Hybrid *in silico/in vitro* target fishing to assign function to "orphan" compounds of food origin – the case of the fungal metabolite atromentin

Luca Dellafiora[1], Georg Aichinger[2], Elena Geib[3], Leyre Sánchez-Barrionuevo[4], Matthias Brock[3], David Cánovas[4], Chiara Dall'Asta[1], Doris Marko[2],

[1] Department of Food and Drug, University of Parma, Via G.P. Usberti 27/A, 43124 Parma, Italy

[2] Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, Waehringer Str. 38, 1090 Vienna, Austria

[3] Fungal Genetics and Biology Groups, School of Life Sciences, University of Nottingham, University Park, NG7 2RD Nottingham, UK

[4] Department of Genetics, Faculty of Biology, University of Sevilla, 41012, Spain

Corresponding author: Chiara Dall'Asta, Department of Food and Drug, University of Parma, Via G.P. Usberti 27/A, 43124 Parma, Italy. Phone: +39 0521 905431. Email: chiara.dallasta@unipr.it.

Authors email list:

Luca Dellafiora luca.dellafiora@unipr.it; Georg Aichinger georg.aichinger@univie.ac.at; Elena Geib Elena.Geib@nottingham.ac.uk; Leyre Sánchez-Barrionuevo leyresan@gmail.com; Matthias Brock Matthias.Brock@nottingham.ac.uk; David Cánovas davidc@us.es; Chiara Dall'Asta chiara.dallasta@unipr.it; Doris Marko doris.marko@univie.ac.at.

Abstract

Many small molecules of food origin may effect human health but lack an adequate description of their biological activity. To fill this knowledge gap, a first-line workflow is needed to assign putative functions, hierarchize the endpoints to test, and guide wet-lab experiments. In this framework, the identification of potential biological targets can be used to probe the activity of orphan compounds using a so-called "target fishing" approach. Here, we present a proof of concept study using an *in silico/in vitro* target fishing approach on the fungal secondary metabolite atromentin. The procedure relies on a computational screening for activity identification coupled with experimental trials for dose-response characterization. Computational results identified estrogen receptors and 17-β-hydroxysteroid dehydrogenase as potential targets. Experiments confirmed a weak estrogenic activity, supporting the reliability of the procedure. Despite limited estrogenicity of atromentin, the proposed inhibition of 17-β-hydroxysteroid dehydrogenase should be considered as a source for endocrine disruptive effects.

1. Introduction

Food may be a source of exposure to many xenobiotics with potential biological or toxic activity. Such compounds may be inherently present as food constituents or they may be introduced, either voluntarily or not, along the food production chain (Rather, Koh, Paek, & Lim, 2017). Considering the huge number of compounds occurring in food, an adequate characterization of any possible biological/toxic activity is still lacking for many of them. Nevertheless, the overall identification of bioactive and toxic molecules in food is crucial to advance the current understanding of the effects on human health and wellbeing. In this respect, the understanding of small molecules profiling and fingerprinting is getting more and more detailed by taking advantage of -omic sciences such as proteomics and metabolomics (e.g. ref. (Scalbert, Brennan, Manach, Andres-Lacueva, Dragsted, Draper, et al., 2014)). However, the assignment of biological properties to such a wide number of molecules is currently challenging, albeit it is the primal step to assess the hazards or benefits to human health. Obviously, the risk assessment of toxicants and the assignment of health claims require the collection of a wealth of evidence including data on occurrence, exposure, bioavailability, bioaccessibility and both, the chronic and acute effects in living organisms. Nevertheless, the use of upstream methods to preliminarily assign a possible toxic or biological activity to molecules of interest, which is framed into the hazard/activity identification process, may allow a more effective setting up of further experimental investigations. Currently, there is an outbreak of novel molecular scaffolds as a result of the efforts of the international scientific community in discovering new fungal secondary metabolites, which makes the assignment of function an exceptional challenge. Thus, prompt work-flows to timely support decision making and to plan effective future investigations need to be implemented.

In this framework, the identification of the biological targets along with proposing possible

mechanisms of action may be an effective way to probe the activity of orphan compounds (i.e. compounds with poorly characterized activity and/or unknown biological targets). In this regard, the so defined "target fishing" is a consolidated approach in medicinal chemistry to discover unexpected targets for pharmacologically active molecules. Target fishing is commonly applied to investigate the polypharmacology of drugs (Cereto-Massagué, Ojeda, Valls, Mulero, Pujadas, & Garcia-Vallve, 2015; Chaudhari, Tan, Huang, & Zhang, 2017), in off-targets identification, and in drug repositioning (Delavan, Roberts, Huang, Bao, Tong, & Liu, 2017). Currently, target fishing studies mainly rely on proteomic-, genomic- and computational-based approaches (Jung & Kwon, 2015). In particular, the integrated use of *in silico* simulations and *in vitro* trials already proved to be effective in target fishing studies and in the bioactivity/toxicity assessment of small molecules (Dellafiora, Dall'Asta, & Galaverna, 2018; Scafuri, Varriale, Facchiano, D'Auria, Raggi, & Marabotti, 2017).

In this work, we investigated whether a target fishing approach can be successfully applied also in the field of food chemistry and toxicology by investigating the potential activity of poorly characterized molecules of food origin. A hybrid *in silico/in vitro* workflow composed by ligand-based virtual screening and structure-based molecular modelling studies coupled to experimental trials is presented. Atromentin (Figure 1), a fungal metabolite from involutin biosynthesis of *Paxillus involutus* (Braesel, Götze, Shah, Heine, Tauber, Hertweck, et al., 2015) with a poorly characterized activity/toxicity (NCBI) was selected as a proof of principle. Atromentin is a low molecular weight fungal pigment precursor with a polyphenolic and benzoquinone structure that can be found in basidiomycetes mushrooms, including some intended for human consumption (Hu & Liu, 2003; Ngoc Quang, Hashimoto, Hitaka, Tanaka, Nukada, Yamamoto, et al., 2003). More recently, it was also found as a secondary metabolite produced by the ascomycete *Aspergillus*

terreus (Hühner, Backhaus, Kraut, & Li, 2018), pointing to a broader distribution among the fungal kingdom. Even though some biological properties have been identified over the years (e.g. anticoagulant, apoptotic and antimicrobial activity) (Brewer, Jen, Jones, & Taylor, 1984; Khanna, Malone, Euler, & Brady, 1965; Kim & Lee, 2009), the body of evidence collected so far is still fragmentary, making atromentin a good candidate for the study we propose. This work identified biological targets of potential concern and possible mechanisms of action, extending the knowledge on atromentin activity and paving the ground for further investigations.

2. Material and methods

2.1 *In silico* modelling

2.1.1 Data collection and ligand-based virtual screening

The atromentin 3D structure was used as template to query a database of compounds with known activity and biological targets, aiming at identifying hits with degrees of physico-chemical similarities. The database was built using the ligands repository available in the RCSB PDB databank (https://www.rcsb.org) (Berman, Westbrook, Feng, Gilliland, Bhat, Weissig, et al., 2000) containing all the ligands bound to the proteins deposited in the databank. In particular, all non-redundant 3D structures of ligands (24,893 entries) were downloaded from the Ligand Expo Download page (http://ligand-expo.rcsb.org) in the chemical table file format (*sdf*; last database access in August, 23rd 2017). Considering that atromentin has a molecular weight of 324.3 g/mol, all compounds were pre-filtered taking only those with a molecular weight ranging between 200 and 500 g/mol (15,248 entries were selected). This subset of compounds was used for the ligand-based virtual screening using the FLAP software (https://www.moldiscovery.com) (Baroni,

Cruciani, Sciabola, Perruccio, & Mason, 2007). The default software setting was used to create the database and virtual screening was performed using the bit string mode to reduce the computation time.

2.1.2 Structure-based molecular modelling

The ligand-based virtual screening aimed at sorting the compounds according to the degree of similarity to atromentin. Notably, structural similarity may allow different compounds to share biological/toxicological endpoints as they may use the common chemical features to interact with the same biological targets (*vide infra*). Therefore, the interaction between atromentin and the biological targets of the best hits identified by the virtual screening was thoroughly investigated using a structure-based molecular modelling study. The structure-based study relied on the pharmacophoric analysis of the ligand binding sites and on docking simulations to investigate the molecular details of the ligand-target interaction.

