Finding New Molecular Targets of Familiar Natural Products Using In Silico Target Prediction

Fabian Mayr[1], Gabriele Möller[2], Ulrike Garscha[3], Jana Fischer[3], Patricia Rodríguez Castaño[4,5], Silvia G. Inderbinen[6], Veronika Temml[1], Birgit Waltenberger[1], Stefan Schwaiger[1], Rolf W. Hartmann[7,8], Christian Gege[9], Stefan Martens[10], Alex Odermatt[6], Amit V. Pandey[4,5], Oliver Werz[11], Jerzy Adamski[2,12,13], Hermann Stuppner[1] and Daniela Schuster[14,15,]*

[1] Institute of Pharmacy/Pharmacognosy, Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innrain 80/82, 6020 Innsbruck, Austria

[2] Research Unit Molecular Endocrinology and Metabolism, Helmholtz Zentrum München, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany

[3] Department of Pharmaceutical/Medicinal Chemistry, Institute of Pharmacy, University Greifswald, Friedrich-Ludwig-Jahn-Straße 17, 17489 Greifswald, Germany

[4] Pediatric Endocrinology, Diabetology, and Metabolism, University Children's Hospital Bern, Freiburgstrasse 15, 3010 Bern, Switzerland

[5] Department of Biomedical Research, University of Bern, Freiburgstrasse 15, 3010 Bern, Switzerland

[6] Division of Molecular and Systems Toxicology, Department of Pharmaceutical Sciences, University of Basel, Klingelbergstrasse 50, 4056 Basel, Switzerland

[7] Helmholtz Institut für Pharmazeutische Forschung, Department für Drug Design und Optimierung, Campus E8.1, 66123 Saarbrücken, Germany

[8] Universität des Saarlandes, Pharmazeutische und Medizinische Chemie, Campus E8.1, 66123 Saarbrücken

[9] Universität Heidelberg, Medizinische Chemie, Im Neuenheimer Feld 364, 69120 Heidelberg, Germany

[10] Research and Innovation Centre, Fondazione Edmund Mach (FEM), Via Mach 1, 38010 San Michele all'Adige, Italy

[11] Department of Pharmaceutical/Medicinal Chemistry, Institute of Pharmacy, Friedrich-Schiller-University Jena, Philosophenweg 14, 07743 Jena, Germany

[12] Lehrstuhl für Experimentelle Genetik, Technische Universität München, Emil-Erlenmeyer-Forum 5, 85356 Freising-Weihenstephan, Germany

[13] Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, 8 Medical Drive, Singapore 117597, Singapore

[14] Institute of Pharmacy, Department of Pharmaceutical and Medicinal Chemistry, Paracelsus Medical University Salzburg, Strubergasse 21, 5020 Salzburg, Austria

[15] Institute of Pharmacy/Pharmaceutical Chemistry, Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innrain 80/82, 6020 Innsbruck, Austria

ABSTRACT

Natural products comprise a rich reservoir for innovative drug leads and are a constant source of bioactive compounds. To find pharmacological targets for new or already known natural products using modern computer-aided methods is a current endeavor in drug discovery. Nature's treasures, however, could be used more effectively. Yet, reliable pipelines for large scale target prediction of natural products are still rare. We have developed an in silico workflow consisting of four independent, stand-alone target prediction tools and evaluated its performance on dihydrochalcones (DHCs) – a well-known class of natural products. Thereby, we revealed four previously unreported protein targets for DHCs, namely 5-lipoxygenase, cyclooxygenase-1, 17β-hydroxysteroid dehydrogenase 3, and aldo-keto reductase 1C3. Moreover, we provide a thorough strategy on how to perform computational target prediction and guidance on using the respective tools.

INTRODUCTION

Finding new chemical entities that alter biological response – the quintessence of drug discovery – is a constant endeavor in pharmaceutical science. In contrast, the need for novel, improved clinical candidates has also remained consistently high, urging drug discovery scientists to explore fresh ground. The integration of chemoinformatic and bioinformatic tools into drug discovery in the early 1990s and the recent advances in big data handling have leveraged access to a myriad of massive public datasets (Campbell et al., 2018; Chen et al., 2017) and powerful tools, e.g. virtual screening (VS) (Sliwoski et al., 2014). In the past decade, the concept of drug repurposing has emerged as an attractive strategy to rededicate approved drugs or partially developed compounds to new molecular targets (Aronson, 2007; Ashburn et al., 2004). This development is, next to the intention to reduce R&D costs, also owed to advances in computational chemistry (Hurle et al., 2013). The latter can be achieved by a so-called 'inverse VS' utilizing techniques like 2D-similarity searches (Keiser et al., 2009), 3D-similarity searches (Rush et al., 2005), and pharmacophore-based VS (Schuster, 2010; Steindl et al., 2006). Many of such tools have been made public in the past decade (Cereto-Massagué et al., 2015; Huang et al., 2001; Sydow et al., 2019), aiming to boost both drug repurposing efforts and drug discovery as a whole.

Natural products are remarkable in many regards, particularly for being the main source of drugs in the past and nowadays by serving as a source for innovative leads (Newman et al., 2016). Natural products bear privileged structural features that were "shaped" by evolution, yielding compounds that that can serve as promising starting points for drug development (Harvey, 2008; Koehn et al., 2005). Further, natural products often show polypharmacological properties, interacting with more than one target (Koeberle et al., 2014). The two groups around Gisbert Schneider and Stuart L. Schreiber found that natural products are more likely to act as true polypharmacological agents rather than unspecific binders – a property instead associated with synthetic compounds (Clemons et al., 2010; Rodrigues et al., 2016).

Based on this concept, tools or workflows, which allow for the accurate prediction of new molecular targets for natural products, are of great interest. We here propose a workflow to specifically search for yet unreported protein targets of known compounds, using a combination of in silico and in vitro methods (Figure 1). The compounds of interest are virtually screened against our in-house resources, representing a panel of 39 drug targets expressed as 387

pharmacophore models (a comprehensive list of all models is provided in Supplementary Information S-1). Additionally, the compounds are subjected to three independent open access target prediction servers in parallel. The results generated by these four diverse methods are combined and those with the highest degree of consent (protein targets predicted by several methods independently of each other) are selected for in vitro evaluation. Moreover, already known protein targets are excluded from further investigations by checking their appearance in PubChem (https://pubchem.ncbi.nlm.nih.gov/), a pertinent open knowledge base for bioactivities (Kim et al., 2018).

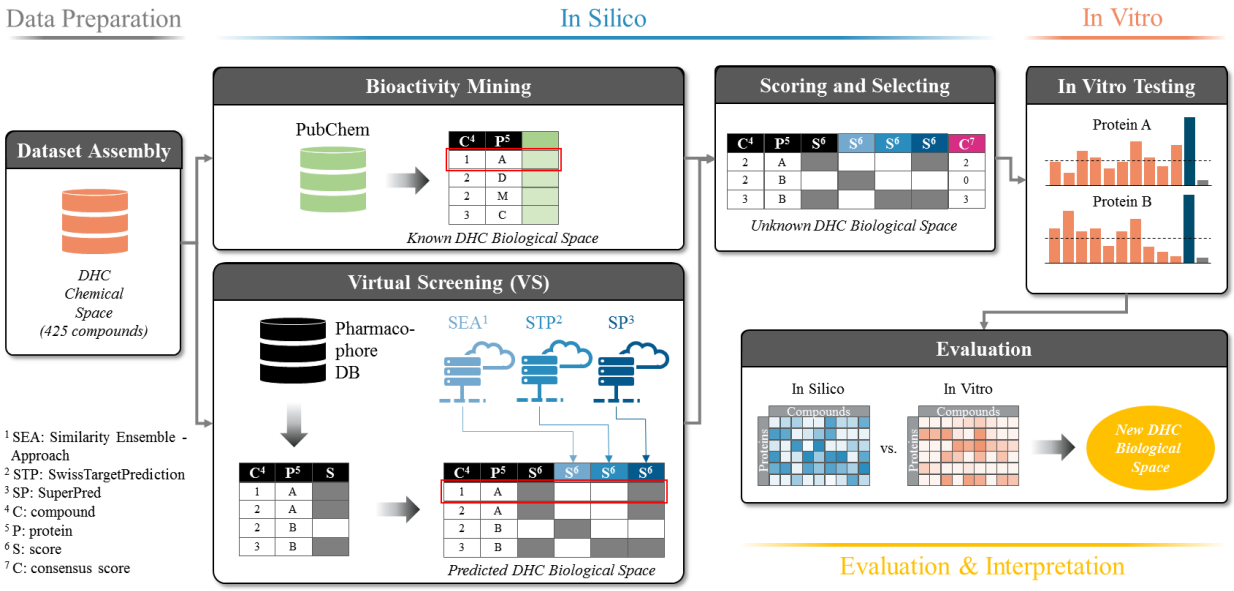


Figure 1. Workflow of the dihydrochalcone (DHC) target prediction campaign. The dataset is assembled (DHC chemical space) and used to retrieve corresponding bioactivity data from PubChem (known DHC biological space) and as input to inverse VS. First, the DHC chemical space is mapped onto the Pharmacophore DB and the resulting matrix extended by the predictions of three individual target prediction servers Similarity Ensemble Approach (SEA), SwissTargetPrediction (STP), and SuperPred (SP), resulting in the predicted DHC biological space. Activities already known from PubChem (the known DHC biological space) are then removed from the predicted DHC biological space and the reduced matrix scored according to consensus predictions of ligand-target interactions (unknown DHC biological space). Protein targets of the unknown DHC biological space are selected according to their consensus score (CS) and chosen for in vitro biological evaluation.

We evaluated our workflow for fitness by performing a target prediction of a complete class of established natural products, namely dihydrochalcones (DHCs). DHCs are readily available from nature, since several representatives are highly accumulated in both fresh and withered leaves of apple trees (Rivière, 2016). On the other hand, phloridzin, one of the most frequently found DHCs served as lead structure for the development of sodium/glucose co-transporter 2 (SGLT2) inhibitors like dapagliflozin, approved drugs for the treatment of type 2 diabetes (Meng et al., 2008). DHCs have recently returned into focus, since their descendants, in clinical use now for about eight years, have shown therapeutic benefits that go beyond SGLT2 inhibition like e.g. in heart failure (Uthman et al., 2018). This instance points towards a high polypharmacological potential of the drugs and its parental template. However, phloridzin research has so far been focused on its anti-diabetic, anti-oxidative, and estrogenic effects. Thus, we prepared a comprehensive virtual library of DHCs (naturally occurring ones and those with modest semi-synthetic modifications), predicted and selected promising, potentially new DHC targets, and tested ten common DHCs in respective in vitro assays.

RESULTS

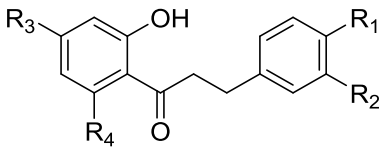
Data Basis, Curation and Technical Setup of In Silico Predictions

To realistically mirror the true diversity of DHCs, we gathered 425 DHCs from literature that were either naturally occurring or roughly resembled physicochemical properties of natural DHCs (molecular weight, ratio glycosides/aglyca, physicochemical properties). Accordingly, we called this virtual library 'DHC chemical space' (see Figure 1, Data Preparation). A panel of ten commonly found DHCs (**1** – **10**, see Table 1) that were physically available to us and intended for in vitro testing were also included in the DHC chemical space. The DHC chemical space was then screened against our historically grown pharmacophore model database (Ph-DB) and the results written to a matrix (e.g. compound 1 is predicted to act on protein A). The DHC chemical space was in parallel also subjected to the three target prediction servers SEA, STP, and SP, each of them predicting potential targets for each of the 425 DHCs. All results were combined into one matrix called 'predicted DHC biological space' (see Figure 1, 'Virtual Screening'). In addition, all of the known bioactivities of the 425 compounds in the DHC chemical space were downloaded from PubChem and the resulting matrix called 'known DHC biological space' (see Figure 1, 'Bioactivity Mining'). The known DHC biological space was additionally depicted as network, as

shown in Figure 2. Next, compound-target interactions present in the known DHC biological space were removed from the predicted DHC biological space and a CS was assigned to each prediction (see Figure 1, 'Scoring & Selecting'). The CS is an expression on how many consents a prediction has in addition to a positive prediction by our Ph-DB. We introduced the requirement of a hit with our own models since they are well-validated and for most targets, experimental testing of hits was available. For instance, if compound 1 was predicted to bind to protein A by our in-house pharmacophore models and two further target prediction servers, the CS was three. Finally, based on the CS and other criteria, the six most promising protein targets were selected and compounds **1** – **10** assayed in vitro.

