UNIVERSITÀ DEGLI STUDI DI NAPOLI "FEDERICO II"

Italian National PhD in Artificial Intelligence

Agrifood and Environment

XXXVIII Cycle

COMPUTER-DRIVEN ANALYSIS OF TASTE AND AROMA COMPOUNDS, BIOACTIVES AND TOXICANTS IN FOOD: A POSSIBLE BLUEPRINT TO PIECE THE ANALYSIS OF RELEVANT CHEMICALS TOGETHER

Supervisor:
Prof. Gianni Galaverna

Co-Supervisor:
Prof. Luca Dellafiora

Course Coordinator:
Prof. Francesco Loreto

Candidate:
Florinda Perugino

Academic Year 2024/2025

Table of Contents

Abstract **1**

Chapter I **4**

General Introduction 4

Chapter II **44**

The bitter side of toxicity: a big data analysis spotted the interaction between trichothecenes and bitter receptors 44

Chapter III **71**

A mechanistic toxicology study to grasp the mechanics of zearalenone estrogenicity: Spotlighting aromatase and the effects of its genetic variability 71

Chapter IV **102**

Virtual display of targets: A new level to rise the current understanding of ochratoxin A toxicity from a molecular standpoint 102

Chapter V **118**

Delving into the possible role of Ochratoxin A in Parkinson's disease: a computational study spotlighting the potential mechanisms of action 118

Chapter VI **146**

Grasping the duality of fumonisins versus ceramide synthase: a computational approach to study their role as inhibitors and substrates from a mechanistic standpoint 146

Chapter VII **182**

Free fatty acid receptors beyond fatty acids: A computational journey to explore peptides as possible binders of GPR120 182

Chapter VIII **209**

A Machine Learning based prefiltering approach to screen GPR120 putative binders 209

Chapter IX **233**

General discussion and perspectives 233

Supplementary material **238**

Acknowledgements **270**

About the author 272

Abstract

Abstract

In the last decades, the world has experienced a rapid development of technology resulting in the integration of computational strategies and Artificial Intelligence (AI) also in food chemistry, biology, food science, and related disciplines. Understanding the interaction between macromolecules (i.e. proteins) and small molecules from a mechanistic point of view represents a critical point to elucidate the biological processes underpinning food safety, food perception, and food technology. The integration of AI-based techniques in computational chemistry has already provided advances in pharmacology. Starting from the assumptions that: i) many pharmacologically relevant targets are also involved in the interaction with molecules of food origin; ii) the biological effects of a specific compound depend on its interaction with one or more binding sites of the target protein, the same computational techniques that are widely applied in pharmaceutical research could be successfully applied also to food science. Among all the compounds relevant to food science and food safety, this work investigates three macro categories: taste-active compounds, bioactive compounds, and toxic compounds (with a focus on mycotoxins). This research project aims to represent a proof-of-principle to refine the current understanding of food-borne molecules leveraging on advanced computational methods including artificial intelligence. For this reason, studies on taste compounds have been conducted focusing not only on their capability to elicit taste and aroma, but also on their potential bioactivity and on their possible "side effects". In this context, the interaction between food-related toxic compounds and taste receptors type 2 (TASRs) was investigated performing a Machine Learning (ML) based virtual screening, providing insights into the potential involvement of extraoral TASRs (especially those expressed in the gut) in the mechanism of toxicity of mycotoxins, such as trichothecenes (i.e. DON and 15-Ac-DON). However, mycotoxins may exert several other adverse effects on human and animal health. Investigating and elucidating effects still largely uncharacterized from a mechanistic point of view may represent a valuable strategy to both strength risk assessment and guarantee food safety. In this perspective, the interaction and possible resulting effects of others concerning mycotoxins within relevant targets has been investigated by applying computational pipelines mostly made up of virtual screening, molecular docking and molecular dynamics simulation. The impact of different single nucleotide variants (SNVs) on the Aromatase enzyme (CYP19A1) was studied revealing that some variants are even more prone to inhibition by zearalenone and α-zearalenol while others are potentially inactive compared to the wild type. Another mycotoxin worthy of major concern is Ochratoxin A (OTA). Its intricate toxicological profile involves multiple

mechanisms including oxidative stress, impaired cellular defense and synthesis, and the induction of carcinogenic, nephrotoxic, and neurotoxic effects. Concerning the effects of OTA on protein synthesis, OGFOD1 - a key player in protein translation - was identified as a potential target for both OTA and its thermal degradation product 2'R-OTA. Also, its potential role in Parkinson's disease has been elucidate from a molecular standpoint proving the ability of OTA and certain congeners to interact with some enzymes involved in the Parkinson's Adverse Outcome Pathway (i.e. Cathepsin L/D and Complex I). Furthermore, the advancement achieved by AI in protein structure prediction enabled the homology modelling of human Ceramide Synthase 5 (CerS5), a known target of fumonisins. This led to the in-depth investigation of the duality of fumonisins toward CerSs revealing that HFB1 can undergo the reaction through both catalytic mechanisms proposed (i.e. one-step and two-step catalytic mechanism), while FB1 preferentially undergoes the reaction through the two-steps mechanism. Concerning bioactive compounds, this work investigated for the first time to the best of our knowledge the capability of L-peptides and D-peptides to interact with GPR120 (Free Fatty Acid Receptor 4). Starting from this evidence, a CatBoost model to apply as a prefilter for docking analysis was developed with the aim of screening and identifying new GPR120 targets from huge chemical compounds libraries. In conclusion, the main goal achieved in this PhD thesis are i) the identification of a connection between taste receptors and toxic and bioactive compounds; ii) the in depth investigation of the mechanic underpinning the toxicity of different mycotoxins; iii) the progressive implementation of AI-based techniques which proved to be a powerful strategy to empower and accelerate the investigation also reducing time and costs.

Chapter I

General Introduction

Chapter I

1.The role of Artificial Intelligence in Food Science

The rapid development of technology has paved the way to the widespread application of Artificial Intelligence (AI) techniques in food science. At a global level, the food industry is increasingly requiring smart strategies to ensure food safety, enhance sustainability, and to fulfill the increasing demand for food driven by the global world population growth (Chhetri, 2023). In this context, AI represents a powerful tool for improving food safety and quality control along the whole productive chain. Moreover, in alignment with the United Nations Agenda 2030, AI has been recognized as a key strategy to achieving the 17 Sustainable Development Goals (SDGs) (Palomares et al., 2021). There are many areas of food science where AI may represent a valuable strategy to improve performance. Starting from the harvesting of the raw material with the development of smart farming moving to the final process such as smart packaging development and final product quality insurance. Additionally, also consumer satisfaction may be monitored with the help of AI performing data-analysis, supporting the development of personalized nutrition and tailored food products. However, the focus of this thesis lies in the identification and the investigation of taste and aroma compounds, bioactive compounds, and toxic compounds (mainly mycotoxins) in food applying *in silico* strategies. Compared to *in vitro* and *in vivo* tests, *in silico* studies agree with the principles of the 3Rs: Replacement, Reduction and Refinement. In fact, they are cheaper, consume less time and do not require animals or other living organism. This computational approach is well framed in the context of New Approach Methodologies (NAMs) which encompasses any technology, methodology, approach, or combination that can provide information on chemical hazard and risk assessment, avoiding the use of animals and including *in silico*, *in vitro* and *ex vivo* approaches (Westmoreland et al., 2022). More in detail, *in silico* experiments are performed *via* computer simulations. They represent an efficient and reliable strategy to obtain information about the chemical and the physical properties of a chemical compound and to investigate its biological activity. From a food science perspective, *in silico* tools represent a valuable strategy to:

- Investigating the mechanism of interaction between macromolecules (e.g. proteins) and small ligands.
- Predicting the role of water in biological interactions.

- Studying the binding affinity of bioactives and toxic compounds toward their biological targets and investigating their activity, stability, and ability to persist within the organism.

AI-based techniques such as Machine Learning (ML) and Deep Learning (DL) have already demonstrated remarkable success in drug discovery processes, where they have successfully and significantly reduced time and costs by accelerating the identification, optimization and validation of active molecules (Das et al., 2024). Considering that the main approach for drug discovery is the investigation of the interaction between a pharmacological or a putative pharmacological target (e.g. protein) and small molecules capable of binding it the same approach could be extended to molecules of food interest. Such an approach may provide an innovative framework for exploring and characterizing foodborne and food-related molecules, providing new insights into their nutritional roles, toxic or bioactive effects, and technological aspects. Within this context, virtual screening has emerged as one of the most important computational methodologies in drug discovery. There are two main approaches for virtual screening: ligand-based virtual screening and structure-based virtual screening. The former relies on the investigation of ligands associated with a specific biological activity (da Rocha et al., 2025) to find similar compounds, the latter on molecular docking to identify potential novel ligands for a specific target. The integration of AI in virtual screening workflows has greatly enhanced predictive performance, allowing the development of complex models based on non-linear relationships between molecular descriptors and biological activities. These upgrades not only increased the accuracy of virtual screening outputs but also expanded the applicability of virtual screening beyond traditional medicinal chemistry and pharmacology, paving the way for computational prediction about foodborne compounds bioactivity, toxic compounds and compounds responsible for taste and aroma.

1.1. Role of artificial intelligence in protein structure prediction

Proteins are macromolecules composed of amino acids which play a fundamental role in all biological processes of life, including catalysis of metabolic reactions, maintenance of pH, osmotic balance and cell structures, as well as transport of molecules and DNA replication. During their synthesis (namely, translation by ribosomes), the linear amino acid sequence eventually folds into complex 3D conformations, also known as tertiary structures, which are crucial for protein activity and stability. Knowledge of protein structure plays a pivotal role in disease prediction, drug discovery and related fields (Mufassirin et al., 2022), including food

science. For this reason, determining their structure represents a critical issue in structural bioinformatics. Traditionally, protein structures have been resolved using experimental techniques such as X-crystallography, Cryo-electron microscopy (Cryo-EM), and nuclear magnetic resonance (NMR) (Zhan et al., 2025). All these methods provide high-resolution structural data, but they have several limitations. In fact, X-ray crystallography requires the formation of high-quality protein crystals, a process that is challenging or impossible in case of membrane or intrinsically disordered proteins. NMR spectroscopy is capable of studying proteins in solutions, but its efficiency is limited to relatively small proteins. Cryo-EM is a recent and powerful alternative also for large proteins, but it is expensive, time consuming and computationally demanding. All the experimentally resolved structures are collected in the Protein Data Bank (PDB; https://www.rcsb.org/), which currently counts over 242,000 experimentally resolved entries (last dataset access October 2025). Despite this large number of determined structures, it represents only a small fraction of the whole proteome. Many proteins, in fact, lack experimental structures, while others are only partially resolved or have regions of poor resolution. For all these reasons, in the last years, time-saving and more affordable protein structure prediction methods have risen in popularity to fill the gap. Computational approaches for protein structure prediction are mainly divided into two categories: Template-Based Modelling (TBM) and Template-Free Modelling (TFM) (Dhingra et al., 2020). The former method predicts a structure from a given sequence using known and available structures of selected template proteins, leveraging on the conservation of structure and 3d organization among phylogenetically related protein. The latter relies on Anfinson's hypothesis that establishes that the native structure of a protein is determined only by the protein's amino acid sequence (Anfinsen, 1973). For the TBM, the template proteins are chosen via homology modelling or comparative approaches (Mufassirin et al., 2022). However, it was thanks to the development of AI that prediction accuracy has increased. TFM approaches have shown great potential with methods such as AlphaFold (Jumper et al., 2021), RoseTTAFold (Baek et al., 2021), trRosetta (Yang et al., 2020), and RaptorX-3DModeling (Peng & Xu, 2011). The successful application of these methods like neural network and support vector machines have been successfully resulted in the award in 2024 Nobel Prize in Chemistry to David Baker, Demis Hassabis, and John M. Jumper for their contributions in the field of using AI methods for the development of artificial intelligence-guided protein structure prediction and for the development of computational protein design (Zhan et al., 2025). Their research led to the generation of a library of nearly 200 million predicted structures (Aguero-Pizzolo et al., 2025). The widest recognized tool is AlphaFold, known for its high accuracy and versatility in

predicting the folding of human proteins (Ohno et al., 2024). From a methodological standpoint, AlphaFold is an AI-based system developed by Google DeepMind that predicts protein 3D structures from its amino acids sequences (https://alphafold.ebi.ac.uk/) utilizing artificial neural network (ANNs) even in absence of known similar structures (Jumper et al., 2021). Neural networks are ML algorithms, consisting of alternating linear and nonlinear components, called layers, that are typically trained using gradient descent on the error of the final predictions (Bonetta & Valentino, 2020). The success of AI-based protein structure prediction was particularly evident in medicine, structural biology, and drug discovery. However, it may be extended to food science and technology. In fact, understanding the structural properties of food proteins, such as enzymes, allergens, bioactive and toxic compounds, could provide useful information for functional food design, protein engineering, digestibility studies, and risk assessment.

1.2 Role of artificial intelligence in structure-based virtual screening

Virtual screening is a promising *in silico* technique used in drug discovery (Maia et al., 2020), and its relevance has risen in recent years with the advent of ultra-large libraries of compounds (Zhou et al., 2024). For structure-based virtual screening, both the 3D structures of target protein and ligands are required. Structure based virtual screening, also known as target-based virtual screening, aims to predict the interaction of small molecules with the molecular target and identify binders within a compounds libraries containing a larger proportion of non-binders (Li et al., 2020) (Figure 1). Molecular docking represents nowadays the best way to achieve this output. In fact, this *in silico* techniques enables the prediction of the binding architecture and according to the software employed it provides a score according to the likelihood of protein-ligand interaction. First, ML may be used to implement docking score functions. The advantage of ML scoring functions compared to classical scoring functions is that they can be trained on active and inactive compounds increasing the accuracy of predictions (Wojcikowski et al., 2017). However, docking remains computationally demanding and wasteful if we consider that only a very small subset of top-scoring compounds is typically considered for experimental evaluation (Gentile et al., 2022). Considering the rapid and exponential growth of compounds libraries, ML guided approach is achieving a key role in structure-based virtual screening. Indeed, the most relevant computational resource required by structure-based virtual screening is the ability to screen ultra-large libraries before performing docking analysis. Luttens et al. succeeded in correlating ML and molecular docking by developing a CatBoost

classifier trained on docking score and Morgan fingerprint able to screen billions of reducing the computational cost by more than 1,000-fold (Luttens et al., 2025). Not only ML, but also DL was successfully employed in structure-based virtual screening, especially to enhance the obtainment of quantitative structure-activity relationship (QSAR) models (Gawehn et al., 2016; Gonczarek et al., 2018).

1.3 Role of artificial intelligence in ligand-based virtual screening

As stated before, ligand-based virtual screening represents one of the most common computational strategies for identifying compounds with a biological, pharmaceutical or technological relevance (Figure 1). Unlike structure-based virtual screening methods, which require a broad knowledge of the 3D structure of the molecular targets, ligand-based approach relies on the principle that compounds with similar structures or physicochemical properties are likely to exhibit similar biological activities and to interact with the same target (Dellafiora et al., 2020). There are several ligand-based approaches that can be used for virtual screening (Chen et al., 2007). Similarity searches according to fingerprint similarity, molecular shape, or pharmacophore models are the most common strategies to perform ligand-based virtual screening (Amendola & Cosconati, 2021). Molecular fingerprints encode structural features as bits in a bit string or as number in a count vector, indicating the presence or absence of specific structural features (Riniker & Landrum, 2013). Pharmacophore modeling, on the other hand, describes the spatial arrangement of functional groups responsible for a molecule's biological activity, providing a more conceptual and mechanistic basis for screening (Pal et al., 2019). Classical ligand-based approaches compare the reference structure with every structure available in a selected database (Hert et al., 2006). The integration of AI, especially ML techniques, enabled the analysis of large chemical datasets making it possible to compute a wide variety of descriptors/features and molecular fingerprints for the ligands. Molecular descriptors like molecular weight, pH values, hydrophobicity, etc. can be used as input data to train the model. Once trained, the model can be applied to predict new molecules. Commonly applied ML models are support vector machines (Jayaraj & Jain, 2019), random forests (Cano et al., 2017), k-nearest neighbors classifier (Mostafa et al., 2022). These traditional ML methods are well-suited for pattern recognition and classification tasks, allowing the discrimination between active and inactive compounds toward very broad arrays of bioactivity. However, in recent years, DL algorithms have been applied to enhance ligand-based virtual screening. Unlike traditional ML models that rely on predefined descriptors, DL algorithms can

automatically extract and learn complex feature representations directly from raw molecular data. This has led to the emergence of Graph Neural Networks (GNNs), Convolutional Neural Networks (CNNs), and Recurrent Neural Networks (RNNs) as powerful tools for molecular property prediction (Alves et al., 2021). For example, GNNs have been proposed for the prediction of bitter peptides (Malavolta et al., 2022; Srivastava et al., 2024).

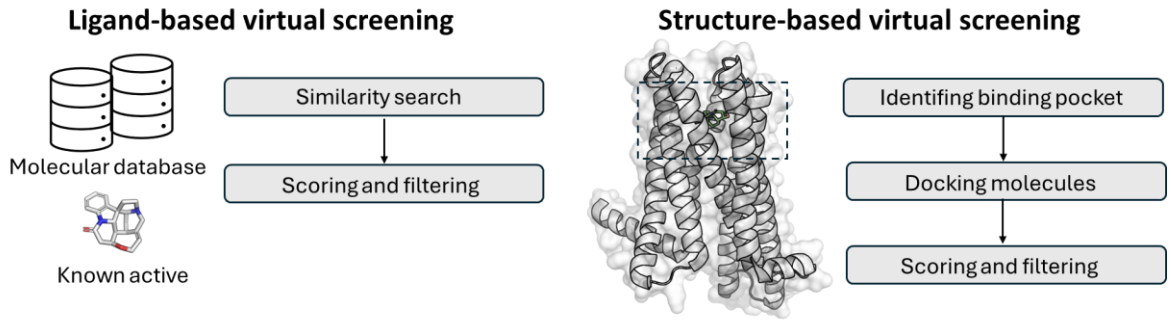


Figure 1. Ligand-based and Structure-based virtual screening general workflow.

1.4 AI-Driven Analysis of Food Chemical Databases

Databases containing structural, chemical and biological information about macromolecules (i.e. proteins) and small molecules are fundamental for *in silico* studies in food chemistry. These databases, in fact, provide access to detailed data on the 3D structure, physicochemical descriptors, and molecular interactions of compounds, thereby enabling virtual screening, molecular modeling, and bioinformatics-based prediction of bioactivity and/or toxicity of food components and contaminants. Typically, such databases (e.g., PDB, PubChem, ChEMBL) are comprehensive collections including molecules from various biological or synthetic origins, broad in scope and therefore not easy to browse specifically for food-related molecules. In recent years, the increasing demand for food-specific molecular data has led to the creation of specialized food databases. Food-related data can generally be categorized into food compositions, flavors, and chemical compounds (Tseng et al., 2023):

i) Food composition databases, containing information on ingredients, macronutrients, micronutrients, and labeling data (e.g. FoodData Central (https://fdc.nal.usda.gov/), InFoods (https://www.fao.org/infoods/infoods/en/), EuroFIR (https://www.eurofir.org/)).

ii) Flavor-related databases, which provide structural, sensory, and molecular data on natural and synthetic flavor compounds, including their receptor targets and perceived taste modalities (e.g. FlavorDB (https://fsbi-db.de/), BitterDB (Ziaikin et al., 2025), VirtualTaste (https://insilico-cyp.charite.de/VirtualTaste/))

iii) Food chemical compound databases which contain chemical entities identified in foods, including phytochemicals, metabolites, and contaminants (e.g. FooDB (www.foodb.ca))

These databases provide a huge volume of food-related data that may be unlocked and widely exploited through AI techniques. AI tools, in fact, can process, integrate, and interpret large and heterogeneous data. In this context, AI serves not only as a tool for data mining but also for pattern recognition, predictive modeling, and knowledge discovery across diverse food data sources. Many studies have successfully employed information derived from various food databases, demonstrating that with abundant and diverse data inputs, ML and DL methods can be effectively applied to extract meaningful insights in food science (Tseng et al., 2023).

2. Taste and aroma compounds

When food is introduced into the oral cavity, it may release a huge number of volatile compounds that may elicit olfactory perceptions upon the interaction with odorant receptors. Moreover, during chewing, food may also release a complex mixture of non-volatile molecules that may be responsible for taste perception upon binding of dedicated taste receptors. Taste and aroma compounds are two of the main food sensory characteristics affecting consumer choice, and diet habits with obvious consequences to consumer health and wellbeing. On the other hand, considering the broad expression of taste and aroma receptors in different tissues, taste and aroma compounds are also likely to have biological activities that might be relevant from a physiological point of view and possibly connecting consumers' behavior, diet habits, and food-related consequences on health. Therefore, the analysis of aroma and taste molecules is fundamental to designing consumer-pleasing and health-protecting foods. There are five basic tastes, i.e., sweet, bitter, salty, sour, and umami. From a molecular standpoint, proteins involved in perception of sweet, bitter, and umami belong to the G protein-coupled receptors (GPCRs) family, while the action of sodium, potassium, or hydrogen ions on ion channels or ion exchangers in the taste buds is responsible for salty and sour taste (Cygankiewicz et al., 2014). GPCRs represent the largest class of cell surface proteins and besides their role in taste and aroma perception, they are involved in a plethora of biological processes such as

neurotransmission, chemoattraction, regulation of appetite, blood pressure, digestion, and reproduction (Cygankiewicz et al., 2014). More in detail, T1R2/T1R3 heterodimer acts as a sweet taste receptor by binding a wide variety of sugars and sugar substitutes (Nelson et al., 2001), whereas T1R1/T1R3 heterodimer acts as a receptor for umami taste (Damak et al., 2003). Since the first structure of sweet receptor was released in 2001 (Nelson et al., 2001), the gustatory mechanism was investigated from a mechanistic point of view via *in silico* tools such as molecular docking to facilitate the discovery of new sweet compounds (Zhu et al., 2024). Recently, crystallographic structures of the human sweet receptor bound to sucralose and aspartame, which are two of the most widely used artificial sweeteners, have been released opening to the possibility of computationally design new taste modulators (Juen et al., 2025). Concerning bitterness, in humans 25 bitter taste receptors (TAS2Rs) have been identified so far. Bitter compounds trigger the aversive response across species and are typically thought to serve as a protective mechanism against the ingestion of toxic and poisoning substances (Di Pizio et al., 2019). However, not all bitter compounds are toxic and not all toxic compounds are bitter. Just thinking about phytocompounds such as polyphenols which are known to be bitter, but also to exert a series of positive effects on human health (e.g. protection of neurons against oxidative stress, improvement of insulin secretion, etc) (Del Rio et al., 2010). Of note, sweet and bitter receptors have been found expressed also in the gut epithelium. Although their role is far from being completely understood, they could be involved in different aspects related to appetite and satiety. The interest in this direction led to the release of the crystallographic structures of two TASRs in the last three years, namely TAS2R46 (Xu et al., 2022) and TAS2R14 (Tao et al., 2024). Thanks to this, *in silico* approaches integrating AI have been applied to the discovery of novel potential bitter compounds. Nowadays, several ML methods have been proposed to classify a compound as bitter or non-bitter, e.g. BitterIntense (Margulis et al., 2021), BitterCNN (Bo et al., 2022), and VirtuousSweetBitter (Maroni et al., 2022). ML models like BitterX (Huang et al., 2016), BitterSweet (Tuwani et al., 2019), and BitterMatch (Margulis et al., 2021) also succeeded in predicting the specific TASRs target for bitter compounds. Moreover, Srivastava at al. developed a new predictor of bitter taste based on Graph Neural Networks which succeeded in predicting bitter peptides, providing both an efficient strategy for the identification of bitter peptides in various food products but also gives a hint into the molecular basis of bitterness (Srivastava et al., 2024). In addition to the "conventional" tastes mentioned above, GPCRs have been proposed as the molecular targets for the debated "fatty taste" perception. Specifically, GPR120, also known as free fatty acid receptor 4 (FFAR4), has been detected both in the human lingual gustatory epithelium and in non-gustatory epithelia and its

contribution in human gustatory fatty acid perception has been hypothesized (Galindo et al., 2012). Another sensation related to taste is kokumi. Kokumi is not a taste, but it is a sensation acting as a taste enhancer with positive effects on mouthfulness, as well as on taste thickness and continuity (Kitajima et al., 2025). It is interesting to note that also kokumi perception has been associated with the activation of a GPCR, specifically the Calcium Sensing Receptor (CaSR). More in detail, the stabilization of the close confirmation of the CaSR's Venus Flytrap Domain (VFT) is the molecular mechanism underpinning kokumi perception (He et al., 2024). Such a sensation is important to determine taste, flavor, and aftertaste qualities of a wide variety of foods with critical implications for food acceptance and consequences on consumers' behavior and diet habits. In this context, the investigation for new kokumi compounds grew during the last years. Traditional methods for generating or identifying novel kokumi compounds require multi-step experimental procedures and are usually time-consuming (Kim et al., 2022; Vasilaki et al., 2022). Therefore, computational approaches have been applied to accelerate and empower this process. Dellafiora et al. developed an *in silico* pipeline made up of ligand based virtual screening, molecular docking and molecular dynamics simulations which succeeded in investigating the capability of gamma glutamyl peptides to act as kokumi active substances activating CaSR. In this framework, the opening of the VFT domain was measured over the time of the molecular dynamics simulations describing γ-Glu-Pro-Ser and γ-Glu-Pro-Ala as potential potent kokumi active compounds (Dellafiora et al., 2022). The same approach was then extended to peptides identified in sparkling wine. The *in silico* investigation performed on a dataset of 50 oligopeptides experimentally detected in wine described 8 dipeptides and 3 tripeptides as potential kokumi active compounds. This combined in vitro-in silico approach was integrated with sensory tests performed on one of the putative kokumi active sequences identified (i.e. Gly-Val) confirming its ability to significantly modify the perception of complex real wine (Perenzoni et al., 2024). Concerning umami, no crystallographic structure of T1R1/T1R3 is available (Wang et al., 2022). However, *in silico* tools like homology modeling, molecular docking and molecular dynamics simulation were applied to construct a computational modeling of T1R1-T1R3 heterodimers (Dang et al., 2019; Lopez Cascales et al., 2010).

3. Bioactive compounds in food

Bioactive compounds are chemically diverse and include several phytochemicals commonly found in food that have biological activity on living organisms. They exert many effects such as antioxidant activity, inhibition or induction of enzymes, inhibition of receptor activities, and induction and inhibition of gene expression with positive results on diseases like obesity, diabetes and hypertension (Cláudia da Costa Rocha et al., 2021). Most bioactive compounds are secondary metabolites of certain organisms like vegetables, fruits, nuts, oils, grains and algae. However, they have been studied also from animal-derived foods like diary product, meat, fish and eggs (Lafarga & Hayes, 2014). Some examples of bioactive compounds are (poly)phenols, carnitine, choline, coenzyme Q, phytosterols, phytoestrogens, glucosinolates, and taurine (Bortolomedi et al., 2022). (Poly)phenols represent the most studied class among phytochemicals with more than 8,000 phenolic structures identified so far for a total of ten classes (Kris-Etherton et al., 2002). Flavonoids, especially flavones, flavanols, and their glycosides, are the most common polyphenols present in the human diet, due to their high presence in food plants. The concentration and activity of these compounds are influenced by several factors, including genetic variability of raw materials, environmental conditions, agricultural practices, processing methods, and storage conditions. Regarding the biological effects, the bioavailability of phytochemical and, generally speaking, of bioactive compounds is critical (i.e. the ability of the compounds to be absorbed by human body). In fact, bioavailability, referring to the possibility of absorbing and having available a given compound, is a major factor determining the overall effect of bioactive compounds in food (Patil et al., 2009). In recent years, the use of computational methods and AI-based tools has improved also the study of bioactive compounds of food origin. In silico approaches such as the above-mentioned molecular docking, QSAR modeling, and virtual screening allow the prediction of biological targets and mechanisms of action, supporting the discovery of novel bioactives from natural sources, including food. Furthermore, their investigation from a molecular standpoint already provided insights about their interaction with new molecular targets (Borgonovi et al., 2025; Dellafiora et al., 2013). Furthermore, bioactive compounds have gained significant attention in the development of functional foods and nutraceuticals, which aim to provide health benefits beyond basic nutrition. During the last decade, (poly)phenols have been studied as potential therapeutic agents against a wide range of diseases, including neurodegenerative diseases, diabetes, cardiovascular dysfunctions, inflammatory diseases, and also ageing (Babbar et al., 2015). For this reason, they may be well framed in nutraceutical field, which

combines nutrition and pharmacology highlighting how food may promote human health, helping to prevent and/or treatment of specific disfunctions. Nutraceuticals are developed in various forms like capsules, tablets, powders, or enriched food products. They are increasingly used as part of functional foods, a category designed to combine nutritional adequacy with specific health-promoting properties. The distinction between functional foods and nutraceuticals mainly lies in their form and regulatory definition: while functional foods are consumed as part of a regular diet, nutraceuticals are often formulated as preparations with standardized doses of active ingredients (Vettorazzi et al., 2020).

3.1 Regulatory aspects

Bioactive compounds are generally defined as biologically reacting substances that stimulate a response in living tissue (Biesalski et al., 2009). However, the reported definition lacks precision. In fact, it does not specify if the effect is beneficial or deleterious, whatever the effect would be beneficial, it does not specify if the type of effect is pharmacological or nutritional, and which is the final form or vehicle in which the bioactive is present in food (Vettorazzi et al., 2020). Despite the growing interest in bioactive compounds, there is still no harmonized definition within European law. resulting in overlaps among different legal frameworks. In fact, according to European law, bioactive compounds intended for food consumption could fall under regulations concerning food supplements, substances for the fortification of food, food for specific groups, and novel foods. Furthermore, depending on the type of bioactive compound, they can be classified as vitamins and minerals, other substances, or botanicals (Vettorazzi et al., 2020). Usually, together with nutraceutical and functional food, bioactive compounds are generally included in the food supplements category defined in the Directive 2002/46/EC ("Directive 2002/46/EC of the European Parliament and of the Council of 10 June 2002 on the Approximation of the Laws of the Member States Relating to Food Supplements,"). According to this Directive, food supplements are “foodstuffs the purpose of which is to supplement the normal diet, and which are concentrated sources of nutrients or other substances with a nutritional or physiological effect, marketed in dose form”. Health and nutrition claims associated with food products or food supplements must comply with Regulation (EC) No 1924/2006 ("Regulation (EC) No 1924/2006 of the European Parliament and of the Council of 20 December 2006 on nutrition and health claims made on foods," 2006), which aims to ensure that all statements made on food labelling and marketing are truthful, scientifically substantiated, and not misleading. The European Food Safety Authority (EFSA) is responsible

for the evaluation and scientific validation of submitted health claims. EFSA's assessment is based on the examination of available scientifical evidence collected so far, including human intervention studies, toxicological evaluations, and bioavailability data, to establish a cause-effect relationship between the consumption of the compound/food and the claimed health benefit (Vero & Gasbarrini, 2012). Only those claims that meet these stringent requirements are included in the EU Register of Nutrition and Health Claims. As mentioned before, bioactive compounds may also fall into the novel foods category if they are "novel" to the European market. In this case, they should follow the Novel Food Regulation (EU) 2015/2283, which governs foods and ingredients that were not consumed to a significant degree within the European Union prior to 15 May 1997. Novel bioactives or extracts derived from unconventional sources such as insects and microalgae in fact require a safety assessment and authorization before they can be commercialized. This ensures that the compound does not represent a health risk to consumers and that its composition and use are clearly defined.

3.2 Bioactive peptides

Among bioactive compounds, bioactive peptides received considerable attention. They are specific polymers or oligomers of amino acids linked by peptide bonds between the carboxyl group of one amino acid and the amino group of the following amino acid. They usually are 3 to 20 residues long, characterized by low molecular weight (< 6,000 Da), and exert positive effects on human body (Ahn et al., 2012). Bioactive peptides are encrypted in the primary sequence of proteins and can be released by the action of proteolytic enzymes, using novel processing technologies such as High-Pressure Processing (HPP) and Pulsed Electric Field (PEF) or during food processing such as cooking, ripening and fermentation (Murtaza et al., 2022). Once released and absorbed, bioactive peptides can interact with specific receptors and regulate different physiological functions (Power et al., 2014). While if undigested and unabsorbed, they pass into the large intestine, where they are metabolized by intestinal microbiota (Segura-Campos et al., 2011). Among bioactive peptides, those of food origin are now gaining popularity because they are promising nutraceutical and functional food ingredients proposed to prevent, manage and occasionally co-treatalong with pharmacological interventions different dysmetabolic and chronic diseases like Diabetes Mellitus Type 2, hypertension, coronary diseases, inflammatory-diseases, and cholesterolaemia (Sánchez & Vázquez, 2017). Bioactive peptides can have a food origin but can also be extracted from proteins and added as ingredients in functional food or food supplements. Traditionally,

foodborne bioactive peptides have been extracted from a variety of animal and vegetable sources including milk, egg, fish, soybean, chickpea and peanuts. They can be commercialized and consumed in the form of food, beverage, tablets, powders, capsules and liquids (Chalamaiah et al., 2019). There are many advantages in using bioactive peptides because they are easy to synthetize, typically free of side/adverse effects and endowed with multi-pharmacology targeting more proteins at once. Furthermore, from a food safety perspective they show a high potency of action, low accumulation in tissues and low toxicity (Skjanes et al., 2021). The classic "bottom-up" approaches to identifying food-derived biopeptides are laborious, time-consuming and expensive, involving the *in vitro* digestion of the protein by different proteases and the following bioactivity assays on the hydrolysates. Usually, the huge number of sequences identified in digested samples make it unfeasible to test all of them experimentally. In this context, the use of *in silico* and bioinformatic approaches may complement experimental techniques investigating the bioactivity of the identified sequences reducing the number of those to further analyze (Agyei et al., 2018; Lammi et al., 2021). Another strategy to identify bioactive peptides is the "top-down" approach. In contrast to the "bottom-up" approach, this method is independent from the food matrices analyzing the intact sequence and not requiring the enzymatic digestion (Smith et al., 2011). In this framework, computational methods may accelerate the screening of peptides and also provide insights on the chemical and mechanistic basis of the identified sequences (Dellafiora et al., 2022). These approaches, in fact, already succeeded in mining bioactive peptides of food origin supporting the identification of active sequences over a variety of molecular targets (Lammi et al., 2021). For example, Pedroni et al. developed an *in silico* pipeline able to mine alpha amylase inhibitory peptides from microalgae and cyanobacteria. Indeed, concerning the capability to regulate glucose metabolism, targeting the pancreatic alpha amylase resulting in its inhibition may represent the key mechanism to regulate glucose levels using bioactive peptides. Alpha amylase is a digestive enzyme responsible for the breakdown of the polysaccharides into oligosaccharides, then converted into monosaccharides by glucosidase enzyme. Glucose is an absorbable monosaccharide that is released into the bloodstream causing a rising of glucose level called "postprandial hyperglycemia". People who suffer from Type 2 Diabetes Mellitus do not have enough insulin to use all the released glucose as a source of energy, so glucose in the bloodstream increases even more resulting in organs damages. Inhibition of alpha-amylase reduces the absorption of starch and carbohydrates in the gastrointestinal tract and low blood glucose level, so inhibiting pancreatic alpha-amylase could represent a valuable strategy for prevention and co-treatment of T2DM. Considering the high protein content (up to 70% of algal

biomass) of microalgae and cyanobacteria, it is likely that they may also contain alpha amylase inhibitory peptides with possible effects on the modulation of glucose metabolism (Pedroni et al., 2022). The above mentioned investigated the proteome of *Chlorella vulgaris*, *Auxenochlorella protothecoides* and *Aphanizomenon flos-aquae* highlighting the presence of alpha amylase inhibitory peptides within them. Furthermore, a mechanistic investigation of their mechanisms of action was performed and potential novel sequences were identified within these proteomes. Of course, the actual release of those sequences in real-world conditions (e.g., via *in vitro* digestive models) should be assessed before moving to *ex vivo* or *in vivo* trials, which may eventually confirm their efficacy in living organisms. With a view to the circular economy, *in silico* methods have been successfully applied to predict bioactive peptides from food waste. Specifically, Morena et al. presented a computational workflow to predict peptide release, bioactivity, and bioavailability from by-products generated from *Solanum lycopersicum* (Morena et al., 2024).

Chemical and structure modifications occurring during food processes have been widely described as a source of bioactive peptides together with the hydrolysis using controlled and commercial peptidases or microorganisms. Eventually, gastrointestinal digestion due to the action of salivary, stomachal, intestinal and pancreatic enzymes constitute the final hydrolysis step in releasing bioactive peptides from food protein (Pepe et al., 2016). Processes such as fermentation, heat and alkali treatments may result in protein racemization with the consequent racemization of L-amino acids to D-amino acids (Friedman, 2010). For this reason, D-amino acids may be used as molecular markers of both thermal and basic treatments, as well as of microbial activity (Genchi, 2017). The presence of D-amino acids within protein sequence may change protein digestibility and consequently the absorption of essential and non-essential amino acids. From a food science perspective, it is important to notice that the inclusion of a D-amino acid in peptides may enhance their resistance to proteases (Yan et al., 2020). However, the effect of racemization on protein digestibility still needs to be fully elucidated. To cover this gap, for example, Accardo et al., evaluated the efficacy of the proteolytic action performed by digestive enzymes acting on proteins embedding D-amino acids. The study revealed that protein cleavability efficiency changed depending on the enzyme and/or amino acid involved (Accardo et al., 2024). Also in this framework, *in silico* investigations have helped to elucidate peptides' behavior and molecular docking and molecular dynamics simulations proved to be able to study the underpinning mechanisms (Pedroni, Perugino, Magnaghi, et al., 2024).

4. Toxic compounds in food

Toxic compounds in food represent a significant concern to food safety and public health perspectives. They may have different origins:

- Natural sources: toxic chemical substances produced by different organisms (e.g. mycotoxins, plant toxins, and marine toxins).
- Food processing: chemical substances naturally formed in food during industrial processes or cooking (e.g. acrylamide and furans).
- Environmental pollutants: chemical substances resulting from industrial or agricultural practices or present in the environment (e.g. heavy metals, persistent organic pollutants, pesticides, veterinary drugs, xenohormones).

As stated above, EFSA is the authority responsible for evaluating and ensuring food safety. Furthermore, to guarantee and protect the health of consumers, the European Union defined a coordinated food alert system called the Rapid Alert System for Food and Feed (RASFF) (https://food.ec.europa.eu/food-safety/rasff_en) (Eissa et al., 2024). This tool allows the rapid exchange of information on food risks between member countries, allowing prompt containment measures, such as recalling and removing contaminated products from the market. Thanks to this platform, consumers, business operators, and authorities around the world can have access to information on recent food recalls and public health warnings in countries. This system is particularly relevant considering that food contamination may occur at any stage along the food supply chain, from the raw material production in field to processing, distribution and final storage. In Europe, specific regulations have been established to define maximum allowable levels of residues and contaminants resulting from the presence of toxic chemical substances in foods (Casado et al., 2024). In the regulatory framework, NAMs may play a crucial role in supporting chemical hazard identification and risk assessment without relying on animal testing. Chemical risk assessment represents a systematic process aimed at estimating from a qualitative and a quantitative point of view the potential harmful effects on human health arising from exposure to a group of chemicals. This process consists of four steps:

- Hazard identification, which is the identification of compounds responsible for adverse effects.
- Hazard characterization, which is the estimation of adverse effects in terms of dose-response that may result from exposure to the hazard.

- Exposure assessment, which is the process that measures or estimates the intensity, the frequency and the duration of exposure to a chemical agent.
- Risk characterization, which merges information about exposure and toxicity to qualitatively and quantitatively estimate if an individual or a population is exposed to a risk.

4.1 Pesticides

According to the current literature, the highest number of food alerts (almost 60% of all the total food alerts collected) reported to date to the RASFF are related to the occurrence of pesticides residues in food (Casado et al., 2024). Pesticides are chemical compounds intentionally used in agriculture to prevent, control or heal pests that can damage crops and consequentially compromise food quality. They comprise a wide range of substances targeting specific organisms, including insecticides, herbicides, fungicides, and rodenticides. However, the use of pesticides may lead to the presence in certain foods of their residues, their metabolites, as well as of their degradation or reaction products (Aagaard et al., 2023). Pesticides can also be classified according to their chemical structure. The most common classes of pesticides employed for the protection of food crops are organophosphorothioates (OPTs) pesticides, carbamates, pyrethroids, organochlorines, triazines, benzimidazoles, and nitro compounds (Sulaiman et al., 2019). According to the recent food alerts reported for the occurrence of pesticides in food, special attention was given to OPTs, such as chlorpyrifos (CPF) and dimethoate. From a chemical standpoint, OPTs are low-molecular weight compounds characterized by the presence of phosphoric acid ester derivatives with at least one oxygen substituted by a sulfur atom. This class of pesticides is mainly used as insecticides causing neurotoxicity upon bioactivation and inhibition of brain and serum acetylcholinesterase (AChE) (Aroniadou-Anderjaska et al., 2023). The reason why they represent a global earth concern is that they may be potentially toxic also to other species, especially mammals and humans, even though their mechanism of action on human enzymes is still not well clarified. OPTs represent a primary risk for such working categories as farmers, but also part of the general population that may be exposed to them using domestic insecticide or because they can be found as residues in food and drinking water (Mojsak et al., 2018). An *in silico* local QSAR model of bioconcentration factor of organophosphate pesticides was developed in 2021 to prioritize the attention toward specific organophosphate and to facilitate risk assessment (Banjare et al., 2021). Concerning their effects in humans, it is well known that certain

cytochromes P450 (CYPs) promote OPTs bioactivation, but this may result in the formation of the corresponding oxon metabolite or in their detoxification (through reaction like dearylation, dealkylation or hydroxylation). However, the underlying mechanisms and chemical underpinnings still need clarifications. Pedroni et al. applied a computational pipeline made up of molecular docking and molecular dynamics simulation using the legacy OPT CPF as a case study to elucidate the reason underlying its diverse metabolic fate. The work pointed out the importance of orientation and of the distance between phosphorotioate and heme, revealing these features as crucial for their bioactivation (Pedroni, Perugino, Dall'Asta, et al., 2024). It is interesting to note how computational strategies could be broader employed and used also to improve risk assessment.

4.2 Mycotoxins

Mycotoxins are secondary metabolites mainly produced by certain filamentous fungi belonging to the genera *Aspergillus*, *Penicillium*, and *Fusarium*. These compounds can contaminate a wide range of food commodities and animal feeds such as grains, nuts, fruits, coffee, spices and vegetables representing a major concern for food safety worldwide (Mihalache et al., 2023). Mycotoxins can contaminate food stuff on fields, during harvest or during storage. Moreover, the phenomenon of carry-over from feed to animals intended for food consumption has not to be overlooked. The health effects of mycotoxins range from acute to chronic toxicity such as immunosuppression, carcinogenesis, endocrine-disrupting activity, gastrointestinal disorders, hepatotoxicity and kidney damage (Eaton & Gallagher, 1994; Mihalache et al., 2023; Tola et al., 2016). From a regulatory perspective, a hundred mycotoxins have been characterized, but only a few of them are found in food and some of them are currently subject to regulation. The European Commission established different maximum limits for mycotoxin exposure in adults and children through Commission Regulation (EC) No. 915/2023. The European Union strictly regulates several mycotoxins, such as aflatoxins (AFB1, B2, G1, G2, and M1), deoxynivalenol (DON), fumonisins (FB1 and FB2), ochratoxin A (OTA), zearalenone (ZEN), and patulin (PAT) in specific foods and food commodities ("COMMISSION REGULATION (EU) 2023/915 of 25 April 2023 on maximum levels for certain contaminants in food and repealing Regulation (EC) No 1881/2006," 2023). Mycotoxins are generally thermostable and food processing operations commonly applied, such as cooking, boiling, baking, frying, baking, and pasteurizing have little effects on mycotoxins degradation. Currently, the most common analytical approaches for determining mycotoxins are divided into two categories: reference

methods for quantitative analysis and rapid methods for first-level screening. However, these methods are often expensive and time-consuming. For this reason, AI techniques may represent a valid way to enhance the speed and improve mycotoxin detection. Among the most applied AI techniques for mycotoxin detection are Image Recognition, Spectroscopy and Sensor Data Analysis, Biosensor, and Predictive Models. In the field of mycotoxins risk assessment, the toxicodynamic of these compounds can be investigated using *in silico* methods to support the first two steps of the process (hazard identification and hazard characterization). Toxicodynamic, in fact, determines the relationship between the concentration of a toxicant at the site of action and the toxic effect at the level of molecule, cell, tissue, organ or organism (Dellafiora, Dall'Asta, et al., 2018). *In silico* approaches have been successfully applied also to toxicokinetic evaluation that describes how the human body deals with a toxic compound in terms of ADME (Absorption, Distribution, Metabolism and Excretion). For example, the investigation from a mechanistic point of view of the impact of missense mutation on certain enzymes provides information about inter-individual differences in toxicokinetic (Dorne et al., 2022).

4.2.1 Deoxynivalenol

Deoxynivalenol (DON) (Figure 2), also known as vomitoxin, is a type 2 trichothecene mycotoxin produced as secondary metabolite by *Fusarium* fungi. It is a contaminant of crops such as wheat, maize, and barley and animal feed worldwide. When ingested in large quantities, DON leads to acute toxicity, while in small quantities over time it may lead to chronic toxicity. Acute intoxication is responsible for damage at the level of gastro-intestinal apparatus, alteration in the normal food and feed intake, and nutrients absorption, immunosuppression and susceptibility to infection (Ganesan et al., 2022). Whereas acute toxicity includes cytotoxicity, immunotoxicity, and neurotoxicity. At the molecular level, trichothecene mycotoxins are known to block protein synthesis. Several analogues and derivatives of DON, including 3-acetyl-DON (3-Ac-DON), 15-acetyl-DON (15-Ac-DON), and DON-3-glucoside (DON-3-Glc). The last one is a so-called “masked” mycotoxin formed during plant metabolism (Berthiller et al., 2013). Despite the knowledge about DON and analogues, many aspects concerning the toxicity and metabolism of DON remain to be clarified (Del Favero et al., 2018). Concerning the capability of in silico approaches to elucidate also toxicological aspects and mechanisms, Dellafiora et al., in depth investigated the ribotoxicity of trichothecenes analyzing the structure-activity relationship of trichothecenes through a computational approach based on

docking simulations and pharmacophoric analysis. The study revealed that modified forms with steric hindrance in position 3 of the sesquiterpenoid core may be responsible for a loss of ribotoxicity. Moreover, positions 4,8 and 15 were found more influenced by hydrophobic/hydrophilic mismatch (Dellafiora, Galaverna, et al., 2017).

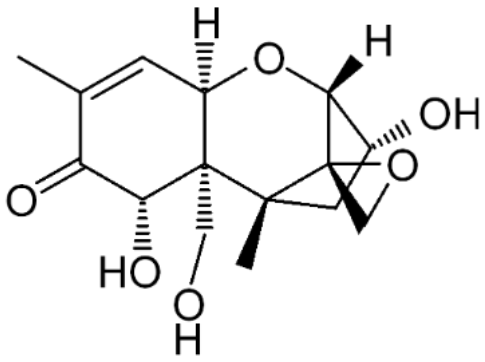


Figure 2. Chemical structure of DON.

4.2.2 Zearalenone

Zearalenone (ZEN) (Figure 3) is a nonsteroidal estrogenic mycotoxin produced by several *Fusarium* species. Its worldwide occurrence in cereal crops and in grain-based products represents a concern to human and animal health due to its endocrine disrupting effects possibly causing precocious puberty, premature thelarche and carcinogenic actions (Zheng et al., 2019). This is due to its chemical structure which resembles 17β-estradiol. In fact, it exerts its estrogenic activity competing with estradiol for activating estrogen receptors (ERs) and inhibiting aromatase (CYP19A1). In 2011, EFSA Panel on Contaminants in the Food Chain established a tolerable daily intake for ZEN of 0.25 μg/kg body weight ((CONTAM), 2016). Also, ZEN metabolites like α-zearalanol (α-ZAL), α-zearalenol (α-ZEL), β-zearalanol (β-ZAL), and β-zearalenol (β-ZEL), which are co-present in contaminated cereals and other crops, are correlated to endocrine disruptor alterations representing a primary health concern (Agahi et al., 2020). Applying an *in silico* pipeline, Cozzini & Dellafiora evaluated the interaction between ZEN and its metabolites and the ERs. Performing an integrated approach of docking simulations and rescoring procedures upon a small training set of well characterized compounds, they defined a range of scores values able to distinguish classes of compounds with and without active estrogenic behavior. Results indicated that ZEN and reductive metabolites show score values included in the cluster of xenoestrogens agonists, revealing that their

approach can be a powerful tool for a preliminary estrogenic evaluation of many compounds (Cozzini & Dellafiora, 2012). *In silico* approach can be also combined with *in vitro* approach. Concerning ZEN and its metabolites, Dellafiora et al., performed a hybrid in silico/in vitro study to investigate the xenoestrogenicity of zearalenone-14-glucoside (ZEN14Glc), compared to ZEN, deepening the underlying molecular mechanisms. Results revealed that ZEN14Glc effectively stimulated a xenoestrogenic response in cells, but such stimulus can be entirely attributable to the hydrolysis phenomenon, as the glycosylated form turned out to be unable to effectively bind and activate the estrogens receptors (Dellafiora, Ruotolo, et al., 2017).

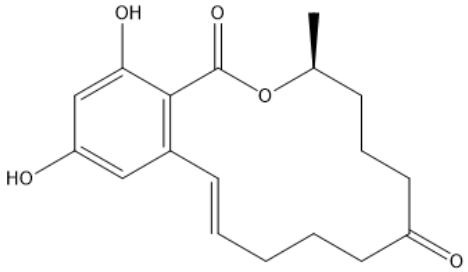


Figure 3. Chemical structure of ZEN.

4.2.3 Ochratoxin A

Ochratoxin A (OTA) (Figure 4) is a low molecular-weight secondary metabolite that is produced by fungal species mainly belonging to *Aspergillus* and *Penicillium* genera (van der Merwe et al., 1965). From a chemical point of view, OTA is a phenylalanine derivative resulting from the condensation of the amino group of L-phenylalanine with the carboxy group of (3R)-5-chloro-8-hydroxy-3-methyl-1-oxo-3,4-dihydro-1H-2-benzopyran-7-carboxylic acid. OTA is among the mycotoxins of most concern to food and feed safety due to many adverse effects on humans and animals including nephrotoxic, hepatotoxic, teratogenic and immunotoxic effects reported in several animal species (van der Merwe et al., 1965). In this respect, OTA has been classified by the International Agency of Research on Cancer (IARC) in Group 2B as a possible human carcinogen (IARC, 1993). Furthermore, from a cellular standpoint damage of protein biosynthesis and energy production, induction of oxidative stress, cells apoptosis and effects on mitosis and cell cycle regulation have been found related to OTA exposure (Garcia-Esparza et al., 2025). Along with OTA, certain food commodities can be contaminated also by degradation

products like 2'R-OTA, an OTA diastereomer. These degradation products are usually less toxic, however certain populations group may be highly exposed to them. Concerning 2'R-OTA, for example, its production occurs during heat processing (e.g. roasting, baking), leading to the high exposure of people like coffee drinkers (Cramer et al., 2008). OTA can also produce different metabolites. The most abundant one is OTα, followed by ochratoxin B (OTB; the non-chlorinated analogue of OTA), as well as by ochratoxin hydroquinone (OTHQ) and ochratoxin quinone (OTQ), two debated potential reactive metabolites of OTA (Dai et al., 2002).

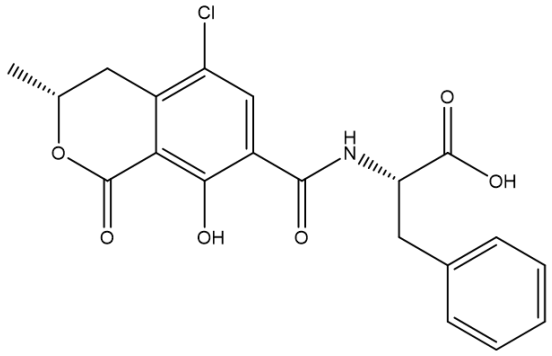


Figure 4. Chemical structure of OTA.

4.2.4 Fumonisins

Fumonisins (Figure 5) are well known mycotoxins produced by *Fusarium verticillioides*, *Fusarium proliferatum* and other *Fusarium* species. They are responsible for the contamination of maize and maize-based products, especially in warm and humid climate conditions (Santiago et al., 2015). FB1 is the most common fumonisin detected in food, not only in maize, but also in beer, rice, sorghum, triticale, cowpea seeds, beans and soybeans (Scott, 2012). Fumonisins adverse effects reported so far are carcinogenicity, hepatotoxicity, nephrotoxicity and embryotoxicity (Missmer et al., 2006). According to this evidence, IARC designated FB_1 in Group 2B as 'possibly carcinogenic to humans'. From a structural standpoint, fumonisins are polar, amphiphilic compounds characterized by a long-chain aliphatic backbone with multiple hydroxyl groups and two tricarballylic acid (TCA) side chains. Their structure resembles sphingoid bases, for this reason they may impair sphingolipid metabolism and rheostat (Seefelder et al., 2003).

Among the modified forms of FB1 which can occur in food contaminated commodities, the hydrolysed FB1 (HFB1; Figure 5) is the most predominant, mainly produced upon alkali food processing (Zimmer et al., 2008). However, not only the hydrolized forms, but also the N-acylated derivatives should be investigated to support the group-based risk assessment (Dellafiora, Galaverna, et al., 2018).

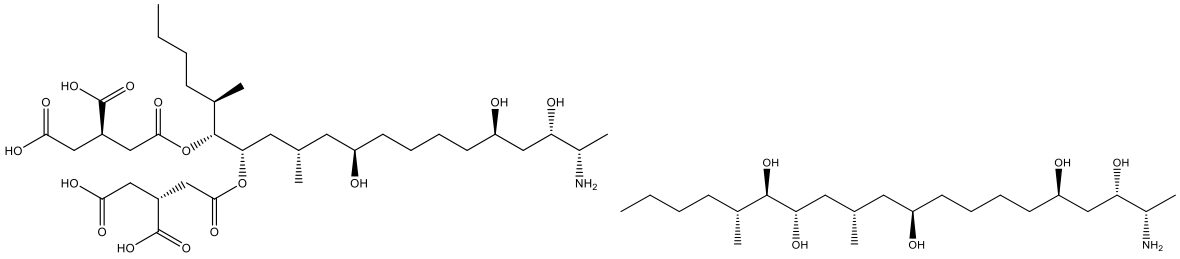


Figure 5. Chemical structure of FB1 (left) and HFB1 (right).

5. Aim and outline of the thesis

This PhD thesis aims to integrate AI techniques into computational approaches for studying the interaction between macromolecules (i.e. proteins) and small molecules, with a particular focus on molecules of food interest either for safety reasons, or to dissect the mechanisms underpinning their capability to elicit taste or bioactivity. Specifically, this work will investigate three major groups of compounds: taste active molecules, bioactive compounds, and toxic compounds (with a focus on mycotoxins). These classes of compounds were studied to expand the current knowledge about their activity as well as to better understand the connection between food/taste and its consequences on human health. In the last decades, in fact, taste receptors have been identified also in extra-oral tissue and recognized as potential pharmacological targets. In this context, their potential involvement in the interaction with food-derived bioactive and toxic compounds deserve further investigation. At the same time, the molecular mechanisms underlying the toxic effects of certain mycotoxins, in particular the interaction with key targets such as CYP450s, OGFOD1, CerS5 and different proteins involved in PD AOP, have been elucidated. This integrative approach aims to deepen the knowledge about the underlying molecular mechanisms, trying to find a connection between taste perception, bioactivity, and toxicity within the food context. A further goal of this work is to integrate AI techniques into the *in silico* pipeline developed, empowering their applicability and throughput. This has been achieved by implementing AI tools in both ligand-based and structure-based virtual screening, as well as in protein structure prediction.

The content of each chapter is briefly described below.

In Chapter 2, the interaction between food-related toxic compounds and TAS2R46 (i.e. one of the human bitter receptors) has been investigated at a molecular level. TASRs are not only present in the oral cavity, but they are expressed also in various extra-oral organs and tissues. Although the role of extra-oral expression of bitter receptors is far from being completely understood, they could be involved in different aspects related to appetite and satiety and are also likely to have biological activities that might be relevant from a physiological/toxicological point of view. In this contest, an AI-based virtual screening procedure was successfully applied to fish bitter toxic compounds from T3DB database (http://www.t3db.ca). The integration of AI has enabled the generation of a model trained on more than 25,000 molecules and the classification as bitter or non-bitter of more than 3,500 molecules. This analysis described

trichothecolone, a member of type 2 trichothecenes class as a possible binder of TAS2R46. This class includes mycotoxins relevant from a food safety perspective. Molecular docking and molecular dynamics were then performed to study the interaction between TAS2R46 and a series of type B trichothecenes. Deoxynivalenol and its 15-acetyl derivative were identified as possible TAS2R46 activators.

In Chapter 3, the impact of a single nucleotide variant (SNVs) of the aromatase enzyme (CYP19A1) on the capability of ZEN and αZEL to inhibit this enzyme was investigated. Aromatase, in fact, is one of the key enzymes involved in the production of estrogens from androgens, and it is known the ability of ZEN to inhibit this enzyme. However, aromatase is also known to display many SNVs (>400) that may affect the inhibitory capability of this mycotoxin. Leveraging a big data-based analysis approach and application of advanced computational modelling, all the current available SNVs were retrieved from the Uniprot database, and the 14 SNVs closer to aromatase binding site were further analyzed through molecular docking and molecular dynamics simulations. Results revealed that T310S variant was likely to have an enhanced susceptibility to ZEN and αZEL compared to wild-type, while D309G and S478F were described as potentially inactive aromatase variants.

In Chapter 4, a virtual pipeline based on molecular docking, molecular dynamics simulations and umbrella simulations was developed to display new Ochratoxin A (OTA) potential targets. Starting from the outcomes of a previously conducted virtual screening campaign, OGFOD1 (known to play a key role in protein synthesis) was identified as OTA potential target. In this work, it was described the capability of OTA and 2'R-OTA diastereomer, a thermal degradation product, to interact and possibly inhibit OGFOD1. This thermal degradation product, in fact, seems to be less toxic compared to OTA, but there are some population groups such as coffee drinkers that may be highly exposed to this compound. This work represents a proof of concept on how computational approaches may broaden the investigation toward new targets never considered before, laying the foundations for a better understanding of OTA toxicity.

In Chapter 5, starting from evidence suggesting that OTA may contribute to Parkinson's disease (PD) pathogenesis, the potential role of OTA and its main metabolites (i.e. OTα, OTB, OTH and OTHQ) at multiple levels along the Parkinson's Adverse Autcome Pathway (AOP) was investigated. Applying a computational pipeline made up of molecular modelling techniques, this study investigated key events in the PD AOP where the current state-of-the-art suggests OTA and its congeners may play a role, i.e. inhibition of cathepsins B, D, and L

(relevant for α-synuclein proteostasis), and Complex I (responsible for mitochondrial dysfunction). The study revealed that none of the compounds under investigation could interact with cathepsin B, but some of them displayed differential binding with cathepsins D and L. Furthermore, ligands binding architectures were analyzed through Markov State Models (MSMs) which described the dynamic of their conformational landscapes within the binding site. These outcomes provided mechanistic insights correlating OTA exposure and PD pathogenesis and revealed a novel and multi-tiered mechanism where OTA and congeners may affect α-synuclein proteostasis and mitochondrial function.

In Chapter 6, a computational pipeline based on molecular docking and molecular dynamics simulations was applied to study the molecular rationale underpinning the dual role of fumonisins focusing on the human Ceramide Synthase 5 (CerS5) as a case study. FB1 and HFB1, in fact, are among the mycotoxins of most concern to food safety. Due to their chemical similarity with sphingoid bases, they may both alter the sphingolipid rheostat by inhibiting the ceramide synthases (CerSs) and act as CerSs substrates. However, the basis of this dual role is still unclear also due to the lack of the structure of the human CerSs at the time of analysis, specifically of CerS5 The rapid development of AI-based tools for protein structure prediction filled this gap and enabled the study of the interaction of FB1 and HFB1 with the CerS5 catalytic and inhibitor binding sites from a molecular standpoint. Furthermore, the two mechanisms of catalysis (i.e. one-step and two step catalytic mechanisms) proposed were investigated. The structure of CerS5 was obtained through an AlphaFold2-based notebook using the crystallographic structure of the yeast CerS as template. AlphaFold2 is a protein modelling system based on ML combining neural networks and homology modelling. Results described HFB1, but not FB1, as an optimal ligand for the catalytic site. Conversely, both fumonisins could fit the hydrophilic Acyl-CoA binding site, pointing to their capability to act as substrate through the two-step mechanism of catalysis and listing the impairment of Acyl-CoA binding among the plausible mechanisms leading to CerS5 inhibition. Moreover, the diffusion of FB1 through the membrane was calculated through molecular dynamics simulations and revealed to be lower than HFB1, providing a general rationale for the lower formation of FB1 N-acyl derivatives described in the literature.

