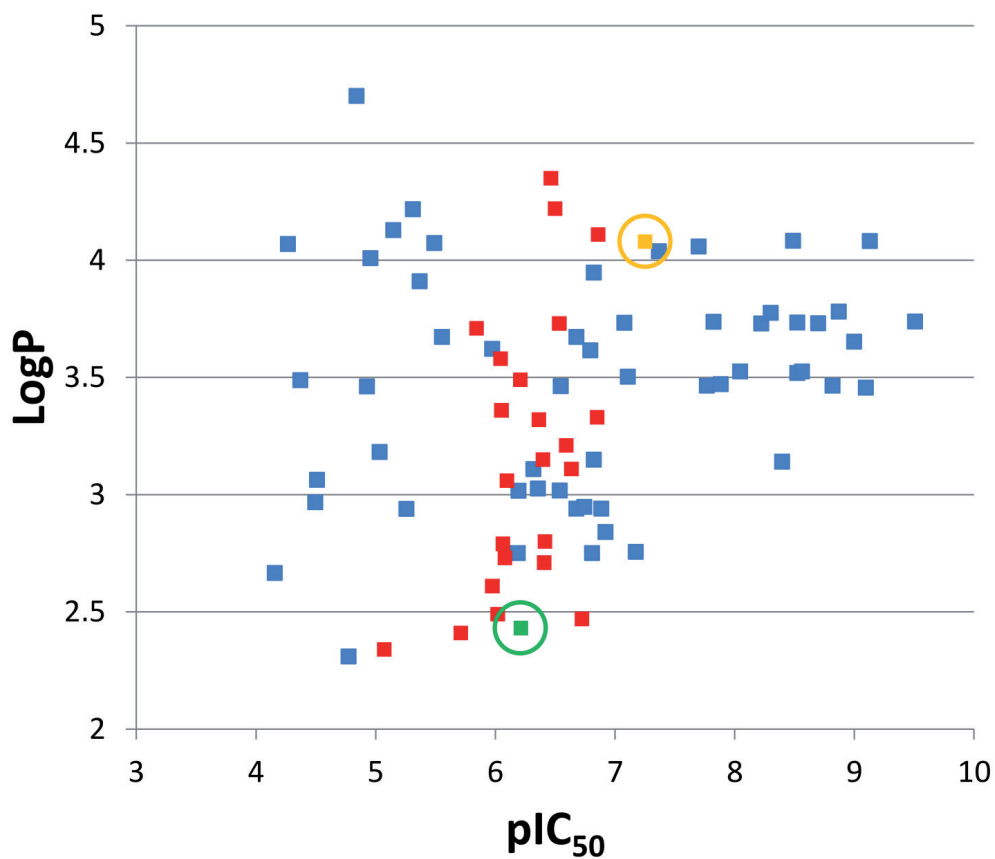
Supplementary Figure 4. The 3-phenylcoumarin derivatives 25-52. The compounds are put in descending order based on their MAO-B activity (Table S1).
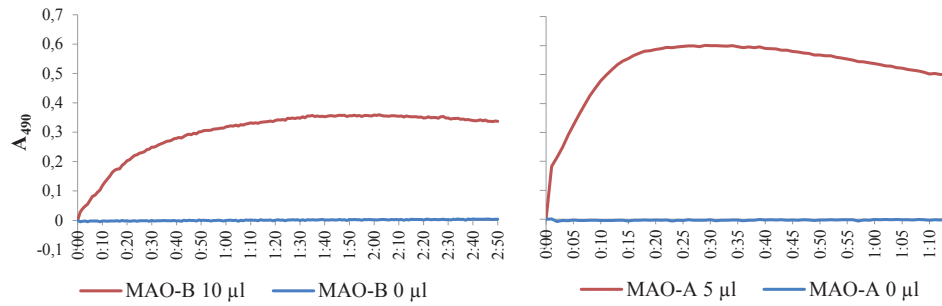
Supplementary Figure 5. MAO-B and MAO-A activity in continuous spectrophotometric assay. The monoamine oxidase (MAO) activity was measured at A_{490} every 15 s, 300 times or until plateau was reached. MAO-A was able to induce significantly faster nearly double the absorbance with the selected dose used throughout the experiments in comparison to the MAO-B. The difference was taken into account in the duration of the experiment as the MAO activity was collected for data analysis at 1 h after the initiation of the assay reaction for MAO-A and at 2 h for MAO-B.

Supplementary Figure 6. The logP *versus* pIC_{50} value for 3-phenylcoumarin-based MAO-B inhibitors. The potency of the 24 novel analogs (red dots; Table 1; Fig. 2), prior 3-phenylcoumarins (blue dots) and the positive control pargyline (green dot) is plotted against the estimated lipophilicity (logP) of the compounds (Table S2). The most potent new derivative **1** (yellow dot) and the pargyline are highlighted with yellow and green circles.

1.2 Supplementary Tables

Supplementary Table 1. Inhibition data on the 3-phenylcoumarin derivatives 25-52.

ID	QPlogPo/w	MAO-A inhibition % (100 μM, 1h)	MAO-B inhibition % (10 μM, 2h)	ER inhibition % (10 μM)	HSD1 inhibition % (1 μM)	CYP1A2 inhibition IC$_{50}$ μM
25	3.16	0.00	67.49	0.00	3.50	35.00
26	2.56	0.00	66.79	N/A	0.00	7.00
27	2.70	43.83	61.91	70.96	12.50	180.00
28	2.64	0.00	59.20	57.12	0.30	114.00
29[*]	2.40	0.00	53.65	74.22	5.40	26.00
30[*]	1.67	0.00	52.68	58.53	48.20	22.00
31[*]	2.37	4.14	49.60	0.00	56.00	7.00
32[*]	4.10	0.00	47.77	N/A	N/A	800.00
33	2.51	3.66	47.64	0.00	76.70	86.00
34	4.15	0.00	46.01	N/A	0.30	30.00
35	2.40	4.08	42.37	N/A	0.00	26.00
36[*]	3.09	0.87	41.12	0.00	3.80	26.00
37	3.15	0.00	39.78	N/A	0.00	480.00
38[*]	1.65	13.11	27.16	18.53	49.10	24.00
39[*]	2.62	12.56	26.44	100.63	0.00	240.00
40	1.65	0.00	25.74	18.58	N/A	N/A
41[*]	2.38	9.52	23.82	101.40	N/A	8.00
42	1.01	48.86	23.30	0.44	N/A	N/A
43	1.70	56.76	22.93	0.00	N/A	N/A
44[*]	1.70	22.39	22.52	96.18	N/A	N/A
45	1.01	43.38	21.38	0.13	N/A	N/A
46	4.76	8.11	19.73	0.00	0.00	240.00
47	2.46	22.33	15.81	N/A	0.00	17.00
48[*]	1.66	0.00	15.50	98.41	83.90	33.00
49	4.29	0.00	10.23	0.66	1.30	N/A
50	3.21	0.00	8.16	N/A	1.50	65.00
51	4.75	0.00	6.66	N/A	31.60	3.00
52[*]	2.43	0.00	0.00	N/A	0.70	80.00

N/A = not available. [*] Compounds **29** (Leitão et al., 2004)**, 30** (Yang et al., 2011), **31** (Leitão et al., 2004), **32** (Pu et al., 2014)**, 36** (Pu et al., 2014), **38** (Kabeya et al., 2008), **39** (CHEMBL1387945), **41** (Wang et al., 2011), **44** (Matos et al., 2013a)**, 48** (Kabeya et al., 2008) and **52** (CHEMBL1559283) had been synthesized independently prior to this study but not tested for MAO-B activity. The compounds are in descending order (**25-52**) based on their MAO-B activity.

Supplementary Table 2. MAO-B inhibition and lipophilicity data on new and prior 3-phenylcoumarins.

ID	MAO-B inhibition IC$_{50}$ nM	QPlogPo/w[1]	pIC$_{50}$	LiPE[2]	Reference
Pargyline	61	2.431	6.21	3.78	
01	56	4.08	7.25	3.17	
02	138	4.11	6.86	2.75	
03	141	3.33	6.85	3.52	(Dobelmann-Mara et al., 2017)
04	317	4.22	6.50	2.28	
05	343	4.35	6.46	2.11	
06	189	2.47	6.72	4.25	
07	888	3.36	6.05	2.69	
08	231	3.11	6.64	3.53	
09	255	3.21	6.59	3.38	(Vilar et al., 2006)
10	400	3.15	6.40	3.25	(Ferino et al., 2013; Prendergast, 2001)
11	798	3.06	6.10	3.04	(Kirkiacharian et al., 1999)
12	955	2.49	6.02	3.53	(Prendergast, 2001)
13	1946	2.41	5.71	3.30	(Bhandri et al., 1949)
14	8476	2.34	5.07	2.73	
15	292	3.73	6.53	2.80	
16	1433	3.71	5.84	2.13	
17	384	2.8	6.42	3.62	(Kirkiacharian et al., 2003)
18	617	3.49	6.21	2.72	
19	866	2.79	6.06	3.27	
20	391	2.71	6.41	3.70	
21	433	3.32	6.36	3.04	
22	831	2.73	6.08	3.35	
23	902	3.58	6.04	2.46	(Chauhan et al., 2016)
24	1058	2.61	5.98	3.37	
M227[3]	440	3.03	6.36	3.33	(Santana et al., 2010)
M228[3]	210	2.94	6.68	3.74	(Santana et al., 2010)
M229[3]	130	2.94	6.89	3.95	(Santana et al., 2010)
M230[3]	N/A	3.72	N/A	N/A	(Santana et al., 2010)
M231[3]	1	3.65	9.00	5.35	(Santana et al., 2010)
M232[3]	4	3.14	8.40	5.26	(Santana et al., 2010)
M233[3]	480	3.11	6.32	3.21	(Santana et al., 2010)
M234[3]	290	3.02	6.54	3.52	(Santana et al., 2010)
M235[3]	640	3.02	6.19	3.18	(Santana et al., 2010)
M236[3]	N/A	3.80	N/A	N/A	(Santana et al., 2010)
M237[3]	2	3.73	8.70	4.97	(Santana et al., 2010)
M238[3]	6	3.73	8.22	4.49	(Santana et al., 2010)
M243[3]	43	4.04	7.37	3.33	(Santana et al., 2010)
M271[3]	78	3.50	7.11	3.60	(Viña et al., 2012a)
M272[3]	3	3.52	8.52	5.00	(Viña et al., 2012a)
M273[3]	5	3.78	8.30	4.53	(Viña et al., 2012a)
M274[3]	20	4.06	7.70	3.64	(Viña et al., 2012a)
M275[3]	150	3.15	6.82	3.67	(Viña et al., 2012a)
M276[3]	3	3.73	8.52	4.79	(Viña et al., 2012a)

M277[3]	1060	3.62	5.97	2.35	(Viña et al., 2012a)
M278[3]	210	3.67	6.68	3.01	(Viña et al., 2012a)
M279[3]	150	3.95	6.82	2.88	(Viña et al., 2012a)
CHEMBL1221929	4300	3.91	5.37	1.46	(Matos et al., 2010)
CHEMBL1221930	11050	4.01	4.96	0.95	(Matos et al., 2010)
CHEMBL1221980	3230	4.07	5.49	1.42	(Matos et al., 2010)
CHEMBL1221981	7120	4.13	5.15	1.02	(Matos et al., 2010)
CHEMBL1221982	4890	4.22	5.31	1.09	(Matos et al., 2010)
CHEMBL1777814	83.48	3.73	7.08	3.35	(Matos et al., 2011b)
CHEMBL1777845	1.35	3.78	8.87	5.09	(Matos et al., 2011b)
CHEMBL1777847	30910	3.06	4.51	1.45	(Matos et al., 2011b)
CHEMBL1777849	16870	2.31	4.77	2.46	(Matos et al., 2011b)
CHEMBL1783714	283.75	3.46	6.55	3.08	(Matos et al., 2009, 2011a; Viña et al., 2012b)
CHEMBL1783715	8.98	3.53	8.05	4.52	(Matos et al., 2009, 2011a)
CHEMBL1783716	160.64	3.62	6.79	3.18	(Matos et al., 2009, 2011a)
CHEMBL1835223	17.05	3.47	7.77	4.30	(Matos et al., 2011a)
CHEMBL1835224	1.52	3.47	8.82	5.35	(Matos et al., 2011a)
CHEMBL1835225	67.1	2.76	7.17	4.42	(Matos et al., 2011a)
CHEMBL1835226	5520	2.94	5.26	2.32	(Matos et al., 2011a)
CHEMBL1835227	14470	4.70	4.84	0.14	(Matos et al., 2011a)
CHEMBL1835228	0.31	3.74	9.51	5.77	(Matos et al., 2011a; Viña et al., 2012b)
CHEMBL1835229	15.01	3.74	7.82	4.09	(Matos et al., 2011a)
CHEMBL1835230	2.73	3.53	8.56	5.04	(Matos et al., 2011a)
CHEMBL1835231	0.74	4.08	9.13	5.05	(Matos et al., 2011a)
CHEMBL1835232	3.25	4.08	8.49	4.41	(Matos et al., 2011a)
CHEMBL1835233	54030	4.07	4.27	0.20	(Matos et al., 2011a)
CHEMBL1835234	N/A	4.11	N/A	N/A	(Matos et al., 2011a)
CHEMBL1835326	650.03	2.75	6.19	3.44	(Matos et al., 2011a)
CHEMBL1835327	120.02	2.84	6.92	4.08	(Matos et al., 2011a)
CHEMBL1835328	180.04	2.95	6.74	3.80	(Matos et al., 2011a)
CHEMBL1917495	69590	2.67	4.16	1.49	(Serra et al., 2012)
CHEMBL1934673	32040	2.97	4.49	1.53	(Serra et al., 2012)
CHEMBL1934674	N/A	3.15	N/A	N/A	(Serra et al., 2012)
CHEMBL1934675	9260	3.18	5.03	1.85	(Serra et al., 2012)
CHEMBL1934676	42680	3.49	4.37	0.88	(Serra et al., 2012)
CHEMBL1934677	2790	3.67	5.55	1.88	(Serra et al., 2012)
CHEMBL510349	11810	3.46	4.93	1.47	(Viña et al., 2012b)
CHEMBL570703	N/A	3.54	N/A	N/A	(Matos et al., 2011a)
CHEMBL570731	0.8	3.46	9.10	5.64	(Matos et al., 2011a)
CHEMBL572233	13.05	3.47	7.88	4.41	(Matos et al., 2009, 2011a)
CHEMBL577099	155.59	2.75	6.81	4.06	(Matos et al., 2011a)
CHEMBL64744	N/A	2.57	N/A	N/A	(Serra et al., 2012)

N/A = not available. [1] QPlogPo/w calculated using Schrödinger Release 2017-1: QikProp (Schrödinger, LLC, New York, NY, 2017). [2] LiPE = pIC_{50} - QPlogPo/w. [3] Numbered as in Matos *et al.* (2013) (Matos et al., 2013b) and designated with letter M. Naming of the rest of the compounds follows the ChEMBL (https://www.ebi.ac.uk/chembl/).

2 References

Bhandri, P. R., Bose, J. L., and Siddiqui, S. (1949). 3-Phenylcoumarin series. I. Synthesis of some new members of the series. *J. Sci. Ind. Res.* 8B, 189–192.

Chauhan, P., Ravi, M., Singh, S., Prajapati, P., and Yadav, P. P. (2016). Regioselective [small alpha]-arylation of coumarins and 2-pyridones with phenylhydrazines under transition-metal-free conditions. *RSC Adv.* 6, 109–118. doi:10.1039/C5RA20954D.

Dobelmann-Mara, L., Riedmueller, S., and Schraub, M. (2017). Compounds for optically active devices. *PCT Int. Appl.*, A1 20170302.

Ferino, G., Cadoni, E., Matos, M. J., Quezada, E., Uriarte, E., Santana, L., et al. (2013). MAO Inhibitory Activity of 2-Arylbenzofurans versus 3-Arylcoumarins: Synthesis, in vitro Study, and Docking Calculations. *ChemMedChem* 8, 956–966. doi:10.1002/cmdc.201300048.

Kabeya, L. M., da Silva, C. H. T. P., Kanashiro, A., Campos, J. M., Azzolini, A. E. C. S., Polizello, A. C. M., et al. (2008). Inhibition of immune complex-mediated neutrophil oxidative metabolism: A pharmacophore model for 3-phenylcoumarin derivatives using GRIND-based 3D-QSAR and 2D-QSAR procedures. *Eur. J. Med. Chem.* 43, 996–1007. doi:10.1016/j.ejmech.2007.07.003.

Kirkiacharian, S., Chidiack, H., Philibert, D., Van De Velde, P., and Bouchoux, F. (1999). Binding affinity to steroid hormone receptors and antiproliferative action on MCF-7 cells of coumarinic derivatives and isoflavonoids. *Ann. Pharm. Fr.* 57, 332–339.

Kirkiacharian, S., Lormier, A. T., Resche-Rigon, M., Bouchoux, F., and Cerede, E. (2003). Synthesis and binding affinity of 3-aryl-7-hydroxycoumarins to human α and β estrogen receptors. *Ann. Pharm. Fr.* 61, 51–56. doi:APF-01-2003-61-1-0003-4509-101019-ART2.

Leitão, A., Andricopulo, A. D., Oliva, G., Pupo, M. T., de Marchi, A. a, Vieira, P. C., et al. (2004). Structure-activity relationships of novel inhibitors of glyceraldehyde-3-phosphate dehydrogenase. *Bioorg. Med. Chem. Lett.* 14, 2199–2204. doi:10.1016/j.bmcl.2004.02.025.

Matos, M. J., Pérez-Cruz, F., Vazquez-Rodriguez, S., Uriarte, E., Santana, L., Borges, F., et al. (2013a). Remarkable antioxidant properties of a series of hydroxy-3-arylcoumarins. *Bioorganic Med. Chem.* 21, 3900–3906. doi:10.1016/j.bmc.2013.04.015.

Matos, M. J., Terán, C., Pérez-Castillo, Y., Uriarte, E., Santana, L., and Viña, D. (2011a). Synthesis and study of a series of 3-arylcoumarins as potent and selective monoamine oxidase B inhibitors. *J. Med. Chem.* 54, 7127–7137. doi:10.1021/jm200716y.

Matos, M. J., Vazquez-Rodriguez, S., Uriarte, E., Santana, L., and Viña, D. (2011b). MAO inhibitory activity modulation: 3-Phenylcoumarins versus 3-benzoylcoumarins. *Bioorganic Med. Chem. Lett.* 21, 4224–4227. doi:10.1016/j.bmcl.2011.05.074.

Matos, M. J., Viña, D., Janeiro, P., Borges, F., Santana, L., and Uriarte, E. (2010). New halogenated 3-phenylcoumarins as potent and selective MAO-B inhibitors. *Bioorganic Med. Chem. Lett.* 20, 5157–5160. doi:10.1016/j.bmcl.2010.07.013.

Matos, M. J., Viña, D., Quezada, E., Picciau, C., Delogu, G., Orallo, F., et al. (2009). A new series of 3-phenylcoumarins as potent and selective MAO-B inhibitors. *Bioorg. Med. Chem. Lett.* 19, 3268–3270. doi:10.1016/j.bmcl.2009.04.085.

Matos, M. J., Viña, D., Vazquez-Rodriguez, S., Uriarte, E., and Santana, L. (2013b). Focusing on New Monoamine Oxidase Inhibitors: Differently Substituted Coumarins As An Interesting Scaffold. *Curr. Top. Med. Chem.* 12, 2210–2239. doi:10.2174/1568026611212200008.

Prendergast, P. T. (2001). Use of flavones, coumarins and related compounds to treat infections. *PCT Int. Appl.*, A2 20010118.

Pu, W., Lin, Y., Zhang, J., Wang, F., Wang, C., and Zhang, G. (2014). 3-Arylcoumarins: Synthesis and potent anti-inflammatory activity. *Bioorg. Med. Chem. Lett.* 24, 5432–5434. doi:10.1016/j.bmcl.2014.10.033.

Santana, L., Orallo, F., Viña, D., Matos Joao Correia, P. C., Quezada, E., Yañes, J., et al. (2010). Use of derivates of 6-substituted 3-phenylcoumarins and preparation of new derivates. Available at: http://www.google.com.pg/patents/WO2010086484A1?cl=en.

Serra, S., Ferino, G., Matos, M. J., Vázquez-Rodríguez, S., Delogu, G., Viña, D., et al. (2012). Hydroxycoumarins as selective MAO-B inhibitors. *Bioorganic Med. Chem. Lett.* 22, 258–261. doi:10.1016/j.bmcl.2011.11.020.

Vilar, S., Quezada, E., Santana, L., Uriarte, E., Yánez, M., Fraiz, N., et al. (2006). Design, synthesis, and vasorelaxant and platelet antiaggregatory activities of coumarin–resveratrol hybrids. *Bioorg. Med. Chem. Lett.* 16, 257–261. doi:10.1016/j.bmcl.2005.10.013.

Viña, D., Matos, M. J., Ferino, G., Cadoni, E., Laguna, R., Borges, F., et al. (2012a). 8-Substituted 3-Arylcoumarins as Potent and Selective MAO-B Inhibitors: Synthesis, Pharmacological Evaluation, and Docking Studies. *ChemMedChem* 7, 464–470. doi:10.1002/cmdc.201100538.

Viña, D., Matos, M. J., Yáñez, M., Santana, L., and Uriarte, E. (2012b). 3-Substituted coumarins as dual inhibitors of AChE and MAO for the treatment of Alzheimer's disease. *Med. Chem. Commun.* 3, 213–218. doi:10.1039/C1MD00221J.

Wang, C., Wu, C., Zhu, J., Miller, R. H., and Wang, Y. (2011). Design, Synthesis and Evaluation of Coumarin-based Molecular Probes for Imaging of Myelination. *J. Med. Chem.* 14, 2331–2340. doi:10.1016/j.jacc.2007.01.076.White.

Yang, J., Liu, G. Y., Dai, F., Cao, X. Y., Kang, Y. F., Hu, L. M., et al. (2011). Synthesis and biological evaluation of hydroxylated 3-phenylcoumarins as antioxidants and antiproliferative agents. *Bioorganic Med. Chem. Lett.* 21, 6420–6425. doi:10.1016/j.bmcl.2011.08.090.

II

**IDENTIFICATION OF ESTROGEN RECEPTOR A LIGANDS
WITH VIRTUAL SCREENING TECHNIQUES**

by

Sanna P. Niinivehmas, Elangovan Manivannan, Sanna Rauhamäki, Juhani
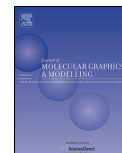Huuskonen & Olli T. Pentikäinen 2016

Journal of Molecular Graphics and Modeling 64: 30–39.

Identification of estrogen receptor α ligands with virtual screening techniques

Sanna P. Niinivehmas [a,1], Elangovan Manivannan [a,b,1], Sanna Rauhamäki [a],
Juhani Huuskonen [c], Olli T. Pentikäinen [a,*]

[a] Department of Biological and Environmental Science & Nanoscience Center, University of Jyvaskyla, P.O. Box 35, FI-40014, Finland
[b] School of Pharmacy, Devi Ahilya University, Indore 452001, Madhya Pradesh, India
[c] Department of Chemistry & Nanoscience Center, University of Jyvaskyla, P.O. Box 35, FI-40014, Finland

ARTICLE INFO

ABSTRACT

Utilization of computer-aided molecular discovery methods in virtual screening (VS) is a cost-effective approach to identify novel bioactive small molecules. Unfortunately, no universal VS strategy can guarantee high hit rates for all biological targets, but each target requires distinct, fine-tuned solutions. Here, we have studied in retrospective manner the effectiveness and usefulness of common pharmacophore hypothesis, molecular docking and negative image-based screening as potential VS tools for a widely applied drug discovery target, estrogen receptor α (ERα). The comparison of the methods helps to demonstrate the differences in their ability to identify active molecules. For example, structure-based methods identified an already known active ligand from the widely-used bechmarking decoy molecule set. Although prospective VS against one commercially available database with around 100,000 drug-like molecules did not retrieve many testworthy hits, one novel hit molecule with pIC$_{50}$ value of 6.6, was identified. Furthermore, our small in-house compound collection of easy-to-synthesize molecules was virtually screened against ERα, yielding to five hit candidates, which were found to be active *in vitro* having pIC$_{50}$ values from 5.5 to 6.5.

1. Introduction

In the modern drug discovery research, computational tools are increasingly used for cost-effective and rapid identification of new drug candidate molecules [1]. Especially, the development of virtual screening (VS) approaches, using computer-aided molecular discovery (CAMD) methods, gain attention because they offer an attractive approach to identify novel bioactive molecules from the chemical libraries [1]. VS protocols can be designed to meet any specific requirement depending on the availability of the known ligands and/or target structure information.

CAMD methods can be classified under two main categories: Ligand-based and protein structure-based methods [2]. Ligand-based methods are in practice the only solution when the structure of the target protein is absent, and cannot be modeled reliably. Ligand-based methods utilize the existing structure-activity data and calculated physicochemical properties of known bioactive molecules. Ligand-based methods include techniques such as three-dimensional quantitative structure-activity relationship (3D-QSAR) [3], two-dimensional fingerprint similarity search [4], and pharmacophore hypothesis [5]. In contrast, protein structure-based methods utilize experimentally solved structural information of a target molecule. From these methods molecular docking is the most commonly used [6]. However, molecular docking is computationally more expensive than ligand-based techniques, and the scoring functions associated with docking are not always effective in prioritizing the active molecules [7]. Although, all of the above mentioned methods have been proven to be useful in hit identification, each method has its pitfalls [8] Accordingly, ligand- and protein structure-based methods have also been integrated, e.g. docking-based ligand alignment has been applied with 3D-QSAR methods [9]. Furthermore, in negative image-based (NIB) screening the ligand binding pocket of the pro-

tein is described as a mimic of the ligand, which again, can be rapidly compared against library of small molecules with typical ligand-based methods [10]. Also the comparison of several methods can be a useful way to obtain more knowledge about the requirements for high affinity binding [11].

Estrogen receptor (ER) is a nuclear hormone receptor which functions as a mediator of estrogen's action in different parts of the body, and is an established target for drug development in infertility management, contraception, menopausal hormone replacement therapy, and endocrine based breast cancer therapy [12]. Among ERα binding ligands, selective ER modulators (SERMs) have received a remarkable impact in pharmaceutical development for osteoporosis and estrogen sensitive cancer. Because of the availability of the X-ray crystal structure-ligand complexes and a large number of chemically diverse active ligands, ERα has been widely used in the development of VS methods [13], and was thus chosen as the test system for comparative evaluation of both ligand- and structure-based VS methods.

In brief, the aim of this study was to compare in retrospective manner the effectiveness and usefulness of common pharmacophore hypothesis, molecular docking, and negative image-based screening. ERα ligands belonging to distinct chemotypes, and their bio-molecular target ERα were used as a test system. Furthermore, we tested the developed models in the prospective discovery of novel ERα ligands. Prospective VS against a commercially available database and an in-house database yielded six novel hit molecules with pIC50 value from 5.5 to 6.6.

2. Experimental

2.1. Biochemical data and preparation of datasets

In this study, we have used two data sets, referred as DS1 and DS2. DS1 contains carefully chosen 101 chemically diverse SERMs from ChEMBL database [14] with homogenous ERα inhibitory activity with pIC50-values (*i.e.*, -log(IC50)) ranging from 5.0 to 9.7 (Supplementary Tables S1 and S2). DS1 ligands belong into 18 different chemotypes (Supplementary Table S1) of which the most active molecules have a benzoxathin or a 2-aryl benzothiophene scaffold. DS2 constitutes of ERα antagonists ($n = 39$) and ERα decoy molecules ($n = 1395$) provided within the Directory of Useful Decoys (DUD) [15]. To complement the usability of DS1, the same DUD decoy molecules ($n = 1395$) were also added to it. The 3D structures of both DS1 and DS2 compounds were prepared using Ligprep 2.5 (Schrödinger Inc., New York, NY), where the stereochemistry of all the compounds was retained, while the protonation state for all the ionizable groups was set at pH 7.4. Conformational search was performed using ConfGen [16] with pre-minimization using MMFF94 force field [17] with distant dependent dielectric solvent model.

2.2. Common pharmacophore hypothesis and 3D-QSAR model generation

The common pharmacophore hypothesis (CPH) based 3D-QSAR models were built with PHASE v3.3 [18]. The CPH model is generated from the overlay of pharmacophoric features that are precalculated for each ligand and ligand conformer. Then CPH is combined with atom-based 3D-QSAR so that the 3D-QSAR model is constructed based on the selected pharmacophore model. CPH is inherently a yes/no classifier sorting compounds to be either binders fulfilling the set pharmacophoric criteria or non-binders violating the criteria. The 3D-QSAR model predicts the activities of all screened molecules fulfilling the pharmacophoric criteria.

For 3D-QSAR model development and validation purpose, DS1 active molecules were further divided into training set and test set with the ratio of 70:30. As PHASE uses only the most active compounds in the training set to create pharmacophore hypothesis, an activity threshold was set for the selection of active ligands (IC50 ≤ 1 nM, *i.e.*, pIC50 ≥ 9) and inactive ligands (pIC50 ≤ 6.5) in the training set. The pharmacophore sites were created for all ligand conformers in DS1. These pharmacophore sites are specific chemical features of ligands defined in PHASE as hydrogen bond acceptors (A), hydrogen bond donors (D), hydrophobic groups (H), negatively charged groups (N), positively charged groups (P) and aromatic rings (R).

2.3. Target structure preparation and molecular docking

The X-ray crystal structure coordinates for human ERα ligand-binding domain in complex with 4-hydroxytamoxifen (PDB 3ERT) was retrieved from the RCSB Protein Data Bank (PDB) and prepared with the Protein Preparation Wizard in MAESTRO v9.2 [19]. During the protein preparation, the bond orders were assigned, and hydrogen atoms and formal charges were added to heterogroups. The water molecules in the ligand-binding area were preserved for docking, and all other water molecules 5 Å beyond heterogroups were deleted. The hydrogen bonding network of binding site residues was optimized by selecting the histidine tautomers and by predicting the ionization states. The optimized protein structure was then subjected to all-atom constrained energy minimization using the IMPREF module of MAESTRO v9.2 with OPLS-2005 force field [19]. The prepared ERα structure was used for the molecular docking simulations. The prepared DS1 and DS2 molecule sets were docked flexibly utilizing GLIDE v5.7 standard precision (SP) and GLIDE extra precision (XP) scoring functions [20]. The default settings in the GLIDE were used for both the grid generation and the flexible docking. In addition, molecular docking with PLANTS software was performed using CHEMPLP scoring function with binding site radius 15 Å from the center of the binding cavity, and the results were cluster with RMSD 2.0 Å [21].

2.4. Molecular dynamics simulations and negative-image creation

The protein flexibility was introduced in the NIB models by constructing the negative images for an ensemble of protein conformations derived from molecular dynamics (MD) simulation snapshots. For MD, the same protocol was used as previously [21]. In short, the preparations for MD simulations were performed as follows (1) The water molecules and the co-crystallized ligand in the crystal structure were removed from ligand binding cavity; (2) Ligand binding cavity was filled with water molecules using VOIDOO/FLOOD [23], (3) TLEAP in ANTECHAMBER 1.27 [24] was used for creating force field parameters for the protein (ff03 [25]), adding hydrogens and solvating the structure into a 13 Å rectangular box of transferable intermolecular potential three-point water molecules (TIP3P [26]). Then the MD simulations were run with NAMD [27], by using the same parameters that we have successfully employed earlier [22]. The snapshots at 400 ps intervals were extracted from the MD trajectories with PTRAJ 10 in the ANTECHAMBER 1.27 [24] and used for NIB model creation. In principle the NIB models were built as described previously [10].

2.5. Electrostatic information and shape comparison

The electrostatic information of target protein was incorporated into the NIB models through assigning an atom-centered MMFF94 charges [17] for the protein atoms. Then the charges of the protein atoms within 1.8 Å radius of each NIB model data point were averaged and the opposite charge values assigned to the corresponding NIB data point. This particular radius is taken into account

due to consideration of atoms within hydrogen bonding distances. The shape comparison and electrostatic matching of the ERα NIB models against the DS1 and DS2 was performed with SHAEP [28], as described in our earlier studies [10]. For electrostatic matching, MMFF94 charges [17] were pre-computed for all molecules in DS1 and DS2.

2.6. Fluorescence polarization and pIC_{50} values

The pIC_{50} values of identified molecules were measured using green PolarScreen™ ER Alpha Competitor Assay (Life Technologies, CA, The United States of America) kit. The protocol provided by the manufacturer was followed. Briefly, dilution series where the final concentration of the test compounds ranged between 0.0007 and 10 000 nM was prepared. The test compounds were combined with 25 nM ERα and 4.5 nM fluormone in the assay buffer and placed on black low volume 384-well assay plate with NBS surface (Corning, NY, The United States of America). After mixing the assay plate, it was incubated for 2 h in room temperature. The fluorescence polarization was then measured using excitation wave length 485 and emission wave length 535 with bandwidths of 25/20 nm. The measurements were performed on 2104 EnVision® Multilabel Plate Reader which had EnVision Workstation version 1.7 (PerkinElmer, MA, The United States of America).

2.7. Synthesis

The coumarin derivatives were synthesized using Perkin–Oglialor condensation reaction. The method was developed from the earlier published procedures and transferred to microwave reactor [29]. The general procedure for the synthesis is shown in Supplementary Fig. S1 with detailed synthesis and characterization.

3. Results and discussion

3.1. Validation and performance evaluation of common pharmacophore hypothesis model

The pharmacophore hypotheses were compiled from six variant lists (A, D, H, N, P, R). The used dataset was divided into training (70%) and test set (30%), where the test set created by PHASE was further verified to contain ligands from all chemotypes and from the most and least active groups. Five combinations of five chemical features were common in all most active ligands ($pIC_{50} \geq 9$). PHASE scores the hypotheses for survival of active ($pIC_{50} \geq 9$) and inactive ligands ($pIC_{50} \leq 6.5$). The high ranking hypotheses were used for CPH based 3D-QSAR model building and validation. APRRR-223 (Fig. 1) was identified as the best hypothesis on statistical grounds ($R^2 = 0.923$, standard deviation = 0.317, Fischer significance $F = 154.7$, and chance correlation $P = 4.47 \times 10^{-44}$). Survival numbers of active and inactive ligands for the best CPH model were 43.1 and 42.5, respectively. Accordingly, it can be said that the CPH-based 3D-QSAR model should be highly reliable and thus give accurate predictions.

The selected model, APRRR-223, was mapped both onto the most active SERM (ChEMBL198803; 0.2 nM), and the least active SERM (ChEMBL380838; 10,000 nM) in the DS1. While the CHEMBL198803 aligns perfectly with the model, the least active, CHEMBL380838, could not match all the pharmacophore features in the hypothesis, *e.g.*, CHEMBL380838 does not have an essential chemical feature to match with the positive pharmacophore site. Thus, the derived pharmacophore hypothesis is capable of differentiating the most active from the least active SERMs. In total, from the 101 DS1 ligands, the CPH-model identifies 61. In the model creation activity limit was set to very high, $pIC_{50} \geq 9$, while inactivity

limit was set to $pIC_{50} \leq 6.5$. In the DS1 there are 49 highly active compounds ($pIC_{50} \geq 9$) of which the model finds 47. However, two active ligands (ChEMBL241301 and ChEMBL391910) in DS1 with $pIC_{50} > 9$ were not mapped into the CPH model and they were consistent outliers for this 3D-QSAR model. The outlying behavior of these ligands is apparent from their very uncommon binding pose compared to other active ligands in the DS1 data set. These two ligands bind in an unusual way to the binding cavity, so that the largest rigid part of the molecule does not bind to the bottom of the binding cavity but closer to the surface of the receptor, and thus changes the folding of the receptor. No inactive compounds were identified by the model.