Model preparations. The models for the alpha and beta isoforms of the estrogen receptor (ERα and ERβ, respectively), the 17-β-hydroxysteroid dehydrogenase (HSD) from *Curvularia lunata* and the human Type I HSD were derived from the crystallographic structures deposited in the RCSB PDB databank ([http://www.rcsb.org)](http://www.rcsb.org) (Berman, et al., 2000) with ID codes 2YJA (Phillips, Roberts, Schade, Bazin, Bent, Davies, et al., 2011), 3OLL (Möcklinghoff, Rose, Carraz, Visser, Ottmann, & Brunsveld, 2010), 4FJ1 (Cassetta, Stojan, Krastanova, Kristan, Brunskole Švegelj, Lamba, et al., 2017) and 3HB5 (Mazumdar, Fournier, Zhu, Cadot, Poirier, & Lin, 2009), respectively. All protein structures were processed using the Sybyl software, version 8.1 (www.certara.com) and the consistency of atom and bond types was checked as described previously (Dellafiora, Galaverna, Dall'Asta, & Cozzini, 2015). All co-crystallized ligands were removed, except for both 17-β-HSD enzymes that were computed in the holo form with the

NADPH cofactor.

Pharmacophoric analysis. The binding site of all models was defined using the Flapsite tool of the software FLAP (Baroni, Cruciani, Sciabola, Perruccio, & Mason, 2007) while the GRID algorithm (Carosati, Sciabola, & Cruciani, 2004) was used to investigate the corresponding pharmacophoric space. The DRY probe was used to describe potential hydrophobic interactions, while the sp2 carbonyl oxygen (O) and the neutral flat amino (N1) probes were used to describe the hydrogen bond acceptor and donor capacity of the target, respectively. All images were obtained using the software PyMol version 1.7 (http://www.pymol.org).

Docking simulations. The interaction between atromentin and the putative targets was investigated with a docking study. The software GOLD (Genetic Optimization for Ligand Docking) was chosen as it already showed good performances in computing ligand-target interactions and in estimating the activity of ligands (e.g. ref. (Dellafiora, Galaverna, & Dall'Asta, 2017). In particular, the higher the computational scores, the better the expected capability of ligands to arrange favourably within the ligand binding site. Software setting and docking protocol reported by Ehrlich (Ehrlich, Dellafiora, Mollergues, Dall'Asta, Serrant, Marin-Kuan, et al., 2015) were used. As an exception, the external scoring functions were omitted as the GOLD's internal scoring function GOLDScore succeeded in analysing the reference set of compounds (*vide infra*). Specifically, GOLDScore fitness considers the external (protein-ligand complex) and internal (ligand only) van der Waals energy, protein-ligand hydrogen bond energy and ligand torsional strain energy. In each docking study, the proteins were kept semi-flexible and the polar hydrogen atoms were set free to rotate. The ligands were set fully flexible. The maximum number of poses to generate for each ligand was set at 50. In all models considered, only the pose with the highest GOLDScore fitness was kept for each ligand, according to Dellafiora and co-workers (Dellafiora,

Galaverna, & Dall'Asta, 2017). Additionally, all analyses were done in triplicate to exclude non-causative scores assignment as GOLD implements a Lamarckian genetic algorithm that may introduce variability in the results.

2.3 *In vitro* experiments

2.3.1 Chemicals and reagents

Cell culture media and supplements were purchased from GIBCO Invitrogen (Karlsruhe, Germany), Sigma–Aldrich (Schnelldorf, Germany) and SARSTEDT AG & CO (Nuembrecht, Germany). 17-β- estradiol was purchased from Sigma-Aldrich (Schnelldorf, Germany), genistein from Extrasynthese (Genay, France) and ICI 182,780 from Tocris (Bristol, United Kingdom). Atromentin was purified ($\geq$ 99% by HPLC; Figure 1S, Supplementing material) from heterologous expression of the *P. involutus invA5* gene in *Aspergillus oryzae* as described below.

2.3.2 Cell culture

The estrogen-sensitive human endometrial adenocarcinoma cell line Ishikawa, expressing ERα and ERβ was selected for this study (Vejdovszky, Schmidt, Warth, & Marko, 2016) and purchased from ECACC (Wiltshire, United Kingdom). Cells were cultivated in Minimum Essential Medium (MEM), supplemented with 5% FBS (Fetal Bovine Serum), 1% L-glutamine and 1% penicillin/streptomycin under humidified conditions (37 °C, 5% CO_2). Experiments were performed in cells between passages 5 and 20 with the use of "assay medium", a phenol red free Dulbecco's Modified Eagle Medium containing the F-12 nutrient solution (DMEM/F-12), supplemented with 5% charcoal-dextran stripped foetal bovine serum to ensure low hormone levels (Fisher Scientific, Braunschweig, Germany). For all assays, the Ishikawa cells were seeded

at a density of 15,000 cells/well into a 96 well-plate 24 h before a 48 h incubation. All measurements were conducted as technical triplicates and repeated in at least three independent cellular preparations (biological replicates). Cells were routinely tested for the absence of mycoplasma with the PCR Mycoplasma Test Kit (PanReac AppliChem, Darmstadt, Germany).

2.3.3 Alkaline phosphatase assay

To evaluate the estrogenic activity, the transcriptional response of the cells was evaluated by measuring the catalytic activity of alkaline phosphatase. The natural estrogen 17-β-estradiol (1 nM) was tested in parallel as positive control. Solvent (1% DMSO) was used as negative control. As an additional control and to confirm that the results of the assay arose from estrogenic stimuli, the most effective concentrations of atromentin was co-incubated with 1 µM of the high affinity estrogen receptor antagonist ICI 182,780.

The alkaline phosphatase assay was conducted as previously described by Lehmann and co-workers (Lehmann, Wagner, & Metzler, 2006). 15,000 cells/well were seeded in 96-well plates and after 48 h of incubation, they were washed three times with PBS, and lysed thermally at −80 °C for 20 min. Cell lysates were then incubated at room temperature for five minutes, followed by addition of 50 µL of AlP buffer (5 mM 4-nitrophenylphosphate, 1 M diethanolamine, 0.24 mM $MgCl_2$, pH 9.8) to each well (5 min). Measurements of absorbance (405 nm) were performed every three minutes, for one hour, at 37 °C, on a PerkinElmer Victor[3V] multiple plate reader. Alkaline phosphatase activity was calculated as the slope of the curve obtained by the measurements over one hour. Final results were normalized to the solvent control.

2.4 Recombinant production and purification of atromentin

A heterologous expression system in *A. oryzae* was developed for recombinant expression

of the *P. involutus invA5* atromentin synthetase gene. The expression system bases on the previously described combination of the terrein biosynthesis regulator TerR and its terrein synthase target promoter P*terA* (Gressler, Hortschansky, Geib, & Brock, 2015). A construct containing the *terR* gene under control of the sugar inducible amylase promoter P*amyB* and the pyrithiamine resistance gene *ptrA* (Gressler, Hortschansky, Geib, & Brock, 2015) was used for transformation of *A. oryzae* RIB40. Strain OP12 containing a single full-length integration of the transcription factor construct was selected as expression platform strain. A synthetic version of the *P. involutus invA5* gene containing an *N*-terminal His-tag (Braesel, et al., 2015) was provided by D. Hoffmeister (Friedrich-Schiller University Jena, Germany) and used as a template for amplification with oligonucleotides invA5_Sm-X_fw (5'- CAT CAC AGC ACC ATG GGC AGC AGC CAT CAT CAT CAT C -3') and invA5_SM-X_rv (5'- ATC ACT GCT GCC ATG GTT ACA GAC CAC GAG CTT CGA G -3') using Phusion Hot Start II polymerase (Thermo Scientific). The PCR product was fused by *in vitro* recombination with an *Nco*I restricted SM-X vector containing a phleomycin resistance cassette (Gressler, Hortschansky, Geib, & Brock, 2015) using the InFusion HD cloning kit (Takara). The resulting *invA5* expression plasmid was used for transformation of the OP12 strain and transformants were regenerated in the presence of 30 µg/mL phleomycin. Independent transformants were screened for atromentin production and strain OP12_invA5_1 was selected for up-scaled atromentin production. *Aspergillus* minimal medium (Gressler, Hortschansky, Geib, & Brock, 2015) containing 2% soluble starch as inducing carbon source and 20 mM glutamine as nitrogen source was used. Eight 50 mL cultures were inoculated with 1×10^6 conidia/mL and incubated for 42 h at 28°C on a rotary shaker. The mycelium was removed by filtration over Miracloth filter gauze (Merck) and the combined culture filtrates were acidified with formic acid (0.2% final concentration). An XAD-16 resin was regenerated in

methanol, extensively washed with water and added to the acidified culture filtrate. The mixture was shaken for 10 min at 150 rpm to allow binding of atromentin to the resin. The clear supernatant was discarded and atromentin was eluted by repeated addition of methanol. Purity of atromentin was analysed by HPLC chromatography over an Eclipse XDB-C18 column, 5 μm, 4.6 × 150 mm (Agilent) on a Dionex UltiMate3000 HPLC (Thermo Fisher Scientific) as previously described (Geib & Brock, 2017). No other contaminating metabolite peaks were observed (Supporting materials, Figure 1S). Purified atromentin (about 120 mg) was dried under reduced pressure and stored as dried compound until further use.