In Chapter 7, the mechanics of GPR120-ligand interaction were investigated to identify new potential GPR120 peptide ligands. In recent years, GPR120, a protein belonging to the GPCR family, drew attention as an interesting pharmacological target to cope with obesity, satiety and diabetes. Long chain fatty acids are GPR120 natural agonists, but also other synthetic molecules

were identified as agonists expanding the chemical space of GPR120's ligands. This work proposed for the first time peptides as possible GPR120 binders. The computational pipeline started with a narrow filtering of hexapeptides based on their chemical similarity with known GPR120 agonists. The *in silico* approach allowed the generation and the ligand-based virtual screening of 531 441 sequences. The best hits were tested through docking studies, molecular dynamics and umbrella sampling simulations, which pointed to G[I,L]FGGG as a promising GPR120 agonist sequence. The presence of both peptides in food-related proteins was thoroughly assessed, revealing they are encrypted proteins of mushrooms, food-grade bacteria and rice. This research was performed through an iterative sequence alignment procedure which enables us to search each of the most promising within the whole SwissProt database (570 420 protein sequences). Simulations on the counterparts with D-amino acids - known to be more resistant to gastrointestinal digestion - were also performed revealing an interaction with GPR120 possibly more favored than that with its L counterpart.

In Chapter 8, the already proved reliability of molecular docking was combined with the efficiency of AI to empower and reinforce the discovery of bioactive compounds. The study aimed at identifying novel GPR120 (Free Fatty Acid Receptor 4) ligands, being also a relevant pharmacological target. Chemical compounds libraries are huge, and the screening of these libraries with traditional structure-based methods would be expensive both in terms of costs and time. For this reason, the integration of a ML model has been proposed to make it affordable. Specifically, a CatBoost model trained on docking score and Morgan fingerprints (Extendended Connectivity Fingerprints) was developed and used to classify potential GPR120 binders and non-binders avoiding docking of unpromising molecules. The model developed is not intended to replace molecular docking but to prefilter compounds and reduce the number of compounds requiring further docking calculations. Key docking limitations (i.e. bias toward larger molecules) were overcome by stratifying the molecules used to build the model according to molecular size. This approach succeeded in reducing time and computational efforts and could be readily adapted to other GPCRs or to other molecular targets relevant from a pharmaceutical and/or from a food science perspective.

In Chapter 9, in this last chapter the main findings of the thesis and the future prospects of the work are presented.

References

(CONTAM), E. P. o. C. i. t. F. C. (2016). Appropriateness to set a group health-based guidance value for zearalenone and its modified forms. *EFSA Journal, 14*(4). doi:10.2903/j.efsa.2016.4425

Aagaard, A., Berny, P., Chaton, P. F., Antia, A. L., McVey, E., Arena, M., . . . Brock, T. (2023). Risk assessment for Birds and Mammals. *EFSA Journal, 21*(2). doi:10.2903/j.efsa.2023.7790

Accardo, F., Prandi, B., Dellafiora, L., Tedeschi, T., & Sforza, S. (2024). How D-amino acids embedded in the protein sequence modify its digestibility: Behaviour of digestive enzymes tested on a model peptide used as target. *Food Chem, 458*, 140175. doi:10.1016/j.foodchem.2024.140175

Agahi, F., Juan, C., Font, G., & Juan-Garcia, A. (2020). In silico methods for metabolomic and toxicity prediction of zearalenone, alpha-zearalenone and beta-zearalenone. *Food Chem Toxicol, 146*, 111818. doi:10.1016/j.fct.2020.111818

Aguero-Pizzolo, S., Bettler, E., & Gouet, P. (2025). [Nobel Prize in chemistry 2024: David Baker, Demis Hassabis et John M. Jumper. The revolution of artificial intelligence in structural biology]. *Med Sci (Paris), 41*(4), 367-373. doi:10.1051/medsci/2025060

Agyei, D., Tsopmo, A., & Udenigwe, C. C. (2018). Bioinformatics and peptidomics approaches to the discovery and analysis of food-derived bioactive peptides. *Anal Bioanal Chem, 410*(15), 3463-3472. doi:10.1007/s00216-018-0974-1

Ahn, C.-B., Je, J.-Y., & Cho, Y.-S. (2012). Antioxidant and anti-inflammatory peptide fraction from salmon byproduct protein hydrolysates by peptic hydrolysis. *Food Research International, 49*(1), 92-98. doi:10.1016/j.foodres.2012.08.002

Alves, L. A., Ferreira, N., Maricato, V., Alberto, A. V. P., Dias, E. A., & Jose Aguiar Coelho, N. (2021). Graph Neural Networks as a Potential Tool in Improving Virtual Screening Programs. *Front Chem, 9*, 787194. doi:10.3389/fchem.2021.787194

Amendola, G., & Cosconati, S. (2021). PyRMD: A New Fully Automated AI-Powered Ligand-Based Virtual Screening Tool. *J Chem Inf Model, 61*(8), 3835-3845. doi:10.1021/acs.jcim.1c00653

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science, 181*(4096), 223-230. doi:10.1126/science.181.4096.223

Aroniadou-Anderjaska, V., Figueiredo, T. H., de Araujo Furtado, M., Pidoplichko, V. I., & Braga, M. F. M. (2023). Mechanisms of Organophosphate Toxicity and the Role of Acetylcholinesterase Inhibition. *Toxics, 11*(10). doi:10.3390/toxics11100866

Babbar, N., Oberoi, H. S., & Sandhu, S. K. (2015). Therapeutic and nutraceutical potential of bioactive compounds extracted from fruit residues. *Crit Rev Food Sci Nutr, 55*(3), 319-337. doi:10.1080/10408398.2011.653734

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., . . . Baker, D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science, 373*(6557), 871-876. doi:10.1126/science.abj8754

Banjare, P., Matore, B., Singh, J., & Roy, P. P. (2021). In silico local QSAR modeling of bioconcentration factor of organophosphate pesticides. *In Silico Pharmacol, 9*(1), 28. doi:10.1007/s40203-021-00087-w

Berthiller, F., Crews, C., Dall'Asta, C., Saeger, S. D., Haesaert, G., Karlovsky, P., . . . Stroka, J. (2013). Masked mycotoxins: a review. *Mol Nutr Food Res, 57*(1), 165-186. doi:10.1002/mnfr.201100764

Biesalski, H. K., Dragsted, L. O., Elmadfa, I., Grossklaus, R., Muller, M., Schrenk, D., . . . Weber, P. (2009). Bioactive compounds: definition and assessment of activity. *Nutrition, 25*(11-12), 1202-1205. doi:10.1016/j.nut.2009.04.023

Bo, W., Qin, D., Zheng, X., Wang, Y., Ding, B., Li, Y., & Liang, G. (2022). Prediction of bitterant and sweetener using structure-taste relationship models based on an artificial neural network. *Food Res Int, 153*, 110974. doi:10.1016/j.foodres.2022.110974

Bonetta, R., & Valentino, G. (2020). Machine learning techniques for protein function prediction. *Proteins, 88*(3), 397-413. doi:10.1002/prot.25832

Borgonovi, S. M., Perugino, F., Dellafiora, L., Annunziata, F., Pedroni, L., Galaverna, G., . . . Di Nunzio, M. (2025). Assessing the impact of food-derived bioactives on digestive proteases by in vitro and in silico approaches. *Food Funct, 16*(8), 2959-2971. doi:10.1039/d4fo04022h

Bortolomedi, B. M., Paglarini, C. S., & Brod, F. C. A. (2022). Bioactive compounds in kombucha: A review of substrate effect and fermentation conditions. *Food Chem, 385*, 132719. doi:10.1016/j.foodchem.2022.132719

Cano, G., Garcia-Rodriguez, J., Garcia-Garcia, A., Perez-Sanchez, H., Benediktsson, J. A., Thapa, A., & Barr, A. (2017). Automatic selection of molecular descriptors using random forest: Application to drug discovery. *Expert Systems with Applications, 72*, 151-159. doi:10.1016/j.eswa.2016.12.008

Casado, N., Berenguer, C. V., Camara, J. S., & Pereira, J. A. M. (2024). What Are We Eating? Surveying the Presence of Toxic Molecules in the Food Supply Chain Using Chromatographic Approaches. *Molecules, 29*(3). doi:10.3390/molecules29030579

Chalamaiah, M., Keskin Ulug, S., Hong, H., & Wu, J. (2019). Regulatory requirements of bioactive peptides (protein hydrolysates) from food proteins. *Journal of Functional Foods, 58*, 123-129. doi:10.1016/j.jff.2019.04.050

Chen, B., Harrison, R. F., Papadatos, G., Willett, P., Wood, D. J., Lewell, X. Q., . . . Stiefl, N. (2007). Evaluation of machine-learning methods for ligand-based virtual screening. *J Comput Aided Mol Des, 21*(1-3), 53-62. doi:10.1007/s10822-006-9096-5

Chhetri, K. B. (2023). Applications of Artificial Intelligence and Machine Learning in Food Quality Control and Safety Assessment. *Food Engineering Reviews, 16*(1), 1-21. doi:10.1007/s12393-023-09363-1

Cláudia da Costa Rocha, A., José de Andrade, C., & de Oliveira, D. (2021). Perspective on integrated biorefinery for valorization of biomass from the edible insect Tenebrio molitor. *Trends in Food Science & Technology, 116*, 480-491. doi:10.1016/j.tifs.2021.07.012

COMMISSION REGULATION (EU) 2023/915 of 25 April 2023 on maximum levels for certain contaminants in food and repealing Regulation (EC) No 1881/2006, (2023).

Cozzini, P., & Dellafiora, L. (2012). In silico approach to evaluate molecular interaction between mycotoxins and the estrogen receptors ligand binding domain: a case study on zearalenone and its metabolites. *Toxicol Lett, 214*(1), 81-85. doi:10.1016/j.toxlet.2012.07.023

Cramer, B., Konigs, M., & Humpf, H. U. (2008). Identification and in vitro cytotoxicity of ochratoxin A degradation products formed during coffee roasting. *J Agric Food Chem, 56*(14), 5673-5681. doi:10.1021/jf801296z

Cygankiewicz, A. I., Maslowska, A., & Krajewska, W. M. (2014). Molecular basis of taste sense: involvement of GPCR receptors. *Crit Rev Food Sci Nutr, 54*(6), 771-780. doi:10.1080/10408398.2011.606929

da Rocha, M. N., de Sousa, D. S., da Silva Mendes, F. R., Dos Santos, H. S., Marinho, G. S., Marinho, M. M., & Marinho, E. S. (2025). Ligand and structure-based virtual screening approaches in drug discovery: minireview. *Mol Divers, 29*(3), 2799-2809. doi:10.1007/s11030-024-10979-6

Dai, J., Park, G., Wright, M. W., Adams, M., Akman, S. A., & Manderville, R. A. (2002). Detection and characterization of a glutathione conjugate of ochratoxin A. *Chem Res Toxicol, 15*(12), 1581-1588. doi:10.1021/tx0255929

Damak, S., Rong, M., Yasumatsu, K., Kokrashvili, Z., Varadarajan, V., Zou, S., . . . Margolskee, R. F. (2003). Detection of sweet and umami taste in the absence of taste receptor T1r3. *Science, 301*(5634), 850-853. doi:10.1126/science.1087155

Dang, Y., Hao, L., Cao, J., Sun, Y., Zeng, X., Wu, Z., & Pan, D. (2019). Molecular docking and simulation of the synergistic effect between umami peptides, monosodium glutamate and taste receptor T1R1/T1R3. *Food Chem, 271*, 697-706. doi:10.1016/j.foodchem.2018.08.001

Das, B., Mathew, A. T., Baidya, A. T. K., Devi, B., Salmon, R. R., & Kumar, R. (2024). Artificial intelligence assisted identification of potential tau aggregation inhibitors: ligand- and structure-based virtual screening, in silico ADME, and molecular dynamics study. *Mol Divers, 28*(4), 2013-2031. doi:10.1007/s11030-023-10645-3

Del Favero, G., Woelflingseder, L., Braun, D., Puntscher, H., Kutt, M. L., Dellafiora, L., . . . Marko, D. (2018). Response of intestinal HT-29 cells to the trichothecene mycotoxin deoxynivalenol and its sulfated conjugates. *Toxicol Lett, 295*, 424-437. doi:10.1016/j.toxlet.2018.07.007

Del Rio, D., Costa, L. G., Lean, M. E., & Crozier, A. (2010). Polyphenols and health: what compounds are involved? *Nutr Metab Cardiovasc Dis, 20*(1), 1-6. doi:10.1016/j.numecd.2009.05.015

Dellafiora, L., Dall'Asta, C., & Galaverna, G. (2018). Toxicodynamics of Mycotoxins in the Framework of Food Risk Assessment-An In Silico Perspective. *Toxins (Basel), 10*(2). doi:10.3390/toxins10020052

Dellafiora, L., Galaverna, G., & Dall'Asta, C. (2017). In silico analysis sheds light on the structural basis underlying the ribotoxicity of trichothecenes-A tool for supporting the hazard identification process. *Toxicol Lett, 270*, 80-87. doi:10.1016/j.toxlet.2017.02.015

Dellafiora, L., Galaverna, G., & Dall'Asta, C. (2018). Mechanisms of Fumonisin B1 Toxicity: A Computational Perspective beyond the Ceramide Synthases Inhibition. *Chem Res Toxicol, 31*(11), 1203-1212. doi:10.1021/acs.chemrestox.8b00188

Dellafiora, L., Gonaus, C., Streit, B., Galaverna, G., Moll, W. D., Vogtentanz, G., . . . Prasad, S. (2020). An In Silico Target Fishing Approach to Identify Novel Ochratoxin A Hydrolyzing Enzyme. *Toxins (Basel), 12*(4). doi:10.3390/toxins12040258

Dellafiora, L., Magnaghi, F., Galaverna, G., & Dall'Asta, C. (2022). A mechanistic investigation on kokumi-active gamma-Glutamyl tripeptides - A computational study to understand molecular basis of their activity and to identify novel potential kokumi-tasting sequences. *Food Res Int, 162*(Pt A), 111932. doi:10.1016/j.foodres.2022.111932

Dellafiora, L., Mena, P., Cozzini, P., Brighenti, F., & Del Rio, D. (2013). Modelling the possible bioactivity of ellagitannin-derived metabolites. In silico tools to evaluate their potential xenoestrogenic behavior. *Food Funct, 4*(10), 1442-1451. doi:10.1039/c3fo60117j

Dellafiora, L., Ruotolo, R., Perotti, A., Cirlini, M., Galaverna, G., Cozzini, P., . . . Dall'Asta, C. (2017). Molecular insights on xenoestrogenic potential of zearalenone-14-glucoside through a mixed in vitro/in silico approach. *Food Chem Toxicol, 108*(Pt A), 257-266. doi:10.1016/j.fct.2017.07.062

Dhingra, S., Sowdhamini, R., Cadet, F., & Offmann, B. (2020). A glance into the evolution of template-free protein structure prediction methodologies. *Biochimie, 175*, 85-92. doi:10.1016/j.biochi.2020.04.026

Di Pizio, A., Ben Shoshan-Galeczki, Y., Hayes, J. E., & Niv, M. Y. (2019). Bitter and sweet tasting molecules: It's complicated. *Neurosci Lett, 700*, 56-63. doi:10.1016/j.neulet.2018.04.027

Directive 2002/46/EC of the European Parliament and of the Council of 10 June 2002 on the Approximation of the Laws of the Member States Relating to Food Supplements. In (pp. 51–57).

Dorne, J., Cirlini, M., Louisse, J., Pedroni, L., Galaverna, G., & Dellafiora, L. (2022). A Computational Understanding of Inter-Individual Variability in CYP2D6 Activity to Investigate the Impact of Missense Mutations on Ochratoxin A Metabolism. *Toxins (Basel), 14*(3). doi:10.3390/toxins14030207

Eaton, D. L., & Gallagher, E. P. (1994). Mechanisms of aflatoxin carcinogenesis. *Annu Rev Pharmacol Toxicol, 34*, 135-172. doi:10.1146/annurev.pa.34.040194.001031

Eissa, F., Sebaei, A. S., & El Badry Mohamed, M. (2024). Food additives and flavourings: Analysis of EU RASFF Notifications from 2000 to 2022. *Journal of Food Composition and Analysis, 130*. doi:10.1016/j.jfca.2024.106137

Friedman, M. (2010). Origin, microbiology, nutrition, and pharmacology of D-amino acids. *Chem Biodivers, 7*(6), 1491-1530. doi:10.1002/cbdv.200900225

Galindo, M. M., Voigt, N., Stein, J., van Lengerich, J., Raguse, J. D., Hofmann, T., . . . Behrens, M. (2012). G protein-coupled receptors in human fat taste perception. *Chem Senses, 37*(2), 123-139. doi:10.1093/chemse/bjr069

Ganesan, A. R., Mohan, K., Karthick Rajan, D., Pillay, A. A., Palanisami, T., Sathishkumar, P., & Conterno, L. (2022). Distribution, toxicity, interactive effects, and detection of ochratoxin and deoxynivalenol in food: A review. *Food Chem, 378*, 131978. doi:10.1016/j.foodchem.2021.131978

Garcia-Esparza, M. A., Mateo, E. M., Robles, J. A., Capoferri, M., Jimenez, M., & Soria, J. M. (2025). Unveiling the Neurotoxic Effects of Ochratoxin A and Its Impact on Neuroinflammation. *Toxins (Basel), 17*(6). doi:10.3390/toxins17060264

Gawehn, E., Hiss, J. A., & Schneider, G. (2016). Deep Learning in Drug Discovery. *Mol Inform, 35*(1), 3-14. doi:10.1002/minf.201501008

Genchi, G. (2017). An overview on D-amino acids. *Amino Acids, 49*(9), 1521-1533. doi:10.1007/s00726-017-2459-5

Gentile, F., Yaacoub, J. C., Gleave, J., Fernandez, M., Ton, A. T., Ban, F., . . . Cherkasov, A. (2022). Artificial intelligence-enabled virtual screening of ultra-large chemical libraries with deep docking. *Nat Protoc, 17*(3), 672-697. doi:10.1038/s41596-021-00659-2

Gonczarek, A., Tomczak, J. M., Zareba, S., Kaczmar, J., Dabrowski, P., & Walczak, M. J. (2018). Interaction prediction in structure-based virtual screening using deep learning. *Comput Biol Med, 100*, 253-258. doi:10.1016/j.compbiomed.2017.09.007

He, Y., Liu, K., Yu, X., Yang, H., & Han, W. (2024). Building a Kokumi Database and Machine Learning-Based Prediction: A Systematic Computational Study on Kokumi Analysis. *J Chem Inf Model, 64*(7), 2670-2680. doi:10.1021/acs.jcim.3c01728

Hert, J., Willett, P., Wilton, D. J., Acklin, P., Azzaoui, K., Jacoby, E., & Schuffenhauer, A. (2006). New methods for ligand-based virtual screening: use of data fusion and machine learning to enhance the effectiveness of similarity searching. *J Chem Inf Model, 46*(2), 462-470. doi:10.1021/ci050348j

Huang, W., Shen, Q., Su, X., Ji, M., Liu, X., Chen, Y., . . . Zhang, J. (2016). BitterX: a tool for understanding bitter taste in humans. *Sci Rep, 6*, 23450. doi:10.1038/srep23450

IARC. (1993). Some Naturally Occurring Substances: Food Items and Constituents, Heterocyclic Aromatic Amines and Mycotoxins.

Jayaraj, P. B., & Jain, S. (2019). Ligand based virtual screening using SVM on GPU. *Comput Biol Chem, 83*, 107143. doi:10.1016/j.compbiolchem.2019.107143

Juen, Z., Lu, Z., Yu, R., Chang, A. N., Wang, B., Fitzpatrick, A. W. P., & Zuker, C. S. (2025). The structure of human sweetness. *Cell, 188*(15), 4141-4153 e4118. doi:10.1016/j.cell.2025.04.021

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., . . . Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583-589. doi:10.1038/s41586-021-03819-2

Kim, J., Ahmad, R., Deb-Choudhury, S., Subbaraj, A., Dalziel, J. E., & Knowles, S. O. (2022). Generation and identification of kokumi compounds and their validation by taste-receptor assay: An example with dry-cured lamb meat. *Food Chem X, 13*, 100218. doi:10.1016/j.fochx.2022.100218

Kitajima, S., Goda, R., Nitta, J., Maruyama, Y., Seki, T., & Kuroda, M. (2025). Volatile medium-chain lactones function as taste modifiers, kokumi substances, by activating the calcium-sensing receptor in the oral cavity. *Food Res Int, 220*, 117022. doi:10.1016/j.foodres.2025.117022

Kris-Etherton, P. M., Hecker, K. D., Bonanome, A., Coval, S. M., Binkoski, A. E., Hilpert, K. F., . . . Etherton, T. D. (2002). Bioactive compounds in foods: their role in the prevention of cardiovascular disease and cancer. *Am J Med, 113 Suppl 9B*, 71S-88S. doi:10.1016/s0002-9343(01)00995-0

Lafarga, T., & Hayes, M. (2014). Bioactive peptides from meat muscle and by-products: generation, functionality and application as functional ingredients. *Meat Sci, 98*(2), 227-239. doi:10.1016/j.meatsci.2014.05.036

Lammi, C., Boschin, G., Bollati, C., Arnoldi, A., Galaverna, G., & Dellafiora, L. (2021). A heuristic, computer-driven and top-down approach to identify novel bioactive peptides: A proof-of-principle on angiotensin I converting enzyme inhibitory peptides. *Food Res Int, 150*(Pt A), 110753. doi:10.1016/j.foodres.2021.110753

Li, H., Sze, K. H., Lu, G., & Ballester, P. J. (2020). Machine-learning scoring functions for structure-based virtual screening. *WIREs Computational Molecular Science, 11*(1). doi:10.1002/wcms.1478

Lopez Cascales, J. J., Oliveira Costa, S. D., de Groot, B. L., & Walters, D. E. (2010). Binding of glutamate to the umami receptor. *Biophys Chem, 152*(1-3), 139-144. doi:10.1016/j.bpc.2010.09.001

Luttens, A., Cabeza de Vaca, I., Sparring, L., Brea, J., Martinez, A. L., Kahlous, N. A., . . . Carlsson, J. (2025). Rapid traversal of vast chemical space using machine learning-guided docking screens. *Nat Comput Sci, 5*(4), 301-312. doi:10.1038/s43588-025-00777-x

Maia, E. H. B., Assis, L. C., de Oliveira, T. A., da Silva, A. M., & Taranto, A. G. (2020). Structure-Based Virtual Screening: From Classical to Artificial Intelligence. *Front Chem, 8*, 343. doi:10.3389/fchem.2020.00343

Malavolta, M., Pallante, L., Mavkov, B., Stojceski, F., Grasso, G., Korfiati, A., . . . Deriu, M. A. (2022). A survey on computational taste predictors. *Eur Food Res Technol, 248*(9), 2215-2235. doi:10.1007/s00217-022-04044-5

Margulis, E., Dagan-Wiener, A., Ives, R. S., Jaffari, S., Siems, K., & Niv, M. Y. (2021). Intense bitterness of molecules: Machine learning for expediting drug discovery. *Comput Struct Biotechnol J, 19*, 568-576. doi:10.1016/j.csbj.2020.12.030

Maroni, G., Pallante, L., Di Benedetto, G., Deriu, M. A., Piga, D., & Grasso, G. (2022). Informed classification of sweeteners/bitterants compounds via explainable machine learning. *Curr Res Food Sci, 5*, 2270-2280. doi:10.1016/j.crfs.2022.11.014

Mihalache, O. A., De Boevre, M., Dellafiora, L., De Saeger, S., Moretti, A., Pinson-Gadais, L., . . . Dall'Asta, C. (2023). The Occurrence of Non-Regulated Mycotoxins in Foods: A Systematic Review. *Toxins (Basel), 15*(9). doi:10.3390/toxins15090583

Missmer, S. A., Suarez, L., Felkner, M., Wang, E., Merrill, A. H., Jr., Rothman, K. J., & Hendricks, K. A. (2006). Exposure to fumonisins and the occurrence of neural tube defects along the Texas-Mexico border. *Environ Health Perspect, 114*(2), 237-241. doi:10.1289/ehp.8221

Mojsak, P., Lozowicka, B., & Kaczynski, P. (2018). Estimating acute and chronic exposure of children and adults to chlorpyrifos in fruit and vegetables based on the new, lower toxicology data. *Ecotoxicol Environ Saf, 159*, 182-189. doi:10.1016/j.ecoenv.2018.05.006

Morena, F., Cencini, C., Calzoni, E., Martino, S., & Emiliani, C. (2024). A Novel Workflow for In Silico Prediction of Bioactive Peptides: An Exploration of Solanum lycopersicum By-Products. *Biomolecules, 14*(8). doi:10.3390/biom14080930

Mostafa, A. A., Alhossary, A. A., Salem, S. A., & Mohamed, A. E. (2022). GBO-kNN a new framework for enhancing the performance of ligand-based virtual screening for drug discovery. *Expert Systems with Applications, 197*. doi:10.1016/j.eswa.2022.116723

Mufassirin, M. M. M., Newton, M. A. H., & Sattar, A. (2022). Artificial intelligence for template-free protein structure prediction: a comprehensive review. *Artificial Intelligence Review, 56*(8), 7665-7732. doi:10.1007/s10462-022-10350-x

Murtaza, M. A., Irfan, S., Hafiz, I., Ranjha, M., Rahaman, A., Murtaza, M. S., . . . Siddiqui, S. A. (2022). Conventional and Novel Technologies in the Production of Dairy Bioactive Peptides. *Front Nutr, 9*, 780151. doi:10.3389/fnut.2022.780151

Nelson, G., Hoon, M. A., Chandrashekar, J., Zhang, Y. F., Ryba, N. J. P., & Zuker, C. S. (2001). Mammalian sweet taste receptors. *Cell, 106*(3), 381-390. doi:Doi 10.1016/S0092-8674(01)00451-2

Ohno, S., Manabe, N., & Yamaguchi, Y. (2024). Prediction of protein structure and AI. *J Hum Genet, 69*(10), 477-480. doi:10.1038/s10038-023-01215-4

Pal, S., Kumar, V., Kundu, B., Bhattacharya, D., Preethy, N., Reddy, M. P., & Talukdar, A. (2019). Ligand-based Pharmacophore Modeling, Virtual Screening and Molecular Docking Studies for Discovery of Potential Topoisomerase I Inhibitors. *Comput Struct Biotechnol J, 17*, 291-310. doi:10.1016/j.csbj.2019.02.006

Palomares, I., Martinez-Camara, E., Montes, R., Garcia-Moral, P., Chiachio, M., Chiachio, J., . . . Herrera, F. (2021). A panoramic view and swot analysis of artificial intelligence for achieving the sustainable development goals by 2030: progress and prospects. *Appl Intell (Dordr), 51*(9), 6497-6527. doi:10.1007/s10489-021-02264-y

Patil, B. S., Jayaprakasha, G. K., Chidambara Murthy, K. N., & Vikram, A. (2009). Bioactive compounds: historical perspectives, opportunities, and challenges. *J Agric Food Chem, 57*(18), 8142-8160. doi:10.1021/jf9000132

Pedroni, L., Perugino, F., Dall'Asta, C., Galaverna, G., Buratti, F. M., Testai, E., & Dellafiora, L. (2024). Computational journey to unveil organophosphorothioate pesticides' metabolism: A focus on chlorpyrifos and CYP2C19 mutational landscape. *Ecotoxicol Environ Saf, 288*, 117354. doi:10.1016/j.ecoenv.2024.117354

Pedroni, L., Perugino, F., Galaverna, G., Dall'Asta, C., & Dellafiora, L. (2022). An In Silico Framework to Mine Bioactive Peptides from Annotated Proteomes: A Case Study on Pancreatic Alpha Amylase Inhibitory Peptides from Algae and Cyanobacteria. *Nutrients, 14*(21). doi:10.3390/nu14214680

Pedroni, L., Perugino, F., Magnaghi, F., Dall'Asta, C., Galaverna, G., & Dellafiora, L. (2024). Free fatty acid receptors beyond fatty acids: A computational journey to explore peptides as possible binders of GPR120. *Curr Res Food Sci, 8*, 100710. doi:10.1016/j.crfs.2024.100710

Peng, J., & Xu, J. (2011). RaptorX: exploiting structure information for protein alignment by statistical inference. *Proteins, 79 Suppl 10*(Suppl 10), 161-171. doi:10.1002/prot.23175

Pepe, G., Sommella, E., Ventre, G., Scala, M. C., Adesso, S., Ostacolo, C., . . . Campiglia, P. (2016). Antioxidant peptides released from gastrointestinal digestion of "Stracchino" soft cheese: Characterization, in vitro intestinal protection and bioavailability. *Journal of Functional Foods, 26*, 494-505. doi:10.1016/j.jff.2016.08.021

Perenzoni, D., Dellafiora, L., Perugino, F., Vrhovsek, U., Piombino, P., Pittari, E., . . . Mattivi, F. (2024). Exploring Putative Kokumi Oligopeptides in Classic Sparkling Wines with a UHPLC-ESI-MS/MS Targeted Protocol. *J Agric Food Chem, 72*(47), 26189-26208. doi:10.1021/acs.jafc.4c08213

Power, O., Nongonierma, A. B., Jakeman, P., & FitzGerald, R. J. (2014). Food protein hydrolysates as a source of dipeptidyl peptidase IV inhibitory peptides for the management of type 2 diabetes. *Proc Nutr Soc, 73*(1), 34-46. doi:10.1017/S0029665113003601

Regulation (EC) No 1924/2006 of the European Parliament and of the Council of 20 December 2006 on nutrition and health claims made on foods, (2006).

Riniker, S., & Landrum, G. A. (2013). Open-source platform to benchmark fingerprints for ligand-based virtual screening. *J Cheminform, 5*(1), 26. doi:10.1186/1758-2946-5-26

Sánchez, A., & Vázquez, A. (2017). Bioactive peptides: A review. *Food Quality and Safety, 1*(1), 29-46. doi:10.1093/fqsafe/fyx006

Santiago, R., Cao, A., & Butron, A. (2015). Genetic Factors Involved in Fumonisin Accumulation in Maize Kernels and Their Implications in Maize Agronomic Management and Breeding. *Toxins (Basel), 7*(8), 3267-3296. doi:10.3390/toxins7083267

Scott, P. M. (2012). Recent research on fumonisins: a review. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess, 29*(2), 242-248. doi:10.1080/19440049.2010.546000

Seefelder, W., Humpf, H. U., Schwerdt, G., Freudinger, R., & Gekle, M. (2003). Induction of apoptosis in cultured human proximal tubule cells by fumonisins and fumonisin metabolites. *Toxicol Appl Pharmacol, 192*(2), 146-153. doi:10.1016/s0041-008x(03)00262-x

Segura-Campos, M., Chel-Guerrero, L., Betancur-Ancona, D., & Hernandez-Escalante, V. M. (2011). Bioavailability of Bioactive Peptides. *Food Reviews International, 27*(3), 213-226. doi:10.1080/87559129.2011.563395

Skjanes, K., Aesoy, R., Herfindal, L., & Skomedal, H. (2021). Bioactive peptides from microalgae: Focus on anti-cancer and immunomodulating activity. *Physiol Plant, 173*(2), 612-623. doi:10.1111/ppl.13472

Smith, K. H., Tejeda-Montes, E., Poch, M., & Mata, A. (2011). Integrating top-down and self-assembly in the fabrication of peptide and protein-based biomedical materials. *Chem Soc Rev, 40*(9), 4563-4577. doi:10.1039/c1cs15064b

Srivastava, P., Steuer, A., Ferri, F., Nicoli, A., Schultz, K., Bej, S., . . . Wolkenhauer, O. (2024). Bitter peptide prediction using graph neural networks. *J Cheminform, 16*(1), 111. doi:10.1186/s13321-024-00909-x

Sulaiman, N. S., Rovina, K., & Joseph, V. M. (2019). Classification, extraction and current analytical approaches for detection of pesticides in various food products. *Journal of Consumer Protection and Food Safety, 14*(3), 209-221. doi:10.1007/s00003-019-01242-4

Tao, L., Wang, D., Yuan, Q., Zhao, F., Zhang, Y., Du, T., . . . Duan, J. (2024). Bitter taste receptor TAS2R14 activation and G protein assembly by an intracellular agonist. *Cell Res, 34*(10), 735-738. doi:10.1038/s41422-024-00995-4

Tola, M., Kebede, B., & Yildiz, F. (2016). Occurrence, importance and control of mycotoxins: A review. *Cogent Food & Agriculture, 2*(1). doi:10.1080/23311932.2016.1191103

Tseng, Y. J., Chuang, P. J., & Appell, M. (2023). When Machine Learning and Deep Learning Come to the Big Data in Food Chemistry. *ACS Omega, 8*(18), 15854-15864. doi:10.1021/acsomega.2c07722

Tuwani, R., Wadhwa, S., & Bagler, G. (2019). BitterSweet: Building machine learning models for predicting the bitter and sweet taste of small molecules. *Sci Rep, 9*(1), 7155. doi:10.1038/s41598-019-43664-y

van der Merwe, K. J., Steyn, P. S., Fourie, L., Scott, D. B., & Theron, J. J. (1965). Ochratoxin A, a toxic metabolite produced by Aspergillus ochraceus Wilh. *Nature, 205*(976), 1112-1113. doi:10.1038/2051112a0

Vasilaki, A., Panagiotopoulou, E., Koupantsis, T., Katsanidis, E., & Mourtzinos, I. (2022). Recent insights in flavor-enhancers: Definition, mechanism of action, taste-enhancing ingredients, analytical techniques and the potential of utilization. *Crit Rev Food Sci Nutr, 62*(32), 9036-9052. doi:10.1080/10408398.2021.1939264

Vero, V., & Gasbarrini, A. (2012). The EFSA health claims 'learning experience'. *Int J Food Sci Nutr, 63 Suppl 1*, 14-16. doi:10.3109/09637486.2011.633899

Vettorazzi, A., Lopez de Cerain, A., Sanz-Serrano, J., Gil, A. G., & Azqueta, A. (2020). European Regulatory Framework and Safety Assessment of Food-Related Bioactive Compounds. *Nutrients, 12*(3). doi:10.3390/nu12030613

Wang, W., Cui, Z., Ning, M., Zhou, T., & Liu, Y. (2022). In-silico investigation of umami peptides with receptor T1R1/T1R3 for the discovering potential targets: A combined modeling approach. *Biomaterials, 281*, 121338. doi:10.1016/j.biomaterials.2021.121338

Westmoreland, C., Bender, H. J., Doe, J. E., Jacobs, M. N., Kass, G. E. N., Madia, F., . . . Cronin, M. T. D. (2022). Use of New Approach Methodologies (NAMs) in regulatory decisions for chemical safety: Report from an EPAA Deep Dive Workshop. *Regul Toxicol Pharmacol, 135*, 105261. doi:10.1016/j.yrtph.2022.105261

Wojcikowski, M., Ballester, P. J., & Siedlecki, P. (2017). Performance of machine-learning scoring functions in structure-based virtual screening. *Sci Rep, 7*, 46710. doi:10.1038/srep46710

Xu, W., Wu, L., Liu, S., Liu, X., Cao, X., Zhou, C., . . . Hua, T. (2022). Structural basis for strychnine activation of human bitter taste receptor TAS2R46. *Science, 377*(6612), 1298-1304. doi:10.1126/science.abo1633

Yan, L., Ke, Y., Kan, Y., Lin, D., Yang, J., He, Y., & Wu, L. (2020). New insight into enzymatic hydrolysis of peptides with site-specific amino acid d-isomerization. *Bioorg Chem, 105*, 104389. doi:10.1016/j.bioorg.2020.104389

Yang, J., Anishchenko, I., Park, H., Peng, Z., Ovchinnikov, S., & Baker, D. (2020). Improved protein structure prediction using predicted interresidue orientations. *Proc Natl Acad Sci U S A, 117*(3), 1496-1503. doi:10.1073/pnas.1914677117

Zhan, Z. H., Hong, J., Li, J. Y., Wang, C., He, L. C., Xu, Z. B., & Zhang, J. (2025). Artificial intelligence-based methods for protein structure prediction: a survey. *Artificial Intelligence Review, 58*(10). doi:ARTN 328

10.1007/s10462-025-11325-4

Zheng, W., Feng, N., Wang, Y., Noll, L., Xu, S., Liu, X., . . . Liu, Z. (2019). Effects of zearalenone and its derivatives on the synthesis and secretion of mammalian sex steroid hormones: A review. *Food Chem Toxicol, 126*, 262-276. doi:10.1016/j.fct.2019.02.031

Zhou, G., Rusnac, D. V., Park, H., Canzani, D., Nguyen, H. M., Stewart, L., . . . DiMaio, F. (2024). An artificial intelligence accelerated virtual screening platform for drug discovery. *Nat Commun, 15*(1), 7761. doi:10.1038/s41467-024-52061-7

Zhu, Z., Zhang, W., Li, Z., Zhao, W., Liu, C., Zhu, B., . . . Yang, Q. (2024). Rethinking Sweetener Discovering: Multiparameter Modeling of Molecular Docking Results

between the T1R2-T1R3 Receptor and Compounds with Different Tastes. *J Agric Food Chem, 72*(13), 7336-7343. doi:10.1021/acs.jafc.4c00407

Ziaikin, E., David, M., Uspenskaya, S., & Niv, M. Y. (2025). BitterDB: 2024 update on bitter ligands and taste receptors. *Nucleic Acids Res, 53*(D1), D1645-D1650. doi:10.1093/nar/gkae1044

Zimmer, I., Usleber, E., Klaffke, H., Weber, R., Majerus, P., Otteneder, H., . . . Martlbauer, E. (2008). Fumonisin intake of the German consumer. *Mycotoxin Res, 24*(1), 40-52. doi:10.1007/BF02985269

The bitter side of toxicity: a big data analysis spotted the interaction between trichothecenes and bitter receptors

This chapter is based on the paper published on:

Lorenzo Pedroni, Florinda Perugino, Ambra Kurtaga, Gianni Galaverna, Chiara Dall'Asta, Luca Dellafiora. The bitter side of toxicity: a big data analysis spotted the interaction between trichothecenes and bitter receptors. (2023). Food Research International, 173, 113284, https://doi.org/10.1016/j.foodres.2023.113284.

Author contribution: Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing.

Chapter II

Abstract

The bitter taste perception evolved in human and animals to rapidly perceive and avoid potential toxic compounds. This is mediated by taste receptors type 2 (TAS2R), expressed in various tissues, which recently proved to be involved in roles beyond the bitter perception itself. With this study, the interaction between food-related toxic compounds and TAS2R46 has been thoroughly investigated via computational approaches, starting with a virtual screening and moving to molecular docking and dynamics simulations. The virtual screening analysis identified trichothecolone and the trichothecenes class it belongs to, which includes mycotoxins widespread in several commodities raising food safety concerns, as possible TAS2R46 binders. Molecular docking and dynamics simulations were performed to further explore the trichotecenes-TAS2R46 interaction. The results indicated that deoxynivalenol and its 15-acetylated derivative could activate TAS2R46. Eventually, this study provided initial evidence supporting the involvement of TAS2R46 in the underpinning mechanisms of deoxynivalenol action highlighting the need of digging into the involvement of TAS2R46 and TAS2Rs in the adverse effects of deoxynivalenol and congeners.

Chapter II

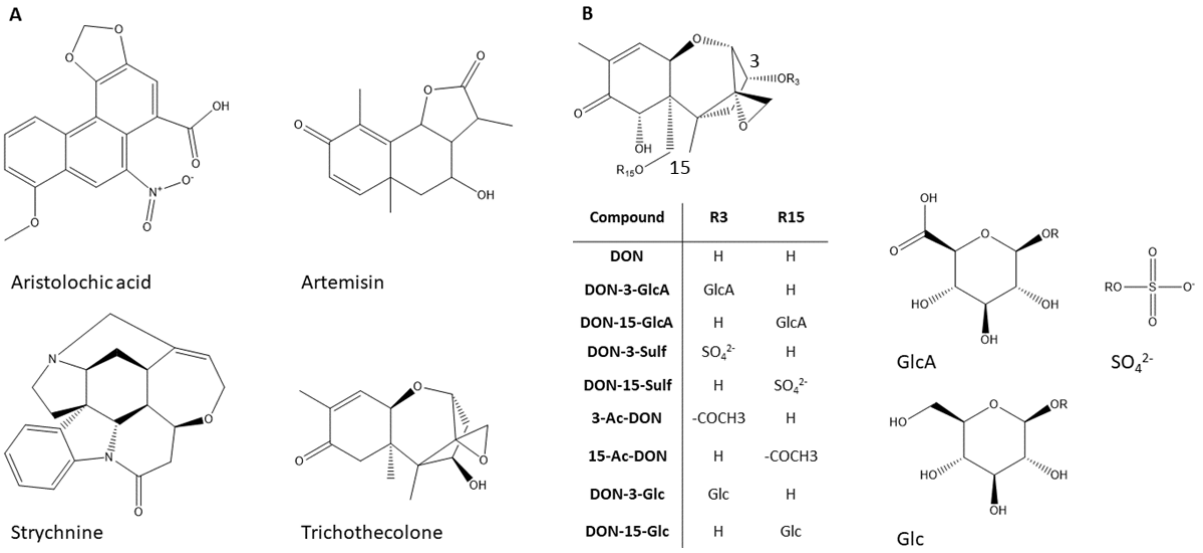
1. Introduction

The capability of certain animals to taste toxic compounds as bitter has been achieved far back along the evolution as a rapid alert system rather successful to avoid the ingestion of noxious compounds (Wooding, Ramirez, & Behrens, 2021). From a mechanistic point of view, the initiating event leading to the perception of bitter taste relies on the interaction between bitter compounds and a series of receptors transducing the signal to the central nervous system. Such receptors include the monomeric G-Protein Coupled Receptor (GPCR) Taste Receptors Type 2 (TAS2Rs) that are broadly expressed in animals, though the number of functional paralogues varies among species (Li & Zhang, 2014). Germane to humans, about 25 functional TAS2Rs are expressed in many tissues besides the lingual epithelium (Q. L. Wang, Liszt, & Depoortere, 2020). The expression of TAS2Rs in non-taste sensing districts was deemed ectopic for a long time, though specialized, testing-unrelated and tissue-dependent roles have been described in recent years (Lu, Zhang, Lifshitz, & ZhuGe, 2017). For this reason, TAS2Rs have gained attention as possible druggable targets to pharmacologically treat a variety of disorders. Evidence pointing to their involvement in certain food-related toxic outcomes have been collected as well (Ekstrand, Young, & Rasmussen, 2017; Kamila & Agnieszka, 2021; Wooding et al., 2021). From a mechanistical standpoint, the interaction between bitter compounds and the extracellular side of TAS2Rs activates the transmembrane receptor and triggers a cascade of conformational changes in the protein structure that reverberate in the cytosolic portion activating the associated G protein. Considering the involvement of TAS2Rs beyond taste perception, the functional relationship between TAS2R and toxic molecules is most likely beyond sensing potentially noxious compounds as bitter tasting and reasonably takes a part in their underpinning mechanisms of action. Therefore, this work describes a computer-driven and knowledge-based big data analysis to study the possible interaction of toxic compounds with TAS2R46 to probe its eventual involvement in their adverse outcome pathways. The study focused specifically on human TAS2R46 since was the unique TAS2R structurally described at the time of analysis with its cryo-EM structure (Xu et al., 2022) recently made available in the Protein Data Bank (PDB; https://www.rcsb.org) (Berman et al., 2000). Indeed, Xu and co-workers described precisely for the first time the topology of a human TAS2R in the active ligand-bound conformation (PDB ID 7XP6). This provided a breakthrough in the mechanistical understanding of TAS2R activation and granted the background for a fruitful application of molecular modelling techniques. Moreover, TAS2R46 is expressed in many tissues other than the lingual epithelium, as per Human Protein Atlas database (https://www.proteinatlas.org, last

access 8th May 2023) (Uhlen et al., 2015), possibly allowing to study effects beyond the bitter taste perception.

The work integrated virtual screening procedures and 3D molecular modelling as this previously succeeded to unveil unexpected mechanisms of action for toxic compounds (Del Favero et al., 2022). Briefly, an AI-based virtual screening method via PyRMD (Amendola & Cosconati, 2021) was used to mine potentially bitter compounds out of a virtual library of toxic molecules derived from T3DB database (http://www.t3db.ca; more than 3 500 toxic molecules, last access on 17th of April) (Lim et al., 2010) with the support of the VirtualTaste Prediction tool (https://virtualtaste.charite.de/VirtualTaste) (Fritz, Preissner, & Banerjee, 2021). This analysis described trichothecolone, a fungal metabolite primarily produced by *Trichotecium* species (Konishi et al., 2003), as a possible binder of TAS2R46. Trichothecolone is a member of the type B trichothecenes class, which includes mycotoxins of utmost importance for food safety. In particular *Fusarium* type B trichothecenes group some of the most prevalent and harmful mycotoxins widespread in the food and feed production chain worldwide and mostly abundant in cereal and cereal-based commodities (Khaneghah, Farhadi, Nematollahi, Vasseghian, & Fakhri, 2020; Tolosa, Rodriguez-Carrasco, Ruiz, & Vila-Donat, 2021). Although trichothecolone is barely reported in food and feed and is not yet regulated, its predicted relevance to TAS2R46 raised the question whether other type B trichothecenes relevant to food safety might target the same bitter receptor. This information would unveil their mechanism of actions eventually improving their risk assessment, and more in general their toxicological understanding. In this respect, recent advances in understanding the toxicity of trichothecenes have described mechanisms which might be consistent with the extra-oral involvement of bitter receptors (Flannery, Clark, & Pestka, 2012; Jia et al., 2022; Q. L. Wang et al., 2020). However, to the best of our knowledge, direct evidence is still missing. Therefore, the interaction with TAS2R46 and trichothecolone and a series of type B trichothecenes relevant to food safety was studied through docking studies and molecular dynamics to gain a thorough analysis from a molecular viewing angle. Specifically, the set under analysis included deoxynivalenol (DON), which was taken as reference for type B trichothecenes germane to food safety, and a selection of DON precursors, i.e. 3- and 15-acetyl DON (mainly produced by specific *F. graminearum* chemotypes), and metabolites either found in food as plant modified forms, i.e. DON-3- and -15-glucoside, or produced upon DON ingestion, i.e. DON-3- and -15-glucuronide, or produced by both plants and animals metabolism, i.e. DON-3- and -15-sulfate (Warth et al., 2015) (Figure 1). This integrated analysis aimed to distinguish TAS2R46's binders from non-binders, in agreement with previous studies (Del Favero et al., 2022; Dellafiora, Magnaghi, Galaverna, &

Dall'Asta, 2022). In brief, DON and 15-acetylated DON were predicted likely to activate TAS2R46 providing for the first time a compelling line of evidence pointing to: i) its involvement in the mechanics of DON action; ii) the need of further investing the role of TAS2Rs in the adverse outcome pathway of DON and congeners.



Compound	R3	R15
DON	H	H
DON-3-GlcA	GlcA	H
DON-15-GlcA	H	GlcA
DON-3-Sulf	SO_4^{2-}	H
DON-15-Sulf	H	SO_4^{2-}
3-Ac-DON	-COCH3	H
15-Ac-DON	H	-COCH3
DON-3-Glc	Glc	H
DON-15-Glc	H	Glc

Figure 1. Chemical structure of compounds analysed via molecular modelling. **A.** Aristolochic acid (CID 2236), artemisin (CID 65030), strychnine (CID 441071) and trichothecolone (CID 107974). **B.** DON scaffold with its chemical substituent in position 3 (R3) and 15 (R15). **2.**

Material and methods

2.1 Data source

2.1.1. Small molecules structures

The chemical structures to instruct and perform the virtual screening procedure were

2022) in the Structured Data File (.sdf) format: strychnine (CID 441071), artemisin (CID 65030), deoxynivalenol (DON; CID 40024), DON-3-Glucoside (DON-3-Glc; CID 71312510), DON-3-Glucoronide (DON-3-GlcA; CID 102202100), DON-15-GlcA (CID 102202102), DON-3-Sulfate (DON-3-Sulf; CID 101899454), DON-15-Sulf (CID 101899453) 3-Acetyl-DON (3-Ac-DON; CID 92043653), 15-Ac-DON (CID 10382483), aristolochic acid (CID 2236), trichothecolone (CID 377304). DON-15-Glc was not available in PubChem as of 17th April 2023 and its 3D structure was derived manually editing the structure of DON-15-GlcA

substituting the carboxylic acid with a hydroxylic group through UCSF Chimera (version 1.16) (Pettersen et al., 2004).

2.1.2 Protein structure

The model of Taste Receptors Type 2 Member 46 (TAS2R46) was derived from the strychnine-bound cryo-EM structure deposited in the Protein Data Bank (PDB; https://www.rcsb.org) with PDB code 7XP6 (Xu et al., 2022). Specifically, the model was derived from the chain E after completing its non-terminal missing parts (residues 157-172) through Modeller (version 10.2) (Sali & Blundell, 1993) implemented in UCSF Chimera (version 1.16) (Pettersen et al., 2004), using the loop/refinement tool with the standard loop modelling protocol and setting at two the number of residues adjacent to missing regions allowed to move, in agreement with previous studies (Dellafiora et al., 2020).

2.2 Virtual screening

The virtual screening was performed using PyRMD (https://github.com/cosconatilab/PyRMD), a fully automated AI-powered ligand-based virtual screening tool (Amendola & Cosconati, 2021). The data retrieved from the FlavorDB database were curated to properly instruct PyRMD. In particular, the 25 595 entries retrieved from FlavorDB in the JSON format were split to obtain two datasets, one made of known bitter molecules (true positive dataset) and the other made of known not bitter molecules (true negative dataset). To this end, compounds classified in FlavorDB as “bitter”, or “phenolic bitter”, or “odourless bitter” were gathered into the true positive dataset (465 molecules), while compounds classified as not exclusively bitter, entries made of multiple chemicals (i.e., coordination complexes or salts), or whole inorganic molecules were excluded from the analysis to avoid ambiguity in the bitterness assignment. Then, the set of compounds classified as not bitter was further curated to provide a useful and representative true negative dataset including only molecules with a molecular weight (MW) comparable to that of true positives, using their mean MW ± SD as reference value (i.e. 302.3 ± 128.9 g/mol). Therefore, not bitter molecules with a MW ranging from 173.3 to 431.3 g/mol were gathered into the true negative dataset (11 982). These two datasets served as input to train and benchmark PyRMD for the sake of binarily classifying molecules from T3DB (http://www.t3db.ca/) (Lim et al., 2010) as either bitter or not bitter. Moreover, the Prediction tool of VirtualTaste (https://virtualtaste.charite.de/VirtualTaste) (Fritz et al., 2021), a webserver for the prediction of organoleptic properties of molecules, was used to identify possible

TAS2R46 interactors among the T3DB molecules predicted as bitter by PyRMD. Only the predictions with a level of confidence above 0.7 were considered.

2.3 Docking simulation

Docking simulations were performed to obtain a plausible binding architecture for each molecule under analysis within the TAS2R46 binding site. The GOLD software (v. 2021) was used as it already proved a high reliability to investigate the protein-ligand interaction (Dellafiora et al., 2020; Maldonado-Rojas & Olivero-Verbel, 2011). The binding site was set as a 10 Å radius sphere around the centroid of the cryo-EM structure strychnine-bound ligand binding site (PDB code 7XP6; 2.1.2). A docking protocol with ligands fully flexible and protein semi-flexible, allowing polar hydrogens to freely rotate, was adopted in agreement with previous studies (Dellafiora et al., 2022). We used the internal GOLDScore scoring function, which was optimised to predict ligand-binding positions from manufacturer declarations (the higher the score, the more likely is the precited binding architecture; https://www.ccdc.cam.ac.uk; as of 17th April 2023).

The strychnine-TAS2R46 structural complex (PDB code 7XP6) described the importance of π-π interaction with Trp88 and polar contacts with Glu265. Therefore, two pharmacophore constraints, "Aromatic" and "H-bond donor" (constraint weight set at 30 units and with a radius of 0.7 Å), were placed in correspondence of the centroid of strychnine's aromatic ring and the quinolinic nitrogen, respectively, to promote a knowledge-based positioning of ligands. For DON and congeners, the trichotecholone's best scored pose was used to set a "Shape overlap" constraint instead (constraint weight of 50 units) to promote their consistent arrangement into the ligand binding site. The best scored pose for each ligand was further analysed through molecular dynamics simulations (see section 2.4). The Root Mean Squared Deviation (RMSD) analysis between the cryo-EM structure and docking pose (excluding hydrogens) of strychnine was performed with DockRMSD webserver (version 1.1; https://zhanggroup.org/DockRMSD) (Bell & Zhang, 2019).

2.4 Molecular dynamics simulation

Molecular dynamics (MD) simulations were performed through GROMACS (version 2019.4) (Abraham et al., 2015) to verify the ligands potential to activate the TAS2R46. The whole system was parametrised with the CHARMM27 all-atom force field (Best et al., 2012) and specifically each ligand parametrisation was performed on the SwissParam webserver

(https://www.swissparam.ch/) (Zoete, Cuendet, Grosdidier, & Michielin, 2011). Input structures were solvated with SPCE waters in a cubic periodic boundary condition. The system was neutralized adding Na^+ and Cl^- as counter ions. Prior to the molecular dynamics production, each system was energetically minimised to both avoid steric clashes and correct improper geometries using the steepest descent algorithm with a maximum of 5000 steps. Each system underwent isothermal (300 K, coupling time 2 ps) and isobaric (1 bar, coupling time 2 ps) 100 ps simulations. Finally, 40 ns long molecular dynamics simulations were run (300 K with a coupling time of 0.1 ps and 1 bar with a coupling time of 2.0 ps).

2.5 Statistical analysis

The statistical analysis to compare the mean Root Mean Squared Deviations (RMSDs) of ligands and protein in each complex analysed has been done with SPSS IBM (v. 27.0, SPSS Inc., Chicago, IL, USA). For each complex, distance values of 8000 frames have been considered, expressed as means ± standard deviation (SD) and compared to each other using one-way ANOVA ($\alpha = 0.05$) with Bonferroni as post hoc test ($\alpha = 0.05$).

3. Results and Discussion

3.1 Virtual screening and case study selection

All the 25 595 molecules stored in the FlavorDB database (last database access 17th April 2023, https://cosylab.iiitd.edu.in/flavordb/) (Garg et al., 2018) were downloaded in the JSON format (.json files) and properly curated for the sake of this study. FlavorDB was meant to provide a consistent set of molecules to train PyRMD, an AI-powered tool for ligand-based virtual screening (https://github.com/cosconatilab/PyRMD) (Amendola & Cosconati, 2021), to mine compounds able to interact with TAS2R46 out of a library of well-known toxicants (taken from T3DB database; http://www.t3db.ca; 3 533 entries as of 17th April 2023) (Lim et al., 2010). The FlavorDB set was curated as per section 2.1 using an *ad hoc* python script (available upon request) to systematically sort bitter from non-bitter compounds, providing a set of "true positive" bitter compounds (465 molecules), and representative set of "true negative" non-bitter compounds (11 982 molecules). The PubChem CID of molecules included in the two lists is reported in Supplementary material, Table S1.

The performance of PyRMD's model, built using the two set of compounds mentioned above, was cross validated through a benchmark procedure. This procedure employed a repeated

stratified k-fold approach with 5 folds and 3 repetitions, with epsilon, beta and alpha parameters set at 0.95, 1 and 0.2, respectively (default setting). The generated model had a true positive rate (TPR) of 76.3 %, a false positive rate (FPR) of 17.4 % and a receiver operating characteristic (ROC) area under the curve (AUC) score of 0.84. Given the quite high TPR and the excellent-for-discrimination ROC AUC score (Hosmer, Lemeshow, & Sturdivant, 2013), the model was considered effective, and it was used to screen the 3 533 toxic molecules included in the T3DB database (last database access 17th of April 2023, http://www.t3db.ca) (Lim et al., 2010). The model classified 1 085 T3DB entries as bitter (PubChem CIDs and T3DB IDs are reported in Supplementary material; Table S2). This selection was further refined by submitting these molecules to the prediction tool of VirtualTaste webserver (https://virtualtaste.charite.de/VirtualTaste/) (Fritz et al., 2021). This webserver has a built-in bitterness prediction model which assigns a confidence score (0 to 1) to the analysed molecules (the closer to 1 the more probable the bitterness of the molecule), along with an estimate of the probability of each entry to interact with specific TAS2Rs. In this study, the confidence threshold was set at 0.7 (to extrapolate molecules with a high potential of being bitter) and the analysis returned a short list of 22 molecules likely interacting with TAS2R46 (Table 1), chosen as a case study being the unique TAS2R with an available experimentally defined structure at the time of analysis. Of note, the high prevalence of compounds already described as bitter (16 out of 22), including a well-known binder of a TAS2R46 binder (quinine), eventually validated the model efficacy to mine compounds able to interact with TAS2Rs. Among those with no evidence of bitterness, the trichothecene trichothecolone (CID 107974, T3DB ID T3D3719) drew the attention for the high relevance of type B trichothecene class to food safety. In more detail, despite trichothecolone is of minor importance for food safety, it served as a bait to focus the analysis over other type B trichothecenes more relevant to food safety and related risk assessment. As a proof of principle, based on the structural similarities to trichothecolone, the analysis covered DON and a series of congeners including fungal, human and plant metabolites being relevant either for food and feed contamination due to their toxicity or for their presence in living organisms upon DON metabolism (Figure 1) (Sun et al., 2022). (Gallo, Ferrara, Calogero, Montesano, & Naviglio, 2020; Khaneghah et al., 2020; Tolosa et al., 2021).

With respect to DON toxicity, depending on the level and frequency of exposure, DON may either cause acute poisoning or chronic toxicity including but not limited to cytotoxicity, immunotoxicity, neurotoxicity, reproductive toxicity, and even carcinogenicity (Hou et al., 2023). Moreover, DON, also commonly known as vomitoxin, has already been reported to adversely affect feed intake in animals such as pigs, mice, cats and dogs also inducing anorexia

(Flannery et al., 2012; Holanda & Kim, 2021; Patience, Myers, Ensley, Jacobs, & Madson, 2014). Several mechanisms have been proposed to explain this phenomenon including but not limited to the modulation of local serotonin, emetic effect, release of proinflammatory cytokines and enhancement of satiety signal with the release of gut satiety peptide YY and GLP-1) (Flannery et al., 2012; Holanda & Kim, 2021). Moreover, although there are no clear and unique connections between the decrease of food intake and TAS2Rs activation, it has been demonstrated that the release of peptide YY and GLP-1, which are involved in the mechanisms of food acceptance, is in part under the control of TAS2Rs (Flannery et al., 2012; Jia et al., 2022; Q. L. Wang et al., 2020). Further analyses are needed to provide final proofs, however a possible functional correlation between the DON-dependent food rejection and its capability to activate TAS2R46 is meaningful, also considered the expression of TAS2R46 in enteroendocrine cells which are inherently exposed to DON (K. S. Kim, Egan, & Jang, 2014). On this basis, a relation between TAS2R46-DON interaction and decreased feed intake is likely to occur and worth to be analysed in further dedicated investigations. In this context, the present work assessed the capability of DON and a selection of congeners to interact with TAS2R46 being structurally similar to trichothecolone and therefore likely interacting with similar protein targets. In more detail, DON may co-occur in contaminated commodities along with a series of modified forms, such as DON glucosides and acetylated forms, which may be produced by mycotoxin producing moulds or by plant metabolism. In addition, extensive phase II metabolism has been described in humans with DON sulphates and glucuronides among the most abundant forms excreted via urines. Therefore, DON and a selection of derivatives representative of the most abundant metabolites found in food or produced by human metabolism underwent a molecular modelling analysis to assess: i) whether DON may interact with TAS2R46; ii) the effects of DON's metabolic modifications on its possible capability to interact with TAS2R46. The list of DON congeners analysed in this study included DON-3-GlcA, DON-15-GlcA, DON-3-Sulf, DON-15-Sulf, 3-Ac-DON, 15-Ac-DON, DON-3-Glc and DON-15-Glc (available CIDs in section 2.1.1).