The 3D-QSAR model constructed along with the CPH-model has the following statistical parameters $Q^2 = 0.822$, RMSE = 0.431 and $R^2 = 0.870$. The high Q^2 value illustrates the accuracy of the model in predicting the ERα activity in the test set. To further evaluate the performance of the 3D-QSAR models, DS2 was also used in the screening process. With compulsory matching of all five pharmacophore features in the hypothesis, out of 39 active ligands and 1395 decoys in the DS2, the hypothesis recognized 16 most active antagonists without selecting any decoys. It must be noted that among the 16 identified DUD antagonists, four already existed in the DS1 and were used in the pharmacophore development and validation. In contrast to DS1, the model was also able to identify chemically distinct antagonists belonging to the chemotypes of naphthalene, dihydronapthalene, triaryl ethylene and tetracyclic scaffolds. With the objective of recognizing more actives from DS2, the secondary screening was performed by relaxing pharmacophore features from the hypothesis one at a time (*i.e.*, all combinations of four features). As a result all 39 active ligands and more than 600 decoys (43%) were recognized. This demonstrates the importance of all five chemical features in APRRR-223 hypothesis (Fig. 1). Accordingly, the hypothesis with all five features is likely to identify most of the highly active ERα ligands, but cannot identify any low-affinity hits in VS campaigns, and if model is relaxed to four features, the distinguisment of active ligands from inactive molecules becomes highly unreliable.

Division of DS1 into training and test molecules as well as detailed activity predictions for DS1 and DS2 ligands are available in Supplementary Table S3.

3.2. Molecular docking

The regression models were derived by correlating the ERα activity and the corresponding docking scores (Fig. 2). The GLIDE SP score based regression model yielded a better correlation coefficient ($R^2 = 0.638$; Fig. 2a) than computationally more demanding GLIDE XP ($R^2 = 0.230$; data not shown). Thus, the GLIDE SP based scoring was used in the model development and validation. In case of PLANTS docking, the regression model yielded the correlation coefficient $R^2 = 0.639$ (Fig. 2b) with CHEMPLP score. Accordingly, both GLIDE SP and PLANTS docking produced a very similar and quite reasonable correlation. Furthermore, with the objective of improving the correlation between the ERα activities and docking results, a hybrid regression model was developed of both GLIDE SP score and PLANTS CHEMPLP score. The combination of normalized docking scores of both GLIDE and PLANTS docking resulted better correlation coefficient $R^2 = 0.787$ (Fig. 2c) than the individual regression models. Docking data for GLIDE SP, PLANTS and hybrid docking model is available in Supplementary Table S3.

The crystal structure of ERα with 4-hydroxytamoxifen reveals the important H-bonding network between Glu353, Arg394, and a water molecule. The visual inspection of the top ranked poses of GLIDE SP and PLANTS CHEMPLP showed a good agreement in binding similarity between the crystal structure ligand and the top-ranked ERα ligands and antagonists. Albeit, GLIDE SP docking
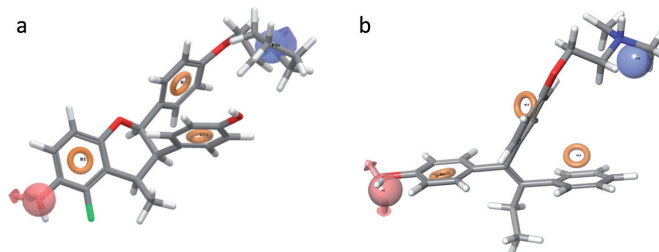
Fig. 1. Common pharmacophore hypothesis (CPH) model showing the essential pharmacophore features for used SERMs. The best CPH model, APRRR-223, represents the pharmacophore features of an ER ligand by way of (a) the reference ligand ChEMBL181369 and (b) 4-hydroxytamoxifen: hydrogen bond acceptor (red spheres with arrows pointing the direction of available lone pairs) and positively charged group at the end of the antagonist-arm (blue sphere with arrow pointing the hydrogen bond direction), and aromatic ring structures in the core of the molecule (orange doughnuts).
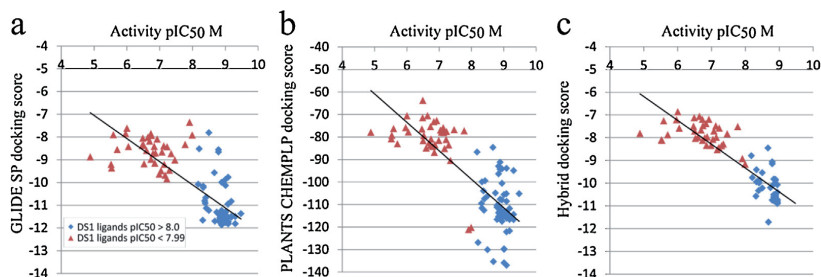
Fig. 2. The predictivity of the docking methods. Correlation of experimental ERα activity of DS1 and DS2 ligands (as pIC_{50}) between (a) GLIDE SP, (b) PLANTS CHEMPLP, and (c) normalized GLIDE SP/PLANTS hybrid docking score are shown.

poses were observed to have closer resemblance to crystal structure complex conformation than PLANTS docking poses. The docking poses of ligands like ChEMBL81 (Raloxifene), ChEMBL328190 (Lasofoxifene) in DS1 and the most active ERα antagonists in DS2, yielded binding modes that resemble closely the binding mode of 4-hydroxytamoxifen, including similar H-bonding network and hydrophobic contacts. The difference between the docking poses of the most active ligands and the weakly active ligands is that the substitution changing the agonist ligand into antagonist (*e.g.*, diethyl amine, pyrrolidine, and piperdine) is able to form an ionic interaction between the amino group of SERMs and carboxylate of Asp351. Note that the ER agonist ChEMBL135 (17β-Estradiol), partial agonists like ChEMBL44 (Genistein), ChEMBL8145 (Daidazin), and ChEMBL30707 (Coumestrol) and other structurally similar ligands in DS1 do not produce ionic interaction with Asp351 due to lack of above mentioned substitution.

A summary of receiver operating characteristics (ROC) curves were given as area under curves (AUC) values. The ROC curves plotted for the docking models and for the SHAEP comparisons with NIB models (see below) are shown in Supplementary Fig. S2. The ROC-AUC show better performance for GLIDE SP docking than PLANTS CHEMPLP docking in both DS1 and DS2 (Supplementary Fig. S2). The ROC plot for GLIDE SP with both DS1 and DS2 yielded high AUC = 0.91 ± 0.01 (Table 1). Correspondingly, in case of PLANTS docking model, the ROC plot for DS1 showed AUC = 0.65 ± 0.01 and for DS2 yielded AUC = 0.73 ± 0.01 (Table 1). Overall, the GLIDE SP docking was able to find actives effectively in both DS1 containing larger number of SERMs and in DS2 having lesser volume of antagonists. Even though the AUC for DS2 in PLANTS is higher than for DS1, the early enrichment for DS1 is considerably higher (Table 1, Supplementary Fig. S2). Even if the efficiency of finding the active ligands over decoys with PLANTS CHEMPLP docking model

was found slightly lower than with GLIDE SP, PLANTS is still suitable for improving the separation of lower activity molecules from the datasets when used in combination with GLIDE. Therefore, the hybrid docking model was also used for plotting ROC and calculating AUC-values. The ROC plot for DS1 yielded AUC = 0.79 ± 0.01 whereas DS2 gained AUC = 0.86 ± 0.01. The hybrid docking model shows a balanced level of screening performance in separating the active molecules over decoys in both DS1 and DS2.

### 3.3. Selection and validation negative image-based models

The X-ray structure analysis of ER-ligand complexes suggests that there are several flexible amino acid residues in the ligand-binding area. In consequence, there is an ensemble of conformations that the ligand-binding area can have, depending on the chemical nature of the bound agonist or antagonist. Therefore, the receptor flexibility was introduced in NIB models by using a collection of conformations of the protein derived from MD simulation. Similarly the electrostatic potentials or charge distribution are important in NIB models and depend heavily on the number of water molecules and the chemical nature of amino acids in the ligand binding area.

Initially, a total of 21 NIB models (numbered from 0 to 20; where 0 is the starting structure) were created for ligand binding pocket of ER at regular intervals from MD simulation. Each of these individual NIB models were screened against DS1 and DS2 using SHAEP. AUC-values for the NIB model search (21 separate models) with DS1 showed AUC ranging from 0.01 to 0.97, whereas with DS2 AUC ranged from 0.38 to 0.85 (data not shown). The highest AUC ≈ 0.97 ± 0.01 with DS1 was obtained for the NIB model 1, *i.e.*, slightly relaxed crystal structure. The above AUC was obtained using the NIB model search with the charge distance of

Table 1
The area under curve (AUC) values, absolute enrichments and number of hits for optimized NIB model 1 with different electrostatics contribution and GLIDE SP, PLANTS CHEMPLP and hybrid docking models with both original crystal structure and relaxed protein conformer.

Method	DS1					DS2				
	AUC	Absolute enrichment (amount of hits)				AUC	Absolute enrichment (amount of hits)			
		0.5%	1%	5%	10%		0.5%	1%	5%	10%
Protein crystal structure (PDB: 3ERT)										
GLIDE SP	0.91 ± 0.01	12.7 (6)	12.8 (13)	11.5 (58)	7.1 (72)	0.91 ± 0.01	15.8 (3)	13.1 (5)	12.3 (24)	7.5 (29)
PLANTS	0.65 ± 0.01	10.6 (5)	6.9 (7)	8.3 (42)	5.0 (51)	0.73 ± 0.01	10.5 (2)	13.1 (5)	6.6 (13)	3.9 (15)
Hybrid	0.79 ± 0.01	8.5 (4)	6.9 (7)	11.5 (58)	6.2 (63)	0.86 ± 0.01	5.3 (1)	10.5 (4)	9.2 (18)	6.2 (24)
Relaxed protein conformer										
GLIDE SP	0.75 ± 0.01	10.6 (5)	5.9 (6)	5.3 (27)	3.9 (40)	0.89 ± 0.01	5.3 (1)	10.5 (4)	10.7 (21)	7.2 (28)
PLANTS	0.62 ± 0.01	14.8 (7)	7.9 (8)	6.5 (33)	4.5 (46)	0.71 ± 0.01	15.8 (3)	10.5 (4)	5.1 (10)	3.6 (14)
Hybrid	0.73 ± 0.01	10.6 (5)	6.9 (7)	6.5 (33)	4.1 (42)	0.85 ± 0.01	5.3 (1)	10.5 (4)	9.7 (19)	5.7 (22)
SHAEP NIB model 1 (relaxed protein conformer)										
Shape only	0.61 ± 0.03	14.8 (7)	11.9 (12)	4.9 (25)	3.0 (30)	0.72 ± 0.05	15.6 (3)	7.9 (3)	4.6 (9)	3.0 (12)
0%	0.58 ± 0.03	14.8 (7)	11.9 (12)	4.9 (24)	2.9 (29)	0.72 ± 0.05	15.6 (3)	10.5 (4)	4.6 (9)	3.0 (12)
10%	0.65 ± 0.03	12.7 (6)	12.8 (13)	4.9 (25)	3.2 (32)	0.72 ± 0.05	15.6 (3)	10.5 (4)	4.6 (9)	3.0 (12)
20%	0.72 ± 0.02	14.8 (7)	11.9 (12)	4.9 (26)	3.3 (33)	0.73 ± 0.05	15.6 (3)	10.5 (4)	4.6 (9)	3.0 (12)
30%	0.82 ± 0.02	14.8 (7)	11.9 (12)	5.9 (30)	5.1 (52)	0.73 ± 0.05	21.0 (4)	10.5 (4)	4.6 (9)	3.3 (13)
40%	0.91 ± 0.02	14.8 (7)	13.8 (14)	10.9 (55)	7.2 (73)	0.73 ± 0.05	21.0 (4)	10.5 (4)	4.6 (9)	3.6 (14)
50%	0.97 ± 0.02	14.8 (7)	14.8 (15)	14.6 (74)	8.8 (89)	0.74 ± 0.05	21.0 (4)	10.5 (4)	4.6 (9)	3.6 (14)
60%	0.96 ± 0.01	14.8 (7)	14.8 (15)	14.6 (74)	8.7 (88)	0.75 ± 0.05	21.0 (4)	10.5 (4)	4.6 (9)	3.6 (14)
70%	0.97 ± 0.01	14.8 (7)	14.8 (15)	14.8 (75)	8.9 (90)	0.74 ± 0.05	21.0 (4)	10.5 (4)	5.1 (10)	4.4 (17)
80%	0.96 ± 0.01	14.8 (7)	14.8 (15)	14.8 (75)	8.8 (89)	0.59 ± 0.05	10.5 (2)	7.9 (3)	4.6 (9)	3.9 (15)
90%	0.93 ± 0.01	14.8 (7)	14.8 (15)	13.6 (69)	7.9 (80)	0.40 ± 0.04	0.0 (0)	0.0 (0)	0.0 (0)	0.0 (0)
100%	0.80 ± 0.03	14.8 (7)	14.8 (15)	12.4 (63)	6.5 (66)	0.72 ± 0.05	0.0 (0)	0.0 (0)	0.0 (0)	0.0 (0)
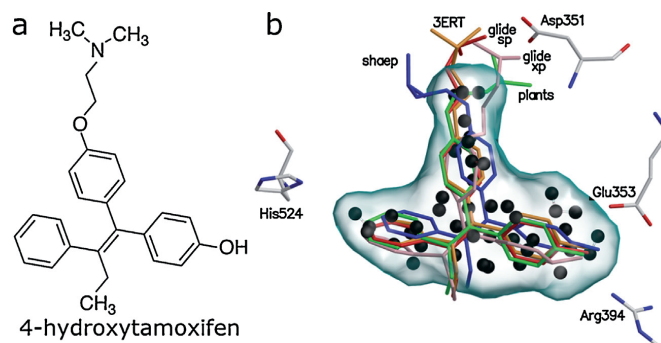
Fig. 3. 4-hydroxytamoxifen, the best NIB model along with the key amino acids of the ERα ligand binding pocket with docking results. Figure shows (a) the 2D structure of 4-hydroxytamoxifen, and (b) its stick model conformations from GLIDE SP (red), GLIDE XP (pink), PLANTS CHEMPLP (green), and NIB model 1 (blue) superimposed with 3ERT crystal structure ligand (orange). For comparison, panel b also shows NIB model 1 as black/white dots, (black dots represent oxygen atoms and white hydrogen atoms of the optimized water molecules, see section *Molecular Dynamics Simulations and Negative-Image Creation*) and pale blue solvent accessible surface. All method are able to produce a representative binding conformation for 4-hydroxytamoxifen. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

1.8 Å and the shape and electrostatic weight of 50%. These parameters for charge distance and electrostatic weighting were chosen based on our previous studies [10]. Correspondingly, NIB model 1 gained AUC ≈ 0.74 ± 0.05 with DS2 in identical search. Furthermore, the binding poses of DS1 ligands on various NIB models were compared with the crystal structure ligand conformation and their corresponding docking pose. These results suggest that NIB model 1 is the best model in every aspect, and thus, although our initial idea was to use all or at least several NIB models together, as previously [10], we decided to use NIB model 1 alone for further evaluation and optimization (Fig. 3).

The selected NIB model 1 was validated by varying the degree of electrostatic weighting from 0 to 100%, with intervals of 10% (Table 1). With this procedure we want to ensure that the electrostatic potentials assigned to the corresponding NIB data points are reasonable and in balance with shape effect. Accordingly, the

best electrostatic weighting should be somewhere around 50%. The shape alone NIB model 1 search resulted an AUC = 0.61 ± 0.03 for DS1. The gradual addition of electrostatic contribution to the NIB model 1 search increased the AUC values from 0.65 ± 0.03 (10% electrostatic contribution) until 0.97 ± 0.01 and 0.96 ± 0.01 (70% and 80% electrostatic contribution, respectively) (Table 1, Supplementary Fig. S3). However, further increase in electrostatic contribution did not improve the AUC values with either of the datasets. The searches with DS1 yielded the poorest AUC-values, when lowest electrostatic contribution was used (shape only to 20%) (Table 1, Supplementary Fig. S3). These results clearly indicate that the shape alone model is unable to effectively distinguish active molecules over decoys. When considering only the very early enrichment, the top 0.5% of DS1, the results were almost equally good regardless of the electrostatic contribution. However, with the early enrichment, the top 1%, the enrichment and number of hits

Table 2
The docking scores in GLIDE SP, PLANTS CHEMPLP and normalized GLIDE SP/PLANTS hybrid docking for the purchased (S1–S5) and synthesized (1–5) molecules. From purchased molecules, only S4 showed binding activity when measuring the pIC_{50} values. NB = non binding.
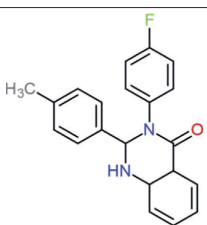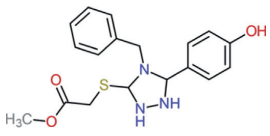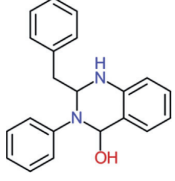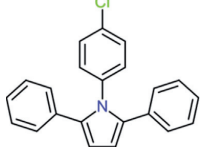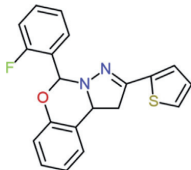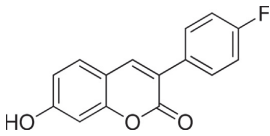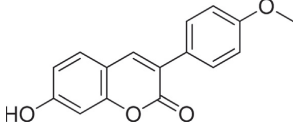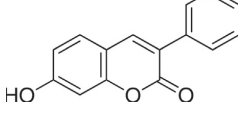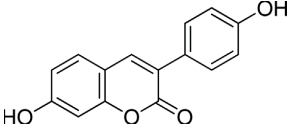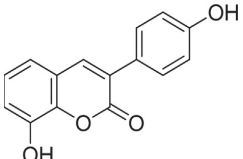
Name	Structure	GLIDE SP	PLANTS CHEMPLP	Hybrid	pIC_{50}
S1		−8.81	−84.67	−8.07	NB
S2		−8.94	−88.37	−8.30	NB
S3		−9.61	−88.85	−8.65	NB
S4		−9.39	−73.80	−7.89	6.6
S5		−9.19	−84.73	−8.27	NB
1		−8.94	−78.05	−7.85	6.1
2		−8.92	−74.07	−7.67	5.9
3		−8.84	−76.04	−7.71	6.5

Table 2 (*Continued*)

Name	Structure	GLIDE SP	PLANTS CHEMPLP	Hybrid	pIC_{50}
4		−8.33	−75.24	−7.56	5.5
5		−9.25	−78.32	−7.89	6.5

were highest when electrostatic contribution was ≥50% (Table 1). With DS2 the best AUCs were gained when electrostatics was weighted 50–70% but altogether the results were very even with all comparisons from shape only up to 70% electrostatic weighting. Notably, with electrostatic weighting ≥90%, no active molecules were found in the top 10%. Scores for DS1 and DS2 ligands with NIB model 1 with 50% electrostatic weighting are available in Supplementary Table S3.

Because NIB model 1 is not based on the conformer introduced by the crystal structure 3ERT but a slightly relaxed protein structure from the minimizations performed before actual MD simulations, the relaxed protein conformer was also used in docking simulations for reliable comparisons. However, neither GLIDE SP nor PLANTS CHEMPLP benefitted from this. All AUC-values decreased compared to docking with original crystal structure (Table 1). The most notable reduction in AUCs occurred in the GLIDE SP docking with DS1 from 0.91 to 0.73. In fact GLIDE SP was able to dock less active molecules from DS1 when using relaxed protein conformer (only 90% of the actives). Correspondingly also absolute enrichments and number of hits decreased (Table 1). Only PLANTS CHEMPLP docking with both DS1 and DS2, showed slender improvement in very early enrichments when relax protein conformer was used.

3.4. Identification of ER ligands with VS

Since NIB-model produces excellent AUC-values, the 20 best recognized decoys selected by these models were examined (electrostatic weighting 50%). Noteworthy, ZINC03826690 (Fig. 4), identified by 13 NIB-models, was also predicted by the docking to have very high activity (GLIDE SP: −11.43; PLANTS CHEMPLP: −118.8), suggesting a pIC_{50} > 8. Indeed, this compound is a known ERα ligand (pIC_{50}: 8.0), and thus inaccurately marked as a decoy in DUD. Furthermore, decoy molecules ZINC02630310 and ZINC03867590 were identified in the top-20 of decoys by several NIB-models, and the docking scores predicted that the pIC_{50}-value for these compounds would be approximately 6.5–7. Unfortunately, these compounds are not commercially available, and thus, cannot be tested. All other decoy molecules were poor based on NIB-models. However, if we consider decoy molecules that are ranked higher than ZINC03826690 by either of the docking methods, we find 12 molecules that all resemble ZINC03826690, indicating that also these molecules could have ER activity (Fig. 4). Unfortunately, these molecules are neither commercially available.

To test the performance of developed models a library of 100,000 drug-like compounds from one commercially available database (SPECS) was screened. Prior to screening, the SPECS database was filtered according to Lipinski's rule of five, which was complemented by allowing the maximum of six rotatable bonds. Unfortunately, we did not identify any potential high affinity molecules (pIC_{50} > 7.0) using combination of four methods pharmacophore, GLIDE SP, PLANTS CHEMPLP and NIB. This is mainly, because the pharmacophore model with all five features is already too specific for this type of VS campaign, and with four features the model does not discriminate inactives with high enough accuracy (see above). On the other hand, if we only take into consideration both docking methods and NIB-method, five hits remain (Table 2: S1–S5). Only one of these molecules (S4) showed activity *in vitro* within concentration range of 0.0007–10 000 nM with pIC_{50} value of 6.6. Accordingly, with blind screening where only the numerical comparisons are used, the results contain still lots of inactive ligands.

Next, small virtual library of easy-to-synthesize molecules with 17beta-estradiol sized coumarin-core was screened and analyzed. Originally the idea of the in-house database of easy-to-synthesize coumarin derivatives is that the compound collection consists of molecules that are synthesizable from cheap starting materials with one step synthesis, excluding possible protecting groups, performed with microwave-assisted organic synthesis in few minutes. For ERα the coumarin derivatives have right size (Table 2); however, hydroxyl as a functional group is the primary reason why these compounds are suitable for ERα. Again, only docking and NIB-method were able to identify hits. Out of these, few molecules were predicted to be active. Visual inspection revealed that some of the compounds could truly be active ER ligands, and were synthesized. Indeed, when the top five molecules were synthesized and tested *in vitro* they had pIC_{50} values from 5.5 to 6.5 (Table 2: 1–5).

Based on the visual inspection of docking results for identified hit molecules the purchased molecule resembles known partial agonists in size, and the synthesized ligands are similar in size with known agonists. In general, the blind follow of numerical estimations (purchased molecules) yielded molecules that were predicted too optimistically. When prior understanding of known ER-ligands was used in the guidance of molecule design (small virtual library of coumarins), also the activities were predicted more accurately (Fig. 5). In general, the GLIDE SP gave over predictions while PLANTS CHEMPLP under predicted the analyzed molecules. To be sure that our compounds are not PAINS (pan assay interference compounds) [30], the PAINS properties of the chosen compounds were studied using PAINS-Remover tool [30]. All chosen compounds survived through the filter.
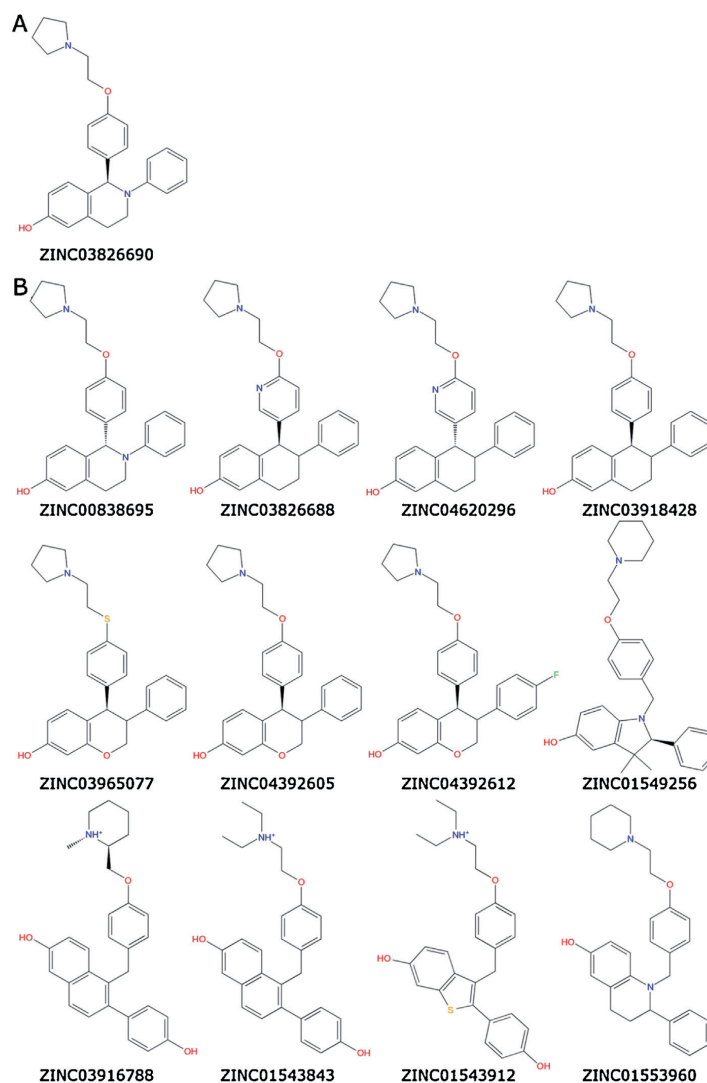
Fig. 4. ZINC03826690 (a) was identified as a high activity molecule by structure-based methods. In fact, this molecule is truly an active molecule (pIC_{50}: 8.0), and thus was marked inaccurately as decoy molecule in DUD. Twelve other decoy (b) molecules that resemble active ZINC03826690 considerably are ranked higher by either GLIDE SP or PLANTS CHEMPLP docking methods indicating that also these molecules might have ER activity.

4. Conclusions

Here, we have investigated the effectiveness and usefulness of ligand-based method CPH, structure-based methods docking and NIB model as possible VS tools for the identification of novel ERα ligands. This comparison of the above mentioned methods as VS tools describes nicely their different ability to identify active binders for ERα and also helps to better understand the relative advantages and limitations of these methods. The NIB was very efficient in prioritizing the active molecules over a large number of DUD decoys that demonstratesthe vitality of the method in structure-based VS for identification of novel ERα ligands. The optimized NIB model yields

better results for early enrichments than other methods, which is important in VS. Furthermore, the NIB model screening is typically faster to perform than other structure-based methods *e.g.*, docking. Thus, NIB is a feasible option for screening large databases with less computational time and efficiency. The CPH-based 3D-QSAR shows that pharmacophore hypothesis finds strictly the most active ligands, but the prediction capacity is highly chemotype dependent. For the prospective VS campaign, the model has too many features, and if relaxed, too many inactives will be identified. The docking (especially GLIDE SP) can be used for the identification of active molecules but above all for improving the estimation of the binding affinities of novel ERα ligands. Accordingly, in a workflow
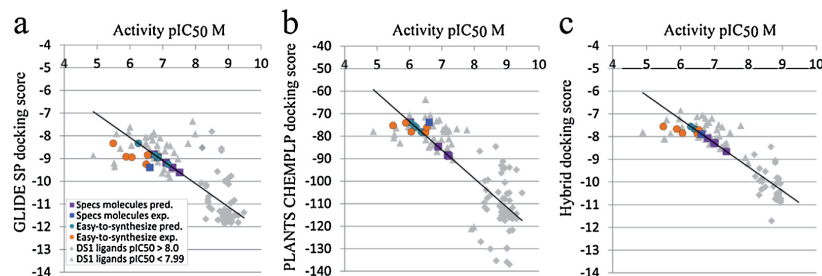
Fig. 5. The correlation of predicted and experimental ERα activity (as pIC_{50}) between docking scores. The predicted and experimental activities are shown for purchased specs molecules and small library of easy-to-synthesize coumarine derivatives as (a) GLIDE SP, (b) PLANTS CHEMPLP, and (c) normalized GLIDE SP/PLANTS hybrid docking scores. The correlation of experimental ERα activity of DS1 and DS2 ligands can be seen on the background.

where large compound collection is explored, the initial screening with fast NIB, and then rescoring of the top-ranked hits with accurate docking with careful visual inspection can be recommended. If the workflow could be applied, not only to identify novel drug candidates but also to predicti binding for the ligands that are developed for other target proteins in drug discovery, for example 17β-hydroxysteroid dehydrogenase6d, the utilized workflow could have applicability in environmental toxicity predictions.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmgm.2015.12.006.

References

[1] (a) G.C. Terstappen, A. Reggiani, In silico research in drug discovery, Trends Pharmacol. Sci. 22 (1) (2001) 23–26;
(b) W.L. Jorgensen, The many roles of computation in drug discovery, Science 303 (5665) (2004) 1813–1818.

[2] Matter, H., Sotriffer, C., Application and Success Stories in Virtual Screening. In Virtual Screening - Principles, Challenges, and Practical Guidelines, 1 ed., Sotriffer, C., Ed. Wiley-VCH Verlag & Co. KGaA: Weinheim, Germany, 2011; Vol. 49, pp 319-358.

[3] J. Verma, V.M. Khedkar, E.C. Coutinho, 3D-QSAR in drug design—a review, Curr. Top. Med. Chem. 10 (1) (2010) 95–115.

[4] M. Sastry, J.F. Lowrie, S.L. Dixon, W. Sherman, Large-scale systematic analysis of 2D fingerprint methods and parameters to improve virtual screening enrichments, J. Chem. Inf. Model. 50 (5) (2010) 771–784.

[5] (a) D. Horvath, Pharmacophore-based virtual screening, Methods Mol. Biol. 672 (2011) 261–298;
(b) A.R. Leach, V.J. Gillet, R.A. Lewis, R. Taylor, Three-dimensional pharmacophore methods in drug discovery, J. Med. Chem. 53 (2) (2010) 539–558.

[6] (a) X.Y. Meng, H.X. Zhang, M. Mezei, M. Cui, Molecular docking: a powerful approach for structure-based drug discovery, Curr. Comput. Aided Drug Des. 7 (2) (2011) 146–157;
(b) J.T. Koivunen, L. Nissinen, J. Kapyla, J. Jokinen, M. Pihlavisto, A. Marjamaki, J. Heino, J. Huuskonen, O.T. Pentikainen, Fluorescent small molecule probe to modulate and explore alpha2beta1 integrin function, J. Am. Chem. Soc. 133 (37) (2011) 14558–14561;
(c) J.T. Koivunen, L. Nissinen, A. Juhakoski, M. Pihlavisto, A. Marjamaki, J. Huuskonen, O.T. Pentikainen, Blockage of collagen binding to integrin alpha 2 beta 1: structure-activity relationship of protein–protein interaction inhibitors, MedChemComm 2 (8) (2011) 764–770;
(d) J. Messinger, L. Hirvela, B. Husen, L. Kangas, P. Koskimies, O. Pentikainen, P.

Saarenketo, H. Thole, New inhibitors of 17beta-hydroxysteroid dehydrogenase type 1, Mol. Cell. Endocrinol. 248 (1–2) (2006) 192–198.

[7] (a) J.B. Cross, D.C. Thompson, B.K. Rai, J.C. Baber, K.Y. Fan, Y. Hu, C. Humblet, Comparison of several molecular docking programs: pose prediction and virtual screening accuracy, J. Chem. Inf. Model. 49 (6) (2009) 1455–1474;
(b) G.L. Warren, C.W. Andrews, A.M. Capelli, B. Clarke, J. LaLonde, M.H. Lambert, M. Lindvall, N. Nevins, S.F. Semus, S. Senger, G. Tedesco, I.D. Wall, J.M. Woolven, C.E. Peishoff, M.S. Head, A critical assessment of docking programs and scoring functions, J. Med. Chem. 49 (20) (2006) 5912–5931.

[8] T. Scior, A. Bender, G. Tresadern, J.L. Medina-Franco, K. Martinez-Mayorga, T. Langer, K. Cuanalo-Contreras, D.K. Agrafiotis, Recognizing pitfalls in virtual screening: a critical review, J. Chem. Inf. Model. 52 (4) (2012) 867–881.

[9] (a) E.M. Nurminen, M. Pihlavisto, L. Lazar, U. Pentikainen, F. Fulop, O.T. Pentikainen, Novel hydrazine molecules as tools to understand the flexibility of vascular adhesion protein-1 ligand-binding site: toward more selective inhibitors, J. Med. Chem. 54 (7) (2011) 2143–2154;
(b) W. Sippl, Receptor-based 3D QSAR analysis of estrogen receptor ligands-merging the accuracy of receptor-based alignments with the computational efficiency of ligand-based methods, J. Comput. Aided Mol. Des. 14 (6) (2000) 559–572;
(c) A.A. Soderholm, P.T. Lehtovuori, T.H. Nyronen, Docking and three-dimensional quantitative structure-activity relationship (3D QSAR) analyses of nonsteroidal progesterone receptor ligands, J. Med. Chem. 49 (14) (2006) 4261–4268.

[10] (a) S.I. Virtanen, O.T. Pentikainen, Efficient virtual screening using multiple protein conformations described as negative images of the ligand-binding site, J. Chem. Inf. Model. 50 (6) (2010) 1005–1011;
(b) S.P. Niinivehmas, S.I. Virtanen, J.V. Lehtonen, P.A. Postila, O.T. Pentikainen, Comparison of virtual high-throughput screening methods for the identification of phosphodiesterase-5 inhibitors, J. Chem. Inf. Model. 51 (6) (2011) 1353–1363;
(c) S.P. Niinivehmas, K. Salokas, S. Latti, H. Raunio, O.T. Pentikainen, Ultrafast protein structure-based virtual screening with Panther, J. Comput. Aided Mol. Des. 29 (10) (2015) 989–1006.

[11] V. Shubina, S. Niinivehmas, O.T. Pentikainen, Reliability of virtual screening methods in prediction of PDE4B inhibitor activity, Curr. Drug Discov. Technol. 12 (2) (2015) 117–126.

[12] Y. Shang, M. Brown, Molecular determinants for the tissue specificity of SERMs, Science 295 (5564) (2002) 2465–2468.

[13] (a) J. Shen, C. Tan, Y. Zhang, X. Li, W. Li, J. Huang, X. Shen, Y. Tang, Discovery of potent ligands for estrogen receptor beta by structure-based virtual screening, J. Med. Chem. 53 (14) (2010) 5361–5365;
(b) A.J. Knox, M.J. Meegan, V. Sobolev, D. Frost, D.M. Zisterer, D.C. Williams, D.G. Lloyd, Target specific virtual screening: optimization of an estrogen receptor screening platform, J. Med. Chem. 50 (22) (2007) 5301–5310;
(c) E. Therrien, P. Englebienne, A.G. Arrowsmith, R. Mendoza-Sanchez, C.R. Corbeil, N. Weill, V. Campagna-Slater, N. Moitessier, Integrating medicinal chemistry, organic/combinatorial chemistry, and computational chemistry for the discovery of selective estrogen receptor modulators with Forecaster, a novel platform for drug discovery, J. Chem. Inf. Model. 52 (1) (2012) 210–224.