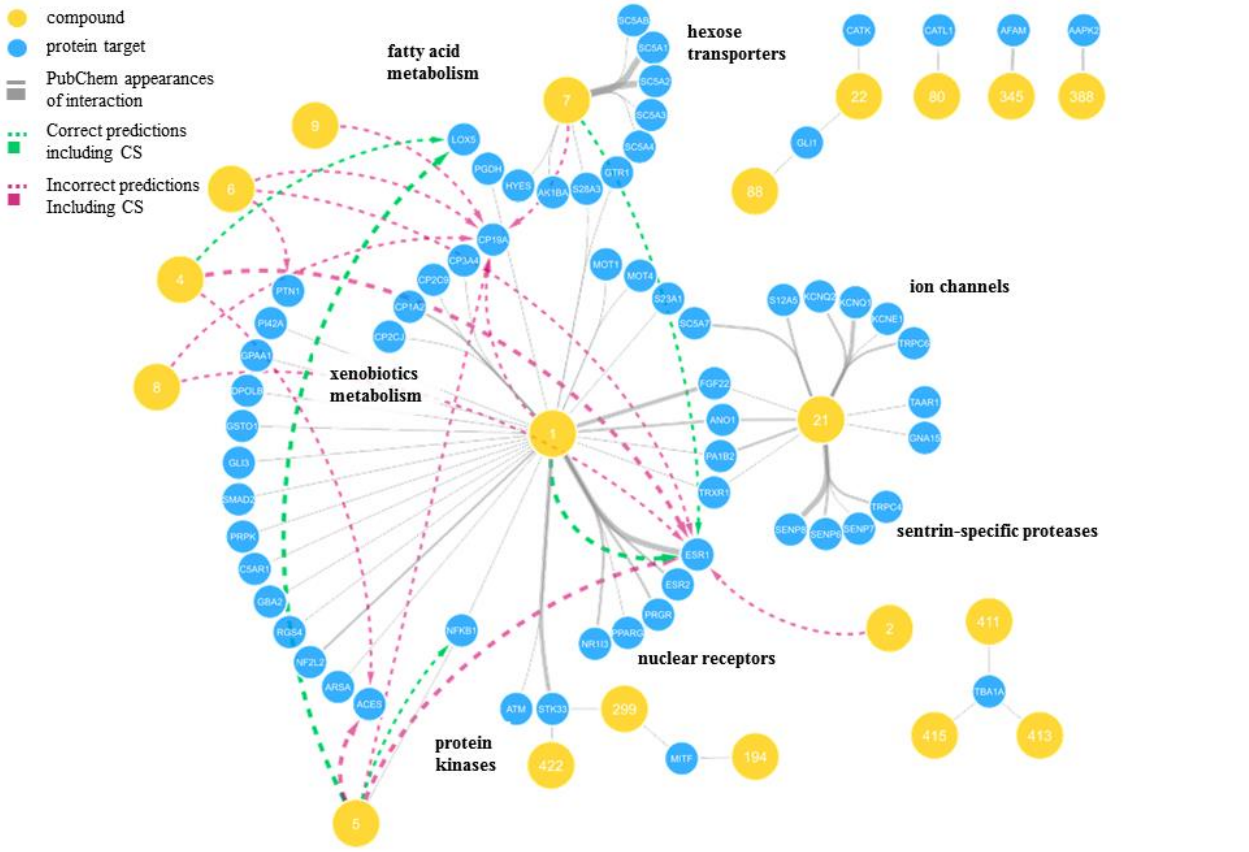
To automate the in silico part of the workflow, operations made for screening our Ph-DB, submission and reconciliation to target prediction servers, and bioactivity mining were performed via custom-made scripts. All input and output files generated in this workflow, including the scripts and a corresponding Jupyter Notebook containing all data manipulations are provided via GitHub (https://github.com/fmayr/DHC_TargetPrediction). For better clarity, the relationships and data flow are schematically shown in Supporting Information S-3.

144 Table 1. Compounds **1** – **10**, which were used for biological evaluation.



No.	Name	R_1	R_2	R_3	R_4
1	phloretin	OH	H	OH	OH
2	3-OH-phloretin	OH	OH	OH	OH
3	2´,6´-dihydroxy-4´-methoxy DHC	H	H	OMe	OH
4	asebogenin	OH	H	OMe	OH
5	calomelanen	OMe	H	OMe	OH
6	sieboldin	OH	OH	O-Glc*	OH
7	phloridzin	OH	H	OH	O-Glc*
8	trilobatin	OH	H	O-Glc*	OH
9	phloretin-2´-xyloglucoside	OH	H	OH	O-Rut†
10	neohesperidin DHC	OMe	OH	O-Neo‡	OH

145 * Glc: glucose (*O*-β-D-glucosyl). † Rut: rutinose (6-*O*-(α-L-rhamnosyl)-D-glucos-1-*O*-β-yl). ‡
146 Neo: neohesperidose (2-*O*-(α-L-rhamnosyl)-D-glucos-1-*O*-β-yl).



ACES	Acetylcholinesterase	*CP2CJ*	CYP2C19	*HYES*	Soluble epoxide hydrolase 2	*PI42A*	PIP4KII-α	*SC5A7*	High-affinity choline transp. 1
AAPK2	AMPK subunit α2	*CP3A4*	CYP3A4	*KCNE1*	Potassium volt.-gated channel E1	*PPARG*	PPAR γ	*SC5AB*	Sodium/myo-inositol cotransp. 1
AFAM	Afamin	*DPOLB*	DNA polymerase β	*KCNQ1*	Potassium volt.-gated channel KQT1	*PRGR*	Progesterone receptor	*SENP6*	Sentrin-specific protease 6
AK1BA	AKR1B10	*ESR1*	Estrogen α	*KCNQ2*	Potassium volt.-gated channel KQT2	*PRPK*	EKC/KEOPS complex sub. TP53RK	*SENP7*	Sentrin-specific protease 7
ANO1	Anoctamin-1	*ESR2*	Estrogen β	*LOX5*	5-lipoxygenase	PTN1	PTP1B	*SENP8*	Sentrin-specific protease 8
ARSA	Arylsulfatase A	*FGF22*	FGF-22	*MITF*	Microphthalmia-associated TF	*RGS4*	Regulator of G-protein signaling 4	*SMAD2*	MAD homolog 2
ATM	Serine-protein kinase	*GBA2*	β-glucosidase 2	*MOT1*	Monocarboxylate transporter 1	*S12A5*	Solute carrier 12A5	*STK33*	Serine/threonine-prot. kinase 33
C5AR1	C5aR	*GLI1*	Zinc finger protein GLI1	*MOT4*	Monocarboxylate transporter 4	*S23A1*	Solute carrier 23A1	*TAAR1*	Trace amine-assoc. receptor 1
CATK	Cathepsin K	*GLI3*	GLI3-190	*NF2L2*	NF erythroid 2-related factor 2	*S28A3*	Solute carrier 28A3	*TBA1A*	Tubulin α-1A chain
CATL1	Cathepsin L1	*GNA15*	G α-15	*NFKB1*	NF κB p105 subunit	*SC5A1*	Sodium/glucose cotransporter 1	*TRPC4*	TRPC4
CP1A2	CYP1A2	*GPAA1*	GPI anchor attachment 1 prot.	*NR1I3*	Nuclear receptor 1I3	*SC5A2*	Sodium/glucose cotransporter 2	*TRPC6*	TRPC6
CP19A	Aromatase	*GSTO1*	Glutathione S-transferase Ω-1	*PA1B2*	PAFAH subunit β	*SC5A3*	Sodium/myo-inositol cotransporter	*TRXR1*	Thioredoxin reductase 1
CP2C9	CYP2C9	*GTR1*	GTP-binding protein GTR1	*PGDH*	15-hydroxyl-PG dehydrogenase	*SC5A4*	Solute carrier 5A4		

Figure 2. Known DHC biological space illustrated as network. Shown are only DHCs with interactions reported in PubChem (grey edges), as well as the respective interactions predicted for these compounds (dashed green arrows for correct predictions and dashed magenta arrows for interactions that either proved incorrect in vitro or were not tested). Blue nodes indicate protein targets, and yellow nodes indicate compounds with respective compound numbers. Grey edges indicate known compound-target interactions, while the line thickness is proportional to the interaction weight (see Materials and Methods).

Predicted and Unknown DHC Biological Space

The final selection of targets to be evaluated in vitro was made based on four criteria (summarized for twelve frequently predicted targets in Table 2): First, interactions of **1** – **10** that were predicted by Ph-DB and had a CS of two or higher (see Figure 3) were included (Table 2, Selection Criterion I). Thereby, 5-lipoxygenase (5-LO) was highlighted for **4** (CS = 2) and **5** (CS = 3) and aromatase was highlighted for **1**, **5**, **6**, **7**, **8**, and **9** (CS = 2). Figure 3 also shows that a handful of other targets achieved high CSs, e.g. acetylcholinesterase (AChE), estrogen receptor α (ERα), protein-tyrosine phosphatase B1 (PTP1B), and nuclear factor κB (NFκB). However, ERα and NFκB were neglected since they were already reported targets for at least one DHC (Dodds et al., 1938; Orlikova et al., 2012), while suitable assays for AChE and PTP1B were not available.

Second, protein targets that were particularly frequently predicted for the whole DHC chemical space (425 compounds) with a positive prediction of Ph-DB and CSs of two or higher (Table 2, Selection Criterion II) were included. We hypothesized that these targets were generally well suited for the DHC scaffold. 17β HSD2 and 17β HSD3 were the fourth and fifth most frequently predicted targets with a CS of 3, behind AChE (assays not available), ERα (already established target for some DHCs), and 5-LO (already selected for in vitro testing, see Figure 3 Supplement 1, D). Moreover, cyclooxygenase 1 (COX-1) and aldo-keto reductase 1C3 (AKR1C3) were the fourth and sixth most frequently predicted targets, respectively, with CSs of two (see Figure 3 Supplement 1, C). Higher ranked targets were all either already known (ERα and ERβ) or already included in our selection (aromatase, 5-LO).

Third, the overall predictions were evaluated for their novelty, their consistency and whether our approach could produce high scores for known ligand-target interactions (Table 2, Selection Criterion III). Novelty means that only unreported targets were selected, while prediction consistency means that predicted targets are more credible, if they are biologically related to one another, e.g. isoenzymes, or proteins that belong to the same pathway. It is actually oftentimes the case that one compound binds to several closely related targets (lack of specificity), which should be reflected in the virtual predictions (Hert et al., 2008; Jalencas et al., 2013). Further, great value is added if, e.g., a closely related target was already reported, or if known targets are enriched in the predictions generated by the in silico workflow. In our case, we observed targets belonging to steroid metabolism (aromatase, 17β HSD2, 17β HSD3, and AKR1C3) or to arachidonic acid (AA) metabolism (5-LO, COX-1). Targets of the steroid metabolism are obviously closely related to

ERs; aromatase, 17β HSD2, and ERs even share the same substrate or ligand, respectively, namely estradiol. The shape and pharmacophore of those targets' binding sites must, therefore, be somewhat similar. On this basis, we hypothesized that the four predicted targets of the steroid metabolism are promising DHC targets, given that ERs are confirmed targets of several DHCs. Similarly, COX-1 and 5-LO share the same substrate, namely AA, implying that also the binding sites of the latter two must be similar to a certain extent. Moreover, **7** has been confirmed to inhibit 15-hydroxyprostaglandin dehydrogenase and soluble epoxide hydrolase – two enzymes in the AA pathway related to COX-1 and 5-LO with again presumably similar binding sites (see Figure 2, A). Steroid metabolism and AA metabolism are on their side interconnected by AKR1C3, which is also commonly referred to as 17β-hydroxysteroid dehydrogenase 5 or prostaglandin F synthase (see Figure 4-Figure supplement 2, A-C). Indeed, this enzyme converts both steroids and AA-like fatty acids using the same binding site (Matsuura et al., 1998). From there it was concluded that those six proteins' binding sites may share substantial similarities and the predictions of the latter can be considered consistent. Finally, we checked if targets of the known DHC biological space could be enriched by our target prediction workflow. Effectively, we observed a clear enrichment of the consensus scored target frequencies (see Figure 3 Supplement 1, C and D and Figure 4-Figure supplement 1) compared to the stand-alone target prediction tools (see Figure 3 Supplement 1, A and B).