Table 1. Molecules identified as likely bitter and interacting with TAS2R46 according to VirtualTaste [1]

Name	Pubchem CID	VirtualTaste Score [1]	Confidence	T3DB ID
4-Hydroxydecenal	6439956	0,776		T3D4193
4-Hydroxynonenal	5283344	0,73		T3D4180
Acetyl tributyl citrate	6505	0,913		T3D4871
Allethrin [2]	11442	0,796		T3D1024
Androstenedione	6128	0,769		T3D4240
Brompheniramine [3]	6834	0,988		T3D4554
Cinerin I [2]	5281547	0,789		T3D1853
Cinerin II [2]	5281548	0,765		T3D1854
Coriamyrtin [4]	433737	0,92		T3D4074
Fexofenadine [5]	3348	0,908		T3D2938
Gelsemine	5390854	0,926		T3D4071
Hypoxanthine [4]	135398638	0,886		T3D4150
Jasmolin I [2]	12304687	0,794		T3D1855
Jasmolin II [2]	12304690	0,773		T3D1856
Pyrethrin I [2]	5281045	0,796		T3D1857
Pyrethrin II [2]	5281555	0,774		T3D1858
Pyrethrum [2]	71310221	0,794		T3D0233
Quinine [4,6,*]	3034034	0,996		T3D2800
S-Bioallethrin [2]	62829	0,796		T3D3920
Trichothecolone	107974	0,713		T3D3719
Tutin [7]	75729	0,831		T3D3089
Xanthine [3]	1188	0,867		T3D4409

Note: [1] VirtualTaste webserver's bitter prediction tool (https://virtualtaste.charite.de/VirtualTaste) (Fritz et al., 2021); [2] Belonging to a class broadly described as bitter for animals according to Wismer (Wismer, 2016); [3] Previously described as bitter according to Rahman and co-workers (Rahman, Zidan, Berendt, & Khan, 2012); [4] Previously described as bitter according to BitterDB (https://bitterdb.agri.huji.ac.il; (Dagan-Wiener et al., 2019); [5] Previously described as bitter according to Suaress and co-workers (Suares & Hiray, 2015); [6] Previously described as bitter according to FlavourDB

(https://cosylab.iiitd.edu.in/flavordb) (Garg et al., 2018); [7] Previously described as bitter according to Easterfield and Aston (Easterfield & Aston, 1901); * Known binder for TAS2R46 according to BitterDB (https://bitterdb.agri.huji.ac.il) (Dagan-Wiener et al., 2019).

3.2 Molecular modelling to study the interaction of DON and metabolites with TAS2R46

An already consolidated molecular modelling procedure integrating docking simulations and molecular dynamics was used to study the interaction of DON and its metabolites based on previous evidence reporting the efficacy of such integrated approach to predict the interaction between low-molecular weight molecules and proteins, including taste receptors (Dellafiora et al., 2022; Dellafiora et al., 2020). However, a fit-for-purpose assessment of procedural performance was done beforehand, as detailed below.

3.2.1 Fit-for-purpose assessment of procedural performances

A fit-for-purpose and knowledge-based validation of procedural performance was done based on the structural information recently made available for the human TAS2R46, which described the topology of its ligand-bound active conformation (PDB code 7XP6) (Xu et al., 2022). The 3D model to perform docking studies and molecular dynamics was derived from the strychnine-bound active conformation (PDB code 7XP6) as detailed in section 2.1. The capability to produce reliable binding poses was assessed first comparing the docking pose of strychnine with its cryo-EM structure's binding architecture. The pose scored 78.4 units pointing to its favourable interaction with the receptor (see section 2.3 and Table S3, Supplementary material, for further details). As shown in Figure 2, the two poses were geometrically comparable with an RMSD of 0.76 Å. The docking pose showed also the key interactions reported by the cryo-EM structure, including the π-π interaction with Trp88 and polar contacts with Glu265 (Sun et al., 2022), supporting the model reliability to predict consistently TAS2R46-ligand binding architectures. Afterwards, it was assessed whether molecular dynamics could provide a geometrical rationale to distinguish TAS2R46's activators from inactive compounds, in agreement with previous studies (Dellafiora et al., 2022; Dellafiora et al., 2020). In this respect, the molecular dynamics simulation of TAS2R46-strychnine complex served as a reference since it describes the active topology of TASR46 (Xu et al., 2022). In more detail, it was inferred that molecular dynamics of TAS2R46-ligand complexes showing deviations from the trends traced when in complex with strychnine may indicate a topology not associated with TAS2R46

activation and/or likely resulting in ligand detachment (transient interaction) and complex dissociation. To test this assumption, artemisin and aristolochic acid were included as additional positive and negative control, respectively, as the former but not the latter proved to activate TAS2R46 (Brockhoff, Behrens, Massarotti, Appendino, & Meyerhof, 2007; Brockhoff, Behrens, Niv, & Meyerhof, 2010). They both underwent docking and molecular dynamics simulations along with strychnine, and the outcomes were compared to those of the latter. Both scored positive according to GOLDScore scoring function suggesting their theoretical capability to interact with TAS2R46 (Table S3; Supplementary material), although aristolochic acid adopted a less centred arrangement into the ligand binding site compared to artemisin and strychnine (Figure 3A). Molecular dynamics provided instead the geometrical criteria to effectively discriminate activators (i.e. strychnine and artemisin) from non-activators (i.e. aristolochic acid). Indeed, the analysis of protein and ligands trajectories and their respective RMSD fluctuations over time described the protein-ligand complexes with strychnine and artemisin as more stable than that with aristolochic acid. This could confirm the assumption mentioned above, and the capability of ligands to properly maintain the strychnine/artemisin-like active topology of TAS2R46-ligand complex over time could be used to discriminate TAS2R46 ‘s activators from transient-interactors and/or non-activators. In particular, the protein RMSD had lower and more stable trends when TAS2R46 was in complex with strychnine or artemisin compared to the complex with aristolochic acid (Figure 3B). This indicated that the latter made the structure to deviate from the active conformation described by Xu (Xu et al., 2022), in agreement with the inactivity reported for aristolochic acid (Brockhoff et al., 2010). This outcome was summarized averaging the RMSDs of protein in each complex (see Figure 3B) which were significantly lower for strychnine and artemisin than in complex with aristolochic acid ($p < 0.01$). Also considering ligands (Figure 3B), the RMSD analysis revealed that strychnine and artemisin had a more stable interaction than aristolochic acid, as reflected by the RMSD mean values of the first two compounds which were significantly lower ($p<0.05$) than that of the latter (see Figure 3B). Overall, this indicated a more stable interaction at the ligand-binding site for strychnine and artemisin compared to aristolochic acid.

Based on the above, the maintenance of a steady-state topology for protein and ligands (which are expected to be stably arranged into the ligand binding site over time) as described for strychnine and artemisin was set as a rationale to distinguish effective activators from non-activators among the list of DON and its metabolites.

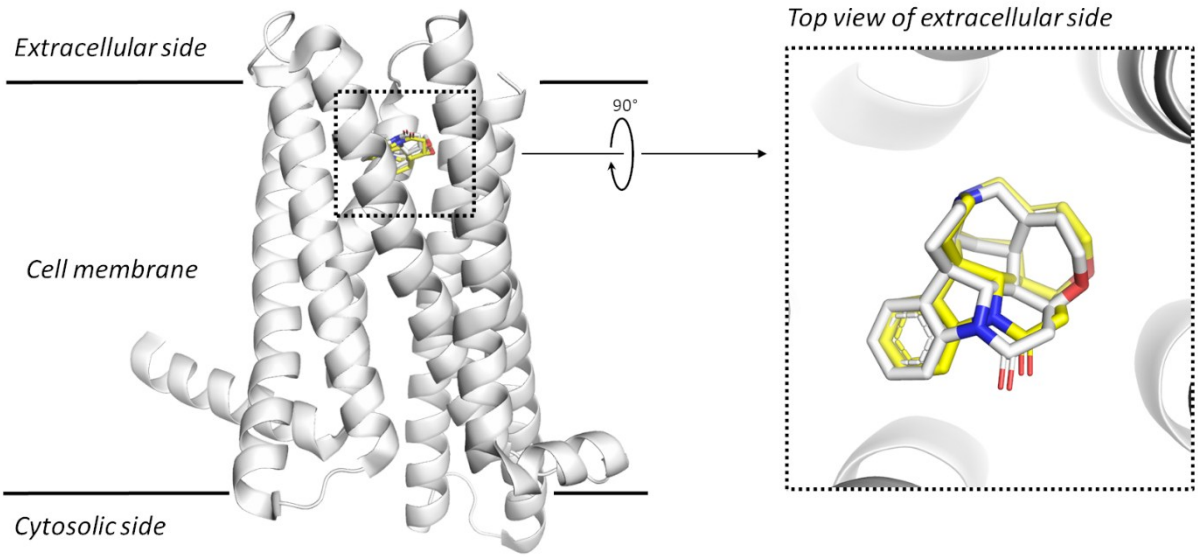


Figure 2. Comparison between cryo-EM structure (white coloured) and docking (yellow coloured) pose of strychnine within human TAS2R46. The protein is represented in cartoon, while strychnine in sticks.

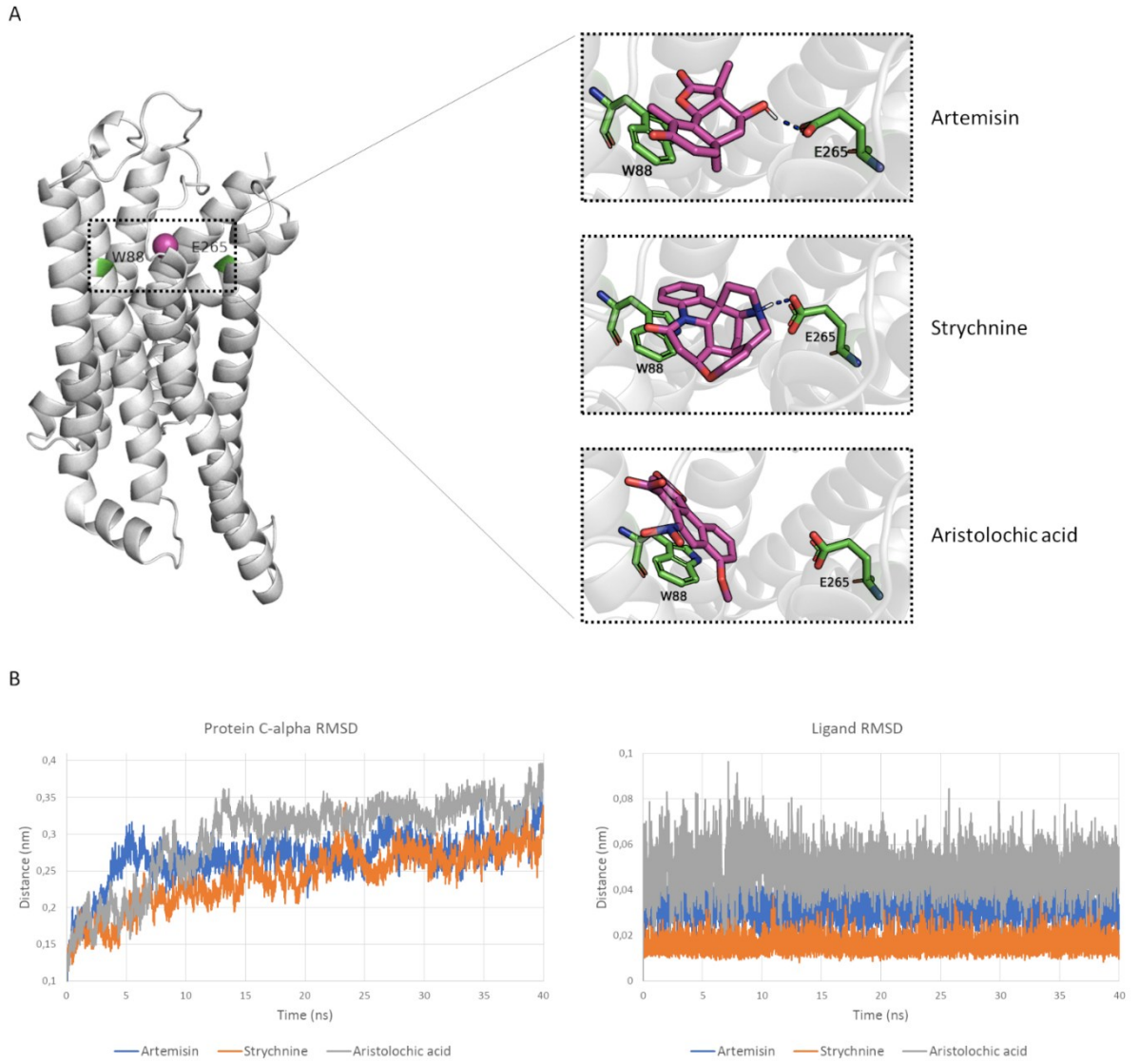


Figure 3. Molecular modelling results of artemisin, strychnine and aristolochic acid. **A.** Docking pose of artemisin, strychnine and aristolochic acid. On the left, TAS2R46 is represented in white cartoon with W88 and E265 shown in green stick and the ligand centroid as a magenta sphere. On the right, close-up of the ligand binding site with W88 and E265 represented as green sticks while ligands as magenta sticks. Blu dashed lines indicate polar contacts. **B**. On the left, protein C-alpha RMSD of TAS2R46 in complex with artemisin (average value 0.212 ± 0.039 nm), strychnine (average value 0.241 ± 0.041 nm) or aristolochic acid (average value 0.297 ± 0.057 nm). On the right, ligand RMSD of artemisin (average value 0.032 ± 0.005 nm), strychnine (average value 0.016 ± 0.004 nm) or aristolochic acid (average value 0.049 ± 0.009 nm).

3.2.2 Analysis of trichothecene mycotoxins interaction with TAS2R46

Once the identification of geometrical parameters to distinguish TAS2R46's activators from non-activators was achieved (see above), DON and a selection of its metabolites (i.e. DON-3-GlcA, DON-15-GlcA, DON-3-Sulf, DON-15-Sulf, 3-Ac-DON, 15-Ac-DON, DON-3-Glc and DON-15-Glc) underwent docking procedures and molecular dynamics. Trichothecolone was included as well since it was mined out the T3DB screening (see above), although its relevance to food safety is still largely overlooked and is substantially a minor problem for food safety compared to other type B trichothecenes. Conversely, DON is among the most prevalent mycotoxins contaminating food and feed commodities and it notoriously undergoes extensive metabolism by plants, animals and mycotoxigenic fungi to form several modified forms. These forms may have a diverse toxicity compared to DON, depending on the nature and position of modifications (Dellafiora, Galaverna, & Dall'Asta, 2017). Specifically, this work focused on acetylated (3- and 15-Ac-DON) and glucosylated forms (DON-3- and DON-15-Glc), which are among the most relevant modified forms produced by fungi or produced by plant metabolism, respectively, contaminating commodities along with DON (Berthiller et al., 2013). DON-3- and DON-15-GlcA, and DON–3- and DON-15-Sulf were considered instead as prevalent circulating forms after DON ingestion and represent its main excreted forms in animals, although DON sulphates can be found also in plant-based contaminated commodities (Vidal, Mengelers, Yang, De Saeger, & De Boevre, 2018).

Concerning docking simulations, DON and all its congeners scored positive pointing to their theoretical capability to interact with TAS2R46 (Table S3; Supplementary material). Molecular dynamics could discriminate instead compounds that are likely to stably interact with and activate TAS2R46, as detailed above.

Trichothecolone. Trichothecolone showed a ligand RMSD trend similar to aristolochic acid, though with a significantly lower ($p < 0.01$) average value (Figure S4, Supplementary material). However, the protein RMSD trend was overall more unstable with a significantly higher ($p < 0.01$) average value (0.310 ± 0.04 nm) than aristolochic acid (0.30 ± 0.06 nm). For this reason, TAS2R46-trichothecolone complex was considered not stable and trichothecolone as not likely to activate TAS2R46.

DON. In line with the geometrical stability criteria defined above (i.e. based on the capability of ligands in complex to TAS2R46 to retrace the dynamics of TAS2R46-strychnin/artemisin complex), DON could be deemed a good interactor thanks to the stability of TAS2R46-DON complex over time (Figure 4). Indeed, DON showed a protein RMSD trend like that of strychnine or artemisin with an average value significantly lower ($p < 0.01$) than

that of artemisin (0.25 ± 0.04 and 0.27 ± 0.03 nm, respectively). Concerning the ligand RMSD, DON showed a trend like that of artemisin, though significantly ($p < 0.01$) higher that of strychnine and artemisin but lower than that of aristolochic acid. Therefore, TAS2R46-DON complex was considered stable and comparable to that with artemisin over time. Therefore, DON was considered likely to activate TAS2R46. This DON's capability is consistent with its toxicity and may explain part of the mechanics underpinning its effects on food acceptance and anorexic action, including the release of GLP-1 (Jia et al., 2020) and other enteroendocrine effects. Indeed, TAS2R46 is expressed in the enteroendocrine cells (K. S. Kim et al., 2014) that are present in districts exposed to DON, such as the gastrointestinal tract (Sun et al., 2022). Moreover, the direct involvement of TAS2R46 activation in enteroendocrine cells to release GLP-1 has been reported (K. S. Kim et al., 2014). Therefore, the enteroendocrine action of DON might be mediated (also) by the activation of TAS2R46, and the present study prioritized this mechanism for further dedicated analysis being likely crucial to mechanistically explain the anorexic effect of DON. TAS2R46 has been found expressed also in brain districts (as per The Human Protein Atlas, https://www.proteinatlas.org, last access 8th May 2023) (Uhlen et al., 2015), and DON proved to permeate the Blood Brain Barrier (BBB) exerting neurotoxicity (Zhang et al., 2020), although the underpinning network of biological targets still needs to be clarified. The DON-TAS2R46 interaction described here may shed light also on the mechanics of this specific aspect. As an example, the DON-dependent reduction of B-cell lymphoma 2 protein (Bcl-2) in the neuroblastoma-derivative cell line PC12 (used as model for neuroendocrine cells) has been linked to cell apoptosis (X. C. Wang et al., 2016), and TAS2R46 activation has been associated to Bcl-2 decrease in many cell lines (Costa, Duarte, Costa-Brito, Goncalves, & Santos, 2023). Therefore, DON-dependent activation of TAS2R46 might have a role in the decrease of Bcl-2, which is among the mechanisms involved in its apoptotic effects, deserving further dedicated investigations with priority to better understand the mechanisms underpinning its neurotoxicity.

Acetyl-DONs. According to the analysis, 3-Ac-DON and aristolochic acid had a similar protein RMSD trend although the average 3-Ac-DON value (0.31 ± 0.04 nm) was significantly higher ($p < 0.01$) than that of aristolochic acid (0.30 ± 0.06 nm). Moreover, the ligand RMSD trend of 3-Ac-DON was different from that of DON, with average values significantly higher ($p < 0.01$) than that of aristolochic acid. On the other hand, 15-Ac-DON and DON had a similar protein RMSD trend with an average value significantly lower ($p < 0.01$) than that of DON and in line with that of strychnine (Figure 4). The ligand RMSD trend was also like that of DON with average values slightly higher ($p < 0.01$) than that of DON. Based on these findings, 15-

Ac-DON-TAS2R46 complex was predicted stable over time and 15-Ac-DON likely to activate TAS2R46, whereas 3-Ac-DON was not. Interestingly, 15-Ac-DON producing *Fusarium* chemotypes are the most occurring worldwide (Crippin, Renaud, Sumarah, & Miller, 2019) while the 3-Ac-DON ones are likely to be more aggressive towards plants (Puri & Zhong, 2010). Indeed, 15-Ac-DON may have a high incidence in contaminated commodities showing a toxicity comparable to DON (Knutsen et al., 2017). It was described to be rapidly converted to DON in the intestine, and its reversion to DON was thought responsible for most of its toxicity *in vivo* (Knutsen et al., 2017). However, the outcome presented in this work described its capability to act against TAS2R46 *per se*, which should be investigated further in dedicated analysis to better detail its own toxicity and related mechanisms of action.

DON Sulphates. The protein RMSD trend of DON-3- and -15-Sulf trends was in line with that of DON, although the average values were significantly higher than DON ($p < 0.01$). Ligands RMSD trends were instead quite different from that of DON with average value significantly and consistently higher ($p < 0.01$) than that of aristolochic acid (Figure S1; Supplementary material). Therefore, the complexes with TAS2R46 were considered not stable and DON sulphates were consequently deemed not good interactors and were considered not likely to activate TAS2R46.

DON glucuronides. Although DON-3- and -15-GlcA showed a DON-like protein RMSD trend with significantly lower average values ($p < 0.01$; Figure S2, Supplementary material), the ligand RMSD average values were significantly higher ($p < 0.01$) pointing to their marked deviation from the starting position. This might suggest the overall instability of TAS2R46 complexes. However, unlike DON sulphates that importantly spread over the binding site along the whole simulation (as testified by the marked variability of their ligand RMSD trend; Figure S2, Supplementary material), the spread of DON glucuronides was limited. Based on this, although TAS2R46 complexes with DON glucuronides were predicted as less stable as shown for DON, their capability to activate TAS2R46 cannot be excluded completely.

DON glucosides. DON, DON-3-Glc and DON-15-Glc had a similar protein RMSD trend, and in line with that of strychnine (Figure S3, Supplementary material), but the average DON-3-Glc value was significantly higher ($p < 0.01$) than that of DON while the average DON-15-Glc value was significantly lower ($p < 0.01$). However, both the glucosides ligand RMSD trends were different from that of DON with average values significantly higher ($p < 0.01$). Hence, TAS2R46 complexes with DON glucosides were predicted as not stable and, consequently, DON glucosides not likely to significantly activate TAS2R46. Nonetheless, as

discussed above for DON glucuronides, DON-15-Glc showed a limited spread at the ligand binding site and considering that protein RMSD trend of TAS246R-DON-15-Glc was in line with that of strychnine, its capability to activate TAS2R46 cannot be excluded completely.

Interestingly, DON sulphates, glucuronides, and glucosides were predicted unable to appreciably interact with and activate TAS2R46. This evidence is important considering the presence of DON glucosides, sulphates and 3-Ac-DON in contaminated commodities. The data presented here could suggest that TAS2R46 activation is not likely involved in the mechanisms of action of these analogues, unless they are reverted to DON as may happen in certain tissues and organs, including gut (Knutsen et al., 2017). Germane to human DON phase II metabolites, like sulphates and glucuronides, the data presented here pointed to their substantial incapability to interact with TAS2R46. However, their activity over TAS2R46 and other TAS2Rs is anyway worth of further dedicated investigations to probe the existence of possible related mechanisms considering the expression of TAS2Rs in urogenital tract, including bladder (as per The Human Protein Atlas, https://www.proteinatlas.org, last access 8th May 2023) (Uhlen et al., 2015) and the exposure of urogenital tissues to those forms (being the main DON metabolites excreted via urine).

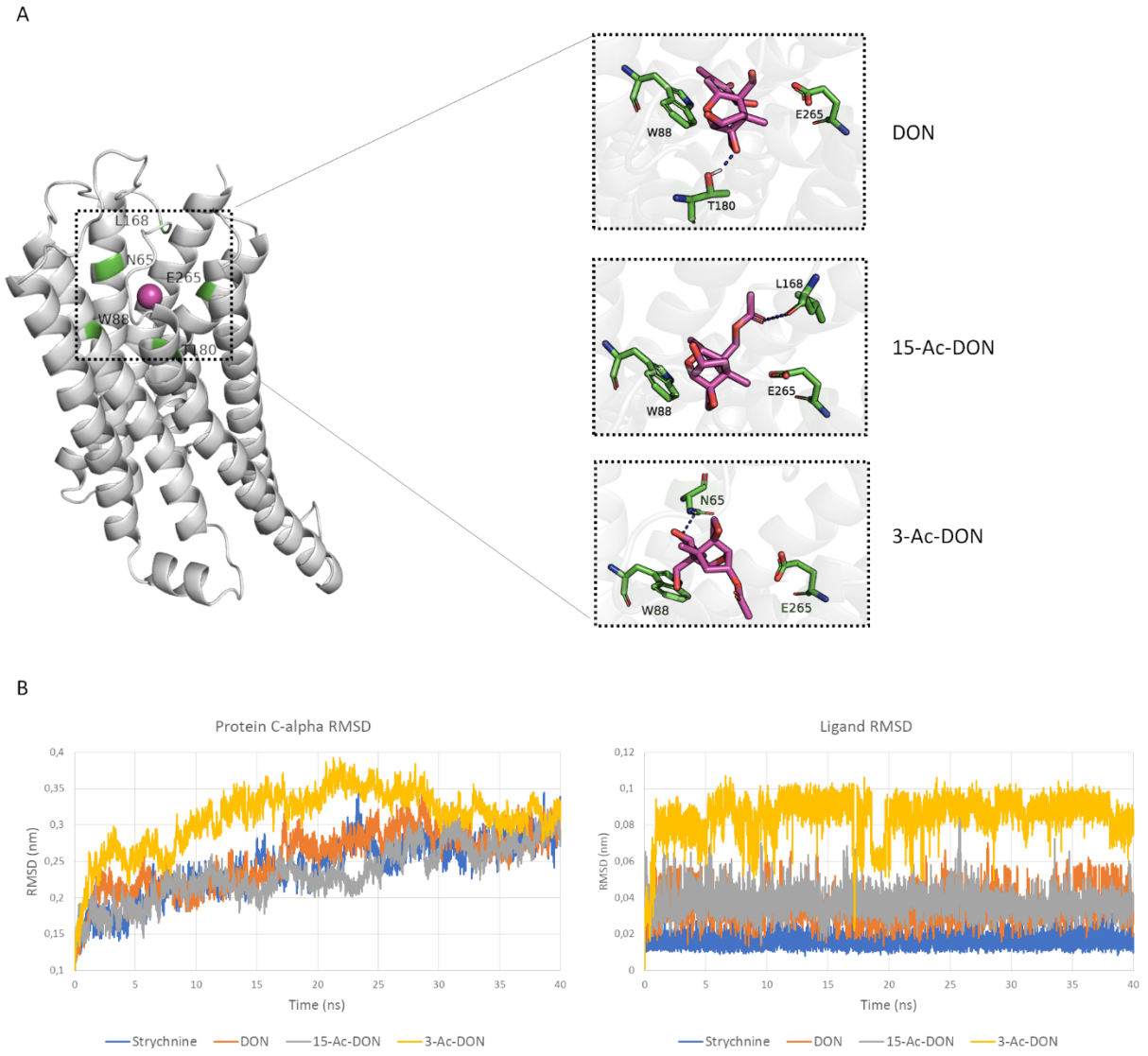


Figure 4. Molecular modelling results of DON, 15-Ac-DON and 3-Ac-DON. **A.** Docking pose of DON, 15-Ac-DON and 3-Ac-DON. On the left, TAS2R46 is represented in white cartoon with the ligand binding site centroid as a magenta sphere. On the right, close-up of the ligand binding site with amino acid side chains and ligands represented in sticks. Blu dashed lines indicate polar contacts. **B**. On the left, protein C-alpha RMSD of TAS2R46 in complex with strychnine (average value 0.241 ± 0.041 nm), DON (average value 0.25 ± 0.04 nm), 15-Ac-DON (average value 0.23 ± 0.04 nm) and 3-Ac-DON (average value 0.31 ± 0.04 nm). On the right, ligand RMSD of strychnine (average value 0.016 ± 0.004 nm), DON (average value 0.034 ± 0.009 nm), 15-Ac-DON (average value 0.037 ± 0.007 nm) and 3-Ac-DON (average value 0.084 ± 0.011 nm).

4. Conclusion

The present work applied a computational workflow that identified some members of the type B trichothecenes of high concern for food safety as possible activators of human TAS2R46. This kind of approach, which previously succeeded as standing-alone upstream method of analysis in the context of mechanistic toxicology, are meant to provide compelling line of evidence to rationally plan further dedicated investigations. Germane to this work, the analysis used trichotecholone as a bait but mainly focused on DON, which is among the most prevalent and abundant mycotoxin in the food chain at a global scale, and a series of congeners including the most relevant analogues found in food and the main circulating and excreted forms in humans upon DON ingestion (i.e., DON glucosides, glucuronides, sulphates and acetylated forms).

Although the extra-oral involvement of bitter receptors in trichothecenes toxicity, especially those expressed in the gut, has been clued in the past, no clear evidence has been provided before to the best of our knowledge. With this respect, this work provided a compelling line of evidence describing the capability of DON to interact with and activate the human TAS2R46 and the need for further dedicated investigations in this direction. This aligns with DON and congeners toxicity extending knowledge on their mechanisms of action and prioritizing TAS2R46 for further investigations aimed at better understanding DON mechanics at the basis of its neurotoxic action and effects on food acceptance.

Concerning DON analogues, 15-Ac-DON was the unique congener among those analysed here found as able to appreciably interact with TAS2R46. Conversely, trichothecolone, 3-Ac-DON, DON sulphates and glycosides (i.e. glucosides and glucuronides) were predicted unable to appreciably interact with and activate TAS2R46.

As a general remark, the outcomes presented could provide a partial mechanistic explanation for DON toxicity. In particular, the capability of DON and 15-Ac-DON to interact with TAS2R46 presented here suggested the need to broadly investigate the capability of trichothecenes to activate bitter receptors, especially those expressed in the extra-oral tissues where trichothecenes have been described exerting toxicity, toward a more aware mechanistic understanding of their action.

References

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX, 1-2*, 19-25.

Amendola, G., & Cosconati, S. (2021). PyRMD: A New Fully Automated Al-Powered Ligand-Based Virtual Screening Tool. *Journal of Chemical Information and Modeling, 61* (8), 3835-3845.

Bell, E. W., & Zhang, Y. (2019). DockRMSD: an open-source tool for atom mapping and RMSD calculation of symmetric molecules through graph isomorphism. *Journal of Cheminformatics, 11*.

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.

Berthiller, F., Crews, C., Dall'Asta, C., De Saeger, S., Haesaert, G., Karlovsky, P., Oswald, I. P., Seefelder, W., Speijers, G., & Stroka, J. (2013). Masked mycotoxins: A review. *Molecular Nutrition & Food Research, 57* (1), 165-186.

Best, R. B., Zhu, X., Shim, J., Lopes, P. E. M., Mittal, J., Feig, M., & MacKerell, A. D. (2012). Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone phi, psi and Side-Chain chi(1) and chi(2) Dihedral Angles. *Journal of Chemical Theory and Computation, 8* (9), 3257-3273.

Brockhoff, A., Behrens, M., Massarotti, A., Appendino, G., & Meyerhof, W. (2007). Broad tuning of the human bitter taste receptor hTAS2R46 to various sesquiterpene lactones, clerodane and labdane diterpenoids, strychnine, and denatonium. *Journal of Agricultural and Food Chemistry, 55* (15), 6236-6243.

Brockhoff, A., Behrens, M., Niv, M. Y., & Meyerhof, W. (2010). Structural requirements of bitter taste receptor activation. *Proceedings of the National Academy of Sciences of the United States of America, 107* (24), 11110-11115.

Costa, A. R., Duarte, A. C., Costa-Brito, A. R., Goncalves, I., & Santos, C. R. A. (2023). Bitter taste signaling in cancer. *Life Sciences, 315*.

Crippin, T., Renaud, J. B., Sumarah, M. W., & Miller, J. D. (2019). Comparing genotype and chemotype of Fusarium graminearum from cereals in Ontario, Canada. *Plos One, 14* (5).

Dagan-Wiener, A., Di Pizio, A., Nissim, I., Bahia, M. S., Dubovski, N., Margulis, E., & Niv, M. Y. (2019). BitterDB: taste ligands and receptors database in 2019. *Nucleic Acids Research, 47* (D1), D1179-D1185.

Del Favero, G., Aichinger, G., Hohenbichler, J., Marko, D., Dall'Asta, C., & Dellafiora, L. (2022). A target fishing study to spot possible biological targets of fusaric acid: Inhibition of protein kinase-A and insights on the underpinning mechanisms. *Food and Chemical Toxicology, 159*.

Dellafiora, L., Galaverna, G., & Dall'Asta, C. (2017). In silico analysis sheds light on the structural basis underlying the ribotoxicity of trichothecenes-A tool for supporting the hazard identification process. *Toxicology Letters, 270*, 80-87.

Dellafiora, L., Magnaghi, F., Galaverna, G., & Dall'Asta, C. (2022). A mechanistic investigation on kokumi-active gamma-Glutamyl tripeptides - A computational study to understand molecular basis of their activity and to identify novel potential kokumi-tasting sequences. *Food Research International, 162*.

Dellafiora, L., Oswald, I. P., Dorne, J. L., Galaverna, G., Battilani, P., & Dall'Asta, C. (2020). An in silico structural approach to characterize human and rainbow trout estrogenicity of mycotoxins: Proof of concept study using zearalenone and alternariol. *Food Chemistry, 312*.

Easterfield, T. H., & Aston, B. C. (1901). XI.—Tutu. Part I. Tutin and coriamyrtin. *Journal of the Chemical Society, Transactions, 79* (0), 120-126.

Ekstrand, B., Young, J. F., & Rasmussen, M. K. (2017). Taste receptors in the gut - A new target for health promoting properties in diet. *Food Research International, 100*, 1-8.

Flannery, B. M., Clark, E. S., & Pestka, J. J. (2012). Anorexia Induction by the Trichothecene Deoxynivalenol (Vomitoxin) Is Mediated by the Release of the Gut Satiety Hormone Peptide YY. *Toxicological Sciences, 130* (2), 289-297.

Fritz, F., Preissner, R., & Banerjee, P. (2021). VirtualTaste: a web server for the prediction of organoleptic properties of chemical compounds. *Nucleic Acids Research, 49* (W1), W679-W684.

Gallo, M., Ferrara, L., Calogero, A., Montesano, D., & Naviglio, D. (2020). Relationships between food and diseases: What to know to ensure food safety. *Food Research International, 137*.

Garg, N., Sethupathy, A., Tuwani, R., Rakhi, N. K., Dokania, S., Iyer, A., Gupta, A., Agrawal, S., Singh, N., Shukla, S., Kathuria, K., Badhwar, R., Kanji, R., Jain, A., Kaur, A.,

Nagpal, R., & Bagler, G. (2018). FlavorDB: a database of flavor molecules. *Nucleic Acids Research, 46* (D1), D1210-D1216.

Holanda, D. M., & Kim, S. W. (2021). Mycotoxin Occurrence, Toxicity, and Detoxifying Agents in Pig Production with an Emphasis on Deoxynivalenol. *Toxins, 13* (2).

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied Logistic Regression, 3rd Edition. In *Applied Logistic Regression, 3rd Edition* (pp. 1-500).

Hou, S. L., Ma, J. J., Cheng, Y. Q., Wang, H. A., Sun, J. H., & Yan, Y. X. (2023). The toxicity mechanisms of DON to humans and animals and potential biological treatment strategies. *Critical Reviews in Food Science and Nutrition, 63* (6), 790-812.

Jia, H., Qin, Z. H., Wei, B., Guo, X. Y., Xiao, H. P., Zhang, H. Y., Li, Z. L., Wu, Q. H., Zheng, R. B., & Wu, W. D. (2022). Substance P and Glucagon-like Peptide-1(7-36) Amide Mediate Anorexic Responses to Trichothecene Deoxynivalenol and Its Congeners. *Toxins, 14* (12).

Kamila, T., & Agnieszka, K. (2021). An update on extra-oral bitter taste receptors. *Journal of Translational Medicine, 19* (1).

Khaneghah, A. M., Farhadi, A., Nematollahi, A., Vasseghian, Y., & Fakhri, Y. (2020). A systematic review and meta-analysis to investigate the concentration and prevalence of trichothecenes in the cereal-based food. *Trends in Food Science & Technology, 102*, 193-202.

Kim, K. S., Egan, J. M., & Jang, H. J. (2014). Denatonium induces secretion of glucagon-like peptide-1 through activation of bitter taste receptor pathways. *Diabetologia, 57* (10), 2117-2125.

Kim, S., Chen, J., Cheng, T. J., Gindulyte, A., He, J., He, S. Q., Li, Q. L., Shoemaker, B. A., Thiessen, P. A., Yu, B., Zaslavsky, L., Zhang, J., & Bolton, E. E. (2022). PubChem 2023 update. *Nucleic Acids Research*.

Knutsen, H. K., Alexander, J., Barregard, L., Bignami, M., Bruschweiler, B., Ceccatelli, S., Cottrill, B., Dinovi, M., Grasl-Kraupp, B., Hogstrand, C., Hoogenboom, L., Nebbia, C. S., Oswald, I. P., Petersen, A., Rose, M., Roudot, A. C., Schwerdtle, T., Vleminckx, C., Vollmer, G., Wallace, H., De Saeger, S., Eriksen, G. S., Farmer, P., Fremy, J. M., Gong, Y. Y., Meyer, K., Naegeli, H., Parent-Massin, D., Rietjens, I., van Egmond, H., Altieri, A., Eskola, M., Gergelova, P., Bordajandi, L. R., Benkova, B., Dorr, B., Gkrillas, A., Gustaysson, N., van Manen, M., Edler, L., & Chain, E. P. C. F. (2017). Risks to human and animal health related to the presence of deoxynivalenol and its acetylated and modified forms in food and feed. *Efsa Journal, 15* (9).

Konishi, K., Iida, A., Kaneko, M., Tomioka, K., Tokuda, H., Nishino, H., & Kumeda, Y. (2003). Cancer preventive potential of trichothecenes from Trichothecium roseum. *Bioorganic & Medicinal Chemistry, 11* (12), 2511-2518.

Li, D. Y., & Zhang, J. Z. (2014). Diet Shapes the Evolution of the Vertebrate Bitter Taste Receptor Gene Repertoire. *Molecular Biology and Evolution, 31* (2), 303-309.

Lim, E., Pon, A., Djoumbou, Y., Knox, C., Shrivastava, S., Guo, A. C., Neveu, V., & Wishart, D. S. (2010). T3DB: a comprehensively annotated database of common toxins and their targets. *Nucleic Acids Research, 38*, D781-D786.

Lu, P., Zhang, C. H., Lifshitz, L. M., & ZhuGe, R. H. (2017). Extraoral bitter taste receptors in health and disease. *Journal of General Physiology, 149* (2), 181-197.

Maldonado-Rojas, W., & Olivero-Verbel, J. (2011). Potential interaction of natural dietary bioactive compounds with COX-2. *Journal of Molecular Graphics & Modelling, 30*, 157-166.

Patience, J. F., Myers, A. J., Ensley, S., Jacobs, B. M., & Madson, D. (2014). Evaluation of two mycotoxin mitigation strategies in grow-finish swine diets containing corn dried distillers grains with solubles naturally contaminated with deoxynivalenol. *Journal of Animal Science, 92* (2), 620-626.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry, 25* (13), 1605-1612.

Puri, K. D., & Zhong, S. B. (2010). The 3ADON Population of Fusarium graminearum Found in North Dakota Is More Aggressive and Produces a Higher Level of DON than the Prevalent 15ADON Population in Spring Wheat. *Phytopathology, 100* (10), 1007-1014.

Rahman, Z., Zidan, A. S., Berendt, R. T., & Khan, M. A. (2012). Tannate complexes of antihistaminic drug: Sustained release and taste masking approaches. *International Journal of Pharmaceutics, 422* (1-2), 91-100.

Sali, A., & Blundell, T. L. (1993). COMPARATIVE PROTEIN MODELING BY SATISFACTION OF SPATIAL RESTRAINTS. *Journal of Molecular Biology, 234* (3), 779-815.

Suares, D., & Hiray, A. (2015). Taste Masked Orodispersible Formulation of Fexofenadine Hydrochloride Using Ion Exchange Resins. *Indian Journal of Pharmaceutical Sciences, 77* (5), 550-555.

Sun, Y., Jiang, J., Mu, P. Q., Lin, R. Q., Wen, J. K., & Deng, Y. Q. (2022). Toxicokinetics and metabolism of deoxynivalenol in animals and humans. *Archives of Toxicology, 96* (10), 2639-2654.

Tolosa, J., Rodriguez-Carrasco, Y., Ruiz, M. J., & Vila-Donat, P. (2021). Multi-mycotoxin occurrence in feed, metabolism and carry-over to animal-derived food products: A review. *Food and Chemical Toxicology, 158*.

Uhlen, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., Alm, T., Edqvist, P. H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J. M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J., & Ponten, F. (2015). Tissue-based map of the human proteome. *Science, 347* (6220).

Vidal, A., Mengelers, M., Yang, S. P., De Saeger, S., & De Boevre, M. (2018). Mycotoxin Biomarkers of Exposure: A Comprehensive Review. *Comprehensive Reviews in Food Science and Food Safety, 17* (5), 1127-1155.

Wang, Q. L., Liszt, K. I., & Depoortere, I. (2020). Extra-oral bitter taste receptors: New targets against obesity? *Peptides, 127*.

Wang, X. C., Xu, W., Fan, M. X., Meng, T. T., Chen, X. F., Jiang, Y. J., Zhu, D. F., Hu, W. J., Gong, J. J., Feng, S. B., Wu, J. J., & Li, Y. (2016). Deoxynivalenol induces apoptosis in PC12 cells via the mitochondrial pathway. *Environmental Toxicology and Pharmacology, 43*, 193-202.

Warth, B., Fruhmann, P., Wiesenberger, G., Kluger, B., Sarkanj, B., Lemmens, M., Hametner, C., Frohlich, J., Adam, G., Krska, R., & Schuhmacher, R. (2015). Deoxynivalenol-sulfates: identification and quantification of novel conjugated (masked) mycotoxins in wheat. *Analytical and Bioanalytical Chemistry, 407* (4), 1033-1039.

Wismer, T. (2016). Chapter 79 - Feline Toxins: Recognition, Diagnosis, Treatment,. In S. E. Little (Ed.), *August's Consultations in Feline Internal Medicine* (Vol. 7, pp. 791-798).

Wooding, S. P., Ramirez, V. A., & Behrens, M. (2021). Bitter taste receptors Genes, evolution and health. *Evolution Medicine and Public Health, 9* (1), 431-447.

Xu, W. X., Wu, L. J., Liu, S. H., Liu, X., Cao, X. L., Zhou, C., Zhang, J. Y., Fu, Y., Guo, Y., Wu, Y. R., Tan, Q. W., Wang, L., Liu, J. L., Jiang, L. Q., Fan, Z. B., Pei, Y., Yu, J. Y., Cheng, J. J., Zhao, S. W., Hao, X. J., Liu, Z. J., & Hua, T. (2022). Structural basis for

strychnine activation of human bitter taste receptor TAS2R46. *Science, 377* (6612), 1298-1303.

Zhang, J. J., You, L., Wu, W. D., Wang, X., Chrienova, Z., Nepovimova, E., Wu, Q. H., & Kuca, K. (2020). The neurotoxicity of trichothecenes T-2 toxin and deoxynivalenol (DON): Current status and future perspectives. *Food and Chemical Toxicology, 145*.

Zoete, V., Cuendet, M. A., Grosdidier, A., & Michielin, O. (2011). SwissParam: A Fast Force Field Generation Tool for Small Organic Molecules. *Journal of Computational Chemistry, 32* (11), 2359-2368.

Chapter III

A mechanistic toxicology study to grasp the mechanics of zearalenone estrogenicity: Spotlighting aromatase and the effects of its genetic variability

This chapter is based on the paper publication on:

Florinda Perugino, Lorenzo Pedroni, Gianni Galaverna, Chiara Dall'Asta, Luca Dellafiora. (2024). A mechanistic toxicology study to grasp the mechanics of zearalenone estrogenicity: Spotlighting aromatase and the effects of its genetic variability. Toxicology, 501 (2024), 153686, https://doi.org/10.1016/j.tox.2023.153686.

Author contribution: Data curation, Formal analysis, Methodology, Software, Writing – original draft, Writing – review & editing

Abstract

Zearalenone (ZEN) is a mycoestrogen produced by *Fusarium* fungi contaminating cereals and in grain-based products threatening human and animal health due to its endocrine disrupting effects. Germane to the mechanisms of action, ZEN may activate the estrogen receptors and inhibit the estrogens-producing enzyme aromatase (CYP19A1). Both show single nucleotide variants (SNVs) among humans associated with a diverse susceptibility of being activated or inhibited. These variations might modify the endocrine disrupting action of zearalenone, requiring dedicated studies to improve its toxicological understanding. This work focused on human aromatase investigating via 3D molecular modelling whether some of the variants reported so far (n=434) may affect the inhibitory potential of ZEN. It has been also calculated the inhibition capability of α-zearalenol, the most prominent and estrogenically potent phase I metabolite of ZEN, toward those aromatase variants with an expected diverse sensitivity of being inhibited by ZEN. The study: i) described SNVs likely associated with a different susceptibility to ZEN and α-zearalenol inhibition - like T310S that is likely more susceptible to inhibition, or D309G and S478F that are possibly inactive variants; ii) proofed the possible existence of inter-individual susceptibility to ZEN; iii) prioritized aromatase variants for future investigations toward a better comprehension of ZEN xenoestrogenicity at an individual level.

Chapter III

1. Introduction

Zearalenone (ZEN; Figure 1) is a mycotoxin produced by fungi belonging to the *Fusarium* genus (mainly *F. culmorum* and *F. graminearum*) (EFSA, 2011; Malir et al., 2023) and it is among those most relevant to food safety due to its toxicity and high incidence in food and feed commodities (Catteuw et al., 2019; Silva et al., 2019). Indeed, ZEN is often present in various food and feed ingredients such as corn and wheat, and some food pose a higher risk to unintended ZEN consumption, including whole meal cereals and certain gluten-free products (Jing et al., 2022; Pfleger & Schwake-Anduschus, 2023). ZEN has shown an endocrine disrupting activity *in vivo* as animal studies reported its capability to impair hormonal balance (Kowalska et al., 2016). In humans, evidence have pointed to a series of disorders associated with the dietary exposure to ZEN, including effects on girls' growth, precocious puberty and thelarche/mastopathy development (Massart et al., 2008; Massart & Saggese, 2010; Rai et al., 2020; Rivera-Nunez et al., 2019). It was also hypothesized that ZEN can be associated with cancer development/progression (Claeys et al., 2020; Ekwomadu et al., 2022; E. K. K. Lo et al., 2023; Rong et al., 2022), though it has been defined as not classifiable as carcinogenic to humans due to the limited evidence in experimental animals (Claeys et al., 2020; IARC, 1993). Germane to the mechanisms of toxicity, ZEN estrogenicity may rely on its concerted capability to activate estrogen receptors (ERs) and inhibit aromatase (CYP19A1) (EFSA, 2011; Y. F. Wang et al., 2014). From a structural standpoint, ZEN is a resorcyclic acid lactone (molecular formula C18H22O5) whose structure resembles that of the endogenous estrogen 17β-estradiol (E2). This similarity allows for a (partial) agonistic behaviour of ZEN versus ERs, which can be bound and activated eventually eliciting an estrogenic response (EFSA, 2011). Similarly, the structural analogies to steroids make ZEN suitable to compete also with aromatase substrates/products resulting in its inhibition (Y. F. Wang et al., 2014). This event may lead to a ZEN-dependent reduction of E2 production being aromatase responsible for the conversion of androgens to estrogens (Rizner & Romano, 2023). The final balance of these two apparently contradictory effects concurs in the determination of the final estrogenic potency of ZEN. Moreover, it has been proved that the sensitivity to ZEN-dependent estrogenic insult may depend on the phase I metabolism of ZEN and specifically on the production of α-zearalenol (αZEL). αZEL is part of the chemical complex found in contaminated foodstuff (De Boevre et al., 2013), but it is also the predominant ZEN phase I metabolite in humans, and it has shown a xenoestrogenic activity way more potent than ZEN (Gupta et al., 2022). With this respect, animal species able to predominantly produce αZEL over the other ZEN phase I metabolites

are typically associated with strong sensitivity to ZEN exposure (Knutsen et al., 2017). Moreover, a close structural analogue of αZEL, i.e. the α-zearalanol (αZAL), has also been proved to inhibit aromatase although with a lower efficiency compared to ZEN (Y. F. Wang et al., 2014), suggesting that αZEL might have a certain degree of aromatase inhibitory potential as well. ERs and aromatase show single nucleotide variants (SNVs) among individuals which may impact their inherent activity and functions (Katzenellenbogen et al., 2018; Simpson, 2000). In this respect, the present work investigated SNVs of aromatase being the ZEN-dependent inhibition of aromatase still largely overlooked though crucial to better characterize the toxicodynamic of ZEN. Specifically, it has been proved that SNVs of aromatase may have a diverse activity and susceptibility of being inhibited compared to the wild-type (WT) enzyme (Kao et al., 1998; Russell et al., 2014). This background information suggests that the inhibitory effects of ZEN might change depending on the aromatase variant, which in turn might impact the disrupting activity of ZEN in individuals depending on the aromatase variant they express. These aspects, the analysis of which targets the toxicodynamic of ZEN from a "personalised" standpoint, are still largely overlooked though they might enforce the background information of ZEN toxicology. The clarification of these aspects is desirable either to boost the understanding of ZEN action among the human population or to improve the assessment of ZEN risk at an individual level. Specifically, the study targeted the whole set of human aromatase SNVs variants available in UniProt (www.uniprot.org) (Bateman et al., 2023) at the time of analysis (i.e. 434 variants; last database access 25th September 2023), focusing on those surrounding the ligand binding pocket (15 SNVs) being most likely to affect the interaction of ZEN and αZEL. From a methodological standpoint, computational approaches already proved to be successful, and their application is getting broader and able to cope with several aspect of the food science field (Ji et al., 2020; Y. Wang et al., 2022; Zhao et al., 2021). Specifically, this study relied on a well-established *in silico* workflow integrating docking studies and molecular dynamics to simulate the binding event and the stability of ligand-aromatase complex over the time as a mean to study the effects of SNVs on aromatase inhibition by ZEN. In this respect, evidence previously collected have proven that this analytical approach succeeded to estimate effectively whether mutations affect the protein-ligand complex formations (Dorne et al., 2022; Louisse et al., 2022), turning out to be particularly suitable for this case study.

Overall, the study described aromatase SNVs likely associated with a different susceptibility to ZEN inhibition, proofing the possible existence of a certain degree of inter-individual variability to ZEN toxicity. Then, it has been also calculated the inhibitory potential of αZEL toward those aromatase variants (i.e. T310I and T310S) which showed an expected diverse (but

not null) inhibition by ZEN compared to WT. This could estimate whether SNV may also affect the susceptibility to αZEL. Such variants have been rationally prioritized for future investigation and the implications of these findings to the current understanding of ZEN toxicology have been also discussed.

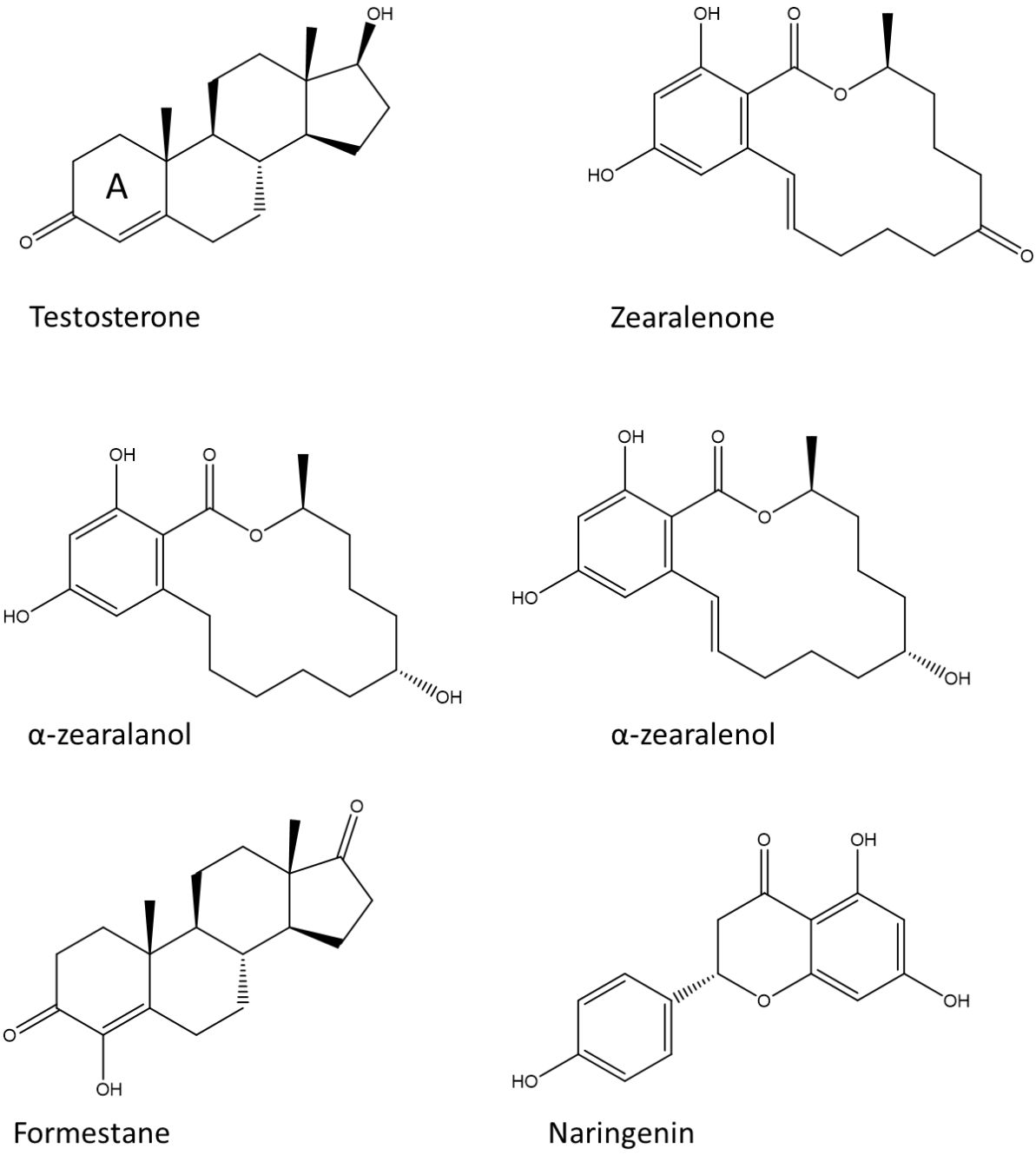


Figure 1. Chemical structure of molecules under analysis.

2. Materials and Methods

2.1 Data source and management

2.1.1 Retrieval of aromatase variants

All the SNVs of human aromatase (UniProt code P11511) were retrieved from the "Variant viewer" section of the UniProt database (www.uniprot.org) (Bateman et al., 2023) which listed 434 variants at the time of analysis (last database access 25th September 2023). However, the analysis focused on those at the ligand binding pocket (15 SNVs) being most likely to influence the interaction of ZEN. The visual inspection of the 3D structure of human aromatase allowed for the definition of such a subset of SNVs (see below) with the following carried forth the analysis: A306T, D309G, D309N, F221S, M374I, M374L, M374T, R115Q, S478F, T310A, T310I, T310S, V370A, V370M and V370F.

2.1.3 Retrieval of protein and ligands structures

The list of molecules under analysis included ZEN, αZEL, testosterone (TES) and naringenin (NAR) with the latter two taken as reference compounds, and formestane. Their 3D structures were retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov) (Kim et al., 2023) in the Structure Data File (.sdf) format (CID: 5281576, 5284645, 6013, 439246 and 11273 respectively).

The 3D model of human WT aromatase was derived from the crystallographic structure stored on the Protein Data Bank (PDB; https://www.rcsb.org; last database access 24th July 2023) (Berman et al., 2000) with PDB code 3S79 (Ghosh et al., 2012). This structure was also used to derive the 3D model for the 15 aromatase variants listed above since their crystallographic structure was not available at the time of analysis (see below).

2.1.3 Protein model preparation

The WT 3D structure of human aromatase (PDB code 3S79) (Ghosh et al., 2012) was processed with UCSF Chimera software (version 1.16) (Pettersen et al., 2004) removing water and the co-crystallized ligand and adding hydrogens, in agreement with previous work. The 3D structure of aromatase variants was generated processing the WT structure with Chimera (version 1.16) (Pettersen et al., 2004), replacing selected amino acids using the Structure Editing/Rotamer tool and choosing the rotamer with the highest computed probability to occur when multiple rotamers were computed, in agreement with previous studies (Louisse et al., 2022).

2.2 Docking Simulations

Docking simulations were performed to provide a plausible binding architecture for each molecule under analysis within the aromatase binding site and to predict the effects of different SNVs on the interaction between ZEN and aromatase. Docking simulations were performed with the GOLD software (version 2021), using the internal GOLDScore scoring function, as already succeeded to calculate the effect of aminoacidic substitutions to the protein-ligand binding (Pedroni, Louisse, Dorne, et al., 2023; Pedroni, Louisse, Punt, et al., 2023). The binding site was set within a 5 Å radius sphere around the centroid of the co-crystallized ligand-binding site. A semi-flexible docking protocol was applied while allowing protein's polar hydrogens free to rotate and considering ligands fully flexible. The Root Mean Square Deviation (RMSD) analysis between the calculated and the crystallographic pose of TES was performed with DockRMSD webserver (version 1.1; https://zhanggroup.org/DockRMSD) (Bell & Zhang, 2019) and considering non hydrogen atoms only.

2.3 Molecular Dynamics Simulations

Molecular dynamics (MD) simulations were performed through GROMACS (version 2021.4) (Abraham et al., 2015) to monitor the geometrical stability of the complex and ligand orientation. All the ligands were processed and parametrised with the SwissParam tool (https://ww.swissparam.ch) (Zoete et al., 2011) and the whole system was parametrised with the CHARMM27 all-atom force field (Brooks et al., 2009). The hydrogen database was modified according to previous works (Dorne et al., 2022; Panneerselvam et al., 2015; Pedroni, Louisse, Dorne, et al., 2023; Zhang et al., 2012) to properly parameterise the heme group. The input complex structures were solvated with SPCE water in a dodecahedron periodic boundary condition and neutralized adding Na^+ and Cl^- as counter ions. Before running MD simulations, each system underwent an energetical minimization to both avoid steric clashes and correct improper geometries using the steepest algorithm with a maximum of 5000 steps. Then, each system underwent isothermal (300 K; coupling time of 2 ps) and isobaric (1 bar; coupling time of 2 ps) 100 ps simulations before undergoing 40 ns long MD simulations (300 K with a coupling time of 0.1 ps and 1 bar with a coupling time of 2 ps).

2.4 Statistical analysis

The statistical analysis to compare SNVs and WT using the mean RMSDs of ligands and/or the average interatomic distances between reacting atom and heme's iron was performed with SPSS IBM (v. 27.0, SPSS Inc., Chicago, IL, USA). For each complex, 8000 frames were considered, expressed as means ± standard deviation (SD) and pairwise compared to the WT using t test ($\alpha = 0.05$).

3. Results and Discussion

The work aimed at assessing whether SNVs of aromatase (CYP19A1) occurring among human population may affect its susceptibility of being inhibited by ZEN. The investigation targeted a selection made from the 434 aromatase SNVs recorded in the UniProt database (www.uniprot.org) (Bateman et al., 2023) at the time of analysis (last database access 25th September 2023) to make the study technically feasible. Specifically, this investigation targeted the substitutions of those amino acids that surrounded the ligand binding site, assuming they are most likely to have an impact on aromatase-ZEN interaction, which were 15 at the time of analysis: A306T, D309G, D309N, F221S, M374I, M374L, M374T, R115Q, S478F, T310A, T310I, T310S, V370A, V370M and V373F (Figure 2A).

3.1 Assessment of procedural performances

The workflow presented here, which is based on the integrated use of docking simulations and MD, already proofed its efficacy to reliably estimate the effect mutations may have on the capability of protein to recruit ligands (Dorne et al., 2022; Louisse et al., 2022).However, a fit-for-purpose validation was done prior to test whether aromatase SNVs may have a diverse susceptibility to ZEN inhibition. In agreement with previous studies, the procedural performances were assessed following a multi-tier approach (Dellafiora et al., 2022; Louisse et al., 2022). Firstly, the capability to reproduce the binding architecture of TES, taken as reference for aromatase ligands, was checked comparing its calculated and crystallographic poses. As shown in Figure 2B, the docking pose was very similar to the crystallographic one with a RMSD value of 0.345 Å, supporting the model reliability to calculate meaningful binding architectures. Afterward, the procedure has been challenged using data from the literature to prove its capability to provide geometrical rationales which may effectively estimate the effects of mutations on aromatase activity and susceptibility to inhibition.