[14] A. Gaulton, L.J. Bellis, A.P. Bento, J. Chambers, M. Davies, A. Hersey, Y. Light, S. McGlinchey, D. Michalovich, B. Al-Lazikani, J.P. Overington, ChEMBL: a large-scale bioactivity database for drug discovery, Nucleic Acids Res. 40 (2012) D1100–D1107.

[15] Y. Liu, X.T. Sang, W.S. Gao, Y.L. Mao, Y.W. Liu, H.F. Liu, Z.Y. Yang, S.Z. Yang, S.X. Zhong, J.F. Huang, [The first case of primary epithelial-myoepithelial carcinoma in the liver], Zhonghua wai ke za zhi [Chin. J. Surg.] 44 (21) (2006) 1477–1479.

[16] K.S. Watts, P. Dalal, R.B. Murphy, W. Sherman, R.A. Friesner, J.C. Shelley, ConfGen: a conformational search method for efficient generation of bioactive conformers, J. Chem. Inf. Model. 50 (4) (2010) 534–546.

[17] T.A. Halgren, Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94, J. Comput. Chem. 17 (5–6) (1996) 490–519.

[18] S.L. Dixon, A.M. Smondyrev, E.H. Knoll, S.N. Rao, D.E. Shaw, R.A. Friesner, PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results, J. Comput. Aided Mol. Des. 20 (10–11) (2006) 647–671.

[19] Schrödinger Release, Protein preparation wizard, Epik version 2.2, Impact version 5.7, Prime version 3.0, Schrödinger, LLC, New York, NY, 2011.

[20] (a) R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, P.S. Shenkin, Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, J. Med. Chem. 47 (7) (2004) 1739–1749;
(b) R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, P.C. Sanschagrin, D.T. Mainz, Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, J. Med. Chem. 49 (21) (2006) 6177–6196.

[21] O. Korb, T. Stutzle, T.E. Exner, Empirical scoring functions for advanced protein-ligand docking with PLANTS, J. Chem. Inf. Model. 49 (1) (2009) 84–96.

[22] (a) P.A. Postila, M. Ylilauri, O.T. Pentikainen, Full and partial agonism of ionotropic glutamate receptors indicated by molecular dynamics simulations, J. Chem. Inf. Model. 51 (5) (2011) 1037–1047;
(b) M. Ylilauri, O.T. Pentikainen, Structural mechanism of *N*-methyl-*D*-aspartate receptor type 1 partial agonism, PLoS One 7 (10) (2012) e47604;
(c) M. Ylilauri, O.T. Pentikainen, MMGBSA as a tool to understand the binding affinities of filamin-peptide interactions, J. Chem. Inf. Model. 53 (10) (2013) 2626–2633.

[23] G.J. Kleywegt, T.A. Jones, Detection, delineation, measurement and display of cavities in macromolecular structures, Acta Crystallogr. Sect. D: Biol. Crystallogr. 50 (Pt. 2) (1994) 178–185.

[24] J. Wang, W. Wang, P.A. Kollman, D.A. Case, Automatic atom type and bond type perception in molecular mechanical calculations, J. Mol. Graphics Modell. 25 (2) (2006) 247–260.

[25] Y. Duan, C. Wu, S. Chowdhury, M.C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. Kollman, A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations, J. Comput. Chem. 24 (16) (2003) 1999–2012.

[26] J. Aaqvist, Ion-water interaction potentials derived from free energy perturbation simulations, J. Phys. Chem. 94 (21) (1990) 8021–8024.

[27] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, Scalable molecular dynamics with NAMD, J. Comput. Chem. 26 (16) (2005) 1781–1802.

[28] M.J. Vainio, J.S. Puranen, M.S. Johnson, ShaEP: molecular overlay based on shape and electrostatic potential, J. Chem. Inf. Model. 49 (2) (2009) 492–502.

[29] S. Kirkiacharian, A.T. Lormier, M. Resche-Rigon, F. Bouchoux, E. Cerede, [Synthesis and binding affinity of 3-aryl-7-hydroxycoumarins to human alpha and beta estrogen receptors], Ann. Pharm. Fr. 61 (1) (2003) 51–56.

[30] (a) J. Baell, M.A. Walters, Chemistry: Chemical con artists foil drug discovery, Nature 513 (7519) (2014) 481–483;
(b) J.B. Baell, G.A. Holloway, New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays, J. Med. Chem. 53 (7) (2010) 2719–2740.

Electronic Supplementary Information

Supplementary material is listed in the order of appearance in the main text.

Supplementary Table S1. Chemical diversity of DS1 ERα ligands used in this study

Chemotype	n	n %	Activity range pIC_{50}	Activity average
2,3 Diaryl tolylchroman-6-ol	8	7.9	9.1 - 9.5	9.2
2,3 Diaryldihydrobenzoxathin	24	23.8	9.1 - 9.7	9.2
2-Arylbenzo[b]thiophene	3	3.0	8.7 - 9.7	9.2
2-Phenyl spironindene	2	2.0	8.4 - 9.0	8.7
2-Phenyl-1H-indole	9	8.9	8.1 - 9.4	8.8
3,4 Diarylisothiochromanol	5	5.0	9.1 - 9.2	9.1
5,6 diaryl tetrahydronapthol	3	3.0	8.9 - 9.2	9.1
6H-benzo[c]chromene-6-one	8	7.9	5.0 -7.0	6.4
Dihydrobenzoxathin	6	5.9	8.5 - 8.6	8.6
Isoflavone	2	2.0	5.7 - 7.0	6.4
Steroid	5	5.0	5.7 - 8.9	6.8
syn-Dihydrobenzoxathiin	4	4.0	8.6 - 9.3	9.0
Tetrahydrofluorenone-fused pyrazole	9	8.9	6.8 -7.6	7.2
Tetrahydrofluorenone-fused triazole	4	4.0	5.7 - 7.5	6.5
Tetrahydrofluorenone	6	6.0	6.9 - 7.5	7.1
9,10-Dihydrophenanthrene	1	1.0	6.6	6.6
Conformationally restricted tetrahydrofluorenone	1	1.0	7.0	7.0
Coumestan	1	1.0	8.0	8.0

Supplementary Table S2   DS1 ligands and their ERα inhibitory activity data used in this study

S.No	Ligand ID	Chemotype (Subtype or Name)	ERα IC_{50} nM	ERα pIC_{50} M
1	ChEMBL180146	2,3 Diaryl tolylchroman-6-ol	0.8	9.1
2	ChEMBL180995	2,3 Diaryl tolylchroman-6-ol	0.9	9.0
3	ChEMBL181248	2,3 Diaryl tolylchroman-6-ol	0.7	9.2
4	ChEMBL181368	2,3 Diaryl tolylchroman-6-ol	0.3	9.5
5	ChEMBL181369	2,3 Diaryl tolylchroman-6-ol	0.9	9.0
6	ChEMBL181404	2,3 Diaryl tolylchroman-6-ol	0.8	9.1
7	ChEMBL359633	2,3 Diaryl tolylchroman-6-ol	0.5	9.3
8	ChEMBL361005	2,3 Diaryl tolylchroman-6-ol	0.8	9.1
9	ChEMBL179852	2,3 Diaryldihydrobenzoxathin	0.7	9.2
10	ChEMBL182794	2,3 Diaryldihydrobenzoxathin	0.7	9.2
11	ChEMBL182980	2,3 Diaryldihydrobenzoxathin	0.5	9.3
12	ChEMBL183092	2,3 Diaryldihydrobenzoxathin	0.8	9.1
13	ChEMBL183263	2,3 Diaryldihydrobenzoxathin	0.5	9.3
14	ChEMBL183333	2,3 Diaryldihydrobenzoxathin	0.6	9.2
15	CHEMBL183371	2,3 Diaryldihydrobenzoxathin	0.9	9.0
16	ChEMBL183388	2,3 Diaryldihydrobenzoxathin	0.9	9.0
17	CHEMBL183467	2,3 Diaryldihydrobenzoxathin	0.4	9.4
18	ChEMBL184202	2,3 Diaryldihydrobenzoxathin	0.7	9.2
19	ChEMBL184360	2,3 Diaryldihydrobenzoxathin	0.6	9.2
20	ChEMBL184367	2,3 Diaryldihydrobenzoxathin	0.6	9.2
21	ChEMBL184421	2,3 Diaryldihydrobenzoxathin	0.6	9.2
22	ChEMBL184598	2,3 Diaryldihydrobenzoxathin	0.8	9.1
23	ChEMBL185083	2,3 Diaryldihydrobenzoxathin	0.4	9.4
24	ChEMBL181862	2,3 Diaryldihydrobenzoxathin	0.5	9.3
25	ChEMBL361601	2,3 Diaryldihydrobenzoxathin	0.4	9.4
26	ChEMBL367350	2,3 Diaryldihydrobenzoxathin	0.7	9.2
27	ChEMBL367574	2,3 Diaryldihydrobenzoxathin	0.7	9.2
28	ChEMBL368688	2,3 Diaryldihydrobenzoxathin	0.9	9.0
29	ChEMBL369545	2,3 Diaryldihydrobenzoxathin	0.5	9.3
30	ChEMBL427324	2,3 Diaryldihydrobenzoxathin	0.3	9.5
31	ChEMBL433769	2,3 Diaryldihydrobenzoxathin	0.3	9.5
32	ChEMBL434525	2,3 Diaryldihydrobenzoxathin	0.7	9.2
33	ChEMBL198803	2-Arylbenzo[b]thiophene	0.2	9.7
34	ChEMBL372337	2-Arylbenzo[b]thiophene	0.6	9.2
35	ChEMBL81	2-Arylbenzo[b]thiophene (Raloxifene)	1.8	8.7
36	ChEMBL267385	2-Phenyl spironindene	1	9.0
37	ChEMBL281499	2-Phenyl spironindene	4.3	8.4
38	ChEMBL46740	2-Phenyl-1H-indole	0.6	9.2
39	ChEMBL240438	2-Phenyl-1H-indole	2	8.7
40	ChEMBL241256	2-Phenyl-1H-indole	6.6	8.2
41	ChEMBL241301	2-Phenyl-1H-indole	1	9.0
42	ChEMBL241303	2-Phenyl-1H-indole	8	8.1
43	ChEMBL391910	2-Phenyl-1H-indole	1	9.0
44	ChEMBL437190	2-Phenyl-1H-indole	4	8.4
45	ChEMBL198914	2-Phenyl-1H-indole (1-Benzyl-sub)	0.5	9.3
46	ChEMBL372808	2-Phenyl-1H-indole (1-Benzyl-sub)	0.4	9.4
47	ChEMBL182690	3,4 Diarylisothiochromanol	0.7	9.2
48	ChEMBL182902	3,4 Diarylisothiochromanol	0.8	9.1
49	ChEMBL183090	3,4 Diarylisothiochromanol	0.8	9.1

50	ChEMBL183399	**3,4 Diarylisothiochromanol**	0.6	9.2
51	ChEMBL304552	**3,4 Diarylisothiochromanol**	0.9	9.0
52	ChEMBL197495	**5,6 diaryl tetrahydronapthol**	0.6	9.2
53	ChEMBL437695	**5,6 diaryl tetrahydronapthol**	0.5	9.3
54	ChEMBL328190	**Tetrahydronaphthalen-2-ol (Lasofoxifene)**	1.3	8.9
55	ChEMBL203072	**6H-benzo[c]chromene-6-one**	129	6.9
56	ChEMBL206500	**6H-benzo[c]chromene-6-one**	229	6.6
57	ChEMBL380717	**6H-benzo[c]chromene-6-one**	716	6.1
58	ChEMBL204922	**6H-benzo[c]chromene-6-one**	159	6.8
59	ChEMBL205934	**6H-benzo[c]chromene-6-one**	785	6.1
60	ChEMBL206547	**6H-benzo[c]chromene-6-one**	101	7.0
61	ChEMBL380838	**6H-benzo[c]chromene-6-one**	10,000	5.0
62	ChEMBL381697	**6H-benzo[c]chromene-6-one**	225	6.6
63	ChEMBL68489	**Dihydrobenzoxathin**	3.1	8.5
64	ChEMBL85090	**Dihydrobenzoxathin**	2.6	8.6
65	ChEMBL85650	**Dihydrobenzoxathin**	3	8.5
66	ChEMBL313941	**Dihydrobenzoxathin**	2.7	8.6
67	ChEMBL315271	**Dihydrobenzoxathin**	3	8.5
68	ChEMBL431611	**Dihydrobenzoxathin**	2.5	8.6
69	ChEMBL8145	**Isoflavone (Daidzein)**	2160	5.7
70	ChEMBL44	**Isoflavone (Genistein)**	92	7.0
71	ChEMBL135	**Steroid (17β)-estra-1,3,5(10)-triene-3,17-diol**	1.35	8.9
72	ChEMBL250110	**Steroid (Androstene-3,5diene)**	110	7.0
73	ChEMBL398456	**Steroid (Androstene-3,5diene)**	1440	5.8
74	ChEMBL77135	**Steroid (Androstenediol)**	210	6.7
75	ChEMBL246138	**Steroid (Androstenediol)**	2240	5.6
76	ChEMBL92660	**syn-Dihydrobenzoxathiin**	2.4	8.6
77	ChEMBL93793	**syn-Dihydrobenzoxathiin**	2	8.7
78	ChEMBL94030	**syn-Dihydrobenzoxathiin**	0.5	9.3
79	ChEMBL94116	**syn-Dihydrobenzoxathiin**	0.51	9.3
80	ChEMBL211349	**Tetrahydrofluorenone-fused pyrazole**	46	7.3
81	ChEMBL211816	**Tetrahydrofluorenone-fused pyrazole**	93	7.0
82	ChEMBL211887	**Tetrahydrofluorenone-fused pyrazole**	82	7.1
83	ChEMBL213829	**Tetrahydrofluorenone-fused pyrazole**	23	7.6
84	ChEMBL215409	**Tetrahydrofluorenone-fused pyrazole**	58	7.2
85	ChEMBL377644	**Tetrahydrofluorenone-fused pyrazole**	176	6.8
86	ChEMBL384268	**Tetrahydrofluorenone-fused pyrazole**	49.2	7.3
87	ChEMBL385680	**Tetrahydrofluorenone-fused pyrazole**	50.9	7.3
88	ChEMBL425439	**Tetrahydrofluorenone-fused pyrazole**	40	7.4
89	ChEMBL214692	**Tetrahydrofluorenone-fused triazole**	34	7.5
90	ChEMBL215802	**Tetrahydrofluorenone-fused triazole**	1906	5.7
91	ChEMBL385949	**Tetrahydrofluorenone-fusedtriazole**	647	6.2
92	ChEMBL441366	**Tetrahydrofluorenone-fusedtriazole**	220	6.7
93	ChEMBL208978	**Tetrahydrofluorenone**	128	6.9
94	ChEMBL209402	**Tetrahydrofluorenone**	30	7.5
95	ChEMBL209429	**Tetrahydrofluorenone**	70.5	7.2
96	ChEMBL209851	**Tetrahydrofluorenone**	141	6.9
97	ChEMBL379821	**Tetrahydrofluorenone**	124	6.9
98	ChEMBL380469	**Tetrahydrofluorenone**	41	7.4
99	ChEMBL205119	**9,10-Dihydrophenanthrene (Effusol)**	240	6.6
100	ChEMBL379008	**Conformationally restricted tetrahydrofluorenone**	97	7.0
101	ChEMBL30707	**Coumestan (Coumestrol)**	11	8.0

Supplementary Table S3   DS1 and DS2 ligands and their numerical data.

Name	Pharm. set[a]	Pred. act.	Shaep	Glide sp[b]	Plants chemplp	Hybrid docking[b]
CHEMBL135	training	7.80	0.34	-9.57	-84.64	-8.45
CHEMBL179852	test	8.79	0.42	-11.44	-107.99	-10.40
CHEMBL180146	test	9.30	0.39	-11.87	-93.58	-9.99
CHEMBL180995	test	9.33	0.42	-11.75	-96.36	-10.05
CHEMBL181248	training	9.07	0.43	-11.56	-94.14	-9.86
CHEMBL181368	training	9.39	0.41	-11.81	-94.91	-10.02
CHEMBL181369	training	9.28	0.40	-11.59	-92.99	-9.82
CHEMBL181404	training	9.28	0.40	-11.82	-94.38	-10.00
CHEMBL181862	training	9.07	0.42	-9.77	-114.18	-9.83
CHEMBL182690	training	9.15	0.42	-11.42	-100.60	-10.07
CHEMBL182794	test	8.61	0.44	-11.75	-115.98	-10.90
CHEMBL182902	training	9.16	0.43	-11.54	-111.37	-10.60
CHEMBL182980	training	9.23	0.42	-11.53	-115.44	-10.77
CHEMBL183090	training	9.23	0.43	-11.81	-107.64	-10.57
CHEMBL183092	training	9.22	0.42	-11.23	-113.24	-10.52
CHEMBL183263	training	8.49	0.45	-11.75	-116.04	-10.90
CHEMBL183333	test	8.99	0.42	-11.28	-129.66	-11.26
CHEMBL183371	training	9.05	0.41	-11.33	-111.76	-10.51
CHEMBL183388	training	9.23	0.41	-11.54	-115.33	-10.77
CHEMBL183399	training	8.92	0.42	-11.55	-113.68	-10.70
CHEMBL183467	training	8.68	0.42	-11.49	-115.44	-10.75
CHEMBL184202	training	9.21	0.41	-11.58	-117.46	-10.88
CHEMBL184360	training	9.07	0.40	-11.54	-113.03	-10.67
CHEMBL184367	test	9.02	0.41	-11.45	-114.00	-10.66
CHEMBL184421	training	9.03	0.42	-9.86	-114.29	-9.88
CHEMBL184598	training	9.06	0.42	-9.99	-112.71	-9.88
CHEMBL185083	training	9.29	0.41	-11.53	-111.75	-10.61
CHEMBL197495	training	9.30	0.43	-11.76	-135.90	-11.77
CHEMBL198803	training	9.51	0.44	-11.37	-105.18	-10.24
CHEMBL198914	test	9.03	0.41	-10.95	-122.16	-10.77
CHEMBL203072	training	6.29	0.39	-7.85	-75.38	-7.19
CHEMBL204922	training	6.55	0.39	-8.01	-70.49	-7.06
CHEMBL205119	training	6.54	0.38	-8.37	-76.42	-7.50
CHEMBL205934	training	6.29	0.37	-7.85	-75.80	-7.21
CHEMBL206500	training	6.55	0.37	-8.05	-71.42	-7.12
CHEMBL206547	training	6.55	0.42	-8.41	-73.18	-7.38
CHEMBL208978	test	5.94	0.40	-8.98	-83.68	-8.12
CHEMBL209402	training	7.34	0.40	-8.73	-90.41	-8.28
CHEMBL209429	training	7.16	0.39	-9.63	-81.51	-8.35
CHEMBL209851	test	6.37	0.37	-8.18	-82.48	-7.66
CHEMBL211349	training	7.32	0.40	-9.68	-85.11	-8.53
CHEMBL211816	test	7.12	0.38	-8.70	-83.37	-7.96
CHEMBL211887	training	7.12	0.40	-8.09	-72.72	-7.20
CHEMBL213829	test	7.12	0.40	-9.01	-76.75	-7.83
CHEMBL214692	test	6.77	0.40	-9.46	-83.54	-8.35

CHEMBL215409	test	7.12	0.37	-8.66	-77.82	-7.70
CHEMBL215802	training	6.43	0.41	-7.90	-76.26	-7.25
CHEMBL240438	test	8.43	0.37	-7.81	-130.08	-9.54
CHEMBL241256	training	8.30	0.40	-7.91	-120.17	-9.16
CHEMBL241301	training	8.89	0.38	-8.56	-110.37	-9.06
CHEMBL241303	training	7.96	0.40	-7.35	-121.08	-8.92
CHEMBL246138	training	4.70	0.36	-9.20	-80.78	-8.10
CHEMBL250110	training	6.08	0.33	-8.72	-84.20	-8.01
CHEMBL267385	test	8.34	0.42	-11.39	-116.61	-10.75
CHEMBL281499	training	8.19	0.42	-10.08	-86.66	-8.79
CHEMBL304552	training	9.07	0.42	-11.65	-110.38	-10.61
CHEMBL30707	training	7.93	0.40	-8.33	-77.22	-7.51
CHEMBL313941	training	8.52	0.40	-10.81	-103.94	-9.91
CHEMBL315271	training	8.59	0.41	-10.22	-111.59	-9.94
CHEMBL328190	training	9.11	0.43	-11.70	-135.33	-11.71
CHEMBL359633	training	9.35	0.41	-11.82	-93.91	-9.98
CHEMBL361005	test	8.16	0.40	-11.77	-91.22	-9.84
CHEMBL361601	training	9.17	0.44	-11.50	-116.64	-10.80
CHEMBL367350	test	9.16	0.44	-11.73	-113.35	-10.78
CHEMBL367574	training	9.27	0.43	-11.49	-114.85	-10.72
CHEMBL368688	test	8.48	0.41	-11.70	-113.62	-10.77
CHEMBL369545	training	8.47	0.43	-11.73	-113.40	-10.78
CHEMBL372337	training	9.29	0.42	-11.32	-105.59	-10.24
CHEMBL372808	test	9.22	0.43	-10.83	-121.70	-10.69
CHEMBL377644	training	7.12	0.41	-8.59	-84.87	-7.97
CHEMBL379008	training	6.40	0.40	-9.15	-86.83	-8.34
CHEMBL379821	training	7.92	0.39	-7.93	-80.90	-7.47
CHEMBL380469	training	7.20	0.39	-9.84	-81.10	-8.43
CHEMBL380717	training	6.31	0.36	-7.61	-70.55	-6.86
CHEMBL380838	training	6.31	0.38	-8.87	-77.88	-7.81
CHEMBL381697	test	7.23	0.39	nd	-63.70	nd
CHEMBL384268	training	7.12	0.39	-9.20	-78.02	-7.98
CHEMBL385680	test	7.12	0.40	-8.39	-75.85	-7.48
CHEMBL385949	test	6.12	0.39	-8.42	-77.79	-7.58
CHEMBL391910	training	9.03	0.37	-8.66	-111.19	-9.15
CHEMBL398456	test	5.50	0.36	-8.55	-83.00	-7.87
CHEMBL425439	training	7.12	0.39	-8.43	-77.57	-7.58
CHEMBL427324	test	8.68	0.44	-11.50	-116.71	-10.81
CHEMBL431611	training	8.59	0.41	-10.67	-107.44	-9.99
CHEMBL433769	test	9.10	0.42	-11.81	-114.38	-10.86
CHEMBL434525	training	9.05	0.42	-11.79	-114.06	-10.84
CHEMBL437190	test	8.10	0.38	-8.52	-126.90	-9.76
CHEMBL437695	training	9.15	0.43	-11.81	-136.98	-11.84
CHEMBL44	training	7.13	0.41	-9.48	-76.93	-8.07
CHEMBL441366	training	5.67	0.40	-8.34	-81.59	-7.71
CHEMBL46740	training	9.09	0.40	-10.13	-129.90	-10.69
CHEMBL68489	training	9.07	0.41	-11.66	-109.52	-10.58
CHEMBL77135	training	6.55	0.37	-9.18	-79.34	-8.03
CHEMBL81	training	8.72	0.41	-11.31	-104.87	-10.20

LIGS01530090	nd	8.05	0.36	-10.25	-74.42	-8.35
LIGS01530602	nd	8.10	0.33	-8.43	-74.98	-7.46
LIGS01530690	nd	8.05	0.36	-8.69	-71.02	-7.42
LIGS01531019	nd	7.62	0.38	-9.98	-86.45	-8.74
LIGS01543842	nd	8.29	0.35	-11.90	-131.63	-11.65
LIGS01545572	nd	8.91	0.34	-6.71	-88.97	-7.21
LIGS01585847	nd	7.74	0.35	-11.21	-90.63	-9.53
LIGS01730387	nd	8.15	0.33	nd	-72.60	nd
LIGS01914469	nd	7.69	0.39	-10.37	-121.26	-10.44
LIGS03793538	nd	8.45	0.38	-10.05	-111.88	-9.87
LIGS03815450	nd	8.29	0.37	-11.50	-108.85	-10.47
LIGS03815451	nd	8.35	0.37	-11.78	-107.12	-10.53
LIGS03815452	nd	8.16	0.30	-10.52	-93.26	-9.30
LIGS03815453	nd	8.44	0.36	-10.66	-127.48	-10.85
LIGS03815454	nd	8.34	0.34	-9.90	-120.68	-10.18
LIGS03815455	nd	8.82	0.37	-10.59	-103.82	-9.79
LIGS03815456	nd	8.91	0.36	-10.83	-107.64	-10.08
LIGS03815457	nd	8.65	0.34	-9.35	-91.54	-8.64
LIGS03815458	nd	8.56	0.36	-10.37	-87.90	-8.99
LIGS03815459	nd	8.45	0.37	-9.65	-106.48	-9.44
LIGS03815460	nd	8.03	0.35	-10.46	-121.65	-10.50
LIGS03815461	nd	8.17	0.36	-10.03	-120.25	-10.23
LIGS03815462	nd	7.95	0.34	-10.90	-118.03	-10.56
LIGS03815463	nd	7.38	0.36	-8.53	-101.14	-8.65
LIGS03815464	nd	7.50	0.39	-9.36	-94.88	-8.79
LIGS03815465	nd	7.48	0.34	nd	-102.92	nd
LIGS03815466	nd	7.46	0.38	-8.80	-101.98	-8.82
LIGS03815467	nd	7.38	0.38	-8.94	-96.29	-8.64
LIGS03815468	nd	7.38	0.39	-9.14	-92.57	-8.58
LIGS03815469	nd	7.48	0.36	nd	-102.75	nd
LIGS03815470	nd	7.57	0.38	-8.54	-102.86	-8.73
LIGS03815471	nd	7.38	0.40	-8.43	-93.19	-8.25
LIGS03815472	nd	7.38	0.38	-9.32	-103.14	-9.13
LIGS03815473	nd	7.38	0.41	-8.50	-94.67	-8.35
LIGS03815474	nd	7.38	0.38	-9.30	-92.40	-8.65
LIGS03815475	nd	7.38	0.37	-9.17	-100.81	-8.95
LIGS03815476	nd	7.38	0.37	-8.13	-97.09	-8.27
LIGS03815477	nd	7.95	0.37	-11.31	-107.80	-10.33
LIGS03815478	nd	8.01	0.38	-10.61	-118.53	-10.44

[a] Training and test sets appear only in pharmacophore/3D-QSAR which was done utilizing DS1
[b] One DS1 molecule and three DS2 molecules were not docked by Glide and thus are also absent from hydrid docking

Detailed synthesis and characterization data for molecules 1-5.
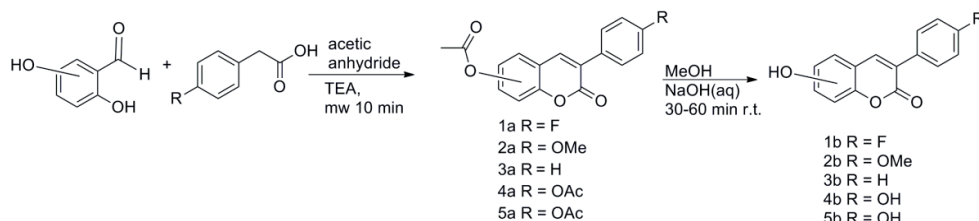
Supplementary Fig. 1. The general procedure for the synthesis of coumarin derivatives

A mixture of salicylaldehyde derivative (2 mmol) and phenyl acetic acid derivative (2.1 mmol), acetic acid anhydride (0.6 ml) and triethylamine (0.36 ml) were placed in a microwave reactor tube and this mixture was heated at 160-170 °C with the CEM Discovery microwave apparatus for 10 min. After cooling, 2 ml of 10% $NaHCO_3$ solution was added and the precipitate was filtered, dried and recrystallized from ETOH/H_2O or acetone/H_2O mixture. The acetyl group(s) were removed by treating the compound with MeOH/NaOH(aq) solution for 30-60 min at r.t.. The solution was acidified with HCl(aq,) and the precipitate was filtered and recrystallized if needed.

Molecule 1

7-acetoxy-3-(4-fluorophenyl)coumarin (1a) [1]

Yield: 92%; m.p. = 204-205 °C (lit. [78] 203-204°C); ^{1}H-NMR (300 MHz, d^6-DMSO) δ: 2.32 (s, 3H, CH$_3$CO), 7.18 (dd, J^3 = 8.4 Hz, J^4 = 2.1 Hz, 1H, H-6), 7.27-7.33 (m, 3H, H-3' and H-6), 7.75-7.81 (m, 3H, H-2' and H-8), 8.25 (s, 1H, H-4); ^{13}C-NMR (100.6 MHz, d^6-DMSO) δ: 20.82, 109.63, 114.95 (C-3', $J^2_{(C-F)}$ = 21.5 Hz), 115.23 (C-3', $J^2_{(C-F)}$ = 21.5 Hz), 117.26, 118.75, 125.11, 129.41, 130.57 (C-2', $J^3_{(C-F)}$ = 8.3 Hz), 130.68 (C-2', $J^3_{(C-F)}$ = 8.3 Hz), 130.88 (C-1', $J^4_{(C-F)}$ = 3.2 Hz),130.92 (C-1', $J^4_{(C-F)}$ = 3.2 Hz), 140.02, 152.68, 153.38, 159.45, 160.55 (C-4', $J^1_{(C-F)}$ = 245.9 Hz), 163.81 (C-4', $J^1_{(C-F)}$ = 245.9 Hz).

7-hydroxy-3-(4-fluorophenyl)coumarin (1b) [1]

Yield: 81%; m.p. = 237-239 °C (lit. [78] 238-239 °C); ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 6.75 (d, J^4 = 2.3 Hz, 1H, H-8), 6.82 (dd, J^3 = 8.5 Hz, J^4 = 2.3 Hz, 1H, H-6), 7.28 (dd, J^3 = 7.8 Hz, $J^{H,F}$ = 7.8, 2H, H-3'), 7.58 (d, J^3 = 8.5 Hz, 1H, H-5), 7.73 (dd, J^3 = 8.9 Hz, $J^{H,F}$ = 5.5 Hz, 2H, H-2'), 8.14 (s, H-4), 10.64 (broad) (s, 1H, O-H); ^{13}C-NMR (100.6 MHz, d^6-DMSO) δ:

101.72, 111.90, 113.44, 114.90 (C-3', $J^2_{(C-F)}$ = 22 Hz), 115.12 (C-3', $J^2_{(C-F)}$ = 22 Hz), 121.12, 129.96, 130.32 (C-2', $J^3_{(C-F)}$ = 8 Hz), 130.40 (C-2', $J^3_{(C-F)}$ = 8 Hz), 131.45 (C-1', $J^4_{(C-F)}$ = 3 Hz), 131.48 (C-1', $J^4_{(C-F)}$ = 3 Hz), 141.07, 154.91, 160.07, 160.63 (C-4', $J^1_{(C-F)}$ = 246 Hz), 161.32, 163.07 (C-4', $J^1_{(C-F)}$ = 246 Hz); ESI-MS: m/z (rel. abund. %) 279 (M + Na$^+$); elemental anal. for $C_{15}H_9FO_3$, calc. C% 7.31, H% 3.54, found C% 70.16, H% 3.76.

Molecule 2

7-acetoxy-3-(4-methoxyphenyl)coumarin (2a) [1]

Yield: 88%; m.p. = 182-183 $^{\circ}$C (lit. [78] 181-182 °C); ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 2.32 (s, 3H, CH$_3$CO), 3.78 (s, 3H, CH$_3$O), 7.05 (d, J^3 = 8.9 Hz, 2H, H-3'), 7.17 (dd, J^3 = 8.4 Hz, J^4 = 2.2 Hz, 1H, H-6), 7.29 (d, J^4 = 2.2 Hz, 1H, H-8), 7.70 (d, J^3 = 8.9 Hz, 2H, H-2'), 7.79 (d, J^3 = 8.9 Hz, H-5), 8.19 (s, 1H, H-4); ^{13}C-NMR (100.6 MHz, d^6-DMSO) δ:20.85, 55.22, 109.57, 113.68, 117.49, 118.68, 125.69, 126.72, 129.19, 129.75, 138.62, 152.36, 153.15, 159.59, 168.81; HRMS: m/z 333.072 (calc. 333.051) for M + Na$^+$; elemental anal. for $C_{18}H_{14}O_5$, calc. C% 69.67, H% 4.55, found C% 69.58, H% 4.52.

7-hydroxy-3-(4-methoxyphenyl)coumarin (2b) [1]

Yield: 84%; m.p. = 214-215 °C (lit. [78] 215-216 °C); ^{1}H-NMR (400 MHz, d^6-DMSO) δ: , 3.78 (s, 3H, CH$_3$O), 6.73 (d, J^4 = 2.3 Hz, 1H, H-8), 6.80 (dd, J^3 = 8.5 Hz, J^4 = 2.3 Hz, 1H, H-6), 6.99 (d, J^3 = 8.9 Hz, 2H, H-3'),, 7.55 (d, J^3 = 8.6 Hz, H-5), 7.64 (d, J^3 = 8.9 Hz, 2H, H-2'), 8.07 (s, 1H, H-4), 10.52 (s, 1H, O-H) ;^{13}C-NMR (100.6 MHz, d^6-DMSO) δ: 55.18, 101.66, 112.10, 113.29, 113.61, 121.84, 127.30, 129.48, 129.70, 139.73, 154.64, 159.14, 160.20, 160.88; HRMS: m/z 290.969 (calc. 291.063) for M + Na$^+$.

Molecule 3

7-acetoxy-3-phenylcoumarin (3a)

Yield: 90%; m.p. 186-188 °C (lit. [78] 187-188 °C); ^{1}H-NMR as described in literature. [1a]

7-acetoxy-3-phenyl coumarin (3b)

Yield: 90%; m.p. 204-206 °C (lit. [78] 205-206 °C); ^{1}H-NMR as described in literature. [1a]

Molecule 4

7-acetoxy-3-(4-acetoxyphenyl)coumarin (4a) [2]

Yield 89%; %; ^{1}H-NMR (300 MHz, d^6-DMSO) δ: 2.30 (s, 1H, CH$_3$CO), 2.38 (s, 1H, CH$_3$CO), 7.19 (dd, J^3 = 8.4 Hz, J^4 = 2.2 Hz, 1H, H-6), 7.24 (t, J^3 = 8.7 Hz, 2H, H-3'), 7.31 (d, J^4 = 2.2 Hz, 1H, H-8), 7.77 (t, J^3 = 8.7 Hz, 2H, H-2'), 7.80 (d, J^3 = 8.6 Hz, 2H, H-5), 8.27 (s,

1H, H-4). ^{13}C-NMR (75.48 MHz, d^6-DMSO) δ: 20.74, 20.82, 116.97, 119.90, 120.77, 121.67, 125.49, 126.75, 129.74, 131.95, 139.84, 146.44, 150.44, 150.77, 159.51, 169.10, 169.23.