Fourth, the availability of a suitable assay was logically a pivotal criterion for targets to be selected (Table 2, Selection Criterion IV).

206 Table 2. Twelve frequently predicted targets for DHC chemical space assessed according to
207 selection criteria I – IV and final selection statement.

Candidate Target	Selection criterion I*	Selection criterion II†	Selection criterion III‡	Selection criterion IV§	*Selected*
17β HSD2	n.a.	5th (CS=3) 12th (CS=2)	**1** - ER α/β	Yes	*Yes*
17β HSD3	n.a.	4th (CS=3) 7th (CS=2)	**1** - ER α/β	Yes	*Yes*
5-LO	**4** (CS=2) **5** (CS=3)	3rd (CS=3) 3rd (CS=3)	**1** - PGDH	Yes	*Yes*
AChE	**4** (CS=2) **5** (CS=3)	1st (CS=3) 10th (CS=2)	n.a.	No	*No*
AKR1C3	n.a.	6th (CS=2)	**1** – AKR1B10	Yes	*Yes*
Aromatase	**1** (CS=2) **5** (CS=2) **6** (CS=2) **7** (CS=2) **8** (CS=2) **9** (CS=2)	1st (CS=2)	**1** - aromatase **1** - ER α/β **1** - several CYPs	Yes	*Yes*
COX-1	n.a.	4th (CS=3)	**1** - PGDH	Yes	*Yes*
ERα	**1** (CS=3) **2** (CS=2) **4** (CS=3) **5** (CS=3) **6** (CS=2) **7** (CS=2) **8** (CS=2)	2nd (CS=3) 2nd (CS=2)	**1** - ER α/β	Yes	*No*
ERβ	n.a.	5th (CS=3) 6th (CS=2)	**1** - ER α/β	Yes	*No*
NFκB	n.a.	8th (CS=2)	**1** - NFκB **5** - NFκB	No	*No*
PPARγ	n.a.	9th (CS=2)	**1** - PPARγ	No	*No*
PTP1B	**6** (CS=2)	10th (CS=2)	n.a.	Yes	*No*

208 * Targets that were predicted with high CSs for compounds **1** – **10**. † Most frequently predicted
209 targets with high CSs for DHC chemical space. ‡ Prediction consistency: Similar or associated
210 targets that are being predicted or similar bioactivities that were already reported. § Availability
211 of a suitable assay.

PH ∩ SEA

Cpd.	Target	SEA	SP	STP	Ph-DB
4	AChE	1.09e-08	nan	nan	0.58
4	5-LO	3.22e-06	nan	nan	0.44
5	NFκB	6.94e-07	nan	nan	0.86

Ph-DB ∩ SEA ∩ STP

Cpd.	Target	SEA	SP	STP	Ph-DB
1	ER α	1.17e-07	nan	1.00	0.95
4	ER α	7.12e-09	nan	0.14	0.96
5	AChE	2.76e-11	nan	0.11	0.59
5	ER α	5.50e-07	nan	0.19	0.97
5	5-LO	3.68e-14	nan	0.11	0.48

Ph-DH ∩ STP

Cpd.	Target	SEA	SP	STP	Ph-DB
1	aromatase	nan	nan	0.10	0.49
2	ER α	nan	nan	0.30	0.95
5	aromatase	nan	nan	0.11	0.76
6	aromatase	nan	nan	0.11	0.64
6	ER α	nan	nan	0.11	0.93
6	PTP1B	nan	nan	0.11	0.94
7	aromatase	nan	nan	0.12	0.56
7	ER α	nan	nan	0.12	0.95
8	aromatase	nan	nan	0.12	0.48
8	ER α	nan	nan	0.12	0.95
9	aromatase	nan	nan	0.11	0.67

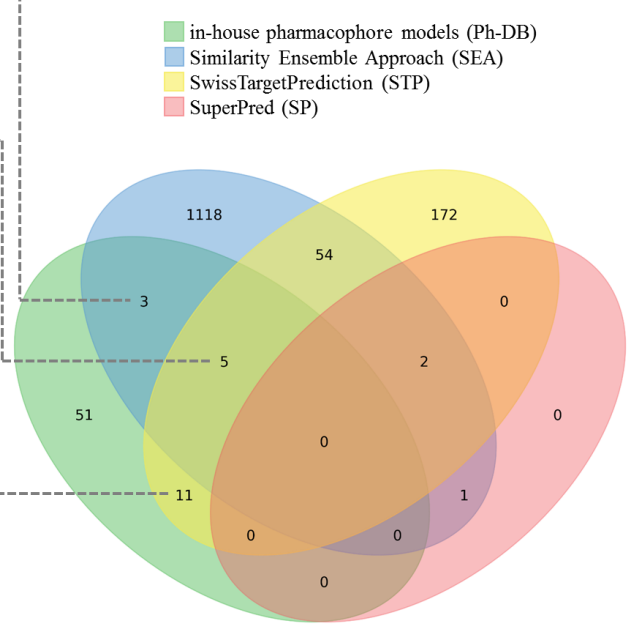
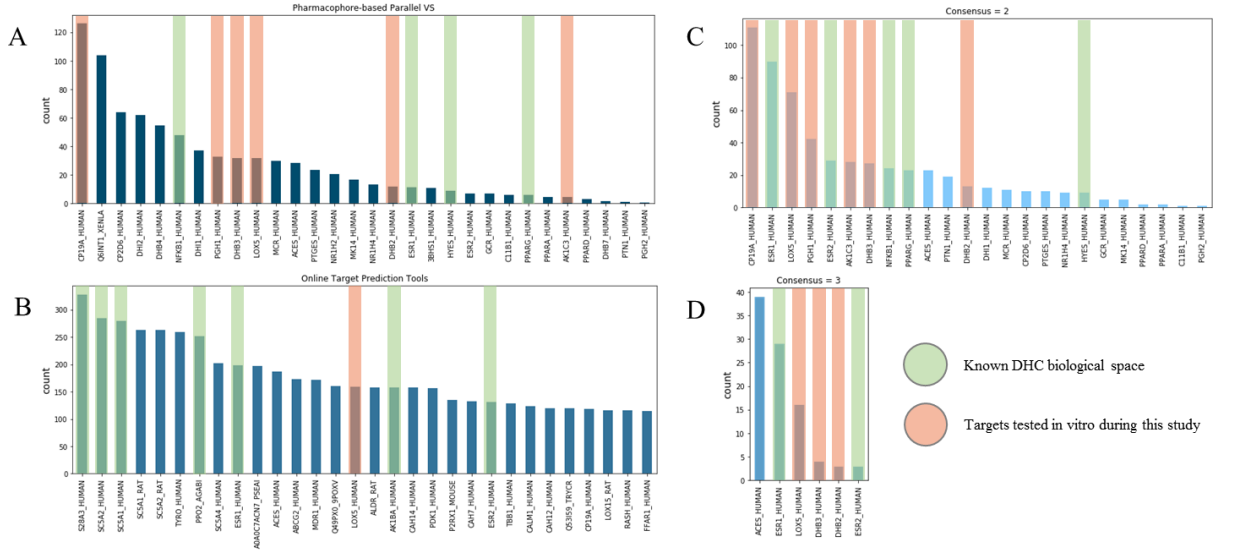


Figure 3. Predicted compound – target interactions (predicted DHC chemical space) illustrated as Venn diagram. Predicted DHC chemical space was filtered for compounds **1** – **10** and a positive prediction by Ph-DB. Sets of predictions by any method are represented as ellipses (green: Ph-DB, blue: SEA, yellow: STP, and red: SP) and consents among different methods as overlaps/intersections each with an integer indicating the size of the intersection. Values indicate the original model fit values, 'not a number' (for short 'nan') indicates no prediction by the respective tool. Fit values of SEA and SP are E-values similar as in the Basic Local Alignment Search Tool (BLAST), meaning the lower the better the model fit (Altschul et al., 1990; Dunkel et al., 2008; Keiser et al., 2007). Fit values for STP and Ph-DB are 0 – 1 normalized probability (STP) or relative pharmacophore fit-scores (Ph-DB) (Gfeller et al., 2013; Wolber et al., 2006). Enrichment of targets reported for DHCs in literature are shown for predictions made with stand-alone Ph-DB (Figure 3-Figure supplement – 1, A), SEA, STP, and SP combined (Figure 3-Figure supplement – 1, B), and all approaches combined and the consensus score applied (Figure 3-Figure supplement – 1, C-D), each as green overlays.



3BHS1_HUMAN	3β hydroxysteroid dehydrogenase 1	*DHB2_HUMAN*	17β HSD2	*MK14_HUMAN*	MAP kinase 14	*Q49PX0_9POXV*	N1L
A0A0C7ACN7_P.	3-oxoacyl-ACP synthase	*DHB3_HUMAN*	17β HSD3	*NFKB1_HUMAN*	NF κB p105 subunit	*Q53I59_TRYCR*	Phosphodiesterase
ABCG2_HUMAN	ATP-binding cassette G2	*DHB4_HUMAN*	17β HSD4	*NR1H2_HUMAN*	Oxysterols receptor β	*Q6INT1_XENLA*	MGC80376 protein
ACES_HUMAN	Acetylcholinesterase	*DHB7_HUMAN*	3-keto-steroid reductase	*NR1H4_HUMAN*	Bile acid receptor	*RASH_HUMAN*	GTPase HRas
AK1BA_HUMAN	Aldo-keto reductase 1B10	*DHI1_HUMAN*	11β HSD1	*P2RX1_MOUSE*	P2X purinoceptor 1	*S28A3_HUMAN*	Solute carrier 28A3
AK1C3_HUMAN	Aldo-keto reductase 1C3	*ESR1_HUMAN*	Estrogen receptor α	*PDK1_HUMAN*	Pyruvate dehydrogenase	*SC5A1_HUMAN*	Sodium/glucose cotransporter 1
ALDR_RAT	Aldo-keto reductase 1B1 (rat)	*ESR2_HUMAN*	Estrogen receptor β	*PGH1_HUMAN*	COX-1	*SC5A1_RAT*	Sodium/glucose cotransporter 1
C11B1_HUMAN	Cytochrome P450 11B1	*FFAR1_HUMAN*	Free fatty acid receptor 1	*PGH2_HUMAN*	COX-2	*SC5A2_HUMAN*	Sodium/glucose cotransporter 2
CAH12_HUMAN	Carbonic anhydrase 12	*GCR_HUMAN*	Glucocorticoid receptor	*PPARA_HUMAN*	PPAR α	*SC5A2_RAT*	Sodium/glucose cotransporter 2
CAH14_HUMAN	Carbonic anhydrase 14	*HYES_HUMAN*	Soluble epoxide hydrolase 2	*PPARD_HUMAN*	PPAR δ	*SC5A4_HUMAN*	Solute carrier family 5 member 4
CAH7_HUMAN	Carbonic anhydrase 7	*LOX15_RAT*	15-lipoxygenase (rat)	*PPARG_HUMAN*	PPAR γ	*TBB1_HUMAN*	Tubulin β chain
CALM1_HUMAN	Calmodulin-1	*LOX5_HUMAN*	5-lipoxygenase	*PPO2_AGABI*	Mushroom tyrosinase	*TYRO_HUMAN*	Tyrosinase
CP19A_HUMAN	Aromatase	*MCR_HUMAN*	Mineralocorticoid receptor	*PTGES_HUMAN*	Prostaglandin E synthase		
CP2D6_HUMAN	Cytochrome P450 2D6	*MDR1_HUMAN*	ATP-dependent translocase ABCB1	*PTN1_HUMAN*	PTP1B		

Figure supplement 1. Enrichment of known DHC targets (known DHC biological space) in differently scored predictions for DHC chemical space.