To do so, the variants V370M, D309N and T310S, taken as reference being previously assessed for activity and susceptibility to inhibition (Kao et al., 1998; J. Lo et al., 2013; Ludwig et al., 1998), were analysed combining data from docking and MD simulations. Specifically, it was monitored the geometrical stability of complexes over time using the RMSD of ligands, proteins and ligands trajectories, and monitoring the number of protein-ligand hydrogen bonds over time, in agreement with previous studies (Dorne et al., 2022; Zhao et al., 2021). V370M and D309N have been previously described as inactive aromatase variants (Kao et al., 1998; J. Lo et al., 2013; Ludwig et al., 1998), while T310S has been described active though more susceptible of being inhibited by NAR compared to the WT enzyme (Kao et al., 1998). Germane to the mechanism of catalysis, aromatase acts turning the A ring of substrates like TES into an aromatic ring through methyl oxidation resulting in the elimination of the methyl group (Di Nardo et al., 2015). To do so and according to previous studies (Dorne et al., 2022), the methyl group undergoing the reaction must be closely and stably oriented toward the catalytic core of the enzyme, namely the heme's iron atom.

3.1.1 Inactive variants (V370M, D309N)

Concerning V370M, in line with this evidence, TES has been docked into the WT aromatase and in the V370M variant. Of note, docking scores may correlate with the capability of ligands to fit the protein pockets and positive scores typically indicate a good fitting (the higher the score, the better the fitting of ligands into the pocket; https://www.ccdc.cam.ac.uk/). In this respect, the docking scores TES recorded into the WT aromatase and the V370M were 63.15 and –16.51 units, respectively. The negative score TES recorded in the V370M variant pointed to the strong impairing effects this substitution had in terms of TES-pocket fitting. Then, both complexes were further analysed with MD simulations. Specifically, the interatomic distance between the carbon of the methyl group undergoing the reaction and the heme's iron atom was measured over time comparing those observed for the WT aromatase with those of the inactive V370M variant. As shown in Figure 3A, docking analysis revealed that the V370M substitution was responsible for a diverse arrangement of TES compared to the WT enzyme where the methyl group undergoing the reaction was not properly oriented as saw in the WT complex. Moreover, MD revealed that the distance between the methyl group in V370M is significantly ($p < 0.001$) farer to the heme's iron over time compared to the WT enzyme (0.699 ± 0.031 and 0.359 ± 0.032 nm, respectively). In line with previous studies (Dorne et al., 2022), this evidence confirmed geometrical rationales such as the arrangement at binding site and the distance of the

atom undergoing the reaction to the heme's iron measured over time as meaningful parameters to estimate the effects of mutations on aromatase capability to turn substrates into products. Besides, to the best of our knowledge, these results provide a first convincing mechanistic explanation for the inactivity of V370M variant. Specifically, our findings suggest that mutations able to "shield" the heme's iron from the atom undergoing the reaction are likely to reduce the enzyme activity. Indeed, the dynamics of interaction observed in V370M suggest that this hindrance possibly occurs when the substitution involves bulky amino acids, thereby preventing the atom undergoing the reaction to get close to the enzyme's catalytic core.

Concerning D309N, at a first instance the docking results and the interaction of TES monitored over time with MD simulations were apparently contradictory to the reported inactivity of this variant. Indeed, the docking scores TES recorded in the WT aromatase and in the D309N variant were comparable (63.15 and 63.99 units, respectively) as well as the complex with the variant was barely comparable to the WT in terms of interatomic distance between the TES reacting atom and heme's iron (0.348 ± 0.030 nm and 0.359 ± 0.032, respectively), and ligand RMSD trends and average values (the latter were 0.030 ± 0.005 and 0.036 ± 0.006, respectively) (Figure 3C). This evidence pointed to the substantial capability of TES to interact with the D309N variant. Interestingly, a deeper analysis of aromatase pharmacodynamics provided a convincing explanation of the reason why D309N may result inactive although TES can stably persist at the catalytic site, as per our outcome. Indeed, TES showed in D309N a stably higher number of hydrogen bonds compared to the WT, with an additional bond between the keto group in position 3 and the N309 polar side chain (Figure 3D). Interestingly, this evidence could explain the mechanisms underpinning the inactivity of D309N variant. Indeed, the interpretative line that an additional hydrogen bond may result in enzyme inactivity, i.e. incapability to turn the substrate TES into products, agrees with the pharmacodynamic of the aromatase inhibitor formestane. Formestane has a steroidal scaffold differing from aromatase substrates uniquely for an additional hydroxyl group in position 4 which has been described forming a hydrogen bond with the side chain of D309 (Di Nardo et al., 2015; Murthy et al., 2005), as also shown in our docking studies (GOLDScore 63.21 units; Figure 3D). The interaction with the D309 residue is critical for the inhibitory mechanisms of formestane being involved in the electron transfer, which is altered to the point of enzyme inhibition when is engaged in hydrogen bonds (Martin et al., 2016; Murthy et al., 2005). Therefore, although the D>N substitution is likely to have an inherently altered electron transfer capacity, the additional hydrogen bond described in this study is also compatible with a disrupting action in electron transfer and could provide a mechanistic explanation for inactivity of D309N variant.

Of note, the outcome of D309N and V370M, beside pointing to the procedural reliability to properly estimate the effects of mutations on aromatase activity, may help shedding light on the mechanics of aromatase deficiency, a possibly severe clinical condition. Indeed, a more informed mechanistic understanding of aromatase SNVs provides potentially relevant pieces of information to integrate its current clinical/pharmacological understanding and plan future dedicated studies.

3.1.2 Susceptible variant (T310S)

Germane to NAR, it is a flavonoid reported as a phytoestrogen responsible for aromatase inhibition while the T310S variant is a functional variant more susceptible to NAR inhibition compared to the WT enzyme (Lephart, 2015). Also in this case, NAR has been docked into the WT aromatase and into the T310S variant (Figure 3B). The GOLDScore of NAR into the WT aromatase and the T310S variant were 55.47 and 57.73 units, respectively. Docking studies revealed that NAR was comparably arranged at the ligand binding site with minor differences between the WT enzyme and its variant. However, the 4% higher score observed for NAR in complex with the T310S variant could point to the better fitting compared to the wild-type enzyme. Then, the geometrical stability of NAR in the WT enzyme and its T310S variant has been monitored over time through MD simulations. As shown in Figure 3B, MD simulations revealed that the interaction of NAR with the T310S variant was significantly more stable than that with the WT enzyme with average RMSD significantly ($p < 0.001$) lower (0.057 ± 0.031 nm in T310S; 0.092 ± 0.011 nm in WT). Of note, as discussed elsewhere (Zhao et al., 2021), an enhanced geometrical stability like that described here for NAR is compatible with an increased capability of ligands to persist at the ligand binding site and, in case of enzymes, this can correlate with a higher inhibitory capacity, in line with the experimental evidence of T310S variant inhibition by NAR (Kao et al., 1998).

Overall, these outcomes confirmed the geometrical rationales provided by the integrated use of docking and MD simulations described above, including the hydrogen bond network monitored over time, as meaningful to study the activity and the susceptibility of being inhibited of aromatase variants. Therefore, the *in silico* workflow presented here turned out to be reliable to estimate the effects of aromatase variants on its susceptibility of being inhibited by ZEN.

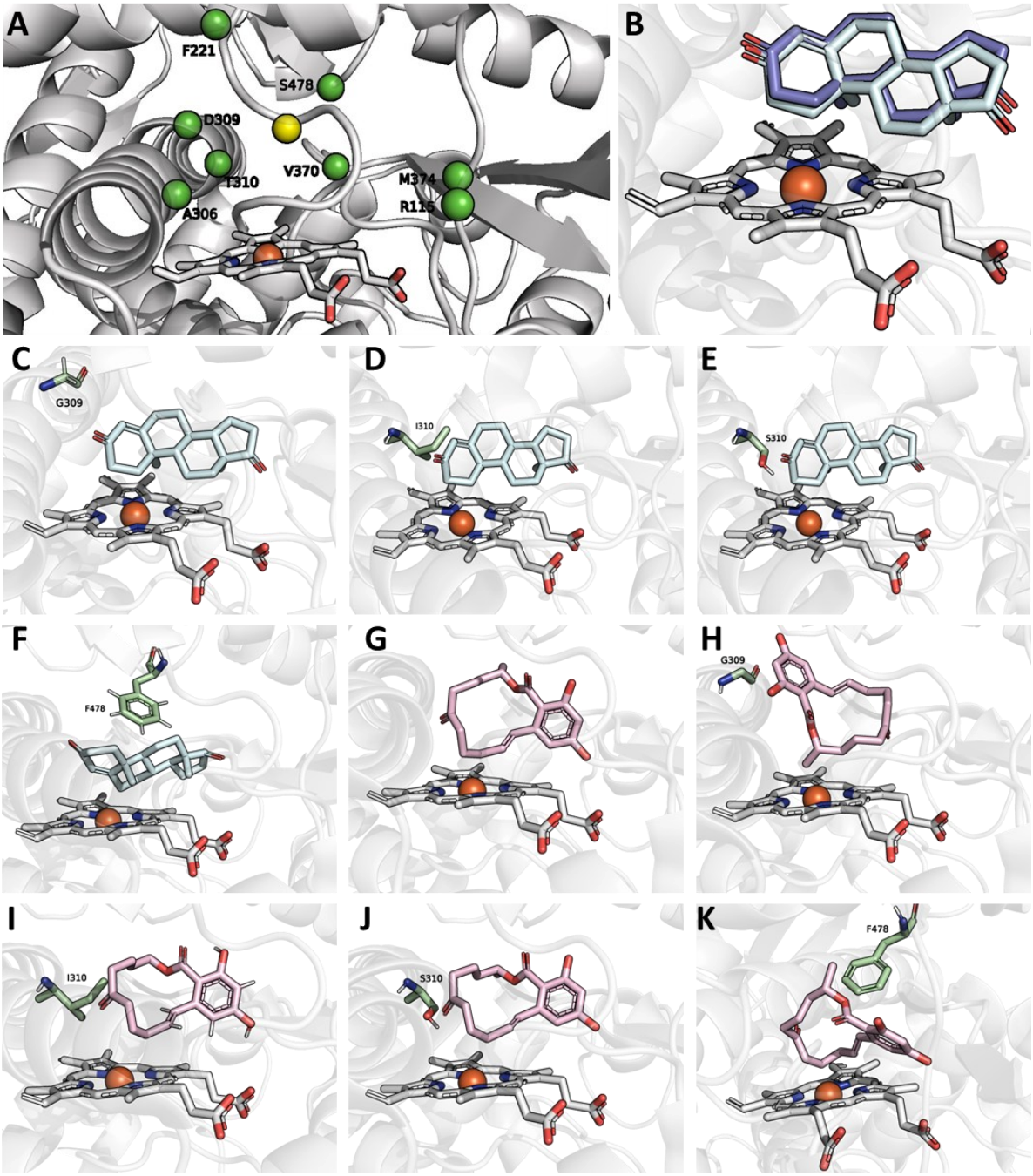


Figure 2. Protein binding site and docking poses of TES (pale cyan sticks) and ZEN (pale pink sticks) in WT and in a selection of variants. Heme is represented as white sticks while the protein is represented as transparent white cartoon. The mutated residue in each variant is represented as pale green sticks. **A**. Close-up of the protein binding site. The yellow sphere represents the centroid set for the docking simulation while the green spheres indicate the positions of residues mutated in the considered variants. **B**. Close-up of the TES binding pose obtained via docking simulation (pale cyan sticks) overlapped to the crystallographic pose of TES (purple sticks; PDB ID 3S79). **C**. TES docking pose in D309G. **D**. TES docking pose in T310I. **E**. TES docking pose in T310S. **F**. TES docking pose in S478F. **G**. ZEN docking pose

in WT. **H**. ZEN docking pose in D309G. **I**. ZEN docking pose in T310I. **J**. ZEN docking pose in T310S. **K**. ZEN docking pose in S478F.

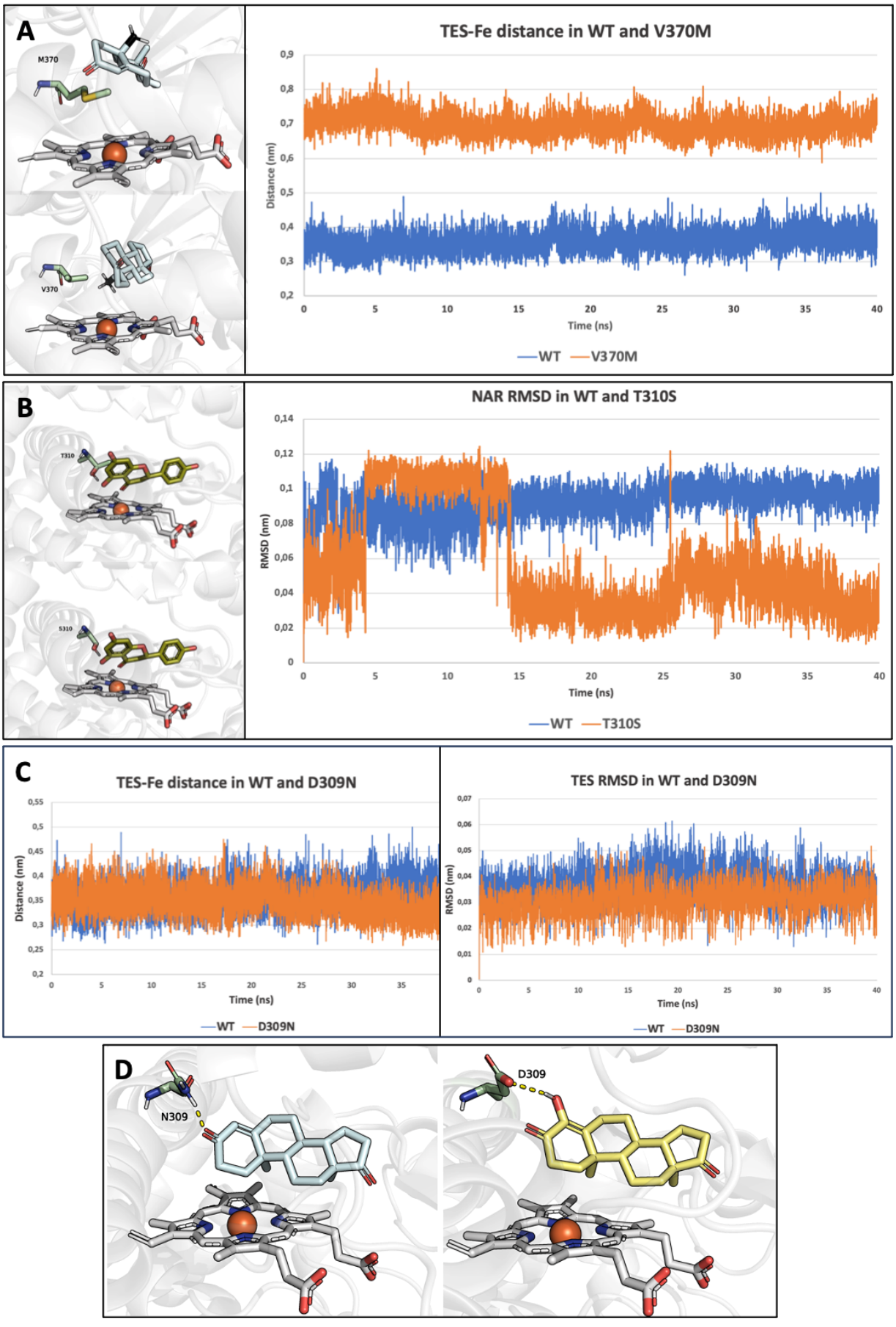

A
M370
V370
TES-Fe distance in WT and V370M
Distance (nm)
Time (ns)
WT
V370M
B
T310
S310
NAR RMSD in WT and T310S
RMSD (nm)
Time (ns)
WT
T310S
C
TES-Fe distance in WT and D309N
Distance (nm)
Time (ns)
WT
D309N
TES RMSD in WT and D309N
RMSD (nm)
Time (ns)
WT
D309N
D
N309
D309

Figure 3. Model validation. Protein is reported as transparent white cartoon while heme in white sticks. The average distances between the atom undergoing reaction and heme's iron atom ± SD are reported between brackets. **A**. Comparison between TES-V370M and TES-WT complex. TES is represented as pale cyan sticks with the methyl group undergoing the reaction as black sticks. On the top-left, TES docking pose in the SNV V370M while below TES docking pose in the WT. The residues V370 and M370 are reported as pale green sticks. On the right, interatomic distances between TES's atom undergoing reaction and heme's iron within WT (blue; 0.359 ± 0.032) and in V370M (orange; 0.699 ± 0.031 nm). **B**. Comparison between NAR-T310S and NAR-WT complex. NAR is represented as green olive sticks. On the top-left, NAR docking pose in the WT while below NAR docking pose in the SNV T310S. The residues T310 and S310 are reported as pale green sticks. On the right, NAR RMSD within WT (blue; 0.092 ± 0.011 nm) and T310S (orange; 0.057 ± 0.031 nm). **C**. On the left, interatomic distances between TES's atom undergoing reaction and heme's iron within WT (blue; 0.359 ± 0.032 nm) and D309N (orange; 0.348 ± 0.030 nm). On the right, TES RMSD within WT (blue; 0.036 ± 0.006 nm) and D309N (orange; 0.030 ± 0.005 nm) **D**. On the left, TES (pale cyan sticks) docking pose in the D309N SNV. The dashed yellow lines represent the hydrogen bond between 309N (pale green sticks) and TES. On the right, formestane (pale yellow sticks) docking pose within WT. The dashed yellow lines represent the hydrogen bond between 309D (pale green sticks) and formestane.

3.2 Study of interaction between ZEN, αZEL and aromatase variants

Once the procedural performances have been validated (see section 3.1), the interaction of ZEN with the following variants was investigated: A306T, D309G, F221S, M374I, M374L, M374T, R115Q, S478F, T310A, T310I, T310S, V370A, and V373F. The interaction of ZEN with V370M and D309N variants was not calculated instead being these variants inactive. Furthermore, it has been calculated the capability of αZEL to inhibit WT and those predicted active variants with a theoretically diverse sensitivity of being inhibited by ZEN with respect to the WT.

3.2.1. Interaction between ZEN and aromatase variants

The interaction of ZEN has been calculated along with that of TES in all the variants listed above via docking simulations to make a short list of complexes to investigate further with MD simulations. The selection criterion was based on the docking score assignment and respective

variations observed in the aromatase variants compared to WT, assuming that the variations magnitude may reflect the impact of amino acid substitution in the capability to bind ZEN, in agreement with previous studies (Dorne et al., 2022; Louisse et al., 2022). Specifically, docking scores may correlate with the capability of ligands to fit the protein pocket (i.e. the higher the score, the better the fitting into the pocket), as previously described (Ji et al., 2020; Louisse et al., 2022). Therefore, all the 13 collected variants were analysed through molecular docking and the docking scores compared to that recorded for the WT enzyme. Then, those complexes having mutations expected to be most effective (scores variations above 10%; threshold arbitrarily set, namely S478F, D309G and T310I) have been further studied with MD simulations (see Table 1; Figure 2B-K). The T310S was also analysed although the % variation was below 10% as it was previously described having a diverse susceptibility to inhibition (Kao et al., 1998).

T310S. As for V370M, T310S was used also to assess model performances as it is more susceptible to NAR inhibition, as described in section 3.1. Concerning the interaction with ZEN (Figure 2J), docking results revealed a score slightly higher (+4.4%) compared to the WT. This might point to a better fitting into the enzyme pocket possibly resulting in an enhanced inhibitory activity. The ligand-T310S complex stability analysed thorough MD simulations highlighted a WT-like behaviour when in complex with TES while ZEN resulted as more stable in T310S than in WT (Figure 4A). Indeed, the T>S substitution likely leads to an enlargement of the pocket volume, due to the lower dimension of serine side chain compared to threonine, which resulted in a better accommodation of ZEN. Based on these findings, T310S could be associated with a possible increased susceptibility towards ZEN inhibition. It must be noted that the xenoestrogenic potential of ZEN depends - though is not limited to - on the concerted activation of ERs and inhibition of aromatase. Therefore, based on this outcome, it cannot be excluded that individuals bearing this variant might have a diverse endocrine disruption when exposed to ZEN compared to those homozygotes for the WT enzyme. Specifically, in subjects bearing the T310S variant, ZEN might have a higher depleting activity of estrogens production, the effects of which at the level of hormones homeostasis and related pathogenesis are worth of future investigations. Despite this SNV recorded a low frequency among human population (minor allele frequency < 0.0006; as per UniProt last access 26th September 2023), its occurrence, expression and frequency should be deeply investigated as fundamental for its relevance to ZEN toxicity in living organisms.

S478F. As shown in Figure 2F, the S>F substitution caused an improper arrangement of TES with respect to the WT where the interatomic distances between the atom undergoing the reaction and the heme's iron are kept significantly ($p < 0.001$) lower (average distance of 0.359 ± 0.032 and 0.455 ± 0.057 nm in WT and S478F, respectively) (Figure 5A). Therefore, this aromatase SNV has been considered not functional for the reason already discussed for V370M. Therefore, ZEN has not been calculated via MD simulation accordingly, although the negative docking score recorded pointed to the substantial incapability of ZEN to fit in. On this basis, S478F might be associated to aromatase deficiency and although it has been found having a low frequency among human population (minor allele frequency of 0.0002; as per UniProt, last access 26th September 2023) it is worthy of further analysis to characterize its occurrence, expression and activity.

D309G. This SNV recorded for TES a docking score comparable to WT, while it obtained a 16.7% score increase with ZEN, pointing to a possible higher susceptibility. However, MD simulations showed ZEN as less stable in D309G than in WT (Figure 5B), indicating a probable reduced ZEN inhibitory activity, while TES had a comparable behaviour. Of note, the D309 is a crucial residue for the electron transfer (see section 3.1) and considering the different physico-chemical properties of glycine with respect to glutamate, the D309G substitution might be related to aromatase deficiency. Moving to the SNV frequency, it is low (minor allele frequency < 0.00008; as per UniProt, last access 26th September 2023) but a better characterization of its occurrence, expression and activity should be carried out anyway.

T310I. The T>I substitution led to a WT-like arrangement of TES (Figure 2D) and ZEN (Figure 2I). However, the former recorded a 30.5% docking score decrease while the latter a 40.1% decrease possibly due to the different polarity of isoleucine and threonine. The MD simulations interestingly led to similar results compared to the WT, indeed both TES and ZEN recorded trends similar to the respective WT complexes (Figure 4B). Nevertheless, seen the drop in the docking score and the substitution of a residue with different physico-chemical properties, T310I was considered as less capable of binding both the natural substrates and ZEN, though its activity and inhibition would deserve a further quantitative assessment. Regarding its frequency, this is low (minor allele frequency < 0.00006; as per UniProt, last access 26th September 2023) but, as for the others considered SNVs, its occurrence, expression and activity should be furtherly analysed as fundamental for its relevance to ZEN toxicity in living organisms.

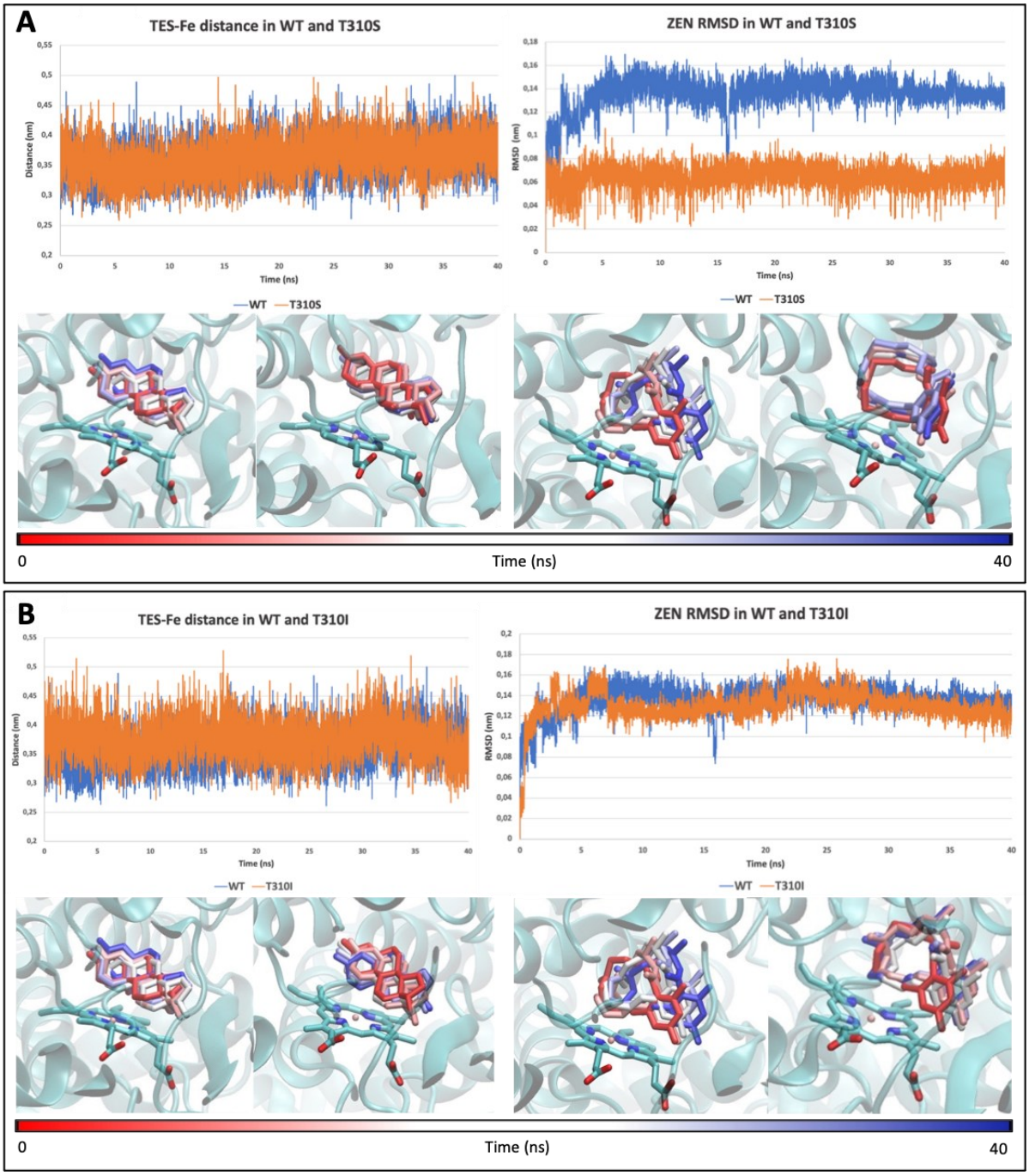


Figure 4. MD results on T310 variants compared to WT. The average distances between the atom undergoing reaction and heme's iron atom ± SD are reported between brackets. **A**. On the top-left, interatomic distances between TES's atom undergoing reaction and heme's iron within WT (blue; 0.359 ± 0.032 nm) and T310S (orange; 0.364 ± 0.033 nm). On the bottom-left, time-step representation (from red, 0 ns, to blue, 40 ns) of the trajectories of TES within WT and T310S. On the top-right, ZEN RMSD within WT (blue; 0.136 ± 0.013 nm) and T310S (orange; 0.065 ± 0.009 nm). On the bottom-right, time-step representation (from red, 0 ns, to blue, 40 ns) of the trajectories of ZEN within WT and T310S. **B**. On the top-left, interatomic distances

between TES's atom undergoing reaction and heme's iron within WT (blue; 0.359 ± 0.032 nm) and T310I (orange; 0.376 ± 0.033 nm). On the bottom-left, time-step representation (from red, 0 ns, to blue, 40 ns) of the trajectories of TES within WT and T310I. On the top-right, ZEN RMSD within WT (blue; 0.136 ± 0.013 nm) and T310I (orange; 0.130 ± 0.015 nm). On the bottom-right, time-step representation (from red, 0 ns, to blue, 40 ns) of the trajectories of ZEN within WT and T310I.

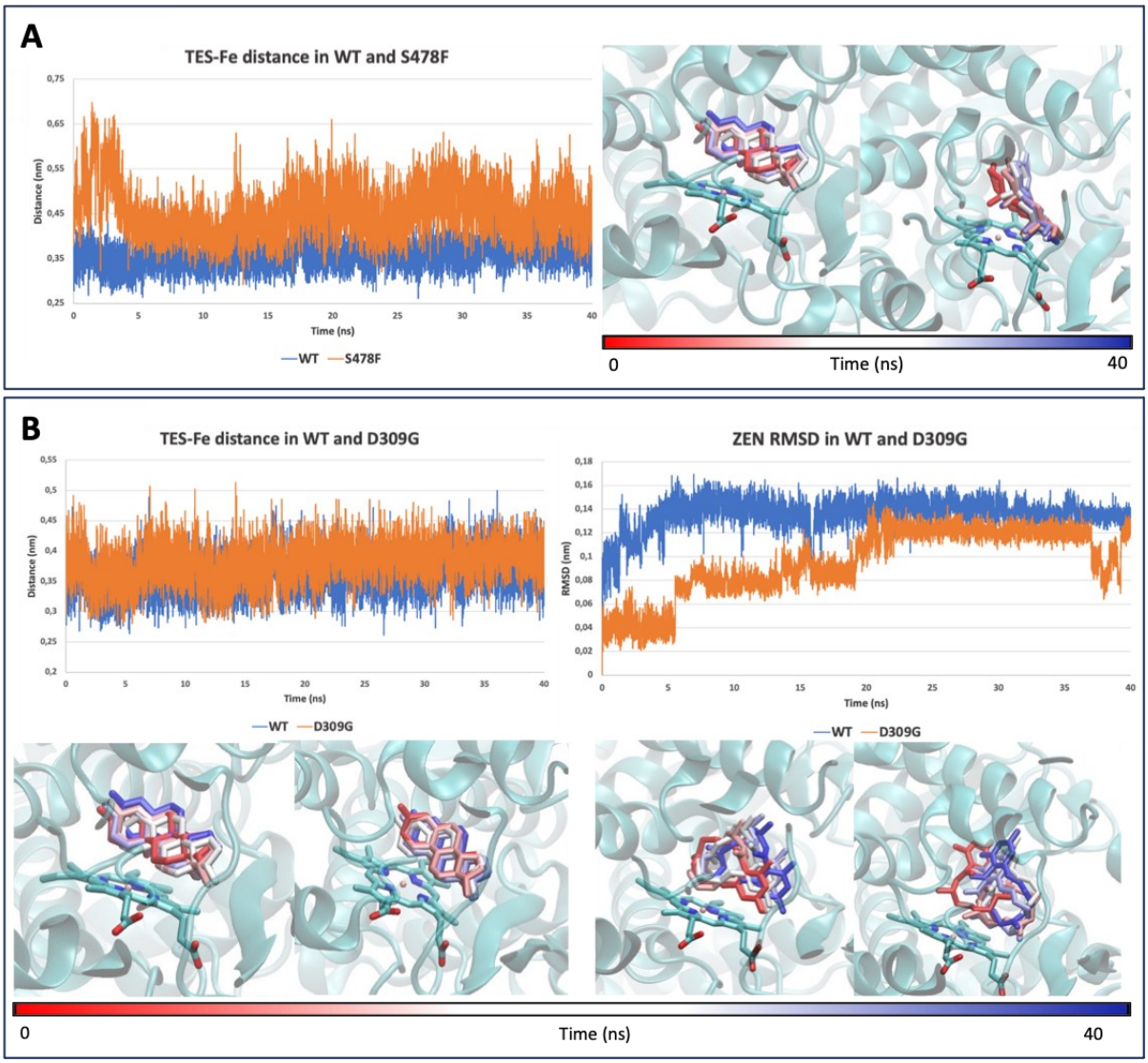


Figure 5. MD results on S478F and D309G compared to WT. **A**. On the left, interatomic distances between TES's atom undergoing reaction and heme's iron within WT (blue; average value ± SD of 0.359 ± 0.032 nm) and S478F (orange; average value ± SD of 0.455 ± 0.057 nm). On the right, time-step representation (from red, 0 ns, to blue, 40 ns) of the trajectories of TES within WT and S478F. **B**. On the top-left, interatomic distances between TES's atom

undergoing reaction and heme's iron within WT (blue; average value ± SD of 0.359 ± 0.032 nm) and D309G (orange; average value ± SD of 0.381 ± 0.032 nm). On the bottom-left, time-step representation (from red, 0 ns, to blue, 40 ns) of the trajectories of TES within WT and D309G. On the top-right, ZEN RMSD within WT (blue; average value ± SD of 0.136 ± 0.013 nm) and D309G (orange; average value ± SD of 0.095 ± 0.029 nm). On the bottom-right, time-step representation (from red, 0 ns, to blue, 40 ns) of the trajectories of ZEN within WT and D309G.

3.3. Interaction between αZEL T310S and T310I aromatase variants

Previous evidence described αZAL, a close analogue of αZEL, as able to inhibit aromatase though with a lower efficacy compared to ZEN (Y. F. Wang et al., 2014), suggesting a certain degree of inhibitory potential for αZEL as well. Moreover, αZEL is a more potent agonist of ER compared to ZEN (Chain, 2014). The stronger activation of ERs and the expected weaker inhibition of aromatase (hypothesized based on αZAL data) are likely part of the mechanisms underpinning the stronger estrogenic insult αZEL may have in certain systems compared to ZEN. Therefore, studying whether SNV may also vary the susceptibility of aromatase to αZEL inhibition is important to better understand the general mechanics of ZEN estrogenicity. This is particularly significant since αZEL is one of the most abundant ZEN phase I metabolite in humans. The analysis was focused on WT aromatase an on those variants who gave evidence of a predicted diverse though not null susceptibility to ZEN compared to the WT (i.e. T310I and T310S).

As shown in Table 1, αZEL recorded a lower score within the WT compared to ZEN (36.55 and 45.92 units, respectively), in line with the weaker inhibitory activity described previously for its structural analogue αZAL (Y. F. Wang et al., 2014). This confirmed that αZEL may eventually have a reduced inhibitory activity, as hypothesized above. The score decrease compared to ZEN may be due to the slightly diverse arrangement αZEL showed into the pocket (Figure 6). Moreover, αZEL recorded a 21.2% higher and 14.1% lower score compared to the WT within T310S and T310I, respectively, suggesting that αZEL may preferentially interact and inhibit T310S than the WT, as described for ZEN above. The analysis of trajectories revealed that αZEL could stably stay at the ligand binding site of WT, T310S and T310I (Figure 6). However, the RMSD trend observed within T310S was more stable compared to that within WT. Concerning T310I, αZEL showed an RMSD trend like that observed within the WT. Therefore, in line with the interpretation made for ZEN, these outcomes could suggest that T310S might be more susceptible to αZEL inhibition compared to the WT. Importantly, this and

the associated reduced production of estrogens need further dedicated investigations to prove whether T310S bearing subjects may "benefit" of a mitigated estrogenic insult when exposed to αZEL or ZEN – considering ZEN prominent conversion to αZEL. The verification of such hypothesis is crucial considering that the estrogenic effect of ZEN and metabolites, including αZEL, results from the concerted activation of ER and aromatase inhibition, as discussed above.

Table 1. Docking results for ZEN and TES into WT aromatase and a selection of its variants

Aromatase	**Ligand**	**GOLDScore**	**% variation to the WT [1]**
Wild-type	TES	63.15	---
	ZEN	45.92	---
	αZEL	36.55	---
A306T	TES	64.45	-2.1
	ZEN	41.65	9.3
D309G	TES	60.96	3.5
	ZEN	53.58	-16.7
F221S	TES	61.96	1.9
	ZEN	48.68	-6.0
M374I	TES	61.6	2.5
	ZEN	46.16	-0.5
M374L	TES	63.05	0.2
	ZEN	45.78	0.3
M374T	TES	62.86	0.5
	ZEN	45.92	0.0
R115Q	TES	60.54	4.1
	ZEN	45.60	0.7
S478F	TES	19.98	68.4
	ZEN	-35.25	---
T310A	TES	60.68	3.9
	ZEN	45.97	-0.1
T310I	TES	39.05	38.2
	ZEN	31.90	30.5
	αZEL	31.40	14.1

V370A	TES	59.12	6.4
	ZEN	43.35	5.6
V373F	TES	65.06	-3.0
	ZEN	48.04	-4.6
T310S [2]	TES	61.20	3.1
	ZEN	47.92	-4.4
	αZEL	44.30	-21.2

Note: [1] the sign "-" indicates scores higher compared to the those observed into the wild-type enzyme. [2] The complex with T310S was analysed through MD simulations although the % variation was below 10% because previous evidence reported the effects of such substitution in terms of susceptibility to inhibition (Kao et al., 1998). Complexes with a % variation above 10% were analysed through MD simulations and are indicated in bold.

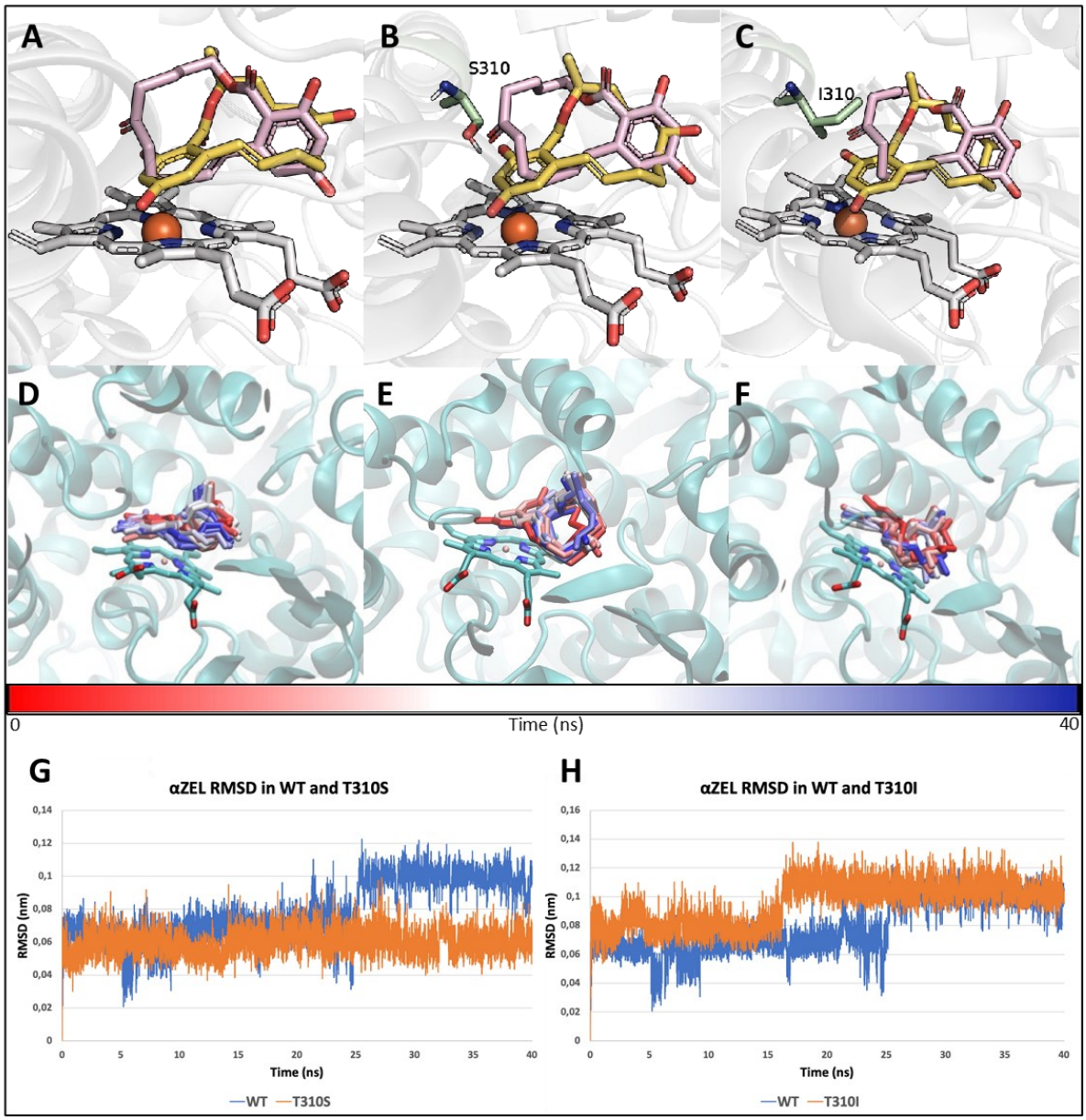


Figure 6. Molecular docking and MD simulations results of αZEL within T310S and T310I compared to WT. **A.** Docking pose superimposition of ZEN (pink sticks) and αZEL (yellow sticks) within WT. **B.** Docking pose superimposition of ZEN (pink sticks) and αZEL (yellow sticks) within T310S. **C.** Docking pose superimposition of ZEN (pink sticks) and αZEL (yellow sticks) within T310I. **D.** Time-step representation (from red, 0 ns, to blue, 40 ns) of the trajectories of αZEL within WT. **E.** Time-step representation (from red, 0 ns, to blue, 40 ns) of the trajectories of αZEL within T310S. **F.** Time-step representation (from red, 0 ns, to blue, 40 ns) of the trajectories of αZEL within T310I. **G.** αZEL RMSD within WT (blue; average value ± SD of 0.078 ± 0.018 nm) and T310S (orange; average value ± SD of 0.059 ± 0.008 nm). **H.** αZEL RMSD within WT (blue; average value ± SD of 0.078 ± 0.018 nm) and T310I (orange; average value ± SD of 0.095 ± 0.016 nm).

4. Conclusions

This work dealt with the analysis of aromatase SNVs with respect to their susceptibility to be inhibited by ZEN and αZEL. The analysis is framed into the context of ZEN toxicodynamic and its variability among the human population. These investigations are fundamental to improve the current understanding of ZEN toxicology, also paving the way for a more informed study of the ZEN-dependent pathogenic potential at an individual level. Our work coped with the SNVs available at the time of analysis focusing on a set of variants that have been thoroughly investigated using a validated and well-consolidated molecular modelling approach. The S478F has been described likely having a severely impaired capacity to receive substrates, suggesting its reduced/null catalytic activity and possible involvement in aromatase deficiency. The possible association to a reduced/null activity has been also described for D309G, for which an impairing effect of the D>G substitution has been postulated based on the key role of D309 for catalysis. The T310I was calculated having a minor role to impair the capability to recruit and interact with TES and ZEN. This might suggest a susceptibility of being inhibited by ZEN comparable to the WT, though expression and activity should be further analyzed for a better characterization of this variant. Finally, the T310S variant was found more susceptible to ZEN inhibition, in line with the effect previously described for NAR, while keeping its activity to convert endogenous substrates (Kao et al., 1998). This suggested that those individuals bearing this variant might be diversly affected by ZEN considering the postulated reduced capability of being inhibited and the comparable activity to convert endogenous substrates. Concerning αZEL, this study described that it may inhibit aromatase, though with an expected lower potency than ZEN, like its close structural analogue αZAL. Moreover, in line with the interpretation made for ZEN, T310S may have an enhanced susceptibility of being inhibited compared to the WT. Taken together the results collected for ZEN and αZEL highlighted that the aromatase inhibitory capacity of ZEN must be investigated along with that of its congeners and metabolites, also with respect to their relative production in living organisms and respective inhibitory activity towards aromatase SNVs. This should be carefully evaluated toward a better understanding and assessment of ZEN and congeners as a group of toxicants rather than single substances, as outlined by EFSA (Alexander et al., 2016).

As a general comment, the present work was meant to prioritize SNVs variant for further analysis toward a better comprehension of ZEN toxicity among the human population. Therefore, this work proposed for further dedicate investigations T310S for the possible increased susceptibility to ZEN and αZEL inhibition, and D309N and S478F for the suspected

loss of function. Also, it provided a mechanistic rationale explaining, at least in part, the inactivity of V370M and D309N variants, concurring to improve the basic understanding of their associations to aromatase deficiency in humans.

References

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX, 1-2*, 19-25. doi:https://doi.org/10.1016/j.softx.2015.06.001

Alexander, J., Barregard, L., Bignami, M., Ceccatelli, S., Cottrill, B., Dinovi, M., . . . Chain, E. P. C. F. (2016). Appropriateness to set a group health-based guidance value for zearalenone and its modified forms. *Efsa Journal, 14*(4). doi:10.2903/j.efsa.2016.4425

Bateman, A., Martin, M. J., Orchard, S., Magrane, M., Ahmad, S., Alpi, E., . . . UniProt, C. (2023). UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research, 51*(D1), D523-D531. doi:10.1093/nar/gkac1052

Bell, E. W., & Zhang, Y. (2019). DockRMSD: an open-source tool for atom mapping and RMSD calculation of symmetric molecules through graph isomorphism. *Journal of Cheminformatics, 11*. doi:10.1186/s13321-019-0362-7

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., . . . Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research, 28*(1), 235-242. doi:10.1093/nar/28.1.235

Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., . . . Karplus, M. (2009). CHARMM: The Biomolecular Simulation Program. *Journal of Computational Chemistry, 30*(10), 1545-1614. doi:10.1002/jcc.21287

Catteuw, A., Broekaert, N., De Baere, S., Lauwers, M., Gasthuys, E., Huybrechts, B., . . . Croubels, S. (2019). Insights into In Vivo Absolute Oral Bioavailability, Biotransformation, and Toxicokinetics of Zearalenone, alpha-Zearalenol, beta-Zearalenol, Zearalenone-14-glucoside, and Zearalenone-14-sulfate in Pigs. *Journal of Agricultural and Food Chemistry, 67*(12), 3448-3458. doi:10.1021/acs.jafc.8b05838

Chain, E. P. o. C. i. t. F. (2014). Scientific Opinion on the risks for human and animal health related to the presence of modified forms of certain mycotoxins in food and feed. *Efsa Journal, 12*(12), 3916. doi:https://doi.org/10.2903/j.efsa.2014.3916

Claeys, L., Romano, C., De Ruyck, K., Wilson, H., Fervers, B., Korenjak, M., . . . Huybrechts, I. (2020). Mycotoxin exposure and human cancer risk: A systematic review of epidemiological studies. *Comprehensive Reviews in Food Science and Food Safety, 19*(4), 1449-1464. doi:10.1111/1541-4337.12567

De Boevre, M., Jacxsens, L., Lachat, C., Eeckhout, M., Di Mavungu, J. D., Audenaert, K., . . . De Saeger, S. (2013). Human exposure to mycotoxins and their masked forms through cereal-based foods in Belgium. *Toxicology Letters, 218*(3), 281-292. doi:10.1016/j.toxlet.2013.02.016

Dellafiora, L., Magnaghi, F., Galaverna, G., & Dall'Asta, C. (2022). A mechanistic investigation on kokumi-active gamma-Glutamyl tripeptides - A computational study to understand molecular basis of their activity and to identify novel potential kokumi-tasting sequences. *Food Research International, 162*. doi:10.1016/j.foodres.2022.111932

Di Nardo, G., Breitner, M., Bandino, A., Ghosh, D., Jennings, G. K., Hackett, J. C., & Gilardi, G. (2015). Evidence for an Elevated Aspartate pK(a) in the Active Site of Human Aromatase. *Journal of Biological Chemistry, 290*(2), 1186-1196. doi:10.1074/jbc.M114.595108

Dorne, J., Cirlini, M., Louisse, J., Pedroni, L., Galaverna, G., & Dellafiora, L. (2022). A Computational Understanding of Inter-Individual Variability in CYP2D6 Activity to Investigate the Impact of Missense Mutations on Ochratoxin A Metabolism. *Toxins, 14*(3). doi:10.3390/toxins14030207

EFSA. (2011). Scientific Opinion on the risks for public health related to the presence of zearalenone in food. *EFSA Journal, 9*(6), 2197. doi:10.2903/j.efsa.2011.2197

Ekwomadu, T., Mwanza, M., & Musekiwa, A. (2022). Mycotoxin-Linked Mutations and Cancer Risk: A Global Health Issue. *International Journal of Environmental Research and Public Health, 19*(13). doi:10.3390/ijerph19137754

Ghosh, D., Lo, J., Morton, D., Valette, D., Xi, J. L., Griswold, J., . . . Davies, H. M. L. (2012). Novel Aromatase Inhibitors by Structure-Guided Design. *Journal of Medicinal Chemistry, 55*(19), 8464-8476. doi:10.1021/jm300930n

Gupta, R. C., Doss, R. B., Lall, R., Srivastava, A., & Sinha, A. (2022). Chapter 49 - Trichothecenes and zearalenone. In R. C. Gupta (Ed.), *Reproductive and Developmental Toxicology (Third Edition)* (pp. 1003-1016): Academic Press.

IARC. (1993). IARC monographs on the evaluation of carcinogenic risks to humans. *IARC monographs, 56*, 599.

Ji, D. W., Xu, M., Udenigwe, C. C., & Agyei, D. (2020). Physicochemical characterisation, molecular docking, and drug-likeness evaluation of hypotensive peptides encrypted in flaxseed proteome. *Current Research in Food Science, 3*, 41-50. doi:10.1016/j.crfs.2020.03.001

Jing, S. Y., Liu, C. M., Zheng, J., Dong, Z. J., & Guo, N. (2022). Toxicity of zearalenone and its nutritional intervention by natural products. *Food & Function, 13*(20), 10374-10400. doi:10.1039/d2fo01545e

Kao, Y. C., Zhou, C. B., Sherman, M., Laughton, C. A., & Chen, S. (1998). Molecular basis of the inhibition of human aromatase (estrogen synthetase) by flavone and isoflavone phytoestrogens: A site-directed mutagenesis study. *Environmental Health Perspectives, 106*(2), 85-92. doi:10.1289/ehp.9810685

Katzenellenbogen, J. A., Mayne, C. G., Katzenellenbogen, B. S., Greene, G. L., & Chandarlapaty, S. (2018). Structural underpinnings of oestrogen receptor mutations in endocrine therapy resistance. *Nature Reviews Cancer, 18*(6), 377-388. doi:10.1038/s41568-018-0001-z

Kim, S., Chen, J., Cheng, T. J., Gindulyte, A., He, J., He, S. Q., . . . Bolton, E. E. (2023). PubChem 2023 update. *Nucleic Acids Research, 51*(D1), D1373-D1380. doi:10.1093/nar/gkac956

Knutsen, H. K., Alexander, J., Barregård, L., Bignami, M., Brüschweiler, B., Ceccatelli, S., . . . Chain, E. P. C. F. (2017). Risks for animal health related to the presence of zearalenone and its modified forms in feed. *Efsa Journal, 15*(7). doi:10.2903/j.efsa.2017.4851

Kowalska, K., Habrowska-Gorczynska, D. E., & Piastowska-Ciesielska, A. W. (2016). Zearalenone as an endocrine disruptor in humans. *Environmental Toxicology and Pharmacology, 48*, 141-149. doi:10.1016/j.etap.2016.10.015

Lephart, E. D. (2015). Modulation of Aromatase by Phytoestrogens. *Enzyme Research, 2015*, 594656. doi:10.1155/2015/594656

Lo, E. K. K., Wang, X. W., Lee, P. K., Wong, H. C., Lee, J. C. Y., Gomez-Gallego, C., . . . Li, J. (2023). Mechanistic insights into zearalenone-accelerated colorectal cancer in mice using integrative multi-omics approaches. *Computational and Structural Biotechnology Journal, 21*, 1785-1796. doi:10.1016/j.csbj.2023.02.048

Lo, J., Di Nardo, G., Griswold, J., Egbuta, C., Jiang, W. H., Gilardi, G., & Ghosh, D. (2013). Structural Basis for the Functional Roles of Critical Residues in Human Cytochrome P450 Aromatase. *Biochemistry, 52*(34), 5821-5829. doi:10.1021/bi400669h

Louisse, J., Dorne, J., & Dellafiona, L. (2022). Investigating the interaction between organic anion transporter 1 and ochratoxin A: An in silico structural study to depict early molecular events of substrate recruitment and the impact of single point mutations. *Toxicology Letters, 355*, 19-30. doi:10.1016/j.toxlet.2021.11.001

Ludwig, M., Beck, A., Wickert, L., Bolkenius, U., Tittel, B., Hinkel, K., & Bidlingmaier, F. (1998). Female Pseudohermaphroditism Associated with a Novel Homozygous G-to-A (V370-to-M) Substitution in the P-450 Aromatase Gene. *11*(5), 657-664. doi:doi:10.1515/JPEM.1998.11.5.657

Malir, F., Pickova, D., Toman, J., Grosse, Y., & Ostry, V. (2023). Hazard characterisation for significant mycotoxins in food. *Mycotoxin Research*. doi:10.1007/s12550-023-00478-2

Martin, G. D. A., Narvaez, J., Bulmer, R., & Durrant, M. C. (2016). Biotransformation and molecular docking studies of aromatase inhibitors. *Steroids, 113*, 95-102. doi:10.1016/j.steroids.2016.07.003

Massart, F., Meucci, V., Saggese, G., & Soldani, G. (2008). High growth rate of girls with precocious puberty exposed to estrogenic mycotoxins. *Journal of Pediatrics, 152*(5), 690-695. doi:10.1016/j.jpeds.2007.10.020

Massart, F., & Saggese, G. (2010). Oestrogenic mycotoxin exposures and precocious pubertal development. *International Journal of Andrology, 33*(2), 369-376. doi:10.1111/j.1365-2605.2009.01009.x

Murthy, J. N., Nagaraju, M., Sastry, G. M., Rao, A. R., & Sastry, G. N. (2005). Active site acidic residues and structural analysis of modelled human aromatase: A potential drug target for breast cancer. *Journal of Computer-Aided Molecular Design, 19*(12), 857-870. doi:10.1007/s10822-005-9024-0

Panneerselvam, S., Yesudhas, D., Durai, P., Anwar, M. A., Gosu, V., & Choi, S. (2015). A Combined Molecular Docking/Dynamics Approach to Probe the Binding Mode of Cancer Drugs with Cytochrome P450 3A4. *Molecules, 20*(8), 14915-14935. doi:10.3390/molecules200814915

Pedroni, L., Louisse, J., Dorne, J., Dall'Asta, C., & Dellafiora, L. (2023). A computational study on the biotransformation of alkenylbenzenes by a selection of CYPs: Reflections on their possible bioactivation. *Toxicology, 488*. doi:10.1016/j.tox.2023.153471

Pedroni, L., Louisse, J., Punt, A., Dorne, J., Dall'Asta, C., & Dellafiora, L. (2023). A Computational Inter-Species Study on Safrole Phase I Metabolism-Dependent Bioactivation: A Mechanistic Insight into the Study of Possible Differences among Species. *Toxins, 15*(2). doi:10.3390/toxins15020094

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry, 25*(13), 1605-1612. doi:10.1002/jcc.20084

Pfleger, F., & Schwake-Anduschus, C. (2023). Relevance of Zearalenone and its modified forms in bakery products. *Mycotoxin Research, 39*(3), 153-163. doi:10.1007/s12550-023-00493-3

Rai, A., Das, M., & Tripathi, A. (2020). Occurrence and toxicity of a fusarium mycotoxin, zearalenone. *Critical Reviews in Food Science and Nutrition, 60*(16), 2710-2729. doi:10.1080/10408398.2019.1655388

Rivera-Nunez, Z., Barrett, E. S., Szamreta, E. A., Shapses, S. A., Qin, B., Lin, Y., . . . Bandera, E. V. (2019). Urinary mycoestrogens and age and height at menarche in New Jersey girls. *Environmental Health, 18*. doi:10.1186/s12940-019-0464-8

Rizner, T. L., & Romano, A. (2023). Targeting the formation of estrogens for treatment of hormone dependent diseases-current status. *Frontiers in Pharmacology, 14*. doi:10.3389/fphar.2023.1155558

Rong, X., Jiang, Y., Li, F., Sun-Waterhouse, D., Zhao, S. C., Guan, X. D., & Li, D. P. (2022). Close association between the synergistic toxicity of zearalenone-deoxynivalenol combination and microRNA221-mediated PTEN/PI3K/AKT signaling in HepG2 cells. *Toxicology, 468*. doi:10.1016/j.tox.2022.153104

Russell, J. F., Wong, J. C., & Grumbach, M. M. (2014). Chapter 3J - Aromatase Deficiency and Aromatase Excess. In M. I. New, O. Lekarev, A. Parsa, T. T. Yuen, B. W. O'Malley, & G. D. Hammer (Eds.), *Genetic Steroid Disorders* (pp. 165-190). San Diego: Academic Press.

Silva, A. S., Brites, C., Pouca, A. V., Barbosa, J., & Freitas, A. (2019). UHPLC-ToF-MS method for determination of multi-mycotoxins in maize: Development and validation. *Current Research in Food Science, 1*, 1-7. doi:10.1016/j.crfs.2019.07.001

Simpson, E. R. (2000). Genetic mutations resulting in loss of aromatase activity in humans and mice. *Journal of the Society for Gynecologic Investigation, 7*(1), S18-S21. doi:10.1177/1071557600007001s07

Wang, Y., Liu, T. J., Xie, J. H., Cheng, M. J., Sun, L. R., Zhang, S., . . . Zhang, N. (2022). A review on application of molecular simulation technology in food molecules interaction. *Current Research in Food Science, 5*, 1873-1881. doi:10.1016/j.crfs.2022.10.012

Wang, Y. F., Wong, T. Y., Chan, F. L., Chen, S., & Leung, L. K. (2014). Assessing the effect of food mycotoxins on aromatase by using a cell-based system. *Toxicology in Vitro, 28*(4), 640-646. doi:10.1016/j.tiv.2014.01.012

Zhang, L., Silva, D. A., Yan, Y. J., & Huang, X. H. (2012). Force field development for cofactors in the photosystem II. *Journal of Computational Chemistry, 33*(25), 1969-1980. doi:10.1002/jcc.23016

Zhao, L., Zhang, M. X., Pan, F., Li, J. Y., Dou, R., Wang, X. Y., . . . Cai, S. B. (2021). In silico analysis of novel dipeptidyl peptidase-IV inhibitory peptides released from Macadamia integrifolia antimicrobial protein 2 (MiAMP2) and the possible pathways involved in diabetes protection. *Current Research in Food Science, 4*, 603-611. doi:10.1016/j.crfs.2021.08.008

Zoete, V., Cuendet, M. A., Grosdidier, A., & Michielin, O. (2011). SwissParam: A Fast Force Field Generation Tool for Small Organic Molecules. *Journal of Computational Chemistry, 32*(11), 2359-2368. doi:10.1002/jcc.21816

Chapter IV

Virtual display of targets: A new level to rise the current understanding of ochratoxin A toxicity from a molecular standpoint

This work is based on the paper published on:

Florinda Perugino, Lorenzo Pedroni, Gianni Galaverna, Chiara Dall'Asta, Luca Dellafiora. (2024). Virtual display of targets: A new level to rise the current understanding of ochratoxin A toxicity from a molecular standpoint. Toxicology, 503, 153686, https://doi.org/10.1016/j.tox.2024.153765.

Author contribution: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis

Chapter IV

Abstract

Ochratoxin A (OTA) is a mycotoxin spread worldwide contaminating several food and feed commodities and rising concerns for humans and animals. OTA toxicity has been thoroughly assessed over the last 60 years revealing a variety of adverse effects, including nephrotoxicity, hepatotoxicity and possibly carcinogenicity. However, the underpinning mechanisms of action have yet to be completely displayed and understood. In this framework, we applied a virtual pipeline based on molecular docking, dynamics and umbrella simulations to display new OTA potential targets. The results collected consistently identified OGFOD1, a key player in protein translation, as possibly inhibited by OTA and its 2'R diastereomer. This is consistent with the current knowledge of OTA's molecular toxicology and may fill some gaps from a mechanistic standpoint. This could pave the way for further dedicated analysis focusing their attention on the OTA-OGFOD1 interaction, expanding the current understanding of OTA toxicity at a molecular level.