7-hydroxy-3-(4-hydroxyphenyl)coumarin (4b) [3]

Yield 85%; ^{1}H-NMR (300 MHz, d^6-DMSO) δ: 6.73 (d, J^4 = 2.2 Hz, 1H, H-8), 6.78-6.84 (m, 3H, H-3', H-6), 7.51-7.57 (m, 3H, H-2´, H-5), 8.01 (s, 1H, H-4), 9.62 (s, 1H, -OH), 10.49 (s, 1H, -OH); ^{13}C-NMR (75.48 MHz, d^6-DMSO) δ: 101.64, 112.17, 113.24, 114.98, 122.23, 125.68, 129.49, 129.55, 139.12, 154.52, 157.46, 160.25, 160.71.

Molecule 5

8-acetoxy-3-(4-acetoxyphenyl)coumarin (5a)

Yield 95%; ^{1}H-NMR (300 MHz, d^6-DMSO) δ: 2.30 (s, 1H, CH$_3$CO), 2.40 (s, 1H, CH$_3$CO), 7.24 (d, J^3 = 8.7 Hz, 2H, H-3'), 7.40 (t, J^3 = 7.8 Hz, 1H, H-6), 7.48 (dd, J^3 = 8.0 Hz, J^4 = 1.7 Hz, 1H, H-7), 7.69 (dd, J^3 = 7.7 Hz, J^4 = 1.7 Hz, 1H, H-5), 7.77 (d, J^3 = 8.7 Hz, 2H, H-2'), 8.32 (s, 1H, H-4); ^{13}C-NMR (75.48 MHz, d^6-DMSO) δ: 20.35, 20.83, 120.69, 121.71, 124.55, 125.15, 126.14, 126.53, 129.75, 131.86, 136.75, 140.48, 144.65, 150.80, 158.71, 168.38, 169.12.

8-hydroxy-3-(4-hydroxyphenyl)coumarin (5b) [4]

Yield 85%; m.p. 247-249 °C (lit. [82] 249-250 °C); ^{1}H-NMR (300 MHz, d^6-DMSO) δ: 6.84 (d, J^3 = 8.7 Hz, 2H, H-3'), 7.04-7.17 (m, 3H, H-5, H-6, H7), 7.60 (d, J^3 = 8.7 Hz, 2H, H-2'), 8.06 (s, 1H, H-4), 9.70 (s, 1H, -OH), 10.14 (s, 1H, -OH); ^{13}C-NMR (75.48 MHz, d^6-DMSO) δ: 115.02, 117.48, 118.31, 120.62, 124.41, 125.29, 126.55, 129.79, 138.89, 141.33, 144.22, 157.91, 159.80.
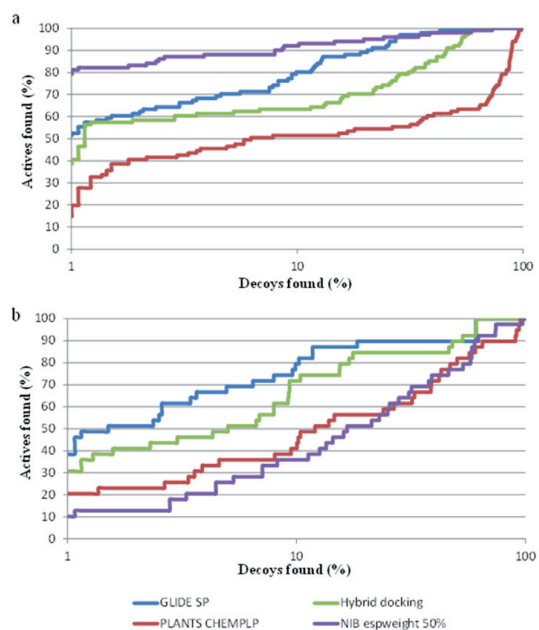
Supplementary Fig. S2 The receiver operating characteristics (ROC) curves of different docking approaches and NIB method on logarithmic scale. ROC curves for GLIDE SP, PLANTS CHEMPLP, hybrid docking models and NIB are shown for both (**a**) DS1 and (**b**) DS2. Figure shows the excellent early enrichments in DS1 (panel **a**), especially with NIB method (purple line).
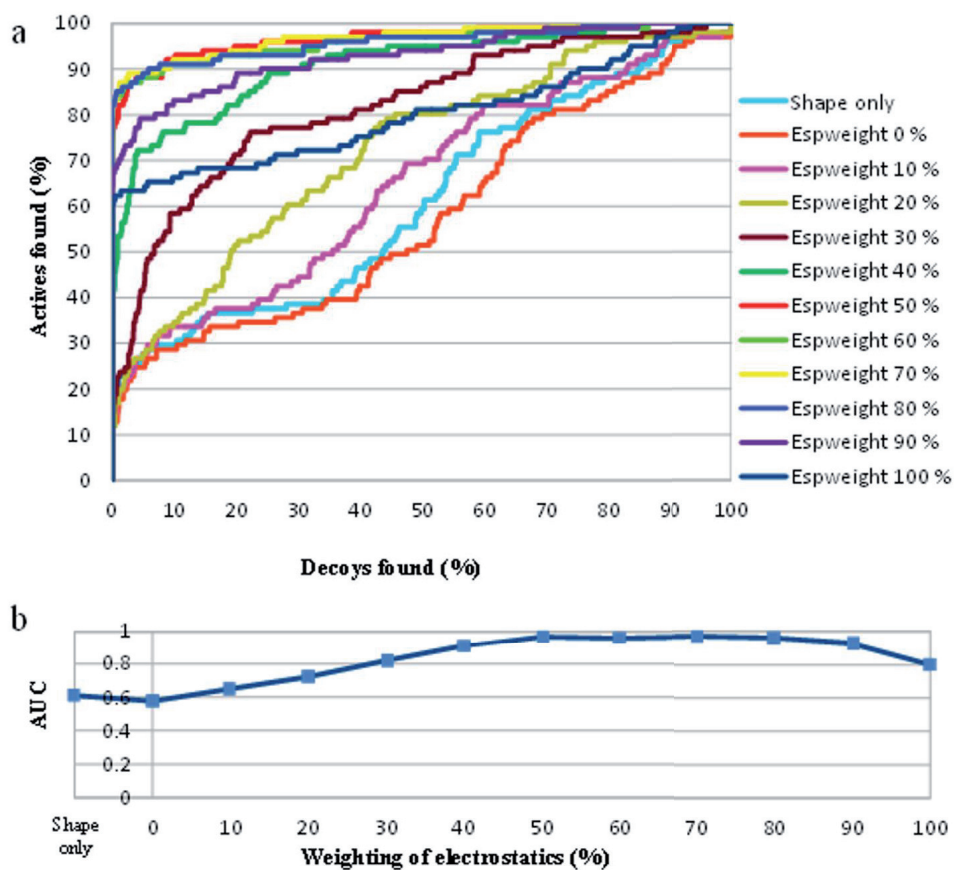
Supplementary Fig. S3 Optimization of the NIB model 1 with varying weighting of electrostatic potentials and shape. ROC curves for NIB model 1 (**a**) show that the best result is achieved using electrostatic weighting of 50 to 80 %. The AUC fluctuates along with the change in electrostatic weighting (**b**) indicating the huge importance of electrostatics in effective distinguishing of active molecules from molecular libraries.

References

1.   (a) Garazd, M. M.; Garazd, Y. L.; Ogorodniichuk, A. S.; Khilya, V. P., Modified coumarins. 29. Synthesis of structural analogs of natural 6-arylfuro[3,2-g]chromen-7-ones. *Chem Nat Compd+* **2009,** *45* (2), 158-163; (b) Kirkiacharian, S.; Lormier, A. T.; Resche-Rigon, M.; Bouchoux, F.; Cerede, E., [Synthesis and binding affinity of 3-aryl-7-hydroxycoumarins to human alpha and beta estrogen receptors]. *Ann Pharm Fr* **2003,** *61* (1), 51-6.

2.   Yang, J.; Liu, G. Y.; Dai, F.; Cao, X. Y.; Kang, Y. F.; Hu, L. M.; Tang, J. J.; Li, X. Z.; Li, Y.; Jin, X. L.; Zhou, B., Synthesis and biological evaluation of hydroxylated 3-phenylcoumarins as antioxidants and antiproliferative agents. *Bioorg Med Chem Lett* **2011,** *21* (21), 6420-5.

3.   (a) Belluti, F.; Uliassi, E.; Veronesi, G.; Bergamini, C.; Kaiser, M.; Brun, R.; Viola, A.; Fato, R.; Michels, P. A. M.; Krauth-Siegel, R. L.; Cavalli, A.; Bolognesi, M. L., Toward the Development of Dual-Targeted Glyceraldehyde-3-phosphate Dehydrogenase/Trypanothione Reductase Inhibitors against Trypanosoma brucei and Trypanosoma cruzi. *Chemmedchem* **2014,** *9* (2), 371-382; (b) Fais, A.; Corda, M.; Era, B.; Fadda, M. B.; Matos, M. J.; Quezada, E.; Santana, L.; Picciau, C.; Podda, G.; Delogu, G., Tyrosinase Inhibitor Activity of Coumarin-Resveratrol Hybrids. *Molecules* **2009,** *14* (7), 2514-2520.

4.   (a) Buuhoi, N. P.; Saintruf, G.; Lobert, B., Oxygen Heterocycles .14. Hydroxylated 3-Aryl- and 3-Pyridyl-Coumarins. *J Chem Soc C* **1969,** (16), 2069-&; (b) Matos, M. J.; Santana, L.; Uriarte, E.; Delogu, G.; Corda, M.; Fadda, M. B.; Era, B.; Fais, A., New halogenated phenylcoumarins as tyrosinase inhibitors. *Bioorg Med Chem Lett* **2011,** *21* (11), 3342-3345.

III

BLOCKING OESTRADIOL SYNTHESIS PATHWAYS WITH POTENT AND SELECTIVE COUMARIN DERIVATIVES

by

Sanna Niinivehmas, Pekka A. Postila, Sanna Rauhamäki, Elangovan Manivannan, Sami Kortet, Mira Ahinko, Pasi Huuskonen, Niina Nyberg, Pasi Koskimies, Sakari Lätti, Elina Multamäki, Risto O. Juvonen, Hannu Raunio, Markku Pasanen, Juhani Huuskonen & Olli T. Pentikäinen 2018

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

Blocking oestradiol synthesis pathways with potent and selective coumarin derivatives

Sanna Niinivehmas[a]*, Pekka A. Postila[a]* (iD), Sanna Rauhamäki[a] (iD), Elangovan Manivannan[a,b], Sami Kortet[a,c], Mira Ahinko[a], Pasi Huuskonen[d], Niina Nyberg[d], Pasi Koskimies[e], Sakari Lätti[a], Elina Multamäki[a], Risto O. Juvonen[d], Hannu Raunio[d], Markku Pasanen[d], Juhani Huuskonen[c] and Olli T. Pentikäinen[a,f] (iD)

[a]Department of Biological and Environmental Science and Nanoscience Center, University of Jyvaskyla, Jyvaskyla, Finland; [b]School of Pharmacy, Devi Ahilya University, Indore, India; [c]Department of Chemistry and Nanoscience Center, University of Jyvaskyla, Jyvaskyla, Finland; [d]School of Pharmacy, University of Eastern Finland, Kuopio, Finland; [e]Forendo Pharma Ltd, Turku, Finland; [f]Institute of Biomedicine, University of Turku, Turku, Finland

ABSTRACT

A comprehensive set of 3-phenylcoumarin analogues with polar substituents was synthesised for blocking oestradiol synthesis by 17-β-hydroxysteroid dehydrogenase 1 (HSD1) in the latter part of the sulphatase pathway. Five analogues produced $\geq$62% HSD1 inhibition at 5 μM and, furthermore, three of them produced $\geq$68% inhibition at 1 μM. A docking-based structure-activity relationship analysis was done to determine the molecular basis of the inhibition and the cross-reactivity of the analogues was tested against oestrogen receptor, aromatase, cytochrome P450 1A2, and monoamine oxidases. Most of the analogues are only modestly active with 17-β-hydroxysteroid dehydrogenase 2 – a requirement for lowering effective oestradiol levels *in vivo*. Moreover, the analysis led to the synthesis and discovery of 3-imidazolecoumarin as a potent aromatase inhibitor. In short, coumarin core can be tailored with specific ring and polar moiety substitutions to block either the sulphatase pathway or the aromatase pathway for treating breast cancer and endometriosis.

Introduction

Despite the recent advances made in early tumour detection, clinical treatments and avoidance of menopausal hormone therapies, breast cancer continues to be the most common invasive cancer, and a second leading cause of cancer death for women[1]. Therefore, potent and selective pharmaceutical agents are actively sought to supplement and/or replace the often-invasive treatments and to lower the medical costs for all breast cancer patients.

A clear majority of breast cancer tumours are oestrogen receptor (ER) positive. The tumour growth is linked to high ER numbers and/or their increased activity due to high 17-β-oestradiol (E_2) levels. Hence, the existing drugs generally aim to block the ER function in breast tissue or limit its function indirectly by lowering the E_2 production. The aromatase pathway produces E_2 from androgen hormones whereas the sulphatase pathway converts oestrone sulphate (E_1S) into oestrone (E_1) and ultimately to E_2. Although the aromatase pathway (active in local E_2 production) is in a lesser role with most breast cancers[2], widely used drugs, such as anastrozole focus on blocking it instead of the more prominent sulphatase pathway.

17-β-hydroxysteroid dehydrogenase 1 (HSD1 or 17-β-HSD1; Figure 1(A)) has a crucial role in the final steps of E_2 biosynthesis via the sulphatase pathway. HSD1 homodimer reduces the C17-keto group of E_1 by acquiring a proton (H^+) from the cofactor nicotinamide adenine dinucleotide phosphate (NADPH) to produce

E_2 (Figure 1(A,B)). In contrary, 17-β-hydroxysteroid dehydrogenase 2 (HSD2 or 17-β-HSD2) promotes the oxidation of the C17-hydroxyl group on E_2 by donating H^+ to the cofactor to produce E_1. HSD1 overexpression is a strong signal for breast cancer – present in ~50% of breast tumours – and, furthermore, HSD2 is known to have an inhibitory effect in the breast tumourigenesis[3,4]. HSD1 is also linked to other cancer types, such as gastric[5] and cervical cancer[6], and, additionally, in endometriosis elevated E_2 production is promoted by increased HSD1 and, inversely, lowered HSD2 expression[7].

A vast number of steroidal ([8–10]; e.g. E2B in Figure 1(C)) and non-steroidal (see e.g.[11–13]) compounds are known to inhibit the HSD1 activity, but none of these promising leads has passed clinical trials so far. There are also several X-ray crystal structures of HSD1 in both ligand-free, substrate-, and inhibitor-bound states to facilitate rational structure-based drug discovery. Here, 3-phenylcoumarin (or 3-arylcoumarin) is shown to be a suitable non-steroidal scaffold for building small-molecule inhibitors targeting HSD1 (Figure 2; Table 1).
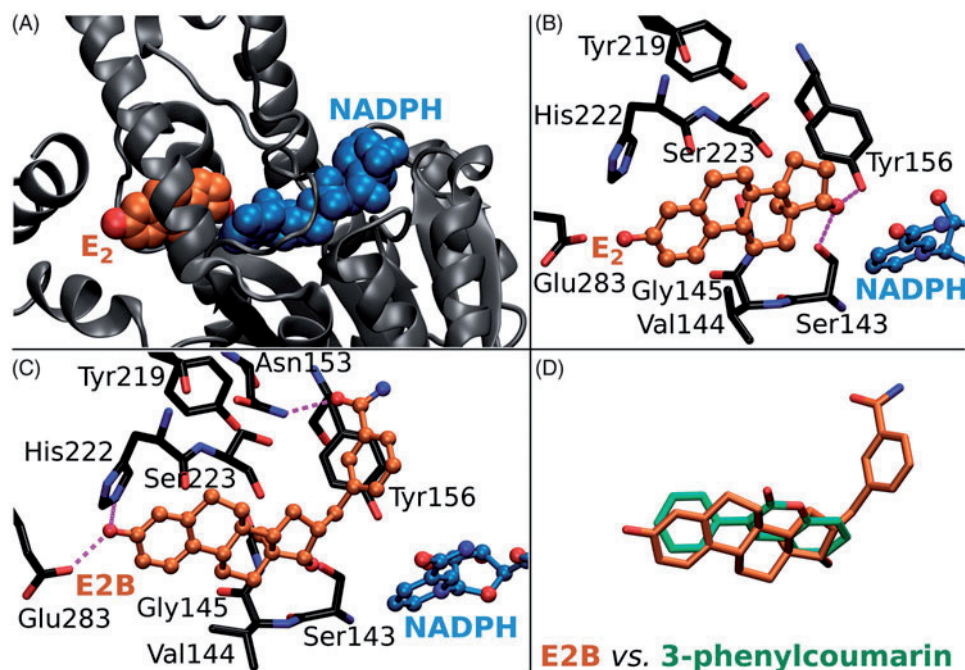
Altogether, nine 3-phenylcoumarin analogues with varying coumarin and 3-phenyl ring substituents (R1–R6 positions; Figure 2) were synthesised (Table 1). Five of the analogues produced $\geq$62% HSD1 inhibition at 5 μM and, furthermore, three of them elicited $\geq$68% inhibition even at 1 μM (estimated $pIC_{50} \geq 6.2$). The docking-based structure-activity relationship (SAR) analysis indicates that the

Figure 1. The ligand binding at the active site of 17-β-hydroxysteroid dehydrogenase 1. (A) Oestradiol (E$_2$; orange backbone) and oxidised cofactor nicotinamide adenine dinucleotide phosphate (NADP; blue) are shown as CPK models in complex with the HSD1 structure (grey cartoon; PDB: 1A27). (B) The H-bonding between E$_2$ (ball-and-stick models with orange backbone) and the residues lining the active site (stick model with black backbone) are shown with magenta dotted lines. The substrate oestrone (E$_1$) acquires a proton (or H$^+$) from NADPH, the reduced form of the cofactor, via the hydroxyl group of Tyr156 (E$_2$ + NADP$^+$ + $\rightleftharpoons$ E$_1$ + NADPH), which is H-bonding with the 17-keto group of the reaction product E$_2$. (C) Inhibitor E2B (ball-and-stick model with orange backbone; PDB: 3HB5)[25] binding at the HSD1 active site blocks E$_2$ binding (B vs. C). (D) The 3-phenyl and coumarin rings of the docked analogues (stick model with green backbone) align in a roughly similar manner inside the active site as the steroid ring of E2B (stick model with orange backbone).
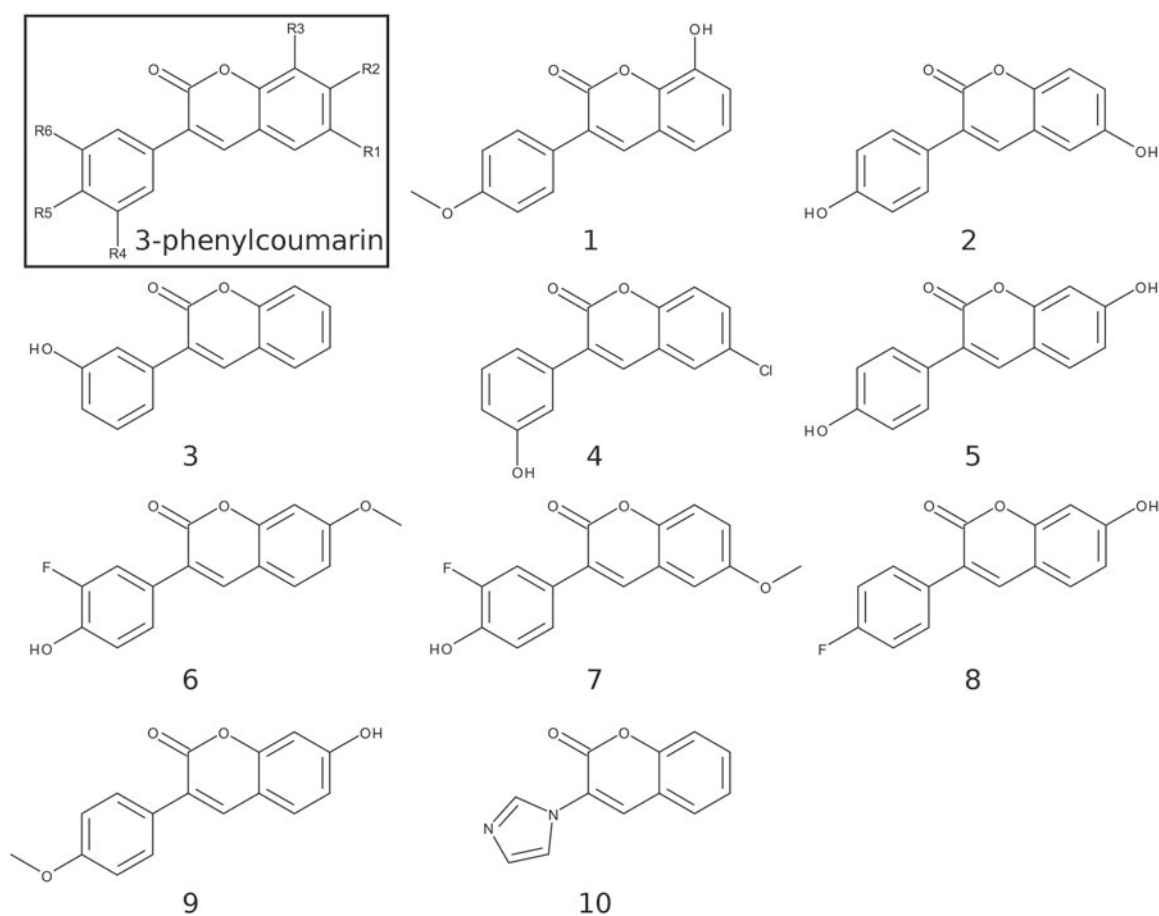
Figure 2. 2D structures of the coumarin derivatives. The 3-phenylcoumarin analogues **1–7** produce HSD1 inhibition at a varying degree, but **8** and **9** were found to be inactive (Table 1). Compound **10** or 3-imidazolecoumarin inhibit aromatase instead of HSD1.

Table 1. Inhibitory profiles of the 3-substituted coumarin compounds.

#	HSD1 inhibition at 100 nM (%)[a]	HSD1 inhibition at 1 µM[b] (%)	HSD1 inhibition at 5 µM (%)	HSD1 inhibition ~pIC$_{50}$	HSD2 inhibition at 1 µM (%)	MAO-A inhibition at 100 µM (%)	MAO-B inhibition at 10 µM (%)	MAO-B inhibition pIC$_{50}$	CYP1A2 inhibition ~pIC$_{50}$	ER binding pIC$_{50}$	Aromatase inhibition ~pIC$_{50}$
1	16	68	88	6.3±0.23	27	4	48	N/A	4.07	N/B	N/I
2	26	69	83	6.2±0.11	7	−6	16	N/A	4.48	N/B	N/I
3	7	47	76	5.9±0.04	42	4	50	N/A	5.15	N/B	N/I
4	47	84	95	6.8±0.05	16	N/A	64	N/A	6.89	N/B	N/I
5	6	18	9	N/I	37	−1	53	N/A	4.66	5.50±0.13	N/I
6	14	23	27	5.4±0.14	31	−1	101	6.39±0.06	4.52	6.15±0.11	N/I
7	16	39	62	5.8±0.12	13	−11	95	6.07±0.04	5.82	N/B	N/I
8	1	1	N/A	N/I	N/A	13	26	N/A	4.89	6.10±0.07	N/I
9	0	3	N/A	N/I	N/A	25	86	6.00±0.03	3.77	5.90±0.10	N/I
10	1	1	N/A	N/I	N/A	N/A	11	N/A	6.94	N/A	7.11

N/I: no inhibition; N/B: no binding; N/A: not available.
[a]Average of two measurements, except for compounds **8** and **9**.
[b]Average of three measurements, except for compounds **8** and **9**. Controls: HSD1 = compound No. 21 (73.1–79.7% 10 nM or 91.3–95.3% 100 nM;[10]); HSD2 = in-house reference compound HM2001 = 3–(4-Chlorophenoxy)-5,7-dihydroxychromen-4-one (62.3% 100 nM or 100.0% 1 µM); MAO-A = clorgyline (101%); MAO-B = pargyline (103%; pIC$_{50}$ = 7.22 ± 0.07); CYP1A2 = fluvoxamine (IC$_{50}$ = 0.1 µM); ER: oestradiol (IC$_{50}$ = 5.7 nM (kit reference)/0.26 ± 0.09 nM or pIC$_{50}$ = 9.58 ± 0.12 (measured); aromatase = finrozole (IC$_{50}$ = 0.2–0.5 µM).

potent analogues mimic steroid binding (Figure 1(D)). A cross-reactivity profile, covering HSD2, monoamine oxidases A (MAO-A) and B (MAO-B), ER, cytochrome P450 1A2 (CYP1A2), and aromatase (or CYP19A1), were built for each analogue. Importantly, the substitution of the 3-phenyl ring with the 3-imidazole ring in the coumarin core, assures strong and selective aromatase inhibition.

In short, the coumarin-based compounds have potential for lowering E$_2$ levels needed in battle against diseases, such as breast cancer or endometriosis by blocking either the aromatase pathway or the sulphatase pathway.

Methods

Chemical procedure

All reactions were carried out using commercial materials and reagents without further purification unless otherwise noted. Reaction mixtures were heated by the CEM Discovery microwave apparatus. All reactions were monitored by thin layer chromatography (TLC) on silica gel plates. ^{1}H NMR and ^{13}C NMR data were recorded on a Bruker Avance 400 MHz spectrometer or Bruker Avance III 300 MHz spectrometer (Bruker, Billerica, MA). Chemical shifts are expressed in parts per million values (ppm) and are designated as singlet (s), broad singlet (br s), doublet (d), double doublet (dd), double double doublet (ddd), and triplet (t). Coupling constants (J) are expressed as values in hertz (Hz). The HRMS mass spectra were recorded using Micromass LCT ESI-TOF equipment (Waters Corporation, Milford, MA). Elemental analyses were done with Elementar Vario EL III elemental analyser (Elementar-Straße 1, Langenselbold, Germany). The 3-phenylcoumarin analogues were synthesised using Perkin–Oglialor condensation reaction. The method was developed from the earlier published procedures and transferred to microwave reactor.

Experimental data for 7-hydroxy-3–(4-fluorophenyl)-2H-chromen-2-one (**5**; Figure 2), 7-hydroxy-3–(4-methoxyphenyl)-2H-chromen-2-one (**8**; Figure 2) and 7-hydroxy-3–(4-hydroxyphenyl)-2H-chromen-2-one (**9**; Figure 2) have been published[14]. However, the synthesis steps are detailed below for other derivatives studied here (**1**, **2**, **3**, **4**, **6**, **7** and **10** Figure 2; Scheme 1). Of these **1–4** have also been synthesised earlier by others prior to this study[15–18]. 2; Scheme 1).

A typical procedure (Scheme 1): A mixture of salicylaldehyde derivative (2 mmol) and phenylacetic acid derivative (2.1 mmol), acetic acid anhydride (0.6 ml), and triethylamine (0.36 ml) were placed in a microwave reactor tube and this mixture was heated at 100–170 °C with microwave apparatus for 10–20 min. After cooling, 2 ml of 10% NaHCO$_3$ solution was added and the precipitate was filtered, dried, and recrystallised from ETOH/H$_2$O or acetone/H$_2$O mixture. The acetyl group(s) were removed by treating the compound with MeOH/NaOH(aq) solution for 30–60 min at r.t. The solution was acidified with HCl(aq,) and the precipitate was filtered and recrystallised if needed.

Based on the elemental analysis and/or ^{1}H-NMR the purity of compounds was >95%.

8-hydroxy-3–(4-methoxyphenyl)-2H-chromen-2-one (1)[15]. In the first step 8-acetoxy-3–(4-methoxyphenyl)-2H-chromen-2-one was obtained. Yield 85%; ^{1}H-NMR (400 MHz, d[6]-DMSO) δ: 2.40 (s, 3H, CH$_3$C(O)O-Ph), 3.80 (s, 3H, CH$_3$O-Ph), 7.02 (d, 2H, J^3 = 8.1 Hz, H-3′, H-5′), 7.37 (t, 1H, J^3 = 7.6 Hz, H-6), 7.43 (d, J^3 = 7.7 Hz, 1H, H-7), 7.64–7.69 (m, 3H, H-5, H-2′, H-6′), 8.11 (s, 1H, H-4); ^{13}C-NMR (100.6 MHz, d[6]-DMSO) δ: 20.33, 55.20, 113.70, 120.85, 124.41, 124.65, 125.87, 126.50, 126.80, 129.82, 136.70, 138.88, 144.38, 158.83, 159.73 and 168.38. HRMS(ESI): calcd for C$_{18}$H$_{14}$O$_5$Na$_1$ [M + Na]$^+$: 333.07389, found 333.07580. Elemental anal. for
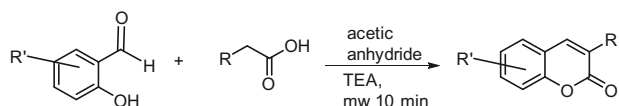
Scheme 1. The synthesis of 3-phenylcoumarin analogues and 3-imidazolecoumarin.

$C_{18}H_{14}O_5$, calc. C% 69.67, H% 4.55, found C% 69.53, H% 4.55. In the second step, 8-hydroxy-3–(4-methoxyphenyl)-2H-chromen-2one was obtained. Yield 81%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 3.80 (s, 6H (CH$_3$O−), 7.01 (d, $J^3 = 8.9$ Hz, 2H, H-3′,H-5′), 7.08 (dd, 1H, $J^3 = 7.0$ Hz, $J^4 = 2.6$ Hz, H-7), 7.12–7.18 (m, 2H, H-5, H-6), 7.70 (d, 2H $J^3 = 8.9$ Hz, H-2′,H-6′), 8.11 (s, 1H, H-4), 10.19 (s, 1H, Ph-OH). ^{13}C-NMR (100.6 MHz, d^6-DMSO) δ: 55.21, 113.64, 117.64 118.39, 120.55, 124.45, 126.22, 126.91, 129.79, 139.52, 141.42, 144.26, 159.54 and159.76. HRMS(ESI)): calcd for $C_{16}H_{12}O_4Na_1$ [M + Na]$^+$: 291.06333, found 291.06180. Elemental anal. for $C_{16}H_{12}O_4$, calc. C% 71.26, H% 4.51, found C% 71.64, H% 4.51.

6-hydroxy-3–(4-hydroxyphenyl)-2H-chromen-2-one (2)[19]. In the first step 4–(6-acetoxy-2-oxo-2H-chromen-3-yl)phenyl acetate was obtained. Yield 90%; ^{1}H-NMR (300 MHz, d^6-DMSO) δ: 2.30 (s, 3H, CH$_3$CO(O)-Ph), 2.31 (s, 3H, CH$_3$CO(O)-Ph), 7.23 (d, 2H, $J^3 = 8.8$ Hz, H-2′, H-6′), 7.40 (dd, $J^3 = 8.9$ Hz, $J^4 = 2.7$ Hz, 1H, H-7), 7.49 (d, 1H, $J^3 = 8.9$ Hz, H-8), 7.55 (d, 1H, $J^4 = 2.6$ Hz, H-5), 7.76 (d, 2H, $J^3 = 8.8$ Hz, H-3′, H-5′), 8.24 (s, 1H, H-4); ^{13}C-NMR (75.5 MHz, d^6-DMSO) δ: 20.73, 20.82, 116.97, 119.90, 120.77, 121.67, 125.48, 126.67, 129.74, 131.95, 139.84, 146.43, 150.43, 150.76, 159.51, 169.10 and 169.22. In the second step, 6-hydroxy-3–(4-hydroxyphenyl)-2H-chromen-2-one was obtained. Yield 85%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 6.83 (d, 2H, $J^3 = 8.8$ Hz, H-3′, H-5′), 6.99 (dd, 1H, $J^3 = 8.8$ Hz, $J^4 = 2.9$ Hz, H-7), 7.06 (d, 1H, $J^4 = 2.8$ Hz, H-5), 7.24 (d, 1H, $J^3 = 8.9$ Hz, H-8), 7.57 (d, 2H, $J^3 = 8.7$ Hz, H-2′, H6′), 8.04 (s, 1H, H-4); ^{13}C-NMR (75.5 MHz, d^6-DMSO) δ: 112.29, 115.00,116.59, 119.15, 120.24, 125.40, 126.71 129.86, 138.51, 146.03, 153.77, 157.90 and 160.13. HRMS(ESI)): calcd for $C_{16}H_{11}F_1O_4Na_1$ [M + Na]$^+$: 277.0477, found 277.0461.

3–(3-hydroxyphenyl)-2H-chromen-2-one (3)[20]. In the first step, 3–(2-oxo-2H-chromen-3-yl)phenyl acetate was obtained. Yield 87%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 2.30 (s, 3H, CH$_3$C(O)O-Ph), 7.20 (ddd, 1H, $J^3 = 9.0$ Hz, $J^4 = 2.2$ Hz, $J^{4'} = 2.3$ Hz, H-6′), 7.39 (t, 1H, $J^3 = 7.6$ Hz, H-5′), 7.44 (d(broad), 1H, $J^3 = 8.3$ Hz, H-4′), 7.49–7.53 (m, 2H, H-6, H-2′), 7.62–7.66 (m, 2H, H-7, H-8) 7.79 (dd, 1H, $J^3 = 8.7$ Hz, $J^4 = 1.5$ Hz, H-5), 8.32 (s, 1H, H-4); ^{13}C-NMR (100 MHz, d^6-DMSO) δ: 20.86, 115.90, 119.38, 121.81, 122.17, 124.68, 125.69, 125.90, 128.81, 129.31, 131.98, 135.99, 141.13, 150.30, 153.02, 159.51 and 169.23. In the second step, 3–(3-hydroxyphenyl)-2H-chromen-2-one was obtained. Yield 74%; ^{1}H-NMR (300 MHz, d^6-DMSO) δ: 6.83 (ddd, 1H, $J^3 = 8.1$ Hz, $J^4 = 2.2$ Hz, $J^{4'} = 2.4$ Hz, H-4′), 7.11–7.18 (m, 2H, H-2′, H-6′), 7.26 (t, 1H, $J^3 = 7.9$, H-5′), 7.37 (ddd, 1H, $J^3 = 7.6$ Hz, $J^4 = 1.1$ Hz, $J^{4'} = 1.1$ Hz, H-6), 7.42 (d, $J^3 = 8.3$ Hz, H-8), 7.61 (ddd, $J^3 = 7.3$ Hz, $J^4 = 1.6$ Hz, $J^{4'} = 2.6$ Hz H-7), 7.83 (dd, 1H, $J^3 = 8.7$ Hz, $J^4 = 1.5$ Hz, H-5), 8.20 (s, 1H, H-4), 9.54 (s, 1H,Ph-OH); ^{13}C-NMR (75.5 MHz, d^6-DMSO) δ: 115.45, 115.59, 115.76, 119.13, 119.43, 124.50, 126.86, 128.60, 129.20, 131.60, 135.79, 140.32, 152.87, 157.06, and 159.54. HRMS (ESI): Calcd for $C_{15}H_{10}O_4Na_1$ [M + Na]$^+$: 261.05276, found 261.04980.