2.5 Statistical analysis

All experiments were performed at least in triplicate ($n \geq 3$; at least three independent biological replicates, each performed as technical triplicates). Data are expressed as the mean ± standard error mean (SEM). Statistical analysis was performed using the Origin Pro 2016G software. The data was analysed by one-way ANOVA ($\alpha = 0.05$), followed by post hoc Fisher's LSD test ($\alpha = 0.05$).

3. Results and Discussion

3.1. Overview of the computational target fishing workflow

The workflow presented in this study aims at assigning function to molecules of interest based on the physico-chemical similarities and mechanistic analogies with already characterized compounds. The rationale behind this approach relies on the principle that similar compounds may compete for the same biological targets. Accordingly, they may have a degree of analogy in

initiating specific molecular events, as well as in eliciting similar biological effects (McKinney, Richard, Waller, Newman, & Gerberick, 2000). Once the reliability of the computational performance has been assessed, the *in silico* workflow presented in this study may serve as a first analytical screening tool to identify the possible activity of orphan compounds. In analogy to the hazard identification in the risk assessment process, the activity identification consists in describing qualitatively the biological effects that orphan compounds might have, without proving their actual incidence *in vivo*. Furthermore, when coupled with *in vitro* trials, it may provide activity characterization in terms of dose-effect response and mode of action. The overall workflow of the study is depicted in Figure 2.

In this work, atromentin was used as a proof of concept. Atromentin 3D structure was used as template for a ligand-based virtual screening querying a database of small molecules with known activity and already described biological targets. Then, the most similar compounds identified by the screening were inspected for the assigned biological targets prioritizing those related to either human macromolecules or pathways of sound pharmacological interest. Subsequently, the interaction with atromentin was assessed thoroughly at an atomistic level in docking studies. Then, *in vitro* experiments were performed to achieve activity characterization.

3.2 Computational results

3.2.1 Ligand-based virtual screening

The RCSB PDB repository (http://www.rcsb.org) (Berman, et al., 2000) was used to generate the database of compounds analysed with the ligand-based virtual screening (the results are reported in Table 1S, Supporting material). The PDB archive is a unique worldwide public repository of information about the 3D structure of biological macromolecules. It contains more

than 130,000 structures resolved mainly by X-ray crystallography, NMR spectroscopy and cryo-electron microscopy. Many macromolecules are in the bound state and therefore their structures are available along with the structure of more than 20,000 different low molecular weight compounds. The so defined "chemical components" groups a wide number of ligands *sensu stricto* (including enzyme substrates and inhibitors), metal ions and cofactors with known activity and well-described dynamics of target interaction. Using atromentin as a template (324.3 g/mol), all compounds with molecular weight ranging between 200 and 500 g/mol (15,248 entries in total) were used to carry out the ligand-based virtual screening. The screening was done using the FLAP software (see material and method for more details) and the output sorted according to the FLAP "distance score", which is an overall estimate of the divergence of compounds from the template (atromentin in this case) in terms of physico-chemical properties (the lower the score the more similar the compounds are).

The RCSB PDB archive collects macromolecules from a wealth of organisms from different kingdoms. Therefore, all ligands not associated to either human macromolecules or pathways of sound pharmacological interest were not taken into account. As shown in Table 1S, the best scored compound among those related to human macromolecules was the ligand having PDB ID 1W3. It is a polyphenol-derivative compound in complex with the human carrier transthyretin. However, the interaction between atromentin and transthyretin was not considered for further investigation as the binding to serum transporters can be reasonably expected as non-specific and it cannot be directly related to the onset of biological stimuli. The next most similar compound identified was the polyphenol genistein (PDB ID: GEN) (Figure 1) found bound to ten protein-ligand complex structures. Among them, the human estrogen receptor isoforms (ERα and ERβ) were considered with highest priority to further investigate the possible interaction with

atromentin as: i) estrogenic compounds of food origin may be of high relevance for human health; ii) a relevant xenoestrogenic activity might be expected on the basis of the molecular similarity between atromentin and the synthetic estrogen diethylstilbestrol (Figure 1); iii) the ligand-based screening highlighted a degree of similarity between atromentin and matairesinol (Table 1S), which is a well-described phytoestrogen (Kiyama, 2016). On this basis, the interaction between atromentin and both human ERα and ERβ was further investigated in molecular modelling studies.

Moreover, the 17-β-hydroxysteroid dehydrogenase (17-β-HSD) from the fungus *Curvularia lunata* was also found among the proteins in the bound state with genistein. This enzyme can be used in medicinal chemistry studies in preliminary screenings for possible inhibitors of the human 17-β-HSD enzymes (Cassetta, et al., 2017). In addition, genistein is also known to inhibit the human Type I 17-β-HSD enzymes (Blomquist, Lima, & Hotchkiss, 2005). Therefore, the interaction between atromentin and the fungal orthologous enzyme as well as the human Type I 17-β-HSD enzyme was assessed *in silico.*

It is worth to mention that many additional targets of interest could have been identified scrolling down the rank list of similar compounds. However, the systematic analysis of all possible targets requires huge computational resources and therefore was not considered in this proof of concept study. Therefore, potential interaction with additional targets not included in this work cannot be ruled out.

3.2.2 Docking studies

On the basis of the virtual screening results, the interaction between atromentin and some potential biological targets was further assessed by docking studies. As shown previously, docking scores may reliably estimate the thermodynamic favours of the host-guest complex formation correlating with the energy of binding, with the inhibitory activity of competitive inhibitors (e.g.

Dellafiora, et al., 2015;) and with the xenoestrogenic activity of low molecular-weight compounds (Dellafiora, Warth, Schmidt, Del Favero, Mikula, Fröhlich, et al., 2018; Ehrlich, et al., 2015). Nevertheless a fit-for-purpose validation of procedural performances was done to assess the case-specific reliability. To this end, the capability to successfully compute a set of reference compounds in comparison to a benchmark ligand was tested for each target.

Estrogen receptors. The results of docking simulation on both the ERα and ERβ are reported in Table 1. The endogenous ligand 17β-estradiol, the polyphenol genistein and the well-known synthetic estrogen diethylstilbestrol were selected as reference compounds to test the performance of the procedure. Genistein was used as the benchmark for computing the relative activity of compounds. The procedure proved to be reliable as: i) it succeeded in calculating the binding architecture of the reference compounds (Figure 2S); ii) reference compounds were correctly ranked for their relative potency in comparison to the benchmark compound (Table 1).

Atromentin showed in both ERs a binding architecture resembling that of the other compounds with a marked shape overlapping with the synthetic estrogen diethylstilbestrol (Figure 3A). However, atromentin recorded the lowest scores among all the compounds tested in both ER isoform models pointing to a worse fitting into the ligand binding site. The pharmacophoric analysis of the ligand pocket and the comparison between the binding pose of atromentin and the mode of binding reported for the ER ligands provided a structural explanation. ER ligands need to interact with the two polar ends of the pocket (formed by Glu353, Arg394 and His524) and have to satisfy the hydrophobic environment with hydrophobic groups (Spyrakis & Cozzini, 2009). However, even though the phenolic groups of atromentin may engage the polar ends of the ERs pocket, the central benzoquinone moiety bears 4 polar groups that locally reduce the lipophilic characteristic of the molecule causing hydrophobic/hydrophilic interferences (Figure 3B).

On the basis of these results, ERs were identified as putative targets of atromentin and a possible estrogenic activity could be expected accordingly. Nevertheless, on account of the low scores recorded in both ER isoform and the non-optimal fitting into the ERs, the interaction with the ERs can be considered as less favoured than that of the benchmark compound analysed. Therefore, a lower efficacy in triggering an ER-dependent stimulus can be expected, as further confirmed *in vitro* (see below).

17-β-HSDs. According to the results of the virtual-screening, the interaction between atromentin and the fungal 17-β-HSD was investigated. In this respect, it has already been reported that the procedure used in this study may successfully estimate the inhibitory activity of competitive inhibitors (Dellafiora, et al., 2015). Therefore, it may also succeed in estimating the possible inhibitory activity of atromentin. In order to assess the procedural performance, genistein and the two other polyphenols kaempferol and biochanin A were taken as reference compounds since both the geometry of binding and the inhibitory activity have already been assessed (Cassetta, et al., 2017). Genistein was accounted as the benchmark compound. The computed binding geometry of reference compounds was in strong agreement with the crystallographic data collected so far (Figure 3S). In addition, as shown in Table 2, reference compounds were correctly ranked for the relative activity in comparison to the benchmark compound genistein. On this basis, the procedure could be considered reliable in providing a plausible geometry of binding for atromentin and in estimating its potential relative activity in comparison to genistein.