1. Introduction

Ochratoxin A (OTA; Figure 1) is a well-known mycotoxin mainly produced by fungi belonging to *Penicillium* and *Aspergillus* genera. It contaminates food and feed worldwide rising concerns for human and animal health (Vettorazzi et al., 2014). Over the last 60 years, OTA toxicity has been widely investigated and thoroughly described. Nowadays, OTA has been shown to be nephrotoxic, hepatotoxic, teratogenic and immunotoxic to several animal species and to cause kidney and liver tumours in mice and rats (Tao et al., 2018). Also, the International Agency for Research on Cancer (IARC) classified it as possibly carcinogenic to humans (group 2B) (IARC, 1993). Along with OTA, certain food commodities can be contaminated also by 2'R-OTA, an OTA diastereomer known to be produced upon thermal treatments like roasting and baking of OTA-containing food (Bryla et al., 2021; Cramer et al., 2015). Of note, 2'R-OTA seems to be less toxic compared to its diastereomer (Cramer et al., 2008) although certain population groups may be highly exposed depending on the diet habits (e.g. coffee drinkers) (Cramer et al., 2015). From a cellular standpoint, OTA adverse effects have been associated to a wealth of events including but not limited to the impairment of protein synthesis and energy production, induction of oxidative/nitrosative stress, apoptosis, and effects on mitosis and cell cycle regulation (Koszegi & Poór, 2016). Specifically, these effects might occur via several pathways such as nucleotide metabolism, RNA synthesis and many more (Liang et al., 2015). Nonetheless, critical pieces of OTA toxicity, including the full elucidation of its biological targets and underpinning mechanisms of action, needs further clarification toward a better understanding of its complex network of interactions. Indeed, the incomplete understanding of OTA toxicity from a molecular standpoint ultimately results in a non-optimal assessment of the risk associated to its actual exposure in everyday scenarios. In order to find novel OTA targets, the present study relied on the outcome of a previous target fishing campaign (Dellafiora et al., 2020) that identified the strong chemical similarity between OTA and the Protein Data Bank (PDB; https://www.rcsb.org) (Berman et al., 2000) ligand UN9 (Figure 1), which is an inhibitor of the human OGFOD1 protein (Horita et al., 2015). OGFOD1 is a protein kinase with a variety of targets and functions, with a well-documented crucial role to ensure the progression of protein synthesis (Thinnes et al., 2019). The chemical analogies between OTA and OGFOD1 inhibitors may enable the former to act likewise, as widely demonstrated for other compounds which have been functionally described based on their chemical similarity with previously characterized molecules (Park et al., 2023). Interestingly, the OGFOD1 inhibition by OTA might uncover some of the still unknown mechanisms of this mycotoxin, including those related

to the protein synthesis impairment. Therefore, we applied a molecular modelling pipeline which previously succeeded to predict the inhibitory potential of ligands (Del Favero et al., 2022; Pedroni et al., 2022; Perugino et al., 2024) to estimate the OGFOD1 inhibition by both OTA and 2'R-OTA (Figure 1). In particular, the procedure eventually estimates the free energy of binding as it already proved to be a valuable mean to calculate the inhibitory activity of compounds being related to binding affinity and inhibitory capacity (Cournia et al., 2017; Kuppusamy et al., 2017). The study revealed that both OTA isomers are likely to interact with OGFOD1. The outcomes collected here and their implications to improve the current understanding of OTA mechanisms of toxicity have been thoroughly discussed.

2. Methods

2.1. Retrieval of 3D information for ligands and protein

The 3D structure of UN9 (N-[1-Chloro-4-Hydroxysoquinolin-3-yl)carbonyl]glycine) was retrieved from the Chemical Component Dictionary of PDB, while the 3D structure of OTA was retrieved from PubChem (Kim et al., 2023) (CID 442530). The 3D structure of 2'R-OTA was generated inverting the stereochemistry of OTA with the *invert* function on PyMol (version 2.5). The 3D protein model was derived from the crystallographic structure with PDB ID 4NHX (Horita et al., 2015) keeping only the domain interacting with the co-crystallized ligand (residues 24-237). The ligands with the PDB code OGA (N-Oxalyglycine) and PD2 (Pyridine-2,4-Dicarboxylic Acid) (Figure 1) were also retrieved from the Chemical Component Dictionary of PDB to serve as additional positive controls being described to inhibit OGFOD1 (Horita et al., 2015). The protonation state of OTA when arranged close to the iron cation was calculated using CHARMM36 via the SwissParam webserver (http://www.swissparam.ch) (Bugnon et al., 2023) and using the multipurpose atom-typer for CHARMM (MATCH) approach. Specifically, the coordinates of OTA calculated by the docking study (see below) were saved as .mol2 file along with those of Fe and used as input for the SwissParam webserver choosing MATCH as selected approach.

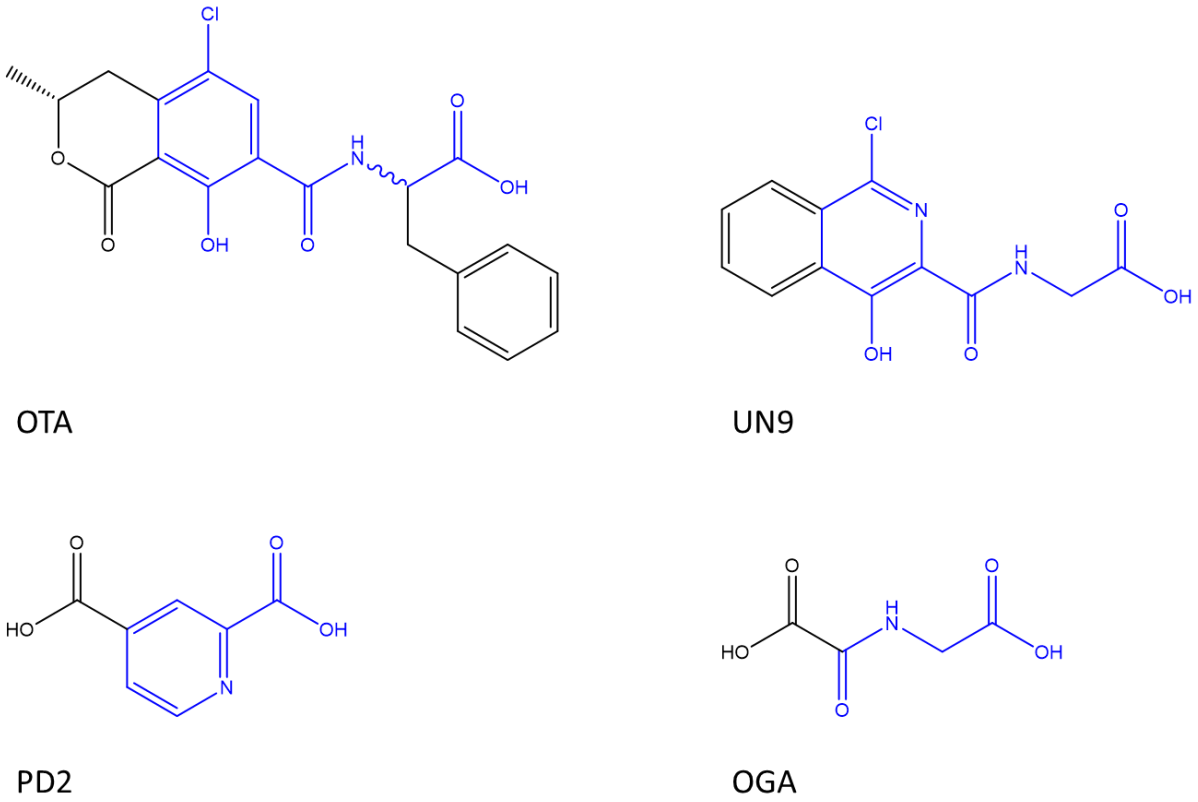


Figure 1. Chemical structure of molecules under analysis. From the top left, N-Oxalyglycine (OGA), Pyridine-2,4-Dicarboxylic Acid (PD2) and N-[1-Chloro-4-Hydroxysoquinolin-3-yl)carbonyl]glycine (UN9). From the bottom left, Ochratoxin A (OTA) and its stereoisomer (2'R-OTA). The OGA, PD2 and UN9 common parts to OTA are coloured in blue.

2.2. Modelling studies

For the sake of a fit-for-purpose validation aimed at confirming the procedure capability to provide reliable architecture of binding, the crystallographic pose of the known inhibitors UN9, OGA and PD2, taken as positive controls, were compared to those calculated by docking simulations. Then, the docked architectures of binding of OTA and 2'R-OTA were also investigated and analysed with respect to those of the positive controls mentioned above. Docking simulations were performed using the software GOLD (Genetic Optimization for Ligand Docking; version 2022) as it already succeeded in estimating reliable ligands architectures of binding (Dellafiora et al., 2020; Pedroni et al., 2022). The internal scoring function GOLDScore was used providing scores proportional to the fitting of ligands within the protein pocket (the higher the score, the better the fitting), while negative scores typically indicate unfavourable interactions and steric clashes, as per manufacturer declarations

(https://www.ccdc.cam.ac.uk). The docking protocol was set to favour the formation of the same hydrogen bonds displayed in the OGFOD1 PDB structure 4NHX to facilitate the generation of reliable architectures of binding. Of note, the Mn ion present within the crystal structure was replaced by Fe, being the natural OGFOD1 cofactor (Horita et al., 2015). The best scored binding poses obtained by the docking simulations were then used as input for molecular dynamics (MD) simulations performed by means of GROMACS (v. 2019.4) (Abraham et al., 2015) to analyse the protein-ligand complexes evolution along a 50 ns interval. Firstly, the complexes were parametrized with the CHARMM36 all-atom force field, and each ligand parametrization was performed on the SwissParam tool (http://swissparam.ch). Protein-ligands complexes were then solvated with SPCE water (roughly 10 000 water molecules each) in a cubic periodic boundary condition and counter ions (11 Na+ per system) were added to neutralize the net system charge. Each system was energetically minimised to avoid steric clashes and correct improper geometries using the steepest descent algorithm with a maximum of 5000 steps. Lastly, each system underwent isothermal (300 K, coupling time 2 ps) and isobaric (1 bar, coupling time 2 ps) 100 ps simulations before running 50 ns long MD simulations. Afterwards, all the complexes underwent umbrella simulations (US) by using GROMACS to study their ΔG. Each complex was oriented such that the binding pocket section exposed to the solvent was along the Y-axis. Then, each ligand was pulled out from the binding site via a steered molecular dynamics simulation at a 0.01 nm/ps pull rate over 500 ps on the Y-axis applying a spring force constant of 750 kJ•mol^{-1} nm^{-2}. An asymmetric distribution of sampling windows was used such that spacing along the reaction coordinate was roughly 0.2 nm. Each US window underwent an NPT equilibration before running a 10 ns simulation with a spring force constant of 750 kJ • mol^{-1} nm^{-2}.

3. Results and Discussion

Although the reliability of the procedure used here was proven in several previous studies (Dellafiora et al., 2020; Pedroni et al., 2022), a fit-for-purpose validation was performed to assess the case-specific effectiveness to produce reliable architectures of binding. To do so, the docking poses of three known OGFOD1 inhibitors, i.e. UN9 (co-crystallized within 4NHM), OGA (co-crystallized within 4NHX) and PD2 (co-crystallized within 4NHY) (Horita et al., 2015), were compared to their respective crystallographic binding architectures.

As shown in Figure 2, the calculated pose of each ligand was almost superimposed to the respective crystallographic architecture supporting the procedural effectiveness to provide

reliable poses. Of note, the docking procedure also positively scored each pose (Dellafiora et al., 2020) with UN9 obtaining the best score (172 units), followed by PD2 (152 units) and OGA (150 units) (Table 1). Interestingly, the calculated rank agreed with the inhibitory potency rank described experimentally (IC_{50} of 0.52 ± 0.08, 1.9 ± 1.0 and 2.3 ± 1.2 μM for UN9, OGA and PD2, respectively), with UN9 as the most potent and the other two with a comparable inhibitory activity (Horita et al., 2015) (Table 1). With respect to the present case study (i.e. an inhibitor-enzyme interaction resulting in the enzyme inhibition), this evidence suggested that the evaluation of ligand-pocket fitting via docking score assignment may estimate and be associated with the capability of ligands to inhibit the enzyme, as widely demonstrated for other similar case studies (Del Favero et al., 2022; Pedroni et al., 2022; Perugino et al., 2024). Moreover, the crystallographic binding architecture of UN9, OGA and PD2 showed a bidentate interaction with the Mn cation, which is expected to be maintained also when the natural cofactor Fe is present. This interaction was kept also in the calculated poses for all the ligands considered (including OTA isomers, see below). Once assessed the procedure reliability, OTA and 2'R-OTA underwent modelling studies via docking and MD simulations to estimate the interaction with the enzyme and its subsequent possible inhibition. Of note, besides the carboxylic moiety that is deprotonated in physiological conditions, OTA may undergo deprotonation of the phenol group at the coumarin-like moiety with a variable pKa depending on the context (Perry et al., 2003). In this study, the protonation state of OTA was estimated when arranged close to Fe cation as calculated by docking. The protonated phenol has been calculated as the most abundant at physiological pH (77% at pH 7.4). Therefore, in the present study, OTA was considered deprotonated and protonated at the carboxylic and phenolic group, respectively. Nevertheless, a more precise estimate of the pKa switch its phenolic group may undergo within the OTA-Fe-OGFOD1 complex is suggested for a thorough calculation of complex thermodynamic – which is however beyond the aim of this work. OTA and 2'R-OTA arranged likewise UN9, OGA and PD2, and the respective poses scored 159 and 168 units (Table 1). The scores in line with those recorded for UN9, OGA and PD2, and the comparable arrangement into the pocket could suggest the capability of both OTA isomers to interact and possibly inhibit OGFOD1. This hypothesis has been further confirmed by MD simulations which showed that their complexes with OGFOD1 were stable over time and comparable to that with UN9 – taken as reference scenario – as demonstrate by the Root-Mean Squared Deviation (RMSD) and trajectories analysis (Figure 2). Of note, with respect to the docking poses, the first minimization steps of MD simulations were crucial to reduce a steric unbalance at the site of interaction of the OTA phenylalanine moiety. Specifically, both OTA isomers and

Leu182, Phe210 and His218 underwent a slight re-arrangement allowing a better fitting of the phenylalanine part. Moreover, the capability of OTA and 2'R-OTA of keeping the bidentate coordination was analysed too. Looking at the crystals, UN9 (PDB ID 4NHM) kept the closest nitrogen to the cation at 2.2 Å and the carbonyl at 2.1 Å, PD2 (PDB ID 4NHY) kept the closest nitrogen to the cation at 2.3 Å and the carboxyl group at 2.1 Å and OGA (PDB ID 4NHX) kept both the carboxyl and carbonyl groups at 2.2 Å from the Mn cation. Regarding OTA, the carbonyl group was kept at 2.03 ± 0.006 Å from the Fe cation while the hydroxyl group at 2.24 ± 0.009 Å, in line with the distances kept by the co-crystallized ligands and with a proper bidentate coordination (Harding, 2006). On the other hand, 2'R-OTA kept the carbonyl group at 0.446 ± 0.019 Å while the hydroxyl one at 0.322 ± 0.019 Å. This suggested that 2'R-OTA might be a weaker OGFOD1 interactor being not able of maintaining distances consistent with the bidentate coordination described for the other compounds (Harding, 2006). Furthermore, the capability of OTA and 2'R-OTA to form and maintain a salt bridge interaction with Arg230 was assessed monitoring the average distance between the carboxylate of the ligands and the arginine guanidinium moiety. Both isomers were found capable of keeping this interaction throughout the course of the simulations, being the distances kept consistent to what expected from a salt bridge (3.04 ± 0.031 Å regarding OTA, 2.75 ± 0.016 regarding 2'R-OTA) (Ban et al., 2019). Eventually, US were performed to estimate the free energy of binding for a thorough estimate of OTA and 2R'-OTA inhibitory potential (OGFOD1-UN9 complex was also computed as reference). The analysis showed that both complexes were stable with ΔG corresponding to roughly -7, -8 and -9 kcal/mol for UN9, 2'R-OTA and OTA, respectively (Figure 2).

The outcomes collected here consistently pointed to the capability of both OTA isomers to interact and likely inhibit OGFOD1. Interestingly, OGFOD1 has been described to catalyse the 3-hydroxylation of Pro62 of small ribosomal subunit uS12 (RPS23), regulating protein translation efficiency (Wehner et al., 2010). Besides, OGFOD1 activity has been also associated to a plethora of other biological functions including the regulation of cell cycle and autophagy, and its inhibition/knockdown may induce cell cycle arrest as well as autophagy (Fujisaki et al., 2023; Singleton et al., 2014). These macroscopic effects are interesting being also associated with OTA toxicity, though with not well clarified mechanisms. In this respect, the hypothesized inhibition of OGFOD1 may fill some knowledge gaps about the mechanism of action of OTA toxicity. In more detail, OTA has been described to induce cycle arrest, apoptosis and autophagy, modulating a series of stress indicators, including the time-dependent increase of phosphorylation of eIF2α (Wang et al., 2022), a key element for translation initiation (Zoll et

al., 2002). Notably, the phosphorylation of eIF2α, which depends also on OGFOD1, turns the protein in its inactive form hindering the translation process (Jiang & Wek, 2005). This apparently contradicts the hypothesised OTA-dependent inhibition of OGFOD1 as OGFOD1 activity may lead to an increase of eIF2α phosphorylation (Wehner et al., 2010). Therefore, at a first instance, the inhibition of OGFOD1 should have led to a reduction rather than to the increase of eIF2α phosphorylation. However, it has been also demonstrated that a depletion of OGFOD1 may increase the phosphorylation of eIF2α in a time-dependent manner, though with unclear mechanisms, and this is associated with the augmented expression of ATF4 and its target protein CHOP (Singleton et al., 2014). This suggested that the OTA-dependent inhibition of OGFOD1 may still be compatible with a certain increase of eIF2α phosphorylation – though the underpinning molecular reasons are unknown. Strikingly, ATF4 has been found overexpressed upon OTA treatment in kidney cells (Yang et al., 2023) and this may have come with the here hypothesized inhibition of OGFOD1 as described elsewhere (Singleton et al., 2014) . Following this line of interpretation, it has been also described that OGFOD1 may increase LC3B lipidation (Singleton et al., 2014), which is also observed upon OTA treatment of kidney cells (Akpinar et al., 2019). In addition, it was demonstrated that OGFOD1 regulates the transcription and post-transcriptional stabilization of cell cycle-related genes such as cyclin dependent kinase 2 (CDK2) (Fujisaki et al., 2023). Interestingly, CDK2 was proposed as one of the key regulators of the G1 cell-cycle arrest induced by low nanomolar concentrations of OTA (Dubourg et al., 2020). This effect could also be related to the possible OTA mediated OGFOD1 inhibition we proposed in this study. From a general point of view, OTA has been previously described to impair the protein synthesis. The inhibition of phenylalanine tRNA synthase and phenylalanine hydroxylase has been proposed among the underlying mechanisms, though the whole network of targets involved is likely more complex and still to get disclosed (Akpinar et al., 2019). Therefore, the inhibition of OGFOD1, which is critical for the termination of protein translation, might take a part in the OTA-dependent inhibition of protein synthesis as an ancillary mechanism worth of future dedicated investigations.

Germane to 2'R-OTA, this diastereomer is important from a food safety perspective as it can be formed upon processing of certain OTA-containing food, and it may reach circulating concentrations even higher compared to OTA (Cramer et al., 2015). The results collected here pointed to its substantial capability to qualitatively act like OTA and a combined assessment of both isomers in terms of OGFOD1 inhibition should be considered for further in-depth analysis considering the exposure of human population to both isomers.

Table 1. Docking scores and reported activity of molecules under analysis

Compound	**Docking Score**	**Experimental Activity (IC_{50})**
UN9	172	0.52 ± 0.08 μM[1]
PD2	152	2.3 ± 1.2 μM[1]
OGA	150	1.9 ± 1.0 μM[1]
OTA	159	Unknown
2'R-OTA	168	Unknown

Note: [1] Data retrieved from the study by Horita et al. (Horita et al., 2015).

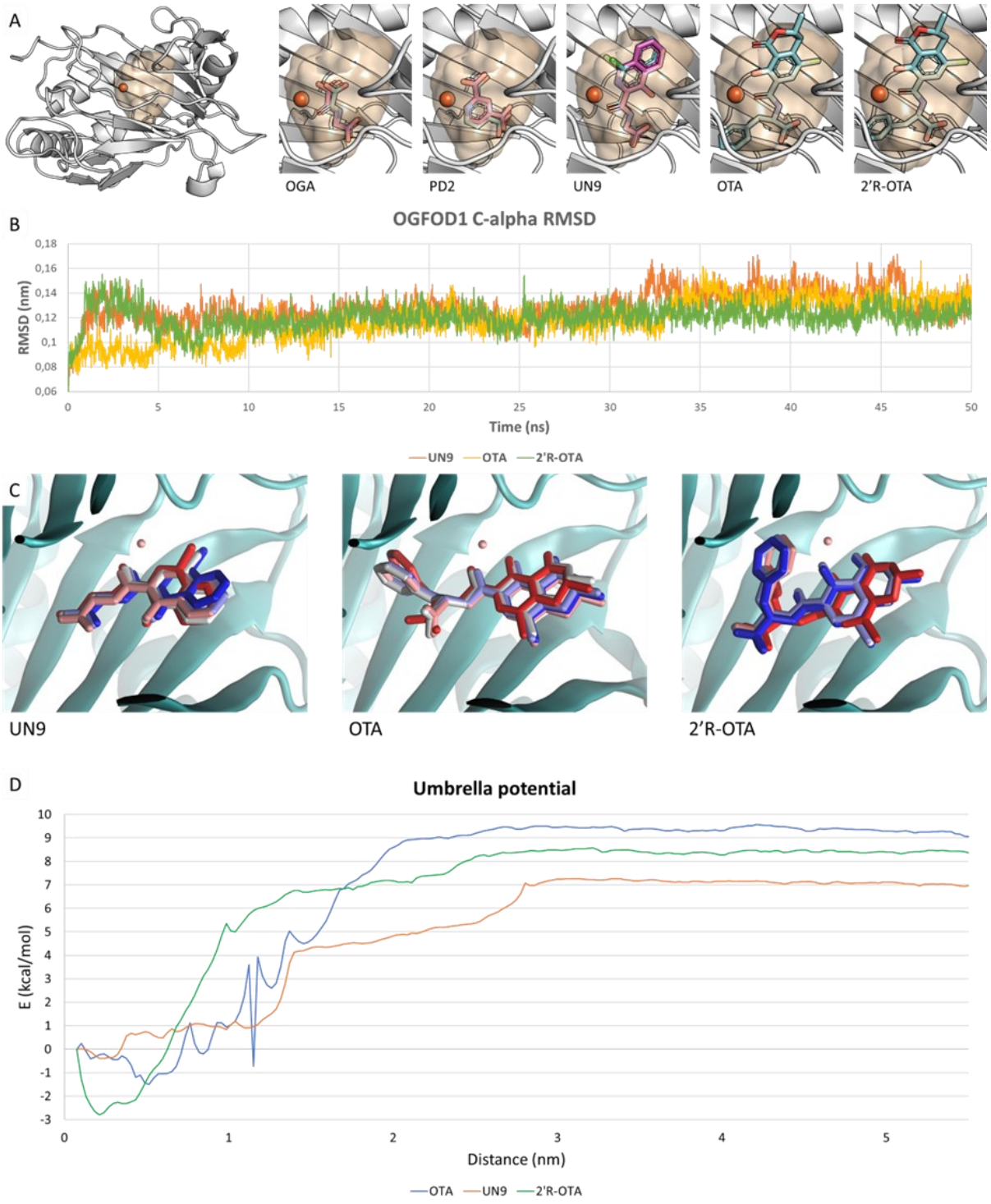


Figure 2. Molecular modelling validation and results of the considered ligands. **A**. On the left the 3D OGFOD1 model represented in white cartoon, with Fe represented as an orange sphere and the binding pocket as wheat semi-transparent surface. On the right, the docked poses are represented as cyan sticks while the crystallographic binding architectures of OGA, PD2 and UN9 are reported as magenta sticks. **B**. Protein C-alpha RMSD of OGFOD1 in complex with UN9 (orange, average value 0.128 ± 0.013 nm), OTA (yellow, average value 0.118 ± 0.016 nm)

and 2'R-OTA (green, average value 0.120 ± 0.010 nm). **C**. Time-step representation (from red, 0 ns, to blue, 40 ns) of the trajectories of UN9, OTA and 2'R-OTA within the OGFOD1 binding site. The protein is represented as cyan cartoon while the Fe as a wheat sphere. **D**. Umbrella potential graph for UN9, 2'R-OTA and OTA with ΔG corresponding to roughly -7, -8 and -9 kcal/mol, respectively.

4. Conclusion

Given the still unclear network of biological targets underpinning OTA toxicity, the present study identified OGFOD1 as a possible key target explaining new possible mechanisms of action of OTA and 2'R-OTA toxicity never considered before. Indeed, the all set of data collected pointed to the capability of OTA and 2'R-OTA to stably interact with OGFOD1, reasonably resulting in its inhibition. This hypothesized inhibition of OGFOD1 may fill some knowledge gaps about the mechanisms of OTA toxicity being the OTA-dependent inhibition of OGFOD1 soundly consistent with the current molecular understanding of OTA toxicity. The strong chemical similarity with known OGFOD1 inhibitors, the soundness of the computational outcomes collected here and their consistency with the current (partial) understanding of OTA mechanism of action provide a compelling line of evidence pointing to the reasonable capability of OTA and its 2'R stereoisomer to inhibit OGFOD1. On this basis, OGFOD1 may be a mechanistic knot interlacing the yet scarcely connected molecular actors and events associated with the OTA-dependent toxicity, including the impairment of protein synthesis and downstream cascades. Therefore, the in-depth analysis of OTA-dependent inhibition of OGFOD1 and the subsequent impact on its interaction network should be prioritized in future dedicated studies toward a more informed description of OTA adverse outcome pathway.

References

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX, 1-2*, 19-25. doi:https://doi.org/10.1016/j.softx.2015.06.001

Akpinar, H. A., Kahraman, H., & Yaman, I. (2019). Ochratoxin A Sequentially Activates Autophagy and the Ubiquitin-Proteasome System. *Toxins, 11*(11). doi:10.3390/toxins11110615

Ban, X. F., Lahiri, P., Dhoble, A. S., Li, D., Gu, Z. B., Li, C. M., . . . Kaustubh, B. (2019). Evolutionary Stability of Salt Bridges Hints Its Contribution to Stability of Proteins. *Computational and Structural Biotechnology Journal, 17*, 895-903. doi:10.1016/j.csbj.2019.06.022

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., . . . Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research, 28*(1), 235-242. doi:10.1093/nar/28.1.235

Bryla, M., Ksieniewicz-Wozniak, E., Stepniewska, S., Modrzewska, M., Waskiewicz, A., Szymczyk, K., & Szafranska, A. (2021). Transformation of ochratoxin A during bread-making processes. *Food Control, 125*. doi:10.1016/j.foodcont.2021.107950

Bugnon, M., Goullieux, M., Röhrig, U. F., Perez, M. A. S., Daina, A., Michielin, O., & Zoete, V. (2023). SwissParam 2023: A Modern Web-Based Tool for Efficient Small Molecule Parametrization. *Journal of Chemical Information and Modeling, 63*(21), 6469-6475. doi:10.1021/acs.jcim.3c01053

Cournia, Z., Allen, B., & Sherman, W. (2017). Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *Journal of Chemical Information and Modeling, 57*(12), 2911-2937. doi:10.1021/acs.jcim.7b00564

Cramer, B., Königs, M., & Humpf, H. U. (2008). Identification and in vitro cytotoxicity of ochratoxin a degradation products formed during coffee roasting. *Journal of Agricultural and Food Chemistry, 56*(14), 5673-5681. doi:10.1021/jf801296z

Cramer, B., Osteresch, B., Muñoz, K. A., Hillmann, H., Sibrowski, W., & Humpf, H. U. (2015). Biomonitoring using dried blood spots: Detection of ochratoxin A and its degradation product 2'R-ochratoxin A in blood from coffee drinkers. *Molecular Nutrition & Food Research, 59*(9), 1837-1843. doi:10.1002/mnfr.201500220

Del Favero, G., Aichinger, G., Hohenbichler, J., Marko, D., Dall'Asta, C., & Dellafiora, L. (2022). A target fishing study to spot possible biological targets of fusaric acid: Inhibition of protein kinase-A and insights on the underpinning mechanisms. *Food and Chemical Toxicology, 159*. doi:10.1016/j.fct.2021.112663

Dellafiora, L., Gonaus, C., Streit, B., Galaverna, G., Moll, W. D., Vogtentanz, G., . . . Prasad, S. (2020). An In Silico Target Fishing Approach to Identify Novel Ochratoxin A Hydrolyzing Enzyme. *Toxins, 12*(4). doi:10.3390/toxins12040258

Dubourg, V., Nolze, A., Kopf, M., Gekle, M., & Schwerdt, G. (2020). Weighted Correlation Network Analysis Reveals CDK2 as a Regulator of a Ubiquitous Environmental Toxin-Induced Cell-Cycle Arrest. *Cells, 9*(1). doi:10.3390/cells9010143

Fujisaki, T., Saito, K., Kikuchi, T., & Kondo, E. (2023). The prolyl hydroxylase OGFOD1 promotes cancer cell proliferation by regulating the expression of cell cycle regulators. *Febs Letters, 597*(8), 1073-1085. doi:10.1002/1873-3468.14547

Harding, M. M. (2006). Small revisions to predicted distances around metal sites in proteins. *Acta Crystallographica Section D-Biological Crystallography, 62*, 678-682. doi:10.1107/s0907444906014594

Horita, S., Scotti, J. S., Thinnes, C., Mottaghi-Taromsari, Y. S., Thalhammer, A., Ge, W., . . . McDonough, M. A. (2015). Structure of the Ribosomal Oxygenase OGFOD1 Provides Insights into the Regio-and Stereoselectivity of Prolyl Hydroxylases. *Structure, 23*(4), 639-652. doi:10.1016/j.str.2015.01.014

IARC. (1993). *Some naturally occurring substances: Food items and constituents, heterocyclic aromatic amines and mycotoxins.* (Vol. 56). Lyon, France: Monographs on the evaluation of carcinogenic risks to humans: IARC Press.

Jiang, H. Y., & Wek, R. C. (2005). Phosphorylation of the α-subunit of the eukaryotic initiation factor-2 (eIF2α) reduces protein synthesis and enhances apoptosis in response to proteasome inhibition. *Journal of Biological Chemistry, 280*(14), 14189-14202. doi:10.1074/jbc.M413660200

Kim, S., Chen, J., Cheng, T. J., Gindulyte, A., He, J., He, S. Q., . . . Bolton, E. E. (2023). PubChem 2023 update. *Nucleic Acids Research, 51*(D1), D1373-D1380. doi:10.1093/nar/gkac956

Koszegi, T., & Poór, M. (2016). Ochratoxin A: Molecular Interactions, Mechanisms of Toxicity and Prevention at the Molecular Level. *Toxins, 8*(4). doi:10.3390/toxins8040111

Kuppusamy, A., Arumugam, M., & George, S. (2017). Combining in silico and in vitro approaches to evaluate the acetylcholinesterase inhibitory profile of some commercially available flavonoids in the management of Alzheimer's disease. *International Journal of Biological Macromolecules, 95*, 199-203. doi:10.1016/j.ijbiomac.2016.11.062

Liang, R., Shen, X. L., Zhang, B. Y., Li, Y. Z., Xu, W. T., Zhao, C. H., . . . Huang, K. L. (2015). Apoptosis Signal-regulating Kinase 1 promotes Ochratoxin A-induced renal cytotoxicity. *Scientific Reports, 5*. doi:10.1038/srep08078

Park, M., Baek, S. J., Park, S. M., Yi, J. M., & Cha, S. (2023). Comparative study of the mechanism of natural compounds with similar structures using docking and transcriptome data for improving in silico herbal medicine experimentations. *Brief Bioinform, 24*(6). doi:10.1093/bib/bbad344

Pedroni, L., Perugino, F., Galaverna, G., Dall'Asta, C., & Dellafiora, L. (2022). An In Silico Framework to Mine Bioactive Peptides from Annotated Proteomes: A Case Study on Pancreatic Alpha Amylase Inhibitory Peptides from Algae and Cyanobacteria. *Nutrients, 14*(21). doi:10.3390/nu14214680

Perry, J. L., Il'ichev, Y. V., Kempf, V. R., McClendon, J., Park, G., Manderville, R. A., . . . Simon, J. D. (2003). Binding of ochratoxin A derivatives to human serum albumin. *Journal of Physical Chemistry B, 107*(27), 6644-6647. doi:10.1021/jp034284w

Perugino, F., Pedroni, L., Galaverna, G., Dall'Asta, C., & Dellafiora, L. (2024). A mechanistic toxicology study to grasp the mechanics of zearalenone estrogenicity: Spotlighting aromatase and the effects of its genetic variability. *Toxicology, 501*, 153686. doi:https://doi.org/10.1016/j.tox.2023.153686

Singleton, R. S., Liu-Yi, P., Formenti, F., Ge, W., Sekirnik, R., Fischer, R., . . . Cockman, M. E. (2014). OGFOD1 catalyzes prolyl hydroxylation of RPS23 and is involved in translation control and stress granule formation. *Proceedings of the National Academy of Sciences of the United States of America, 111*(11), 4031-4036. doi:10.1073/pnas.1314482111

Tao, Y. F., Xie, S. Y., Xu, F. F., Liu, A. M., Wang, Y. X., Chen, D. M., . . . Yuan, Z. H. (2018). Ochratoxin A: Toxicity, oxidative stress and metabolism. *Food and Chemical Toxicology, 112*, 320-331. doi:10.1016/j.fct.2018.01.002

Thinnes, C. C., Lohans, C. T., Abboud, M. I., Yeh, T. L., Tumber, A., Nowak, R. P., . . . Schofield, C. J. (2019). Selective Inhibitors of a Human Prolyl Hydroxylase

(OGFOD1) Involved in Ribosomal Decoding. *Chemistry-a European Journal, 25*(8), 2019-2024. doi:10.1002/chem.201804790

Vettorazzi, A., González-Peñas, E., & de Cerain, A. L. (2014). Ochratoxin A kinetics: A review of analytical methods and studies in rat model. *Food and Chemical Toxicology, 72*, 273-288. doi:10.1016/j.fct.2014.07.020

Wang, Y., Cui, J. F., Zheng, G. N., Zhao, M., Hao, Z. F., Lian, H. G., . . . Wang, J. (2022). Ochratoxin A induces cytotoxicity through ROS-mediated endoplasmic reticulum stress pathway in human gastric epithelium cells. *Toxicology, 479*. doi:10.1016/j.tox.2022.153309

Wehner, K. A., Schütz, S., & Sarnow, P. (2010). OGFOD1, a Novel Modulator of Eukaryotic Translation Initiation Factor 2α Phosphorylation and the Cellular Response to Stress. *Molecular and Cellular Biology, 30*(8), 2006-2016. doi:10.1128/mcb.01350-09

Yang, S. A., Rhee, K. H., Yoo, H. J., Pyo, M. C., & Lee, K. W. (2023). Ochratoxin A induces endoplasmic reticulum stress and fibrosis in the kidney via the HIF-1α/miR-155-5p link. *Toxicol Rep, 10*, 133-145. doi:10.1016/j.toxrep.2023.01.006

Zoll, W. L., Horton, L. E., Komar, A. A., Hensold, J. O., & Merrick, W. C. (2002). Characterization of mammalian eIF2A and identification of the yeast homolog. *Journal of Biological Chemistry, 277*(40), 37079-37087. doi:10.1074/jbc.M207109200

Delving into the possible role of Ochratoxin A in Parkinson's disease: a computational study spotlighting the potential mechanisms of action

This chapter is based on the paper published on:

Lorenzo Pedroni, Florinda Perugino, Chiara Dall'Asta, Lydia Alvarez-Erviti, Angela Mally, Ariane Vettorazzi, Luca Dellafiora. (2025). Delving into the possible role of Ochratoxin A in Parkinson's disease: a computational study spotlighting the potential mechanisms of action. Food Research International, 221, 117532, https://doi.org/10.1016/j.foodres.2025.117532.

Author contribution: Methodology, Investigation

Chapter V

Abstract

Ochratoxin A (OTA) is a widespread food contaminant relevant to food safety that has various toxic effects including nephrotoxicity and neurotoxicity. Evidence suggests OTA may contribute to Parkinson's disease (PD) pathogenesis acting at multiple levels along the adverse outcome pathway (AOP), though molecular targets and underpinning mechanisms remain unclear. Using a computational pipeline based on molecular modelling techniques, which are well-established methods to elucidate protein-ligand interactions, this study investigated key events in the PD AOP where the current state-of-the-art suggests OTA and its congeners may play a role, i.e. inhibition of cathepsins B, D, and L (relevant for α-synuclein proteostasis), and Complex I (responsible for mitochondrial dysfunction). The study revealed that neither OTA nor its congeners could interact with cathepsin B, but some of them displayed differential binding with cathepsins D and L. The debated metabolites ochratoxin hydroquinone and quinone could favourably interact with the catalytic residue of the former and the latter, respectively. Additionally, OTA and ochratoxin quinone showed binding to Complex I in a rotenone-like fashion, suggesting its possible inhibition. These outcomes well-aligned with the current comprehension of OTA-PD relationship by providing deep mechanistic insights correlating OTA exposure and PD pathogenesis and revealing a novel and multi-tiered mechanism where OTA and metabolites may affect α-synuclein proteostasis and mitochondrial function. This could not only explain previous experimental observations, but also lays a knowledge-based groundwork for targeted investigations into OTA's role in PD pathogenesis.

1. Introduction

The mycotoxin ochratoxin A (OTA; Figure 1), a secondary metabolite of *Aspergillus* and *Penicillium* fungi (Zhang et al., 2024), is a well-known food contaminant raising food safety concerns (Kerstner & Garda-Buffon, 2024). From a chemical point of view, it consists of a chloro-methyl-dihydro-isocoumarin moiety linked by an amide bond to a L-phenylalanine (Figure 1). OTA exerts several toxic effects in animal species, may cause kidney and liver tumours in rodents (Schrenk et al., 2020) and it was classified as a possible human carcinogen by IARC (group 2B) (IARC, 1993). Evidence also highlights its potential neurotoxicity (Nie et al., 2025; Obafemi et al., 2023) along with its possible role in Parkinson's Disease (PD) (Sava et al., 2006; Serrano-Civantos et al., 2025). Serrano-Civantos et al. reviewed the available literature reporting that *in vitro* and *in vivo* exposure to OTA causes key events related to the PD's Adverse Outcome Pathway (AOP; Figure 2) like mitochondrial dysfunction, impaired proteostasis, degeneration of dopaminergic neurons, and neuroinflammation (Bal-Price et al., 2018; Serrano-Civantos et al., 2025). Additionally, Izco et al. demonstrated *in vivo* (oral exposure of mice) and *in vitro* (intestinal and neuronal cell models) that low-dose and sub-toxic levels of OTA triggered PD-related key features like altered α-synuclein (α-syn) proteostasis and associated accumulation (Izco et al., 2021). They also reported OTA-induced α-syn aggregates in the murine *substantia nigra compacta*, compatible with a PD-like pathological progression (Veras et al., 2024). Importantly, Serrano-Civantos et al. highlighted gaps on PD's molecular initiating events (MIE), which involve the disruption of mitochondrial Complex I, as well as on the mechanisms responsible for the impaired α-syn proteostasis (Serrano-Civantos et al., 2025). Of note, α-syn proteostasis is a complex and finely regulated process that is yet to be fully understood. However, the lysosomal pathway is critical (Kakuda et al., 2024) and, although the specific lysosomal components involved and the degraded α-syn state (i.e. aggregate or non-aggregate) differ across studies, cathepsins (e.g. cathepsin D, B and L) have been identified among the repertoire of lysosomal proteases capable of cleaving and degrading α-syn (Stefanis et al., 2019). Cathepsins are lysosomal proteases involved in many physiological processes with different mechanisms of catalysis: cathepsin L and B are cysteine-based proteases while cathepsin D is an aspartic-based one (Anes et al., 2022).

Strikingly, a previous work estimated the capability of OTA to interact with the catalytic site of cathepsin A, a broad-spectrum serine protease, in a substrate-like manner (Dellafiora et al., 2020), while previous evidence suggested the possible hydrolysis of OTA to form ochratoxin α (OTα) by the cysteine-based cathepsin C (Tao et al., 2018). Although cathepsin A and C have

been excluded as crucial for the α-syn proteostasis (Sampognaro et al., 2023), the question arose as to whether OTA could instead interact with cathepsin D, B and L possibly resulting in their inhibition. In this framework, we applied an already validated computational pipeline, integrated with machine learning, to elucidate the capability of OTA to stably persist at the catalytic site of cathepsin D, B and L. The ability to model whether small molecules can remain stable (see section 2.3.) at an enzyme's catalytic site is a crucial factor to successfully access their inhibitory capacity (e.g. (Florinda Perugino et al., 2024; Zhao et al., 2021)), and it may also enable the investigation of the inhibitory potential of OTA (F. Perugino et al., 2024). Moreover, in line with the shortage of data concerning the MIE of PD mentioned above, we tested also the capability to competitively interact and inhibit Complex I. Of note, evidence has described that alterations of lysosomal functions (crucial for α-syn proteostasis), Complex I impairment and altered mitochondrial biogenesis may contribute to the loss of dopaminergic neurons (Guerra et al., 2019; Thomas et al., 2012), which is part of PD AOP's key events (Figure 2). Although it remains to be clarified whether these are interconnected through a cascade of events or may occur simultaneously and synergistically, this evidence emphasized the reasonableness of a combined action against the functionality of cathepsins, inherently linked to lysosomal functions, and Complex I. To account for the spectrum of OTA congeners that may co-occur in real-world scenarios, the investigation also included its main metabolite OTα, ochratoxin B (OTB; the non-chlorinated analogue of OTA), as well as two debated potential reactive metabolites of OTA, i.e. ochratoxin hydroquinone (OTHQ) and ochratoxin quinone (OTQ) (Dai et al., 2002; Dekant et al., 2021; Schrenk et al., 2020) (Figure 2). Of note, these computational approaches are gaining significant traction in toxicological investigations as part of NAMs (New Approach Methodologies) offering a robust, independent and stand-alone analytical framework which uniquely enables investigation of biological/toxicological consequences of protein-ligand complex formation at a fine molecular resolution (Cattaneo et al., 2023; Östberg et al., 2024; Undiano et al., 2023; Veras et al., 2024).

Overall, the study revealed that OTA and its congeners exhibit distinct interaction profiles with cathepsins D, B, and L, as well as Complex I. This differential interaction suggests that both the parent toxin and its metabolites may contribute to the disruption of normal proteolytic and mitochondrial functions, thereby providing compelling evidence to explain, at least in part, the mechanistic basis of OTA role across some of the key events of PD's AOP.

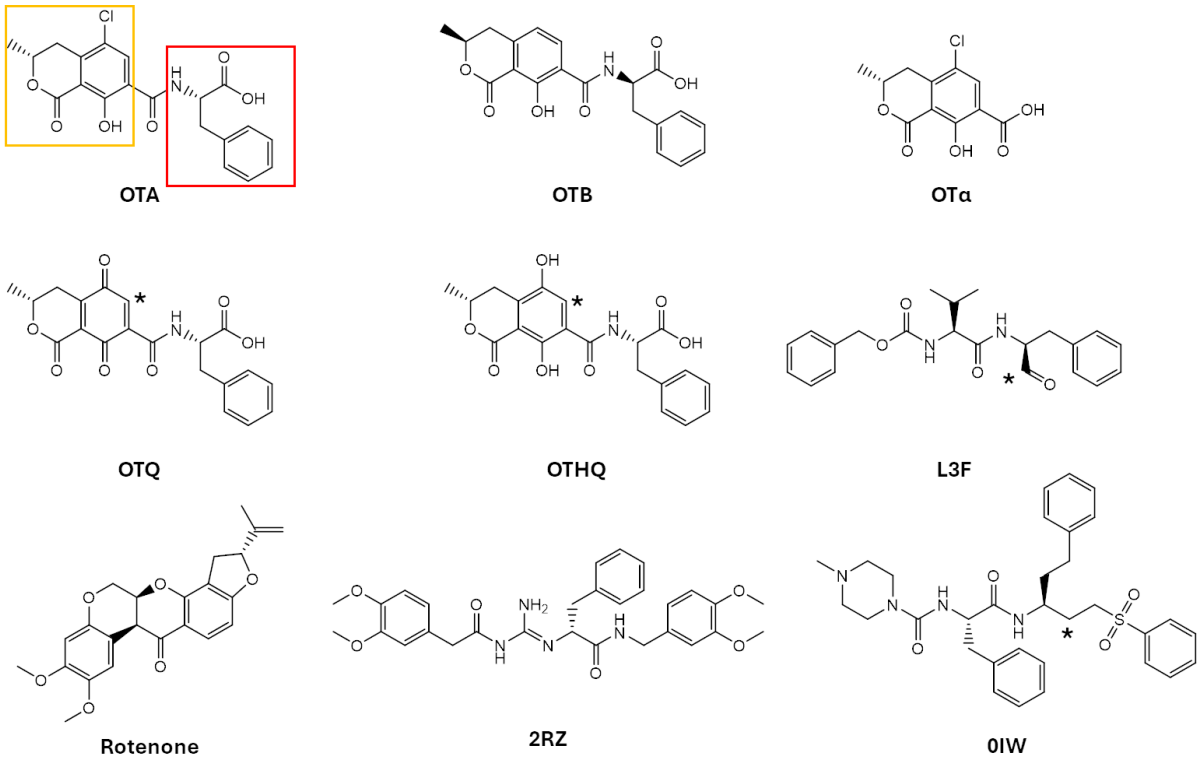


Figure 1. Chemical sketches of molecules under analysis. The asterisk indicates the region intended to undergo nucleophilic reactions. Regarding OTA and its congeners, the yellow and red boxes indicate the chloro-methyl-dihydro-isocoumarin group and the L-phenylalanine moiety, respectively (OTα lacks the phenylalanine moiety).

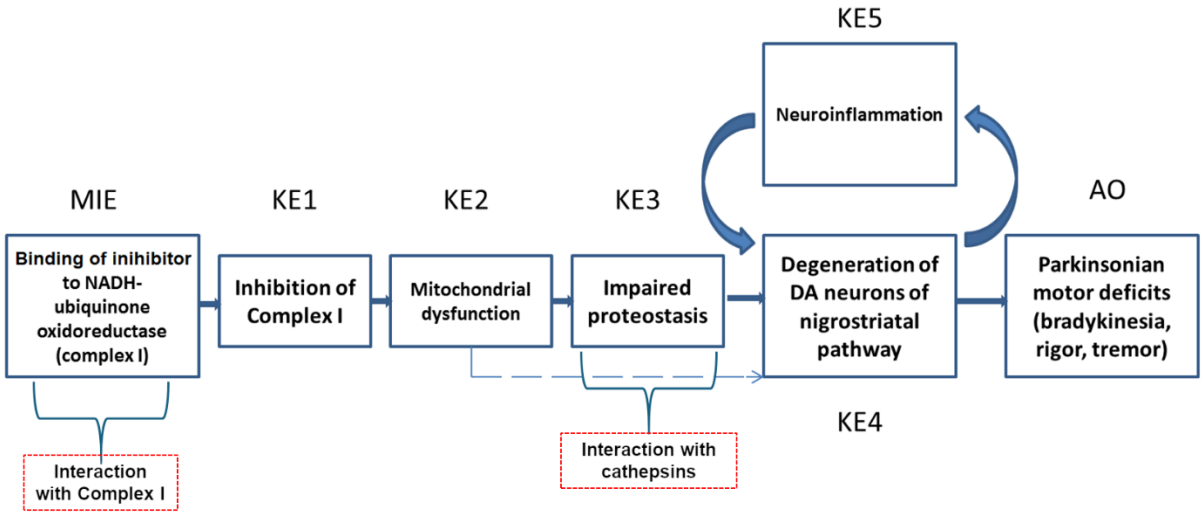


Figure 2. Parkinson's disease adverse outcome pathway (AOP). The red dotted boxes indicate the biological targets with respect the key event they belong to. Information shown are adapted from the PD's AOP (https://aopwiki.org/aops/3) (Bal-Price et al., 2018).

2. Material and methods

2.1. Data source

The 3D structures for all the analysed proteins were collected from the Protein Data Bank (PDB; https://www.rcsb.org/). The 3D structure of cathepsin B (UniProt AC P07858) was derived from PDB ID 3AI8 (Grädler et al., 2014), cathepsin D (UniProt AC P07339) from PDB ID 4OD9 (Grädler et al., 2014), cathepsin L (UniProt AC P07711) from PDB ID 2XU3 (Hardegger et al., 2011), and Complex I (Uniprot AC P49821) from PDB ID 5XTB (R. Y. Guo et al., 2017). Regarding the cathepsins, all co-crystallized ligands, ions and waters were removed keeping only the protein chain. Moving to Complex I, only the protein chains C and Q were kept for the subsequent molecular docking and dynamics simulations. Regarding OTA and congeners, they were collected in the SDF format from PubChem (https://pubchem.ncbi.nlm.nih.gov/) under the following IDs: OTA (CID 442530), OTB (CID 20966), OTα (CID 107911), OTQ (CID 101013464), OTHQ (CID 101944009). A set of known inhibitors for the investigated proteins was also collected for benchmarking purposes. Specifically, CID 52946595 for Cathepsin B (namely, 0IW) (H. F. Wang et al., 2023), CID 9843057 for Cathepsin L oxidized to an aldehyde (namely, L3F) as per Falke et al. (Falke et al., 2024), CID 76871908 for Cathepsin D (namely, 2RZ) (Grädler et al., 2014), and CID 6758 (namely, rotenone) for Complex I (J. K. Gu et al., 2022). All the ligands were converted to a 3D MOL2 format using Open Babel (v. 3.1.1) (O'Boyle et al., 2011) setting the pH to 7.

2.2. Molecular docking

Docking simulations were performed to predict the architecture of binding of ligands at the target proteins' defined binding sites. The software GOLD (Genetic Optimization for Ligand Docking; version 2024) was used to perform docking simulations as it has previously succeeded in estimating protein-ligand interactions (L Pedroni et al., 2023). Concerning cathepsins, the binding site was defined in a 5 Å-radius sphere around the centroid of the respective co-crystallized ligands. For human Complex I, the binding site was set as a 10 Å-radius sphere around the centroid of the rotenone co-crystallized in *Sus scrofa* Complex I (PDB ID 7VBZ) (J. K. Gu et al., 2022). This was done after superimposing the *Sus scrofa* Complex I structure onto the human one. Ten poses for each ligand under investigation were generated at the defined binding site and scored according to the internal scoring function PLPscore (the higher the score, the more reliable the binding architecture, as per manufacturer declaration https://www.ccdc.cam.ac.uk).

2.3. Molecular dynamics

Molecular dynamics (MD) simulations were performed to refine the molecular docking outcomes while monitoring the geometrical stability of each complex under investigation. MD simulations were conducted using GROMACS (v. 2022.6) (Abraham et al., 2015). Each complex was parametrized adopting the CHARMM36 all-atom force field (Best et al., 2012). While proteins were handled using GROMACS internal libraries, the ligands' CHARMM36 topology and parameters were generated via the SwissParam web-tool (http://swissparam.ch/) (Bugnon et al., 2023). All complexes were embedded in a dodecahedron periodic box, solvated with SPC/E waters and an ion concentration (Na^+ and Cl^-), sufficient to neutralize the net charge. This was followed by energy minimization (steepest descent algorithm, maximum of 5 000 steps). Then, an isothermal (300K, coupling times of 2 ps) and isobaric (1 bar, coupling time 2 ps) equilibration step lasting 100 ps were conducted. This was done to obtain properly equilibrated systems prior to the final 200 ns long MD production step. Of note, the MD production step for Complex I was extended to 300 ns. For all systems under investigation, the stability of protein-ligand complexes was assessed by combining the expert visual inspection of MD trajectories and ligand root mean square deviation analysis (RMSD). The trajectories were analysed regarding ligand movements over time with respect to the binding site. Concerning the RMSD analysis, low and steady RMSD values indicate a geometrically stable interaction within the protein's binding pocket. Last, the software named MMPBSA.py (Miller et al., 2012), which is a post-processing end-state streamlining method to calculate free energy via MM/PBSA, was used to evaluate the interaction free energy of those protein-ligand complexes who showed a stable RMSD trend. For each analysed MD production, a sample of 50 snapshots was collected starting from the half-time of the MD trajectory. It should be noted that the intrinsic approximations of the MM/PBSA methodology can introduce systematic and system-dependent offsets in the absolute scale of binding free energies, although the approach remains well suited for analysing and rationalizing trends, especially within congeneric series of ligands and targets (Genheden & Ryde, 2015). Accordingly, for each protein-ligand complex who showed a stable RMSD trend, binding free energies were expressed as percentage relative to the free energy of binding of the given benchmarked ligand, thereby reducing systematic bias and yielding results that are more interpretable and directly comparable.

2.4. Analysis of conformational states

Markov State Models (MSMs) were performed based on a custom Python script (Lorenzo Pedroni & Dellafiora, 2025) to characterize proteins conformational landscape revealing metastable states, transition pathways, and kinetic barriers. In agreement with previous studies, the script integrated multiple libraries including MDTraj (v 1.10.3; to handle trajectory processing), scikit-learn (1.6.1; for machine learning operations like features extraction, dimensionality reduction (PCA), and clustering (KMeans)), and matplotlib and seaborn (v 3.9.2 and 0.13.2, respectively; for data visualization) (Harrigan et al., 2017; Noé et al., 2009; Vanden-Eijnden, 2014). The trajectory data were first pre-processed by extracting multiple structural and dynamic features including dihedrals, distances, radius of gyration (Rg), solvent accessible surface area (SASA), root-mean-square deviation (RMSD), hydrogen bonds, salt bridges, and secondary structure (ss) metrics (Pérez-Hernández et al., 2013). To make the analysis computationally affordable, each MD trajectory was reduced via striding procedure to 1000 frames (stride equal to 0.1% of total MD trajectory frames). This ensured efficient sampling while maintaining trajectory resolution. Features were standardized to zero mean and unit variance using the StandardScaler (a function in the scikit-learn Python library that standardizes features by removing the mean and scaling to unit variance) prior to PCA application, ensuring that all structural metrics contributed equally to the dimensionality reduction regardless of their original numerical scales. Principal Component Analysis (PCA) was used to reduce dimensionality including 10 principal components (PCs) to cover as much variance as possible (Pérez-Hernández et al., 2013). K-means clustering was performed dynamically selecting the number of initializations between 3 and 10 based on the MD characteristics to cover it complexity and granting at the same time computational efficiency. Of note, the clustering was applied to identify metastable states, with the optimal number of clusters determined based on the consensus of multiple scoring metrics, including silhouette, Calinski-Harabasz, and Davies-Bouldin scores. Clustering quality was assessed by calculating silhouette coefficients, which measure how similar objects are to their assigned cluster compared to other clusters (Husic & Pande, 2018). Representative structures for each identified cluster were extracted based on the frame closest to each cluster centroid in the reduced PCA space, using Euclidean distance in the PCA-transformed feature space as the similarity metric (Shao et al., 2007). Structures were aligned to a reference frame using protein backbone atoms prior to representative selection. These representative structures were saved as PDB files for further structural analysis. The relative populations of each conformational state were calculated from the frame distribution

across clusters. The transition probability matrices were estimated maintaining a consistent physical timescale target with respect to the striding. The timescale was set equal to the actual time interval derived after striding procedure (100 ps in this case) ensuring that the lag time corresponded to physically meaningful timescales in the molecular system, allowing for accurate estimation of transition kinetics (Noé & Fischer, 2008). The resulting matrix of transition probability was row-normalized to ensure proper stochastic matrix properties. Transition probabilities were weighted using a Markov chain Monte Carlo approach with 100 bootstrap iterations, providing confidence intervals for each transition probability. To validate the Markovian assumption, implied timescales were calculated across a range of lag times and checked for convergence, confirming that the selected lag time was sufficient for the system to exhibit memory-less behaviour (Noé & Clementi, 2017). Based on previous studies (Noé et al., 2009), the transition path analysis to identify probable pathways between metastable states employed flux calculations using the equation 1.

$$Fij = \pi i\ Pij - \pi j\ Pji \quad \text{(equation 1)}$$

where πi represents the equilibrium probability of state i, πj represents the equilibrium probability of state j Pij is the transition probability from state i to state j, and Fij is the resulting net flux between states.

Transition states are determined using committor probabilities, with the reaction coordinate (RC) determined as per equation 2 following established principles of transition path theory (Vanden-Eijnden, 2014).

$$RC = q^{+}/(q^{+} + q^{-}) \approx 0.5 \quad \text{(equation 2)}$$

q^{+} represents the forward committor (probability of reaching the target state before returning to the source state) and q^{-} is the backward committor.

3. Results

3.1 Cathepsins

Benchmarking. The procedure reliability was validated by calculating the interaction of a benchmark ligand for each protein under analysis. The interaction was calculated at the binding site indicated by crystallographic evidence, thereby providing meaningful references and enabling the validation of predictions. Cathepsin B was tested with the known inhibitor 0IW (CID 52946595) (H. F. Wang et al., 2023). This ligand recorded a positive docking score (Table 1) and showed a stable interaction along the whole 200 ns MD production. Notably, the computed docking pose and the maintained proximity to the catalytic residue Cys29 throughout

the MD production were in line with the available crystallographic data (Figure S1, Supplementary material). Cathepsin D was tested with the known inhibitor 2RZ (Grädler et al., 2014), obtaining a positive docking score (Table 1) and a negative free energy of binding pointing to the energetic favour of interaction (Table S1, Supplementary material), showing a computed pose in line with structural data (Figure S1; Supplementary material) and demonstrating a stable interaction, which engaged the Asp231 and Asp33, during the whole MD production (Figure 3). Cathepsin L was tested with the known inhibitor L3F (i.e. CID 9843057 oxidized to an aldehyde as per Falke et al. (Falke et al., 2024)) obtaining a positive docking score (Table 1) and a negative free energy of binding pointing to the energetic favour of interaction (Table S1, Supplementary material), with a computed pose in line with structural data (Figure S1 ; Supplementary material), while maintaining a good stability and a proper positioning to the Cys25's side chain, in line with structural evidence. These results eventually confirmed procedural performances and effectiveness for the considered cathepsins, as well as enabled to benchmark and validate the results obtained for OTA and its analogues (see below). It should be noted that for cathepsin B, the free energy of binding for the benchmark ligand was not calculated since neither OTA nor its analogues were observed to interact stably. Therefore, there was no need for comparisons with the benchmark ligand (see below).

Cathepsin B. Neither OTA nor any congener tested here was found stable (see section 2.3.) when interacting with the binding site of cathepsin B (UniProt AC P07858). Although positive docking scores were recorded (Table 1) – which theoretically indicate a physico-chemical ligand-pocket match (as per manufacturer declaration, https://www.ccdc.cam.ac.uk) – the whole set of OTA and congeners detached from the designated binding site within 100 ns of MD production (Figure S2; Supplementary material).

Cathepsin D. OTA, OTB and OTQ remained close to the interaction site of cathepsin D (UniProt AC P07339) over time (Figure S3; Supplementary material). However, despite the positive docking scores (Table 1) and their capability to persist at the catalytic site, they could not stably engage the catalytic residue Asp231, which is a recurrent feature of effective cathepsin D inhibitors, as per structural studies (e.g. (Grädler et al., 2014)). Conversely, OTHQ was stable and maintained a steady interaction, similar to that observed for the known inhibitor 2RZ, and close to the catalytic residue Asp231 during the whole MD production (Figure 3). The interaction was further assessed via MM-PBSA, a post-processing end-state method to calculate free energies (see section 2.3), obtaining a negative free energy of binding pointing to the energetic favour of interaction though corresponding to the 30% of the tested known inhibitor 2RZ (Table S1, Supplementary material). The binding architecture of OTHQ was further

analysed via MSM to describe the dynamic of its conformational landscape within the binding site. The PCA performed during the MSM construction captured 70% of the system's conformational variance, indicating an effective sampling of the relevant conformational space. MSM clustering identified three predominant metastable states: State 1 (50% occupancy) and 2 (45% occupancy) were both stable with 56 and 55% self-transition probability, respectively, while State 3 (5% occupancy) was most likely to transit to State 2 (52% probability). In the representative structures of State 2, OTHQ formed a hydrogen bond with Asp231 with an oxygen-to-oxygen distance of 2.6 Å and favourable geometry (Figure 3). Transition probability analysis indicated infrequent interconversion of States 1 and 2, whereas State 3 frequently transitioned to State 2. Therefore, State 2 could roughly represent half of the trajectory suggesting that once OTHQ establishes its interaction at the catalytic site it could consistently engage Asp231. Concerning OTα, although it scored positive at the interaction site (Table 1), it detached from the binding site around 80 ns from the beginning of the MD production (Figure S3; Supplementary material).

Cathepsin L. Regarding cathepsin L (UniProt AC P07711), OTA and all tested congeners obtained positive docking scores (Table 1). OTA and OTHQ maintained a steady stable interaction during the whole MD production. However, for both compounds the phenylalanine moiety (Figure 1) displayed a certain degree of mobility (Figure S4; Supplementary material). Conversely, OTQ deserves special attention as it not only demonstrated good stability within the active site during MD production but also maintained a promising distance from the catalytic residue Cys25 (Figure 3). This positioning potentially enables a nucleophilic reaction since OTQ is a chemical compound known to be reactive against thiols (Schrenk et al., 2020). The OTQ-Cathepsin L complex was further analysed with MM-PBSA, obtaining a negative free energy of binding corresponding to 68% of the tested known inhibitor L3F pointing to the energetic favour of interaction (Table S1, Supplementary material), and with MSM. The MSM's PCA captured around 80% of the system's conformational variance, indicating a thorough sampling of the relevant conformational space. MSM clustering identified two predominant metastable states: State 1 was stable with 94.8% self-transition probability, while State 2 was computed to transit to State 1 (100% probability). In the representative structure of State 1, the reactive region of OTQ (i.e. the atom meant to be part of nucleophilic reactions) was 3.8 Å from the Cys25's sulphur atom (Figure 3). Transition probability analysis indicated infrequent interconversion between these two states, suggesting that once OTQ establishes its interaction with the catalytic site, its closeness to Cys25 may favour the nucleophilic attack. On the other hand, OTα and OTB poorly interacted with the active site showing instability and eventually

detaching from the cathepsin L's catalytic site (Figure S4; Supplementary material).