6-chloro-3–(3-hydroxyphenyl)-2H-chromen-2-one (4)[21]. In the first step, 3–(6-chloro-2-oxo-2H-chromen-3-yl)phenyl acetate was obtained. Yield 85%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 2.30 (s, 3H, CH$_3$C(O)O-Ph), 7.22 (ddd, 1H, $J^3 = 8.0$ Hz, $J^4 = 2.2$ Hz, $J^{4'} = 2.3$ Hz, H-6′), 7.48–7.52 (m, 3H, H-8, H-2′, H-5′), 7.62 (m, 1H, H-4′), 7.67 (dd, 1H, $J^3 = 8.9$ Hz, $J^4 = 2.6$ Hz, H-7), 7.88 (d, 1H, $J^4 = 2.6$ Hz, H-5), 8.27 (s, 1H, H-4).); ^{13}C-NMR (100 MHz, d^6-DMSO) δ: 20.85, 117.96, 120.78, 121.85, 122.47, 125.93, 126.88, 127.70,

128.31, 129.41, 131.45, 135.67, 139.82, 150.30, 151.66, 159.10 and169.22. In the second step, 6-chloro-3–(3-hydroxyphenyl)-2H-chromen-2-one was obtained. Yield 80%; %; ^{1}H-NMR (400 MHz, d^6-DMSO δ:), 6.84 (ddd, 1H, $J^3 = 8.0$ Hz, $J^4 = 2.4$ Hz, $J^{4'} = 2.3$ Hz, H-6′), 7.10–7.15 (m, 2H), 7.27 (t, 1H, $J^3 = 7.9$ Hz, H-5′), 7.47 (d, 1H, $J^3 = 8.9$ Hz, H-8), 7.65 (dd, 1H, $J^3 = 8.3$ Hz, $J^4 = 2.6$ Hz, H-7), 7.90 (d, 1H, $J^4 = 2.5$ Hz, H-5), 8.17 (s, 1H, H-4), 9.57 (s, 1H, Ph-OH); ^{13}C-NMR (75.5 MHz, d^6-DMSO) δ: 115.43, 115.85, 117.79, 119.13, 120.84, 127.53, 127.99, 128.18, 129.26, 131.09, 135.44, 139.03, 151.50, 157.08 and 159.11. HRMS (ESI): Calcd for $C_{15}H_9Cl_1O_3Na_1$ [M + Na]$^+$: 295.01379, found 295.01380.

3–(3-fluoro-4-hydroxyphenyl)-7-methoxy-2H-chromen-2-one (6). In the first step, 2-fluoro-4–(7-methoxy-2-oxo-2H-chromen-3-yl)phenyl acetate was obtained. Yield 75%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 2.35 (s, 3H, CH$_3$C(O)O-Ph), 3.88 (s, 3H, CH$_3$O-Ph), 6.99 (dd, 1H, $J^3 = 8.6$ Hz, $J^4 = 2.4$ Hz, H-6), 7.05 (d, 1H, $J^4 = 2.4$ Hz, H-8), 7.37 (t, 1H, $J = 8.3$ Hz, H-6′), 7.62 (d, $J = 8.5$ Hz, 1H, H-5′), 7.68 (d, $J = 8.6$ Hz, 1H, H-5), 7.74 (dd, 1H, $J^{H-F} = 12.1$ Hz, $J^4 = 2.0$ Hz, H-3′), 8.31 (s, 1H, H-4); ^{13}C-NMR (100 MHz, d^6-DMSO) δ: 20.19, 55.97, 100.25, 112.79, 116.35 (d, $J^{C-F} = 20$ Hz), 121.02, (d, $J^{C-F} = 1.9$ Hz), 123.83, 124.79 (d, $J^{C-F} = 3.2$ Hz), 129.86, 134.24 (d, $J^{C-F} = 7.7$ Hz), 137.20 (d, $J^{C-F} = 13.1$ Hz), 141.55, 153.00 ($J^{C-F} = 246.1$ Hz), 154.92, 159.65 and 162.69, 168.19. In the second step, 3–(3-fluoro-4-hydroxyphenyl)-7-methoxy-2H-chromen-2-one was obtained. Yield 70%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 3.87 (s, 3H, CH$_3$O-Ph), 6.96–7.03 (m, 3H, H-6, H-8, H-5′), 7.41 (d, 1H, $J = 8.4$, H-6′), 7.57 (dd, 1H, $J^{H-F} = 13.1$ Hz, $J^4 = 2.2$ Hz (H-H), 1H, H-2′), 7.66 (d, 1H, $J^3 = 8.4$ Hz, H-5), 8.18 (s, 1H, H-4), 10.09 (s, 1H, Ph-OH). ^{13}C-NMR (75.5 MHz, d^6-DMSO) δ: 55.91, 100.16, 112.61, 113.04, 115.95 (d, $J^{C-F} = 20$ Hz), 117.37 (d, $J^{C-F} = 3.3$ Hz), 121.78 ($J^{C-F} = 2.0$ Hz), 124.54 (d, $J^{C-F} = 3.0$ Hz), 126.08 (d, $J^{C-F} = 7.0$ Hz), 129.49, 139.62, 145.0 ($J^{C-F} = 13$ Hz), 150.46 (d, $J^{C-F} = 240$ Hz), 154.52, 159.87 and 162.19. HRMS (ESI): Calcd for $C_{16}H_{11}F_1O_4Na_1$ [M + Na]$^+$: 309.0539, found 309.0553.

3–(3-fluoro-4-hydroxyphenyl)-6-methoxy-2H-chromen-2-one (7). In the first step, 2-fluoro-4–(6-methoxy-2-oxo-2H-chromen-3-yl)phenyl acetate was obtained. Yield 66%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 2.33 (s, 3H, CH$_3$C(O)O-Ph), 3.82 (s, 3H (CH$_3$O-Ph), 7.23 (dd, 1H, $J^3 = 9.0$ Hz, $J^4 = 3.0$ Hz, H-7), 7.30 (d, 1H, $J^4 = 3.0$ Hz, H-5), 7.35 (d, 1H, $J^3 = 9.2$ Hz, H-8), 7.61 (d, 1H, $J^3 = 8.5$ Hz, H-5′), 7.75 (dd, 1H, $J^{H-F} = 12.0$ Hz, $J^4 = 1.7$ Hz (H-H), 1H, H-2′), 8.30 (s, 1H, H-4); ^{13}C-NMR (100.6 MHz, d^6-DMSO) δ: 20.22, 55.69, 110.83, 116.67, 117.02, 119.66, 123.96, 125.10, 135.96, 141.18, 147.44, 151.78, 154.23, 155.70, 159.53 and 168.21. In the second step, 3–(3-fluoro-4-hydroxyphenyl)-6-methoxy-2H-chromen-2-one was obtained. Yield 71%; ^{1}H-NMR (400 MHz, d^6-DMSO) δ: 3.81 (s, 3H (CH$_3$O-Ph), 7.02 (dd, 1H, $J^3 = 9.2$ Hz, H-6′), 7.18–7.28 (m, 1H, H-5, H-7), 7.35 (d, $J^3 = 9.0$ Hz, H-8), 7.42 (d, 1H, $J^3 = 8.4$ Hz, H-5′), 7.57 (dd, 1H, $J^{H-F} = 13.0$ Hz, $J^4 = 2.2$ Hz (H-H), 1H, H-2′), 8.17 (s, 1H, H-4), 10.19 (s, 1H, Ph-OH); ^{13}C-NMR (100.6 MHz, d^6-DMSO) δ: 55.66, 110.59, 116.67, 117.02, 119.66, 123.96, 125.10, 135.96, 141.18, 147.44, 151.78, 154.23, 155.70, 159.53 and 168.21. HRMS (ESI): Calcd for $C_{16}H_{11}F_1O_4Na_1$ [M + Na]$^+$: 309.0539, found 309.0521.

3–(1H-imidazol-1-yl)-2H-chromen-2-one (10). Yield: 39% light brown solid; $R_f = 0.18$ (EtOAc); ^{1}H-NMR (300 MHz, d^6-DMSO) δ: 7.10 (br s, 1H, H-4′), 7.44 (apparent td, $J^3 = 7.5$ Hz, $J^4 = 1.0$ Hz, 1H, H-6), 7.51 (d, $J^3 = 8.3$ Hz, 1H, H-8), 7.64–7.70 (m, two overlapping signals, 2H, H-7 and H-5′), 7.77 (dd, $J^3 = 7.7$ Hz, $J^4 = 1.5$ Hz, 1H, H-5), 8.16 (br s, 1H, H-2′), 8.34 (s, 1H, H-4). ^{13}C-NMR (75 MHz, d^6-DMSO) δ: 116.06 (**C**-H8), 118.51 (H5-**C**-C-H4), 119.57 (**C**-H5′), 123.37 (N-**C**-C=O), 125.09 (**C**-H6), 128.63 (**C**-H4′), 128.80 (**C**-H5), 131.87 (**C**-H7), 132.97 (**C**-H4), 137.12 (N-**C**(-H2′)=N), 151.78 (H8-C-**C**-O), 156.83 (**C**=O). IR (KBr): 1727, 1708, 1630, 1608, 1486, 1318, 1083 and 760. ESI-MS: m/z (rel. abund. %): calculated for (M + Na$^+$) = 235.0478, measured 235.0476, $\Delta = 0.2$ mDa. Elemental analysis for

$C_{12}H_8N_2O_2$: calc. C% 67.92, H% 3.80, N% 13.20, found C% 67.49, H% 3.72 and N% 13.13. Mp. 180–182 °C.

17-β-Hydroxysteroid dehydrogenase 1 and 2

The inhibition was determined by HPLC using recombinant human HSD1 and HSD2 proteins as described in a prior study[10]. In short, recombinant human HSD1 and HSD2 were produced in Sf9-insect cells. The assay was performed in a final volume of 0.2 ml buffer (20 mM KH_2PO_4, 1 mM EDTA, pH 7.4) containing 0.1 mg/ml protein, 1 mM cofactor (NADPH for HSD1, NAD for HSD2), 30 nM substrate oestrone or oestradiol, 800,000 cpm/ml of tritium labelled oestrone ([3H]-E1) or oestradiol ([3H]-E2), and inhibitors concentrations in the range of 0.1–5.0 mM. Triplicate samples were incubated for 25 min at the room temperature. After incubation, the reaction was stopped by addition of 20 ml 10% trichloroacetic acid per sample. After incubation the substrate and the product of enzymatic conversion [3H]-E1 and [3H]-E2 were separated and quantified by HPLC (Alliance 2790, Waters, Milford, MA) connected to an online counter (Packard Flow Scintillation Analyser; Perkin Elmer Inc., Waltham, CA). The ratio of [3H]-E1 converted to [3H]-E2 or vice versa determines the conversion percentage of the samples. Inhibition was measured in three concentrations (100 nM, 1, and 5 μM) in order to follow the progression of inhibition efficiencies. Inhibition efficiencies of the tested inhibitors were calculated by comparing the conversion percentages of the samples including inhibitors with those of conversion controls (without inhibitors). The pIC$_{50}$ average values and their standard errors were estimated from three measurements at 1 μM.

Aromatase

Aromatase (CYP19A1) activity was measured as described previously[22] by using human placental microsomes and 50 nM [3H]-androstenedione as a substrate and inhibitor concentrations in the range of 60–1000 nM. Aromatase activities were measured as released [3H]-H_2O in Optiphase Hisafe 2 scintillation liquid (Perkin Elmer, Waltham, MA) with a Wallac 1450 MicroBeta Trilux scintillation counter (Perkin Elmer, Waltham, MA). As a positive control for aromatase inhibition, 1 μM finrozole (generous gift from Olavi Pelkonen, University of Oulu, Finland) was used.

Monoamine oxidase A and B

The protein in addition to the reagents for the chromogenic solution (vanillic acid (4-hydroxy-3-methoxybenzoic acid, 97% purity), 4-aminoantipyrine (reagent grade), horseradish peroxidase, and the substrate tyramine hydrochloride (minimum 99% purity)) as well as the potassium phosphate buffering agents (potassium phosphate dibasic trihydrate (≥99% ReagentPlusTM) and potassium phosphate monobasic (minimum 98% purity, molecular biology tested)) were all purchased from Sigma-Aldrich (St. Louis, MO). The protocol of continuous spectrophotometric assay by Holt et al. was first used to determine the activity of the proteins[23]. The assay was performed in 0.2 M potassium phosphate buffer pH 7.6 on 94-well plates (NuncTM 96 F microwell plate without a lid, Nunc A/S, Roskilde, DK) with chromogenic solution containing 250 μM vanillic acid, 125 μM 4-aminoantipyrine and 2 U/ml horseradish peroxide in the total assay volume of 200 μl. The protein was first incubated for 30 min at 37 °C in the chromogenic solution and then the substrate tyramine was introduced at 0.5 mM final plate concentration completing the assay volume. The activity measurement using multilabel reader (VictorTM X4, 2030 Multilabel Reader, PerkinElmer, Waltham, MA) at A$_{490}$

immediately followed and the plates were read 300 times every 15 s using 1 s exposure time. The assay should produce absorbance change of ∼0.35[23]. The more active MAO-A produced over 0.5 change in absorbance reaching the assay maximum in 30 min with 25 μg of protein (enzymatic activity 5.25 units) per well while MAO-B produced the expected 0.35 change in absorbance with 50 μg of protein (enzymatic activity 3.2 units) per well and reached the assay maximum in 2 h. These protein concentrations were selected to be used to analyse the molecules 1–9. The analysis conditions followed the above-described assay protocol[23] and the activity of tested molecules was measured at 100 μM for MAO-A and at 10 μM for MAO-B. The analysis was performed as single point measurements and the signal was read by the same instrument at the expected assay maximum indicated by the activity measurements, at 30 min for MAO-A and at 2 h for MAO-B, respectively. Clorgyline was used as MAO-A and pargyline as MAO-B inhibitor control. Both of the control inhibitors provided 100% inhibition at the assay concentration of the test molecules. In addition, pIC$_{50}$ values were determined for MAO-B inhibition using duplicated dilution series and the pIC$_{50}$ value calculated for MAO-B inhibition by pargyline was 6.21. The observed activity was calculated as inhibition percentage (Table 1). The pIC$_{50}$ values were calculated with GraphPad Prism version 5.03 (GraphPad Software Inc., San Diego, CA).

Oestrogen receptor

The pIC$_{50}$ values of the molecules (Table 1) were measured using green PolarScreen™ ER Alpha Competitor Assay (Life Technologies, Carlsbad, CA) kit, following the protocol provided by the manufacturer as previously described[14]. The final concentration of the molecules ranged between 0.0007 and 10,000 nM in the prepared dilution series. The molecules were combined with 25 nM ERα and 4.5 nM fluormone in the assay buffer and placed on black low volume 384-well assay plate with NBS surface (Corning Inc., Corning, NY). After mixing the assay plate, it was incubated for 2 h at the room temperature. The fluorescence polarisation was then measured using excitation wave length 485 and emission wave length 535 with bandwidths of 25/20 nm on a 2104 EnVision® Multilabel Plate Reader which had EnVision Workstation version 1.7 (PerkinElmer, Waltham, MA).

Cytochrome P450 1A2

Inhibition of CYP1A2 activity was determined using commercial heterologously expressed human CYP1A2 enzyme (Corning Inc., Corning, NY) essentially as described previously[24].

Molecular docking

The small-molecule ligands (Figure 2), including their probable tautomeric states and 3D conformers, were built using LIGPREP, CONFGEN, and MACROMODEL modules in MAESTRO 2016–3 (Schrödinger, LLC, New York, NY, 2016) to match pH 7.4. The compounds were docked to the X-ray crystal structures of HSD1 (PDB: 3HB5[25]; Figures 3 and 4), aromatase (PDB: 3EQM[26]; Figure 6(C)), MAO-B (PDB: 2V61[27]; Figure 6(A)) and CYP1A2 (PDB: 2HI4[28]; Figure 6(B,C)) with the PANTHER protocol[29], where the ligand-binding site is described as a negative image, and the shape and electrostatic potentials of the Panther-models and ligand conformations are compared using SHAEP[30].
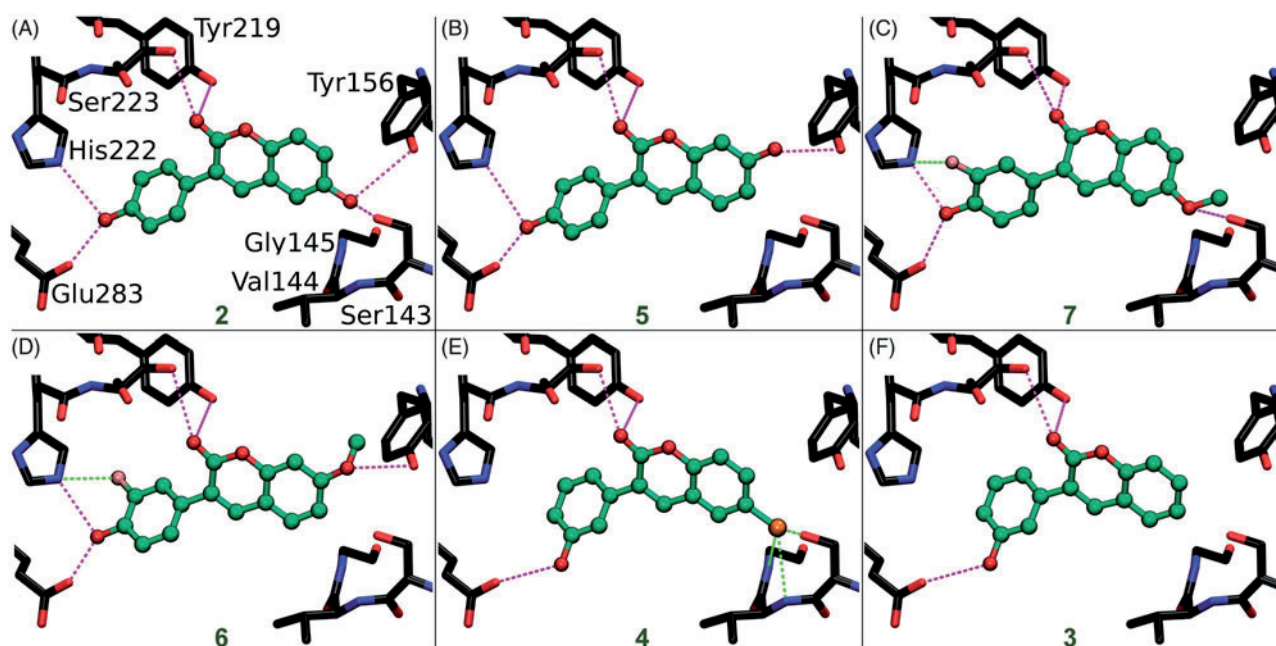
Figure 3. The canonical binding modes of 3-phenylcoumarin analogues inhibiting 17-β-hydroxysteroid dehydrogenase 1. The H-bonding (magenta dotted lines) and halogen bonding/favourable electrostatic (green dotted lines) interactions of (A) compounds **2**, (B) **5**, (C) **7**, (D) **6**, (E) **4**, and (F) **3** shown as suggested by docking. The active site residues of HSD1 enzyme (stick models with black backbone) bonding with the 3-phenylcoumarin analogues (ball-and-stick models with green backbone) are shown. The fluorine atom in **6** (D) and **7** (C) as well as the chlorine atom in **4** (E) are shown with pink and orange colour, respectively. Note that the His222 side chain is set epsilon protonated in order to facilitate H-bonding with the analogues. This is the opposite arrangement, if compared to the delta protonation of His222 suggested by the original E2B-bound HSD1 X-ray crystal structure (PDB: 3HB5)[25].
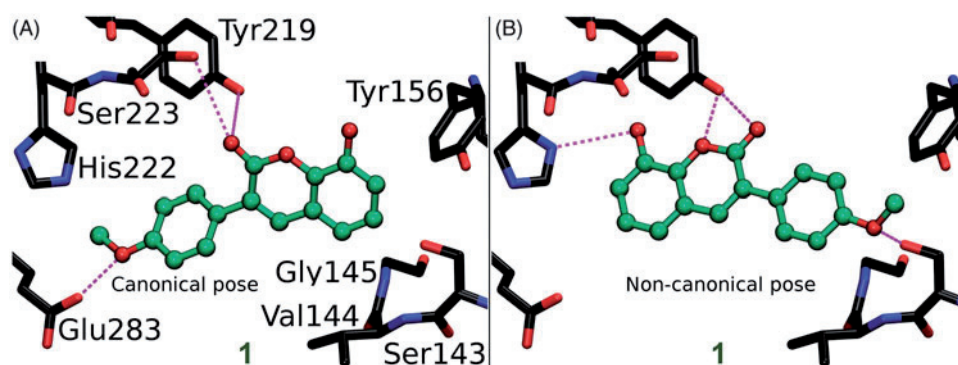
Figure 4. The canonical vs. non-canonical binding mode of compound **1**. (A) The "canonical" binding mode at the HSD1 active site, likely adopted by the other 3-phenylcoumarin analogues (Figure 3), is not suggested for compound **1** based on the docking-based SAR analysis; (B) instead, an alternative "non-canonical" pose is proposed for this potent inhibitor. Note that the His222 side chain is set delta protonated to facilitate H-bonding with the analogue's R3-hydroxyl. See Figure 3 for further details.

Figure preparation

Figures 2 and 5 showing 2D structures of the 3-phenylcoumarin scaffold and the analogues are drawn with BIOVIA Draw 2016 (Dassault Systèmes, San Diego, CA, 2016). Figures 1, 3, 4 and 6 are prepared using BODIL[31] and VMD 1.9.2[32].

Results and discussion

Computer-aided drug discovery

Whether the small-molecule design originates from automated virtual screening schemes, expert *de novo* work[33] or combination of the two, the computer-aided drug discovery (CADD) requires experimental verification[14,34]. This is achieved by pairing biochemical activity testing with, for example X-ray crystallographic studies[35], site-directed mutagenesis experiments[36,37], and/or "mutating" the lead compounds into diverse libraries of closely-related analogues

using organic synthesis[38–40]. The latter approach was applied here to demonstrate that 3-phenylcoumarin (Figure 2) is a suitable non-steroidal scaffold for building potent and selective HSD1 inhibitors.

Inhibitor design hypothesis

Based on a detailed analysis of the known inhibitors, 3-phenylcoumarin was chosen as a suitable scaffold for designing non-steroidal HSD1-specific inhibitors *de novo*. The analogue ring system alignment at the active site of HSD1 would mimic the hydrophobic packing of the steroid ring (e.g. inhibitor E2B; PDB: 3HB5 (25); Figure 1(D)). The coumarin ring would align in an orientation that allows its C2-carbonyl to form direct hydrogen bonds (or H-bonds) with the hydroxyl groups of Tyr219 (or Tyr219^{OH}) and/or Ser223 (or Ser223^{OH}; Figures 3 and 4). The coumarin ring could flip also sideways, if Arg258 side chain would rotate into the active site to interact with the C2-carbonyl. The probability of this
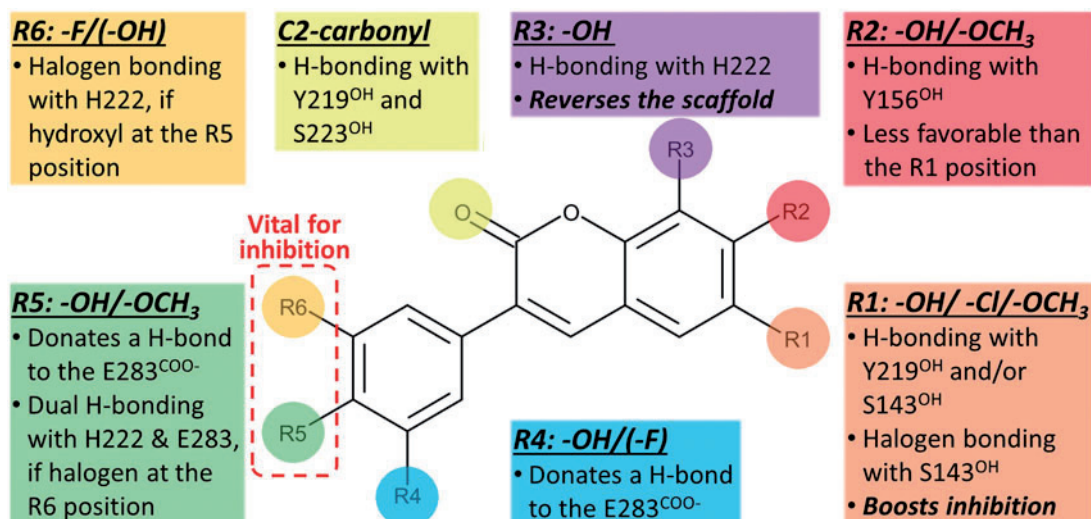
Figure 5. The docking-based structure-activity relationship analysis of the 3-phenylcoumarin analogues with 17-β-hydroxysteroid dehydrogenase 1.
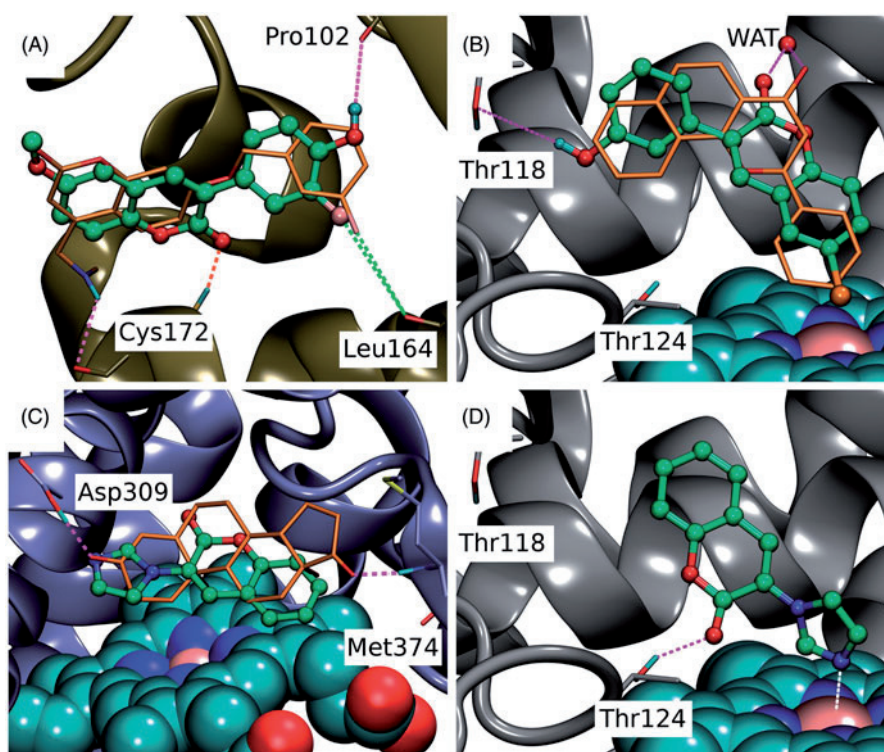

Figure 6. The binding of coumarin derivatives with aromatase, monoamine oxidase B and CYP1A2. (A) With the MAO-B (yellow cartoon), the docked pose of **6** demonstrates the analogous hydrophobic packing characteristic of the 3-phenylcoumarin analogues with the inhibitor C18 (stick model with orange backbone; PDB: 2V61)[27]. Notably, the R6-positioned polar group, fluorine in particular, improves the inhibition by forming a halogen bond with the Leu164°. (B) The docked pose of **4** (ball-and-stick model with green backbone) at the active site of CYP1A2 (grey cartoon) mimics α-naphthoflavone (stick model with orange backbone; PDB: 2HI4)[28]. Additionally, the R1-chlorine packs against the haeme and the C2-carbonyl and R4-hydroxyl, respectively, H-bond with crystal water (wat) and the Thr118^{OH}. (C) Based on docking, **10** (ball-and-stick model with green backbone) aligns similarly on top of the haeme (CPK model with cyan carbon atoms) in the active site of aromatase (magenta cartoon) as the androstenedione (stick model with orange backbone). Unlike the 3-phenylcoumarins the compound **10** has an acceptor group or the N3' in the imidazole ring capable of H-bonding with the neutral Asp309 (PDB: 3EQM)[26] and, thus, 3-imidazolecoumarin is a potent aromatase inhibitor. Alternatively, the N3' of **10** could be coordinated with the haeme (not shown). (D) The coumarin ring of **10** is aligned in a way that its C2-carbonyl accepts an H-bond from the Thr124^{OH}. Moreover, the deprotonated and electronegative N3' of 3-imidazole ring is likely coordinated with the positively charged iron in the haeme (CPK model with cyan carbon atoms). See Figure 3 for further details.

rotamer adjustment is difficult to estimate due to missing density data on the relevant X-ray crystal structure (PDB: 1EQU)[41]. Beyond this hypothesis, the plan was to establish and improve the 3-phenylcoumarin binding and HSD1 inhibition by introducing a number of polar (hydroxyl/methoxy/halogen) moieties for the 3-phenyl ring's R4–R6 and the coumarin ring's R1–R3 positions (Figure 2; Table 1).

Inhibition of 17-β-hydroxysteroid dehydrogenase 1 by the 3-phenylcoumarin analogues

The activity measurements (Table 1) indicated that the 3-phenylcoumarin is indeed a suitable scaffold for building HSD1 inhibitors. The dissimilarities in the inhibition levels between the analogues arise from their R1–R6 substituents (Figure 1).

Five analogues produced ≥62% HSD1 inhibition at 5 μM (Table 1). Moreover, analogues **1**, **2** and **4** produced ≥68% inhibition (estimated $pIC_{50} \geq 6.2$) at 1 μM. The most potent inhibitor **4** produced 47% inhibition even at 100 nM. Rest of the analogues elicited much weaker inhibition at 100 nM. The inhibition was consistently, regardless of the concentration, more modest for analogues **3**, **5**, **6** and **7** than for the three most potent analogues. In contrast, analogues **8** and **9** did not block the HSD1 (Table 1).

Scaffold hopping: 3-phenylcoumarin vs. steroid alignment

Due to the plasticity of the catalytic site, full understanding of the structural basis of the HSD1 inhibition or the selectivity is challenging. The ring systems of the 3-phenylcoumarin could mimic the steroid ring positioning in four different ways, if only the hydrophobic packing is considered. To address this issue, a specifically tailored docking protocol was utilised[29,42,43] for predicting how the analogues bind and elicit the inhibition (Figure 3). This docking-based SAR analysis point out how the R1–R6 moieties (Figure 2) affect the HSD1 binding (Table 1) and inhibition (Figures 3 and 4).

Coumarin (2H-chromen-2-one) contains a bicyclic structure of phenyl ring fused to a six-member ring with 1- and 2-positioned oxygen atom and carbonyl group, respectively (Figure 2). The 3-phenyl is tilted in relation to the coumarin ring as indicated by the small-molecule X-ray crystallography (CSD: QECNUJ)[44]. The binding of the 3-phenylcoumarin analogues is predicted to mimic closely the pose and hydrophobic packing of the E2B's steroid ring at the active site of HSD1 (Figure 1(D)). The ring positioning is likely highly similar or "canonical" for the HSD1 analogues (Figure 3), except for **1** (Figure 4). Moreover, both the Ser223[OH] and the Tyr219[OH] are predicted to H-bond with the coumarin's C2-carbonyl (Figures 3 and 4).

R1 position is important for strong 3-phenylcoumarin inhibition

A docking-based SAR analysis (Figure 5) explains the atomistic determinants of the HSD1 inhibition for each analogue.

The strong potency of **2** (Figure 2; Table 1) reflects its ability to form well-coordinated H-bonds between the proximal R1/R5-hydroxyl groups and the residues lining both ends of the binding site (Figure 3(A)). The R1-hydroxyl of the coumarin ring H-bonds with the main chain oxygen of Tyr156 (or Tyr156[O]) and the Ser143 side chain. Furthermore, the main chain nitrogen of Val144 (or Val144[N]), Gly145[N], and Cys186[O] are favourably positioned in relation to the analogue's R1-hydroxyl group. In turn, the R5-hydroxyl H-bonds with the carboxyl group of Glu283 (or Glu283[COO-]) and, reciprocally, accept an H-bond from the epsilon position of His222.

When the R1-hydroxyl of **2** is switched to the R2 position at the coumarin ring in **5** (Figure 2), the HSD1 inhibition lowers dramatically (Table 1). This highlights the importance of the R1 position for the 3-phenylcoumarin binding as the overall alignment of **2** and **5** is likely similar despite the switch (Figure 3(A,B)). Although the R2-hydroxyl is able to H-bond with the Tyr156[OH], it is evident that the R1-hydroxyl of **2** form stronger interactions with the close-by residues than the R2-hydroxyl (Figure 3(A,B)). The R5-hydroxyl of **5** assumes the same dual H-bonding role with Glu283 and His222 as the equivalent hydroxyl of **2**; assuring inhibition despite the R1/R2-hydroxyl switch (Table 1).

Replacing the R1-hydroxyl with a methoxy lowers the HSD1 inhibition considerably (Figure 5). This effect is apparent when **7** (Figure 2) is compared to **2** (Table 1). Although the R1-methoxy is H-bonding with the Ser143[OH] in the docked pose (Figure 3(C)), it

cannot coordinate as many or as strong interactions in this position as a hydroxyl (Figure 3(A,C)). However, the addition of R6-fluorine next to the hydroxyl offsets in part the negative effect of the R1 substitution. The fluorine is able to form a halogen bond with His222 (Figure 3(C)). In addition, the R5-hydroxyl of **7** function in the same dual H-bonding role with the side chains of Glu283 and His222 (compare to **2** and **5**; Figure 3(A–C)).

When the R1-methoxy of **7** is shifted to the R2 position in **6** (Figure 2), the inhibition is moderately reduced (Table 1). This effect is analogous to the weakening of inhibition seen in response to the R1/R2-hydroxyl switch between **2** and **5** (Figure 1(A,B)). The R5-hydroxyl of **6** H-bonds with both His222 and Glu283 and the adjacent R6-fluorine halogen bonds with His222 (Figure 3(D)). Despite the proximity of the R2-methoxy to several H-bond donors, such as the Tyr156[OH] and the Ser143[OH], the group cannot form as coordinated polar interactions as the R1-methoxy of **7** (Figure 3(C,D)).

The importance of the R1 position is highlighted with **4** (Figure 2) – the most potent HSD1 inhibitor of the analogues set (Table 1). Although the R4-hydroxyl of **4** donates an H-bond only to the Glu283[COO-] (Figure 3(E)), the inhibition is strong (Table 1). This is likely due to the hydroxyl/chlorine substitution at the R1 position (Figure 3(A,E); Table 1) allowing the R1-chlorine to halogen bond with the Ser143[OH] and potentially with the Tyr156[OH] (Figure 3(E)). Besides, the protons of the Val144[N] and the Gly145[N] cater to the halogen's negative charge. The inability of **4** to form H-bonds with both Glu283 and His222 is, therefore, likely offset by the analogue's ability to halogen bond (Table 1).

The relatively poor potency of **3** (Figure 2; Table 1) correlates with its limited ability to H-bond (Figure 2). Although **3** is almost identical to the most potent inhibitor **4**, it lacks the R1-chlorine (Figure 3(E,F)). The Glu283[COO-] and the Ser223[OH] form H-bonds with the R6-hydroxyl and the C2-carbonyl, respectively (Figure 3(F)). The Tyr219[OH], in turn, is potentially H-bonding with the C2-carbonyl. This underlines the importance of proximal groups capable of bonding at the coumarin's R4-R6 positions for the HSD1 inhibition (Figure 2; Table 1).