As shown in Table 2, atromentin recorded a score close to that of genistein pointing to a comparable capability to fit into the enzyme pocket. Concerning the geometries of binding, the crystallographic data collected so far revealed a unique binding pose for genistein and kaempferol with the B ring directed toward the binding site entrance (Figure 3C). Conversely, biochanin A

may have two possible modes of arranging into the binding pocket: one retraces the mode of binding of genistein and kaempferol (the one predicted by the molecular modelling), while the other presents an opposite direction of the B ring that is more embedded into the binding site close to the nicotinamide moiety of the prostetic group NADPH (Cassetta, et al., 2017). This eventually suggests that polyphenols may have alternative modes in arranging within the enzyme binding site. Atromentin retraced the binding mode of genistein keeping the same π-π interaction with the side chain of Tyr212. However, the longer length of the chemical structure of atromentin caused a more pronounced sinking of the aromatic ring towards the bottom of the pocket engaging via hydrogen bond of the side chain of Tyr167 (Figure 3D). In addition, atromentin satisfied the pharmacophoric fingerprint of the binding site, which consists in a central lipophilic environment with spaces able to receive hydrophilic groups on the two sides. Indeed, the central hydrophobic core of the molecule is embedded in the lipophilic space, while the polar groups are placed in the hydrophilic environment (Figure 3E). On the basis of these results the interaction between atromentin and the fungal 17-β-HSD can be considered as possible and a certain degree of inhibitory activity can be expected accordingly. It is noteworthy however that the inhibitory activity of genistein towards the fungal enzyme is quite weak and therefore, on the basis of computational scores, a mild activity can be expected also for atromentin.

Subsequently, the interaction between atromentin and the human Type I 17-β-HSD was assessed in account of: i) the positive interaction between atromentin and the fungal orthologous; ii) the similarity found by the virtual screening between atromentin and genistein, which is a known inhibitor of the human Type I 17-β-HSD (Schuster, Nashev, Kirchmair, Laggner, Wolber, Langer, et al., 2008). Genistein, the strong inhibitor 3-{[(9beta,14beta,16alpha,17alpha)-3,17-dihydroxyestra- 1,3,5(10)-trien-16-yl]methyl}benzamide (PDB ID E2B) and flavanone were

selected as reference compounds to assess the procedural performance. Genistein was accounted as the benchmark. As seen for the fungal enzyme, the procedure succeeded in calculating the geometry of binding (Figure S2) and in ranking the reference compounds in accordance to the inhibitory activity reported so far (Table 2).

As shown in Table 2, atromentin recorded a score higher than genistein pointing to an overall better fitting into the pocket. The pharmacophoric fingerprint of the pocket was similar to that of the fungal enzyme (i.e. a central lipophilic environment with hydrophilic regions on the sides), albeit the hydrophilic regions were found to be less extended. The geometry of binding of genistein into the human enzymes has not been resolved yet. However, in accordance to the geometry of binding proposed for polyphenols by Schuster and co-workers (Schuster, et al., 2008), genistein was found contacting via hydrogen bonds Glu284 with the hydroxyl groups on the A ring and the Ser142 with the hydroxyl group on the B ring (Figure 3F). The consistence with the geometry of binding previously reported may strengthen further the reliability of the predicted geometries. Also an additional contact with Tyr218 was found, while the aromatic parts were arranged into the lipophilic environment of the pocket. Similar to genistein, atromentin engaged via hydrogen bonds with Glu284 and Tyr218, but not Ser142. In addition, as seen in the fungal enzyme, the more pronounced sinking into the pocket allowed atromentin to contact Tyr155 via hydrogen bonding (Figure 3G). Also in this case, the hydrophobic aromatic scaffolds were found to be embedded in the lipophilic environment, while the polar groups were found mainly arranged within a hydrophilic environment. Therefore, atromentin succeeded in fitting the pharmacophoric requirements of the Type I human 17-β-HSDs pocket.

Taking into account the computational score and the pharmacophoric fitting into the enzyme pocket, the interaction between the human Type I 17-β-HSD enzyme and atromentin can

be considered as possible. A degree of inhibitory activity can be hypothesized accordingly. In particular, a higher inhibitory activity than genistein might be hypothesized on the basis of the higher score recorded. It is worth to note that genistein already has a good inhibitory activity towards the human Type I 17-β-HSD in comparison to the other polyphenols tested (Poirier, 2003; Schuster, et al., 2008). Notably, the human Type 1 17-β-HSD is a relevant enzyme in human physiology as it plays critical role in balancing the conversion of estrone to 17β-estradiol and *vice versa* (Mazumdar, Fournier, Zhu, Cadot, Poirier, & Lin, 2009). Along with the other human isoforms, Type 1 17-β-HSD has been accounted in drug development for the therapeutic treatment of a number of hydroxysteroid-sensitive pathologies (Poirier, 2003). Therefore, the applied workflow identified the possible inhibitory activity towards Type 1 17-β-HSD as an activity to be tested with priority in further studies to deepen the possible biological roles of atromentin.

3.3 Experimental assessment of estrogenic activity

The *in silico* prediction of the possible estrogenic activity of atromentin was supported by an *in vitro* assessment of estrogenicity with a reporter gene assay in Ishikawa cells using the estrogenic impact on alkaline phosphatase expression as read-out. In particular, the dose-response activity of atromentin was assessed at concentrations ranging from 0.05 to 250 µM and compared to that of the benchmark compound genistein at concentrations ranging from 0.001 to 1 µM. The activity of the endogenous ligand 17β-estradiol (1 nM) was taken as a reference. Both atromentin and genistein showed a dose-response effect (Figure 4A). However, no significant activity for atromentin was detected below 100 µM, while genistein showed significant response starting from 0.05 µM ($P < 0.001$). In addition, the estrogen receptor-dependent activity of atromentin was tested by co-incubating the most effective concentration with the ER antagonist ICI 182,780 (1 µM). As shown in Figure 4B, the co-incubation was able to significantly suppress the activity of atromentin

($P < 0.001$) supporting the ER-dependent mechanism of action identified by the *in silico* study. On the basis of this result, the mild estrogenic activity of atromentin *via* the ER-dependent mechanism found *in silico* was confirmed *in vitro.*

4. Conclusions

A target fishing study using ligand-based virtual screening and structure-based analysis, coupled with *in vitro* trials, has been presented. In particular, once the reliability of the computational procedure was assessed, the workflow allowed the identification of potential activities for further testing. However, when coupled with *in vitro* experiments, the procedure may also provide an activity characterization in terms of dose-response effects. The analysis of the orphan fungal secondary metabolite atromentin was selected as a proof of principle. Ligand-based screening identified several putative biological targets and the interaction with some of those relevant for human health had been investigated thoroughly with docking studies. According to the computational results, estrogen receptors and 17-β-HSD were identified possible relevant targets of atromentin. Therefore, in the framework of activity identification, degrees of estrogenicity via ERs activation and inhibition of 17-β-HSD can be expected. The computational data on estrogenicity were supported by the *in vitro* assessment using an alkaline phosphatase activity assay in Ishikawa cells, which confirmed the weak potency of atromentin. Nevertheless, in spite of the mild activity observed, broader effects on the enzymes that regulate estrogens production cannot be excluded on account of the results obtained for 17-β-HSDs. Therefore, in the light of deepening possible biological/toxic roles of atromentin in living organisms, the effects on estrogen production and regulation, along with those on the estrogen-mediated responses, have to be investigated with priority.

In addition, the results collected proved the procedure reliable in assigning biological functions. Accordingly, hybrid *in silico/in vitro* target fishing studies can be accounted as a valuable tool to investigate the unexpected activity of molecules of food origin, as well as to assign possible biological targets to orphan compounds. Therefore, we propose this approach as a first-line analytical tool to widen the current understanding of the bioactive/toxic composition of food constituents.

Conflict of Interest

The authors declare no conflict of interest.

References

- Baroni, M., Cruciani, G., Sciabola, S., Perruccio, F., & Mason, J. S. (2007). A common reference framework for analyzing/comparing proteins and ligands. Fingerprints for Ligands and Proteins (FLAP): theory and application. *Journal of Chemical Information*

and Modeling, 47(2), 279-294.

▪ 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*(1), 235-242.

▪ Blomquist, C. H., Lima, P. H., & Hotchkiss, J. R. (2005). Inhibition of 3alpha-hydroxysteroid dehydrogenase (3alpha-HSD) activity of human lung microsomes by genistein, daidzein, coumestrol and C(18)-, C(19)- and C(21)-hydroxysteroids and ketosteroids. *Steroids, 70*(8), 507-514.

▪ Braesel, J., Götze, S., Shah, F., Heine, D., Tauber, J., Hertweck, C., Tunlid, A., Stallforth, P., & Hoffmeister, D. (2015). Three Redundant Synthetases Secure Redox-Active Pigment Production in the Basidiomycete *Paxillus involutus*. *Chemistry & Biology, 22*(10), 1325-1334.

▪ Brewer, D., Jen, W. C., Jones, G. A., & Taylor, A. (1984). The antibacterial activity of some naturally occurring 2,5-dihydroxy-1,4-benzoquinones. *Canadian Journal of Microbiology, 30*(8), 1068-1072.