New DHC Biological Space

Following the definition of the six protein targets issued for biological evaluation, the respective in vitro assays were performed using compounds **1** – **10** (see Figure 4 and Table 3). Biological activities were expressed as percent inhibition at 10 µM compound concentration relative to the mock control (= 0%). Three independent experiments (n = 3) were conducted and the mean inhibition plus/minus standard deviation depicted. Mean inhibition values that were below 30% were regarded as inactive, negative inhibition values and relative standard deviations larger than 20% were regarded as ambiguous assay results and thus as inactive. Unfortunately, 17β HSD2 measurements yielded ambiguous assay results for all compounds, which were regarded as inactive, as well as aromatase, where all ten compounds were inactive. For 5-LO, COX-1, AKR1C3, and 17β HSD3 at least one of the ten compounds exhibited weak inhibitory activities towards the respective target. Thus, **4** and **5** showed weak inhibition of AKR1C3, **2** showed weak inhibition of 17β HSD3 and **1**, **2**, **3**, and **5** showed weak inhibition of COX-1. The ten DHCs showed the best results of this study in 5-LO inhibition, where **1**, **2**, **7**, **8**, and **10** showed weak to moderate inhibitory activities, **1** and **2** even reaching mean inhibition values of 85.4 ± 9.3% and 99.2 ± 1.2%, respectively. In the course of this study, six further protein targets were evaluated in vitro due to availability of the respective assays, rather than based on in silico predictions. Most of the measured activities were not affected, however, to not withhold those results to the community, the results are shown in Figure 4-Figure supplement – 3.

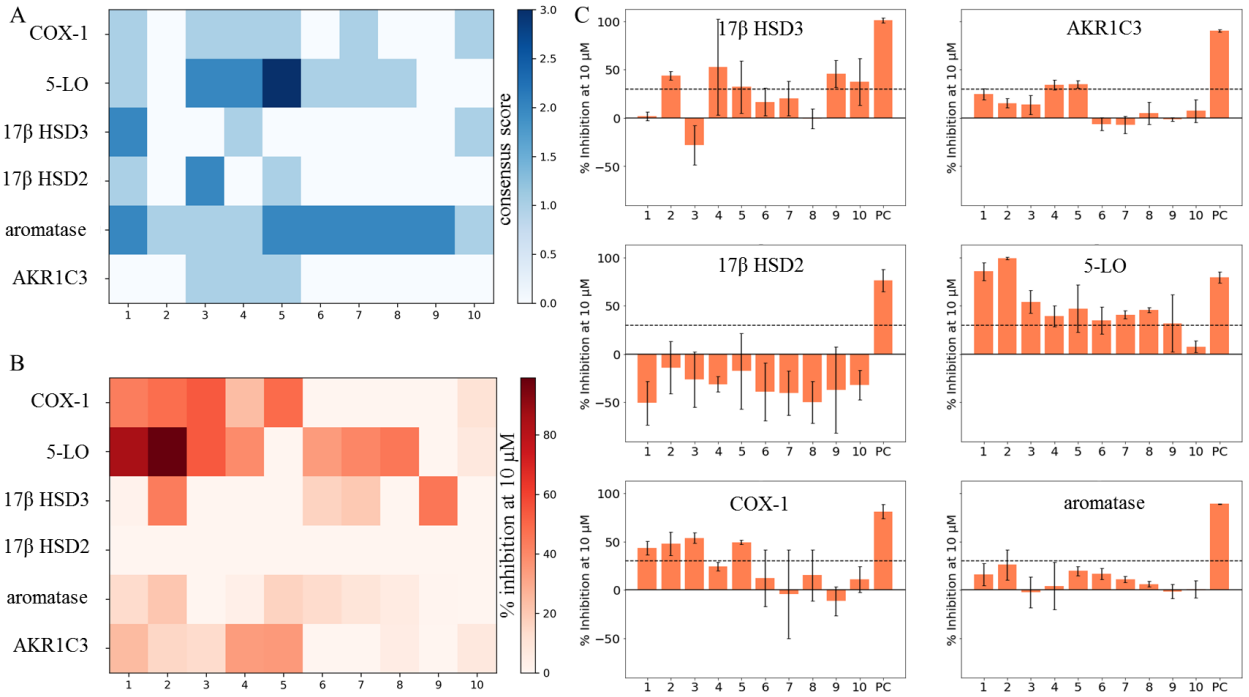


Figure 4. Comparison of unknown DHC biological space (COX-1, 5-LO, 17β HSD3, 17β HSD2, aromatase, and AKR1C3) and actual in vitro test results of compounds **1** – **10**. (A) CSs of compounds **1** – **10** on all of the six targets of the unknown DHC biological space plotted as heatmap. (B) Means (n = 3) of percent inhibition at 10 µM (0 – 100%) of compounds **1** – **10** on all of the six targets of the unknown DHC biological space plotted as heatmap. Observations with mean inhibition values smaller than 30% or relative standard deviations greater than 20% were regarded as inactive. (C) Bar charts of the six targets of the unknown DHC biological space showing compounds **1** – **10** with the respective means (n = 3) of percent inhibition at 10 µM (0 – 100%) and standard deviation. A cut-off of 30% mean inhibition at 10 µM was chosen (black dashed line), separating active from inactive observations. DMSO was used to measure baseline enzyme activities, on which samples were normalized (not shown) and positive controls (PC) were used as indicated in Materials and Methods.

Metrics Computed for the Isolated and Combined Target Prediction Tools

	Metric	AKR1C3	17β HSD3	17β HSD2	Aromatase	5-LO	COX-1	Mean
Consenensus Score (CS)	Recall	1.00	0	0	0	0.86	0.75	0.44
	Precision	0.67	0	0	0	0.86	0.50	0.34
	Relative EF	0.67	0	0	0	0.86	0.50	0.34
pharmacophore-based parallel VS (Ph-DB)	Recall	0	0	0	0	0.57	0.50	0.18
	Precision	0	0	0	0	0.80	0.50	0.22
	Relative EF	0	0	0	0	0.80	0.50	0.22
Similarity Ensemble Approach (SEA)	Recall	1.00	0	0	0	0.29	0	0.22
	Precision	0.67	0	0	0	0.67	0	0.22
	Relative EF	0.67	0	0	0	0.67	0	0.22
SwissTarget Prediction (STP)	Recall	0	0	0	0	0.29	0.25	0.27
	Precision	0	0	0	0	0.67	0.50	0.20
	Relative EF	0	0	0	0	0.67	0.50	0.20
SuperPred (SP)	Recall	0	0	0	0	0	0	0
	Precision	0	0	0	0	0	0	0
	Relative EF	0	0	0	0	0	0	0

True/False Positives/Negatives

True Positives (TPs) refer to actives that were predicted as active, while False Positives (FPs) refer to inactives that were predicted as active. Vice versa, True Negatives (TNs) are inactives predicted as active, while False Negatives (FNs) are actives predicted as inactive.

Recall/Sensitivity

Proportion of true actives a model is able to retrieve from the screening dataset (0 - 1).

$$recall = \frac{TP}{TP + FN}$$

Precision

Proportion of true actives in a hitlist produced by a model (0 - 1).

$$precision = \frac{TP}{TP + FP}$$

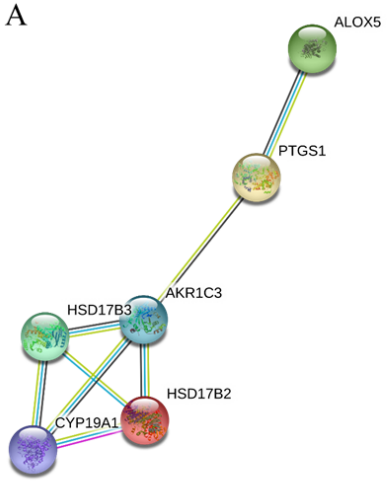
Relative Enrichment Factor

Ability of a model to enrich a hitlist with true positive predictions (0 - 1).

$$relative\ EF = \frac{\frac{TP}{TP + FP}}{\frac{actives}{inactives}} * \frac{1}{\frac{actives}{inactives}}$$

Figure supplement - 1. Metrics computed for the isolated and combined target prediction tools.

A



B

Node 1		Node 2		STRING Score
STRING name	alternative name	STRING name	alternative name	
AKR1C3	AKR1C3	CYP19A1	aromatase	0.952
AKR1C3	AKR1C3	HSD17B2	17β HSD2	0.924
AKR1C3	AKR1C3	HSD17B3	17β HSD3	0.935
AKR1C3	AKR1C3	PTGS1	COX-1	0.452
ALOX5	5-LO	PTGS1	COX-1	0.973
CYP19A1	aromatase	AKR1C3	AKR1C3	0.952
CYP19A1	aromatase	HSD17B2	17β HSD2	0.962
CYP19A1	aromatase	HSD17B3	17β HSD3	0.958
HSD17B2	17β HSD2	AKR1C3	AKR1C3	0.924
HSD17B2	17β HSD2	CYP19A1	aromatase	0.962
HSD17B2	17β HSD2	HSD17B3	17β HSD3	0.979
HSD17B3	17β HSD3	AKR1C3	AKR1C3	0.935
HSD17B3	17β HSD3	CYP19A1	aromatase	0.958
HSD17B3	17β HSD3	HSD17B2	17β HSD2	0.979
PTGS1	COX-1	AKR1C3	AKR1C3	0.452
PTGS1	COX-1	ALOX5	5-LO	0.973

C

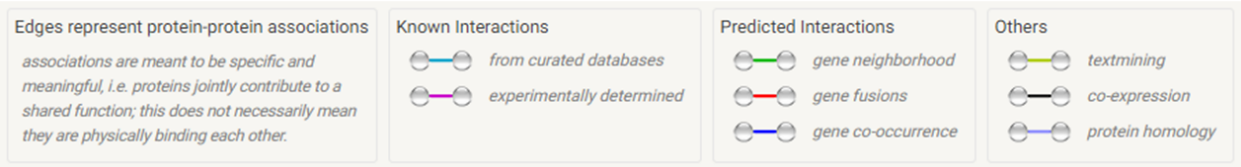


Figure supplement - 2. Six potential protein targets for DHCs input to STRING, which creates a network of both direct and functional protein-protein interactions (Szklarczyk et al., 2018).

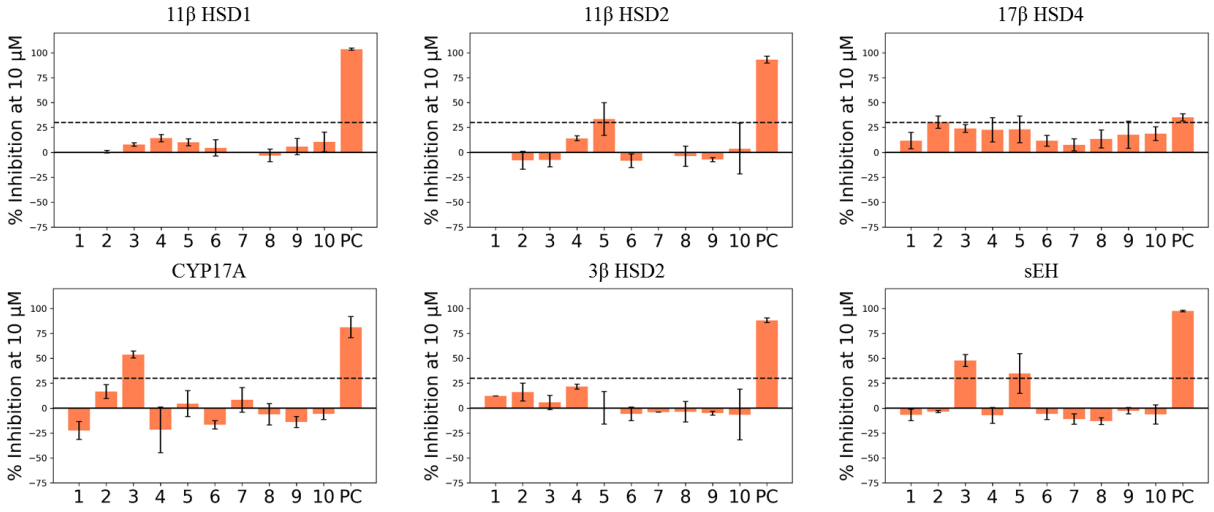


Figure supplement - 3. Results of other protein targets tested in the course of this study due to assay availability shown as bar charts. Compounds **1** – **10** indicated with the respective means (*n* = 3) of percent inhibition at 10 µM (0 – 100%) and standard deviation. Black dashed line again indicates arbitrarily chosen 30% activity cut-off. DMSO was used to measure baseline enzyme activities, on which samples were normalized (not shown) and positive controls (PC) were used as indicated in Materials and Methods.

Table 3. In vitro inhibitory activities of compounds **1** – **10** towards targets of DHC biological space, expressed as percent inhibition (0 – 100%) at 10 µM compound concentration relative to mock control. Shown is the mean of three independent experiments (n = 3) plus/minus standard deviation. Different compounds were used as positive controls (PC) as indicated in Materials and Methods. Highly negative values of 17β HSD2 assays are believed to be technical artefacts, as enzyme activation seems unlikely.