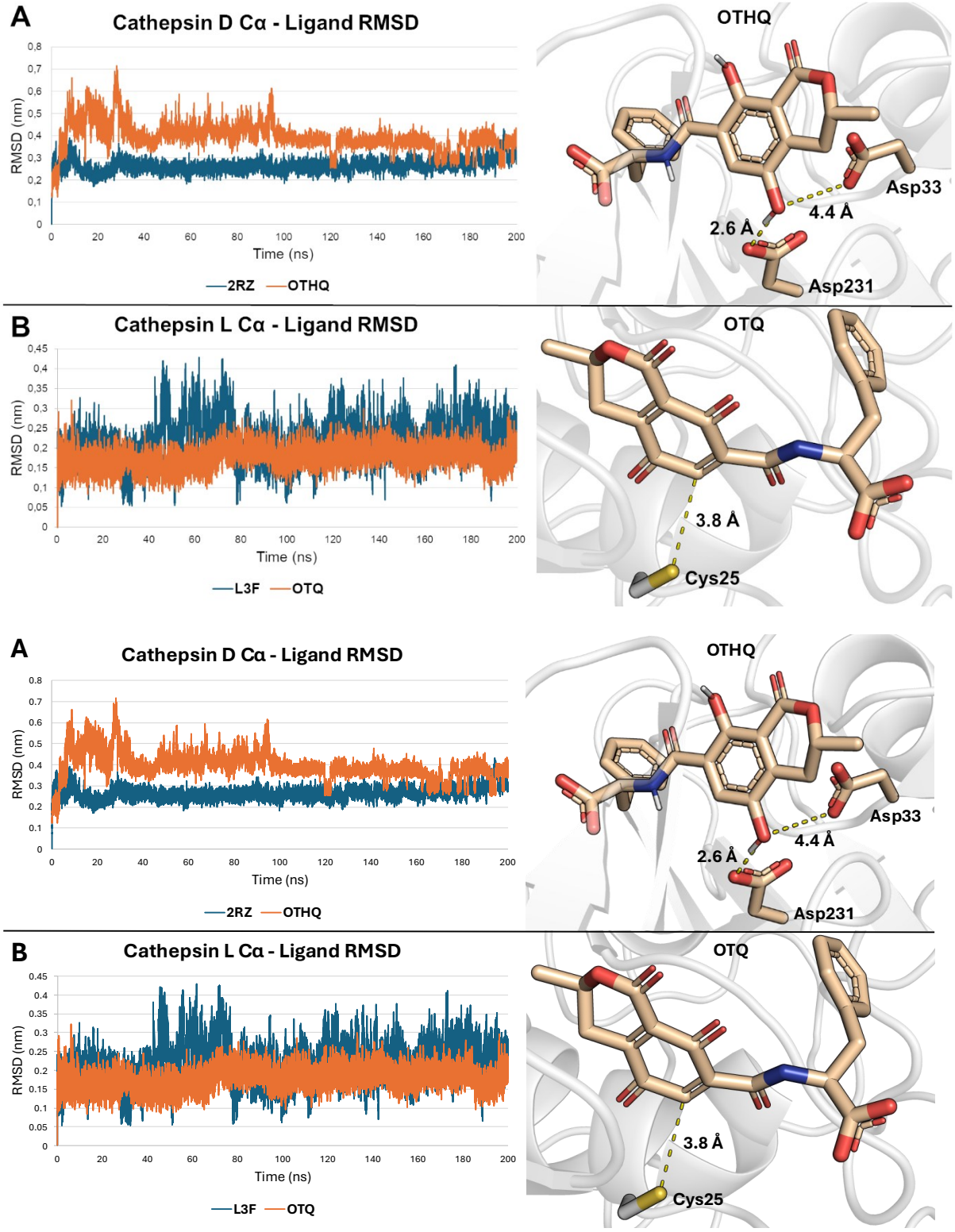


Figure 3. MD results showing RMSD plots and the most representative structure after MSM analysis. The proteins are represented in cartoon, ligands and amino acids side chains are represented in sticks. The dashed yellow lines indicate interatomic distances. **A.** OTHQ – with polar hydrogens displayed – within cathepsin D. On the left, the RMSD plot compared with the

benchmarked known inhibitor 2RZ. On the right the representative binding mode as per MSM analysis. **B.** OTQ within cathepsin L. On the left, the RMSD plot with the benchmarked known inhibitor L3F. On the right the representative binding mode as per MSM analysis.

3.2 Complex I

A benchmarking procedure was performed using the well-known inhibitor rotenone (CID 6758) (J. K. Gu et al., 2022) to validate the reliability of the approach. Rotenone was tested for its interaction at the interface of two mitochondrial Complex I chains (UniProt AC O75251 and O75306) (Figure S5; Supplementary material) – hereafter referred to as Complex I – at the binding site described by crystallographic studies (Gu et al., 2022). It obtained a positive docking score (Table 1), with a computed pose in line with structural data (Figure S1; Supporting material), a negative free energy of binding (Table S1, Supplementary material), and stably interacted at the defined binding site during the whole MD simulation (Figure 4). Taken together, these findings eventually confirmed procedural performances also on Complex I. OTA and its congeners were also tested for their interaction at the same interaction site. All tested congeners obtained positive docking scores (Table 1). Both OTA and OTQ, maintained a steady interaction with the defined binding site throughout the whole MD production similarly to rotenone (Figure 4), obtaining negative free energy of binding values, corresponding to the 93% (OTA) and 69% (OTQ) of the values obtained for the benchmarked inhibitor rotenone (Table S1, Supplementary material), pointing to the energetic favour of both interactions. This could suggest that OTA and OTQ may result in appreciable inhibition of Complex I via a mechanism qualitatively comparable to that of rotenone. Conversely, as reflected by RMSD trends and trajectories analysis, OTHQ, OTα, and OTB displayed a poor stability, with OTα eventually detaching from Complex I (Figure S6; Supplementary material).

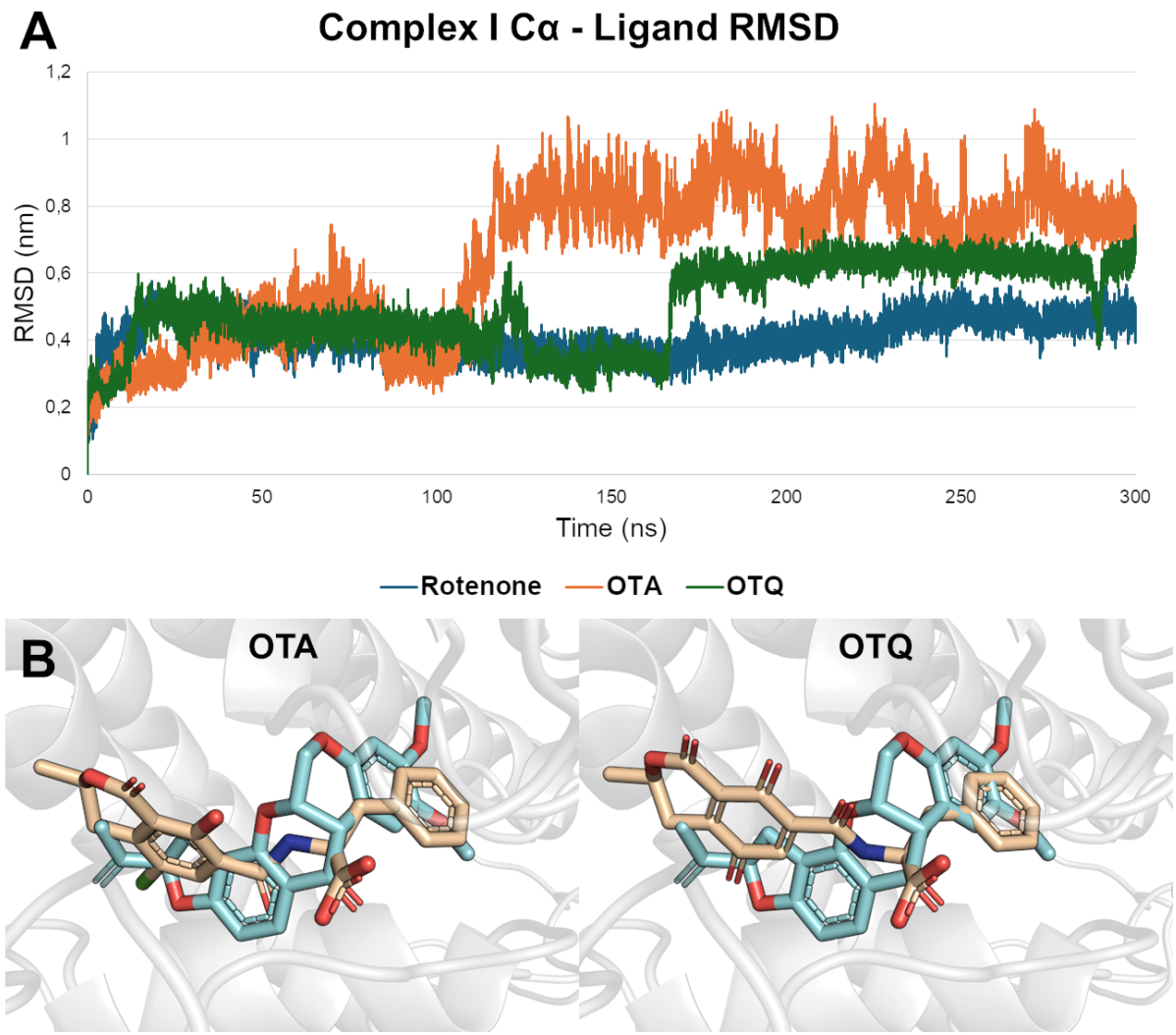


Figure 4. MD results showing RMSD plots and input conformation for MD simulations for OTA and OTQ. The protein is represented in cartoon while ligands are represented in sticks. **A.** RMSD plots compared with the benchmarked known inhibitor rotenone. B. On the right the representative binding mode as per MSM analysis. **B.** Input conformation for MD simulations for OTA and OTQ superimposed to that of rotenone (cyan sticks).

Table 1. Docking scores [1]

	Cathepsin B	Cathepsin D	Cathepsin L	Complex I
OTA	79.82	55.68	52.96	63.93
OTα	65.94	44.70	38.58	40.24
OTB	80.95	59.32	50.30	68.75
OTHQ	89.56	52.40	51.97	64.87
OTQ	87.06	59.01	50.46	63.27
Rotenone [2]	np	np	np	49.15
L3F [2,3]	np	np	59.14	np
2RZ [2,3]	np	75.77	np	np
0IW [2,3]	124.42	np	np	np

Note: [1] the higher the score the better the interaction (as per manufacturer declaration; https://www.ccdc.cam.ac.uk); [2] used as positive control in the benchmarking procedure for each system tested positive, either with OTA or congeners; [3] PDB ID; np stands for "calculation not performed"

3. Discussion

In the context of using NAMs to deepen the relationship between OTA exposure and PD, this work addressed the possible role of OTA and some congeners (OTα, OTB, OTHQ, and OTQ) in two key events of PD's AOP (Bal-Price et al., 2018): i) the impairment of cathepsins B, D and L activity (key actors in α-syn proteostasis); ii) the inhibition of Complex I, the upstream event resulting in mitochondrial dysfunction. Aligning with previous studies (e.g.(Florinda Perugino et al., 2024; F. Perugino et al., 2024)), molecular docking and dynamics simulations were benchmarked using four known ligands: 2RZ for cathepsin D, L3F for cathepsin L, 0IW for cathepsin B, and rotenone for Complex I. These computations for reference ligands aligned well with the structural data available (Figure S1; Supplementary material), showing stable interactions at the designated binding site, in line with their mechanisms of inhibition (e.g. the engagement of Asp231/33 and the proximity of reactive moieties to Cys25's for cathepsin D and L inhibitors, respectively), and they all recorded negative free energy of binding values pointing to the energetic favour of their interaction with the respective targets (Table S1, Supplementary material). This could eventually confirm the methodological accuracy. Concerning OTA and its congeners, all the tested molecules obtained positive docking scores (Table 1) suggesting physico-chemical ligand-binding site compatibility. However, in line with previous evidence (Florinda Perugino et al., 2024; F. Perugino et al., 2024), MD simulations

were crucial to refine docking outcomes.

Concerning cathepsin B, neither OTA nor its congeners demonstrated stable binding at the catalytic site, excluding this enzyme as a possible target. Conversely, cathepsin D and L emerged as convincing targets. The lack of stable interactions of OTA and its congeners with cathepsin B, but not with cathepsin L, despite both being cysteine proteases, might arise from structural differences that have been already widely investigated. Of note, as reported by Fujishima et al. (Fujishima et al., 1997), cathepsin B shows an occluding loop that is known to affect enzyme activity and substrate specificity (Illy et al., 1997) and could limit access and proper binding of OTA and its congeners. On the other hand, the wider space of the ligand binding cleft of cathepsin L might better adapt and accommodate a broader range of compounds determining the proper accommodation of OTQ.

Concerning cathepsin D, OTα detached from the catalytic site, while OTA, OTB, and OTQ could interact, although not replicating key features observed for known inhibitors like the consistent engagement of Asp33 and/or Asp231 (Grädler et al., 2014). Therefore, they are unlikely to prevail over natural substrates or inhibitors. Notably, a weak inhibitory potential due to steric hindrance cannot be excluded. OTHQ appeared to stably interact with Asp231 via hydrogen bonding, as the benchmarked inhibitor (2RZ), and recorded negative free energy of binding as well accounting for the 30% of the free energy of binding calculated for 2RZ. The MSM analysis supported this observation, describing State 2 (around 50% occupancy, 55% self-transition probability), where OTHQ forms a hydrogen bond with Asp231 as a significant portion of the conformational landscape. The relatively high self-transition probability highlights its conformational stability upon this binding mode achievement. Although OTHQ is known to be unstable in the presence of oxygen (Schrenk et al., 2020), its (meta-)stable existence at the interaction site cannot be excluded *a priori*. Indeed, although hydroquinones are generally barely stable and prone to oxidation (Bolton et al., 2000; Ghanbarzadeh et al., 2015), they demonstrated stability under certain circumstances, including when bound to proteins (Deri et al., 2016; W. C. Guo et al., 2008). These findings, despite the evidence supporting OTA/OTHQ conversion remains limited, underpin OTHQ potential role in disrupting cathepsin D activity and related α-syn proteostasis, though with an expected potency lower than that of the benchmarked inhibitor based on the lower calculated free energy of binding.

Concerning cathepsin L, OTα and OTB were unstable or not resembled the benchmarked inhibitor L3F and thus excluded as possible inhibitors. Conversely, the MD simulation revealed

that OTA and OTHQ maintained a stable interaction while the phenylalanine moiety rearranged during the MD production step, though in a similar fashion to one of the two phenyl moieties of the benchmarked inhibitor. However, given that cathepsin L is a cysteine-based protease, OTA, and possibly OTHQ, might undergo a substrate-like processing as it was already proposed for another cysteine-based protease, i.e. cathepsin C (Tao et al., 2018). If this is the case, OTA and OTHQ could still interfere with the reaction yield by competing with the natural substrates, as demonstrated for other substrate-like enzymatic inhibitors (e.g. (Iwaniak et al., 2014)). For this reason, further dedicated experiments should investigate whether OTA and OTHQ primarily act as substrates or inhibitors. In addition to the already known OTA's interference with protein translation, this could add a crucial piece of information to broaden the understanding of OTA's mechanisms of toxicity. On the other hand, OTQ exhibited a stable benchmarked ligand-alike interaction maintaining a distance to the catalytic Cys25 suitable to undergo nucleophilic reaction (Figure 3). Indeed, previous evidence reported that OTQ could undergo nucleophilic reaction with thiol groups (Schrenk et al., 2020), further corroborating the hypothesis that a similar reaction could occur in this context. Despite covalent binding cannot be directly observed via MD simulations performed in this study - nucleophilic reactions may happen at a fs scale (Zhong et al., 1997) – the consistent proximity over the simulation timeframe (i.e. nanoseconds) supports this hypothesis, in line with previous studies (Cianni et al., 2020; Silva et al., 2021). Of note, although this approach may reliably estimate the likelihood of nucleophilic attacks (L. Pedroni et al., 2025), it is not meant to replace more accurate pipelines based on specific forcefields (e.g. ReaxFF (van Duin et al., 2001)) and in general QM/MM approaches. However – in this context – it can support a theoretical yet plausible interpretation within the limits of classical MD, especially considering the system size and the adopted timescale in the order of hundreds of nanoseconds. Therefore, based on the above, OTQ was deemed as a possible covalent inhibitor worthy of additional focused analyses. Also, the negative free energy of binding recorded for this compound, which accounted for the 68% of the free energy of binding calculated for the known tested inhibitor L3F (with a Ki in pM range (Falke et al., 2024)), furtherly substantiates in this sense pointing to the energetic favour of its interaction. Last, this hypothesis is strongly supported by the MSM results, which identified a dominant metastable conformation (State 1) characterized by 94.8% self-transition probability. Moreover, the reactive region of OTQ is at an optimal distance from the Cys25 reactive side chain, suggesting a pre-reactive conformation primed for nucleophilic attack. Additionally, the computed probable transition of State 2 to State 1 further emphasized the energetic favours of the latter binding mode, supporting the hypothesized potential of OTQ to

act as a covalent inhibitor. A similar behaviour could not be excluded for OTHQ, given the intrinsic reactivity of hydroquinones, although its behaviour retraced that of OTA, as it did not maintain an optimal distance favouring a possible nucleophilic reaction with the sulphur atom of Cys25.

Of note, cathepsin D and L might be differently engaged regarding α-syn degradation, with the former mostly involved in the degradation of full-length α-syn, while the latter possibly degrading α-syn fibrils (Stefanis et al., 2019). Based on this assumption, OTA and congeners may exert a multifaceted role with respect to α-syn proteostasis considering the underlined impairment of multiple cathepsins and their possible co-occurrence in a real-world context, which are worthy of in-depth dedicated analysis.

Last, strong evidence links mitochondrial Complex I dysfunction/disruption to progressive parkinsonism, specifically leading to the gradual loss of the dopaminergic phenotype (Bhat et al., 2016; Q. Li et al., 2019; Serrano-Civantos et al., 2025). Interestingly, OTA already proved to induce mitochondrial dysfunction, impair proteostasis and cause dopaminergic neurons degeneration (Serrano-Civantos et al., 2025). However, a key knowledge gap remained in understanding whether OTA could possibly bind – and thus inhibit – Complex I. In this framework, our *in silico* results revealed that both OTA and OTQ maintained a rotenone-like interaction with Complex I and showed negative free energies of binding in line with this known inhibitor (accounting for the 93% and 69% of free energy of binding calculated for rotenone, respectively), providing mechanistic insight bridging the previously highlighted gap. Aligning with the capability of OTA to affect mitochondrial function mentioned above, this finding supports the hypothesis that OTA may inhibit Complex I in a rotenone-like manner possibly exerting its role in PD by directly binding to Complex I. This would also be in line with other natural compounds, either found in plants or produced by microorganisms, which are able to bind and inhibit Complex I, leading to progression of neurodegenerative disorders (Höllerhage et al., 2009; Schiller & Zickermann, 2022), also causing oxidative stress through ROS formation and general energy depletion (Bates et al., 1994). It is interesting to note that rotenone has been reported to inhibit not only Complex I but also cathepsins (A. L. Wang et al., 2009) suggesting a potential multi-target action along the PD pathway. A similar effect was observed for 1-metil-4-fenil-1,2,3,6-tetraidropiridina (MPTP), a toxin mainly known to inhibit Complex I, but also reported to reduce cathepsin D levels (R. Gu et al., 2024). This evidence points out the simultaneous inhibition of cathepsins and Complex I as possible, aligning with the multiple hit hypothesis of PD (Carvey et al., 2006), which suggests that a combination of multiple chemical, genetic and environmental risk factors may combinatorially triggers key events

leading to the development of PD. In this regard, two pieces of evidence emerge as particularly relevant: i) different compounds may interact with the same target (e.g. OTQ and OTA computed to interact with Complex I); ii) a single compound may act on multiple targets (e.g. OTQ saw to interact with cathepsins and Complex I, similarly to rotenone and MPTP). These observations well align not only with the multi-hit hypothesis of PD, but also with the previously described crosstalk between mitochondria and endo-lysosomal pathway in the pathophysiology of genetic PD (Plotegher & Duchen, 2017). Importantly, it should be noted that the existence of OTQ and OTHQ in biological systems remains a topic of scientific debate. Dekant et al. (Dekant et al., 2021) found no evidence for the formation of OTA-derived mercapturic acids that would indicate metabolic processing of reactive OTQ intermediates. They observed that while signals corresponding to OTHQ-GSH could be detected in liver S9 fraction incubations without NADPH, this phenomenon disappeared in the presence of NADPH. This evidence, along with the finding that OTHQ is very unstable in the presence of oxygen and rapidly oxidised to OTQ, which itself decomposes in aqueous solution to several products (Gillman et al., 1999) and may react with nucleophilic biological molecules (Schrenk et al., 2020) suggests the transient nature of OTHQ/OTQ *in vivo*. These two OTA's congeners might be formed at best to a very low extent at a systemic level thereby preventing the formation of detectable conjugates when analysing tissues or tissue preparations. However, OTHQ/OTQ are real chemical species that can be generated under certain circumstances, as demonstrated *in vitro* through peroxidase activation or photochemical reactions. Furthermore, de Medeiros et al. reported the formation of OTQ from OTA under acidic condition (de Medeiros et al., 2025) and lysosomes, the organelles presenting cathepsins, are known to present an acidic environment (S. S. Li et al., 2019). This may determine favourable conditions to form OTQ that might transiently reach concentrations *in situ* at the enzymes interaction site that might be high enough to play a biological role. The *in situ* production of OTQ was also hypothesised by Dai et al. (Dai et al., 2002) to explain the photochemical reaction of OTHQ with cysteine resulting in the production of the cysteinyl conjugate. In line with this interpretation, our findings suggest that under the specific conditions that make OTHQ and OTQ (meta-)stable, like those acidic of lysosomes, they may potentially contribute to the mechanistic basis of OTA toxicity, even if their presence *in vivo* is likely to be context dependent.

Overall, this study mechanistically described how OTA and some of its debated congeners (OTHQ and OTQ) may contribute to PD pathogenesis by interfering with multiple targets of the PD AOP. These findings bridge important knowledge gaps in understanding the mechanism of OTA neurotoxicity, particularly regarding α-syn proteostasis impairment and mitochondrial

dysfunction, both key events of PD's AOP. This, along with evidence collected in previously published studies, strongly supports the hypothesis that OTA might favour PD onset and progression by both interfering with α-syn proteostasis – e.g. disrupting cathepsin D and L activity – and with more basic cellular function, like the impairment of mitochondrial function via Complex I inhibition. Importantly, based on the multi-hit hypothesis of PD and acknowledging the alignment of such a hypothesis with the outcomes collected here and those available to date connecting OTA to PD, further studies will need to quantitatively and precisely assess the impact of OTA relative to other environmental and genetic risk factors associated with PD onset and development. With this respect, while the present study is computational, it provides a solid mechanistic groundwork that might guide future and further experimental follow-ups. *In vitro* enzyme inhibition assays – for cathepsin D and L – and targeted mitochondrial assays – for Complex I – would be valuable to confirm and quantify the proposed interactions and functional impacts of OTA and its congeners. In addition, techniques such as surface plasmon resonance could also be employed to study the real-time kinetics of these molecular interactions. Although such experiments lie beyond the scope of this work and deserve dedicated studies to link molecular evidence to *in vivo* implications, the mechanistic findings presented here may provide well-defined molecular targets and foster further hypothesis for focused experimental follow-ups and/or designing specific intervention strategies against OTA-induced neurotoxicity.

References

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX, 1-2*, 19-25. doi:10.1016/j.softx.2015.06.001

Anes, E., Pires, D., Mandal, M., & Azevedo-Pereira, J. M. (2022). Spatial localization of cathepsins: Implications in immune activation and resolution during infections. *Frontiers in Immunology, 13*. doi:10.3389/fimmu.2022.955407

Bal-Price, A., Leist, M., Schildknecht, S., Tschudi-Monnet, F., Paini, A., & Terron, A. (2018). *Adverse Outcome Pathway on Inhibition of the mitochondrial complex I of nigro-striatal neurons leading to parkinsonian motor deficits*. Retrieved from http://dx.doi.org/10.1787/b46c3c00-en

Bates, T. E., Heales, S. J., Davies, S. E., Boakye, P., & Clark, J. B. (1994). Effects of 1-methyl-4-phenylpyridinium on isolated rat brain mitochondria: evidence for a primary involvement of energy depletion. *J. Neurochem., 63*(2), 640-648. doi:10.1046/j.1471-4159.1994.63020640.x

Best, R. B., Zhu, X., Shim, J., Lopes, P. E. M., Mittal, J., Feig, M., & MacKerell, A. D. (2012). Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone phi, psi and Side-Chain chi(1) and chi(2) Dihedral Angles. *Journal of Chemical Theory and Computation, 8*(9), 3257-3273. doi:10.1021/ct300400x

Bhat, P. V., Pandareesh, M. D., Khanum, F., & Tamatam, A. (2016). Cytotoxic Effects of Ochratoxin A in Neuro-2a Cells: Role of Oxidative Stress Evidenced by N-acetylcysteine. *Frontiers in Microbiology, 7*. doi:10.3389/fmicb.2016.01142

Bolton, J. L., Trush, M. A., Penning, T. M., Dryhurst, G., & Monks, T. J. (2000). Role of quinones in toxicology. *Chemical Research in Toxicology, 13*(3), 135-160.

Bugnon, M., Goullieux, M., Röhrig, U. F., Perez, M. A. S., Daina, A., Michielin, O., & Zoete, V. (2023). SwissParam 2023: A Modern Web-Based Tool for Efficient Small Molecule Parametrization. *Journal of Chemical Information and Modeling, 63*(21), 6469-6475. doi:10.1021/acs.jcim.3c01053

Carvey, P. M., Punati, A., & Newman, M. B. (2006). Progressive dopamine neuron loss in Parkinson's disease: The multiple hit hypothesis. *Cell Transplantation, 15*(3), 239-250. doi:10.3727/000000006783981990

Cattaneo, I., Astuto, M. C., Binaglia, M., Devos, Y., Dorne, J., Agudo, A. F., . . . Terron, A. (2023). Implementing New Approach Methodologies (NAMs) in food safety assessments: Strategic objectives and actions taken by the European Food Safety Authority. *Trends in Food Science & Technology, 133*, 277-290. doi:10.1016/j.tifs.2023.02.006

Cianni, L., Rocho, F. D., Rosini, F., Bonatto, V., Ribeiro, J. F. R., Lameira, J. N., . . . Montanari, C. A. (2020). Optimization strategy of single-digit nanomolar cross-class inhibitors of mammalian and protozoa cysteine proteases. *Bioorganic Chemistry, 101*. doi:10.1016/j.bioorg.2020.104039

Dai, J., Park, G., Wright, M. W., Adams, M., Akman, S. A., & Manderville, R. A. (2002). Detection and characterization of a glutathione conjugate of ochratoxin A. *Chemical Research in Toxicology, 15*(12), 1581-1588. doi:10.1021/tx0255929

de Medeiros, J. S., Oliveira, J. P., Knidel, H., da Silva, T. M. N., Rodrigues, J. G. A., & Ferreira, R. Q. (2025). Optimized Electrochemical Sensor for Ochratoxin A Quantification in Coffee and Wheat Matrices. *ACS Omega, 10*(26), 28252-28268. doi:10.1021/acsomega.5c03171

Dekant, R., Langer, M., Lupp, M., Chilaka, C. A., & Mally, A. (2021). In Vitro and In Vivo Analysis of Ochratoxin A-Derived Glucuronides and Mercapturic Acids as Biomarkers of Exposure. *Toxins, 13*(8). doi:10.3390/toxins13080587

Dellafiora, L., Gonaus, C., Streit, B., Galaverna, G., Moll, W. D., Vogtentanz, G., . . . Prasad, S. (2020). An In Silico Target Fishing Approach to Identify Novel Ochratoxin A Hydrolyzing Enzyme. *Toxins, 12*(4). doi:10.3390/toxins12040258

Deri, B., Kanteev, M., Goldfeder, M., Lecina, D., Guallar, V., Adir, N., & Fishman, A. (2016). The unravelling of the complex pattern of tyrosinase inhibition. *Scientific Reports, 6*. doi:10.1038/srep34993

Falke, S., Lieske, J., Herrmann, A., Loboda, J., Karnicar, K., Günther, S., . . . Meents, A. (2024). Structural Elucidation and Antiviral Activity of Covalent Cathepsin L Inhibitors. *Journal of Medicinal Chemistry, 67*(9), 7048-7067. doi:10.1021/acs.jmedchem.3c02351

Fujishima, A., Imai, Y., Nomura, T., Fujisawa, Y., Yamamoto, Y., & Sugawara, T. (1997). The crystal structure of human cathepsin L complexed with E-64. *Febs Letters, 407*(1), 47-50. doi:10.1016/s0014-5793(97)00216-0

Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery, 10*(5), 449-461. doi:10.1517/17460441.2015.1032936

Ghanbarzadeh, S., Hariri, R., Kouhsoltani, M., Shokri, J., Javadzadeh, Y., & Hamishehkar, H. (2015). Enhanced stability and dermal delivery of hydroquinone using solid lipid nanoparticles. *Colloids and Surfaces B-Biointerfaces, 136*, 1004-1010. doi:10.1016/j.colsurfb.2015.10.041

Gillman, I. G., Clark, T. N., & Manderville, R. A. (1999). Oxidation of ochratoxin A by an Fe-porphyrin system: Model for enzymatic activation and DNA cleavage. *Chemical Research in Toxicology, 12*(11), 1066-1076. doi:10.1021/tx9901074

Grädler, U., Czodrowski, P., Tsaklakidis, C., Klein, M., Werkmann, D., Lindemann, S., . . . Leuthner, B. (2014). Structure-based optimization of non-peptidic Cathepsin D inhibitors. *Bioorganic & Medicinal Chemistry Letters, 24*(17), 4141-4150. doi:10.1016/j.bmcl.2014.07.054

Gu, J. K., Liu, T. Y., Guo, R. Y., Zhang, L. X., & Yang, M. J. (2022). The coupling mechanism of mammalian mitochondrial complex I. *Nature Structural & Molecular Biology, 29*(2), 172-182. doi:10.1038/s41594-022-00722-w

Gu, R., Bai, L. P., Yan, F., Zhang, S., Zhang, X. W., Deng, R. H., . . . Bai, J. (2024). Thioredoxin-1 decreases alpha-synuclein induced by MPTP through promoting autophagy-lysosome pathway. *Cell Death Discovery, 10*(1). doi:10.1038/s41420-024-01848-0

Guerra, F., Girolimetti, G., Beli, R., Mitruccio, M., Pacelli, C., Ferretta, A., . . . Bucci, C. (2019). Synergistic Effect of Mitochondrial and Lysosomal Dysfunction in Parkinson's Disease. *Cells, 8*(5). doi:10.3390/cells8050452

Guo, R. Y., Zong, S., Wu, M., Gu, J. K., & Yang, M. J. (2017). Architecture of Human Mitochondrial Respiratory Megacomplex I2III2IV2. *Cell, 170*(6), 1247-1257. doi:10.1016/j.cell.2017.07.050

Guo, W. C., Siegel, D., & Ross, D. (2008). Stability of the Hsp90 Inhibitor 17AAG Hydroquinone and Prevention of Metal-Catalyzed Oxidation. *Journal of Pharmaceutical Sciences, 97*(12), 5147-5157. doi:10.1002/jps.21394

Hardegger, L. A., Kuhn, B., Spinnler, B., Anselm, L., Ecabert, R., Stihle, M., . . . Diederich, F. (2011). Systematic Investigation of Halogen Bonding in Protein-Ligand Interactions. *Angewandte Chemie-International Edition, 50*(1), 314-318. doi:10.1002/anie.201006781

Harrigan, M. P., Sultan, M. M., Hernandez, C. X., Husic, B. E., Eastman, P., Schwantes, C. R., . . . Pande, V. S. (2017). MSMBuilder: Statistical Models for Biomolecular Dynamics. *Biophysical Journal, 112*(1), 10-15.

Höllerhage, M., Matusch, A., Champy, P., Lombès, A., Ruberg, M., Oertel, W. H., & Höglinger, G. U. (2009). Natural lipophilic inhibitors of mitochondrial complex I are candidate toxins for sporadic neurodegenerative tau pathologies. *Experimental Neurology, 220*(1), 133-142. doi:10.1016/j.expneurol.2009.08.004

Husic, B. E., & Pande, V. S. (2018). Markov State Models: From an Art to a Science. *Journal of the American Chemical Society, 140*(7), 2386-2396. doi:10.1021/jacs.7b12191

IARC. (1993). IARC monographs on the evaluation of carcinogenic risks to humans. *IARC monographs, 56*, 599.

Illy, C., Quraishi, O., Wang, J., Purisima, E., Vernet, T., & Mort, J. S. (1997). Role of the occluding loop in cathepsin B activity. *Journal of Biological Chemistry, 272*(2), 1197-1202. doi:10.1074/jbc.272.2.1197

Iwaniak, A., Minkiewicz, P., & Darewicz, M. (2014). Food-Originating ACE Inhibitors, Including Antihypertensive Peptides, as Preventive Food Components in Blood Pressure Reduction. *Comprehensive Reviews in Food Science and Food Safety, 13*(2), 114-134. doi:10.1111/1541-4337.12051

Izco, M., Vettorazzi, A., Forcen, R., Blesa, J., de Toro, M., Alvarez-Herrera, N., . . . Alvarez-Erviti, L. (2021). Oral subchronic exposure to the mycotoxin ochratoxin A induces key pathological features of Parkinson's disease in mice six months after the end of the treatment. *Food and Chemical Toxicology, 152*. doi:10.1016/j.fct.2021.112164

Kakuda, K., Ikenaka, K., Kuma, A., Doi, J., Aguirre, C., Wang, N., . . . Mochizuki, H. (2024). Lysophagy protects against propagation of α- synuclein aggregation through ruptured lysosomal vesicles. *Proceedings of the National Academy of Sciences of the United States of America, 121*(1). doi:10.1073/pnas.2312306120

Kerstner, F., & Garda-Buffon, J. (2024). Mycotoxins in plant-based beverages: An updated occurrence. *Food Research International, 194*. doi:10.1016/j.foodres.2024.114863

Li, Q., Dong, Z., Lian, W. G., Cui, J. F., Wang, J., Shen, H. T., . . . Cui, H. J. (2019). Ochratoxin A causes mitochondrial dysfunction, apoptotic and autophagic cell death and also induces mitochondrial biogenesis in human gastric epithelium cells. *Archives of Toxicology, 93*(4), 1141-1155. doi:10.1007/s00204-019-02433-6

Li, S. S., Zhang, M., Wang, J. H., Yang, F., Kang, B., Xu, J. J., & Chen, H. Y. (2019). Monitoring the Changes of pH in Lysosomes during Autophagy and Apoptosis by Plasmon

Enhanced Raman Imaging. *Analytical Chemistry, 91*(13), 8398-8405. doi:10.1021/acs.analchem.9b01250

Miller, B. R., McGee, T. D., Swails, J. M., Homeyer, N., Gohlke, H., & Roitberg, A. E. (2012). MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *Journal of Chemical Theory and Computation, 8*(9), 3314-3321. doi:10.1021/ct300418h

Nie, T., Li, J. F., You, L., & Wu, Q. H. (2025). Environmental mycotoxins: A potential etiological factor for neurodegenerative diseases? *Toxicology, 511*. doi:10.1016/j.tox.2025.154056

Noé, F., & Clementi, C. (2017). Collective variables for the study of long-time kinetics from molecular trajectories: theory and methods. *Current Opinion in Structural Biology, 43*, 141-147. doi:10.1016/j.sbi.2017.02.006

Noé, F., & Fischer, S. (2008). Transition networks for modeling the kinetics of conformational change in macromolecules. *Current Opinion in Structural Biology, 18*(2), 154-162. doi:10.1016/j.sbi.2008.01.008

Noé, F., Schütte, C., Vanden-Eijnden, E., Reich, L., & Weikl, T. R. (2009). Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proceedings of the National Academy of Sciences of the United States of America, 106*(45), 19011-19016. doi:10.1073/pnas.0905466106

O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. *J. Cheminform., 3*(1), 33. doi:10.1186/1758-2946-3-33

Obafemi, B. A., Adedara, I. A., & Rocha, J. B. T. (2023). Neurotoxicity of ochratoxin A: Molecular mechanisms and neurotherapeutic strategies. *Toxicology, 497*. doi:10.1016/j.tox.2023.153630

Östberg, L. J., Höög, J. O., & Persson, B. (2024). Computational analysis of human medium-chain dehydrogenases/ reductases revealing substrate- and coenzyme-binding characteristics. *Chemico-Biological Interactions, 390*. doi:10.1016/j.cbi.2024.110876

Pedroni, L., Dall'Asta, C., Galaverna, G., & Dellafiora, L. (2025). Computational Perspectives on Amoxicillin and Staphylococcus Aureus in Mirror Life. *Global Challenges, 9*(9), e00051. doi:10.1002/gch2.202500051

Pedroni, L., & Dellafiora, L. (2025). MSM4MD - Markov state models for Molecular Dynamics simulations: Zenodo; 10.5281/zenodo.15126414. Retrieved from http://dx.doi.org/10.5281/ZENODO.15126414

Pedroni, L., Perugino, F., Kurtaga, A., Galaverna, G., Dall'Asta, C., & Dellafiora, L. (2023). The bitter side of toxicity: A big data analysis spotted the interaction between trichothecenes and bitter receptors. *Food Research International, 173*, 113284. doi:10.1016/j.foodres.2023.113284

Pérez-Hernández, G., Paul, F., Giorgino, T., De Fabritiis, G., & Noé, F. (2013). Identification of slow molecular order parameters for Markov model construction. *Journal of Chemical Physics, 139*(1). doi:10.1063/1.4811489

Perugino, F., Pedroni, L., Galaverna, G., Dall'Asta, C., & Dellafiora, L. (2024). A mechanistic toxicology study to grasp the mechanics of zearalenone estrogenicity: Spotlighting aromatase and the effects of its genetic variability. *Toxicology, 501*(153686), 153686. doi:10.1016/j.tox.2023.153686

Perugino, F., Pedroni, L., Galaverna, G., Dall'Asta, C., & Dellafiora, L. (2024). Virtual display of targets: A new level to rise the current understanding of ochratoxin A toxicity from a molecular standpoint. *Toxicology, 503*. doi:10.1016/j.tox.2024.153765

Plotegher, N., & Duchen, M. R. (2017). Crosstalk between Lysosomes and Mitochondria in Parkinson's Disease. *Frontiers in Cell and Developmental Biology, 5*. doi:10.3389/fcell.2017.00110

Sampognaro, P. J., Arya, S., Knudsen, G. M., Gunderson, E. L., Sandoval-Perez, A., Hodul, M., . . . Kao, A. W. (2023). Mutations in α-synuclein, TDP-43 and tau prolong protein half-life through diminished degradation by lysosomal proteases. *Molecular Neurodegeneration, 18*(1). doi:10.1186/s13024-023-00621-8

Sava, V., Reunova, O., Velasquez, A., & Sanchez-Ramos, J. (2006). Can low level exposure to ochratoxin-A cause parkinsonism? *Journal of the Neurological Sciences, 249*(1), 68-75. doi:10.1016/j.jns.2006.06.006

Schiller, J., & Zickermann, V. (2022). Binding of Natural Inhibitors to Respiratory Complex I. *Pharmaceuticals, 15*(9). doi:10.3390/ph15091088

Schrenk, D., Bodin, L., Chipman, J. K., del Mazo, J., Grasl-Kraupp, B., Hogstrand, C., . . . Bignami, M. (2020). Risk assessment of ochratoxin A in food. *EFSA Journal, 18*(5). doi:10.2903/j.efsa.2020.6113

Serrano-Civantos, M., Beraza, E., alvarez-Erviti, L., de Cerain, A. L., & Vettorazzi, A. (2025). Potential role of ochratoxin A in Parkinson's disease: a systematic review of current evidence. *Archives of Toxicology, Epub ahead of print*. doi:10.1007/s00204-025-03994-5

Shao, J. Y., Tanner, S. W., Thompson, N., & Cheatham, T. E. (2007). Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. *Journal of Chemical Theory and Computation, 3*(6), 2312-2334. doi:10.1021/ct700119m

Silva, L. R., Guimaraes, A. S., do Nascimento, J., Nascimento, I. J. D., da Silva, E. B., McKerrow, J. H., . . . da Silva, E. F. (2021). Computer-aided design of 1,4-naphthoquinone-based inhibitors targeting cruzain and rhodesain cysteine proteases. *Bioorganic & Medicinal Chemistry, 41*. doi:10.1016/j.bmc.2021.116213

Stefanis, L., Emmanouilidou, E., Pantazopoulou, M., Kirik, D., Vekrellis, K., & Tofaris, G. K. (2019). How is alpha-synuclein cleared from the cell? *Journal of Neurochemistry, 150*(5), 577-590. doi:10.1111/jnc.14704

Tao, Y. F., Xie, S. Y., Xu, F. F., Liu, A. M., Wang, Y. X., Chen, D. M., . . . Yuan, Z. H. (2018). Ochratoxin A: Toxicity, oxidative stress and metabolism. *Food and Chemical Toxicology, 112*, 320-331. doi:10.1016/j.fct.2018.01.002

Thomas, R. R., Keeney, P. M., & Bennett, J. P. (2012). Impaired Complex-I Mitochondrial Biogenesis in Parkinson Disease Frontal Cortex. *Journal of Parkinsons Disease, 2*(1), 67-76. doi:10.3233/jpd-2012-11074

Undiano, E., Millán-Pacheco, C., Ayala, M., & Monroy-Noyola, A. (2023). Computational study to find the goat serum albumin (GSA) binding site as A-esterase. *Chemico-Biological Interactions, 381*. doi:10.1016/j.cbi.2023.110564

van Duin, A. C. T., Dasgupta, S., Lorant, F., & Goddard, W. A., III. (2001). ReaxFF: A reactive force field for hydrocarbons. *Journal of Physical Chemistry A, 105*(41), 9396-9409. doi:10.1021/jp004368u

Vanden-Eijnden, E. (2014). Transition Path Theory. In G. R. Bowman, V. S. Pande, & F. Noé (Eds.), *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation* (pp. 91-100): Springer.

Veras, J. P. C., França, V. L. B., Carvalho, H. F., & Freire, V. N. (2024). Noncovalent binding of carbofuran to acetylcholinesterase from Homo sapiens, Danio rerio, Apis mellifera and Caenorhabditis elegans: Homology modelling, molecular docking and dynamics, and quantum biochemistry description. *Chemico-Biological Interactions, 388*. doi:10.1016/j.cbi.2023.110826

Wang, A. L., Lukas, T. J., Yuan, M., Du, N., Tso, M. O., & Neufeld, A. H. (2009). Autophagy and Exosomes in the Aged Retinal Pigment Epithelium: Possible Relevance to Drusen

Formation and Age-Related Macular Degeneration. *Plos One, 4*(1). doi:10.1371/journal.pone.0004160

Wang, H. F., Yang, Q., Liu, X. C., Xu, Z. L., Shao, M. L., Li, D. X., . . . Yang, H. T. (2023). Structure-based discovery of dual pathway inhibitors for SARS-CoV-2 entry. *Nature Communications, 14*(1). doi:10.1038/s41467-023-42527-5

Zhang, L. Y., Zhang, X. K., Chen, X. X., Zhang, W., Zhao, L. H., Wang, Z. X., & Guo, Y. P. (2024). Biodegradation of ochratoxin A by Brevundimonas diminuta HAU429: Characterized performance, toxicity evaluation and functional enzymes. *Food Research International, 187*. doi:10.1016/j.foodres.2024.114409

Zhao, L., Zhang, M. X., Pan, F., Li, J. Y., Dou, R., Wang, X. Y., . . . Cai, S. B. (2021). In silico analysis of novel dipeptidyl peptidase-IV inhibitory peptides released from Macadamia integrifolia antimicrobial protein 2 (MiAMP2) and the possible pathways involved in diabetes protection. *Current Research in Food Science, 4*, 603-611. doi:10.1016/j.crfs.2021.08.008

Zhong, D., Ahmad, S., Cheng, P. Y., & Zewail, A. H. (1997). Femtosecond nucleophilic substitution reaction dynamics. *Journal of the American Chemical Society, 119*(9), 2305-2306.

Grasping the duality of fumonisins versus ceramide synthase: a computational approach to study their role as inhibitors and substrates from a mechanistic standpoint

This chapter is based on the paper published on:

Florinda Perugino, Lorenzo Pedroni, Gianni Galaverna, Chiara Dall'Asta, Luca Dellafiora. (2025). Grasping the duality of fumonisins versus ceramide synthases: a computational approach to study their role as inhibitors and substrates from a mechanistic standpoint. Helyion, 11, 13, e43725. 10.1016/j.heliyon.2025.e43725 .

Author contribution: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation.

Abstract

Fumonisins are *Fusarium* mycotoxins contaminating cereals worldwide. Fumonisin B1 (FB1) and its hydrolysed form (HFB1) are among those of most concern to food safety. They chemically resemble sphingoid bases and may alter the sphingolipid rheostat by inhibiting the ceramide synthases (CerSs). They may be also CerSs substrates, but the basis of this dual role is still unclear though relevant for their toxicological understanding and risk assessment. A validated computational pipeline including docking and molecular dynamics was applied to study the molecular rationale underpinning such a dual role focusing on the human CerS5 as a case study in the light of the two mechanisms of catalysis proposed so far involving either a one- or two-steps mechanisms. The interaction of FB1 and HFB1 with the proposed CerS5 lateral site of interaction, which was previously described for CerS2 and the yeast orthologue, was studied describing HFB1, but not FB1, as an optimal candidate substrate to undergo the reaction through the one-step catalytic mechanism. Moreover, the diffusion of FB1 through the membrane was calculated to be lower than HFB1, hampering FB1 from getting to the lateral site providing a possible rationale for the lower formation of FB1 N-acyl derivatives. Conversely, both fumonisins could fit the hydrophilic acyl-CoA binding pointing to their capability to act as substrates through the two-steps mechanism of catalysis as well as resulting in CerS5 inhibitors when preventing the interaction of acyl- CoA. Concerning the interaction with CerS5, FB1 and HFB1 interacted more steadily compared to their N-acyl derivatives. Overall, these results may extend the current understanding of fumonisins mechanisms describing the diverse chemistry of FB1 and HFB1 and their N-acyl derivatives as able to diversely interact with CerS5. The results collected may have a possible impact on their cumulative risk assessment which is worthy of further dedicated investigations.

1. Introduction

Fumonisins, mycotoxins contaminating cereal grains and derived products worldwide, are primarily produced as secondary me- tabolites by certain Fusarium fujikuroi complex species, though some Tolypocladium and Aspergillus producing species were also identified (Li et al., 2024; Sultana et al., 2019). Fumonisin B1 (FB1; Figure 1) is the most prevalent and toxic among the fumonisins group (Verstraete, 2013), comprising 70–80 % of the total fumonisins (Gao et al., 2023). FB1 is commonly found in a variety of contaminated food and feed threatening human and animal health and causing economic loss globally (Chen et al., 2021). Occurrence studies revealed that FB1 can be found in contaminated commodities along with analogues and modified forms. Among them, the hydrolysed FB1 (HFB1; Fig. 1), which is predominantly produced upon alkali food processing, was found at concentration levels higher than the non-hydrolysed form in certain food products (Chen et al., 2021). From a toxicological standpoint, FB1 was classified as possibly carcinogenic to humans (IARC group 2B) (International Agency for Research on Cancer & Cancer, 2002), and much evidence in vitro and in vivo showed that FB1 can be toxic to animals at a multi-organ level (Wang et al., 1991; Yu et al., 2021). HFB1 showed a similar toxicological profile though it seemed less toxic than FB1 (Chain et al., 2018). With respect to the molecular mechanisms of action, the structural similarity of FB1 and HFB1 to sphingoid bases (Figure 1) enables their interaction with the ceramide synthases (CerSs). Such interaction may inhibit the enzymes (Khan et al., 2024) and disrupt the sphingolipids metabolism and rheostat (Guerre et al., 2024; Seefelder et al., 2003). As an example, Wang et al. demonstrated that FB1 may reduce the activity of CerSs in rat microsomes by 50 % at 0.1 μM (Wang et al., 1991). Furthermore, FB1 was also found to inhibit de novo sphingolipid biosynthesis by neuronal cells with IC50 in the low μM range resulting in a limiting ceramide synthesis that differentially affects the formation of sphingomyelin and glycosphingolipid (Merrill et al., 1993). However, other mechanisms involving other proteins and enzymes related to the metabolism of sphingolipids were hypothesized too (Dellafiora, Galaverna, & Dall'Asta, 2018). Besides the inhibition of CerSs, FB1 and HFB1 may also act as CerSs substrates forming N-acylated derivatives with fatty acids of various lengths (Seiferlein et al., 2007) (Figure 1) pointing to their dual role either as CerSs inhibitors or substrates (Harrer et al., 2015). The mechanisms and route of formation of N-acylated forms, as well as their toxicology, still need further clarifications though they tested highly toxic in certain systems (Chain et al., 2018). Importantly, the reason why FB1 and HFB1 can act either as CerSs inhibitors or substrates is still poorly understood, as well as the entry path of sphingoid bases to the catalytic site of CerSs. In fact, in CerS6 the catalysis has been proved to occur

through a two-steps (ping-pong) catalytic mechanism involving the acyl-CoA substrate binding site (Pascoa et al., 2025), while a lateral binding site has been observed in CerS2 (Zelnik et al., 2023) and yeast CerS (T. Xie et al., 2023) suggesting that the substrate may bind human CerSs also orthogonally resulting in a one-step mechanism of catalysis. Within such a blurred scenario, key aspects related to fumonisins mechanisms of toxicity, crucial for their risk assessment, should be clarified toward a better understanding of their toxicity in living organisms (Lambert & Lipscomb, 2007). In particular, the precise understanding of the mechanisms of action is crucial to group fumonisins while perfoming cumulative risk assessment. In this respect, although HFB1 was described to be a detoxified form (Chain et al., 2018), the current cumulative risk assessment does not consider the FB1 modified forms, such as HFB1 and N-acyl derivatives, due to the insufficient data available and the limited understanding of the underpinning mechanisms of action (Chain et al., 2018). Indeed, risk assessors stated the urgent need for further data, including those relevant to toxicodynamics and modes of action of FB1 modified forms such as HFB1 and N-acylated forms (Chain et al., 2018).To clarify the molecular basis of the dual role of FB1 and HFB1 described above and investigate their role as inhibitor and/or substrates with respect to the two mechanisms of catalysis proposed, the present study applied a previously validated 3D molecular modelling pipeline (Pedroni et al., 2024). Of note, computational approaches are getting widely used to study protein-ligand interactions (e.g. Refs. (Moussa et al., 2021; L. Xie et al., 2024)). Regarding the study of enzymes action and inhibition, 3D molecular modelling – as we used in this work – may provide an effective means to evaluate whether chemicals may act as substrates or inhibitors for a variety of enzymes investigating the interaction with the designated binding sites (Dellafiora et al., 2015; Xia et al., 2024). Specifically, this study took advantage from the recent release of cryo-EM structures of yeast CerS (Schafer et al., 2025; T. Xie et al., 2023). Indeed, the yeast structure, the closest structurally resolved orthologous at the time of the analysis, was used as template to derive the human ceramide synthase 5 (CerS5) through advanced homology modelling to derive a reliable protein 3D structure, as previously described (Ben Boubaker et al., 2023). The yeast enzyme preferentially synthetizes C:26-phytoceramide starting from phytosphingosine and hexacosanoyl-CoA (C:26) (T. Xie et al., 2023). Nevertheless, the relatively high identity and similarity percentage, both crucial to successfully derive homology models (Jumper et al., 2021), enabled the modelling of the human CerS5, for which a structural characterization is still lacking. Human CerSs differ in tissue-specific expression pattern and in fatty acyl-CoA substrate chain length preference (Mesicek et al., 2010). CerS5 is mainly expressed in lung epithelial cells, but it has been also

detected in brain tissues (M. Zhang et al., 2023) where it is involved in the synthesis of C:16-ceramide starting from either sphingosine or sphinganine and hexadecanoyl-CoA (C:16).

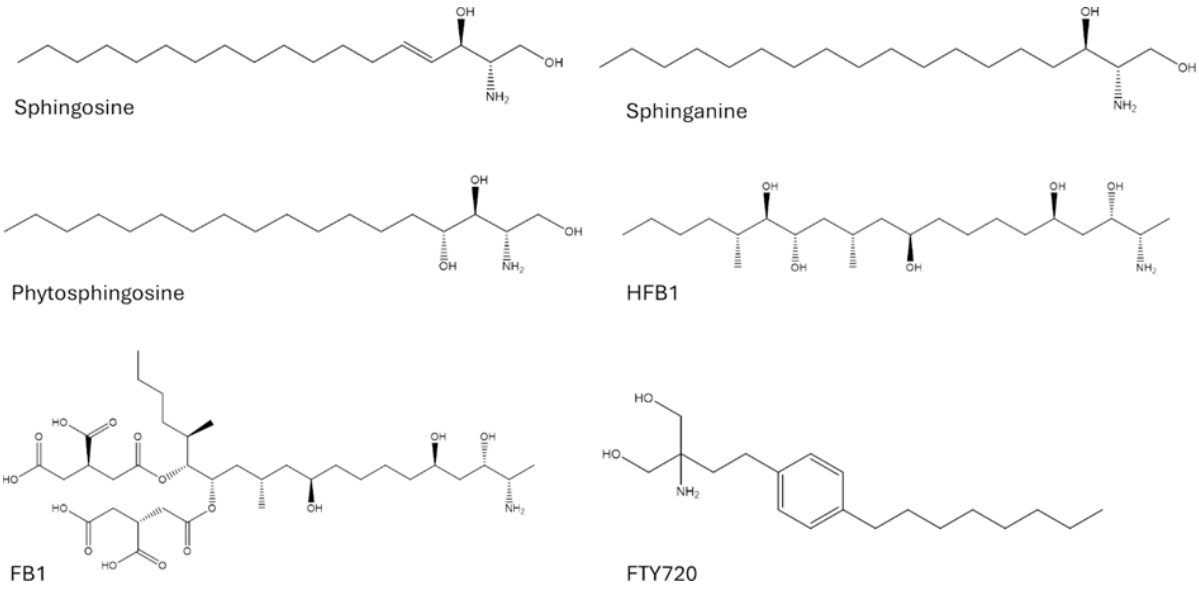


Figure 1. Chemical structure of molecules under analysis.

2. Material and methods

2.1. 3D data retrieval for ligands and protein

The public open database PubChem (https://pubchem.ncbi.nlm.nih.gov) (Kim et al., 2023) was used to retrieve the structural information of the ligands under analysis (last database access October 7, 2024). The 3D structures of sphingosine (CID 5280335) and FTY720 (CID 107969) were retrieved in the.sdf format and subsequently converted to Tripos.mol2 format through OpenBabel (O'Boyle et al., 2011), version 3.1.1, setting the protonation state as per pH 7 (i.e. setting -p 7 option) to obtain the proper protonation state under physiological conditions, while keeping the other options as per default setting. In the lack of 3D information for FB1, HFB1, sphinganine and phytosphingosine, their 2D structures were retrieved (CID 2733487, 21119850, 91486 and 122121, respectively) (Kim et al., 2023) in the.sdf format and converted in the 3D Tripos.mol2 structure through OpenBabel (O'Boyle et al., 2011) setting the protonation state as per pH 7 and using the –3dgen option to generate the 3D structures. The

3D structures of C:16 and C:26 N-acyl-FB1 were retrieved from the Protein Data Bank (PDB; https://www.rcsb. org), a global archive that provides data free of charge and without restrictions on usage, from structure 8IZD and 8ZB1, respectively and modified with PyMol (version 2.5.0) builder tool to obtain C:16 and C:26 N-acyl-HFB1 structures. The consistency of atom type assignment and stereochemistry of generated 3D structures was ensured upon visual inspection via PyMol software (version 2.5). The 3D structures of yeast CerS (Uniprot AC P28496) were retrieved from PDB (ww, 2019) with code 8IZD (T. Xie et al., 2023), 8QTN and 8QTR (Schafer et al., 2025). Specifically, the model for yeast enzyme was derived from the PDB structure 8IZD removing all co-purified low-molecular weight molecules but the cofactor hexacosanoyl-CoA. The pairwise sequence alignment of human and yeast CerSs was done with the EMBL-EBI webserver using with default setting the global alignment based on the Needleman-Wunsch algorithm (https://www.ebi.ac.uk/jdispatcher/psa) (Madeira et al., 2022). The modelling of human CerS5 was performed with ColabFold (v1.5.5) (Mirdita et al., 2022), an AlphaFold2-based notebook, starting from its primary sequence retrieved from the public and unrestricted UniProt repository (https://www.uniprot.org; UniProt AC Q8N5B7). Briefly, ColabFold as per manufacturer declaration, is an easy-to-use software for protein structure and complex prediction using AlphaFold2 and Alphafold2-multimer (Jumper et al., 2021) notebook (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb). For the sake of the project, the EM-cryo CerS structure having PDB code 8IZD, the closest orthologous to human CerS5 at the time of analysis, was used as template, while using the other parameters as per default setting. Once obtained the model for human CerS5, the inclusion of acyl-CoA to obtain a CerS5-cofactor binary complex was achieved editing the structure with PyMol (version 2.5) as follows. The PDB structure having code 8IZD (see above) was superimposed to the CerS5 model, the hexacosanoyl-CoA (C:26) merged within the CerS5 cofactor binding site and edited at the fatty chain by removing the last ten carbons to obtain hexadecanoyl-CoA (C:16), which is the optimal cofactor for CerS5 (Levy & Futerman, 2010). The hexacosanoid acid (C:26) and palmitic acid (C:16) bound state binding site of yeast CerS and human CerS5, respectively, were derived from the structures 8Y2M (Z. Zhang et al., 2024) and 8QZ7 (Pascoa et al., 2025) using PyMol (version 2.5) as following: the acylated PDB structures (8Y2M and 8QZ7) were superimposed to the yeast and CerS5 model and the hexacosanoid acid (C:26) and palmitic acid (C:16) merged within the yeast and CerS5 binding site, respectively. In agreement with previous studies (Louisse et al., 2022), the so obtained structure was further relaxed and energetically minimized to avoid steric clashes and

to correct improper geometries by means of GROMACS (Abraham et al., 2015), version 2021.4, with CHARMM36 all-atom force field parameters support (Huang & MacKerell, 2013). In particular, the steepest descent minimization algorithm was used with an energy step size of 0.01, a maximum of 5000 steps allowed and an energy threshold to stop minimization set at 10.0 kJ/mol. The 3D structure of human CerS6 (UniProt AC Q6ZMG9) was retrieved for a subsequent comparison from PDB with the code 8QZ7 (https://www.rcsb.org/structure/8QZ7) (Pascoa et al., 2025) keeping only the protein chain.