▪ Carosati, E., Sciabola, S., & Cruciani, G. (2004). Hydrogen bonding interactions of covalently bonded fluorine atoms: from crystallographic data to a new angular function in the GRID force field. *Journal of Medicinal Chemistry, 47*(21), 5114-5125.

▪ Cassetta, A., Stojan, J., Krastanova, I., Kristan, K., Brunskole Švegelj, M., Lamba, D., & Lanišnik Rižner, T. (2017). Structural basis for inhibition of 17β-hydroxysteroid dehydrogenases by phytoestrogens: The case of fungal 17β-HSDcl. *The Journal of Steroid Biochemistry and Molecular Biology, 171*, 80-93.

▪ 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.

- Chaudhari, R., Tan, Z., Huang, B., & Zhang, S. (2017). Computational polypharmacology: a new paradigm for drug discovery. *Expert Opinion on Drug Discovery, 12*(3), 279-291.

- Delavan, B., Roberts, R., Huang, R., Bao, W., Tong, W., & Liu, Z. (2017). Computational drug repositioning for rare diseases in the era of precision medicine. *Drug Discovery Today, in press*.

- Dellafiora, L., Dall'Asta, C., & Galaverna, G. (2018). Toxicodynamics of Mycotoxins in the Framework of Food Risk Assessment-An In Silico Perspective. *Toxins, 10*(2), E52.

- Dellafiora, L., Galaverna, G., & Dall'Asta, C. (2017). An in silico perspective on the toxicodynamic of tetrodotoxin and analogues - A tool for supporting the hazard identification. *Toxicon, 138*, 107-118.

- Dellafiora, L., Galaverna, G., Dall'Asta, C., & Cozzini, P. (2015). Hazard identification of cis/trans-zearalenone through the looking-glass. *Food and Chemical Toxicology, 86*, 65-71.

- Dellafiora, L., Marchetti, M., Spyrakis, F., Orlandi, V., Campanini, B., Cruciani, G., Cozzini, P., & Mozzarelli, A. (2015). Expanding the chemical space of human serine racemase inhibitors. *Biorganic & Medicinal Chemistry Letters, 25*(19), 4297-4303.

- Dellafiora, L., Warth, B., Schmidt, V., Del Favero, G., Mikula, H., Fröhlich, J., & Marko, D. (2018). An integrated in silico/in vitro approach to assess the xenoestrogenic potential of Alternaria mycotoxins and metabolites. *Food Chemistry, 248*, 253-261.

- Ehrlich, V. A., Dellafiora, L., Mollergues, J., Dall'Asta, C., Serrant, P., Marin-Kuan, M., Lo Piparo, E., Schilter, B., & Cozzini, P. (2015). Hazard assessment through hybrid in vitro/in silico approach: The case of zearalenone. *ALTEX, 32*(4), 275-286.

- Geib, E., & Brock, M. (2017). ATNT: an enhanced system for expression of polycistronic secondary metabolite gene clusters in *Aspergillus niger*. *Fungal Biology & Biotechnology, 4*, 13.

- Gressler, M., Hortschansky, P., Geib, E., & Brock, M. (2015). A new high-performance heterologous fungal expression system based on regulatory elements from the *Aspergillus terreus* terrein gene cluster. *Frontiers in Microbiology, 6*, 184.

- Hu, L., & Liu, J. K. (2003). p-Terphenyls from the basidiomycete *Thelephora aurantiotincta*. *Z. Naturforsch., 58*(5-6), 452-454.

- Hühner, E., Backhaus, K., Kraut, R., & Li, S. M. (2018). Production of α-keto carboxylic acid dimers in yeast by overexpression of NRPS-like genes from *Aspergillus terreus*. *Applied Microbiology and Biotechnology, 102*(4), 1663-1672.

- Jung, H. J., & Kwon, H. J. (2015). Target deconvolution of bioactive small molecules: the heart of chemical biology and drug discovery. *Archives of Pharmacal Research, 38*(9), 1627-1641.

- Khanna, J. M., Malone, M. H., Euler, K. L., & Brady, L. R. (1965). Atromentin, anticoagulant from Hydnellum diabolus. *Journal of Pharmaceutical Sciences, 54*(7), 1016-1020.

- Kim, J. H., & Lee, C. H. (2009). Atromentin-induced apoptosis in human leukemia U937 cells. *Journal of Microbiology and Biotechnology, 19*(9), 946-950.

- Kiyama, R. (2016). Biological effects induced by estrogenic activity of lignans. *Trends in Food Science & Technology, 54*, 186-196.

- Laplante, Y., Cadot, C., Fournier, M. A., & Poirier, D. (2008). Estradiol and estrone C-16 derivatives as inhibitors of type 1 17beta-hydroxysteroid dehydrogenase: blocking of ER+

breast cancer cell proliferation induced by estrone. *Bioorganic & Medicinal Chemistry, 16*(4), 1849-1860.

- 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 17beta-HSD type 1: a lead compound for breast cancer therapy. *The Biochemical Journal, 424*(3), 357-366.

- McKinney, J. D., Richard, A., Waller, C., Newman, M. C., & Gerberick, F. (2000). The practice of structure activity relationships (SAR) in toxicology. *Toxicological Sciences, 56*(1), 8-17.

- Möcklinghoff, S., Rose, R., Carraz, M., Visser, A., Ottmann, C., & Brunsveld, L. (2010). Synthesis and crystal structure of a phosphorylated estrogen receptor ligand binding domain. *Chembiochem, 11*(16), 2251-2254.

- NCBI. PubChem Compound Database (https://pubchem.ncbi.nlm.nih.gov) CID=99148; Last database access 7th March 2018.

- Ngoc Quang, D., Hashimoto, T., Hitaka, Y., Tanaka, M., Nukada, M., Yamamoto, I., & Asakawa, Y. (2003). Thelephantins D-H: five p-terphenyl derivatives from the inedible mushroom Thelephora aurantiotincta. *Phytochemistry, 63*(8), 919-924.

- Phillips, C., Roberts, L., Schade, M., Bazin, R., Bent, A., Davies, N., Moore, R., Pannifer, A., Pickford, A., Prior, S., Read, C., Scott, A., Brown, D., Xu, B., & Irving, S. (2011). Design and structure of stapled peptides binding to estrogen receptors. *Journal of the American Chemical Society, 133*(25).

- Poirier, D. (2003). Inhibitors of 17 beta-hydroxysteroid dehydrogenases. *Current Medicinal Chemistry, 10*, 453-477.

- Rather, I. A., Koh, W. Y., Paek, W. K., & Lim, J. (2017). The Sources of Chemical

Contaminants in Food and Their Health Implications. *Frontiers in Pharmacology, 8*, 830.

- Scafuri, B., Varriale, A., Facchiano, A., D'Auria, S., Raggi, M. E., & Marabotti, A. (2017). Binding of mycotoxins to proteins involved in neuronal plasticity: a combined in silico/wet investigation. *Scientific Reports, 7*(1), 15156.

- Scalbert, A., Brennan, L., Manach, C., Andres-Lacueva, C., Dragsted, L. O., Draper, J., Rappaport, S. M., van der Hooft, J. J., & Wishart, D. S. (2014). The food metabolome: a window over dietary exposure. *The American Journal of Clinical Nutrition, 99*(6), 1286-1308.

- Schopfer, U., Schoeffter, P., Bischoff, S. F., Nozulak, J., Feuerbach, D., & Floersheim, P. (2002). Toward selective ERbeta agonists for central nervous system disorders: synthesis and characterization of aryl benzthiophenes. *Journal of Medicinal Chemistry, 45*(7), 1399-1401.

- Schuster, D., Nashev, L. G., Kirchmair, J., Laggner, C., Wolber, G., Langer, T., & Odermatt, A. (2008). Discovery of nonsteroidal 17beta-hydroxysteroid dehydrogenase 1 inhibitors by pharmacophore-based screening of virtual compound libraries. *Journal of Medicinal Chemistry, 51*(14), 4188-4199.

- Spyrakis, F., & Cozzini, P. (2009). How computational methods try to disclose the estrogen receptor secrecy--modeling the flexibility. *Current Medicinal Chemistry, 16*(23), 2987-3027.