Compound	aromatase	17β HSD2	17β HSD3	AKR1C3	5-LO	COX-1
1	13.8 ± 2.0	-50.7 ± 22.7	1.7 ± 4.5	24.8 ± 5.9	85.4 ± 9.3	43.5 ± 7.2
2	21.1 ± 11.7	-14.1 ± 27.1	43.8 ± 4.7	15.5 ± 4.9	99.2 ± 1.2	48.1 ± 12.0
3	-1.0 ± 12.0	-26.3 ± 28.8	-28.0 ± 20.5	13.8 ± 10.0	54.1 ± 11.6	53.9 ± 5.3
4	3.5 ± 19.0	-31.2 ± 8.1	52.7 ± 49.6	34.4 ± 5.1	39.2 ± 11.1	24.4 ± 4.2
5	17.0 ± 2.0	-17.6 ± 39.0	32.1 ± 27.2	35.2 ± 3.8	47.2 ± 24.2	49.5 ± 1.9
6	13.8 ± 4.4	-39 ± 29.9	16.7 ± 14.5	-6.1 ± 6.8	34.8 ± 14.2	12.34 ± 29.0
7	9.4 ± 3.5	-40.5 ± 22.9	20.2 ± 17.7	-7.2 ± 8.9	40.8 ± 4.1	-4.2 ± 45.6
8	5.9 ± 3.5	-49.8 ± 21.4	-0.6 ± 10.0	5.3 ± 11.5	45.5 ± 2.7	15.3 ± 26.4
9	0.67 ± 4.0	-37.1 ± 44.3	45.8 ± 14.2	-1.3 ± 1.9	31.8 ± 29.4	-11.4 ± 15.1
10	0 ± 5.8	-32.2 ± 15.4	37.5 ± 23.9	7.4 ± 11.8	7.7 ± 6.0	11.1 ± 13.2
PC	70.2 ± 0.5*	76.1 ± 11.4†	101.2 ± 2.4‡	90.5 ± 1.2§	79.26 ± 5.95¶	81.3 ± 7.5#

* 10 nM anastrozole (CAS: 120511-73-1). † 1 µM ML376 (CAS: 1340482-23-6). ‡ 1 µM compound 24 (CAS: 873206-61-2) (Möller et al., 2009), § 1 µM compound 2-9 (CAS: 745028-76-6) (Schuster et al., 2011), ¶ 3 µM zileuton (CAS: 111406-87-2). # 10 µM indomethacin (CAS: 53-86-1).

DISCUSSION

In the current study, we present an in silico target prediction workflow capable of prioritizing new molecular targets for known chemical entities, here exemplified on DHCs. Even though in silico target prediction is common today for synthetic compounds (Cereto-Massagué et al., 2015; Keiser et al., 2007; Keiser et al., 2009; Lounkine et al., 2012), it is still challenging for natural products (Reker et al., 2014; Rodrigues et al., 2015; Rollinger, 2009; Rollinger et al., 2009). To our knowledge, this is the first study performing in silico target prediction on a natural product class while systematically combining diverse established tools. We have recently observed that, especially when predicting targets for natural products, none of the established tools (including our own) performed perfectly, but owed to different methods, all of them performed differently (Mayr et al., 2019b). Exploiting the complementarity of the methods' strengths should therefore correct for this shortcoming and result in better predictive performance. This strategy might be of great interest to the community since it underlines the benefits of predicting targets of a closely related compound series, rather than single compounds and it provides a thorough use case of publicly available tools and how to interpret its predictions. Also, we have recently predicted DHCs to be inhibitors of mushroom tyrosinase, albeit they turned out to be alternative substrates of the latter – an unexpected form of bioactivity that an in silico target prediction cannot distinguish from competitive inhibitors (Mayr et al., 2019a).

Finally, six potential DHC targets were selected of which four could be experimentally confirmed (5-LO, COX-1, 17β HSD3, and AKR1C3) as molecular targets of at least one of the ten DHCs **1 – 10**. These four proteins are all new DHC targets and thus expanded the DHC biological space, which we were aiming for (see Figure 2). In terms of accuracy, our yield of novel targets (four out of six) is clearly superior to the yield of a "random selection" (test a random panel of protein targets towards compounds **1** - **10**), which typically lies below 1% (Doman et al., 2002; Ferreira et al., 2010; Polgár et al., 2005; Young et al., 2005). 17β HSD2 and aromatase could not be confirmed as targets, however, a thorough literature search revealed that **1** was reported once in literature to inhibit aromatase with an IC_{50} value ≥ 50 µM (Le Bail et al., 2001). Activities at such high concentrations are considered inactive by PubChem and were thus not retrieved by our bioactivity mining approach. This activity is indeed negligible, however, this value seems consistent with the result generated by us (13.8 ± 2.0% inhibition at 10 µM, see Table 3). The observed activities were all in typical ranges for non-optimized lead structures discovered by VS. Ripphausen et al.

316 conducted a survey in 2010, showing that hit-compounds identified by VS have defined potency
317 endpoints (IC_{50}, EC_{50}, K_i, or K_d) of 4 to 19 µM in average, which is arguably high. However, the
318 true value of VS lies in the ability to identify new chemotypes as leads, or in its turn, to identify
319 new targets for known compounds in target prediction (Ripphausen et al., 2010). The exciting fact
320 of this study is that our workflow could successfully prioritize new molecular targets of well-
321 known compounds with noteworthy accuracy, opening up new avenues for DHC research (see
322 Figure 4-Figure supplement 1).

323 The results can now be utilized in manifold ways: First, our findings revealed polypharmacology
324 of some DHCs as a sideline. Thus, **1** and **2** seem to have promising anti-inflammatory properties,
325 by simultaneously inhibiting 5-LO and COX-1. Investigating whether this property was inherited
326 to gliflozin drugs could be the subject of another study. Potential anti-inflammatory properties of
327 gliflozin drugs could contribute to the currently observed beneficial effects of those drugs in e.g.
328 heart failure (Dutka et al., 2019). Several authors have described anti-inflammatory properties of
329 gliflozin drugs on a functional level, however a distinct mode of action and association to distinct
330 molecular targets is still to be elucidated (Hattori, 2018; Iannantuoni et al., 2019; Xu et al., 2018).

331 Second, to remain with polypharmacology, e.g. **4** and **5** could serve as lead to develop dual
332 inhibitors of AKR1C3 and 5-LO. Analogously, **2** could be used as a starting point for the
333 development of a dual 17β HSD3/5-LO inhibitor. Both options would represent a new compound
334 class with potentially interesting pharmacological properties. These agents could be beneficial e.g.
335 in the treatment of prostate cancer since AKR1C3 and 17β HSD3 are frequently linked to this
336 condition (Margiotti et al., 2002; Vicker et al., 2009) and malignant tumors are usually surrounded
337 by a pro-inflammatory microenvironment (Neuwirt et al., 2020). Apart from the two just
338 mentioned possibilities, several other scenarios of utilization are reasonable.

339 Third, our target prediction workflow revealed the great potential of DHCs as lead structures. As
340 mentioned above, DHCs are incredibly well accessible from biomass like apple leaves, in
341 concentrations up to 20% dry mass (Gaucher et al., 2013). According to the principle of
342 bioprospection, meaning the harnessing of resources from nature for medical purposes, apple
343 leaves could play an interesting role in future lead optimization campaigns.

Fourth, the established in silico prediction workflow can be applied to several other compound classes alike DHCs. Therefore, this study blueprints a strategy to predict targets for known compounds, by mostly using open-source platforms, thereby empowering a great number of researchers to actively re-dedicate their compounds. This technology however is greatly enabled by the growing body of known bioactivities, hinting towards increasing accuracies reachable with target prediction campaigns in the near future, as this knowledge keeps on expanding.

MATERIALS AND METHODS

Dataset Assembly. A dataset representing the chemical space of DHCs was gathered, containing naturally occurring DHCs as well as semi-synthetic derivatives of the latter. First, compounds **1** – **10** were included. Then, a thorough literature search was conducted in SciFinder, using substructure search function for 1,2-diphenyl-propan-1-one. The SciFinder query was conducted on the 8th of August 2017 and yielded 5457 DHCs. The resulting compounds were checked manually for their origin (natural product vs. semi-synthetic), created in ChemDraw Professional (version 16.0.0.82 (68), PerkinElmer, Waltham, MA, USA) with assigned stereochemistry, and double checked with SciFinder using the 'search SciFinder' function in ChemDraw. For compounds, especially natural products, with lacking absolute configuration, all possible stereoisomers were included. Finally, the dataset 'DHC_full' contained 425 natural or semi-synthetic DHCs (corrected for stereochemistry) and was converted with a custom-built Pipeline Pilot protocol (version 9.5.0.831, Dassault Systèmes BIOVIA, Vélizy-Villacoublay, France) to an sd-file (DHC_full.sdf) and a csv-file (DHC_full.csv). The Pipeline Pilot protocol is outlined in Supplementary Information S-1. The files are provided on GitHub (see Supplementary Information S-2).

Bioactivity Mining. The mining of bioactivity data of DHCs was done through a script coded in Python 3 (version 3.7.3, https://www.python.org/) called 'bioactivity_network_generator_SMILES.py'. This script iterates over a two-column csv-table (name, SMILES) while fetching one SMILES code at a time. The script interacts with various application programming interfaces (APIs) of PubChem (PUG REST, https://pubchem.ncbi.nlm.nih.gov/) (Kim et al., 2018), UniProt (https://www.uniprot.org/) (The UniProt Consortium, 2018), and Reactome (https://reactome.org/) (Fabregat et al., 2017). At the first stage, the SMILES code is posted as a query to the PubChem API, which returns the desired PubChem compound ID (CID), if present. The CID is then used to post a second query to the PUG REST API, fetching all assay IDs (AIDs) associated with the posted CID, for which the assay result was flagged as 'active'. The produced list of AIDs is then annotated with the respective gene names from PubChem. This step eliminates all AIDs that are not associated to a single protein (e.g. cell-based assays). The gene name from PubChem can be translated to UniProt names and the respective entry names used in UniProt using the UniProt KB API. In a last step, the UniProt entry

name was used to retrieve associated human pathways from the Reactome API. The gathered bioactivity data of all provided SMILES codes was processed in Pandas (version 0.24.2, https://pandas.pydata.org/) (McKinney, 2010) and finally written to a csv-file in the simple interaction file (sif) format, which can be visualized in e.g. Cytoscape (https://cytoscape.org/) (Shannon et al., 2003). Interaction weights were calculated by using the '*group by*' method implemented in Pandas chained with the '*count()'* method. Interaction weights thus are an integer representing the appearances of a particular interaction in PubChem. The script 'bioactivity_network_generator_SMILES.py' is provided on GitHub (see Supplementary Information S-2).

Pharmacophore-based Parallel Virtual Screening. Pharmacophore-based parallel VS was performed using the historically grown in-house pharmacophore model databased built and maintained by Prof. Daniela Schuster. The database consists of 387 ligand-based and structure-based pharmacophore models for 39 protein targets built in two different software environments, namely Discovery Studio (version 4.5.0.15071, Dassault Systèmes BIOVIA, Vélizy-Villacoublay, FR) and various versions of LigandScout (Inte:Ligand, Vienna, AT). For the Discovery Studio models, parallel screening was performed using the 'ligand profiler' protocol (for settings refer to Supplementary Information Table S-4). Screening of the LigandScout models was performed using the 'iscreen.exe' program via command line and the databases created for every LigandScout version by using the 'idbgen.exe' via command line for the respective LigandScout versions. Thus, Omega (OpenEye Scientific Software, Santa Fe, NM, USA) with 'best' settings was used. The targets represented by these models are predominantly targets belonging to the arachidonic acid (AA) cascade as well as corresponding downstream signaling and steroid metabolism and signaling. The targets are often associated with inflammation, neoplasm, or are popular off-targets. A detailed compilation of all models is provided in Supplementary Information Table S-3.