2.2. Docking studies

As described elsewhere (Pedroni et al., 2024), molecular docking provides the architectures of binding of ligands at the designated binding site of a target protein and was performed in this study with the software GOLD (Genetic Optimization for Ligand Docking; version 2022). Each simulation was computed in triplicate to check the score assignment consistency, and scores were reported as mean ± coefficient of variation. In this respect, the higher the score, the better the expected interaction, as indicated in the manufacturer declarations, while negative and/or variable score assignment among replicates indicate the incapability of ligands to dock at the designated binding site, as previously described (Dellafiora, Galaverna, Cruciani, et al., 2018; Dellafiora, Oswald, et al., 2020). Concerning the docking of phytosphingosine at the catalytic lateral site of yeast model, the binding site was defined in a 5 Å-radius sphere centered close to the thioester of hexacosanoyl-CoA. Default parameters were used and PLPScore was chosen as the scoring function being found to well-predict the ligands architecture of binding (the higher the score, the more likely the predicted pose, as per manufacturer declaration; https://www.ccdc.cam.ac.uk). Ten poses were generated and only the best scored one was carried forth to the analysis. Based on crystallographic data, the positioning of phytosphingosine at the lateral channel accessing the catalytic site was facilitated by setting the lipid fragment retraced in the PDB structure 8IZD as a template with a 50 units constraint weight. Concerning the docking of FB1, HFB1 and sphingosine and sphinganine within the human CerS5 model, similar settings were used. With respect to the lateral binding site, the placement of amino group was set awarding the occupancy of a 3.5 Å-radius sphere close to the thioester using the “region” option with a score contribution of 100 units. Concerning the docking at the CoA binding site proposed for the ping-pong mechanism, the binding site was set within 10 Å-radius sphere at the centroid of the pocket and setting the architecture of FB1 described for yeast CerS (Z. Zhang et al., 2024) as template. For the

docking at the inhibitor site for yeast CerS, CerS5 and CerS6 the binding site was set within a 10 Å-radius sphere at the centroid of the inhibitor pocket when in complex with FB1 and HFB1, while was set at 20 and 25 Å-radius sphere at the centroid of the inhibitor pocket when in complex respectively with C:16 and C:26 N-acyl FB1 and N-acyl HFB1 due to their length. Eventually, the architecture of FB1 described in a previous study was set (Schafer et al., 2025) as a template (constraint weight of 100 units). The comparison of docking scores was statistically analysed using the Student's t-test post hoc performed with SPSS IBM (v. 29.0, SPSS Inc., Chicago, IL, USA).

2.3. Molecular dynamics

Conventional molecular dynamics (CMD) simulations were used to monitor the geometrical stability of each CerS-ligand complex over time. Molecular dynamics (MD) simulations were performed with the GROMACS software (version 2021.4) using the best scored docking pose for each ligand as input structure in each system. Since CerSs are transmembrane proteins, prior to running the MD simulation, each complex was embedded into a dipalmitoylphosphatidylcholine (DPPC) membrane using the CHARMM-GUI (Jo et al., 2008) membrane builder tool (https://www.charmm-gui.org), setting the box size as small as possible, which parametrized the whole system with CHARMM36 all-atom force field support (Best et al., 2012). CHARMM-GUI is an online interface designed to facilitate the construction of biologically realistic membrane systems. Each input membrane-protein-ligand complex was solvated with TIP3P waters in a cubic periodic boundary condition, and counter ions (Na+ and Cl-, 0.1 M) were added to neutralize the system. Concerning the study of CerS-ligand complex stability, prior to MD simulation, the minimization and equilibration steps provided by CHARMM-GUI were the following: a maximum of 5000 steps for minimization to avoid steric clashes, and 250000 steps for isothermal (300 K) and 975000 steps for isobaric (1 bar) equilibrations. Once the minimization and equilibrations steps were completed, a 50 ns simulation (300 K with a coupling time of 0.1 ps and 1 bar with a coupling time of 2 ps) was run. Concerning the study of ligands inward pathway toward the designated substrate binding site, unbound systems were created and subsequently subjected to pull-in MD using GROMACS (version 2021.4), in agreement with previous studies (Hong et al., 2023; Subramanian et al., 2015). Specifically, for each ligand, the best scored docking pose generated as described above was detached from the acyl-CoA along the Y axis of about 4 nm – taking as the reference atoms ligands amino group and acyl-CoA thioester carbon. Each system was

then embedded within a DPPC membrane in a cubic box (12 nm each side), solvated and neutralized adding counter ions, as described above. Subsequently, each system was energetically minimized using the steepest descent algorithm, with a maximum of 50 000 steps and a max force of 1000 kJ/mol/nm, and equilibrated undergoing an isobaric (1 bar, coupling time 2 ps) 100 ps simulation. Next, for each system, a 500 ps steered pull-in dynamic along the Y-axis and toward the inner part of the substrate binding site was performed setting the pull rate at 0.015 nm/ps and the spring force constant at 500 kJ mol—1 nm—2. To provide an estimate of FB1 and HFB1 capability to permeate and diffuse across the lipid membrane, pull-in MD through the membrane and CMD within the membrane were performed. For both simulations, the DPPC membrane was built using the CHARMM- GUI membrane builder tool. Concerning the pull-in MD, FB1 and HFB1 were placed 2 nm above the membrane before the systems were solvated, neutralized and energetically minimized as described above. After that, fumonisins were pulled from the starting position toward the inner part of the membrane along the z axis setting the pull rate at 0.00015 nm/ps and the spring constant force at 1500 kJ ·mol-1·nm-2 for 600 ps. CMD simulations were then performed to check the stability and the capability to diffuse of FB1 and HFB1 within the membrane. For this reason, both molecules were placed within the lipid bilayer. After that fumonisins-membrane complexes were solvated, neutralized and energetically minimized as described above before running a 25 ns simulation. For the sake of clarity, the workflow used to derive the human CerS5 model and to study the interaction of ligands under analysis is represented in Figure 2, which represents the key steps in a flowchart.

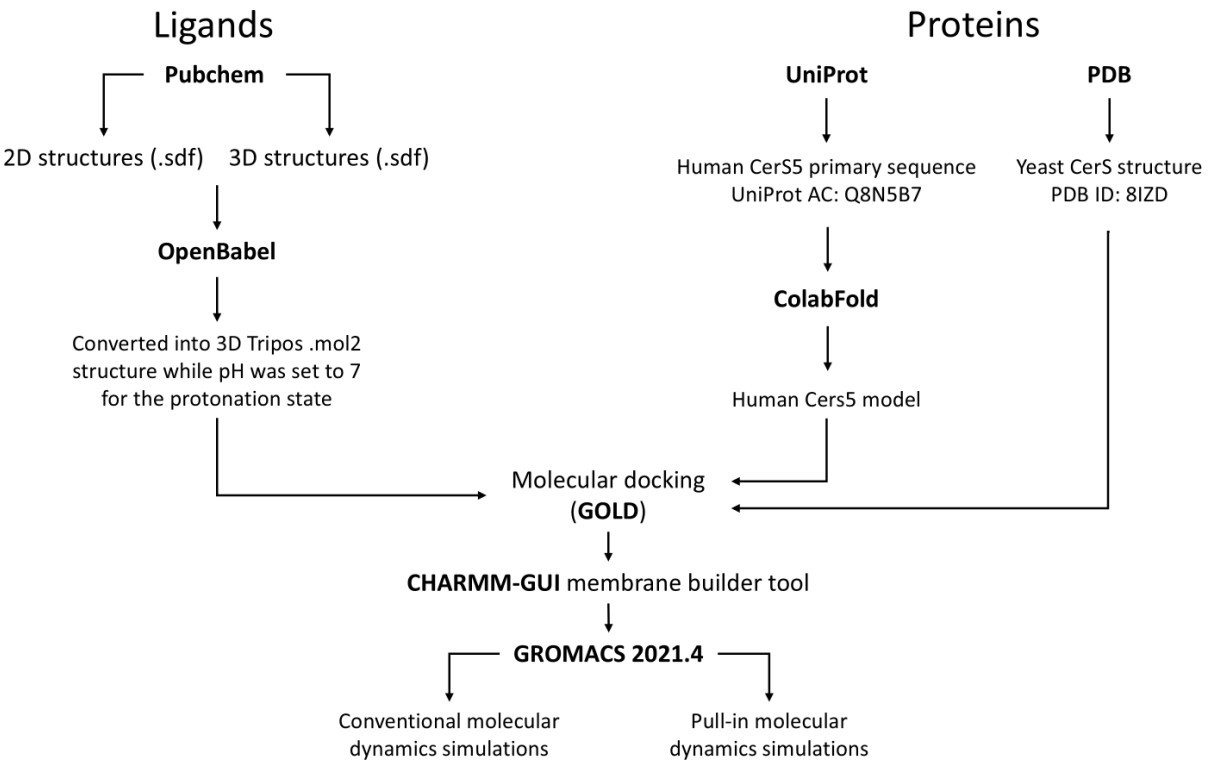


Figure 2. Key steps of the workflow used.

3. Results and discussion

The precise mechanics of CerS activity, including the aspects related to its inhibition by fumonisins, have been uncertain for de- cades. In particular, the paucity of structural data with respect to such a complex system prevented us from drawing a clear picture of CerSs mechanics in terms of substrate and inhibitor interaction sites and recruitment, making unclear and controversial how and where fumonisins can target the enzyme. This missing information made also unclear the reason why FB1 and its derived forms such as HFB1 and N-acylated derivatives could act either as CerSs substrates or inhibitors, blurring the overall understanding of their mechanisms of toxicity. However, the recent release of validated structural information for yeast CerS (Schafer et al., 2025; T. Xie et al., 2023) and for the human CerS6 (Pascoa et al., 2025) offered a breakthrough to understand CerSs functionality. This could also enable the proficient application of 3D molecular modelling ap- proaches to deepen the mechanics of CerSs activity, including those related to fumonisins toxicity (as per this study). Here a consolidated pipeline integrating docking analysis and MD simulations was applied (Dellafiora, Gonaus, et al., 2020; Pedroni et al., 2024). As a general comment, although docking simulations may provide initial binding affinity estimates, they have inherent limitations due to force field approximations and conformational sampling that can lead to possible artifacts and preventing reliable quantitative comparisons among ligands and estimates of complex stability over time. For this reason, docking analysis was coupled to MD simulations as crucial to refine docking outcomes, in agreement with previous studies (Louisse et al., 2023; Louisse et al., 2022). The study first verified the lateral entry path and site of interaction of substrates related to the one-step mechanism proposed by Xie et al. for the yeast CerS (T. Xie et al., 2023) and human CerS2 (Zelnik et al., 2023), as well as the interaction at the acyl-CoA binding sites observed in yeast CerS and CerS6 (Pascoa et al., 2025; Z. Zhang et al., 2024). To do so, as a proof-of-principle validation, the capability of phytosphingosine (the reference ligand for yeast CerS) to reach and interact with the lateral binding site of yeast CerS was addressed. Similarly, it was also accounted the ping-pong mechanism and the associated inhibition mechanisms studying the respective capability of FB1 and N-acyl-FB1 (known inhibitors of the yeast CerS) to interact at the acyl-CoA binding site. Once the procedure was validated, the human CerS5 was modelled and subjected to the same pipeline to study the dual role of FB1 and HFB1 as CerS substrate/inhibitor. In particular, the work: i) studied the capability of sphingosine and sphinganine (natural substrates of CerS5), FB1 and HFB1 to dock at the lateral catalytic site; ii) investigated their inward

pathway and related capability to move toward the inner part of the lateral catalytic site; iii) studied the mode of interaction of FB1, HFB1 and their N-acyl derivatives at the acyl-CoA binding site collecting outcomes related to the ping-pong mechanism of catalysis and to the mechanism of inhibition.

3.1. Modelling of the human CerS5

The recent release of validated 3D information made possible the application of advanced methodologies to derive the 3D model of human CerS. CerS5 was chosen as the case study among the six human CerSs as it showed the highest degree of sequence identity and similarity to the yeast orthologous (i.e. 19 % and 33 %, respectively; see Material and Methods for further details). The CerS5 model was obtained through an AlphaFold2-based notebook (see Material and Methods for further details). In brief, AlphaFold2 is a modelling system based on machine learning combining neural networks and homology modelling that provides highly reliable 3D protein models (Yang et al., 2023). For the sake of this study, the yeast CerS (PDB code 8IZD) was used as the homologous template in the model generation. The human CerS5 model scored a good confidence value with a pLDDT >0.7, which indicates a good backbone prediction as per manufacturer declaration (https://alphafold.ebi.ac.uk) (Tunyasuvunakool et al., 2021). Therefore, after the inclusion of the acyl-CoA, the model was used for subsequent docking and MD studies upon minimization to avoid improper arrangement of side chains and steric clashes (see Material and Methods for further details), in agreement with previous study (Louisse et al., 2022).

3.2. Modelling the interaction at the substrate lateral binding site

The molecular modelling pipeline used here already succeeded in predicting the substrate-enzyme interaction on a variety of different systems (Dellafiora, Gonaus, et al., 2020; Pedroni et al., 2024). However, a fit-for-purpose validation was performed to check the case specific reliability. To that end, the yeast CerS-phytosphingosine complex was taken as reference for the one-step catalytic mechanism and the yeast CerS-FB1 complex for the ping-pong mechanism. The stability of the complexes was studied through docking and MD simulations. Afterward, the analysis addressed the human CerS5.

3.2.1. Yeast CerS

Phytosphingosine was docked at the designated lateral substrate binding site of the yeast CerS. The binding architecture could satisfy the physico-chemical requirements of the pocket considering the relatively high docking score recorded, as reported in Table 1 (104.710 ± 0.001 units; the higher the score, the better the expected interaction, see Material and Methods section for further details). Of note, docking simulations were run in triplicate to check the consistency and stability of score assignment. The very low variability reflected the unambiguous score assignment supporting the consistency of the predicted architecture of binding, as previously described (Dellafiora, Galaverna, Cruciani, et al., 2018; Dellafiora, Oswald, et al., 2020). As shown in Fig. 3A, phytosphingosine could arrange at the substrate side entrance proposed in the recent structural investigations retracing the interaction of the co-purified lipid (as per PDB structure 8IZD). The further refinement of docking outcome through CMD simulations also confirmed the hypothesis describing a stable interaction of phytosphingosine over the time that showed the atom undergoing the reaction (i.e. the amine nitrogen) stably close to the thioester's carbon of the acyl-CoA, i.e. the donor of the acyl group (average distance 0.70 ± 0.13 nm; Figure 3C). Once the capability to stably and favourably interact with the designated lateral substrate binding site was verified, the inward pathway and the capability to reach the inner region of this site were studied through pull-in MD simulations. As shown in Fig. 3B, the phytosphingosine nitrogen took around 320 ps to reach the distance observed at the beginning of CMD simulations (i.e. 0.73 nm; Figure 3B). Taken together, these results ultimately confirmed the procedure reliability and corroborated both the hypothesized substrate lateral binding site and the inward pathway proposed for CerS substrates in previous studies (T. Xie et al., 2023; Zelnik et al., 2023).

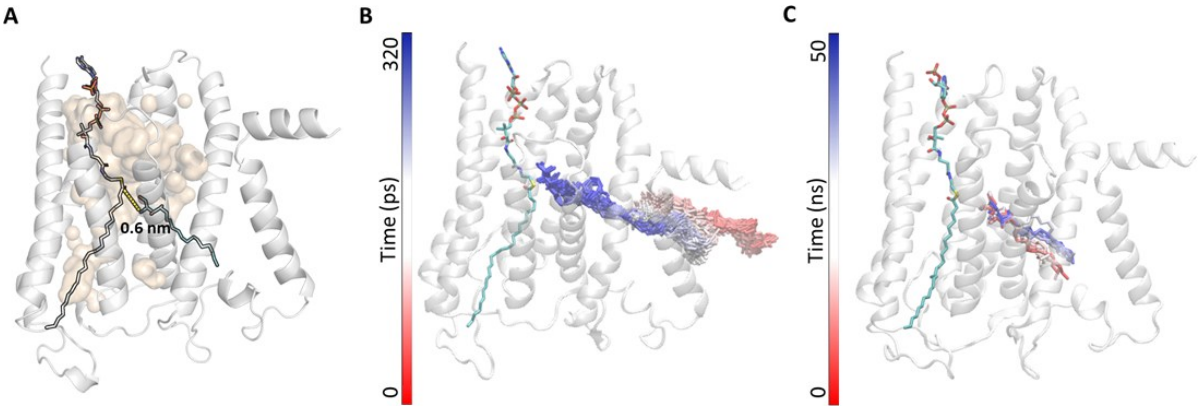


Figure 3. Docking and MD results for phytosphingosine at yeast CerS lateral binding site. **A.** Docking pose of phytosphingosine within the yeast CerS lateral binding site. The protein is represented in white cartoon, while phytosphingosine and acyl-CoA are represented in cyan and white sticks, respectively. The dashed yellow line represents the interatomic distance between the ligand's amine nitrogen and the thioester's carbon of the acyl- CoA (0.6 nm). The yellow surface indicates the space available to receive ligands. **B.** Pull-in MD results for phytosphingosine within the yeast CerS. The protein is represented in white cartoon, while the acyl-CoA is represented in cyan sticks. The time-step trajectory (from red, 0 ps, to blue, 320 ps) of the ligand is represented in sticks. **C.** CMD results for phytosphingosine within the yeast CerS. The protein is represented in white cartoon, while the acyl-CoA is represented in cyan stick. The time-step trajectory (from red, 0 ns, to blue, 50 ns) of the ligand is represented in sticks.

3.2.2. Human CerS5

Regarding human CerS5 and addressing the one-step mechanism of catalysis, sphingosine, sphinganine, FB1 and HFB1 were docked at the lateral binding site and those interactions that scored positive were refined by CMD simulations. The outcomes of docking and CMD simulations of sphingosine and sphinganine were in line with those observed for the yeast CerS-phytosphingosine complex recording a positive and consistent score among the replicates (152.320 ± 0.060 units; 139.330 ± 0.061 units, respectively; Table 1), as well as showing stable interaction and distances between the amine nitrogen and the thioester's carbon of the acyl-CoA (average distance 1.00 ± 0.17 nm and 0.78 ± 0.08 nm, respectively). Concerning the pull-in MD, in line with phytosphingosine and yeast CerS, sphingosine's and

sphinganine's nitrogen took respectively around 320 ps and 360 ps to reach the distance observed at the beginning of the CMD simulation (i.e. 0.62 nm for both) (Figure 4A–B). Taken together, these data eventually confirmed the applicability of the pipeline to the CerS5 model and confirmed that CerS5 may conserve the lateral substrate binding site and inward pathway of sphingoid bases described for the yeast enzyme and proposed for the human CerS2 in a previous study (Zelnik et al., 2023).

Table 1. Docking scores *

Yeast CerS – lateral binding site (one-step mechanism)	
Phytosphingosine	104.710 ± 0.001
FB1	105.067 ± 0.022
HFB1	98.943 ± 0.098
Yeast CerS – acyl-CoA binding site (ping-pong mechanism)	
Phytosphingosine	108.770 ± 0.006
FB1	162.817 ± 0.018
HFB1	120.993 ± 0.014
Yeast CerS – acyl-CoA binding site (inhibitor mechanism)	
FB1	158.333 ± 0.049
HFB1	137.427 ± 0.033
N-acyl-FB1	155.357 ± 0.123
N-acyl-HFB1	162.513 ± 0.044
Human CerS5 – lateral binding site (one-step mechanism)	
Sphingosine	152.320 ± 0.060
Sphinganine	139.330 ± 0.061
FB1	-116.617 ± 0.331
HFB1	96.760 ± 0.026

Human CerS5 – acyl-CoA binding site (ping-pong mechanism)	
Sphingosine	115.713 ± 0.011
Sphinganine	117.303 ± 0.004
FB1	173.707 ± 0.024
HFB1	124.767 ± 0.014
Human CerS5 – acyl-CoA binding site (inhibitor mechanism)	
FTY720	123.600 ± 0.007
FB1	179.957 ± 0.013
HFB1	140.133 ± 0.002
N-acyl-FB1	177.637 ± 0.039
N-acyl-HFB1	177.517 ± 0.040
Human CerS6 – acyl-CoA binding site (inhibitor mechanism)	
FB1	163.840 ± 0.055
HFB1	126.260 ± 0.024
N-acyl-FB1	179.253 ± 0.046
N-acyl-HFB1	174.830 ± 0.046

* Each simulation was computed in triplicate and scores are reported as mean ± coefficient of variation; the higher the score, the better the expected interaction, see Material and Methods section for further details

With respect to HFB1 and FB1, based on docking results, the former but not the latter was predicted to favourably arrange at the designated lateral substrate binding site. Indeed, HFB1 recorded a positive score (96.760 ± 0.026 units; Table 1) pointing to its capability to satisfy the physico-chemical requirements of the designated binding site. Of note, the score of HFB1 was significantly lower compared to that of sphingosine and sphinganine ($p < 0.05$; Student's t-test), though it showed a similar binding pose. This was reasonably due to the polar-hydrophobic interferences of the four hydroxyl groups on the HFB1 scaffold which arranged within the prevalently lipophilic environment of the binding site. This may explain, at least in part, the lower affinity of fumonisins to CerSs compared to sphingosine previously described in experimental trials (Humpf et al., 1998). Concerning the refinement of docking outcome with CMD simulations, HFB1 showed stable interaction and distances between the amine nitrogen and the thioester's carbon of the acyl-CoA (average distance 0.74 ± 0.13 nm), in line with the trends observed for sphingosine and sphinganine (Figure 4). Concerning the pull-in MD, the HFB1 nitrogen took around 380 ps to reach the distance observed at the beginning of CMD simulation (i.e. 0.57 nm; Figure 4C). Of note, HFB1 took more time to reach the CerS inner site compared to sphingosine, sphinganine and pythosphingosine, plausibly reflecting its lower capacity to diffuse through the membrane due to its lower lipophilicity, as discussed above. This differential capability to diffuse through the membrane may concur to explain the lower capability of CerS to transform HFB1 compared to the natural substrates, as discussed above. Taken together, these results consistently pointed to the likelihood of HFB1 as CerS5 substrate. Concerning FB1, as mentioned above, the negative score obtained (-116.617 ± 0.331 units) and its variability (more than 30 % from the average value) suggested its substantial incapability to satisfy the physico-chemical properties of the pocket and consequently its unfavourable arrangement at the designated lateral binding site (see Material and Methods for further details). This may suggest that FB1 may not be efficiently transformed by CerS5 through a one-step mechanism, though a certain degree of transformation could not be excluded completely in case FB1 can use a diverse route to access the catalytic site or it undergoes the reaction via the two-steps (ping-pong) mechanism. The incapability of CerS5 to efficiently transform FB1 via the lateral entry path was also confirmed by pull-in MD simulations as FB1 could not reach the inner part of the site within the considered timeframe (500 ps) (1.6 nm the lowest distance observed between the FB1 nitrogen and the acyl-CoA thioester carbon; Figure 6B). Therefore, FB1 was not calculated via CMD at the inner site due to the physico-chemical mismatch with the pocket requirements (as per the docking outcome), as well as due to the incapability to reach the inner site (as per pull-in MD outcome). Of note,

these results agreed with previous evidence that excluded such a route of access for FB1 in CerS6. For the sake of comparison with the yeast enzyme, FB1 and HFB1 were docked also at the lateral binding site of the yeast CerS. Interestingly, both recorded positive scores (105.067 ± 0.022 and 98.943 ± 0.098 units, respectively) pointing to their capability to favourably interact with that site (Figure 5). Of note, FB1 should be hampered to undergo such a pathway due to the limited capability to diffuse through the membrane (see paragraph below). However, CMD were used to refine docking outcome of both compounds showing that both could stably interact there (Figure 5). Considering that the diffusion in the membrane is a crucial feature (see below), this result pointed to the possible existence of differences among paralogues and orthologues in terms of accessibility to the catalytic site by certain ligands. In other words, the incapability of a certain ligand to interact with the lateral site of a specific CerS does not exclude the possibility to interact with the lateral site of other paralogues and/or orthologues requiring case-by-case investigations.

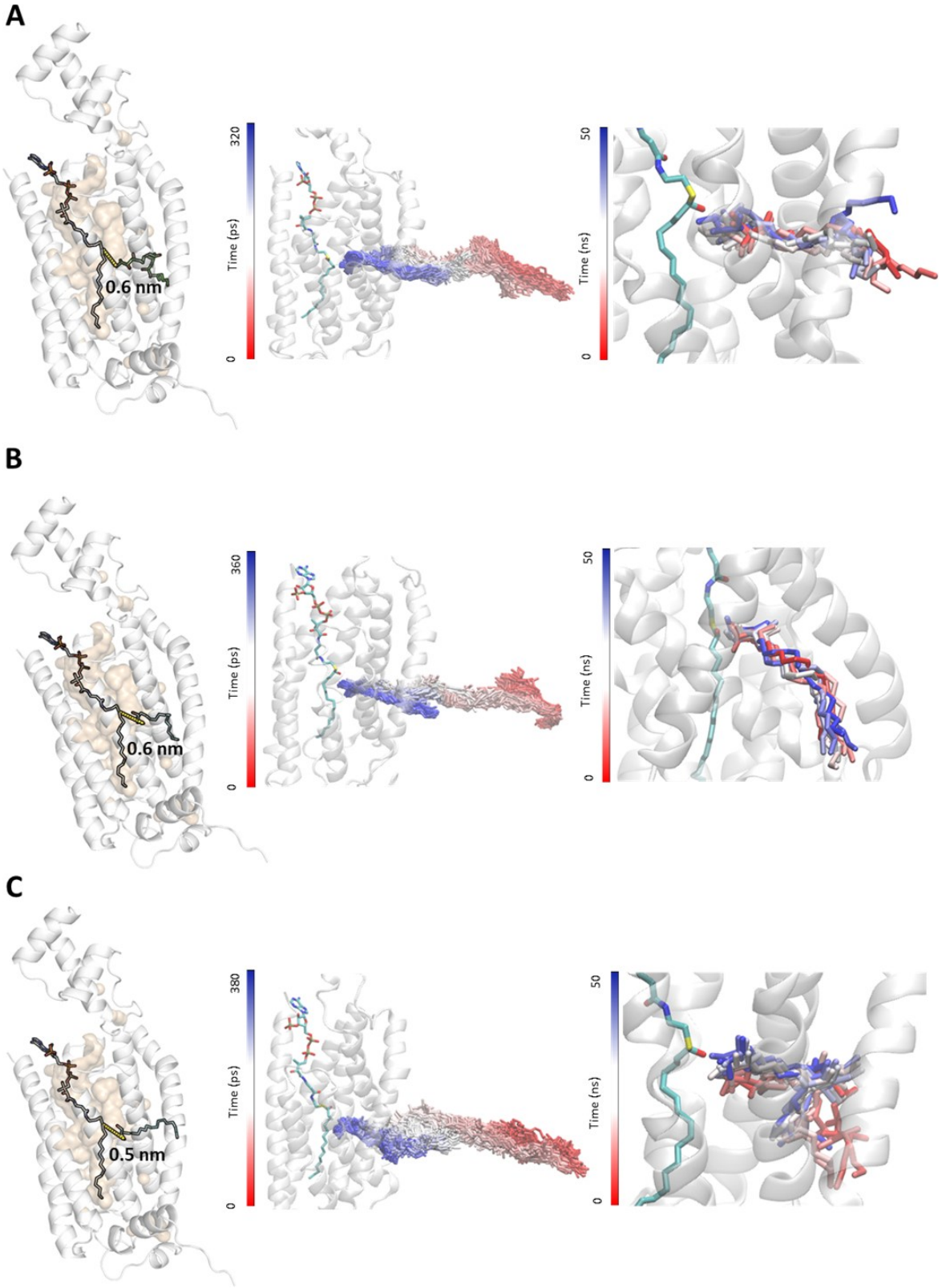


Figure 4. Docking and MD outcomes of human CerS5 at lateral binding site. Protein is represented in white cartoon, while the ligands and the co- factor are represented in sticks. The dashed yellow lines represent the interatomic distance between the ligand's amine nitrogen and the thio- ester's carbon of the acyl-CoA. The from-red-to-blue color switch

indicates the stepwise changes of the ligands' coordinates along the simulation. **A.** Sphingosine outcomes. From left, docking pose (0.6 nm the distance between the ligand's amine nitrogen and the thioester's carbon of the acyl- CoA), pull-in MD results (from red, 0 ps, to blue, 320 ps) and CMD trajectories (from red, 0 ns, to blue, 50 ns). **B.** Sphinganine outcomes. From left, docking pose (0.6 nm the distance between the ligand's amine nitrogen and the thioester's carbon of the acyl-CoA), pull-in MD results (from red, 0 ps, to blue, 360 ps) and CMD trajectories (from red, 0 ns, to blue, 50 ns). **C.** HFB1 outcomes. From left, docking pose (0.5 nm the distance between the ligand's amine nitrogen and the thioester's carbon of the acyl-CoA), pull-in MD results (from red, 0 ps, to blue, 380 ps) and CMD trajectories (from red, 0 ns, to blue, 50 ns).

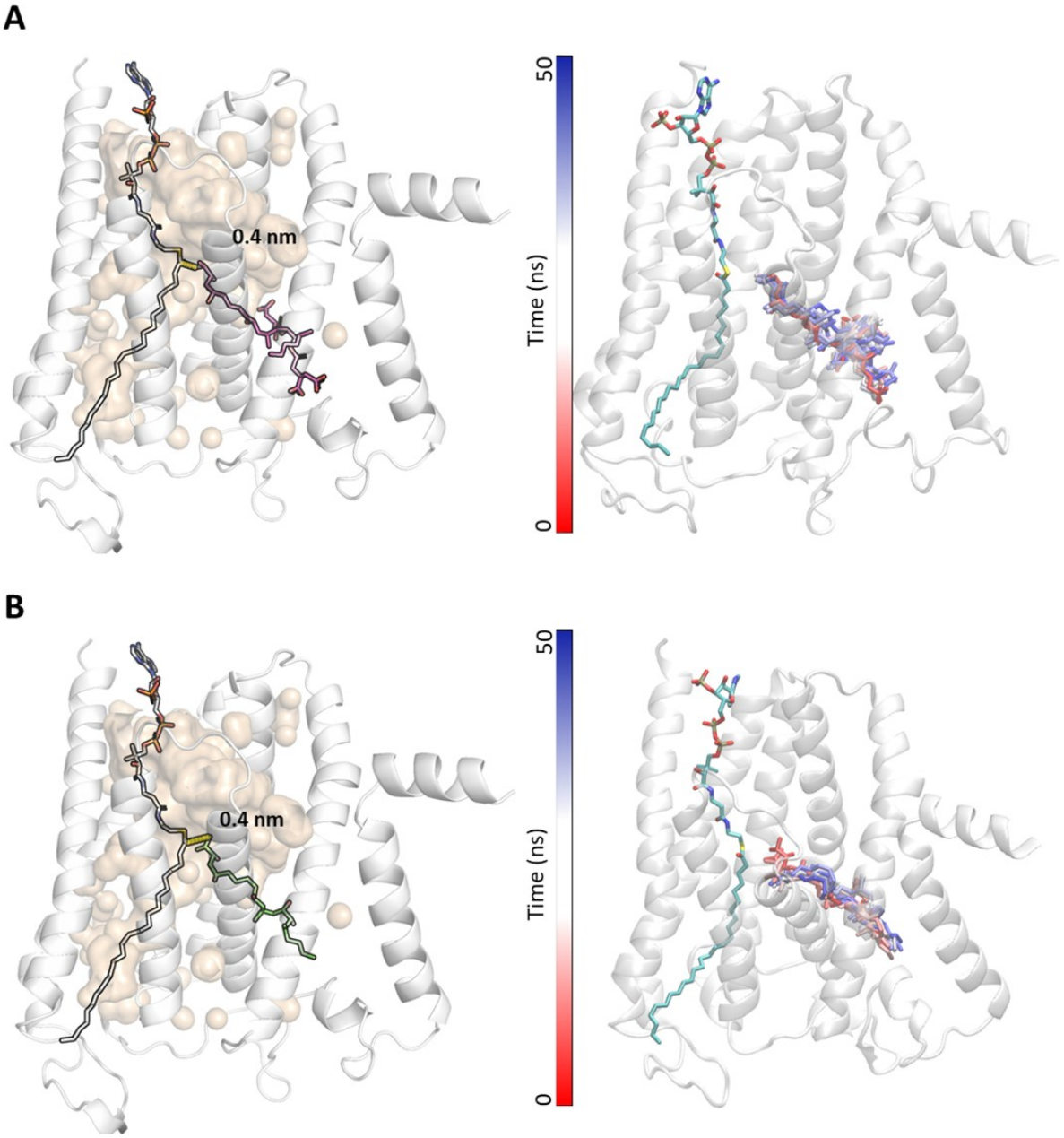


Figure 5. Docking and CMD results of FB1 and HFB1 within yest CerS lateral binding site. Protein is represented in white cartoon, while the ligands and the co-factor are represented in sticks. The dashed yellow lines represent the interatomic distance between the ligand's amine nitrogen and the thioester's carbon of the acyl-CoA. The from-red-to-blue color switch indicates the stepwise changes of the ligands' coordinates along the simulation. **A.** FB1 outcomes. On the left, docking pose of FB1 (0.4 nm the distance between the ligand's amine nitrogen and the thioester's carbon of the acyl- CoA). On the right, CMD trajectories (from red, 0 ns, to blue, 50 ns). **B.** HFB1 outcomes. On the left, docking pose of HFB1 (0.4 nm the distance between the ligand's amine nitrogen and the thioester's carbon of the acyl-CoA). On the right, CMD trajectories (from red, 0 ns, to blue, 50 ns).

3.3. Capability of FB1 and HFB1 to diffuse through the membrane

Interestingly, in pull-in MDs, the progression overtime of HFB1 toward the lateral substrate site compared to FB1 suggested that the latter could have a less favoured capability to move through the membrane. Indeed, as shown in Figure 6A, from around 150 ps onward HFB1 accelerated its progression toward the protein covering a higher distance over time compared to FB1, which had a rather steady progression. This suggested that HFB1 and FB1 might have a diverse capability to penetrate and diffuse through the membrane. This hypothesis was verified through dedicated pull-in MD and CMD simulations. Specifically, pull-in MD simulations investigated whether FB1 and HFB1 could move from the solvent to penetrate within the membrane. As shown in Figure 6C, pull-in MDs revealed that HFB1 (LogP —0.5 as per PubChem) but not FB1 (LogP 2.7 as per PubChem) could efficiently penetrate the membrane. Indeed, HFB1 permeated the membrane within 5 ns, while FB1 remained stuck at the surface till the end of the simulation. The higher hydrophilicity of the latter and its capability to interact with the positively charged heads of membrane lipids through the tricarballylic acid (TCA) groups were proposed as a rationale to explain such a different outcome. In addition, the capability of FB1 and HFB1 to diffuse through the membrane was envaulted via CMDs. The analysis revealed the higher mobility (namely, higher capability to diffuse) of HFB1 compared to FB1 (Figure 6D). Also in this case, the reduced mobility of FB1 was due to polar interaction between the TCA groups and the polar heads of the membrane. Taken together, these results pointed to the less suitability of FB1 to enter and move through the membrane's lipophilic environment, likely due to the higher polarity caused by the TCAs. This may reasonably determine a lower capability to reach the membrane-embedded lateral substrate site of CerSs providing an ancillary mechanistic rationale to explain the lower production of N-acyl-FB1 compared to N-acyl-HFB1 previously described (Harrer et al., 2013). These findings aligned also with the outcome of Schertz et al. that underlined either the lower absorption or metabolism of FB1 compared to HFB1 (Schertz et al., 2018). Although the ADME (Absorption, Distribution, Metabolism and Excretion) encompasses several mechanisms and events, the impaired capability of FB1 to diffuse thorough the membrane observed in this work may provide a compelling up-stream mechanistic rationale for the differences between FB1 and HFB1 in terms of absorption and metabolism described experimentally (e.g. Ref. (Harrer et al., 2013)).

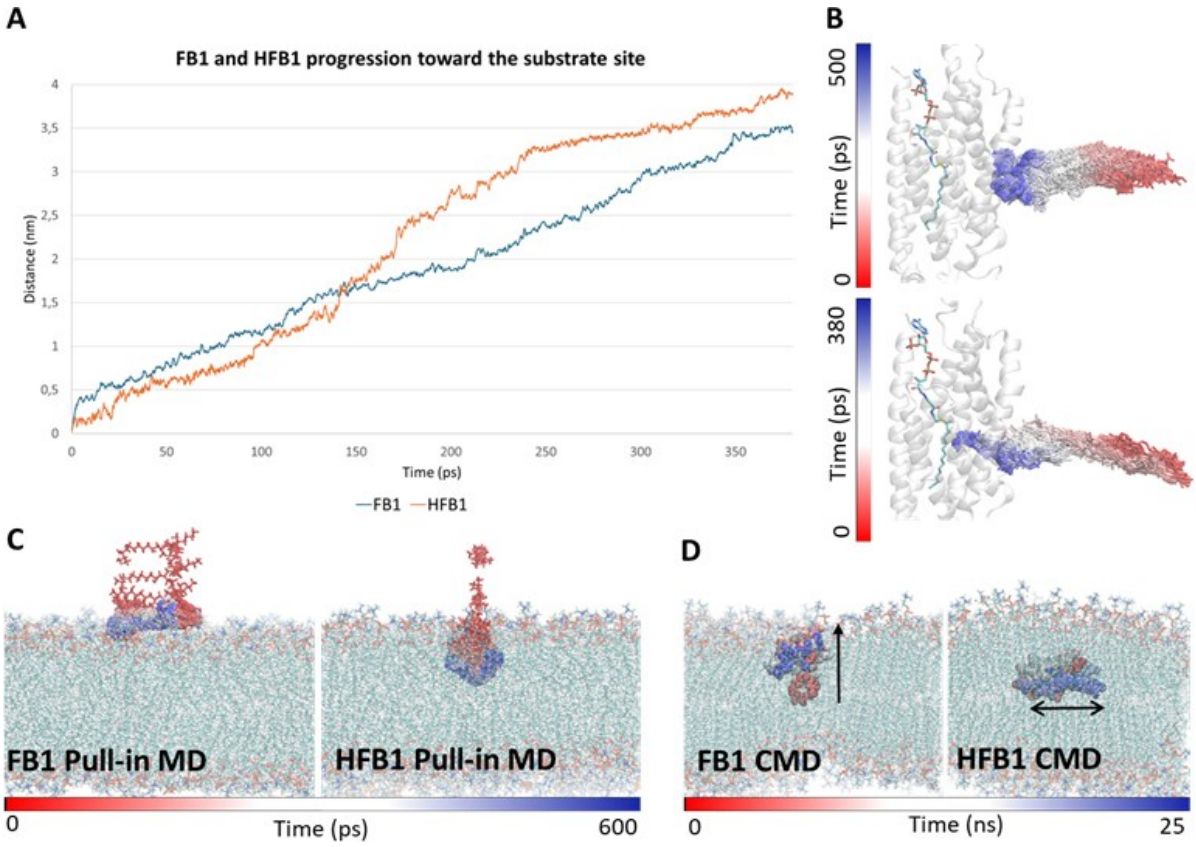


Figure 6. Molecular modelling outcomes of FB1 and HFB1 within the membrane. **A.** Pull-in MD plot of FB1 and HFB1 progression toward the substrate site from the beginning of the simulation (0 ps) until when HFB1 nitrogen reached the distance observed at the beginning of CMD simulation (380 ps). **B.** Pull-in MD results of FB1 (top) and HFB1 (bottom) within the human CerS5. The protein is represented in white cartoon, while the cofactor is represented in cyan sticks. The time-step trajectories (from red, 0 ps, to blue, 500 ps and 380 ps, respectively) of the ligand are represented in sticks. **C.** Pull-in MD results for FB1 (left) and HFB1 (right) within the DPPC membrane. The time-step trajectories (from red, 0 ps, to blue, 600 ps) of the ligands and the bilayer are represented in sticks. **D.** CMD results of FB1 (left) and HFB1 (right) within the DPPC membrane. The membrane bilayer is represented in sticks, while the time-step trajectories (from red, 0 ns, to blue, 25 ns) of ligands are represented in spheres. The black arrows indicate the movements of FB1 and HFB1 within the membrane.

3.4. Modelling the interaction at the acyl-CoA binding site

Previous studies suggested a ping-pong mechanism of catalysis for FB1 in CerS6 where the co-factor's acyl chain is transferred to the enzyme and subsequently to FB1 as it accesses the catalytic site through the acyl-CoA channel (Pascoa et al., 2025). Considering that the binding site of sphingoid bases remains debated and that it may vary among different isoforms and organisms, the capability of sphingosine and sphinganine to interact with the acyl-CoA binding site of CerS5, and that of phytosphingosine to interact with the acyl-CoA binding site of yeast CerS, was assessed via molecular docking. Results confirmed the capability of all the sphingoid bases to interact with the acyl-CoA binding tunnel (115.713 ± 0.011, 117.303 ± 0.004 and 108.770 ± 0.006 units, respectively; Table 1) revealing the capability of CerS5 and yeast CerS to synthetize ceramide also via ping-pong mechanism. The outcome collected (Table 1) underlined the possible existence of ping-pong mechanism for yeast CerS as supported by Schafer et al. and Zhang et al. (Schafer et al., 2025; Z. Zhang et al., 2024). To investigate the role of fumonisins in such a mechanism of catalysis, FB1 and HFB1 were docked at the acyl-CoA binding site of CerS5 bound to C:16 fatty acid. Docking analysis revealed the capability of both FB1 and HFB1 to interact with the designed binding site of CerS5 (173.707 ± 0.024 and 124.767 ± 0.014 units, respectively; Table 1) while correctly orienting the amine nitrogen toward the carboxyl group of the C:16 acyl residue (Figure 7A). Subsequently, CMD were performed to further refine the docking outcomes and the distance between fumonisins' amine nitrogen and the fatty acid carboxyl group was monitored along the simulation. FB1 and HFB1 showed a comparable average distance (0.36 ± 0.01 nm and 0.35 ± 0.03 nm, respectively). This was in line with FB1 in yeast, computed as a reference to validate the model (162.817 ± 0.01; average distance 0.33 ± 0.02 nm) (Figure 7B). For comparative purposes, the inter- action of HFB1 at the yeast CerS acyl-CoA binding site bound to C:26 was calculated as well. Although the docking score pointed to a favoured interaction (120.993 ± 0.014 units), the average distance calculated between HFB1's amine nitrogen and the fatty acid carboxyl group was higher and more variable (0.40 ± 0.09 nm) than the distance observed for FB1 (0.33 ± 0.02 nm) suggesting that the N-acylation of HFB1 via a two-step mechanism might be not preferential in yeast. Regarding CerS5, HFB1 showed a less stable geometry of interaction than FB1 possibly pointing to its less-favoured interaction with the acyl-CoA binding site (Figure 7B). This evidence integrated the lower docking score compared to that recorded for FB1 within CerS5 acyl-CoA binding site bound to C:16 (124.767 ± 0.014 and 173.707 ± 0.024, respectively; Table 1). In principle, this suggested that

HFB1 is likely less prone to undergo the ping-pong mechanism of catalysis compared to FB1, though the diffusion through the membrane is an upstream factor limiting the one-step reaction. Furthermore, it has been shown that FB1 and its N-acylated form can block the activity of yeast CerS competitively preventing the binding of acyl-CoA (Schafer et al., 2025; Z. Zhang et al., 2024). Therefore, it was investigated whether FB1, HFB1 and their N-acylated derivates could also act similarly for human CerS5. The known inhibitor FTY720 in complex with CerS5 (Berdyshev et al., 2009) and C:26 N-acyl-FB1 in complex with yeast CerS were also calculated for comparison. Docking results revealed the capability of FB1, HFB1, FTY720, N-acyl-FB1 and N-acyl-HFB1 to adopt a binding architecture comparable to that of FB1 in the yeast CerS and CerS6 (Pascoa et al., 2025; Schafer et al., 2025). They all satisfied the physico-chemical requirements of the designated pocket, as highlighted by the relatively high docking scores recorded (179.957 ± 0.013, 140.133 ± 0.002 and 123.600 ± 0.007, 177.637 ± 0.039 and 177.517 ± 0.040 units, for FB1, HFB1, FTY720, N-acyl-FB1 and N-acyl-HFB1, respectively; Table 1). However, CMD outcomes for yeast CerS refined those of docking studies (Table 1) showing a less stable interaction of FB1 compared to N-acyl-FB1 and N-acyl-HFB1 (Figure 7C), corroborating the results collected by Zhang et al. (Hu & Cao, 2024; Z. Zhang et al., 2024), reporting N-acyl-FB1 as a stronger inhibitor than FB1. The refinement of docking outcome aligned with previous evidence demonstrating that CMD simulations may refine the conclusion drawn based solely on the docking score assignment eventually improving the prediction (e.g. Ref. (Louisse et al., 2023)). Regarding CerS5, FB1 and the FB1 portion of its N-acyl derivative showed a similar stability, while the FB1 N-acyl derivative showed the acyl moiety stable at its respective region of the binding site. Considering this, the higher inhibitory activity of N-acyl derivatives with respect to FB1 and HFB1 described experimentally (Humpf et al., 1998; Z. Zhang et al., 2024) could be explained, at least in part, by the additional contributions of binding offered by the stable interaction of N-acyl moiety. Conversely, HFB1 showed a more stable interaction compared to its N-acyl-derivative (Figure 7C). These results pointed to the possibility that HFB1 N-acyl derivative may represent a weaker inhibitor for CerS5. Interestingly, this finding is apparently in contrast to what reported by Humpf et al. (Humpf et al., 1998) describing the higher inhibitory activity of N-acyl-HFB1 compared to FB1 and HFB1 tested in HT-29 cells. Of note, this cell model has a prominent expression of CerS2 and CerS6, while CerS5 is among the less abundant ones (as per The Human Protein Atlas database; https://www. proteinatlas.org (Uhlen et al., 2015)) suggesting that HT-29 cells do not provide a representative system to check the inhibition of CerS5. However, to investigate further this point, we also calculated

the interaction of FB1 and HFB1 and their C:16 N-acyl derivatives with CerS6 (Figure 7D). The outcome showed that the interaction of FB1 and N-acyl-FB1 was like that observed within CerS5, with N-acyl-FB1 more stable than FB1 and likely showing an inhibitory potential higher than FB1 for the same reasons described above. Concerning HFB1 and N-acyl-HFB1, the latter was more stable overall than the former. Therefore, N-acyl-HFB1 was expected having a higher inhibitory potential compared to HFB1 based on its stability and the additional contributions of binding brought by the acyl moiety. Similarly, N-acyl-HFB1 could be expected to have a higher inhibitory potential than FB1. These data aligned with previous findings (Seiferlein et al., 2007) and further validated the procedure applied. On the other hand, taken together, these findings suggested that the relative inhibitory potency of FB1/HFB1 and their N-acyl derivatives may be CerS isoform-specific and may depend on the cell type and organism, depending on the level and isoforms of the CerSs expressed.

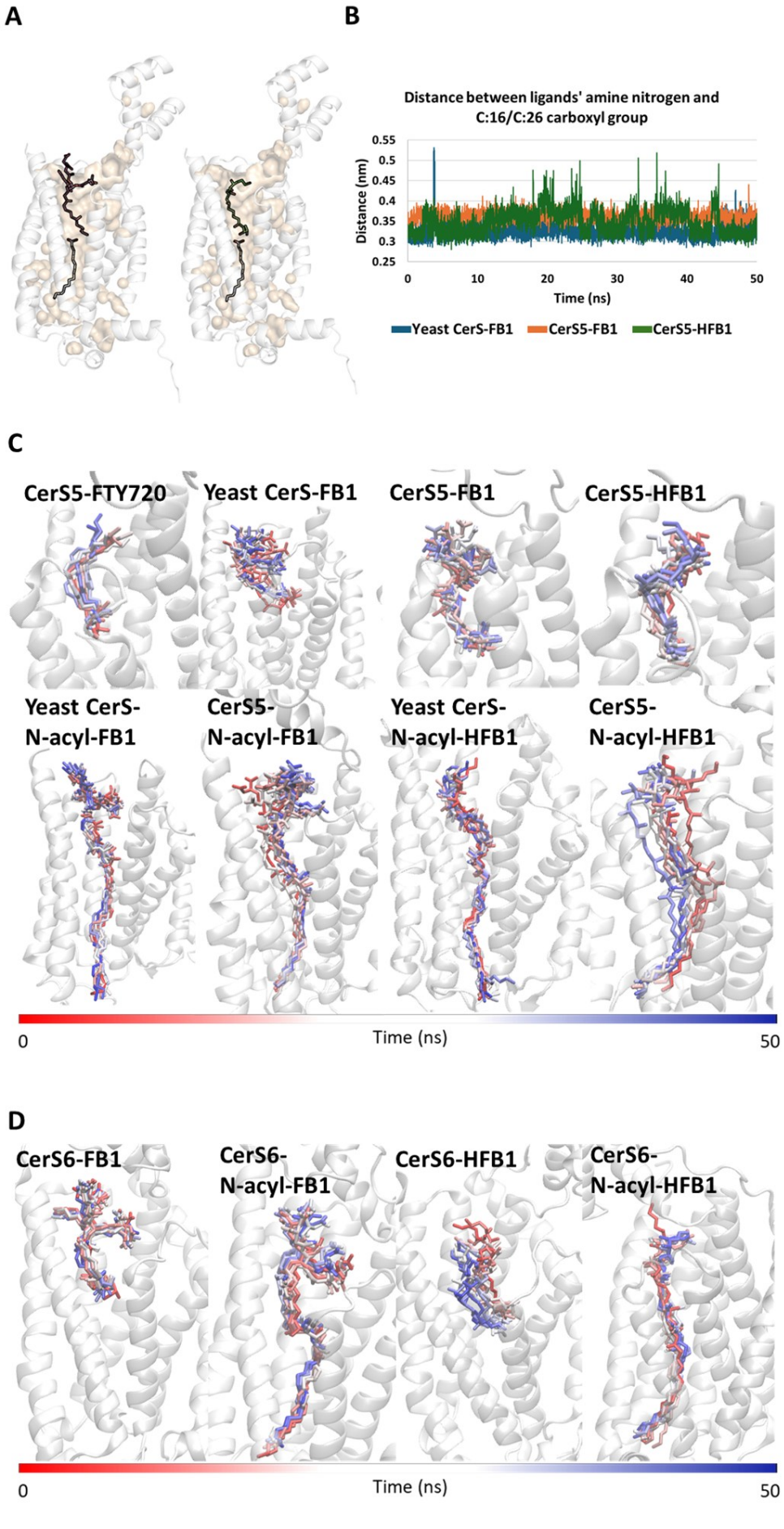

A
B
Distance between ligands' amine nitrogen and C:16/C:26 carboxyl group
Distance (nm)
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0
10
20
30
40
50
Time (ns)
Yeast CerS-FB1
CerS5-FB1
CerS5-HFB1
C
CerS5-FTY720
Yeast CerS-FB1
CerS5-FB1
CerS5-HFB1
Yeast CerS-N-acyl-FB1
CerS5-N-acyl-FB1
Yeast CerS-N-acyl-HFB1
CerS5-N-acyl-HFB1
0
Time (ns)
50
D
CerS6-FB1
CerS6-N-acyl-FB1
CerS6-HFB1
CerS6-N-acyl-HFB1
0
Time (ns)
50

Figure 7. Molecular modelling outcomes of yeast CerS, CerS5 and CerS6 at acyl-CoA binding site involved in ping-pong mechanism and CerSs inhibition. **A.** Molecular docking poses of FB1 (left) and HFB1 (right) involved in the ping-pong mechanism within CerS5 acyl-CoA binding site. Protein is represented as cartoon except for the acyl-CoA binding site which is represented as surface. FB1, HFB1 and C:16 acyl residue are represented as sticks. **B.** Distance plot of FB1 amine nitrogen – C:26 carboxyl group within yeast CerS and FB1/HFB1 amine nitrogen – carboxyl group within CerS5 along CMD simulations. **C.** Time-step trajectories of FTY720, FB1, HFB1, N-acyl-FB1 and N-acyl-HB1 within CerS5 and time-step trajectories of FB1, N-acyl-FB1 and N-acyl-HFB1 within yeast CerS. Proteins are represented as cartoon, while ligands are represented as sticks. The from-red-to-blue colour switch indicates the stepwise changes of the ligands' coordinates along the simulation (50 ns). **D.** Time-step trajectories of FB1, HFB1, N-acyl-FB1 and N-acyl-HB1 within CerS6. Protein is represented as cartoon, while ligands are represented as sticks. The from-red-to- blue colour switch indicates the stepwise changes of the ligands' coordinates along the simulation (50 ns).

4. Conclusion

In the lack of verified structural data for the human CerS5, the present study modelled its 3D structure as a proof-of-principle to investigate the interaction of FB1 and HFB1 exploring their dual role as substrates, either via the one-step or ping-pong mechanism, or inhibitors. The outcome collected suggested the existence of possible differences among FB1 and HFB1 in terms of toxicodynamics, specifically with respect to their diverse interaction with CerS5, which are worthy of further dedicated investigations. This is crucial to drive a more aware cumulative risk assessment for this class of food contaminants which still excludes modified forms like HFB1 and N- acylated forms due to a general shortage of data and scarce mechanistic understanding (Chain et al., 2018). As a general comment, this computational approach is grounded in experimental evidence and provides molecular explanations for observed phenomena whose mechanistic rationale was uncharacterized. Concerning the mechanisms of catalysis of CerS5, the present study corroborated the proposed substrates' lateral binding site and entry path proposed for the yeast enzyme and CerS2 (Zelnik et al., 2023) showing its likely existence also in CerS5. Concerning the activity of fumonisins as inhibitors, previous studies have demonstrated that FB1 reduces CerS activity and inhibits de

novo sphingolipid biosynthesis with concentrations in the low μM range. Our computational results align with these findings. Indeed, we show that both FB1 and HFB1 can specifically interact with and inhibit CerS5 competitively preventing the interaction of Acyl-CoA, specifically adding this CerS isoform to the list of targets responsible for the observed effects. This provides a molecular basis for their experimentally observed inhibitory effects, well-aligning with the most recent structural findings on CerS5 homologues (Pascoa et al., 2025; Schafer et al., 2025; Z. Zhang et al., 2024), though underlying possible differences in terms of catalytic site accessibility. Furthermore, experimental evidence confirms that both fumonisins can serve as CerS substrates forming N-acylated derivatives, with HFB1 showing enhanced metabolic conversion compared to FB1 (Harrer et al., 2013). Our modeling results corroborate this differential behavior, revealing that HFB1, but not FB1, can efficiently access the lateral catalytic site for the one-step mechanism, while both can undergo the ping-pong mechanism. Assuming that both mechanisms of catalysis may be possible, the possibility of HFB1 to have multiple pathways to the catalytic site of certain CerS isoforms may provide an additional mechanistic rationale for the enhanced conversion to N-Acyl derivatives compared to FB1 observed experimentally. Additionally, the simulation of membrane diffusion provides a mechanistic explanation for experimental observations of lower FB1 absorption and metabolism compared to HFB1, as we demonstrate FB1's impaired capability to diffuse through lipid membranes due to its TCA groups. These computational insights not only validate experimental findings but also reveal the underlying molecular mechanisms. Such mechanistic insights are essential for translating molecular interactions into toxicological understanding supporting evidence-based risk assessment frameworks for fumonisins and their modified forms, as claimed by risk assessors. Additionally, the experimental asset presented here provides a means for prioritization which is well aligned with the increasing emphasis on NAMs in toxicology and risk assessment to ultimately reduce animal testing and globally move towards a mechanistic understanding of toxicity as highlighted in the upcoming European Commission road map for the phasing out of animal studies.

References

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX, 1-2*, 19-25. doi:10.1016/j.softx.2015.06.001

Ben Boubaker, R., Tiss, A., Henrion, D., & Chabbert, M. (2023). Homology Modeling in the Twilight Zone: Improved Accuracy by Sequence Space Analysis. *Methods Mol Biol, 2627*, 1-23. doi:10.1007/978-1-0716-2974-1_1

Berdyshev, E. V., Gorshkova, I., Skobeleva, A., Bittman, R., Lu, X., Dudek, S. M., . . . Natarajan, V. (2009). FTY720 inhibits ceramide synthases and up-regulates dihydrosphingosine 1-phosphate formation in human lung endothelial cells. *J Biol Chem, 284*(9), 5467-5477. doi:10.1074/jbc.M805186200

Best, R. B., Zhu, X., Shim, J., Lopes, P. E., Mittal, J., Feig, M., & Mackerell, A. D., Jr. (2012). Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J Chem Theory Comput, 8*(9), 3257-3273. doi:10.1021/ct300400x

Chain, E. P. o. C. i. t. F., Knutsen, H. K., Barregard, L., Bignami, M., Bruschweiler, B., Ceccatelli, S., . . . Alexander, J. (2018). Appropriateness to set a group health-based guidance value for fumonisins and their modified forms. *EFSA J, 16*(2), e05172. doi:10.2903/j.efsa.2018.5172

Chen, J., Wei, Z., Wang, Y., Long, M., Wu, W., & Kuca, K. (2021). Fumonisin B(1): Mechanisms of toxicity and biological detoxification progress in animals. *Food Chem Toxicol, 149*, 111977. doi:10.1016/j.fct.2021.111977

Dellafiora, L., Galaverna, G., Cruciani, G., Dall'Asta, C., & Bruni, R. (2018). On the Mechanism of Action of Anti-Inflammatory Activity of Hypericin: An In Silico Study Pointing to the Relevance of Janus Kinases Inhibition. *Molecules, 23*(12). doi:10.3390/molecules23123058

Dellafiora, L., Galaverna, G., & Dall'Asta, C. (2018). Mechanisms of Fumonisin B1 Toxicity: A Computational Perspective beyond the Ceramide Synthases Inhibition. *Chem Res Toxicol, 31*(11), 1203-1212. doi:10.1021/acs.chemrestox.8b00188

Dellafiora, L., Galaverna, G., Dall'Asta, C., & Cozzini, P. (2015). Hazard identification of cis/trans-zearalenone through the looking-glass. *Food Chem Toxicol, 86*, 65-71. doi:10.1016/j.fct.2015.09.009

Dellafiora, L., Gonaus, C., Streit, B., Galaverna, G., Moll, W. D., Vogtentanz, G., . . . Prasad, S. (2020). An In Silico Target Fishing Approach to Identify Novel Ochratoxin A Hydrolyzing Enzyme. *Toxins (Basel), 12*(4). doi:10.3390/toxins12040258

Dellafiora, L., Oswald, I. P., Dorne, J. L., Galaverna, G., Battilani, P., & Dall'Asta, C. (2020). An in silico structural approach to characterize human and rainbow trout estrogenicity of mycotoxins: Proof of concept study using zearalenone and alternariol. *Food Chem, 312*, 126088. doi:10.1016/j.foodchem.2019.126088

Gao, Z., Luo, K., Zhu, Q., Peng, J., Liu, C., Wang, X., . . . Zhang, H. (2023). The natural occurrence, toxicity mechanisms and management strategies of Fumonisin B1 : A review. *Environ Pollut, 320*, 121065. doi:10.1016/j.envpol.2023.121065

Guerre, P., Lassallette, E., Beaujardin-Daurian, U., & Travel, A. (2024). Fumonisins alone or mixed with other fusariotoxins increase the C22-24:C16 sphingolipid ratios in chicken livers, while deoxynivalenol and zearalenone have no effect. *Chem Biol Interact, 395*, 111005. doi:10.1016/j.cbi.2024.111005

Harrer, H., Humpf, H. U., & Voss, K. A. (2015). In vivo formation of N-acyl-fumonisin B1. *Mycotoxin Res, 31*(1), 33-40. doi:https://doi.org/10.1007/s12550-014-0211-5.

Harrer, H., Laviad, E. L., Humpf, H. U., & Futerman, A. H. (2013). Identification of N-acyl-fumonisin B1 as new cytotoxic metabolites of fumonisin mycotoxins. *Mol Nutr Food Res, 57*(3), 516-522. doi:10.1002/mnfr.201200465

Hong, X. Z., Han, Z. G., Yang, J. K., & Liu, Y. H. (2023). The Motion Paradigm of Pre-Dock Zearalenone Hydrolase Predictions with Molecular Dynamics and the Docking Phase with Umbrella Sampling. *Molecules, 28*(11). doi:10.3390/molecules28114545

Hu, K., & Cao, Y. (2024). Ping, pong, and freeze: Structural insights into the inhibition of ceramide synthase by Fumonisin B1. *Structure, 32*(9), 1296-1298. doi:10.1016/j.str.2024.08.010

Huang, J., & MacKerell, A. D., Jr. (2013). CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem, 34*(25), 2135-2145. doi:10.1002/jcc.23354

Humpf, H. U., Schmelz, E. M., Meredith, F. I., Vesper, H., Vales, T. R., Wang, E., . . . Merrill, A. H., Jr. (1998). Acylation of naturally occurring and synthetic 1-deoxysphinganines by ceramide synthase. Formation of N-palmitoyl-aminopentol produces a toxic metabolite of hydrolyzed fumonisin, AP1, and a new category of ceramide synthase inhibitor. *J Biol Chem, 273*(30), 19060-19064. doi:10.1074/jbc.273.30.19060

International Agency for Research on Cancer, W. H. O., & Cancer, I. A. f. R. o. (2002). Fumonisin B1. *IARC monographs on the evaluation of carcinogenic risks to humans, some traditional herbal medicines, some mycotoxins, naphthalene and styrene, 82*, 301-366.