- Vejdovszky, K., Schmidt, V., Warth, B., & Marko, D. (2016). Combinatory estrogenic effects between the isoflavone genistein and the mycotoxins zearalenone and alternariol in vitro. *Molecular Nutrition & Food Research, in press*.
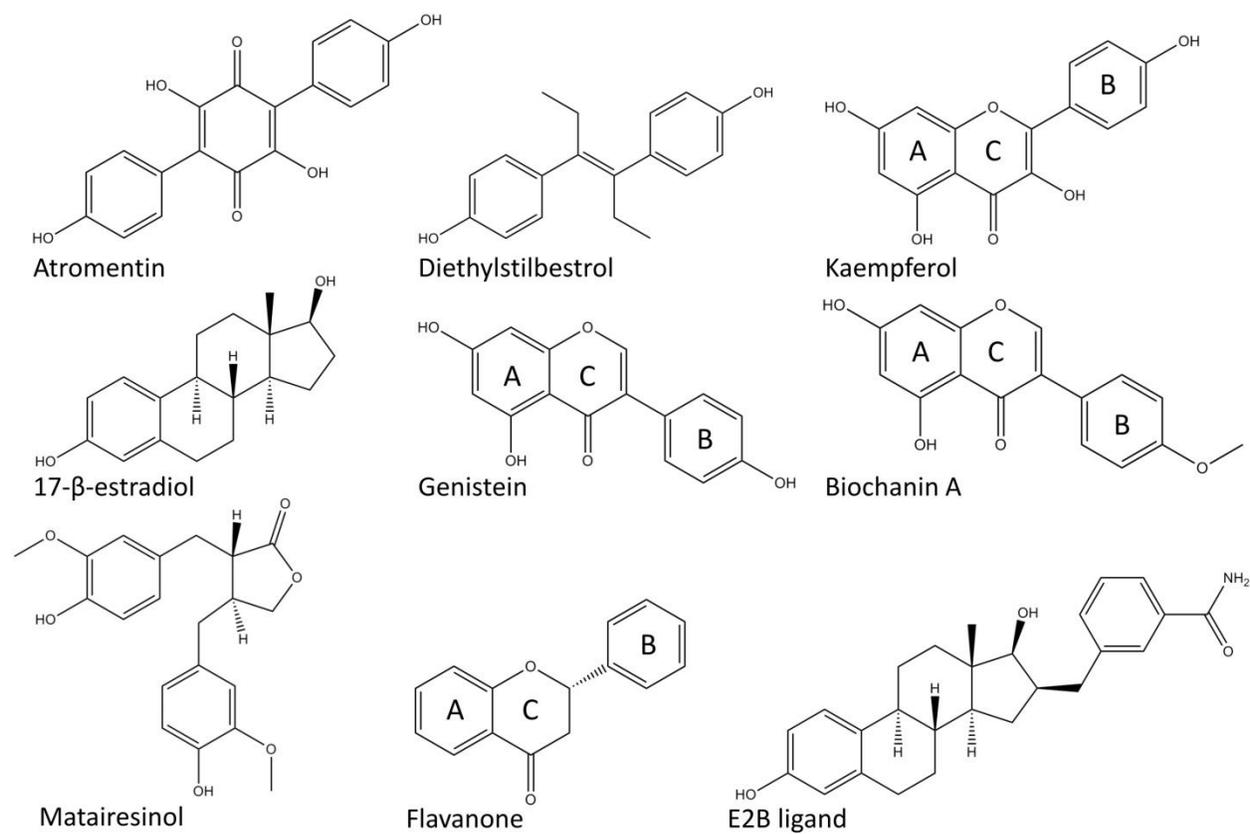
Figures and captions

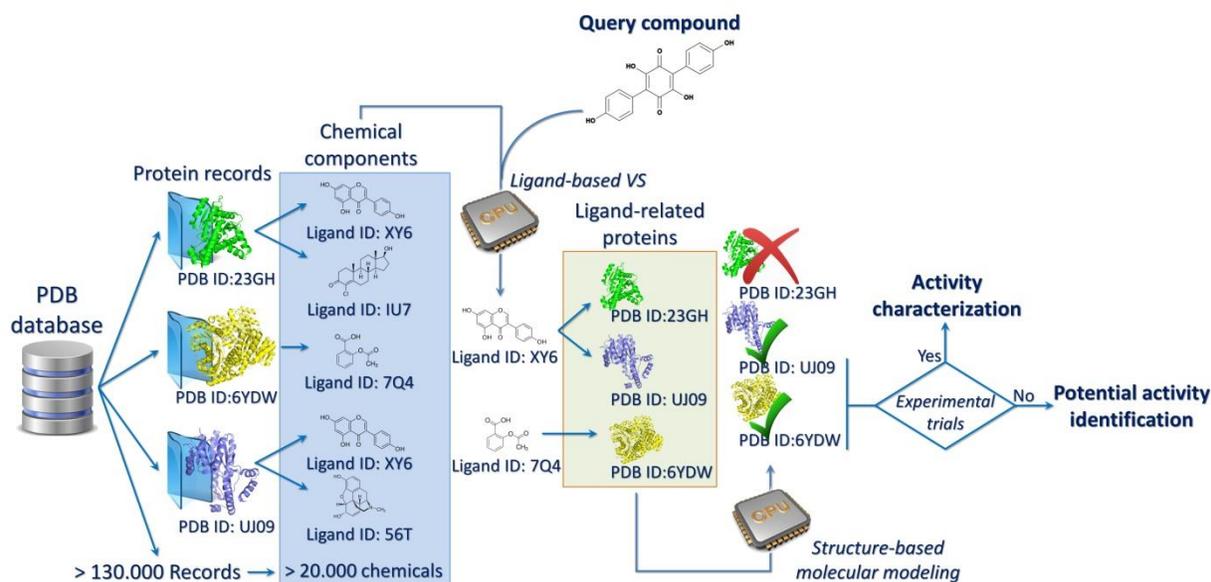**Figure 1.** Structures of the molecules mentioned in the work.

Figure 2. Schematic representation of the workflow used in the target fishing study. The low-molecular weight compounds belonging to the so called "chemical components" repository of the RCSB PDB archive (https://www.rcsb.org/) are accounted for a ligand-based virtual screening study. This first-line screening aims at identifying compounds sharing a degree of physico-chemical similarity with the query compound atromentin. Each low-molecular weight compound occurs as ligand in at least one macromolecule-ligand complex. Therefore, once the most similar compounds are identified, the interaction between the respective macromolecular structures and atromentin is thoroughly investigated with docking studies. The prediction of the atromentin-target interaction may lead to the identification of a potential activity, while the experimental assessment will allow characterization of the biological activity.
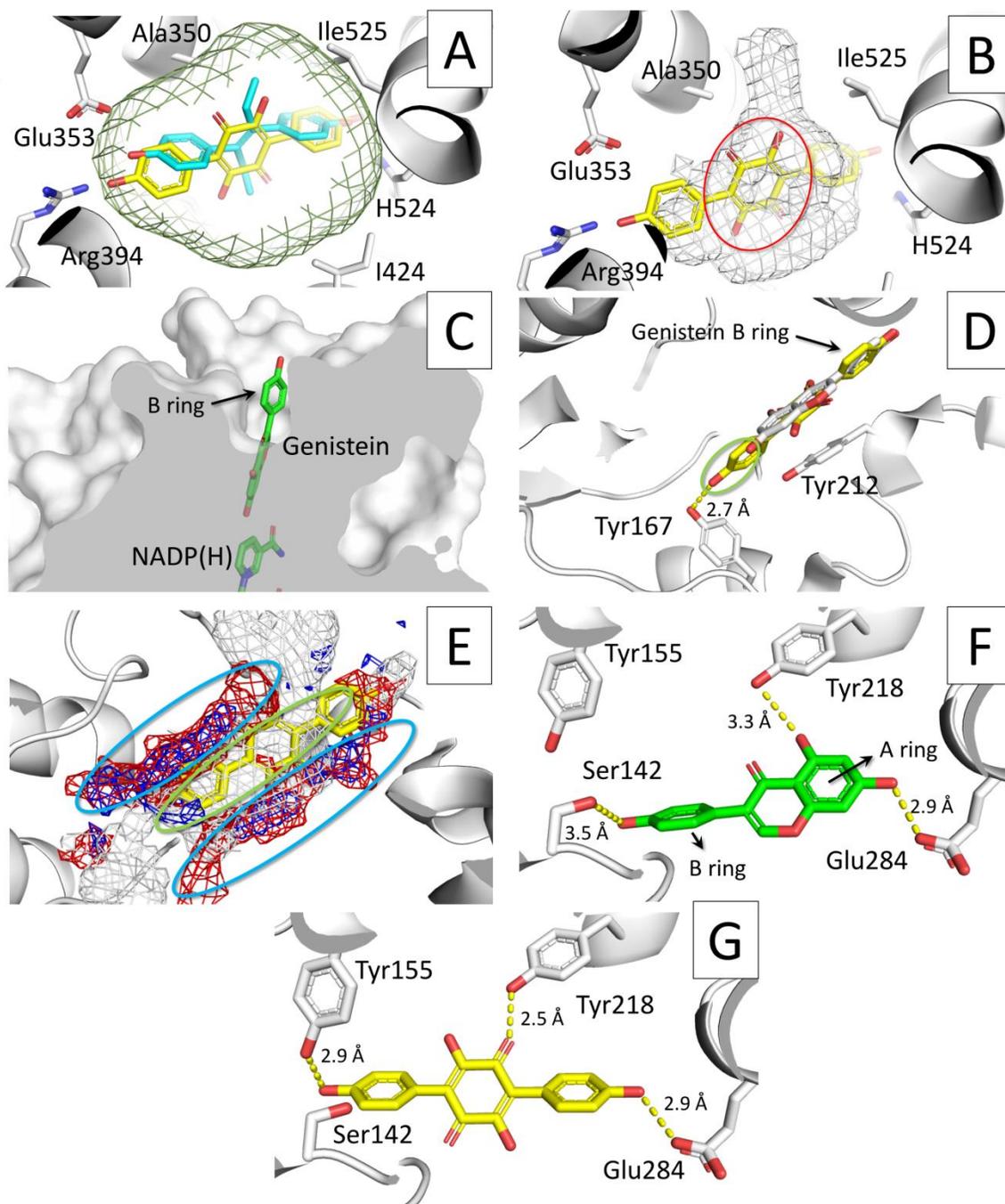
Figure 3. Binding geometry of atromentin into the ER pocket (A and B), and atromentin and genistein into fungal and human 17-β-HSDs (C-G). Due to the high conservation between ER isoforms, only the interaction with the ERα pocket is shown. Proteins are represented in in cartoon (A, B, D-G) or in cut surface (C), while ligands and amino acids side chains are represented in