Target Prediction with Publicly Available Tools. Next to the pharmacophore-based parallel screening, three target prediction tools that are available as web servers were used, namely SEA (http://sea16.docking.org/) (Keiser et al., 2007), STP (http://www.swisstargetprediction.ch/) (Daina et al., 2019; Gfeller et al., 2014; Gfeller et al., 2013), and SP (http://prediction.charite.de/) (Dunkel et al., 2008; Nickel et al., 2014). SEA is a 2D ligand-based, similarity ensemble method. Each target present in SEA is described by a set of its known ligands of various size. An input

ligand is then compared against all ligands of all target sets via Tanimoto similarity of the ECFP4 fingerprints. For each target, the Tanimoto similarities are summed up and z-scores calculated. Since the authors computed the distribution of z-scores obtained between random similarity ensembles, the z-scores of a screening ligand to each target can be used to calculate expectation values (E-values). Those E-values, similar as in the BLAST algorithm, express the likelihood that the observed similarity happened due to coincidence. SEA uses bioactivity data derived from ChEMBL (Gaulton et al., 2017) and is maintained by the University of California, San Francisco (UCSF). SP operates in a very similar way to SEA, being a 2D similarity ensemble approach and using ECFP4 fingerprints as well. Bioactivities that were used to build reference target sets in SP were derived from ChEMBL (Gaulton et al., 2012), Binding DB (Liu et al., 2007), and SuperTarget (Günther et al., 2008; Hecker et al., 2012). SP is maintained by the structural bioinformatics group of the Charité - University Medicine Berlin in Germany. In contrast to SEA and SP, STP makes use of the ligand-based similarity ensembles principle as well, however it is a hybrid method between 2D and 3D. 2D similarity is computed via Tanimoto similarity using FP2 fingerprints, while 3D similarity is described as Manhattan distance between the Electroshape (Armstrong et al., 2010) vectors. Finally, a logistic regression classifies the input ligand based on 2D and 3D similarities. STP derived its bioactivity data also from ChEMBL. It was developed by the Swiss Institute of Bioinformatics (SIB). The online servers described above were accessed via a web scraper script called 'TarPredCrawler.py' written in python 3. The script uses selenium (https://www.seleniumhq.org/, version 3.141.0) to send post requests of smiles codes to the four servers and downloads the resulting prediction as a table. The script 'TarPredCrawler.py.py' is provided on GitHub (https://github.com/fmayr/DHC_TargetPrediction).

Biochemical Assays. Aromatase assays were performed as previously described (Pandey et al., 2007). Briefly, genes for human wild-type aromatase and NADPH P450 oxidoreductase were transfected into *E. coli*, to express both proteins in the recombinant form and proteins were purified using multiple chromatograpgic procedures as described previously. Liposomes containing both enzymes were formed for the assay of enzymatic activities. aromatase activity was quantified by measuring the release of tritiated water after incubation with 1β-^{3}H androstenedione, a method introduced by Lephart and Simpson (Lephart et al., 1991). Ten nM Anastrozole (CAS: 120511-73-1) were used as positive control.

Inhibitory activities towards AKR1C3, 17β HSD3, and 17β HSD2 were assayed as described in Schuster et al. (Schuster et al., 2011). Briefly, AKR1C3 and 17β HSD2 were transformed into *E. coli* BL21 (DE3) and 17β HSD3 transfected into HEK293 cells. For assaying inhibitory activities towards 17β HSD2, bacterial suspensions were used, while cell suspensions were used to assay 17β HSD3 and bacterial lysates were used to assay AKR1C3. The protein containing lysates and suspensions, were incubated with tritiated substrates and cofactors (21 nM 17β-estradiol (6,7-^{3}H) and 750 nM NAD^{+} for 17β HSD2 and 10,6 nM 4-androstene-3,17-dione (1,2,6,7-^{3}H) and 600 µM NADPH for 17β HSD3 and AKR1C3) in the presence of test compounds in a final concentration of 10 µM (compounds supplied in DMSO; 1% final DMSO in the assay). After a defined incubation time, substrates and products were extracted using solid phase extraction (SPE) and analyzed by RP-HPLC and online scintillation counting. Quantification of relative conversion occurred via chromatographic peak integration and the percentage of inhibition was calculated relative to a mock control (1% DMSO). As positive controls compound 2-9 (CAS: 745028-76-6) (Schuster et al., 2011) was used for AKR1C3 assays, compound 24 (CAS: 873206-61-2) (Möller et al., 2009) for 17β HSD3 assays and compound 19 (CAS: 1340482-23-6) (Wetzel et al., 2011) for 17β HSD2 assays, all in 1 µM concentration. Inhibitory activities towards 5-LO and COX-1 were determined as described earlier by Schaible et al (Schaible et al., 2014). and Koeberle et al. (Koeberle et al., 2008), respectively. Briefly, polymorphonuclear leukocytes (for 5-LO) and human platelets (for COX-1) were freshly isolated from the blood of healthy volunteers, pre-incubated with the potential inhibitors and stimulated with 2.5 µM Ca^{2+}-ionophore A23187 or arachidonic acid, respectively. The reaction was stopped, substrates and products isolated and analysed on RP-HPLC. 5-LO products included LTB_4, its trans-isomers, 5-HPETE, and 5-HETE, while the COX-1 product was quantified as 12-HHT. Again, quantification occurred via chromatographic peak integration and the percentage calculated relative to a mock control. Indomethacin (CAS: 53-86-1) in 10 µM concentration was used as positive control for COX-1 assays and Zileuton (CAS: 111406-87-2) in 3 µM concentration for 5-LO assays.

Activities for assaying 11β hydroxysteroid dehydrogenase 1 (11β HSD1) and 11β hydroxysteroid dehydrogenase 2 (11β HSD2) were determined as previously described by Kratschmar et al. (Kratschmar et al., 2011). Briefly, lysates of HEK-293 cells stably expressing human 11β HSD1 were incubated with 200 nM cortisone (including 10 nM [1,2-^{3}H]-cortisone), 500 µM NADPH and the test substance. For 11β HSD2, lysate of HEK-239 cells stably expressing human 11β HSD2

was incubated with 50 nM cortisol (including 10 nM [1,2,6,7-^{3}H]-cortisol), 500 µM NAD^{+} and the test compounds. Conversion of radiolabeled substrate was determined and compared to enzyme activity in the control sample. 18β-Glycyrrhetinic acid (CAS: 471-53-4) was used as positive control for both enzymes. Inhibitory activities for 3β hydroxysteroid dehydrogenase 1 (3β HSD1) and cytochrome P450 17A1 (CYP17A) were assayed as described before by Samadari et al. and Udhane et al (Samandari et al., 2007; Udhane et al., 2017). Activities were measured in cell-based assays, using Human adrenocortical NCI-H295R cells obtained from American Type Culture Collection (ATCC; CRL-2128). The cells were treated with tritiated substrates and the product mix separated by thin layer chromatography (TLC) and the resulting spots subsequently densiometrically quantified. Trilostane (CAS: 13647-35-3) was used as positive control. Inhibitory activities towards soluble epoxide hydrolase (sEH) were assayed using the purified enzyme as described by Wixtrom et al. and Morisseau et al. (Morisseau et al., 2000; Wixtrom et al., 1988). A baculovirus was used to transduce sEH into Sf9 insect cells, which were subsequently lysed and and the enzyme purified using affinity chromatography. Enzyme inhibition could then by quantified using the purified sEH and substrate which turned into a fluorophore by the latter, which can be read at 465 nm after excitation at 300 nm (Waltenberger et al., 2016). AUDA (CAS: 479413-70-2) was used as positive control. Inhibitory activities towards 17β HSD4 were assayed according to the description in Schuster et al. (Schuster et al., 2011). Briefly, a plasmid coding for 17β hydroxysteroid dehydrogenase 4 (17β HSD4) was transformed into *E. coli* BL21 (DE3) Codon Plus RP (Stratagene). Subsequently, bacterial suspensions were prepared and incubated in the presence of 21 nM 17β-estradiol (6,7-^{3}H), 750 nM NAD^{+}, and 10µM test compound (1% DMSO final). After a defined incubation time, substrate and product were extracted using solid phase extraction (SPE) and analyzed with RP-HPLC in a Beckman-Coulter system and online scintillation counting. Enzymatic conversion was calculated by integrating substrate and product peaks and calculating percent inhibition relative to a control assay without inhibitor (1% DMSO). Compound 19 (CAS: 1340482-23-6) from (Wetzel et al., 2011) served as positive control.

Materials. Compounds **1** – **10** were purchased at TransMIT GmbH (PlantMetaChem, Gießen, Germany) with the following product numbers: **1**: P 036; **2**: H 031; **3**: D 017; **4**: A 020; **5**: D 018; **6**: S 025; **7**: P 037; **8**: T 017; **9**: P 064; **10**: N 019. Purity was assessed by HPLC-DAD (280 nm) found to be above 95% for all compounds.

ACKNOWLEDGEMENTS

The research was funded by GECT Euregio Tirol–Südtirol–Trentino (IPN55). F.M. and D.S. thank OpenEye Scientific Software and Inte:Ligand for their academic licenses. D.S. is an Ingeborg Hochmair professor at the University of Innsbruck.

AUTHOR CONTRIBUTIONS

Fabian Mayr, Stefan Martens, Birgit Waltenberger, Veronika Temml, Stefan Schwaiger, Hermann Stuppner, and Daniela Schuster designed the study. Fabian Mayr, Veronika Temml and Daniela Schuster performed the in silico screening and evaluated the results. Stefan Martens provided physical samples. Rolf W. Hartmann discovered and provided the positive control for 17β HSD2 testing, Christian Gege discovered and provided the positive control for 17β HSD3 testing. Gabriele Möller, Jerzy Adamski, Ulrike Garscha, Jana Fischer, Oliver Werz, Silvia G. Inderbinen, Alex Odermatt, Patricia Rodríguez Castaño, and Amit Pandey planned and performed the in vitro assays. Fabian Mayr and Daniela Schuster wrote the manuscript with the contributions of all authors.

COMPETING INTERESTS

The authors declare no competing financial interests.

DATA AVAILABILITY

Table 4. Files used and produced during this study. Every file is freely available at GitHub (https://github.com/fmayr/DHC_TargetPrediction). For greater clarity, a file scheme is provided in Supplementary Information S-2 describing all dependencies.

File Name	Contains	Subfolder in GitHub
DHC_full.csv	DHC chemical space as csv-file (name, smiles).	/dataset
DHC_full.sdf	DHC chemical space as 3D-molecule files.	/dataset
DHC_full_lit_network.csv	Known DHC biological space and result of bioactivity mining. Ready to be imported to Cytoscape.	/bioactivity%20mining
DHC_full_online.csv	Result produced by online target prediction servers (SEA, STP, SP) and DHC_full.csv as input.	/TarPredCrawler
DHC_full_LS_mergedhits.csv	Csv-file of hitlists produced by LigandScout models in Ph-DB.	/pharmacophore-based parallel VS
DHC_full_ligandprofiler.csv	Csv-file of hitlists produced by Discovery Studio model in Ph-DB.	/pharmacophore-based parallel VS
DHC_full_inhouse.csv	Joined results of LigandScout and Discovery Studio outputs.	/pharmacophore-based parallel VS
DHC_full_pivoted.csv	Joined results of SEA, STP, SP, and Ph-DB predictions for DHC chemical space.	
DHC_10_pivoted.csv	DHC_full_pivoted.csv filtered for compounds **1** – **10**.	
DHC_10_network.csv	DHC_10_pivoted.csv joined with DHC_full_lit_network.csv. Network file ready to be imported to Cytoscape. Contains known and predicted compound-target associations.	
Bioactivity_network_generator_SMILES.py	Python script used for literature mining. For installation instruction see README.	
TarPredCrawler.py	Python script used for submitting and collecting results from SEA, STP, and SP.	
DHC_targetpreidction_datatreatment.ipynb	Jupyter Notebook containing all data treatment and plotting performed in this study.	