R5/R6-hydroxyl group is critical for the 3-phenylcoumarin inhibition

Analogues **2**–**7** (Figure 3; Table 1) collectively indicate that a halogen or hydroxyl at the R1 position (Figure 2) improves the HSD1 inhibition of the 3-phenylcoumarins (Table 1; Figure 5). The absolute position or even the presence of this group is, however, not essential for inhibition (Figure 2; Table 1). In contrast, if one considers only those six analogues (Figure 3), excluding **1** (Figure 4), that produces inhibition at 1 or 5 μM (Table 1), placing a hydroxyl group at the R4 or R5 position is a necessity (Figure 2).

In **8** (Figure 2), there is a hydroxyl group at the R2 position of the coumarin ring the same way as in **5**, but the lack of a hydroxyl group in the 3-phenyl ring renders the analogue unable to bond with Glu283 and His222. The loss of this dual contact is not fully compensated by the R5-fluorine and, as a result, the HSD1 inhibition is non-existent (Table 1). Further evidence of the importance of R4/R5-hydroxyl is provided by the inability of **9** (Figure 2) to prevent the HSD1 activation. The R5-methoxy of **9** cannot establish as strong H-bonding coordination for the 3-phenyl as a hydroxyl group in the "canonical" pose would. In this respect, **1** (Figure 2) is a noteworthy exception. Although the analogue's 3-phenyl ring contains only R5-methoxy group and no hydroxyl moiety (Figure 2), it still induces strong inhibition (Table 1).

R3-hydroxyl reverses the 3-phenylcoumarin binding

The binding of **1** is predicted to differ markedly (Figure 4) from other 3-phenylcoumarin analogues (Figure 3) producing HSD1 inhibition (Table 1; Figure 5). Instead of the "canonical" pose (Figure 4(A)), the coumarin and 3-phenyl ring systems of **1** are suggested to have reverse order or "non-canonical" positioning of at the site (Figure 4(B)) when compared to the other analogues (Figure 3). Even though this flip represents a profound change for the scaffold, it imposes only few drawbacks.

The ring systems of **1** pack against the same residues as they would in the "canonical" pose. Importantly, the R3-hydroxyl accepts an H-bond from the delta position of His222 (Figure 4(B)). This interaction is not feasible, when the hydroxyl is switched to the R2 position to produce the otherwise identical (but inactive) analogue **9** (Figure 2). Moreover, the C2-carbonyl and the heterocyclic oxygen in the coumarin can H-bond with the Tyr219^{OH} in this "non-canonical" pose. Due to the flip, the Glu283^{COO-} cannot H-bond with **1** (Figure 4); however, the inward pose of the residue is not required for binding (PDB: 3KLM[45]). In this "non-canonical" pose, the R5-methoxy H-bonds with the Ser143^{OH} and, additionally, the Val144^N and Tyr156^{OH} are favourably oriented towards the polar group (Figure 4(B)).

Cross-reactivity of the 3-phenylcoumarin analogues

It is not enough that drug candidates bind into their target proteins to elicit desired effects *in situ*. One also needs to consider their absorption, distribution, metabolism, and excretion (ADME) properties, toxicity, off-target effects, and overall selectivity. For example coumarins are known to produce hepatotoxic effects with a certain subgroup of humans – a phenomenon likely emerging from problems in the 7-hydroxylation of coumarins by the genetically polymorphic CYP2A6 enzyme[46,47]. Although no animal testing was performed in this study, the cross-reactivity of the 3-phenylcoumarin analogues was tested against ER, HSD2, CYP1A2, MAO-A, MAO-B, and aromatase using *in vitro* assays (Table 1).

Oestrogen receptor (ER) antagonists/agonists or selective oestrogen receptor modulators, such as tamoxifen and raloxifene are used routinely in treatment against ER-positive breast cancer. Potent HSD1 inhibitors could have a dual function as ER antagonists but they should not have a dual role as ER agonists promoting breast tissue tumourigenesis. Thus, the effect of the HSD1 inhibitor analogues was studied against both ER and of the potent HSD1 inhibitor analogues, only **5** was found to produce moderate ER inhibition. Of the more modest HSD1 inhibitor analogues **6** yielded reasonable ER inhibition. In addition, compounds **8** and **9**, which do not inhibit HSD1 activity, inhibited ER. The molecular basis for this is clear based on a prior study with the ER-specific compounds[14]: the 3-phenylcoumarin scaffold must have R2-functional group, e.g. R2-hydroxyl moiety, at its coumarin ring system to produce the inhibitory effect.

17-β-hydroxysteroid dehydrogenase 2 (HSD2), which is the enzymatic counterpart of HSD1, converts E_2 to E_1. Accordingly, to avoid counterproductive effects, it is paramount that any potential drugs aiming to lower the E_2 production should not effectively block the HSD2 activity as a side effect. The activity testing indicates that none of the 3-phenylcoumarin analogues produce >50% HSD2 inhibition at $1 \mu M$ as the inhibition remains at a range from 7 to 42% (Table 1). Notably, the most potent HSD1 inhibitor analogues block the HSD2 only at a moderate level (**1** at 27%, **2** at 7%, and **4** at 16%; Table 1). If concentrating on the HSD2 activity, **2** is the most selective HSD1 inhibitor analogue while **4** is a close runner-up. Although **4** is the more potent HSD1 inhibitor of the

two (or of all the tested analogues), the close to optimal H-bonding coordination with the R1- and R5-hydroxyls of **2** inside the HSD1 active site (Figure 3(A)) could be the underlying reason for its higher selectivity. However, the lack of 3D structural data on HSD2 or its homologous proteins, especially regarding the enzyme's binding site, make it difficult to resolve this issue.

Monoamine oxidases (MAO) A and B are inhibited to some degree by the 3-phenylcoumarin analogues and this effect is notable for the MAO-B (see e.g.[18,48,49]). For that reason, the inhibition levels of the analogues were studied here against both enzyme subtypes (Table 1). Analogous to earlier studies[18,49–51] analogues showing HSD1 inhibition also blocked the MAO-B activity at $10 \mu M$. However, of the HSD1 inhibitor analogues presented in this study, only **6** (Figure 6(A)) and **7** have pIC$_{50}$ above 6 (IC$_{50} < 1 \mu M$). Based on the docking, the R6-fluorine and R5-hydroxyl of **6** form a halogen bond and an H-bond with the Leu164^O and the Pro102^O, respectively. Interestingly, **4** has been tested with MAO-A and MAO-B earlier (C6 in[18]). Although **4** reached 64% inhibition at $10 \mu M$ in our studies, it has shown more promising activity in a study by Delogu et al.[18]. Overall, the results suggest that the MAO-B inhibition would not be a critical issue for the new analogues or at least for the most potent of them.

Cytochrome P450 (CYP) enzymes metabolise majority of oestrogens first in the liver. In this vital process, CYP1A2 enzyme has a prominent role[52] and, therefore, its unintended inhibition by a small-molecule could promote upswing in the effective E_2 levels. Because the ultimate goal of any HSD1 inhibitor, including the 3-phenylcoumarins presented in this study, is to lower the E_2 levels *in vivo*, their ability to block the CYP1A2 was studied as well. All of the analogues block CYP1A2 activity at some concentration, however, only the most potent HSD1 inhibitor **4** blocks its function at an alarming level (Table 1). The ligands that bind into the narrow and hydrophobic active site of CYP1A2 can be either substrates that are metabolised by the enzyme or inhibitors that block its function. As the substrates can be metabolised at different positions, it is unpractical to offer just one binding pose for each analogue. Regardless, for example the binding pose of **4**, which is the strongest CYP1A2 inhibitor of the analogue set (Table 1), likely reminds the validated pose of α-naphthoflavone (Figure 6(B))[28]. Based on the docking, the R1-chlorine of **4** packs against the haeme and the 3-phenyl ring is sandwiched between the side chains of Phe226 and Phe260 (not shown). Moreover, the C2-carbonyl of **4** forms an H-bond with a crystal water the same way as is seen for α-naphthoflavone and the Thr118^{OH} accepts an H-bond from the R4-hydroxyl (Figure 6(B)).

Aromatase (CYP19A1) inhibitors are used in breast cancer treatments, but unlike in the case of ER, their potential ability to bind into both HSD1 and aromatase could not be harmful. Aromatase inhibitors are predominantly used with post-menopausal breast cancer patients, because the E_2 production via the aromatase pathway happens locally rather than relying on the ovaries[53]. In contrast, although the 3-phenylcoumarin scaffold mimics the steroidal core, and fit into the active site of the aromatase, analogues **1–9** do not produce aromatase inhibition (Table 1). A closer inspection indicates that this lack of activity is due to the inability of the polar R1–R6 groups to produce favourable interactions at the aromatase's active site. On the one hand, the R1-positioned chlorine (**4**; Figure 2), methoxy (**7**; Figure 2), or hydroxyl group (**2**; Figure 2) could bond with the proton of Met374^N. On the other hand, while the R4-hydroxyl groups of **2**, **3** and **4** are within the H-bonding range from the Asp309 side chain, this key residue is in a neutral state at pH 7.4 (PDB: 3EQM)[26] and, therefore, ready to donate a proton instead of accepting one.

3-Imidazolecoumarin inhibits aromatase potently

The analysis of analogues **1**–**9** (see above) indicated that the coumarin-based compounds with flat ring systems at the 3-position could fit into the active site of the aromatase. However, a simple H-bond acceptor at the R4 position, such as a carbonyl group (of androstenedione in Figure 6(C)) would be needed to avoid the detrimental clash of proton donors between the bound ligand and the neutral Asp309 side chain at the active site. Instead of trying to "mutate" 3-phenylcoumarin core further to facilitate aromatase inhibition, a new kind of coumarin-derivative **10**, in which the 3-phenyl ring is substituted with a 3-imidazole was synthesised (Figure 2).

There are two potential binding poses at the aromatase's active site for **10**. First, the deprotonated N3′ of the 3-imidazole ring could accept an H-bond from the neutral Asp309 side chain (Figure 6(C)). Second, the N3′ could coordinate directly with the haeme. Although the latter option was not put forward by the docking (not shown), the imidazole group is known to bind strongly with the haeme groups and induced-fit effects could help to accommodate it at the site. Nevertheless, the activity testing shows that **10** inhibits strongly the aromatase ($pIC_{50} = 7.11$; Table 1).

Furthermore, cross-reactivity testing of **10** indicates that the compound is blocking neither HSD1 nor MAO-B but it has a stronger inhibitory effect with CYP1A2 than with any of the 3-phenylcoumarin analogues (Table 1). The coumarin ring of **10** is likely to be flipped in a reverse pose inside the active site of CYP1A2 in comparison to the 3-phenylcoumarin analogues (Figure 6(B,D)). Importantly, in this pose the deprotonated and electronegative N3′ of imidazole would be coordinated with the positively charged iron in the middle of the haeme; meanwhile, the C2-carbonyl of **10** accepts an H-bond from the Thr124^{OH} (Figure 6(D)).

3-Phenylcoumarins are not pan-assay interference compounds

The cross-reactivity data demonstrates that coumarin with C3-substituted phenyl or imidazole ring does not belong to the pan assay interference compounds (PAINS) category, but that it is a privileged structure, which can be fine-tuned or tailored to function selectively with various targets. The PAINS filtering[54], performed using CANVAS module of MAESTRO, supported this conclusion (no compounds filtered out).

Coumarins are a widely studied group of compounds with structural and pharmacological variability. Thus, it is not surprising that also 3-phenylcoumarins have been studied against other targets elsewhere. Some of the compounds published here have been previously tested for inhibitory activity against for HIV-1 replication (**1**; C17 in[15]), immune complex-mediated neutrophil oxidative metabolism (**2**; CHEMBL486894; C13 in[16]), glyceraldehyde-3-phosphate dehydrogenase (**3**; CHEMBL71407; C18 in[17]), and MAO-A and -B (**4**; C6 in[18]). All these compounds showed moderate ability to inhibit their intended targets. This further shows that 3-phenylcoumarins have interesting pharmacologic properties and that they have a broad utilisation range over therapeutic target proteins.

Conclusions

3-Phenylcoumarin (Figure 2) is established here as a non-steroidal scaffold for building potent small-molecule HSD1 inhibitors. The 3-phenyl and coumarin rings are suggested to adopt similar hydrophobic packing at the active site as the established steroidal compounds (Figure 1(D)). Five of the 3-phenylcoumarin analogues produced ≥62% HSD1 inhibition at 5 μM (Table 1). Moreover, three of the analogues produced C68% inhibition even at 1 μM (**1**,

2, and **4**; Figure 2; Table 1). The approximated pIC_{50} value at 1 μM for the three best analogues was ≥6.2. Housing polar moieties at the R5 and/or R6 positions in the 3-phenyl ring is generally critical for establishing the 3-phenylcoumarin binding and inhibition with HSD1 (Figure 5; Table 1). Introducing yet another polar group at the R1 position (Figure 5) in the coumarin ring boosts the HSD1 inhibition even further (e.g. **4**; Figure 3(E); Table 1). Moreover, inserting a hydroxyl group at the R3 position is expected to reverse the 3-phenylcoumarin binding at the active site (Figure 5) in comparison to the other analogues (**1**; Figures 3 and 4(B)) but without doing away with the inhibition (Table 1). A thorough cross-reactivity analysis highlights the fact that the 3-phenylcoumarin analogues block HSD2 only at moderate levels (Table 1), which is an essential feature for any potential drug candidates aiming to combat the E_2-linked diseases, such as breast cancer and endometriosis. In addition, substituting the 3-phenyl with an imidazole changed the scaffold selectivity completely as the resulting compound **10** blocked potently the aromatase instead of the HSD1. To sum up, the coumarin core can be tailored to block the E_2 synthesis by either the sulphatase pathway or the aromatase pathway by adding either a 3-phenyl or a 3-imidazole ring, respectively.

ORCID

Pekka A. Postila 🆔 http://orcid.org/0000-0002-2947-7991
Sanna Rauhamäki 🆔 http://orcid.org/0000-0002-3014-3120
Olli T. Pentikäinen 🆔 http://orcid.org/0000-0001-7188-4016

References

1. DeSantis C, Ma J, Bryan L, Jemal A. Breast cancer statistics, 2013. CA Cancer J Clin 2014;64:52–62.
2. Pasqualini J, Chetrite G, Blacker M, et al. Concentrations of estrone, estradiol, and estrone sulfate and evaluation of sulfatase and aromatase activities in pre- and postmenopausal breast cancer patients. J Clin Endocrinol Metab 1996; 81:1460–4.
3. Vihko P, Härkönen P, Soronen P, et al. 17β-Hydroxysteroid dehydrogenases – their role in pathophysiology. Mol Cell Endocrinol 2004;215:83–8.
4. Hanamura T, Niwa T, Gohno T, et al. Possible role of the aromatase-independent steroid metabolism pathways in hormone responsive primary breast cancers. Breast Cancer Res Treat 2014;143:69–80.
5. Frycz BA, Murawa D, Wysocki-Borejsza M, et al. Expression of 17β-hydroxysteroid dehydrogenase type 1 in gastric cancer. Biomed Pharmacother 2013;67:651–7.
6. Tomaszewska A, Roszak A, Pawlik P, et al. Increased 17ß-hydroxysteroid dehydrogenase type 1 levels in primary cervical cancer. Biomed Pharmacother Biomédecine Pharmacothérapie 2015;72:179–83.

7. Dassen H, Punyadeera C, Kamps R, et al. Estrogen metabolizing enzymes in endometrium and endometriosis. Hum Reprod 2007;22:3148–58.

8. Maltais R, Ayan D, Trottier A. Discovery of a non-estrogenic irreversible inhibitor of 17β-hydroxysteroid dehydrogenase type 1 from 3-substituted-16β-(m-carbamoylbenzyl)-estradiol derivatives. J Med Chem 2013;57:204–22.

9. Maltais R, Trottier A, Delhomme A, et al. Identification of fused 16β,17β-oxazinone-estradiol derivatives as a new family of non-estrogenic 17β-hydroxysteroid dehydrogenase type 1 inhibitors. Eur J Med Chem 2015;93:470–80.

10. Messinger J, Husen B, Koskimies P, et al. Estrone C15 derivatives-a new class of 17beta-hydroxysteroid dehydrogenase type 1 inhibitors. Mol Cell Endocrinol 2009;301:216–24.

11. Abdelsamie AS, van Koppen CJ, Bey E, et al. Treatment of estrogen-dependent diseases: design, synthesis and profiling of a selective 17ß-HSD1 inhibitor with sub-nanomolar IC50 for a proof-of-principle study. Eur J Med Chem 2017; 127:944–57.

12. Lilienkampf A, Karkola S, Alho-Richmond S, et al. Synthesis and biological evaluation of 17beta-hydroxysteroid dehydrogenase type 1 (17beta-HSD1) inhibitors based on a thieno[2,3-d]pyrimidin-4(3H)-one core. J Med Chem 2009; 52:6660–71.

13. Messinger J, Hirvelä L, Husen B, et al. Novel non-steroidal inhibitors of human 11beta-hydroxysteroid dehydrogenase type 1. Mol Cell Endocrinol 2006;104:137–50.

14. Niinivehmas SP, Manivannan E, Rauhamäki S, et al. Identification of estrogen receptor ligands with virtual screening techniques. J Mol Graph Model 2016;64:30–9.

15. Olmedo D, Sancho R, Bedoya LM, et al. 3-phenylcoumarins as inhibitors of HIV-1 replication. Molecules 2012;17:9245–57.

16. Kabeya LM, da Silva CHTP, Kanashiro A, et al. Inhibition of immune complex-mediated neutrophil oxidative metabolism: a pharmacophore model for 3-phenylcoumarin derivatives using GRIND-based 3D-QSAR and 2D-QSAR procedures. Eur J Med Chem 2008;43:996–1007.

17. Leitão A, Andricopulo AD, Oliva G, et al. Structure-activity relationships of novel inhibitors of glyceraldehyde-3-phosphate dehydrogenase. Bioorg Med Chem Lett 2004; 14:2199–204.

18. Delogu GL, Serra S, Quezada E, et al. Monoamine oxidase (MAO) inhibitory activity: 3-phenylcoumarins versus 4-hydroxy-3-phenylcoumarins. Chem Med Chem 2014;9:1672–6.

19. Buu-Hoi NP, Saint-Ruf G, Lobert B. Oxygen heterocycles. Part XIV. Hydroxylated 3-aryl- and 3-pyridyl-coumarins. J Chem Soc C Org 1969;16:2069–70.

20. Kirkiacharian S, Chidiack H, Philibert D, et al. Binding affinity to steroid hormone receptors and antiproliferative action on MCF—7 cells of coumarinic derivatives and isoflavonoids. Ann Pharm Fr 1999;57:332–9.

21. Quezada E, Delogu G, Picciau C, et al. Synthesis and vasorelaxant and platelet antiaggregatory activities of a new series of 6-halo-3-phenylcoumarins. Molecules 2010;15:270–9.

22. Pasanen M. Human placental aromatase activity: use of a C18 reversed-phase cartridge for separation of tritiated water or steroid metabolites in placentas from both smoking and non-smoking mothers in vitro. Biol Res Pregnancy Perinatol 1985;6:94–9.

23. Holt A, Sharman DF, Baker GB, Palcic MM. A continuous spectrophotometric assay for monoamine oxidase and related enzymes in tissue homogenates. Anal Biochem 1997; 244:384–92.

24. Korhonen LE, Rahnasto M, Mähönen NJ, et al. Predictive three-dimensional quantitative structure-activity relationship of cytochrome P450 1A2 inhibitors. J Med Chem 2005; 48:3808–15.

25. Mazumdar M, Fournier D, Zhu DW, et al. Binary and ternary crystal structure analyses of a novel inhibitor with 17beta-HSD type 1: a lead compound for breast cancer therapy. Biochem J 2009;424:357–66.

26. Ghosh D, Griswold J, Erman M, Pangborn W. Structural basis for androgen specificity and oestrogen synthesis in human aromatase. Nature 2009;457:219–23.

27. Binda C, Wang J, Pisani L, et al. Structures of human monoamine oxidase B complexes with selective noncovalent inhibitors: safinamide and coumarin analogs. J Med Chem 2007;50:5848–52.

28. Sansen S, Yano JK, Reynald RL, et al. Adaptations for the oxidation of polycyclic aromatic hydrocarbons exhibited by the structure of human P450 1A2. J Biol Chem 2007; 282:14348–55.

29. Niinivehmas SP, Salokas K, Lätti S, et al. Ultrafast protein structure-based virtual screening with Panther. J Comput Aided Mol Des 2015;29:989–1006.

30. Vainio MJ, Puranen JS, Johnson MS. ShaEP: molecular overlay based on shape and electrostatic potential. J Chem Inf Model 2009;49:492–502.

31. Lehtonen JV, Still DJ, Rantanen VV, et al. BODIL: a molecular modeling environment for structure-function analysis and drug design. J Comput Aided Mol Des 2004;18:401–19.

32. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14:33–8.

33. Koivunen JT, Nissinen L, Käpylä J, et al. Fluorescent small molecule probe to modulate and explore a2b1 integrin function. J Am Chem Soc 2011;133:14558–61.

34. Virtanen SI, Niinivehmas SP, Pentikäinen OT. Case-specific performance of MM-PBSA, MM-GBSA, and SIE in virtual screening. J Mol Graph Model 2015;62:303–18.

35. Frydenvang K, Lash LL, Naur P, et al. Full domain closure of the ligand-binding core of the ionotropic glutamate receptor iGluR5 induced by the high affinity agonist dysiherbaine and the functional antagonist 8,9-dideoxyneodysiherbaine. J Biol Chem 2009;284:14219–29.

36. Postila PA, Kaszuba K, Kuleta P, et al. Atomistic determinants of co-enzyme Q reduction at the Qi-site of the cytochrome bc1 complex. Sci Rep 2016;6:33607.

37. Kuleta P, Sarewicz M, Postila P, et al. Identifying involvement of Lys251/Asp252 pair in electron transfer and associated proton transfer at the quinone reduction site of Rhodobacter capsulatus cytochrome bc1. Biochim Biophys Acta Bioenerg 2016;1857:1661–8.

38. Lash LL, Sanders JM, Akiyama N, et al. Novel analogs and stereoisomers of the marine toxin neodysiherbaine with specificity for kainate receptors. J Pharmacol Exp Ther 2008; 324:484–96.

39. Postila PA, Swanson GT, Pentikäinen OT. Exploring kainate receptor pharmacology using molecular dynamics simulations. Neuropharmacology 2010;58:515–27.

40. Leanne Lash-Van Wyhe L, Postila PA, Tsubone K, et al. Pharmacological activity of C10-substituted analogs of the high-affinity kainate receptor agonist dysiherbaine. Neuropharmacology 2010;58:640–9.

41. Sawicki MW, Erman M, Puranen T, et al. Structure of the ternary complex of human 17beta-hydroxysteroid dehydrogenase type 1 with 3-hydroxyestra-1,3,5,7-tetraen-17-one (equilin) and NADP+. Proc Natl Acad Sci USA 1999;96:840–5.

42. Virtanen SI, Pentikäinen OT. Efficient virtual screening using multiple protein conformations described as negative images of the ligand-binding site. J Chem Inf Model 2010;50:1005–11.

43. Niinivehmas SP, Virtanen SI, Lehtonen JV, et al. Comparison of virtual high-throughput screening methods for the identification of phosphodiesterase-5 inhibitors. J Chem Inf Model 2011;51:1353–63.

44. Matos MJ, Santana L, Uriarte E3. Phenyl-coumarin. Acta Crystallogr Sect E Struct Reports Online 2012;68:0–6.

45. Aka JA, Mazumdar M, Chen CQ, et al. 17beta-hydroxysteroid dehydrogenase type 1 stimulates breast cancer by dihydrotestosterone inactivation in addition to estradiol production. Mol Endocrinol 2010;24:832–45.

46. Abraham K, Wöhrlin F, Lindtner O, et al. Toxicology and risk assessment of coumarin: focus on human data. Mol Nutr Food Res 2010;54:228–39.

47. Raunio H, Rahnasto-Rilla M. CYP2A6: genetics, structure, regulation, and function. Drug Metabol Drug Interact 2012;27:73–88.

48. Matos MJ, Vazquez-Rodriguez S, Uriarte E, et al. MAO inhibitory activity modulation: 3-phenylcoumarins versus 3-benzoylcoumarins. Bioorganic Med Chem Lett 2011;21:4224–7.

49. Rauhamäki S, Postila PA, Niinivehmas S, et al. Structure-activity relationship analysis of 3-phenylcoumarin-based monoamine oxidase B inhibitors. Front Chem 2018;6:41.

50. Patil PO, Bari SB, Firke SD, et al. A comprehensive review on synthesis and designing aspects of coumarin derivatives as monoamine oxidase inhibitors for depression and Alzheimer's disease. Bioorganic Med Chem 2013;21:2434–50.

51. Matos MJ, Terán C, Pérez-Castillo Y, et al. Synthesis and study of a series of 3-arylcoumarins as potent and selective monoamine oxidase B inhibitors. J Med Chem 2011;54:7127–37.

52. Tsuchiya Y, Nakajima M, Yokoi T. Cytochrome P450-mediated metabolism of estrogens and its regulation in human. Cancer Lett 2005;227:115–24.

53. Simpson ER. Sources of estrogen and their importance. J Steroid Biochem Mol Biol 2003;86:225–30.

54. Baell JB, Holloway GA. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. J Med Chem 2010;53:2719–40.

IV

DISCOVERY OF RETINOIC ACID-RELATED ORPHAN RECEPTOR ΓT INVERSE AGONISTS VIA DOCKING AND NEGATIVE IMAGE-BASED SCREENING

by

Sanna Rauhamäki, Pekka A. Postila, Sakari Lätti, Sanna Niinivehmas, Elina Multamäki, Klaus Liedl & Olli T. Pentikäinen 2018

Submitted manuscript

V

MOLECULAR DOCKING-BASED DESIGN AND DEVELOPMENT OF A HIGHLY SELECTIVE PROBE SUBSTRATE FOR UDP-GLUCURONOSYLTRANSFERASE 1A10

by

Risto O. Juvonen, Sanna Rauhamaki, Sami Kortet, Sanna Niinivehmas, Johanna Troberg, Aleksanteri Petsalo, Juhani Huuskonen, Hannu Raunio, Moshe Finel & Olli T. Pentikainen 2018

Molecular Docking-Based Design and Development of a Highly Selective Probe Substrate for UDP-glucuronosyltransferase 1A10

Risto O. Juvonen,*[,†] Sanna Rauhamäki,[‡] Sami Kortet,[‡,§] Sanna Niinivehmas,[‡] Johanna Troberg,[∥] Aleksanteri Petsalo,[†] Juhani Huuskonen,[§] Hannu Raunio,[†] Moshe Finel,[∥] and Olli T. Pentikäinen*[,‡,⊥]
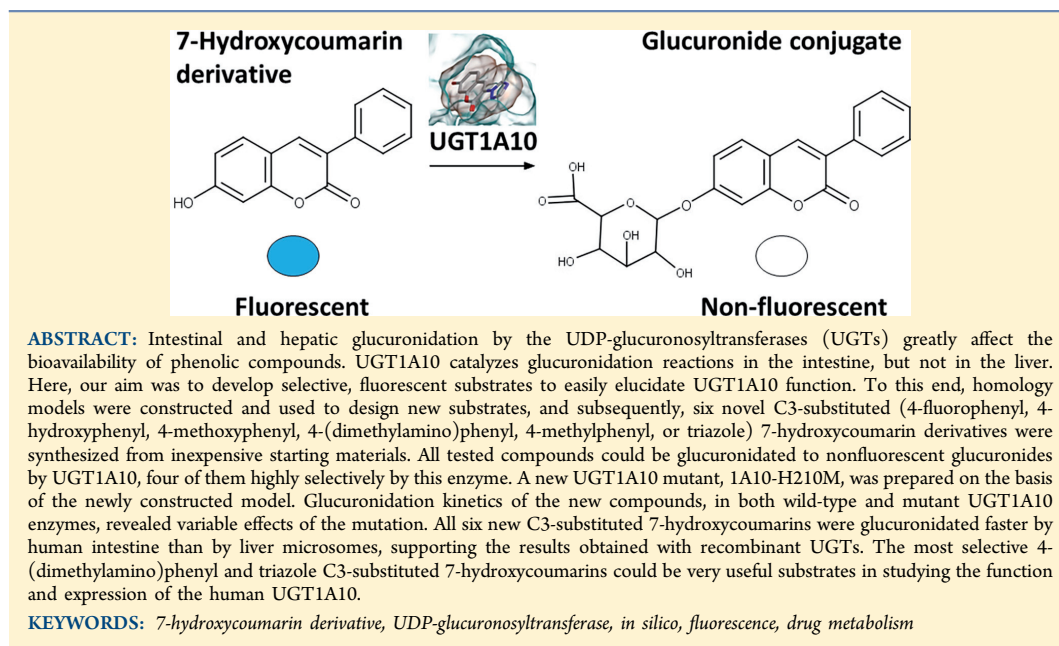
[†]School of Pharmacy, Faculty of Health Sciences, University of Eastern Finland, Box 1627, FI-70211 Kuopio, Finland

[‡]Department of Biological and Environmental Science and [§]Department of Chemistry, University of Jyvaskyla, P.O. Box 35, FI-40014 Jyvaskyla, Finland

[∥]Division of Pharmaceutical Chemistry and Technology, Faculty of Pharmacy, University of Helsinki, P.O. Box 56, FI-00014 Helsinki, Finland

[⊥]Institute of Biomedicine, Faculty of Medicine, University of Turku, FI-20014 Turku, Finland

S *Supporting Information*

ABSTRACT: Intestinal and hepatic glucuronidation by the UDP-glucuronosyltransferases (UGTs) greatly affect the bioavailability of phenolic compounds. UGT1A10 catalyzes glucuronidation reactions in the intestine, but not in the liver. Here, our aim was to develop selective, fluorescent substrates to easily elucidate UGT1A10 function. To this end, homology models were constructed and used to design new substrates, and subsequently, six novel C3-substituted (4-fluorophenyl, 4-hydroxyphenyl, 4-methoxyphenyl, 4-(dimethylamino)phenyl, 4-methylphenyl, or triazole) 7-hydroxycoumarin derivatives were synthesized from inexpensive starting materials. All tested compounds could be glucuronidated to nonfluorescent glucuronides by UGT1A10, four of them highly selectively by this enzyme. A new UGT1A10 mutant, 1A10-H210M, was prepared on the basis of the newly constructed model. Glucuronidation kinetics of the new compounds, in both wild-type and mutant UGT1A10 enzymes, revealed variable effects of the mutation. All six new C3-substituted 7-hydroxycoumarins were glucuronidated faster by human intestine than by liver microsomes, supporting the results obtained with recombinant UGTs. The most selective 4-(dimethylamino)phenyl and triazole C3-substituted 7-hydroxycoumarins could be very useful substrates in studying the function and expression of the human UGT1A10.

KEYWORDS: *7-hydroxycoumarin derivative, UDP-glucuronosyltransferase, in silico, fluorescence, drug metabolism*

INTRODUCTION

The extents of absorption and first-pass metabolism in the intestine and liver strongly affect the bioavailability of drugs and other orally ingested xenobiotic compounds.[1,2] Although hepatic metabolism is the major determinant of first-pass metabolism for most drugs, intestinal metabolism is critical for the bioavailability of certain compounds, particularly those that could be directly conjugated, like phenols and flavonoids.[3−5] Such xenobiotics typically contain a nucleophilic functional group, usually a hydroxyl group, that can accept an endogenous conjugating moiety, particularly glucuronic acid or sulfone.

More rarely, the conjugating group is methyl, acetyl, or amino acid.[6]
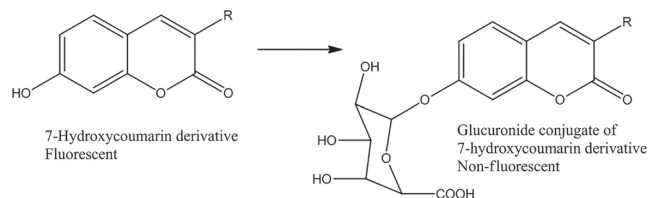
The UDP-glucuronosyltransferase enzymes (UGTs, EC 2.4.1.17) catalyze about 35% of the drug conjugation reactions and are abundantly expressed among the intestinal conjugating enzymes.[4,5,7] They catalyze transfer of the glucuronic acid moiety from UDP-glucuronic acid cofactor onto hydroxyl,

Figure 1. Glucuronidation of 7-hydroxycoumarin derivatives. Fluorescent 3-substituted 7-hydroxycoumarins are glucuronidated to nonfluorescent glucuronide conjugates by UGT enzymes. The decrease in fluorescence can be measured conveniently in different kinds of experimental setups.

amine, carboxylic acid, thiol, or thioacid groups of the aglycone substrates, reactions that are commonly called glucuronidation.[8] There is a significant difference in the expression profile of individual UGTs between the intestine and liver.[9−11] While about 10 different UGTs are abundantly expressed in the liver, only UGT1A1, UGT1A10, UGT2B7, and UGT2B17 are expressed in the small intestine to significant amounts at the protein level.[10,12−14] Unfortunately, the high activity and importance of the intestinal UGT1A10 was (and still is) often underestimated due to common use of poorly active commercial UGT1A10.[13]

UGT1A10 glucuronidates many drugs and xenobiotics.[13,15−17] Glucuronidation of estriol at the 3-OH and of estrone could be used as a selective reaction for UGT1A10, but the measurements require chromatographic separation of the resulting glucuronide from the substrate.[16,18] Likewise, dopamine is a UGT1A10-selective substrate, but its low affinity[19] has limited its use to qualitative measurements only.[14] Accordingly, availability of more convenient marker substrates for UGT1A10 would foster evaluation of glucuronidation reactions, particularly if the assays were easy and fast to perform.

The interactions between substrates and xenobiotic metabolizing enzymes are increasingly being studied by in silico modeling.[20−23] Currently, the lack of crystal structures of the N-terminal domain of any mammalian UGT impedes the structure−activity relationship and mechanistic studies. UDP-glucuronic acid binds at the C-terminal domain, which is highly homologous among different UGTs and evolutionary conserved. The structure of the latter domain has been solved by X-ray crystallography.[24] However, that available crystal structure does not provide sufficient data to predict binding of acceptor substrates, e.g., a drug molecule, as substrates bind to the N-terminal domain, which is more variable than the C-terminal domain among UGT enzymes. In silico models for this domain were constructed previously,[20,21] but we need better ones in order to analyze substrate−enzyme interactions in UGTs at the atomic level.

The hydroxyl group at C7 on the coumarin scaffold renders the compound fluorescent and glucuronidation of this hydroxyl abolishes the fluorescence (Figure 1).[25] Substituents at positions such as C3 or C4 of the coumarin do not quench the 7-hydroxycoumarin derivatives' fluorescence but modify its intensity, depending on the substituent's chemical nature. The 7-hydroxyl group on coumarin is also a good functional group for glucuronidation by many UGTs. Therefore, fluorescent 7-hydroxycoumarin derivatives provide an opportunity to design novel fluorescent substrates for UGTs using molecular modeling as the starting point.

In this study, we first constructed homology models for all the human UGT enzymes of subfamily 1A and used them to design fluorescent 7-hydroxycoumarin derivatives in silico. We then synthesized six compounds and developed a convenient multiwell plate assay protocol, based on fluorescence decrease, to test their glucuronidation rate. The results led to the identification of several 3-substituted 7-hydroxycoumarins as selective substrates for the human UGT1A10. A UGT1A10 mutant was prepared based on the model and as a test for it. Subsequently, glucuronidation kinetic analyses of the 3-substituted 7-hydroxycoumarins, by wild-type and mutant UGT1A10, were carried out using the same multiwell plate assay protocol.