sticks. Hydrogen bonds are indicated with yellow dashed lines. **A.** Binding pose of atromentin (yellow) overlapping diethylstilbestrol (cyan). The green mesh represents the shape of the binding site. **B.** Binding pose of atromentin. The white mesh indicates the hydrophobic environment of the pocket. The red circle highlights the unfavourable arrangement of the polar core of atromentin. **C.** Crystallographic pose of genistein into the fungal enzyme (PDB code 4FJ1; (Cassetta, et al., 2017)). **D.** Computational pose of atromentin (yellow) into the fungal enzyme superimposed to the crystallographic pose of genistein (white). The ring indicates the molecular region sinking towards the bottom of the binding site. **E.** Binding pose of atromentin in comparison to the pharmacophoric fingerprint of the biding site. White, red and blue contours identify regions sterically and energetically favourable for hydrophobic, H-bond acceptors and H-bond donor groups, respectively. Blue and green rings indicate the position of the polar and hydrophobic regions of atromentin, respectively. **F.** Computed architecture of genistein. **G.** Computed architecture of atromentin.
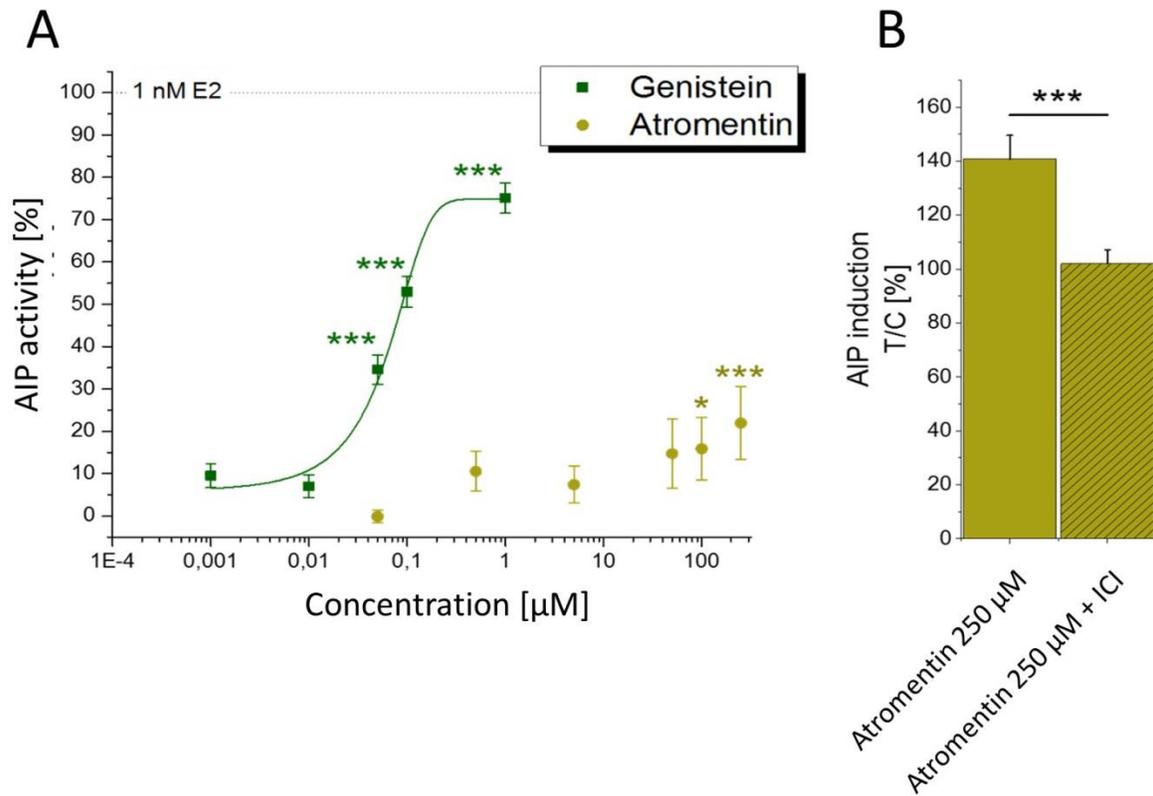
Figure 4. Estrogenic activity of atromentin assessed *in vitro*. **A.** Estrogenic activity of atromentin in comparison to genistein as measured after 48 h of incubation with the AlP assay in Ishikawa cells. The graph shows the increase of AlP induction as mean values + SEM of at least 4 independent experiments related to the effect of the positive control (1 nM E2), which was set to 100%. Significant differences to the lowest applied concentration of the respective compound were calculated by one-way ANOVA, followed by Fisher's LSD post-hoc testing, and are indicated with "*" ($P < 0.05$), "**" ($P < 0.01$) or "***" ($P < 0.001$).

B. AlP results for 250 µM atromentin, incubated with and without 1 µM ICI 182,780. Results are expressed as means + SEM of at least 4 independent experiments. The significant difference between those results, which proves its ER-dependency of the observed inductive effect, was calculated by Student's t-test ($P < 0.001$).

Supporting materials

Hybrid *in silico/in vitro* target fishing to assign function to "orphan"

compounds of food origin – the case of the fungal metabolite atromentin

Luca Dellafiora[1], Georg Aichinger[2], Elena Geib[3], Leyre Sánchez-Barrionuevo[4], Matthias Brock[3], David Cánovas[4], Chiara Dall'Asta[1], Doris Marko[2],

[1] Department of Food and Drug, University of Parma, Via G.P. Usberti 27/A, 43124 Parma, Italy

[2] Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, Waehringer Str. 38, 1090 Vienna, Austria

[3] Fungal Genetics and Biology Groups, School of Life Sciences, University of Nottingham, University Park, NG7 2RD Nottingham, UK

[4] Department of Genetics, Faculty of Biology, University of Sevilla, 41012, Spain

Corresponding author: Chiara Dall'Asta

       Department of Food and Drug, University of Parma, Via G.P. Usberti 27/A, 43124 Parma,

       Italy

**Table 1S.** Results of the ligand-based virtual screening

Distance score	Ligand ID	Ligand name	Related structure ID	
0.12	1G4	(3S,6S)-3,6-bis(4-hydroxybenzyl)piperazin-2-one	4IPS	Cytochrome P450 from *Mycobacterium tuberculosis*
0.13	69T	4-(5-((4-hydroxyphenyl)amino)-1H-pyrazol-3-yl)phenol	5IBJ	Cytochrome P450 from *M. tuberculosis*
0.14	KRA	1-(4-hydroxyphenyl)-4-phenyl-butane-2,3-dione	Record not available	---
0.14	1W3	7-hydroxy-3-[(E)-2-(4-hydroxy-3,5-dimethylphenyl)ethenyl]-4-methyl-2H-chromen-2-one	4KY2	Transthyretin complex from *Homo sapiens*
0.15	1G7	(3S,6S)-3-(3,4-dihydroxybenzyl)-6-(4-hydroxybenzyl)piperazine-2,5-dione	4IPW	Cytochrome P450 from *M. tuberculosis*
0.16	GEN	Genistein	5KDA	Aromatic prenyltransferase from *Aspergillus terreus*
			5AUZ	Death-associated protein kinase 1 from *H. sapiens*
			5AV4	Death-associated protein kinase 1 from *H. sapiens*
			<u>4FJ1</u>	<u>17-β-HSD from *Curvularia lunata*</u>
			3KGT	Transthyretin complex from *H. sapiens*
			3KGU	Transthyretin complex from *H. sapiens*
			<u>2QA8</u>	<u>ERα from *H. sapiens*</u>
			<u>1X7J</u>	<u>ERβ from *H. sapiens*</u>
			<u>1X7R</u>	<u>ERα from *H. sapiens*</u>
			<u>1QKM</u>	<u>ERβ from *H. sapiens*</u>
0.16	1CQ	4,4'-{3-[(4-hydroxyphenyl)amino]-1H-pyrazole-4,5-diyl}diphenol	4KTL	Cytochrome P450 from *M. tuberculosis*
0.16	MAX	Matairesinol	2BGM	Secoisolariciresinol Dehydrogenase from *Podophyllum peltatum*
0.16	RP4	(1R,4R,5R)-1,4,5-trihydroxy-3-(3-phenylsulfanylphenyl)cyclohex-2-ene-1-carboxylic acid	2CJF	Type II dehydroquinase from *Streptomyces coelicolor*
0.16	64O	N-(4-{[4'-(carbamoylamino)-6-hydroxybiphenyl-3-yl]oxy}-3,5-dichlorophenyl)propanamide	5HK9	CFTR inhibitory factor from *Pseudomonas aeruginosa*