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol., 215*(3), 403-410. doi: https://doi.org/10.1016/S0022-2836(05)80360-2

Armstrong, S. M., Morris, G. M., Finn, P. W., Sharma, R., Moretti, L., Cooper, R. I., & Richards, W. G. (2010). ElectroShape: fast molecular similarity calculations incorporating shape, chirality and electrostatics. *J. Comput. Aid. Mol. Des., 24*, 789–801. doi: 10.1007/s10822-010-9374-0

Aronson, J. K. (2007). Old drugs – new uses. *Br. J. Clin. Pharmacol., 64*(5), 563-565. doi: 10.1111/j.1365-2125.2007.03058.x

Ashburn, T. T., & Thor, K. B. (2004). Drug repositioning: identifying and developing new uses for existing drugs. *Nat. Rev. Drug Discovery, 3*(8), 673-683. doi: 10.1038/nrd1468

Campbell, I. B., Macdonald, S. J. F., & Procopiou, P. A. (2018). Medicinal chemistry in drug discovery in big pharma: past, present and future. *Drug Discovery Today, 23*(2), 219-234. doi: https://doi.org/10.1016/j.drudis.2017.10.007

Cereto-Massagué, A., Ojeda, M. J., Valls, C., Mulero, M., Pujadas, G., & Garcia-Vallve, S. (2015). Tools for in silico target fishing. *Methods, 71*, 98-103. doi: http://dx.doi.org/10.1016/j.ymeth.2014.09.006

Chen, C., Huang, H., & Wu, C. H. C. Chen, H. Huang, & C. H. Wu. (2017). Protein bioinformatics databases and resources. *Fundamentals of protein bioinformatics*. Vol. 1558, pp. 3-39. doi: 10.1007/978-1-4939-6783-4_1New York, NY: Humana Press.

Clemons, P. A., Bodycombe, N. E., Carrinski, H. A., Wilson, J. A., Shamji, A. F., Wagner, B. K., . . . Schreiber, S. L. (2010). Small molecules of different origins have distinct distributions of structural complexity that correlate with protein-binding profiles. *Proc. Natl. Acad. Sci. U. S. A., 107*(44), 18787. doi: 10.1073/pnas.1012741107

Daina, A., Michielin, O., & Zoete, V. (2019). SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. *Nucleic Acids Res., 47*(W1), W357-W364. doi: 10.1093/nar/gkz382

Dodds, E. C., Lawson, W., & Dale, H. H. (1938). Molecular structure in relation to oestrogenic activity. Compounds without a phenanthrene nucleus. *Proc. R. Soc. B, 125*(839), 222-232. doi: 10.1098/rspb.1938.0023

554 Doman, T. N., McGovern, S. L., Witherbee, B. J., Kasten, T. P., Kurumbail, R., Stallings, W. C.,
555 . . . Shoichet, B. K. (2002). Molecular docking and high-throughput screening for novel
556 inhibitors of protein tyrosine phosphatase-1B. *J. Med. Chem., 45*(11), 2213-2221. doi:
557 10.1021/jm010548w
558 Dunkel, M., Günther, S., Ahmed, J., Wittig, B., & Preissner, R. (2008). SuperPred: drug
559 classification and target prediction. *Nucleic Acids Res., 36*(suppl_2), W55-W59. doi:
560 10.1093/nar/gkn307
561 Dutka, M., Bobiński, R., Ulman-Włodarz, I., Hajduga, M., Bujok, J., Pająk, C., & Ćwiertnia, M.
562 (2019). Various aspects of inflammation in heart failure. *Heart Failure Rev.* doi:
563 10.1007/s10741-019-09875-1
564 Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., . . . D'Eustachio,
565 P. (2017). The reactome pathway knowledgebase. *Nucleic Acids Res., 46*(D1), D649-
566 D655. doi: 10.1093/nar/gkx1132
567 Ferreira, R. S., Simeonov, A., Jadhav, A., Eidam, O., Mott, B. T., Keiser, M. J., . . . Shoichet, B.
568 K. (2010). Complementarity between a docking and a high-throughput screen in
569 discovering new cruzain inhibitors. *J. Med. Chem., 53*(13), 4891-4905. doi:
570 10.1021/jm100488w
571 Gaucher, M., Dugé de Bernonville, T., Lohou, D., Guyot, S., Guillemette, T., Brisset, M.-N., &
572 Dat, J. F. (2013). Histolocalization and physico-chemical characterization of
573 dihydrochalcones: Insight into the role of apple major flavonoids. *Phytochemistry, 90*, 78-
574 89. doi: https://doi.org/10.1016/j.phytochem.2013.02.009
575 Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., . . . Overington, J. P.
576 (2012). ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids*
577 *Res., 40*(D1), D1100-D1107. doi: 10.1093/nar/gkr777
578 Gaulton, A., Hersey, A., Nowotka, M., Bento, A. P., Chambers, J., Mendez, D., . . . Leach, A. R.
579 (2017). The ChEMBL database in 2017. *Nucleic Acids Res., 45*(D1), D945-D954. doi:
580 10.1093/nar/gkw1074
581 Gfeller, D., Grosdidier, A., Wirth, M., Daina, A., Michielin, O., & Zoete, V. (2014).
582 SwissTargetPrediction: a web server for target prediction of bioactive small molecules.
583 *Nucleic Acids Res., 42*(W1), W32-W38. doi: 10.1093/nar/gku293

584 Gfeller, D., Michielin, O., & Zoete, V. (2013). Shaping the interaction landscape of bioactive
585 molecules. *Bioinformatics, 29*(23), 3073–3079. doi: 10.1093/bioinformatics/btt540
586 Günther, S., Kuhn, M., Dunkel, M., Campillos, M., Senger, C., Petsalaki, E., . . . Preissner, R.
587 (2008). SuperTarget and Matador: resources for exploring drug-target relationships.
588 *Nucleic Acids Res., 36*(suppl_1), D919-D922. doi: 10.1093/nar/gkm862
589 Harvey, A. L. (2008). Natural products in drug discovery. *Drug Discovery Today, 13*(19), 894-
590 901. doi: https://doi.org/10.1016/j.drudis.2008.07.004
591 Hattori, S. (2018). Anti-inflammatory effects of empagliflozin in patients with type 2 diabetes and
592 insulin resistance. *Diabetol. Metab. Syndr., 10*, 93-93. doi: 10.1186/s13098-018-0395-5
593 Hecker, N., Ahmed, J., von Eichborn, J., Dunkel, M., Macha, K., Eckert, A., . . . Preissner, R.
594 (2012). SuperTarget goes quantitative: update on drug–target interactions. *Nucleic Acids*
595 *Res., 40*(D1), D1113-D1117. doi: 10.1093/nar/gkr912
596 Hert, J., Keiser, M. J., Irwin, J. J., Oprea, T. I., & Shoichet, B. K. (2008). Quantifying the
597 relationships among drug classes. *J. Chem. Inf. Model., 48*(4), 755-765. doi:
598 10.1021/ci8000259
599 Huang, Y.-W., Pineau, I., Chang, H.-J., Azzi, A., Bellemare, V. r., Laberge, S., & Lin, S.-X.
600 (2001). Critical residues for the specificity of cofactors and substrates in human estrogenic
601 17β-hydroxysteroid dehydrogenase 1: variants designed from the three-dimensional
602 structure of the enzyme. *Mol. Endocrinol., 15*(11), 2010-2020. doi:
603 10.1210/mend.15.11.0730
604 Hurle, M. R., Yang, L., Xie, Q., Rajpal, D. K., Sanseau, P., & Agarwal, P. (2013). Computational
605 drug repositioning: from data to therapeutics. *Clin. Pharmacol. Ther., 93*(4), 335-341. doi:
606 10.1038/clpt.2013.1
607 Iannantuoni, F., M de Marañon, A., Diaz-Morales, N., Falcon, R., Bañuls, C., Abad-Jimenez, Z.,
608 . . . Rovira-Llopis, S. (2019). The SGLT2 inhibitor empagliflozin ameliorates the
609 inflammatory profile in type 2 diabetic patients and promotes an antioxidant response in
610 leukocytes. *J. Clin. Med., 8*(11), 1814. doi: 10.3390/jcm8111814
611 Jalencas, X., & Mestres, J. (2013). On the origins of drug polypharmacology. *MedChemComm,*
612 *4*(1), 80-87. doi: 10.1039/C2MD20242E

613 Keiser, M. J., Roth, B. L., Armbruster, B. N., Ernsberger, P., Irwin, J. J., & Shoichet, B. K. (2007).
614 Relating protein pharmacology by ligand chemistry. *Nat. Biotechnol., 25*(2), 197-206. doi:
615 doi:10.1038/nbt1284
616 Keiser, M. J., Setola, V., Irwin, J. J., Laggner, C., Abbas, A. I., Hufeisen, S. J., . . . Roth, B. L.
617 (2009). Predicting new molecular targets for known drugs. *Nature, 462*, 175. doi:
618 https://doi.org/10.1038/nature08506
619 Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., . . . Bolton, E. E. (2018). PubChem
620 2019 update: improved access to chemical data. *Nucleic Acids Res., 47*(D1), D1102-
621 D1109. doi: 10.1093/nar/gky1033
622 Koeberle, A., Siemoneit, U., Bühring, U., Northoff, H., Laufer, S., Albrecht, W., & Werz, O.
623 (2008). Licofelone suppresses prostaglandin E_2 formation by interference with the
624 inducible microsomal prostaglandin E_2 synthase-1. *J. Pharmacol. Exp. Ther., 326*(3), 975.
625 doi: 10.1124/jpet.108.139444
626 Koeberle, A., & Werz, O. (2014). Multi-target approach for natural products in inflammation.
627 *Drug Discovery Today, 19*(12), 1871-1882. doi:
628 https://doi.org/10.1016/j.drudis.2014.08.006
629 Koehn, F. E., & Carter, G. T. (2005). The evolving role of natural products in drug discovery. *Nat.*
630 *Rev. Drug Discovery, 4*(3), 206-220. doi: 10.1038/nrd1657
631 Kratschmar, D. V., Vuorinen, A., Da Cunha, T., Wolber, G., Classen-Houben, D., Doblhoff, O., .
632 . . Odermatt, A. (2011). Characterization of activity and binding mode of glycyrrhetinic
633 acid derivatives inhibiting 11β-hydroxysteroid dehydrogenase type 2. *J. Steroid Biochem.*
634 *Mol. Biol., 125*(1), 129-142. doi: https://doi.org/10.1016/j.jsbmb.2010.12.019
635 Le Bail, J.-C., Pouget, C., Fagnere, C., Basly, J.-P., Chulia, A.-J., & Habrioux, G. (2001).
636 Chalcones are potent inhibitors of aromatase and 17β-hydroxysteroid dehydrogenase
637 activities. *Life Sci., 68*(7), 751-761. doi: https://doi.org/10.1016/S0024-3205(00)00974-7
638 Lephart, E. D., & Simpson, E. R. (1991). Assay of aromatase activity. *Methods in Enzymology*.
639 Vol. 206, pp. 477-483. doi: https://doi.org/10.1016/0076-6879(91)06116-K: Academic
640 Press.
641 Liu, T., Lin, Y., Wen, X., Jorissen, R. N., & Gilson, M. K. (2007). BindingDB: a web-accessible
642 database of experimentally determined protein–ligand binding affinities. *Nucleic Acids*
643 *Res., 35*(suppl_1), D198-D201. doi: 10.1093/nar/gkl999

644 Lounkine, E., Keiser, M. J., Whitebread, S., Mikhailov, D., Hamon, J., Jenkins, J. L., . . . Urban,
645 L. (2012). Large-scale prediction and testing of drug activity on side-effect targets. *Nature, 486*(7403), 361-367. doi: 10.1038/nature11159
647 Margiotti, K., Kim, E., Pearce, C. L., Spera, E., Novelli, G., & Reichardt, J. K. V. (2002).
648 Association of the G289S single nucleotide polymorphism in the HSD17B3 gene with
649 prostate cancer in italian men. *The Prostate, 53*(1), 65-68. doi: 10.1002/pros.10134
650 Matsuura, K., Shiraishi, H., Hara, A., Sato, K., Deyashiki, Y., Ninomiya, M., & Sakai, S. (1998).
651 Identification of a principal mRNA species for human 3α-hydroxysteroid dehydrogenase
652 isoform (AKR1C3) that exhibits high prostaglandin D2 11-ketoreductase activity. *J. Biochem., 124*(5), 940-946. doi: 10.1093/oxfordjournals.jbchem.a022211
654 Mayr, F., Sturm, S., Ganzera, M., Waltenberger, B., Martens, S., Schwaiger, S., . . . Stuppner, H.
655 (2019a). Mushroom tyrosinase-based enzyme inhibition assays are not suitable for
656 bioactivity-guided fractionation of extracts. *J. Nat. Prod., 82*(1), 136-147. doi:
657 10.1021/acs.jnatprod.8b00847
658 Mayr, F., Vieider, C., Temml, V., Stuppner, H., & Schuster, D. A. D. Kinghorn, H. Falk, S.
659 Gibbons, J. i. Kobayashi, Y. Asakawa, & J.-K. Liu. (2019b). Open-access activity
660 prediction tools for natural products. Case study: hERG blockers. *Progress in the*
661 *Chemistry of Organic Natural Products 110: Cheminformatics in Natural Product*
662 *Research*. pp. 177-238. doi: 10.1007/978-3-030-14632-0_6Cham: Springer International
663 Publishing.
664 McKinney, W. (2010). *Data structures for statistical computing in python.* Paper presented at the
665 Proceedings of the 9th Python in Science Conference.
666 Meng, W., Ellsworth, B. A., Nirschl, A. A., McCann, P. J., Patel, M., Girotra, R. N., . . . Washburn,
667 W. N. (2008). Discovery of dapagliflozin: a potent, selective renal sodium-dependent
668 glucose cotransporter 2 (SGLT2) inhibitor for the treatment of type 2 diabetes. *J. Med. Chem., 51*(5), 1145-1149. doi: 10.1021/jm701272q
670 Möller, G., Deluca, D., Gege, C., Rosinus, A., Kowalik, D., Peters, O., . . . Hillisch, A. (2009).
671 Structure-based design, synthesis and in vitro characterization of potent 17β-
672 hydroxysteroid dehydrogenase type 1 inhibitors based on 2-substitutions of estrone and D-
673 homo-estrone. *Bioorg. Med. Chem. Lett., 19*(23), 6740-6744. doi:
674 https://doi.org/10.1016/j.bmcl.2009.09.113