■ EXPERIMENTAL SECTION

Materials. Alamethicin, trichloroacetic acid, UDPGA sodium salt, 7-hydroxycoumarin (99%), 7-hydroxy-(4-trifluoromethyl)coumarin (99%), and bovine serum albumin were from Sigma-Aldrich (Mannheim, Germany). Formic acid (99%) and MgCl$_2$ were from Riedel-de Haen (Vantaa, Finland). Acetonitrile (Ultra gradient HPLC grade), methanol (HPLC gradient grade), and glycine were from J.T. Baker (Deventer, The Netherlands). Ethanol (≥99.5%, Etax Aa) was from Altia (Helsinki, Finland). Water was deionized by Milli-Q gradient A10.

Methods. Modeling. To enable structure-based design of UGT1A10 selective substrates, all nine UGT1A-enzymes were modeled. Sequences of human UGT1A enzymes were gathered from the UniProt Knowledgebase at www.uniprot.org (UniProt Consortium, 2015). The accession codes for the retrieved UGT1As were: Q9HAW8 (1A10), O60656 (1A9), Q9HAW9 (1A8), Q9HAW7 (1A7), P19224 (1A6), P35504 (1A5), P22310 (1A4), P35503 (1A3), and P22039 (1A1). To identify template protein structures for homology modeling purposes, the retrieved UGT sequences were used in blast searches against the protein data bank (PDB) structures. Based on the results of these searches, structures 2O6L,[24] 3HBF,[26] 3WC4,[27] and 2C1Z[28] were selected as templates for homology modeling. The sequence alignment for the modeling was produced in two steps. First a protein structure-based sequence alignment was derived for the selected four protein structures by using Vertaa in BODIL[29] and the 2C1Z-structure as a template, since it gave the best match for UGT1A10 and contained both the N- and C-termini. In the second step the above listed nine UGT1A sequences were aligned, using BODIL, against the structural alignment, using STRMAT110 matrix[30] with 40 as the gap penalty. The alignment was adjusted for occasional variations in sequence length and used to create models for each UGT1A as well as in model construction that was performed using Modeler version 9.15.[31]

Molecular Docking. The ligands that were selected for docking studies were prepared by using LigPrep (version 3.3, Schrödinger, LLC, New York, 2015). The shapes and

Scheme 1. General Procedure for the Synthesis of Coumarin Derivatives

1 R = 4-fluorophenyl
2 R = 4-hydroxyphenyl
3 R = 4-(dimethylamino)phenyl
4 R = p-tolyl
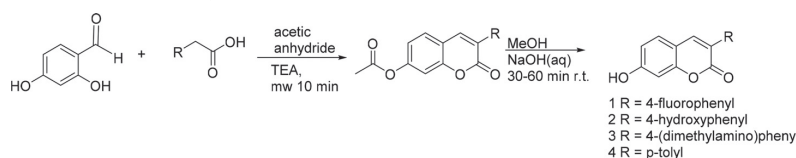
electrostatic properties of the substrate binding sites of modeling-produced UGT1A enzymes were analyzed with Panther,[32] and molecular docking was performed with PLANTS.[33]

Synthesis. All the synthesis reactions were carried out using commercial materials and reagents without further purification, unless otherwise noted. Reaction mixtures were heated using the CEM Discovery microwave apparatus. All reactions were monitored by thin-layer chromatography (TLC) on silica gel plates. ^{1}H NMR and ^{13}C NMR data were recorded on a Bruker Avance 400 MHz spectrometer or Bruker Avance III 300 MHz spectrometer. Chemical shifts are expressed in parts per million values (ppm) and are designated as s (singlet), br s (broad singlet), d (doublet), dd (double douplet), t (triplet). Coupling constants (J) are expressed as values in hertz (Hz). The mass spectra were recorded using Micromass LCT ESI-TOF equipment. Elemental analyses were done with Elementar Vario EL III elemental analyzer. All compounds tested present more than 95% purity.

General Procedure for the Synthesis of Coumarin Derivatives. The coumarin derivatives **1−6** were synthesized using the Perkin-Oglialor condensation reaction (Scheme 1). The method was developed from the previously published procedures and transferred to a microwave reactor.[34]

Typical procedure: A mixture of salicylaldehyde derivative (2 mmol) and phenyl acetic acid derivative (2.1 mmol), acetic acid anhydride (0.6 mL), and triethylamine (0.36 mL) was placed in a microwave reactor tube and heated at 100−170 °C in the microwave apparatus for 10−20 min. After cooling, 2 mL of 10% NaHCO$_3$ solution was added, and the precipitate was filtered, dried, and recrystallized from EtOH/H$_2$O or acetone/ H$_2$O mixture. The acetyl group(s) were removed by treating the compound with MeOH/NaOH(aq) solution for 30−60 min at rt. The solution was acidified with HCl (aq), and the precipitate was filtered and recrystallized if needed. Experimental data for 7-acetoxy-3-(4-fluorophenyl)-2*H*-chromen-2one, 7-hydroxy-3-(4-fluorophenyl)-2*H*-chromen-2-one (**1**), 7-acetoxy-3-(4-acetoxyphenyl)-2*H*-chromen-2one, 7-hydroxy-3-(4-hydroxyphenyl)-2*H*-chromen-2-one (**2**), and 7-acetoxy-3-(4-methoxyphenyl)coumarin, 7-hydroxy-3-(4-methoxyphenyl)-coumarin (**3**) were already published elsewhere.[34]

7-Hydroxy-3-(4-(dimethylamino)phenyl)-2H-chromen-2-one (4).[35] In the first step, 7-acetoxy-3-(4-(dimethylamino)-phenyl)-2*H*-chromen-2one was obtained. Yield: 70%. ^{1}H NMR (400 MHz, DMSO-d_6) δ: 2.31 (s, 3H, CH$_3$CO), 2.95 (s, 6H, (CH$_3$)$_2$N−), 6.77 (d, J^3 = 9.0 Hz, 2H, H-2′, H-6′), 7.14 (dd, J^3 = 8.4 Hz, J^4 = 2.2 Hz, 1H, H-5), 7.26 (d, J^4 = 2.2 Hz, 1H, H-8), 7.63 (d, J^3 = 9.0 Hz, 2H, H-3′, H-5′) 7.76 (d, J^3 = 8.5 Hz, 1H, H-5), 8.11 (s, 1H, H-4). ^{13}C NMR (100.6 MHz, DMSO-d_6) δ: 20.85, 39.84, 109.44, 111.58, 117.76, 118.57, 121.57, 126.00, 128.82, 129.11, 136.46, 150.45, 151.90, 152.77, 159.74, 168.85.

In the second step, 7-hydroxy-3-(4-(dimethylamino)phenyl)-2*H*-chromen-2-one (**4**) was obtained. Yield: 85% yellow solid. ^{1}H NMR (400 MHz, DMSO-d_6) δ: 2.94 (s, 6H, (CH$_3$)$_2$N-),

6.72 (d, J^4 = 2.3 Hz, 1H, H-8), 6.75 (d, J^3 = 9.0 Hz, 2H, H-2′, H-6′), 6.79 (dd, J^3 = 8.4 Hz, J^4 = 2.3 Hz, 1H, H-5),), 7.55 (d, J^3 = 8.5 Hz, 1H, H-5), 7.58 (d, J^3 = 9.0 Hz, 2H, H-3′, H-5′), 7.99 (s, 1H, H-4). ^{13}C NMR (100.6 MHz, DMSO-d_6) δ: 39.92, 101.59, 112.33, 113.16, 122.30, 122.32, 129.34, 137.83, 150.07, 154.27, 160.30, 160.41. ESI-MS: m/z (rel abund) 304 [M + Na$^+$]. Anal. Calcd for C$_{17}$H$_{15}$N$_1$O$_3$: C, 72.58; H, 5.37; N, 4.98. Found: C, 72.45; H, 5.40; N, 5.15.

7-Hydroxy-3-(p-tolyl)-2H-chromen-2-one (5).[36] In the first step, 7-acetoxy-3-(*p*-tolyl)-2*H*-chromen-2one was obtained. Yield: 70%. ^{1}H-NMR (300 MHz, d[6]-DMSO) δ: 2.34 (s, 3H, CH$_3$-Ph), 6.75 (d, J^4 = 2.2 Hz, 1H, H-8), 6.81 (dd, J^3 = 8.5 Hz, J^4 = 2.2 Hz, 1H, H-6), 7.23 (d, J^3 = 8.0 Hz, 2H, H-2′, H-6′), 7.58 (d, J^3 = 8.5 Hz, 1H, H-5), 7.59 (d, J^3 = 8.1 Hz, 2H, H-3′, H5′), 8.10 (s, 1H, H-4), 10.54 (s, 1H, Ph-OH). δ: 2.32 (s, 3H, CH$_3$CO), 2.35 (s, 3H, CH$_3$−Ph), 7.17 (dd, J^3 = 8.4 Hz, J^4 = 2.2 Hz 1H, H-6), 7.27 (d, J^3 = 8.0 Hz, 2H, H-2′, H-6′), 7.29 (d, J^4 = 2.2 Hz, 1H, H-8), 7.63 (d, J^3 = 8.2 Hz, 2H, H-3′, H-5′) 7.80 (d, J^3 = 8.5 Hz, 1H, H-5), 8.21 (s, 1H, H-4). ^{13}C NMR (100.6 MHz, DMSO-d_6) δ: 20.77, 20.82, 109.56, 117.37, 118.66, 125.99, 128.25, 128.76, 129.30, 131.60, 138.09, 139.37, 152.51, 153.28, 159.47, 168.74.

In the second step, 7-hydroxy-3-(*p*-tolyl)-2*H*-chromen-2one (**5**) was obtained. Yield: 85% brownish solid. ^{1}H NMR (300 MHz, DMSO-d_6) δ. ^{13}C NMR (75 MHz, DMSO-d_6) δ: 20.74, 101.65, 112.01, 113.30, 122.07, 128.04, 128.69, 129.80, 132.14, 137.37, 140.38, 154.77, 160.05, 161.05. ESI-MS: m/z (rel abund) 275 [M + Na$^+$]. Anal. Calcd for C$_{16}$H$_{12}$O$_3$: C, 76.18; H, 4.79. Found: C, 75.95; H, 4.83.

7-Hydroxy-3-(1H-1,2,4-triazol-1-yl)-2H-chromen-2-one (6). In the first step, 2-oxo-3-(1*H*-1,2,4-triazol-1-yl)-2*H*-chromen-7-yl acetate was obtained. Yield: 65% slightly brown solid. ^{1}H NMR (400 MHz, DMSO-d_6) δ: 2.33 (s, 3H, CH$_3$-CO), 7.27 (dd, J^3 = 8.5 Hz, J^4 = 2.2 Hz, 1H, H-6), 7.44 (d, J^4 = 2.2 Hz 1H, H-8), 7.98 (d, J^3 = 8.5 Hz, 1H, H-5), 8.33 (s, 1H, H-3′), 8.61 (s, 1H, H-4), 9.20 (s, 1H, H-5′). ^{13}C NMR (100.6 MHz, DMSO-d_6) δ: 20.87, 110.02, 116.19, 119.50, 123.04, 130.10, 131.81, 144.40, 152.06, 152.21, 153.12, 155.73, 168.76. FT-IR (KBr, cm^{-1}): 3416, 1723, 1617, 1208, 1129. Mp: 190−192 °C. ESI-MS: m/z (rel abund): calcd for [M + Na$^+$] 294.0485, measured 294.0499, Δ = −1.4 mDa.

In the second step, the acetyl group was removed by treating the above-mentioned compound with MeOH/K$_2$CO$_3$ solution. Yield: 92% slightly brown solid. ^{1}H NMR (400 MHz, DMSO-d_6) δ: 6.85 (d, J^4 = 2.3 Hz, 1H, H-8), 6.90 (dd, J^3 = 8.5 Hz, J^4 = 2.3 Hz, 1H, H-6), 7.74 (d, J^3 = 8.6 Hz, 1H, H-5), 8.27 (s, 1H, H-3′), 8.49 (s, 1H, H-4), 9.10 (s, 1H, H-5′), 10.83 (br s, 1H, OH). ^{13}C NMR (100.6 MHz, DMSO-d_6) δ: 102.07, 110.42, 114.23, 119.72, 130.70, 134.11, 144.24, 151.78, 153.96, 156.37, 162.01. FT-IR (ATR, cm^{-1}): 1724, 1623, 1511, 1328, 1129 811. Mp: > 312 °C dec. ESI-MS: m/z (rel abund) calcd for [M − H$^+$] 228.0415, measured 228.0408. Δ = 0.7 mDa.

UGTs and Microsomes. Recombinant human UGTs 1A1, 1A3, 1A6, 1A7, 1A8, 1A9, 1A10, 2A1, 2A2, 2A3, 2B4, 2B10,

2B7, and 2B17 were produced, as His-tagged proteins, in baculovirus-infected insect cells as previously described.[37−39] The relative expression level of each of these recombinant UGTs was evaluated by immunodetection, using monoclonal antibody against the His-tag, as described elsewhere.[40] A numeric value of 1.0 was given to the expression level of UGT1A8 and the relative expression level of each of the other UGTs was related to this value. Normalized activities were obtained by dividing the glucuronidation rate values by the relative expression level of the tested UGT. In addition, UGTs 1A4, 2B10, and 2B15 were purchased from Corning Life Sciences (New York) and are marked, in Figure 5, with a "C" to indicate that they are commercial enzymes. The expression levels of the UGTs in the commercial samples could not be determined, so their protein concentration was used to calculate the reaction rate.

A commercial HLM pool (cat no:452210, BD Gentest, Bedford, MA) and human intestine microsomes (HIM) pool (lot. no 1110189, XenoTech, Kansas City, KS) were purchased. Pig liver samples were prepared from untreated female pigs that were used for practicing surgical procedures at the University of Kuopio (currently: University of Eastern Finland, Kuopio campus). Other animal liver microsomes were prepared as described previously.[41] The Ethics Committee for Animal Experiments, University of Kuopio, approved these experiments.

Mutagenesis of UGT1A10. The UGT1A10 mutant 1A10-H210 M was prepared according to the QuikChange methodology, using the cloned UGT1A10 in pFastBac[42] as a template and the following two oligonucleotides:

1. 5′-ACTTTCAAGGAGAGAGTATGGAACATGATCGT-GCACTTGGAGGACCATTT-3′

2. 5′- AAATGGTCCTCCAAGTGCACGATCATGTTCC-ATACTCTCTCCTTGAAAGT-3′

The entire coding sequence of UGT1A10 in the mutant clone was sequenced, and subsequently, recombinant baculovirus was prepared and used to express the mutant enzyme in baculovirus-infected SF9 insect cells.

Absorbance and Fluorescence Spectra of C3-Substituted 7-Coumarins. The absorbance spectra of 10 μM coumarin derivatives at 100 mM Tris−HCl pH 7.4 containing 10% dimethyl sulfoxide were measured using a Hitachi U-2000 Spectrophotometer (Tokyo, Japan). Excitation and emission fluorescence spectra of 0.1 μM coumarin derivatives in 100 mM Tris-HCl pH 7.4 were measured using a Shimadzu RF-5000 spectrophotofluorometer (Tokyo, Japan). The excitation spectra was from 200 to 420 at 460 nm emission and the emission spectra was from 400 to 600 at 390 nm excitation (data in Supplementary Table S1 and Figures S1 and S2). The effect of pH on fluorescence intensity was determined at 405 nm excitation and 460 nm emission, in the presence of 1.5% trichloroacetic acid, 100 mM phosphate buffers at pH 5, 6, 7, 8, and 9, or 1.6 M glycine−NaOH pH 10.4.

Glucuronidation Reactions. The incubation mixtures for glucuronidation assays contained 100 mM Tris-HCl buffer pH 7.4, 2.5 mM MgCl₂, 0.5 mM UDPGA, recombinant UGT, or microsomes as the enzyme source and 0−15 μM of the test aglycone substrate. When the incubation mixtures contained microsomes, alamethicin was used at a final concentration of 12.5 μg/mL, but it was not included in the recombinant enzyme assays.[43] In the first experiments three negative control samples were tested, namely (i) without the substrate 7-hydroxycoumarin derivative, (ii) without the cofactor UDPGA,

or (iii) without the enzyme source. In subsequent experiments, the control samples lacked the enzyme source since it gave the highest fluorescence background. Preliminary experiments were done under different conditions than most of the later assays, namely in 2.5 mL of 100 mM phosphate buffer pH 7.4, containing 500 μg pig liver microsomes, 1.0 μM compound **6** and 0.5 mM UDPGA, at room temperature. Fluorescence spectra (excitation 200−420 nm; emission 400−600 nm) or fluorescence decrease at 390 nm excitation and 460 nm emission, were measured in these preliminary experiments, using a Shimadzu RF-5000 spectrophotofluorometer.

Most of the glucuronidation assays were carried out in 96 multiwell plate format, and incubations were carried out in 100 μL and at 37 °C, in the presence of Tris−HCl buffer pH 7.4, UDPGA, and the tested 7-hydroxycoumarin derivative, at the indicated concentrations. Fluorescence decline in the multiwell plate experiments was monitored every other minute, for 40 min, using an excitation filter at 405 nm and detection at 460 nm, in a Victor2 1420 Multilabel counter (PerkinElmer, Life Sciences, Turku, Finland). The fluorescence values were transformed into molarity using the aglycone substrates for making the respective standard curves. Slopes of the substrate concentration decrease per minute were calculated using linear regression analysis, in which the linear part of the kinetic assay indicated the glucuronidation rate. Enzyme catalyzed glucuronidation rate was calculated by subtracting the blank value from the full reaction value. The intra-assay variability of the kinetic assays was 6% when compound **6** was used as the aglycone substrate. Kinetic analyses were also performed in the same 96 multiwell plates format, with excitation filter at 405 nm and detection at 460 nm, using 6−8 different substrate concentrations per substrate and two different protein concentrations for both the wild-type and mutant UGT1A10. The higher concentrations, 13.5 mg/L for UGT1A10 and 12.0 mg/L for the UGT1A10 mutant, were used when the substrates were HFC, compounds **4** and **6**, whereas the lower protein concentrations, 6.75 and 6.0 mg/L for UGT1A10 and the mutant, respectively, were used when the substrates were compounds **1**, **2**, **3**, and **5**.

In end-point determinations, the glucuronidation reactions were stopped by the addition of 150 μL 1.6 M glycine-NaOH buffer, pH 10.4, followed by fluorescence measurements at 405 nm excitation and 460 nm emission. There was a good correlation between kinetic and end-point assays (data not shown).

HPLC−MS Analysis of 7-Hydroxy-3-triazolecoumarin Glucuronide. For further analysis of 7-hydroxy-3-triazolecoumarin glucuronide, 10 μM of compound **6** were incubated in 100 μL of 100 mM Tris−HCl pH 7.4 buffer containing 5 mM MgCl₂, 1 mM UDPGA, and either 30−40 μg recombinant UGT, 30 μg HLM, or 20 μg HIM for 1 h at 37 C. The reactions were stopped by the addition of 300 μL of methanol and centrifugation, and then the supernatant was divided into two. One part, 100 μL supernatant, was mixed with 150 μL of 1.6 M glycine−NaOH pH 10.4 and subjected to fluorescence measurements as described above for the 96-well plate. The other part, 250 μL supernatant, was stored at −20 C until analysis by HPLC−MS.

An Agilent 1200 Series Rapid Resolution LC System (Agilent Technologies, Waldbronn, Germany) was used for the chromatographic separation, equipped with a reversed-phase C8 column (Brownlee Supra, 3 μm, 50 × 2.1 mm, PerkinElmer). The mobile phases were (A) 0.1% aqueous

formic acid and (B) acetonitrile. A linear gradient from 10% to 90% B in five min was applied, followed by one min isocratic elution with 90% B and column re-equilibration, yielding a total analysis time of nine min. The flow rate was 0.3 mL/min, injection volume 3 μL and column oven temperature 30 °C. For the MS detection, a Finnigan LTQ ion trap mass spectrometer (Thermo, San Jose, CA) was used in the positive electrospray (ESI) mode. A divert valve was used to direct eluent flow to the waste for 1 min at the beginning and at the end of the gradient run. The MS analysis was carried out using the following parameters: spray voltage 4 kV, sheath gas 30 (instrument units), aux gas 15, sweep gas 3, capillary temperature 250 °C, capillary voltage 43 V, tube lens 100 V. The collision energy used for MS/MS was 30 V. The acquired full scan MS range was m/z 60−600. Ions used for the detection of the aglycone and glucuronide conjugate were m/z 230 $[M + H]^+$ and m/z 406 $[M + H + 176]^+$, respectively. Data was acquired and processed using Xcalibur software package.

■ RESULTS

Modeling UGT1A Enzymes. Template structures for homology modeling of the UGT enzymes of subfamily 1A were selected based on sequence similarity using the blast option in Uniprot. Structures that contained both the sugar—nucleotide cofactor and a small aglycone molecule were beneficial for the model, and therefore, structures (pdb-codes) 2C1Z/X, 2O6L, 3HBF, and 3WC4 were selected. The conserved lipophilic region that spans the endoplasmic reticulum membrane and ends in the "cytoplasmic tail" were excluded from the model because the focus of the model construction was on the catalytic site and its surrounding area. In addition, other structures outside the catalytic site of the model, including loops and amino acid side chain conformations, were not optimized. Optimizing efforts were considered detrimental for the quality of the models as the sequence similarity between the templates and the UGT sequences was very low. In other words, the models were not considered to be good enough for molecular dynamics simulations in order to evaluate binding energy or other kinetic parameters. On the other hand, an essential part of the model building was to predict binding of the 7-hydroxycoumarin derivatives so that glucuronidation will occur at the 7-hydroxyl group, and for this purpose the models were adequate.

The homology models were used to design novel and UGT-selective fluorescent 7-hydroxycoumarin substrate molecules, whose fluorescence will decrease upon glucuronidation (Figure 1). Docking of the 7-hydroxycoumarin scaffolds (Figure 2A-C) into different UGT1A models yielded several possible conformations. One of them oriented the 7-hydroxy group toward the bound UDP-glucuronic acid (UDPGA) cofactor at the catalytic site, and simultaneously, the coumarin core was optimally placed into a nearby cavity in the catalytic site (Figure 2D). In our models, the coumarin scaffold could be stabilized through its C2-carbonyl oxygen, accepting a hydrogen bond from Q101 in UGT1A8 (Figure 2C), UGT1A9, and UGT1A10 (Figure 2A). Even stronger stabilizing interactions could be formed between the coumarin carbonyl at C2 and R102 of UGT1A6 (Figure 2C). On the other hand, D103 in UGT1A1 (Figure 2B), E104 in UGT1A3 and UGT1A4, or P101 in UGT1A7, are not capable of forming similar stabilizing hydrogen bond interactions at this site, suggesting that the coumarin core would not be an optimal substrate for these UGTs. Furthermore, R103 of UGT1A7, UGT1A8, and
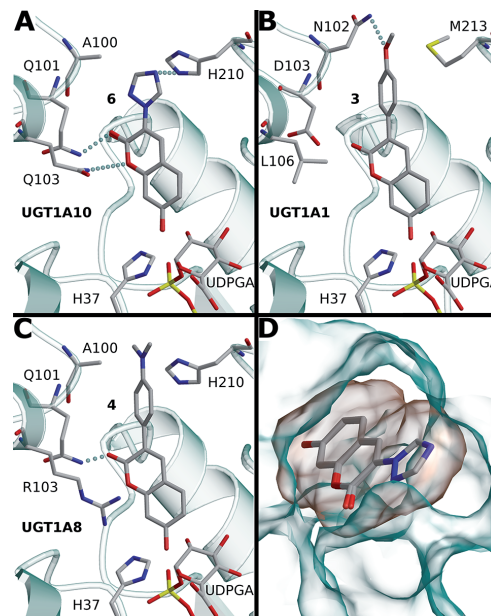

Figure 2. Docking of 7-hydroxycoumarin and its derivatives into the UGT1A10 model. (A) Molecular docking placed the 7-hydroxycoumarin with the 7-hydroxy facing the catalytic site formed by H37 and UDPGA, enabling the glucuronidation reaction and the subsequent decrease in fluorescence. (B) In UGT1A1, D103, and L106 are not capable of forming similar beneficial interactions with the coumarin core as the UGT1A10 model. Although N102 might be able to form beneficial interactions with certain compounds, the cavity might not be large enough for more sizable substitutions due to N102 and M213. (C) UGT1A8 has otherwise the same amino acid residues aligning the cavity as UGT1A10, except for R103. Due to its size, R103 might impair the binding of the coumarin core. (D) **6** fills the binding cavity (orange solvent accessible surface) of the UGT1A10 model (green solvent accessible surface). Docking suggests that the 7-hydroxycoumarin and **6** have a similar binding mode at their identical core. In addition, the C3 substituent of **6** fits tightly to the additional space at the outward facing end of the binding cavity.

UGT1A9, counteracted the stabilizing interactions and partially blocked the binding of the coumarin scaffold to these UGTs due to its size (Figure 2C). Unlike significantly bulkier R103 of UGT1A8 (Figure 2C), the Q103 of UGT1A10 at this site could donate a hydrogen bond to the oxygen at position 1 of the coumarin scaffold (Figure 2A).

Based on these homology models, there was an additional, but variable in size, free space available in the active site of each UGT1A enzyme, next to position 3 of the coumarin core. When combined with the 7-hydroxycoumarin docking site analyses (see above), this additional space was large enough in UGT1A1 and UGT1A10 to accommodate five- or six-membered ring substituents (Figure 2D). In contrast, a bulkier phenylalanine (F212) is present in the UGT1A6 model at this position, clearly hindering binding of any 3-substituted 7-hydroxycoumarins. The properties of this binding site were quite different also between UGT1A1 and UGT1A10. While UGT1A1 has a hydrophobic methionine (M213) facing this site (Figure 2B), UGT1A10 has a histidine at the
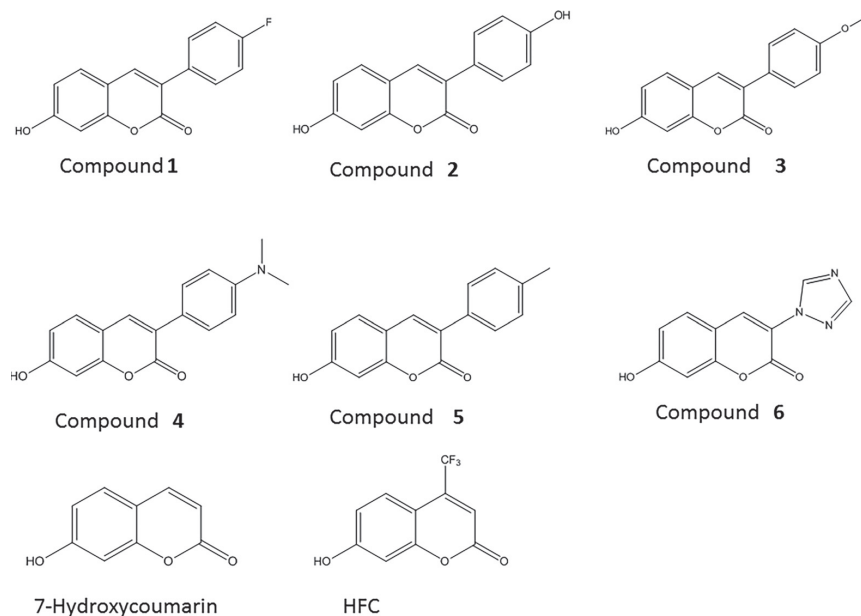
Figure 3. Novel (**1−6**) and control 7-hydroxycoumarin and 7-hydroxy-4-trifluoromethylcoumarin (HCF) substrates of UGT1A enzymes in this study.
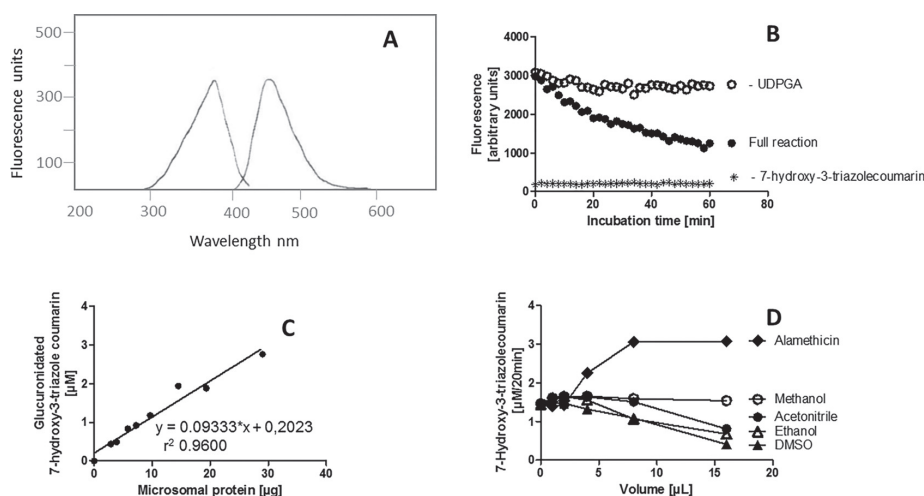

Figure 4. Decrease of 7-hydroxycoumarin fluorescence during glucuronidation. Excitation and emission fluorescence spectra of 0.1 μM 7-hydroxy-3-triazolecoumarin at 100 mM phosphate pH 7.4 buffer (panel A); decrease in 10 μM 7-hydroxy-3-triazolecoumarin fluorescence in the presence of 0.7 mg/mL pig liver microsomal protein, 0.5 mM UDPGA, 5 mM $MgCl_2$, and 100 mM Tris-HCL pH 7.4 (panel B); effect of the amount of microsomal protein on the decrease in fluorescence (panel C); effect of solvents and alamethicin on the decrease in fluorescence (panel D). Corresponding results were obtained with other C7-substituted 7-hydroxycoumarin derivatives.

corresponding site (H210; Figure 2A). On the opposite side of this additional space, UGT1A10 has an alanine (A100), while UGT1A1 has an asparagine (N102) aligned to the same position. This may mean that in UGT1A10 there is enough space for rather large 3-substituted 7-hydroxycoumarins, whereas UGT1A1 would have problems accommodating larger

subsitutions at this position of the 7-hydroxycoumarin scaffold. At the other end of the UGT1As "spectrum", UGT1A6 has a glutamic acid (E101) aligned at the same site, which would limit the cavity even further.

Based on docking of a virtual library into the UGT1A10 model, the most promising molecules were 7-hydroxycoumarin
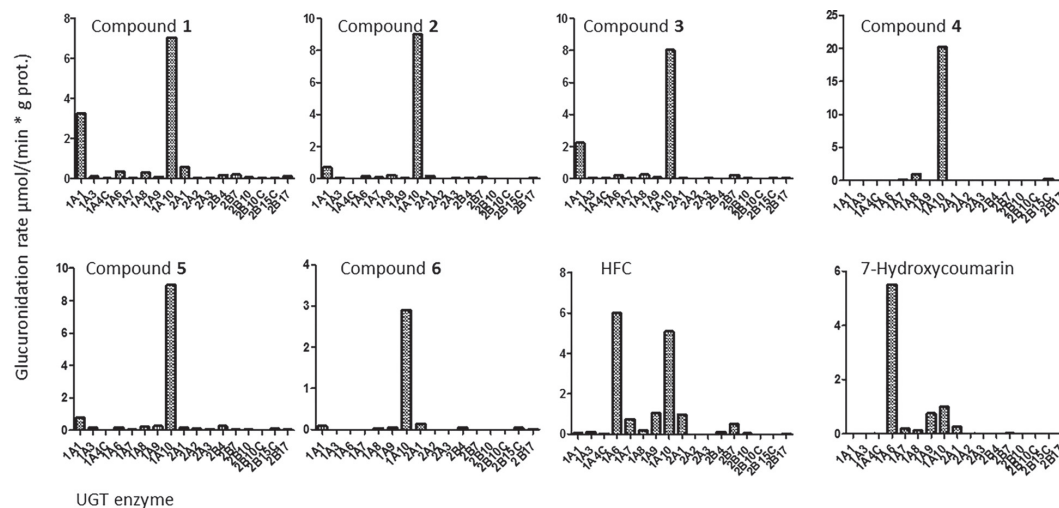
Figure 5. Glucuronidation of C3-substituted 7-hydroxycoumarins by human UGTs. Glucuronidation was determined at 10 μM aglycone concentration.

derivatives with the following substituents at the C3 position: 4-fluorophenyl (**1**), 4-hydroxyphenyl (**2**), 4-methoxyphenyl (**3**), 4-(dimethylamino)phenyl (**4**), 4-methylphenyl (**5**) or triazole (**6**). Among these, the triazole (**6**) derivative could be stabilized by H210 in UGT1A10, while the corresponding methionine in UGT1A1 (M213) would not be able to form hydrogen bonds with it. In contrast, the addition of a phenyl moiety to the C3-position would allow binding by both UGT1A1 and UGT1A10, as the hydrophobic methionine in UGT1A1 is quite an ideal companion, while the histidine in UGT1A10 could change its conformation according to the donor functionality, as with triazole. Based on these observations and considerations, the six new coumarin derivatives were designed and synthesized (Figure 3: compounds **1**-**6**). The selected compounds were made from inexpensive starting materials, producing 3-substituted coumarins with 5- or 6-membered rings.

Glucuronidation of the 3-Substituted 7-Hydroxycoumarins. The absorbance and fluorescence spectra of compounds **1**−**6** were very similar showing, however, some differences in intensity as well as excitation and emission peaks among them (Figure 4A, Supplementary Figures 1 and 2, and Supplementary Table 1). The fluorescence was pH dependent, emitting strongly in neutral and alkaline solutions.

In the first glucuronidation experiments, the fluorescence intensity of all of the C3-substituted 7-hydroxycoumarin derivatives, compounds **1**−**6**, decreased when they were incubated in the presence of pig liver microsomes and UDPGA in a buffer containing 100 mM Tris-HCl pH 7.4 and 5 mM MgCl$_2$ (Figure 4B). No significant fluorescence decrease was observed in the negative controls, namely in the absence of either microsomes, a coumarin derivative such as **6**, or UDPGA. The fluorescence changes were linearly dependent on the amount of microsomes (Figure 4C), indicating that the biosynthesis of C3 substituted 7-hydroxycoumarin derivatives to nonfluorescent glucuronide conjugates was catalyzed by one or more UDP-glucuronosyltransferase enzymes in the microsomes. The presence of alamethicin increased glucuronidation rates in this microsomal sample, whereas the addition of more

than 5% (v/v) dimethyl sulfoxide, acetonitrile, or ethanol decreased it (Figure 4D). An example for glucuronide formation under such incubation condition, using **6** as the substrate, is presented in Figure S3.

The above results with **6** and pig liver microsomes suggested that the glucuronidation rates of the new 7-hydroxycoumarin derivatives could be accurately determined under our standardized assay conditions, using different enzyme sources. Subsequently, the glucuronidation rates of all the new C3 substituted 7-hydroxycoumarin derivatives by nearly all the human UGTs, including enzymes that are not commercially available, were determined (Figure 5). For comparison and additional controls, the glucuronidation of 7-hydroxycoumarin and 4-(trifluoromethyl)-7-hydroxycoumarin (HFC), that carry no C3 substitution, by the recombinant enzymes were also measured (Figure 5).