Note → The ten top scored ligands are shown. The records related to either relevant macromolecules and pathway of sound pharmacological interest are underlined

0

1 **Purified atromentin from heterologous *invA5* expression in *Aspergillus oryzae* OP12.**

2   An HPLC analysis was performed to test the purity of atromentin purified from the culture

3 broth of *A. oryzae* OP12. The methanolic fraction was loaded on an Eclipse XDB-C18 column and

4 the absorption spectrum was recorded at 254 nm. As shown in Figure 1S, atromentin elutes at 9.6

5 min and shows its typical UV/Vis spectrum. Only very minor impurities were observed and based

6 on this analysis the fraction was judged as $\geq 99\%$ pure.
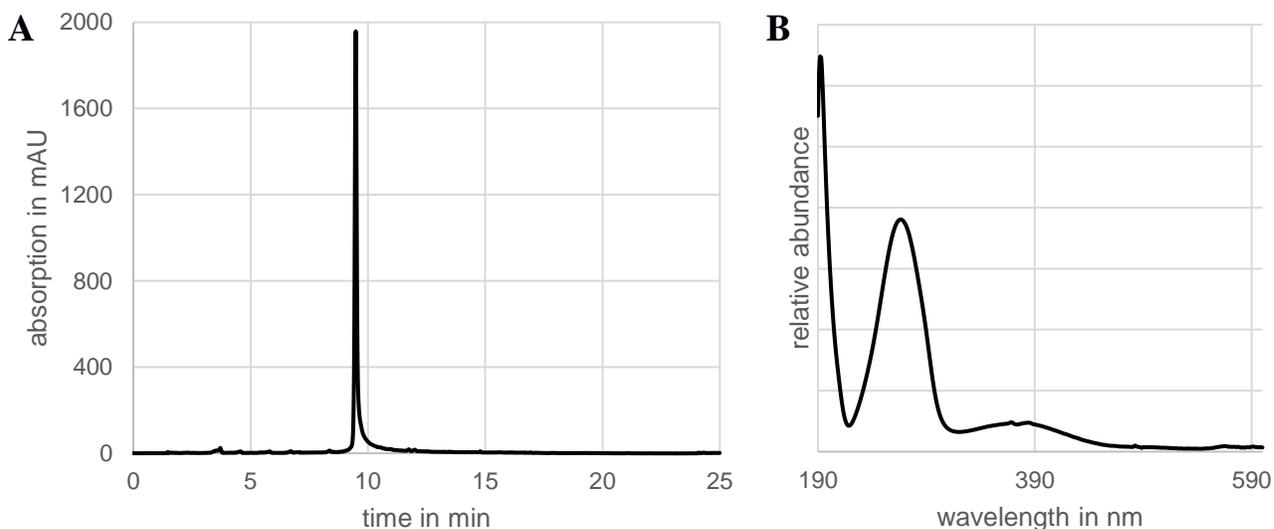
7

8 **Figure 1S**: HPLC analysis of purified atromentin by reverse phase chromatography over an

9 Eclipse XDB-C18 column. (A) Chromatogram recorded at 254 nm. (B) UV/Vis spectrum of

10 atromentin.

11

12 **Comparison between computational geometries of binding and crystallographic poses of**

13 **compounds**

14   To check the procedural reliability in predicting the geometries of binding, the computed

15 poses were compared to those obtained in crystallographic studies. As shown in Figure 2S and 3S,

16 all calculated geometries of binding were found to be consistent with the crystallographic poses,

17 confirming the procedural reliability.
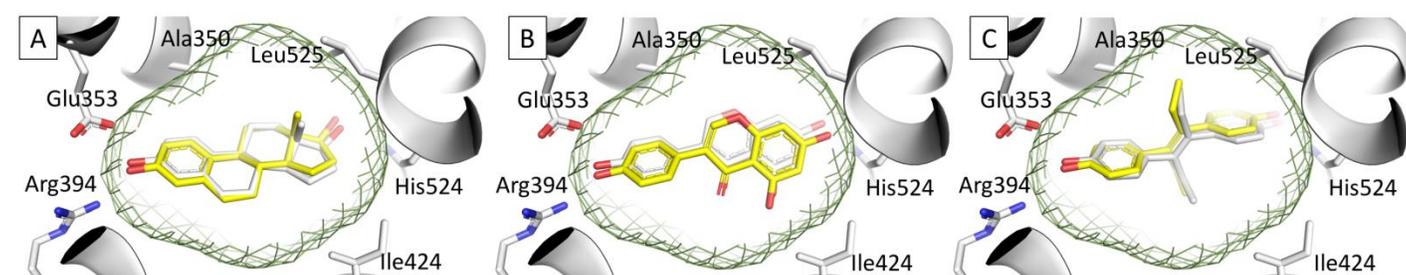
18

19 **Figure 2S.** Comparison between the computational geometries of binding (yellow) and the

20 crystallographic poses (white) of 17β-estradiol (A), genistein (B) and diethylstilbestrol (C) into

21 the ER ligand binding pocket. All ligands and side chain residues are represented in stick while

22 the protein is shown in cartoon. The green mesh retraces the shape of the binding site. The

23 crystallographic pose of ligands derived from the crystallographic ERα structure with PDB code

24 2YJA. The PDB codes used for comparison were 2YJA [24], 1X7R [51] and 3ERD [52],

25 respectively.

26
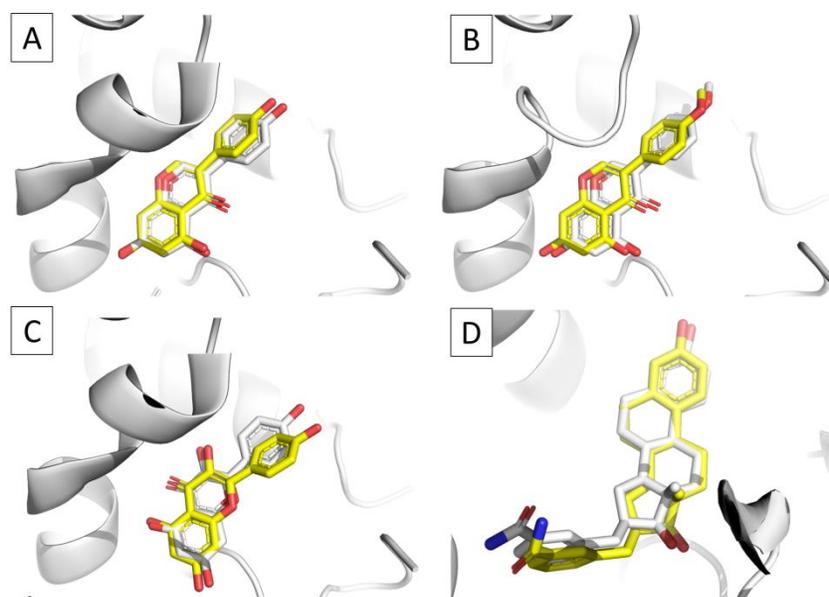
Figure 3S. Comparison between the computational geometries of binding (yellow) and the crystallographic poses (white) of genistein (A), biochanin A (B) and kaempferol (C) into the pocket of 17-β-HSD enzymes from *Curvularia lunata* (A, B and C) or *Homo sapiens* (D). All ligands are shown in sticks while proteins are represented in cartoon. The crystallographic pose of ligands derived from the crystallographic 17-β-HSD structures with PDB codes 4FJ1, 4FJ2, 3QWH [26] and 3HB5 [27], respectively

Reference not reported in the main text

51. Manas ES, Xu ZB, Unwalla RJ, Somers WS. Understanding the selectivity of genistein for human estrogen receptor-beta using X-ray crystallography and computational methods. *Structure*. 2004 Dec;12(12):2197-207

52. Shiau AK, Barstad D, Loria PM, Cheng L, Kushner PJ, Agard DA, Greene GL. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell*. 1998 Dec 23;95(7):927-37