675 Morisseau, C., Beetham, J. K., Pinot, F., Debernard, S., Newman, J. W., & Hammock, B. D.
676 (2000). Cress and potato soluble epoxide hydrolases: Purification, biochemical
677 characterization, and comparison to mammalian enzymes. *Arch. Biochem. Biophys.*,
678 *378*(2), 321-332. doi: https://doi.org/10.1006/abbi.2000.1810
679 Neuwirt, H., Bouchal, J., Kharaishvili, G., Ploner, C., Jöhrer, K., Pitterl, F., . . . Eder, I. E. (2020).
680 Cancer-associated fibroblasts promote prostate tumor growth and progression through
681 upregulation of cholesterol and steroid biosynthesis. *Cell Commun. Signaling, 18*(1), 11.
682 doi: 10.1186/s12964-019-0505-5
683 Newman, D. J., & Cragg, G. M. (2016). Natural products as sources of new drugs from 1981 to
684 2014. *J. Nat. Prod., 79*(3), 629-661. doi: 10.1021/acs.jnatprod.5b01055
685 Nickel, J., Gohlke, B.-O., Erehman, J., Banerjee, P., Rong, W. W., Goede, A., . . . Preissner, R.
686 (2014). SuperPred: update on drug classification and target prediction. *Nucleic Acids Res.*,
687 *42*(W1), W26-W31. doi: 10.1093/nar/gku477
688 Orlikova, B., Schnekenburger, M., Zloh, M., Golais, F., Diederich, M., & Tasdemir, D. (2012).
689 Natural chalcones as dual inhibitors of HDACs and NF-κB. *Oncol. Rep., 28*, 797-805. doi:
690 https://doi.org/10.3892/or.2012
691 Pandey, A. V., Kempná, P., Hofer, G., Mullis, P. E., & Flück, C. E. (2007). Modulation of human
692 CYP19A1 activity by mutant NADPH P450 oxidoreductase. *Mol. Endocrinol., 21*(10),
693 2579-2595. doi: 10.1210/me.2007-0245
694 Polgár, T., Baki, A., Szendrei, G. I., & Keserűu, G. M. (2005). Comparative virtual and
695 experimental high-throughput screening for glycogen synthase kinase-3β inhibitors. *J.*
696 *Med. Chem., 48*(25), 7946-7959. doi: 10.1021/jm050504d
697 Reker, D., Perna, A. M., Rodrigues, T., Schneider, P., Reutlinger, M., Mönch, B., . . . Schneider,
698 G. (2014). Revealing the macromolecular targets of complex natural products. *Nat. Chem.*,
699 *6*(12), 1072-1078. doi: 10.1038/nchem.2095
700 Ripphausen, P., Nisius, B., Peltason, L., & Bajorath, J. (2010). Quo vadis virtual Screening? A
701 comprehensive survey of prospective applications. *J. Med. Chem., 53*, 8461-8467. doi:
702 10.1021/jm101020z
703 Rivière, C. R. Atta ur. (2016). Chapter 7 - dihydrochalcones: occurrence in the plant kingdom,
704 chemistry and biological activities. *Studies in Natural Products Chemistry*. Vol. 51, pp.
705 253-381. doi: https://doi.org/10.1016/B978-0-444-63932-5.00007-3: Elsevier.

706 Rodrigues, T., Reker, D., Kunze, J., Schneider, P., & Schneider, G. (2015). Revealing the
707 macromolecular targets of fragment-like natural products. *Angew. Chem., Int. Ed., 54*(36),
708 10516-10520. doi: 10.1002/anie.201504241
709 Rodrigues, T., Reker, D., Schneider, P., & Schneider, G. (2016). Counting on natural products for
710 drug design. *Nat. Chem., 8*, 531. doi: 10.1038/nchem.2479
711 Rollinger, J. M. (2009). Accessing target information by virtual parallel screening—The impact
712 on natural product research. *Phytochem. Lett., 2*(2), 53-58. doi:
713 https://doi.org/10.1016/j.phytol.2008.12.002
714 Rollinger, J. M., Schuster, D., Danzl, B., Schwaiger, S., Markt, P., Schmidtke, M., . . . Stuppner,
715 H. (2009). In silico target fishing for rationalized ligand discovery exemplified on
716 constituents of *Ruta graveolens*. *Planta Med., 75*(03), 195-204. doi: 10.1055/s-0028-
717 1088397
718 Rush, T. S., Grant, J. A., Mosyak, L., & Nicholls, A. (2005). A shape-based 3-D scaffold hopping
719 method and its application to a bacterial protein−protein interaction. *J. Med. Chem., 48*(5),
720 1489-1495. doi: 10.1021/jm040163o
721 Samandari, E., Kempná, P., Nuoffer, J.-M., Hofer, G., E. Mullis, P., & E. Flück, C. (2007). Human
722 adrenal corticocarcinoma NCI-H295R cells produce more androgens than NCI-H295A
723 cells and differ in 3β-hydroxysteroid dehydrogenase type 2 and 17,20 lyase activities. *J.*
724 *Endocrinol., 195*(3), 459-472. doi: 10.1677/JOE-07-0166
725 Schaible, A. M., Filosa, R., Temml, V., Krauth, V., Matteis, M., Peduto, A., . . . Werz, O. (2014).
726 Elucidation of the molecular mechanism and the efficacy in vivo of a novel 1,4-
727 benzoquinone that inhibits 5-lipoxygenase. *Br. J. Pharmacol., 171*(9), 2399-2412. doi:
728 10.1111/bph.12592
729 Schuster, D. (2010). 3D pharmacophores as tools for activity profiling. *Drug Discovery Today:*
730 *Technol., 7*(4), e205-e211. doi: https://doi.org/10.1016/j.ddtec.2010.11.006
731 Schuster, D., Kowalik, D., Kirchmair, J., Laggner, C., Markt, P., Aebischer-Gumy, C., . . .
732 Adamski, J. (2011). Identification of chemically diverse, novel Inhibitors of 17 beta
733 hydroxysteroid dehydrogenase type 3 and 5 pharmacophore-based virtual screening. *J.*
734 *Steroid Biochem. Mol. Biol., 125*(1-2), 148-161.

735 Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., . . . Ideker, T. (2003).
736 Cytoscape: a software environment for integrated models of biomolecular interaction
737 networks. *Genome Res., 13*(11), 2498-2504. doi: 10.1101/gr.1239303
738 Sliwoski, G., Kothiwale, S., Meiler, J., Lowe, E. W., & Barker, E. L. (2014). Computational
739 methods in drug discovery. *Pharmacol. Rev., 66*(1), 334-395. doi:
740 https://doi.org/10.1124/pr.112.007336
741 Steindl, T., Schuster, D., Laggner, C., & Langer, T. (2006). Parallel Screening: A novel concept
742 in pharmacophore based modeling and virtual screening. *J. Chem. Inf. Model., 45*(3), 716-
743 724. doi: 10.1021/ci6002043
744 Sydow, D., Burggraaff, L., Szengel, A., van Vlijmen, H. W. T., Ijzerman, A. P., van Westen, G. J.
745 P., & Volkamer, A. (2019). Advances and challenges in computational target prediction.
746 *J. Chem. Inf. Model., 59*(5), 1728-1742. doi: 10.1021/acs.jcim.8b00832
747 Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., . . . Mering,
748 Christian v. (2018). STRING v11: protein–protein association networks with increased
749 coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic*
750 *Acids Research, 47*(D1), D607-D613. doi: 10.1093/nar/gky1131
751 The UniProt Consortium. (2018). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids*
752 *Res., 47*(D1), D506-D515. doi: 10.1093/nar/gky1049
753 Udhane, S. S., Parween, S., Kagawa, N., & Pandey, A. V. (2017). Altered CYP19A1 and CYP3A4
754 activities due to mutations A115V, T142A, Q153R and P284L in the human P450
755 oxidoreductase. *Front. Pharmacol., 8*(580). doi: 10.3389/fphar.2017.00580
756 Uthman, L., Baartscheer, A., Schumacher, C. A., Fiolet, J. W. T., Kuschma, M. C., Hollmann, M.
757 W., . . . Zuurbier, C. J. (2018). Direct cardiac actions of sodium glucose cotransporter 2
758 inhibitors target pathogenic mechanisms underlying heart failure in diabetic patients.
759 *Frontiers in Physiology., 9*(1575). doi: 10.3389/fphys.2018.01575
760 Vicker, N., Sharland, C. M., Heaton, W. B., Gonzalez, A. M. R., Bailey, H. V., Smith, A., . . .
761 Potter, B. V. L. (2009). The design of novel 17β-hydroxysteroid dehydrogenase type 3
762 inhibitors. *Mol. Cell. Endocrinol., 301*(1), 259-265. doi:
763 https://doi.org/10.1016/j.mce.2008.08.005
764 Waltenberger, B., Garscha, U., Temml, V., Liers, J., Werz, O., Schuster, D., & Stuppner, H.
765 (2016). Discovery of potent soluble Epoxide hydrolase (sEH) Inhibitors by

766 pharmacophore-based virtual screening. *J. Chem. Inf. Model.*, *56*, 747-762. doi:
767 10.1021/acs.jcim.5b00592
768 Wetzel, M., Marchais-Oberwinkler, S., Perspicace, E., Möller, G., Adamski, J., & Hartmann, R.
769 W. (2011). Introduction of an electron withdrawing group on the hydroxyphenylnaphthol
770 scaffold improves the potency of 17β-hydroxysteroid dehydrogenase type 2 (17β-HSD2)
771 inhibitors. *J. Med. Chem.*, *54*(21), 7547-7557. doi: 10.1021/jm2008453
772 Wixtrom, R. N., Silva, M. H., & Hammock, B. D. (1988). Affinity purification of cytosolic epoxide
773 hydrolase using derivatized epoxy-activated sepharose gels. *Anal. Biochem.*, *169*(1), 71-
774 80. doi: https://doi.org/10.1016/0003-2697(88)90256-4
775 Wolber, G., Dornhofer, A. A., & Langer, T. (2006). Efficient overlay of small organic molecules
776 using 3D pharmacophores. *J. Comput. Aided Mol. Des.*, *20*, 773-788. doi: 10.1007/s10822-
777 006-9078-7
778 Xu, C., Wang, W., Zhong, J., Lei, F., Xu, N., Zhang, Y., & Xie, W. (2018). Canagliflozin exerts
779 anti-inflammatory effects by inhibiting intracellular glucose metabolism and promoting
780 autophagy in immune cells. *Biochem. Pharmacol. (Amsterdam, Neth.)*, *152*, 45-59. doi:
781 https://doi.org/10.1016/j.bcp.2018.03.013
782 Young, S. M., Bologa, C., Prossnitz, E. R., Oprea, T. I., Sklar, L. A., & Edwards, B. S. (2005).
783 High-throughput screening with HyperCyt® flow cytometry to detect small molecule
784 formylpeptide receptor ligands. *J. Biomol. Screen.*, *10*(4), 374-382. doi:
785 10.1177/1087057105274532

786