The human UGTs screen revealed that all six newly synthesized C3 substituted 7-hydroxycoumarins were glucuronidated by UGT1A10 more rapidly than by other UGT forms. Furthermore, **2**, **4**, **5**, and **6** were selective for UGT1A10, as other UGTs catalyzed their glucuronidation at very low rates. The remaining two new substrates, **1** and **3**, were glucuronidated, in addition to UGT1A10, also by UGT1A1, at rates of about 40% and 20%, respectively, of the corresponding UGT1A10 rate (Figure 5). The glucuronidation profiles of the new derivatives differed significantly from the control substrates, 7-hydroxycoumarin and HFC, which were glucuronidated primarily by UGT1A6 (7-hydroxycoumarin), or UGT1A6 and UGT1A10 (HFC), as well as by few other UGTs at lower rates (Figure 5).

In addition to recombinant UGTs, the glucuronidation rates of all the new C3 7-hydroxycoumarin derivatives, along with 7-hydroxycoumarin and HFC, were measured in human liver microsomes (HLM) and human intestinal microsomes (HIM). The results showed that all the new compounds were glucuronidated by HIM at higher rates than by HLM (Figure 6). Furthermore, **2**, **4**, and **6** were not glucuronidated by HLM at all, or only at very low rates, whereas **3** and **5** were
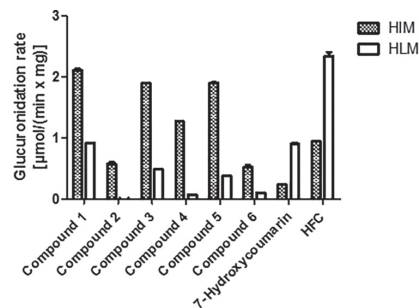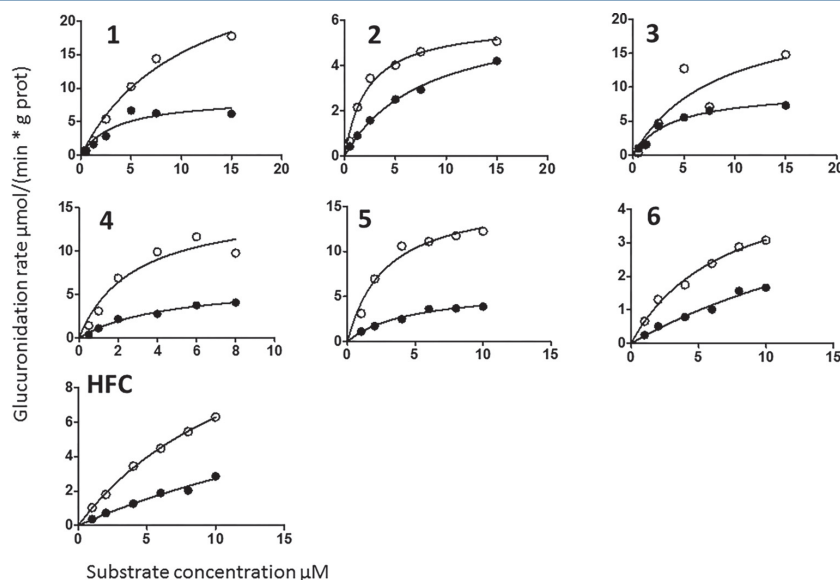
Figure 6. Glucuronidation of C3-substituted 7-hydroxycoumarins by human intestinal and hepatic microsomes. Glucuronidation was determined at 10 μM aglycone concentration.

conjugated by HLM to about 20−25% of the rate exhibited by HIM. Compound **1** was glucuronidated by HLM at about 40% of the corresponding rate in HIM (Figure 6). Kinetics of compound **6** glucuronidation confirmed that glucuronidation is more specific in intestinal than hepatic microsomes, as its K_m value was 82 (49−115) μM and V_{max} 4.0 (2.6−5.4) μmol/ (min* g prot) in human hepatic microsomes and its K_m value was 12 (8.8−15) μM and V_{max} 3.5 (2.9−4.0) μmol/(min * g prot) in human intestinal microsomes. Adding albumin to the incubation mixture increased both K_m and V_{max} values in instestinal microsomes. In sharp contrast to glucuronidation of the new 7-hydroxycoumarin derivatives, the "parent compound" 7-hydroxycoumarin, as well as HFC, were glucuronidated to much higher rates by HLM than by HIM (Figure 6).

Effect of H210 to M210 Mutation on UGT1A10 Activity. The developed model, by being explicit about the role of certain key UGT1A10 residues in the binding of the designed 7-hydroxycoumarin derivatives, also allowed testing it by mutagenesis. Hence, we prepared mutant 1A10-H210M, a mutant of UGT1A10 in which H210 was changed to methionine, the corresponding residue in UGT1A1. The mutant was expressed in insect cells, similarly to all the other recombinant UGTs that were used in this study, and its activity toward the new 7-hydroxycoumarin derivatives was tested. This was combined with kinetic analyses of the glucuronidation of these compounds by both UGT1A10 and UGT1A10-H201M, side by side (Figure 7 and Table 1). The results demonstrated clear effects of the mutation on the glucuronidation kinetics of most compounds. A decrease of V_{max} occurred in all but **2** and **6**, in which the K_m values were considerably increased (Table 1). Changes in the K_m values of the other compounds were variable; however, an increase in the case of HFC, no change for **5**, and a decrease in **1**, **3**, and **4** (Table 1).

■ DISCUSSION

In this study, we constructed predictive homology models for the human UGT1A enzymes in order to design selective substrates for them. The models indicated several key characteristics that differ among the active sites of individual UGT1A enzymes. These differences were exploited for the design and synthesis of six new 3-substituted 7-hydroxycoumarin derivatives. All the new compounds were good substrates for UGT1A10 and four of them, namely **2**, **4**, **5**, and **6**, were selective substrates for this UGT (Figure 5). Among the clear advantages of these new 7-hydroxcoumarin derivatives as UGT1A10 substrates are their extensive fluorescence and simple synthesis from low-cost starting materials that make

Figure 7. Michaelis−Menten kinetics of UGT1A10 (open circle) and UGT1A10 mutant (closed circle) catalyzed 7-hydroxyl glucuronidation for the 3-substituted 7-hydroxycoumarins (**1**−**7** and HFC). The data are from one experiment, and the analyzed data are shown in Table 1.

Table 1. Michaelis−Menten (MM) Kinetic Constants of UGT1A10- and UGT1A10mutant-Catalyzed 7-Hydroxyl Glucuronidation for the 3-Substituted 7-Hydroxycoumarins

compd	UGT1A10				UGT1A10mutant			
	K_m (95% confidence interval) (μM)	V_{max} (95% confidence interval) (μmol/min/g prot)	V_{max}/K_m (L/min/g prot)	nonlinear MM model R^2	K_m (95% confidence interval) (μM)	V_{max} (95% confidence interval) (μmol/min/g prot)	V_{max}/K_m (L/min/g prot)	nonlinear MM model R^2
1	10.7 (1.7–19.6)	31 (17–46)	2.9	0.9813	3.5 (0–9.3)	8.7 (3.4–14)	2.5	0.8726
2	2.2 (1–3.4)	5.9 (4.9–6.9)	2.7	0.9788	7.8 (5.5–10.2)	6.3 (5.4–7.3)	0.81	0.9965
3	8.3 (0–31.4)	22.3 (0–53.4)	2.7	0.8028	3.8 (0.97–6.7)	9.5 (6.8–12.3)	2.5	0.9695
4	2.8 (0–6.5)	15.3 (7.3–24)	5.5	0.9215	4.7 (1.5–7.9)	6.5 (4.4–8.6)	1.4	0.9846
5	2.8 (0.7–4.9)	16.2 (12–20)	5.8	0.9571	4.6 (1.5–7.8)	5.9 (4.1–7.6)	1.3	0.9750
6	7.0 (3.1–11)	5.3 (3.7–8)	0.76	0.9879	26.8 (0–80)	6.3 (0–15.9)	0.24	0.9733
HFC	14 (10.7–17.3)	15.1 (12.7–17.5)	1.1	0.9991	32.9 (0–102)	11.7 (0–31.7)	0.36	0.9781

them suitable for quick and convenient activity measurements in a high-throughput format.

Currently, no high-resolution, or even low-resolution, structure of a full-length UGT is available from X-ray crystallography or cryo-electron microscopy. Although their 3D-model construction is challenging, homology models of UGT1A1, 1A3, 1A4, 1A5, 1A6, 1A7, 1A8, 1A9, and 1A10 were constructed in this study. These models suggested that the active site of UGT1A10 is sufficiently different from the other UGTs to enable design and synthesis of selective substrates for it.

We had two goals in designing these substrates: they should be selective for UGT1A10, and determination of their glucuronidation should be based on easily measurable fluorescence change during the assay. As a starting point, the 7-hydroxycoumarin scaffold was selected, since it is a common UGT substrate and its derivatives have intense fluorescence.[44] In addition, 7-hydroxycoumarin and 7-hydroxy-4-(trifluoromethyl)coumarin have both been shown to be glucuronidated by several different UGTs.[25] The hydroxyl group of 7-hydroxycoumarin was oriented toward the UDPGA in the active site of the UGT1A10 model, indicating that there is space in the active site for an additional six or five ring substituent at position C3 of the 7-hydroxycoumarin scaffold (Figure 2). Thus, we synthesized six new UGT1A10 substrates, all of which were 7-hydroxycoumarin derivatives with various substituents at position C3 of the coumarin scaffold.

All of the new 7-hydroxycoumarin derivatives are highly fluorescent, and their fluorescence decreases upon enzymatic glucuronidation (see Figure 4 for an example with compound **6**). A concern in these assays was nonlinearity of the fluorescence at substrate concentrations above 20 μM. However, fluorescence changes of these substrates were selective, sensitive and quantitative enough for measurements below this concentration limit, as the amount of UGT1A10 enzyme could be adjusted to yield linear glucuronidation rates.

The selectivities of the 7-hydroxycoumarin derivatives for UGT1A10 differed substantially. While some of them, such as **2**, **4**, and **6**, were highly selective, others, like **1** and **3**, were also glucuronidated by UGT1A1 at considerable rates (Figure 5). Previous studies have shown that fluorescent derivatives of N-butyl-4-phenyl-1,8-naphthalimide are UGT1A1 selective substrates.[45]

Since UGT1A10 is an extrahepatic enzyme expressed at high levels in the intestine,[11] the results obtained with recombinant UGTs could be tested with HLM and HIM. Although these microsomal preparations contain several different UGTs each, HLM lacks a functional UGT1A10 while HIM contains it.[11] Indeed, the glucuronidation rates in HLM of **4** and **6**, as well as **2**, were very low or below the detection limit. Glucuronidation of **1**, **3** and **5** took place in HLM at rates up to 40% of the rates in HIM (Figure 6). These results also suggest which of the new compounds would be most useful for studies on UGT1A10 activity in samples from human tissues that express or may express this enzyme. The current results point at **2**, **4** and **6** as good candidates, **2** due to the lack of detectable activity in HLM, **4** based on the combination of high rate with high selectivity, and **6** due to its high selective fluorescence intensity (Figure 6 and Supplementary Figures S1 and S2). It may be added here that we recently reported that the commonly used commercial UGT1A10 has low activity.[13] In this study we have used the UGT1A10-H preparation, not the commercial UGT1A10. Researches should not expect to get similar results

to those reported here for UGT1A10 when working with the commonly used commercial UGT1A10. On the other hand, experiments with commercial HIM and HLM are expected to reproduce the current results.

It is interesting to understand why **6** has a remarkably lower (less than 30%) intrinsic glucuronidation clearance (V_{max}/K_m) by UGT1A10 than the other new 7-hydroxycoumarin derivatives, **1−5** (Table 1). Examination of the substituents at C3 suggests that **6** with a triazole derivative at C3 was glucuronidated less efficiently than derivatives containing hydrophobic or other types of hydrophilic substituents. The reason for this could be that the triazole causes stronger interactions than the other substituents in the active site of UGT1A10, resulting in slower release of the formed glucuronide.

The present modeling work indicates specific residues that are expected to lead to substrate selectivity of UGT1A10. We have tested one of the predicted residues, H210, by changing it to M210, the corresponding residue in UGT1A1, expressed the mutant UGT1A10-M210 and studied the glucuronidation kinetics of all the new compounds by both the wildtype (i.e., UGT1A10-H210) and mutant enzymes. The results revealed changes in the K_m and/or V_{max} values of the glucuronidation kinetics of all the compounds, but with clear differences among them (Figure 7 and Table 1). Another unexpected observation was the effect of the mutation on HFC glucuronidation rate, which may suggest that the effect of the mutation is larger than expected by the model, or that substrate(s) binding by UGT1A10 (also) involves an induced fit mechanism.

In conclusion, in this study, 3D molecular models of the UGT1As were constructed and used for the design and synthesis of six new fluorescent UGT substrates. A new multiwell-based method that takes advantage of the fluorescence of the compounds and their fluorescence decrease upon glucuronidation was established to measure glucuronidation rates. Of the new compounds, 4-(dimethylamino)phenyl (**4**) and triazole (**6**) C3-substituted 7-hydroxycoumarins appeared to be the most selective substrates for UGT1A10, an important and often underestimated extrahepatic human UGT. It is concluded that the selectivity of the new coumarin derivatives for UGT1A10 depends on the chemical character of their substituent at C3. These new compounds should enable better, faster, and easier determination of UGT1A10 activity in tissues than was earlier possible. In addition, their further chemical modification could stimulate the development of new tools to explore the active site of different UGT enzymes.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.molpharmaceut.7b00871.

Supplementary figures and tables (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: risto.juvonen@uef.fi (in vitro).
*E-mail: olli.pentikainen@utu.fi (modeling and synthesis).

ORCID ◉

Risto O. Juvonen: 0000-0001-9240-7673
Moshe Finel: 0000-0003-1775-854X
Olli T. Pentikäinen: 0000-0001-7188-4016

Author Contributions

R.O.J, H.R., and O.T.P designed the study. S.R., S.N., and O.T.P. performed the modeling. S.K. and J.H. performed the synthesis of molecules. A.P. performed HPLC−MS analysis. R.O.J. performed the experimental measurements. M.F. provided the UGT enzymes and with J.T. prepared the mutant. All the authors contributed to writing of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Jones, C. R.; Hatley, O. J.; Ungell, A. L.; Hilgendorf, C.; Peters, S. A.; Rostami-Hodjegan, A. Gut Wall Metabolism. Application of Pre-Clinical Models for the Prediction of Human Drug Absorption and First-Pass Elimination. *AAPS J.* **2016**, *18* (3), 589−604.

(2) Gonzalez, F. J.; Coughtrie, M.; Tukey, R. H., Drug Metabolism. In *Goodman & Gilmans's The Pharmacological Basis of Therapeutics*, 12th ed. ed.; Brunton, L. L., Chabner, B., Knollman, B., Eds.; McGraw-Hill: New York, 2011; pp 123−143.

(3) Mizuma, T. Intestinal glucuronidation metabolism may have a greater impact on oral bioavailability than hepatic glucuronidation metabolism in humans: A study with raloxifene, substrate for UGT1A1, 1A8, 1A9, and 1A10. *Int. J. Pharm.* **2009**, *378* (1−2), 140−141.

(4) Ritter, J. K. Intestinal UGTs as potential modifiers of pharmacokinetics and biological responses to drugs and xenobiotics. *Expert Opin. Drug Metab. Toxicol.* **2007**, *3* (1), 93−107.

(5) Wu, B. J.; Kulkarni, K.; Basu, S.; Zhang, S. X.; Hu, M. First-Pass Metabolism via UDP-Glucuronosyltransferase: a Barrier to Oral Bioavailability of Phenolics. *J. Pharm. Sci.* **2011**, *100* (9), 3655−3681.

(6) Testa, B.; Kramer, S. D. The biochemistry of drug metabolism–an introduction: part 5. Metabolism and bioactivity. *Chem. Biodiversity* **2009**, *6* (5), 591−684.

(7) Kaivosaari, S.; Finel, M.; Koskinen, M. N-glucuronidation of drugs and other xenobiotics by human and animal UDP-glucuronosyltransferases. *Xenobiotica* **2011**, *41* (8), 652−69.

(8) Rowland, A.; Miners, J. O.; Mackenzie, P. I. The UDP-glucuronosyltransferases: their role in drug metabolism and detoxification. *Int. J. Biochem. Cell Biol.* **2013**, *45* (6), 1121−32.

(9) Court, M. H.; Zhang, X. L.; Ding, X. X.; Yee, K. K.; Hesse, L. M.; Finel, M. Quantitative distribution of mRNAs encoding the 19 human UDP-glucuronosyltransferase enzymes in 26 adult and 3 fetal tissues. *Xenobiotica* **2012**, *42* (3), 266−277.

(10) Ohno, S.; Nakajin, S. Determination of mRNA Expression of Human UDP-Glucuronosyltransferases and Application for Localization in Various Human Tissues by Real-Time Reverse Transcriptase-Polymerase Chain Reaction. *Drug Metab. Dispos.* **2009**, *37* (1), 32−40.

(11) Sato, Y.; Nagata, M.; Tetsuka, K.; Tamura, K.; Miyashita, A.; Kawamura, A.; Usui, T. Optimized Methods for Targeted Peptide-Based Quantification of Human Uridine 5 '-Diphosphate-Glucuronosyltransferases in Biological Specimens Using Liquid Chromatography-Tandem Mass Spectrometry. *Drug Metab. Dispos.* **2014**, *42* (5), 885−889.

(12) Hu, D. G.; Meech, R.; McKinnon, R. A.; Mackenzie, P. I. Transcriptional regulation of human UDP-glucuronosyltransferase genes. *Drug Metab. Rev.* **2014**, *46* (4), 421−58.

(13) Troberg, J.; Jarvinen, E.; Ge, G. B.; Yang, L.; Finel, M. UGT1A10 Is a High Activity and Important Extrahepatic Enzyme: Why Has Its Role in Intestinal Glucuronidation Been Frequently Underestimated? *Mol. Pharmaceutics* **2017**, *14*, 2875.

(14) Oda, S.; Kato, Y.; Hatakeyama, M.; Iwamura, A.; Fukami, T.; Kume, T.; Yokoi, T.; Nakajima, M. Evaluation of expression and glycosylation status of UGT1A10 in Supersomes and intestinal epithelial cells with a novel specific UGT1A10 monoclonal antibody. *Drug Metab. Dispos.* **2017**, *45* (9), 1027−1034.

(15) Hoglund, C.; Sneitz, N.; Radominska-Pandya, A.; Laakonen, L.; Finel, M. Phenylalanine 93 of the human UGT1A10 plays a major role in the interactions of the enzyme with estrogens. *Steroids* **2011**, *76* (13), 1465−1473.

(16) Kallionpaa, R. A.; Jarvinen, E.; Finel, M. Glucuronidation of estrone and 16alpha-hydroxyestrone by human UGT enzymes: The key roles of UGT1A10 and UGT2B7. *J. Steroid Biochem. Mol. Biol.* **2015**, *154*, 104−111.

(17) Itaaho, K.; Mackenzie, P. I.; Ikushiro, S.; Miners, J. O.; Finel, M. The configuration of the 17-hydroxy group variably influences the glucuronidation of beta-estradiol and epiestradiol by human UDP-glucuronosyltransferases. *Drug Metab. Dispos.* **2008**, *36* (11), 2307−15.

(18) Sneitz, N.; Vahermo, M.; Mosorin, J.; Laakkonen, L.; Poirier, D.; Finel, M. Regiospecificity and stereospecificity of human UDP-glucuronosyltransferases in the glucuronidation of estriol, 16-epiestriol, 17-epiestriol, and 13-epiestradiol. *Drug Metab. Dispos.* **2013**, *41* (3), 582−91.

(19) Itaaho, K.; Court, M. H.; Uutela, P.; Kostiainen, R.; Radominska-Pandya, A.; Finel, M. Dopamine is a low-affinity and high-specificity substrate for the human UDP-glucuronosyltransferase 1A10. *Drug Metab. Dispos.* **2009**, *37* (4), 768−75.

(20) Raunio, H.; Kuusisto, M.; Juvonen, R. O.; Pentikainen, O. T. Modeling of interactions between xenobiotics and cytochrome P450 (CYP) enzymes. *Front. Pharmacol.* **2015**, *6*, 123.

(21) Kirchmair, J.; Goller, A. H.; Lang, D.; Kunze, J.; Testa, B.; Wilson, I. D.; Glen, R. C.; Schneider, G. Predicting drug metabolism: experiment and/or computation? *Nat. Rev. Drug Discovery* **2015**, *14* (6), 387−404.

(22) Dai, Z. R.; Feng, L.; Jin, Q.; Cheng, H.; Li, Y.; Ning, J.; Yu, Y.; Ge, G. B.; Cui, J. N.; Yang, L. A practical strategy to design and develop an isoform-specific fluorescent probe for a target enzyme: CYP1A1 as a case study. *Chem. Sci.* **2017**, *8* (4), 2795−2803.

(23) Dai, Z. R.; Ge, G. B.; Feng, L.; Ning, J.; Hu, L. H.; Jin, Q.; Wang, D. D.; Lv, X.; Dou, T. Y.; Cui, J. N.; Yang, L. A Highly Selective Ratiometric Two-Photon Fluorescent Probe for Human Cytochrome P450 1A. *J. Am. Chem. Soc.* **2015**, *137* (45), 14488−95.

(24) Miley, M. J.; Zielinska, A. K.; Keenan, J. E.; Bratton, S. M.; Radominska-Pandya, A.; Redinbo, M. R. Crystal structure of the cofactor-binding domain of the human phase II drug-metabolism enzyme UDP-glucuronosyltransferase 2B7. *J. Mol. Biol.* **2007**, *369* (2), 498−511.

(25) Rahikainen, T.; Hakkinen, M. R.; Finel, M.; Pasanen, M.; Juvonen, R. O. A high throughput assay for the glucuronidation of 7-hydroxy-4-trifluoromethylcoumarin by recombinant human UDP-glucuronosyltransferases and liver microsomes. *Xenobiotica* **2013**, *43* (10), 853−61.

(26) Modolo, L. V.; Li, L.; Pan, H.; Blount, J. W.; Dixon, R. A.; Wang, X. Crystal structures of glycosyltransferase UGT78G1 reveal the molecular basis for glycosylation and deglycosylation of (iso)-flavonoids. *J. Mol. Biol.* **2009**, *392* (5), 1292−302.

(27) Hiromoto, T.; Honjo, E.; Tamada, T.; Noda, N.; Kazuma, K.; Suzuki, M.; Kuroki, R. Crystal structure of UDP-glucose:anthocyanidin 3-O-glucosyltransferase from Clitoria ternatea. *J. Synchrotron Radiat.* **2013**, *20*, 894−8.

(28) Offen, W.; Martinez-Fleites, C.; Yang, M.; Kiat-Lim, E.; Davis, B. G.; Tarling, C. A.; Ford, C. M.; Bowles, D. J.; Davies, G. J. Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *EMBO J.* **2006**, *25* (6), 1396−405.

(29) Lehtonen, J. V.; Still, D. J.; Rantanen, V. V.; Ekholm, J.; Bjorklund, D.; Iftikhar, Z.; Huhtala, M.; Repo, S.; Jussila, A.; Jaakkola, J.; Pentikainen, O.; Nyronen, T.; Salminen, T.; Gyllenberg, M.; Johnson, M. S. BODIL: a molecular modeling environment for structure-function analysis and drug design. *J. Comput.-Aided Mol. Des.* **2004**, *18* (6), 401−19.

(30) Johnson, M. S.; Overington, J. P. A structural basis for sequence comparisons. An evaluation of scoring methodologies. *J. Mol. Biol.* **1993**, *233* (4), 716−38.

(31) Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234* (3), 779−815.

(32) Niinivehmas, S. P.; Salokas, K.; Latti, S.; Raunio, H.; Pentikainen, O. T. Ultrafast protein structure-based virtual screening with Panther. *J. Comput.-Aided Mol. Des.* **2015**, *29* (10), 989−1006.

(33) Korb, O.; Stutzle, T.; Exner, T. E. Empirical scoring functions for advanced protein-ligand docking with PLANTS. *J. Chem. Inf. Model.* **2009**, *49* (1), 84−96.

(34) Niinivehmas, S. P.; Manivannan, E.; Rauhamaki, S.; Huuskonen, J.; Pentikainen, O. T. Identification of estrogen receptor alpha ligands with virtual screening techniques. *J. Mol. Graphics Modell.* **2016**, *64*, 30−9.

(35) Kirkiacharian, S.; Lormier, A. T.; Resche-Rigon, M.; Bouchoux, F.; Cerede, E. [Synthesis and binding affinity of 3-aryl-7-hydroxycoumarins to human alpha and beta estrogen receptors]. *Ann. Pharm. Fr.* **2003**, *61* (1), 51−6.

(36) Buuhoi, N. P.; Ekert, B.; Royer, R. Bz-Hydroxylated-3-Aryl-Coumarins and 3,4-Diaryl-Coumarins. *J. Org. Chem.* **1954**, *19* (9), 1548−1552.

(37) Kurkela, M.; Garcia-Horsman, J. A.; Luukkanen, L.; Morsky, S.; Taskinen, J.; Baumann, M.; Kostiainen, R.; Hirvonen, J.; Finel, M. Expression and characterization of recombinant human UDP-glucuronosyltransferases (UGTs). UGT1A9 is more resistant to detergent inhibition than other UGTs and was purified as an active dimeric enzyme. *J. Biol. Chem.* **2003**, *278* (6), 3536−44.

(38) Kuuranne, T.; Aitio, O.; Vahermo, M.; Elovaara, E.; Kostiainen, R. Enzyme-assisted synthesis and structure characterization of glucuronide conjugates of methyltestosterone (17 alpha-methylan-drost-4-en-17 beta-ol-3-one) and nandrolone (estr-4-en-17 beta-ol-3-one) metabolites. *Bioconjugate Chem.* **2002**, *13* (2), 194−9.

(39) Sneitz, N.; Court, M. H.; Zhang, X.; Laajanen, K.; Yee, K. K.; Dalton, D.; Ding, X.; Finel, M. Human UDP-glucuronosyltransferase UGT2A2: cDNA construction, expression, and functional character-ization in comparison with UGT2A1 and UGT2A3. *Pharmacogenet. Genomics* **2009**, *19* (12), 923−34.

(40) Kurkela, M.; Patana, A. S.; Mackenzie, P. I.; Court, M. H.; Tate, C. G.; Hirvonen, J.; Goldman, A.; Finel, M. Interactions with other human UDP-glucuronosyltransferases attenuate the consequences of the Y485D mutation on the activity and substrate affinity of UGT1A6. *Pharmacogenet. Genomics* **2007**, *17* (2), 115−26.

(41) Lang, M. A.; Gielen, J. E.; Nebert, D. W. Genetic evidence for many unique liver microsomal P-450-mediated monooxygenase activities in heterogeneic stock mice. *J. Biol. Chem.* **1981**, *256* (23), 12068−75.

(42) Kuuranne, T.; Kurkela, M.; Thevis, M.; Schanzer, W.; Finel, M.; Kostiainen, R. Glucuronidation of anabolic androgenic steroids by recombinant human UDP-glucuronosyltransferases. *Drug Metab. Dispos.* **2003**, *31* (9), 1117−24.

(43) Zhang, H.; Tolonen, A.; Rousu, T.; Hirvonen, J.; Finel, M. Effects of cell differentiation and assay conditions on the UDP-glucuronosyltransferase activity in Caco-2 cells. *Drug Metab. Dispos.* **2011**, *39* (3), 456−64.

(44) Lavis, L. D.; Raines, R. T. Bright building blocks for chemical biology. *ACS Chem. Biol.* **2014**, *9* (4), 855−66.

(45) Lv, X.; Feng, L.; Ai, C. Z.; Hou, J.; Wang, P.; Zou, L. W.; Cheng, J.; Ge, G. B.; Cui, J. N.; Yang, L. A Practical and High-Affinity Fluorescent Probe for Uridine Diphosphate Glucuronosyltransferase 1A1: A Good Surrogate for Bilirubin. *J. Med. Chem.* **2017**, *60* (23), 9664−9675.

Supplementary information

Molecular docking-based design and development of a highly selective probe substrate for

UDP-glucuronosyltransferase 1A10

[1]Risto O. Juvonen, [2]Sanna Rauhamäki, [2,3]Sami Kortet, [2]Sanna Niinivehmas, [4]Johanna Troberg, [1]Aleksanteri Petsalo, [3]Juhani Huuskonen, [1]Hannu Raunio, [4]Moshe Finel, [2,5]Olli T. Pentikäinen

1 School of Pharmacy, Faculty of Health Sciences, University of Eastern Finland, Box 1627, FI-70211 Kuopio, Finland

2 University of Jyvaskyla, Department of Biological and Environmental Science, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland

3 University of Jyvaskyla, Department of Chemistry, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland

4 Division of Pharmaceutical Chemistry and Technology, Faculty of Pharmacy, University of Helsinki, P.O. Box 56, FI-00014 University of Helsinki, Finland

5 Institute of Biomedicine, Faculty of Medicine, University of Turku, FI-20014 University of Turku, Finland
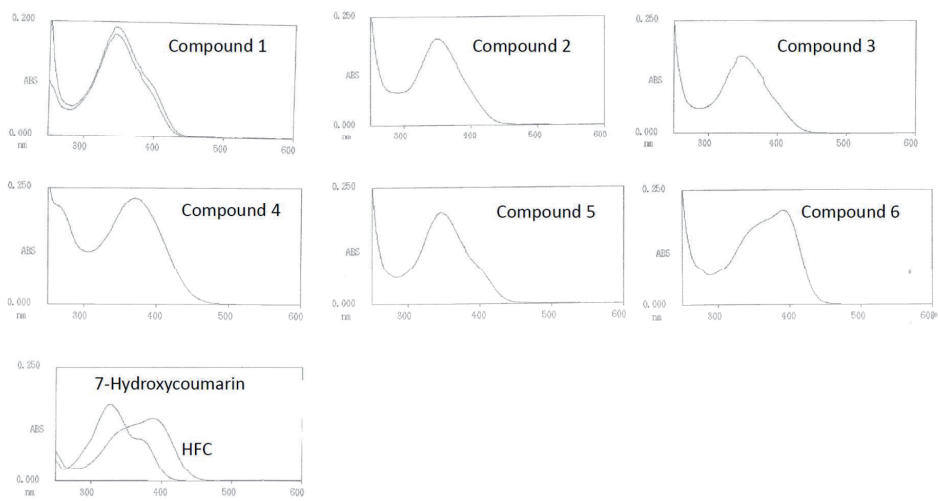
Figure S1. Absorbance spectra of 7-hydroxycoumarin derivatives. Absorbance of 10 µM coumarin derivatives in 100 mM phosphate, 10 % dimethylsulfoxide were measured from 600 nm to 200 nm.
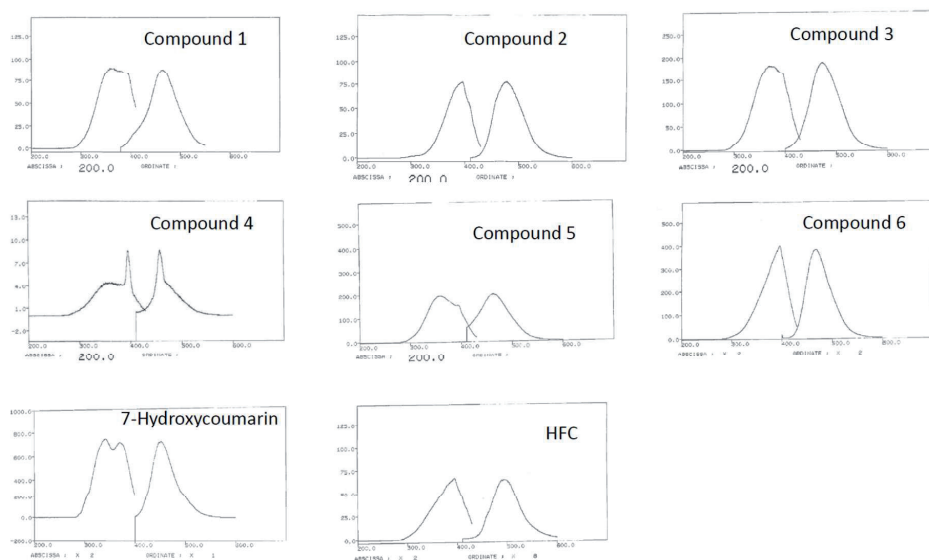
Figure S2. Excitation and emission fluorescence spectra of 100 nM 7-hydroxycoumarin derivatives in 100 mM phosphate were measured at maximums of emission and excitation, respectively.

Table S1. Absorbance and fluorescence of 7-hydroxycoumarin derivatives.

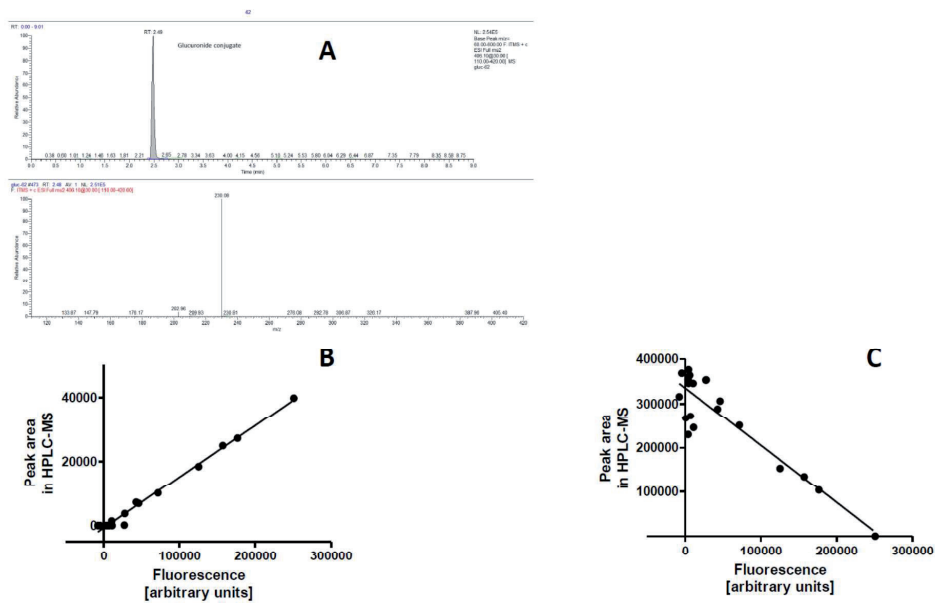| Compound	Amax [nm]	ε abs [cm-1*M-1]	Fluorescence EXmax [nm]	Fluorescence EMmax [nm]
1	344	19 100	360 -396	466
2	352	20 000	395	475
3	350	17 300	368	472
4	370	22 900	372, 396	457
5	347	19 500	358	462
6	392	20 500	396	494

Figure S3. Identification of 7-hydroxy-3-triazolecoumarin glucuronide conjugate and its correlation with decrease in fluorescence. Panel A shows the HPLC chromatogram and the mass spectrum of the glucuronide conjugate formed in 1 h incubation with UGT1A10, 10 μM 7-hydroxy-3-triazolecoumarin and 1 mM UDPGA. Panel B shows the correlation of decrease in fluorescence of 7-hydroxy-3-triazolecoumarin glucuronide with its peak area in HPLC-MS analysis (y = 0.159 * x -693, r^2 0.9911) and panel C with the peak area of 7-hydroxy-3-triazolecoumarin aglycone (y = -1.298 * x +335485, r^2 0.8180) in HPLC-MS analysis in glucuronidation reaction with different human recombinant UGTs.

4