Jo, S., Kim, T., Iyer, V. G., & Im, W. (2008). CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem, 29*(11), 1859-1865. doi:10.1002/jcc.20945

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., . . . Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583-589. doi:10.1038/s41586-021-03819-2

Khan, R., Anwar, F., & Ghazali, F. M. (2024). A comprehensive review of mycotoxins: Toxicology, detection, and effective mitigation approaches. *Heliyon, 10*(8), e28361. doi:10.1016/j.heliyon.2024.e28361

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., . . . Bolton, E. E. (2023). PubChem 2023 update. *Nucleic Acids Res, 51*(D1), D1373-D1380. doi:10.1093/nar/gkac956

Lambert, J. C., & Lipscomb, J. C. (2007). Mode of action as a determining factor in additivity models for chemical mixture risk assessment. *Regul Toxicol Pharmacol, 49*(3), 183-194. doi:10.1016/j.yrtph.2007.07.002

Levy, M., & Futerman, A. H. (2010). Mammalian ceramide synthases. *IUBMB Life, 62*(5), 347-356. doi:10.1002/iub.319

Li, T., Li, J., Wang, J., Xue, K. S., Su, X., Qu, H., . . . Jiang, Y. (2024). The occurrence and management of fumonisin contamination across the food production and supply chains. *J Adv Res, 60*, 13-26. doi:10.1016/j.jare.2023.08.001

Louisse, J., Dellafiora, L., van den Heuvel, J., Rijkers, D., Leenders, L., Dorne, J. C. M., . . . Koenderink, J. B. (2023). Perfluoroalkyl substances (PFASs) are substrates of the renal human organic anion transporter 4 (OAT4). *Arch Toxicol, 97*(3), 685-696. doi:10.1007/s00204-022-03428-6

Louisse, J., Dorne, J., & Dellafiora, L. (2022). Investigating the interaction between organic anion transporter 1 and ochratoxin A: An in silico structural study to depict early molecular events of substrate recruitment and the impact of single point mutations. *Toxicol Lett, 355*, 19-30. doi:10.1016/j.toxlet.2021.11.001

Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., . . . Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res, 50*(W1), W276-W279. doi:10.1093/nar/gkac240

Merrill, A. H., van Echten, G., Wang, E., & Sandhoff, K. (1993). Fumonisin B1 inhibits sphingosine (sphinganine) N-acyltransferase and de novo sphingolipid biosynthesis in cultured neurons in situ. *Journal of Biological Chemistry, 268*(36), 27299-27306. doi:10.1016/s0021-9258(19)74249-5

Mesicek, J., Lee, H., Feldman, T., Jiang, X., Skobeleva, A., Berdyshev, E. V., . . . Kolesnick, R. (2010). Ceramide synthases 2, 5, and 6 confer distinct roles in radiation-induced apoptosis in HeLa cells. *Cell Signal, 22*(9), 1300-1307. doi:10.1016/j.cellsig.2010.04.006

Mirdita, M., Schutze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nat Methods, 19*(6), 679-682. doi:10.1038/s41592-022-01488-1

Moussa, N., Hassan, A., & Gharaghani, S. (2021). Pharmacophore model, docking, QSAR, and molecular dynamics simulation studies of substituted cyclic imides and herbal medicines as COX-2 inhibitors. *Heliyon, 7*(4), e06605. doi:https://doi.org/10.1016/j.heliyon.2021.e06605.

O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. *J Cheminform, 3*, 33. doi:10.1186/1758-2946-3-33

Pascoa, T. C., Pike, A. C. W., Tautermann, C. S., Chi, G., Traub, M., Quigley, A., . . . Sauer, D. B. (2025). Structural basis of the mechanism and inhibition of a human ceramide synthase. *Nat Struct Mol Biol, 32*(3), 431-440. doi:10.1038/s41594-024-01414-3

Pedroni, L., Doherty, D. Z., Dall'Asta, C., Galaverna, G., Bell, S. G., & Dellafiora, L. (2024). Computational methods meet in vitro techniques: A case study on fusaric acid and its possible detoxification through cytochrome P450 enzymes. *Ecotoxicol Environ Saf, 273*, 116167. doi:10.1016/j.ecoenv.2024.116167

Schafer, J. H., Clausmeyer, L., Korner, C., Esch, B. M., Wolf, V. N., Sapia, J., . . . Frohlich, F. (2025). Structure of the yeast ceramide synthase. *Nat Struct Mol Biol, 32*(3), 441-449. doi:10.1038/s41594-024-01415-2

Schertz, H., Danicke, S., Frahm, J., Schatzmayr, D., Dohnal, I., Bichl, G., . . . Kluess, J. (2018). Biomarker Evaluation and Toxic Effects of an Acute Oral and Systemic Fumonisin Exposure of Pigs with a Special Focus on Dietary Fumonisin Esterase Supplementation. *Toxins (Basel), 10*(7). doi:10.3390/toxins10070296

Seefelder, W., Humpf, H. U., Schwerdt, G., Freudinger, R., & Gekle, M. (2003). Induction of apoptosis in cultured human proximal tubule cells by fumonisins and fumonisin metabolites. *Toxicol Appl Pharmacol, 192*(2), 146-153. doi:10.1016/s0041-008x(03)00262-x

Seiferlein, M., Humpf, H. U., Voss, K. A., Sullards, M. C., Allegood, J. C., Wang, E., & Merrill, A. H., Jr. (2007). Hydrolyzed fumonisins HFB1 and HFB2 are acylated in vitro and in vivo by ceramide synthase to form cytotoxic N-acyl-metabolites. *Mol Nutr Food Res, 51*(9), 1120-1130. doi:10.1002/mnfr.200700118

Subramanian, N., Condic-Jurkic, K., Mark, A. E., & O'Mara, M. L. (2015). Identification of Possible Binding Sites for Morphine and Nicardipine on the Multidrug Transporter P-Glycoprotein Using Umbrella Sampling Techniques. *J Chem Inf Model, 55*(6), 1202-1217. doi:10.1021/ci5007382

Sultana, S., Kitajima, M., Kobayashi, H., Nakagawa, H., Shimizu, M., Kageyama, K., & Suga, H. (2019). A Natural Variation of Fumonisin Gene Cluster Associated with Fumonisin Production Difference in Fusarium fujikuroi. *Toxins (Basel), 11*(4). doi:10.3390/toxins11040200

Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Zidek, A., . . . Hassabis, D. (2021). Highly accurate protein structure prediction for the human proteome. *Nature, 596*(7873), 590-596. doi:10.1038/s41586-021-03828-1

Uhlen, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., . . . Ponten, F. (2015). Proteomics. Tissue-based map of the human proteome. *Science, 347*(6220), 1260419. doi:10.1126/science.1260419

Verstraete, F. (2013). Risk management of undesirable substances in feed following updated risk assessments. *Toxicol Appl Pharmacol, 270*(3), 230-247. doi:10.1016/j.taap.2010.09.015

Wang, E., Norred, W. P., Bacon, C. W., Riley, R. T., & Merrill, A. H., Jr. (1991). Inhibition of sphingolipid biosynthesis by fumonisins. Implications for diseases associated with Fusarium moniliforme. *J Biol Chem, 266*(22), 14486-14490.

ww, P. D. B. c. (2019). Protein Data Bank: the single global archive for 3D macromolecular structure data. *Nucleic Acids Res, 47*(D1), D520-D528. doi:10.1093/nar/gky949

Xia, M., Zheng, J., Chen, S., Tang, Y., Wang, S., Ji, Z., . . . Liu, Y. (2024). Bisphenol a alternatives suppress human and rat aromatase activity: QSAR structure-activity relationship and in silico docking analysis. *Food Chem Toxicol, 183*, 114257. doi:10.1016/j.fct.2023.114257

Xie, L., Huang, B., Zhao, X., & Zhu, N. (2024). Exploring the mechanisms underlying effects of bisphenol a on cardiovascular disease by network toxicology and molecular docking. *Heliyon, 10*(10), e31473. doi:10.1016/j.heliyon.2024.e31473

Xie, T., Fang, Q., Zhang, Z., Wang, Y., Dong, F., & Gong, X. (2023). Structure and mechanism of a eukaryotic ceramide synthase complex. *EMBO J, 42*(24), e114889. doi:10.15252/embj.2023114889

Yang, Z., Zeng, X., Zhao, Y., & Chen, R. (2023). AlphaFold2 and its applications in the fields of biology and medicine. *Signal Transduct Target Ther, 8*(1), 115. doi:10.1038/s41392-023-01381-z

Yu, S., Jia, B., Liu, N., Yu, D., Zhang, S., & Wu, A. (2021). Fumonisin B1 triggers carcinogenesis via HDAC/PI3K/Akt signalling pathway in human esophageal epithelial cells. *Sci Total Environ, 787*, 147405. doi:10.1016/j.scitotenv.2021.147405

Zelnik, I. D., Mestre, B., Weinstein, J. J., Dingjan, T., Izrailov, S., Ben-Dor, S., . . . Futerman, A. H. (2023). Computational design and molecular dynamics simulations suggest the mode of substrate binding in ceramide synthases. *Nat Commun, 14*(1), 2330. doi:10.1038/s41467-023-38047-x

Zhang, M., Li, Z., Liu, Y., Ding, X., Wang, Y., & Fan, S. (2023). The ceramide synthase (CERS/LASS) family: Functions involved in cancer progression. *Cell Oncol (Dordr), 46*(4), 825-845. doi:10.1007/s13402-023-00798-6

Zhang, Z., Fang, Q., Xie, T., & Gong, X. (2024). Mechanism of ceramide synthase inhibition by fumonisin B(1). *Structure, 32*(9), 1419-1428 e1414. doi:10.1016/j.str.2024.06.002

Free fatty acid receptors beyond fatty acids: A computational journey to explore peptides as possible binders of GPR120

This chapter is based on the paper published on:

Lorenzo Pedroni, Florinda Perugino, Fabio Magnaghi, Chiara Dall'Asta, Gianni Galaverna, Luca Dellafiora. (2024). Free fatty acid receptors beyond fatty acids: A computational journey to explore peptides as possible binders of GPR120. Current Research in Food Science, 8, 100710, https://doi.org/10.1016/j.crfs.2024.100710.

Author contribution: Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing.

Abstract

Free fatty acids receptors, with members among G protein-coupled receptors (GPCRs), are crucial for biological signaling, including the perception of the so called “fatty taste”. In recent years, GPR120, a protein belonging to the GPCR family, drew attention as an interesting pharmacological target to cope with obesity, satiety and diabetes. Apart from long chain fatty acids, which are GPR120 natural agonists, other synthetic molecules were identified as agonists expanding the chemical space of GPR120’s ligands. In this scenario, we unveiled peptides as possible GPR120 binders toward a better understanding of this multifaced and relevant target. This study analyzed a virtual library collecting 531 441 low-polar hexapeptides, providing mechanistic insights into the GPR120 activation and further extending the possible chemical space of GPR120 agonists. The computational pipeline started with a narrow filtering of hexapeptides based on their chemical similarity with known GPR120 agonists. The best hits were tested through docking studies, molecular dynamics and umbrella sampling simulations, which pointed to G[I,L]FGGG as a promising GPR120 agonist sequence. The presence of both peptides in food-related proteins was thoroughly assessed, revealing they may occur in mushrooms, food-grade bacteria and rice. Simulations on the counterparts with D-amino acids were also performed. Umbrella sampling simulations described that GdIFGGG may have a better interaction compared to its all-L counterpart (-13 kCal/mol ΔG and –6 kCal/mol ΔG, respectively). Overall, we obtained a predictive model to better understand the underpinning mechanism of GPR120-hexapeptides interaction, hierarchizing novel potential agonist peptides for further analysis and describing promising food sources worth of further dedicated investigations.

1. Introduction

Heterotrimeric G protein-coupled receptors (GPCRs), also known as 7-transmembrane domain receptors, are the largest family of cell-surface and signal transducing proteins (Civciristov et al., 2018) (up to 800 members identified so far) (Guo et al., 2022). They are integral transmembrane proteins and mediate most cellular responses to hormones and neurotransmitters. GPCRs also play a crucial role in taste and smell perceptions as they take a part in transducing the signal for aroma compounds as well as for sweet, bitter, umami, and kokumi taste (Ahmad & Dalziel, 2020). They have been also recently described as involved in the perception of the still debated fatty taste upon binding of free fatty acids (FFAs) (Galindo et al., 2012). Several FFAs receptors (FFARs) were identified among GPCRs. Specifically, long-chain (> 12 carbon atoms) saturated and unsaturated fatty acids bind and activate GPR40 and GPR120, while short-chain fatty acids (< 6 carbon atoms; acetic acid, propionic acid and butyric acid, mainly) activate GPR41 and GPR43 (Kimura et al., 2020). A lot of studies focused on GPR120, also named Free Fatty Acid Receptor 4 (FFAR4; UniProt ID Q5NUL3), being an interesting drug target for the management and treatment of a series of disorders including Type 2 diabetes mellitus, obesity and those related to the sense of satiety (Burns & Moniri, 2010). GPR120 is relevantly expressed in enteroendocrine cells where it is involved in the secretion of glucagon-like-peptide-1 (GLP1) and cholecystokinin, known to play important roles in satiety and general food intake (Hirasawa et al., 2005). Further evidence pointing to its involvement in antidiabetic action came in 2010 with the discovery of ω-3 fatty acids as suitable ligands (Hirasawa et al., 2005), which may associate with anti-inflammatory effects and consequent insulin sensitizing action (Oh et al., 2010). From a molecular standpoint, there are two human GPR120 isoforms, a long one (377 aa) and a short one (361 aa), differing from each other for the insertion of 16 amino acids between positions 231 and 247 of the third intracellular loop. However, there is no evidence pointing to functional differences between the two isoforms which are meant to be functionally comparable (Pal et al., 2021). As said before, FFAs are among the natural agonists of GPR120. However, the growing interest in this pharmacological target is driving the identification of new agonists for further developments either from a medicinal chemistry or food science standpoints – the former to derive pharmacologically active compounds, the latter to study the chemistry of fatty taste perception and/or to develop nutraceuticals and functional foods. Several GPR120 agonists were identified over the years, such as TUG-891 and GW9508 (Hudson et al., 2013). Specifically, they enhance gastric inhibitory peptide (GIP) and GLP-1 secretion, suggesting that GPR120 may represent a valid

antidiabetic drug target (Carullo et al., 2021; McKillop et al., 2023), as well as a target heading to anti-inflammatory processes (Anbazhagan et al., 2016). Nowadays, besides non-natural pharmaceuticals, peptides are rising in popularity as therapeutics to target proteins and pathways involved in the onset and progression of certain diseases (Davenport et al., 2020) as well as to design food supplements and nutraceuticals (de Castro & Sato, 2015). Indeed, bioactive peptides derived from food proteins upon processing (e.g. via hydrolysis and fermentation) were shown to have potential applications as ingredients in functional foods owing to their potential health-promoting characteristics (Udenigwe & Aluko, 2012) . Similarly, food rich in bioactive peptides may bring benefit when properly consumed (Udenigwe & Aluko, 2012). This makes the identification of bioactive sequences and their organism sources relevant from a food science standpoint. Germane to GPCRs, several agonist peptides have been already described (Davenport et al., 2020; Krumm & Grisshammer, 2015). Nonetheless, the discovery of new agonist peptides is relevant either to expand the chemical space of GPCRs ligands (fundamental for drug development) or to find valuable food-grade protein sources to develop nutraceuticals and functional foods. In this framework, the present study aimed at investigating the mechanics of GPR120-ligand interaction to identify new potential GPR120 peptide ligands. To do so, a virtual library of 531 441 hexapeptides has been set up and screened to mine sequences showing chemical analogies to known GPR120 ligands (namely, oleic acid and TUG-891), based on the assumption that similar compounds may activate the same target (McKinney et al., 2000). Afterward, a selection of the best hits underwent a 3D molecular modelling study integrating docking, conventional molecular dynamics (CMD) and umbrella sampling (US) simulations to thoroughly estimate the receptor activation, in agreement with previous studies (Lammi et al., 2021; Vidal-Limon et al., 2022).

Overall, this study developed a predictive model to better understand the underpinning mechanism of GPR120-hexapeptides interaction, hierarchizing novel potential agonist peptides and their food-related sources for further analysis. The effect of D-amino acids in the activity of those sequences proposed to be active has been also evaluated.

2.Materials and Methods

2.1 GPR120 model construction

The 3D model of GPR120 was derived from the crystallographic structure available on Protein Data Bank (https://www.rcsb.org) (Berman et al., 2000) with the PDB code 8ID6 (Chain A) (Mao et al., 2023). The structure was chosen based on its good resolution (2.80 Å), although it

was uncomplete in some parts (residues 1-22, 33-34, 70-74 and 145-150). The structural continuity was obtained by homology modelling using the SwissModel tool (https://swissmodel.expasy.org/) (Waterhouse et al., 2018). SwissModel is a structural bioinformatic web server for generating 3D models of proteins using a comparative approach. The 3D structure 8ID6 was used as template while the FASTA sequence of GPR120 (Uniprot code: Q5NUL3) was used as input. After that, the predicted structure was superimposed to the AlphaFold (https://alphafold.ebi.ac.uk/) (Jumper et al., 2021) structure and residues 15-30 were substituted with those of the AlphaFold model as their reconstructions seemed more plausible, in agreement with previous studies (Dellafiora et al., 2020).

2.2 Retrieval of ligands 3D information

The structures of TUG-891 and FFAs were retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov/) (Kim et al., 2023) in the Structured Data File (.sdf): TUG-891 (CID: 57522038), oleic acid (CID: 445639), isobutyric acid (CID: 6590) and propionic acid (CID: 1032) (Figure 1). Then, they were converted to the Tripos Mol2 format (.mol2) using Open Babel (O'Boyle et al., 2011) to enable further analysis.

2.3 Construction of peptides libraries

The 3D hexapeptides structures analyzed in this study were built using an *ad hoc* python script interfaced with PyMol (opensource v. 2.3.0; https://pymol.org), in agreement with a previous study (Dellafiora et al., 2022). All the possible combinations of 9 nonpolar amino acids (i.e. glycine, alanine, leucine, isoleucine, valine, methionine, proline, phenylalanine and tryptophan) were generated, for a total of 531 441 combinations. The protonation state was set at pH 7 with C- and N- terminal as deprotonated and protonated, respectively. The entire set of generated peptides was collected in the Tripos Mol2 format. The list of peptides including D-amino acid was generated by manually editing the respective amino acid position using the *invert* PyMol's function.

2.4 Ligand-based virtual screening of the library

The hexapeptides library underwent a ligand-based virtual screening to mine peptides according to their similarity to oleic acid and TUG-891, taken as reference for GPR120 natural and synthetic ligands, respectively. The ligand-based virtual screening was performed using the LiSiCA (Ligand Similarity using Clique Algorithm) algorithm (Legnik et al., 2015) as

previously succeeded to rank entries and efficiently estimate their bioactivity (Del Favero et al., 2022; Dellafiora et al., 2022). This algorithm provides a fast ligand-based virtual screening platform to quantify chemical similarities between a reference template (oleic acid and TUG-891 in this case) and a database of target compounds. LiSiCA expresses similarities using the Tanimoto coefficient's metric, a gold standard to quantify chemical analogies (from 0 to 1, with 1 meaning identical molecules). LiSiCA's default parameters were used. This fast virtual screening aimed to identify a small selection of hexapeptides within the whole library to carry forth to the molecular modelling steps.

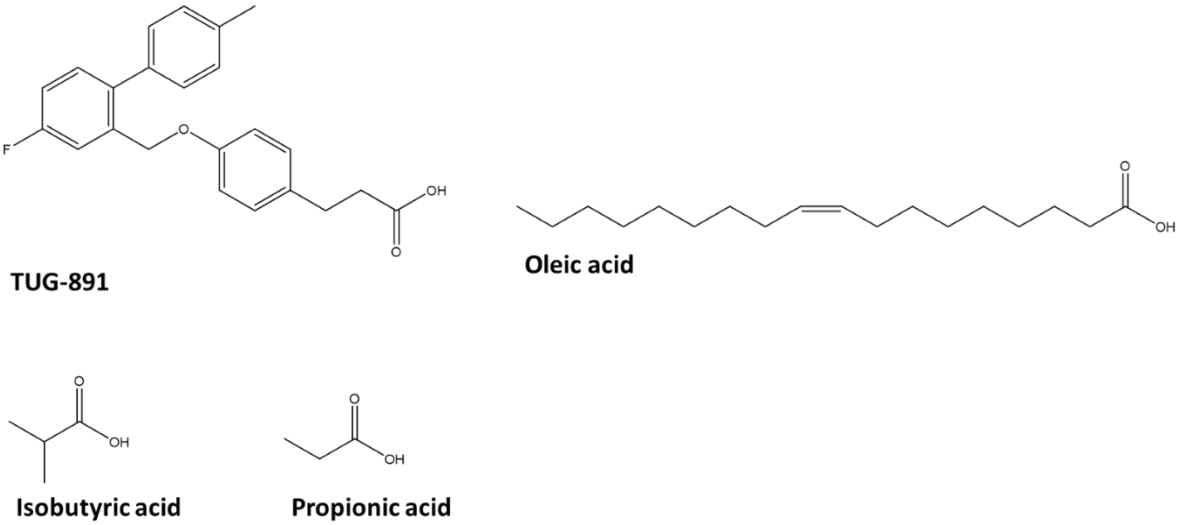


Figure 1. Chemical structures of positive (TUG-891 CID: 57522038 and oleic acid CID: 445639) and negative (isobutyric acid CID: 6590 and propionic acid CID: 1032) controls.

2.5 Docking simulations

Molecular docking simulations aimed to provide a plausible binding architecture for TUG-891, FFAs and hexapeptides obtained from the virtual screening and were performed using the GOLD software (v. 2021). The binding site was defined within a 10-Å radius sphere from the centroid of the pocket. The water #501 (from PDB structure 8ID4) was included and set forming hydrogen bonds (default settings) as it was described involved in polar interactions with ligands (Mao et al., 2023). Additionally, a spatial constraint has been used to better reproduce the binding architecture as per crystallographic data setting the similarity shape overlap option (weight 100) with respect to the crystallographic pose of oleic acid in 8ID6. According to

previous study (Pedroni et al., 2023), a semi-flexible docking approach was applied keeping ligands fully flexible and the protein semi-flexible allowing polar hydrogens free to rotate. 10 poses for each ligand were generated and the internal score function PLPScore was used to estimate the fitting of each ligand within GPR120 binding site (the higher the score, the more likely the precited binding architecture; https://www.ccdc.cam.ac.uk) as it proved to reliably reproduce and predict architectures of binding (Del Favero et al., 2022).

2.6 Conventional Molecular Dynamics simulations

CMD simulations were performed using GROMACS (version 2021.4) (Abraham et al., 2015) to study ligands-GPR120 stability over time. Before running CMD, the protein structure was embedded into a dipalmitoylphosphatidylcholine (DPPC) membrane using the CHARMM-GUI (Jo et al., 2008; Lee et al., 2016) membrane builder tool (https://www.charmm-gui.org). More in detail, CHARMM-GUI is a web-based user interface which allows to build realistic biological membrane systems. The whole system was parametrized with CHARMM36 all-atom force field (Lee et al., 2016). TUG-891, oleic acid, isobutyric acid and propionic acid parametrization was performed on the SwissParam webserver (https://www.swissparam.ch) (Zoete et al., 2011). Each complex was solvated with SPCE waters in a cubic periodic boundary condition and the system was neutralized by adding counter ions (Na^+ or Cl^-). After that, each system was energetically minimized to avoid steric clashes and to correct improper geometries using the steepest descent algorithm with a maximum of 5 000 steps. Next, all the systems underwent isothermal (300 K, coupling time 2 ps) and isobaric (1 bar, coupling time 2 ps) 100 ps simulations before undergoing a 40 ns CMD simulation each.

2.7 Umbrella sampling simulation

The protein-peptide complex with the highest PLPScore was chosen for further investigations through US. Simulations were performed using GROMACS (version 2021.4) to estimate protein-peptide binding free energy and the whole system was parametrized with CHARMM36 all-atom force field (Huang & MacKerell, 2013). The input structure was placed in a rectangular box with dimensions sufficient to provide space for the pulling simulations to take place along the Y-axis. The box was then solvated with SPCE water and neutralized by adding counter ions (Na^+ and Cl^-). Subsequently the system was energetically minimized with a maximum of 50 000 steps and underwent isothermal (310 K, coupling time 0.1 ps) and isobaric (1 bar, coupling time 2 ps). The peptide was then pulled from the protein binding site at 0.01 nm/ps pull rate

over the course of 500 ps on the Y-axis using a spring force constant of 1500 kJ•mol^{-1} nm^{-2}. The frames obtained from the pulling were used as initial coordinates for binding free energy calculations through the US simulations. An asymmetric distribution of sampling windows was used such that spacing along the reaction coordinate was 0.2 nm. Each US window underwent an NPT equilibration before running a 10 ns simulation with a spring force constant of 1500 kJ mol^{-1} nm^{-2}.

2.8 Identification of potential proteins source of food origin and possible release

An iterative sequence alignment procedure was set up based on an *ad hoc* bash script. This was meant to search each of the most promising hexapeptides (namely, GIFGGG and GLFGGG) within the whole SwissProt database (570 420 protein sequences; last database access 6th December 2023). Briefly, SwissProt is the manually curated part of UniProt (https://www.uniprot.org) (Bateman et al., 2023) collecting all the protein sequences referenced as "reviewed" (i.e. derived from scientific literature or, if computationally derived, curated and evaluated by expert analysis).

3. Results and Discussion

3.1. Model validation

Previous works demonstrated that the procedural workflow used here may reliably estimate whether a given molecule may bind and activate a designated receptor (Dellafiora et al., 2022; Pedroni et al., 2023; Wu et al., 2022). Specifically, docking simulations, which aimed at providing plausible binding architectures scored based on their fitting into the pocket, were integrated to CMD and US simulations to study the complex stability and energy over time. Of note, monitoring the geometrical stability of the receptor-ligand complex over time may efficiently estimate the activity of ligands. Specifically, a ligand is meant to be an agonist when its interaction with the receptor is stable over time, preserving at the same time the overall structure of the given activated receptor. Conversely, unstable complexes, e.g with ligands that can not persist at the designated binding site and/or when the overall protein structure loses the native (active) structure, point to the incapability of ligands to appreciably activate the receptor. The capability of *in silico* methods to achieve such analysis for GPCRs and peptides has been already demonstrated (Dellafiora et al., 2022; Pedroni et al., 2023). However, a fit-for-purpose assessment of procedural performances has been done to define the reference scenario for active

ligands and for molecules unable to appreciably bind and activate the receptor. In this respect, oleic acid and TUG-891 were taken as positive controls being well-known GPR120 agonists, while propionic and isobutyric acid were taken as negative controls, being unable to bind and activate GPR120 (Galindo et al., 2012). Once defined the set of reference compounds, they were docked at the receptor binding site and the calculated poses of positive controls (oleic acid and TUG-891) were compared to the respective crystallographic poses (as per PDB structure with code 8ID6 and 8ID8, respectively). As shown in Figure 2, all the molecules considered engaged the receptor with a similar binding pose, arranging the carboxylic acid moiety as per the crystallographic pose of oleic acid (PDB code 8ID6). Particularly, for positive controls, calculated and crystallographic poses were comparable proving the model effectiveness to reliably reproduce the binding pose of ligands. Of note, as shown in Table 1, the negative controls recorded scores lower than the known GPR120 ligands TUG-891 and oleic acid. Although this might have been due to a diverse number of atoms (short-chain FFAs have inherently a lower number of atoms compared to the mid-long counterpart), the score assignment could suggest which molecules were able to best fit the receptor pocket (the higher the score, the better the fitting, in agreement with the manufacturer declarations; https://www.ccdc.cam.ac.uk). Then, the four complexes underwent CMD simulations to monitor their stability over time. As shown in Figure 2 and S1 (see Supplementary Materials), the negative controls (i.e. isobutyric acid and propionic acid) could not stably interact at the ligand binding site showing a full detachment from the designated interaction site within 25 ns. Conversely, the positive controls (i.e. oleic acid and TUG-891), remained stable at the ligand binding site, as shown by the root mean squared deviation (RMSD) analysis of the carboxylate moiety and the respective trajectories. They also kept the overall receptor structure stable. Of note, RMSD fluctuation was monitored only for the carboxylate moiety with respect to the whole protein to enable a comparative analysis among the ligands considered (including peptides), as they differ in shape and number of atoms (see Figure 2).

Taken together, these outcomes have shown that the integrated use of docking analysis and CMD can reliably distinguish GPR120 ligands from non-ligands, with geometrical parameters like RMSD and trajectories analysis as probative features to enable ligands discrimination and function assignment.

Table 1. Virtual screening and docking scores

Ligands		**LiSiCA Score**[a]	**PLPScore**[b]
FFAs and synthetic agonist	Propionic acid	n.p.	95.4
	Isobutyric acid	n.p.	98.0
	Oleic acid	n.p.	143.8
	TUG-891	n.p.	162.0
Total L peptides	**GLFGGG**	**0.281**	**145.0**
	GIFGGG	**0.281**	**151.5**
	GGPGWG	0.280	135.4
	GIGGGA	0.282	140.2
	GIGWGA	0.282	128.9
	GIPGWG	0.284	119.3
	GLGGGA	0.282	131.9
	GLGWGA	0.282	132.5
	GLPGWG	0.284	115.7
Peptides with D-amino acids	GdIdFGGG	n.p.	142.7
	GdIFGGG	n.p.	**156.9**
	GdLdFGGG	n.p.	121.4
	GdLFGGG	n.p.	147.2
	GLdFGGG	n.p.	146.4
	GIdFGGG	n.p.	146.2

Note: [a] average LiSiCA score using alternatively TUG-891 and Oleic acid as reference compounds; [b] The higher the score, the better the pocket fitting, as per manufacturer declaration (https://www.ccdc.cam.ac.uk); peptides considered for molecular dynamics are in bold; n.p. stands for not performed.

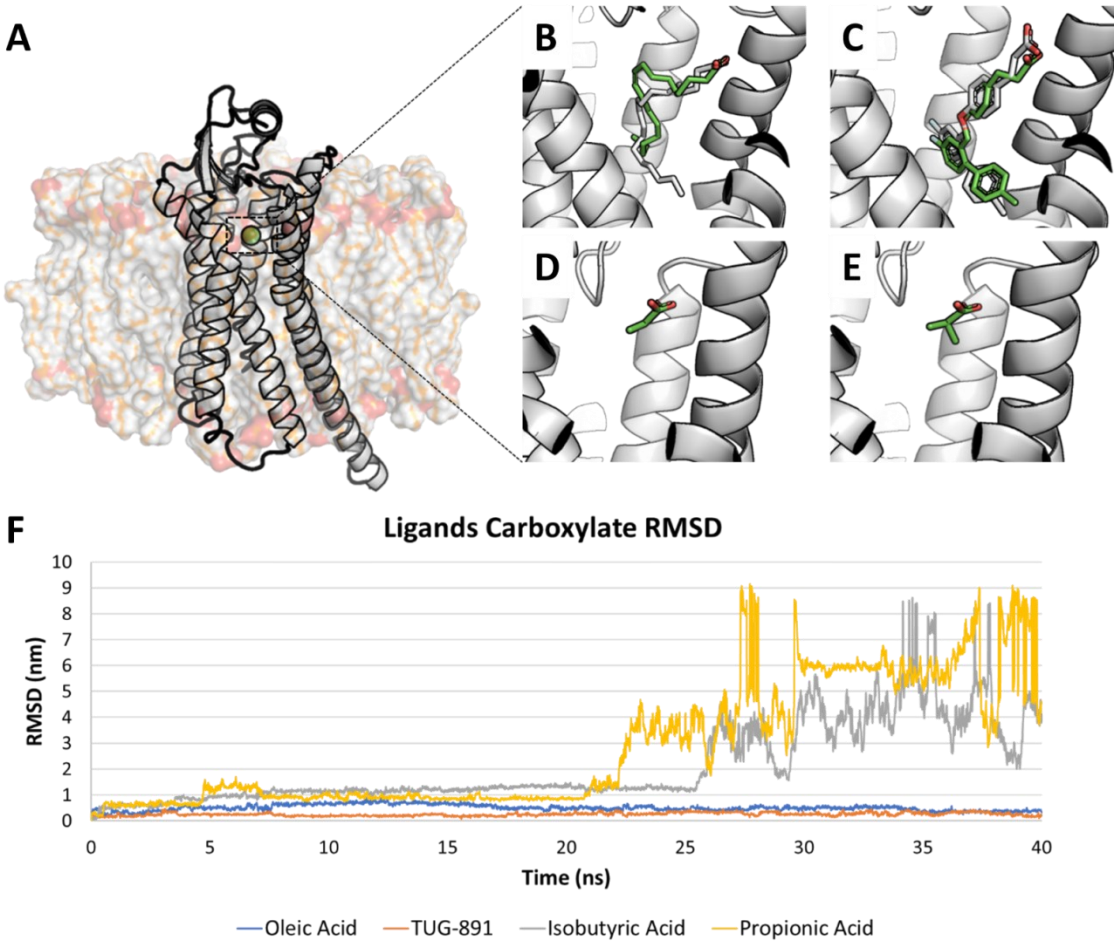


Figure 2. Molecular docking and dynamics results for GPR120 model validation. **A**. GPR120 3D model represented as white cartoon within membrane represented as semi-transparent surface. The green sphere represents the centroid set for the docking simulation. **B**. Oleic acid best docking pose (green sticks) superimposed to oleic acid retrieved from PDB ID 8ID6 (white sticks). **C**. TUG-891 best docking pose (green sticks) superimposed to TUG-891 retrieved from the PDB ID 8ID8 (white sticks). **D**. Propionic acid best docking pose (green sticks) **E**. Isobutyric acid best docking pose (green sticks). **F**. Ligands carboxylate RMSD representing the distance of the carboxylate moiety from the protein. Of note, both isobutyric and propionic acid detached from the GPR120 binding pocket within 25 ns.

3.2. Analysis of peptides library

3.2.1 Virtual Screening

The 3D hexapeptides library analyzed in this study was built using an *ad hoc* python script interfaced with PyMol generating all the possible combinations of the 9 non-polar amino acids (see section 2.3). This was done to obtain peptides with a partial chemical similarity in terms

of steric properties with the native GPR120 binders (i.e. long chain fatty acids). The hexapeptides library screening was achieved using the LiSiCA algorithm (Legnik et al., 2015) setting default parameters (see 2.4). Two different screening procedures were run setting alternatively oleic acid (i.e. the natural substrate retrieved from the crystallographic structure with PDB ID 8ID6) and TUG-891 (i.e. the synthetic agonist retrieved from the crystallographic structure with PDB ID 8ID8) as reference compounds (Figure 1). As expected, the significant chemical difference between the hexapeptides and the reference compounds led to relatively low scores. This was not considered a limiting factor for the sake of this study since this step was crucial to obtain a feasible number of hexapeptides to enable the subsequent molecular modelling analysis. The average score from the two screening procedures was computed for each hexapeptide obtaining their overall ranking via an *ad hoc* python script. Arbitrarily, only hexapeptides scoring equal or higher than 0.28 units were considered to focus the analysis on those most similar to the reference compounds. This resulted in a list of 9 hexapeptides: GLFGGG, GIFGGG, GGPGWG, GIGGGA, GIGWGA, GIPGWG, GLGGGA, GLGWGA and GLPGWG (Table 1).

3.2.1 Analysis of best hits

Based on the above, and on the assumption that molecules sharing analogies in terms of molecular structure may target the same protein (McKinney et al., 2000), the following peptides underwent docking analysis to estimate their actual capability to interact with GPR120: GLFGGG, GIFGGG, GGPGWG, GIGGGA, GIGWGA, GIPGWG, GLGGGA, GLGWGA and GLPGWG. As shown in Table 1, GLFGGG and GIFGGG recorded the highest docking scores (PLPScore of 145.0 and 151.5 units, respectively). This suggested that they may be those most able to fit the receptor pocket among those considered in docking studies (the higher the score, the better the interaction). Afterward, they underwent CMD to check the peptide-receptor stability over time to compare the outcome with the reference scenario obtained for positive and negative controls (see section 3.1), predicting their possible activity thereby. As shown in Figure 3 and S2 (see Supplementary Materials), the RMSD and trajectory analysis described that they both kept a steady state interaction at the designated binding site of GPR120 keeping the overall protein structure stable over time. This pointed to the possible capability of both peptides to effectively interact and activate GPR120, as proved in previous studies for many other peptide-activated receptors. Specifically, peptides have been already described as able to activate GPCRs involved in taste perception, as found for e.g. umami, bitter and kokumi

(Dellafiora et al., 2022; Wang et al., 2022; Yu et al., 2022). To further verify this hypothesis, a US simulation was performed to estimate the free energy of binding for a thorough assessment of complex stability. Since both the hexapeptides showed comparable results in terms of CMD simulations, the US simulation was performed only on the GIFGGG peptide as it obtained the highest docking score (i.e. 151.5 units). Once GIFGGG was pulled out for 5 nm, 24 windows with a 0.2 nm interval were chosen to perform the US simulation (see section 2.6). This enabled the collection of histograms and the umbrella potential graph reported in Figure 4. The peak distribution in the histogram reflected the good sampling obtained via the parameters set (section 2.6). Indeed, the evenly spaced windows were sufficient to cover all the regions GIFGGG passed by, exhibiting a satisfactory overlap. The analysis of the umbrella potential graph showed a ΔG corresponding to roughly -6 kcal/mol, pointing to its energetically favorable interaction with GPR120, in line with other studies on protein-peptide complexes (Cool & Lindert, 2022). Of note, the capability of hexapeptides to activate a GPCR described here is in line with previous finding describing identical length peptides able to activate a variety of GPCRs, including those associated with taste perception (Calderón et al., 2023; Thibeault et al., 2020; Wang et al., 2022). This evidence eventually supported the appreciable binding of GIFGGG with GPR120.

This outcome is crucial since the identification of peptides able to interact and activate GPR120 would have many implications in terms of receptors pharmacology and related medicinal chemistry, as well as from a food science standpoint. As an example, the possible existence of GPR120 peptide agonists may significantly expand the chemical space of GPR120 ligands. In this respect, agonist peptides have not been described yet for GPR120. This is reasonably due to its relatively recent de-orphanization, though the existence of agonist peptides could be expected based on the growing evidence of peptides with agonists activity towards other GPCRs related to GPR120 (Davenport et al., 2020). The characterization of agonist peptides is also important toward a more informed understanding of fatty acid receptors pharmacology (including GPR120) to unlock their ultimate therapeutic potential and pharmacological intervention routes (Milligan et al., 2017). Regarding this case study, GPR120 has a role in several diseases, including cancer, inflammatory conditions, central nervous system disorders and Type 2 diabetes mellitus (Carullo et al., 2021). Moreover, GPR120 is abundantly expressed in entero-endocrine cells, adipocytes, taste buds of circumvallate, fungiform, and foliate papillae, and through the gastrointestinal tract (Carullo et al., 2021; Lu et al., 2018; Miyamoto et al., 2016). This could suggest a first-line role in the complex network of molecular events underpinning the action of food bioactive constituents. Therefore, from a pharmacological

standpoint, the two sequences identified here could suggest a promising scaffold to investigate further and derive compounds with enhanced pharmacological properties (as sub-optimal pharmacodynamics and pharmacokinetics are expected for them, as typically reported for naturally occurring peptides; (Davenport et al., 2020)). From a food science perspective, the identification of peptides able to activate GPR120 deserves attention considering the multifaced role it has been ascribed to, and consequent plausible implications in the food-health binomial. Specifically, further investigations of GLFGGG and GIFGGG shall reveal to what extent they effectively activate GPR120 in a real-world scenario. This may promote a thorough identification of GPR120 agonist peptides of food origin from a foodomics perspective – referred to as the comprehensive, high-throughput fingerprinting of food bioactives with a role to improve human nutrition and wellbeing (Capozzi & Bordoni, 2013). In this respect, the identification of GPR120 agonist peptides should be systematically pursued considering that they typically have an individual limited pharmacological activity, but the complex mixture of peptides of food origin may reach pharmacologically relevant concentrations as a chemical complex rather than single substances. In addition, the possible involvement of GPR120 in the perception of the so called “fatty taste” (Iwasaki et al., 2021) may lead to consider selected peptides as useful scaffold to derive fatty taste elicitors. On the other side, a deeper understanding of peptides-GPR120 interaction and subsequent activation may shed light on the chemistry of taste toward a more informed profiling of taste active components of food. The above makes the systematic and comprehensive study of GPR120 agonist peptides of food origin critical. In this respect, the sequences identified here described hexapeptides rich in glycine as the prototype scaffold for agonist peptides of GPR120 and provided a compelling line of evidence pointing to the need of moving pioneering steps to instruct knowledge-based research of new peptide-based GPR120 ligands.

3.3 Identification of possible sources of food origin for GLFGGG and GIFGGG

As discussed above, the identification of possible food sources of GLFGGG and GIFGGG is crucial to move the outcome presented to a real-world scenario. This may be important either to characterize better the chemical complex of certain foods possibly responsible for the biological outcomes associated with their consumption, or to identify valuable sources of bioactives to derive e.g. functional foods or nutraceuticals (Lammi et al., 2021).

Germane to the identification of possible sources of GLFGGG and GIFGGG, both sequences have been searched within the whole SwissProt database (570 420 protein sequences; last

database access 6th) (Bateman et al., 2023) collecting information about "reviewed" sequences (i.e. derived from scientific literature or, if computationally derived, curated and evaluated by an expert analysis). The focus on the SwissProt database, rather than extending the analysis over the whole UniProt database, aimed at pursuing the identification of reasonably existing proteins, excluding those hypotheticals from the analysis (Bateman et al., 2023)

As shown in Table 2, 36 diverse UniProt entries from several organisms, including procaryotes and eukaryotes, were found to contain GLFGGG (27 entries) or GIFGGG (9 entries). The relative shortage of proteins found is in line with the particularity of the two sequences under analysis. Indeed, polyglycine segments are not functionally neutral and they are typically associated with poorly organized protein regions and/or signals which may impact protein function, folding, transport and turnover (Bykov & Asher, 2010; Endow et al., 2016). For this reason, its frequency is expected to be low. Among those identified, three were of possible interest from a food science standpoint: endoglucanase 4 from *Oryza sativa* (rice), ferredoxin--NADP reductase from *Lactobacillus delbrueckii subsp. Bulgaricus* (lactic acid bacteria common in foodstuff), and ligninase C from *Trametes versicolor* (a mushroom not edible *per se* but considered for food supplements production). These are possible valuable sources from a food perspective as rice is massively consumed worldwide with many uses also as food supplement, *L. delbrueckii* is used in a large variety of fermented milk products, and *T. versicolor*, which is not considered edible for humans though it is used to manufacture food supplements (Barros et al., 2016). However, it was not possible to retrieve from the literature clear information about the level of expression of rice endoglucanase 4 and bacterial ferredoxin—NADP reductase within the respective organism of origin. Therefore, the relevance of GIFGGG/GLFGGG-containing proteins and related organisms from a food production standpoint could not be ascertained. Conversely, ligninase C is a key enzyme of *T. versicolor* expected to be highly expressed as it is crucial for lignin degradation, a fundamental step in the physiology and nourishment of this fungus (Jonsson & Nyman, 1992). Of note, *T. versicolor* has been already described as a promising organism rich in health-promoting components, and it has been also considered as ingredients for food supplements (Barros et al., 2016; Janjusevic et al., 2017; Janjusevic et al., 2018). Therefore, besides considering the pure sequence obtained via chemical synthesis, the consumption of food supplements enriched with or made of *T. versicolor* might be a possible valuable source of GIFGGG worth of further dedicated analysis.

Table 2. Possible protein sources of GLFGGG and GIFGGG

UniProt ID	Organism	Protein
Q96L46[a]	*Homo sapiens*	Calpain small subunit 2
A5U7B6[a] P9WG18[a] P9WG19[a]	*Mycobacterium tuberculosis*	Cell division protein FtsX
Q7TX91[a]	*M. bovis*	
O32882[a]	*M. leprae*	
P27747[a]	*Cupriavidus necator*	Dihydrolipoyllysine-residue acetyltransferase component of acetoin cleaving system
A6L2J8[b]	*Phocaeicola vulgatus*	Dipeptidyl-peptidase 7
Q5LB17[b]	*Bacteroides fragilis*	Dipeptidyl-peptidase 7
B4UCY7[a]	*Anaeromyxobacter sp. (strain K)*	DNA mismatch repair protein MutS
B8JA66[a] Q2IIJ3[a]	*A. dehalogenans*	
Q6Z715[b]	*Oryza sativa subsp. japonica*	Endoglucanase 4
Q049B3[b] **Q1G967[b]**	*Lactobacillus delbrueckii subsp. bulgaricus*	Ferredoxin--NADP reductase
Q6Q8A5[a]	*Nicotiana tabacum*	Hexokinase-2
P13645[b]	*H. sapiens*	Keratin, type I cytoskeletal 10
P20013[b]	*Trametes versicolor*	Ligninase C
P91928[a]	*Drosophila melanogaster*	MICOS complex subunit Mic60
Q9VCH5[a]	*D. melanogaster*	Nuclear pore complex protein Nup98-Nup96
O82230[a]	*Arabidopsis thaliana*	Nucleoid-associated protein At2g24020
Q9M098[a]	*A. thaliana*	Nucleoid-associated protein At4g30620
G0SAK3[a]	*Chaetomium thermophilum*	Nucleoporin NUP145
Q9UTK4[a]	*Schizosaccharomyces pombe*	Nucleoporin nup189
B2HE92[a]	*M. marinum*	PE cleavage protein A
L0T4W6[b]	*M. tuberculosis*	PE-PGRS family protein PE_PGRS4
Q2NT27[b]	*Sodalis glossinidius*	Phenylalanine--tRNA ligase beta subunit
E1XTG6[a]	*Salmonella phage ViI*	Portal protein
P0CQ46[a] P0CQ47[a]	*Cryptococcus neoformans*	Protein SEY1

Q9WU70[a]	*Rattus norvegicus*	
Q5T5C0[a]	*H. sapiens*	Syntaxin-binding protein
Q8K400[a]	*Mus musculus*	
Q54LC9[a]	*Dictyostelium discoideum*	Uncharacterized Golgi apparatus membrane protein-like protein
Q9HDZ8[a]	*S. pombe*	Uncharacterized protein C589.06c
Q9XJR4[a]	*Pseudoalteromonas phage PM2*	Uncharacterized protein Gp-j

Note. [a] Presence of GLFGGG; [b] Presence of GIFGGG. In bold those who have been considered relevant from a food perspective

3.4 D-amino acid containing derivatives of GIFGGG

It has been previously demonstrated that the inclusion of D-amino acid in peptides may significantly enhance their resistance to proteases activity (Yan et al., 2020). Therefore, calculations of D-amino acid containing derivatives of GIFGGG and GLFGGG were added for the sake of estimating the possible activity of derivatives likely more resistant to gastrointestinal digestion. Particularly, all the possible combinations including one or two D-amino acids in the two sequences were analyzed (Table 1). As shown in Table 1, one out of the six combinations possible (namely, GdIFGGG) recorded a docking score higher than the parental peptide suggesting a better fitting into the pocket and possibly a higher activity toward GPR120. Therefore, the GdIFGGG-GPR120 complex underwent CMD and US to estimate the geometrical stability over time and the free energy of binding. As shown in Figure 3 and S2 (see Supplementary Materials), the RMSD of the carboxylate moiety with respect to the protein was higher than the corresponding L-peptide. However, this was due to an initial adjustment of the starting binding pose which eventually stabilized on a trend comparable to its L counterpart. Moreover, as shown in the protein RMSD plot (Figure 3), it kept the overall protein structure as stable as the other complexes with positive controls (i.e. oleic acid and TUG-891). Regarding the US simulation, once GdIFGGG was pulled out for 5 nm, 26 windows with an approximate 0.2 nm interval were chosen to perform the US simulation (see section 2.6). The collection of histograms and the umbrella potential graph are shown in Figure 4. The peak distribution in the histogram reflected the total coverage for the pulling simulation obtained via the parameters set (section 2.6). Indeed, the evenly spaced windows were sufficient to cover all the regions GdIFGGG passed by, exhibiting a sufficient overlap. The analysis of the umbrella potential graph showed a ΔG corresponding to roughly -13 kcal/mol, pointing to an interaction with GPR120 possibly more favored than that with its L counterpart. This was in line with other

studies on protein-peptide complexes (Cool & Lindert, 2022). Moreover, previous evidence described that hexapeptides may bind with a high affinity GPCRs with a high affinity showing values of free energy of binding in the range of that obtained here for GdIFGGG (Calderón et al., 2023). The finding that D-amino acid containing sequence may have an enhanced activity is important in the light of the common production of D-amino acid during food processing. Indeed, they can be either present in raw materials (such as some fruits and vegetables) or produced upon processing like fermentation or the application of thermal treatments (Genchi, 2017; Marcone et al., 2020). Germane to our study, the design of proper treatments of T. versicolor in food supplements production may result in an enriched fraction of D-amino acid containing peptides, which are worth of future dedicated investigations.

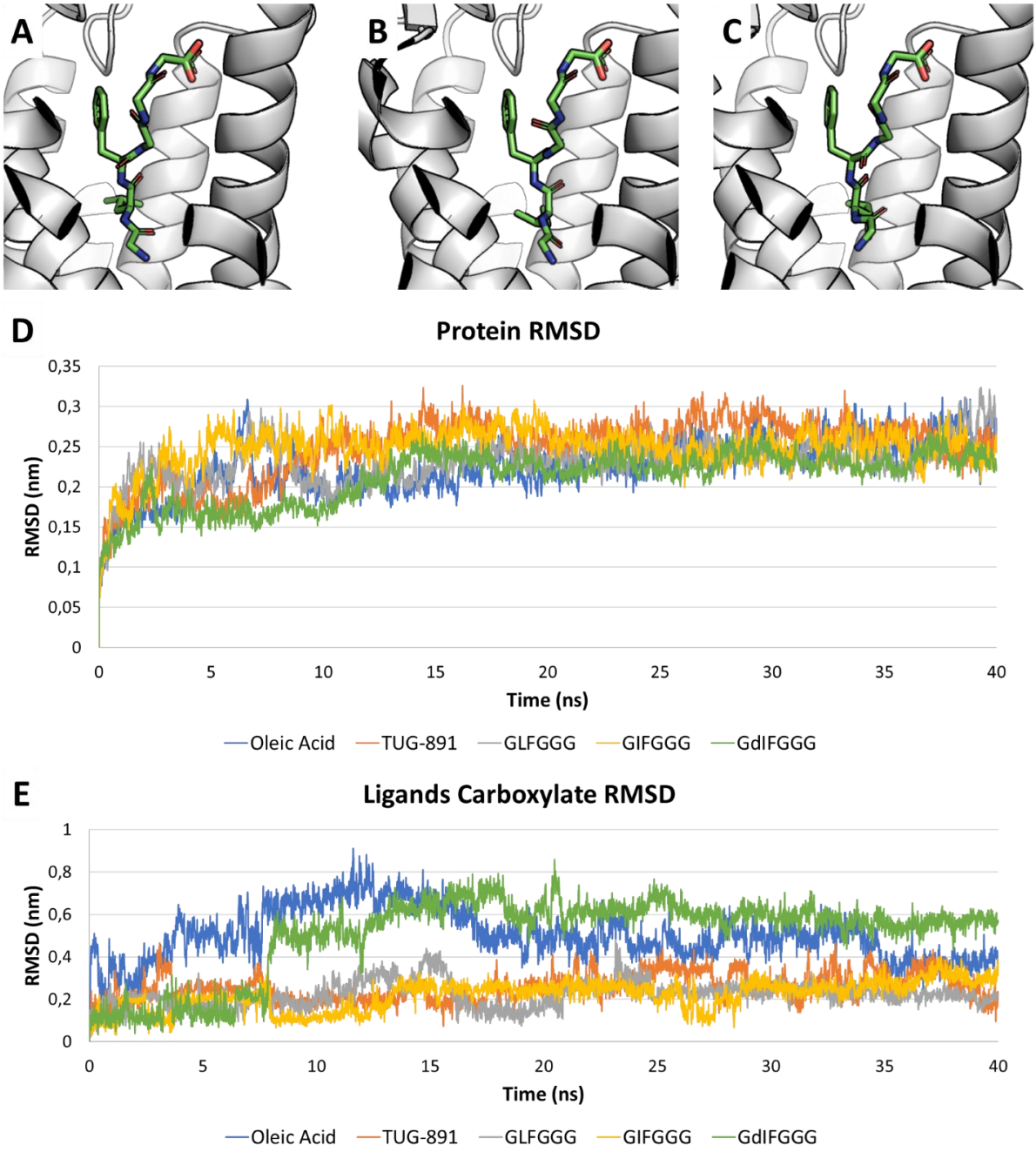


Figure 3. Molecular docking and CMD simulations results for GLFGGG, GIFGGG and GdIFGGG. **A**. Best docking pose for GLFGGG represented as green sticks. **B**. Best docking pose for GIFGGG represented as green sticks. **C**. Best docking pose for GdIFGGG represented as green sticks. **D**. Protein RMSD graph comparing the positive controls (i.e. oleic acid and TUG-891) with GLFGGG, GIFGGG and GdIFGGG. As appreciable from the graph the protein RMSD trend are comparable. **E**. Ligands carboxylate RMSD representing the distance of the carboxylate moiety from the protein. All the tested ligands show a stable trend within the GPR120's binding pocket although GdIFGG (green) showed an initial movement which stabilized around 8 ns from the beginning of the CMD.

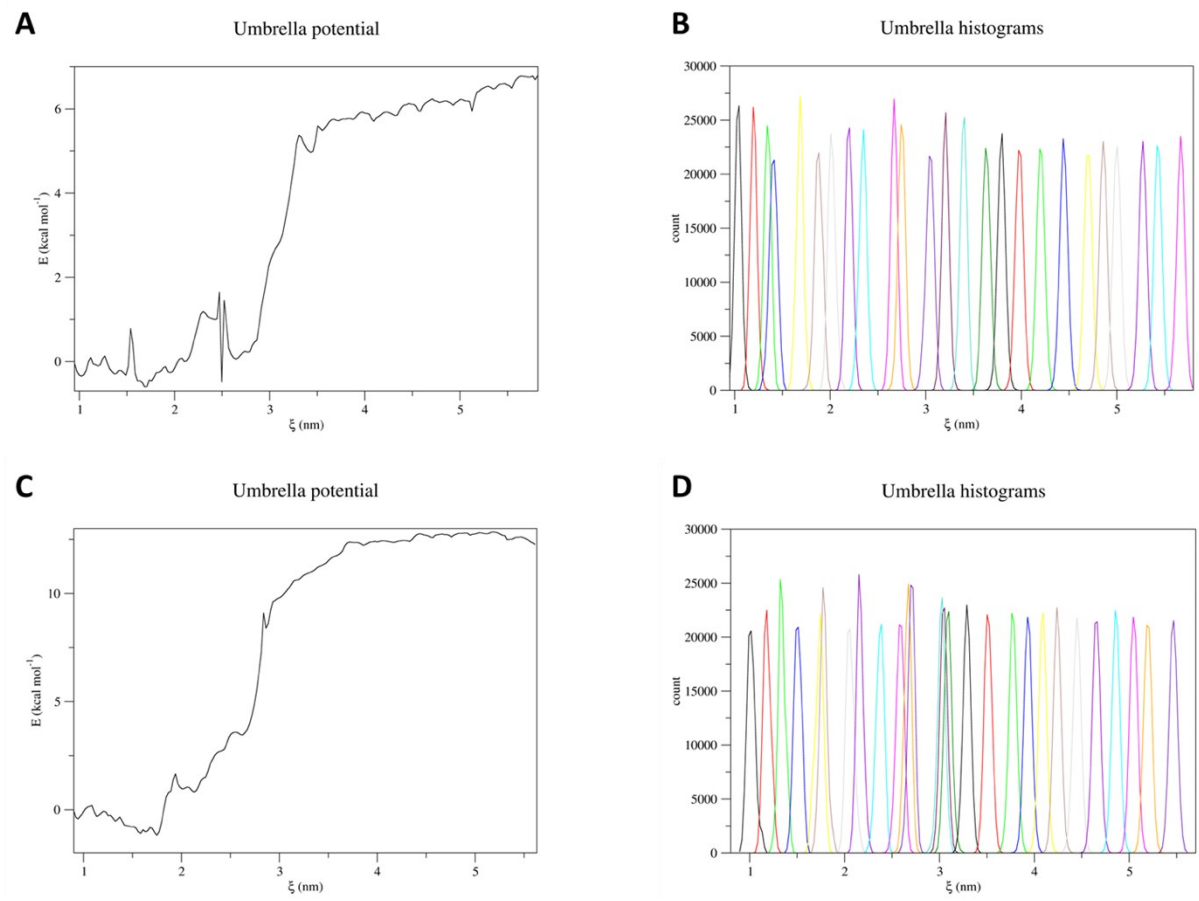


Figure 4. US results of the GIFGGG and GdIFGGG peptides. **A**. GIFGGG umbrella potential graph. **B**. GIFGGG histograms distribution showing total coverage of the pulling procedure. **C**. GdIFGGG umbrella potential graph. **D**. GdIFGGG histograms distribution showing total coverage of the pulling procedure.

4. Conclusion

The present work identified from a virtual library of more than 500 000 entries two hexapeptides, i.e. GLFGGG and GIFGGG, with a potential agonist activity toward GPR120 which are worth of further dedicated investigations. GPR120 is an intriguing target relevant to pharmaceutical and food science being involved in the fatty taste perception as well as in a wealth of biological outcomes reasonably germane to the mechanics of food bioactives. The identification of peptides as possible agonists of GPR120 is consistent with evidence for other GPCRs and may expand the chemical space of GPR120 ligands. In this respect, our results provided a convincing rationale supporting peptides as useful scaffold to derive new agonists in further dedicated studies. This may be important for food science as GPR120-agonist peptides and "natural" peptidomimetic compounds (such as D-amino acids containing derivatives) might be considered for food supplements/functional foods production or taste elicitors development. The outcomes presented also highlighted that certain food-related proteins may be a valuable source for the identified sequences. Finally, the analysis also covered D-amino acid containing sequences, which are reasonably more resistant to gastrointestinal digestion, and described GdIFGGG having a better interaction compared to its all-L counterpart (-13 kCal/mol ΔG and –6 kCal/mol ΔG, respectively). This may be important in the light of designing nutraceuticals or functional foods targeting GPR120.

References

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX, 1-2*, 19-25. doi:https://doi.org/10.1016/j.softx.2015.06.001

Ahmad, R., & Dalziel, J. E. (2020). G Protein-Coupled Receptors in Taste Physiology and Pharmacology. *Frontiers in Pharmacology, 11*. doi:10.3389/fphar.2020.587664

Anbazhagan, A. N., Priyamvada, S., Gujral, T., Bhattacharyya, S., Alrefai, W. A., Dudeja, P. K., & Borthakur, A. (2016). A novel anti-inflammatory role of GPR120 in intestinal epithelial cells. *American Journal of Physiology-Cell Physiology, 310*(7), C612-C621. doi:10.1152/ajpcell.00123.2015

Barros, A. B., Ferrao, J., & Fernandes, T. (2016). A safety assessment of *Coriolus versicolor* biomass as a food supplement. *Food & Nutrition Research, 60*. doi:10.3402/fnr.v60.29953

Bateman, A., Martin, M. J., Orchard, S., Magrane, M., Ahmad, S., Alpi, E., . . . UniProt, C. (2023). UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research, 51*(D1), D523-D531. doi:10.1093/nar/gkac1052

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., . . . Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research, 28*(1), 235-242. doi:10.1093/nar/28.1.235

Burns, R. N., & Moniri, N. H. (2010). Agonism with the omega-3 fatty acids α-linolenic acid and docosahexaenoic acid mediates phosphorylation of both the short and long isoforms of the human GPR120 receptor. *Biochemical and Biophysical Research Communications, 396*(4), 1030-1035. doi:10.1016/j.bbrc.2010.05.057

Bykov, S., & Asher, S. (2010). Raman Studies of Solution Polyglycine Conformations. *Journal of Physical Chemistry B, 114*(19), 6636-6641. doi:10.1021/jp100082n

Calderón, J. C., Plut, E., Keller, M., Cabrele, C., Reiser, O., Gervasio, F. L., & Clark, T. (2023). Extended Metadynamics Protocol for Binding/Unbinding Free Energies of Peptide Ligands to Class A G-Protein-Coupled Receptors. *Journal of Chemical Information and Modeling, 64*(1), 205-218. doi:10.1021/acs.jcim.3c01574

Capozzi, F., & Bordoni, A. (2013). Foodomics: a new comprehensive approach to food and nutrition. *Genes and Nutrition, 8*(1), 1-4. doi:10.1007/s12263-012-0310-x

Carullo, G., Mazzotta, S., Vega-Holm, M., Iglesias-Guerra, F., Vega-Pérez, J. M., Aiello, F., & Brizzi, A. (2021). GPR120/FFAR4 Pharmacology: Focus on Agonists in Type 2 Diabetes Mellitus Drug Discovery. *Journal of Medicinal Chemistry, 64*(8), 4312-4332. doi:10.1021/acs.jmedchem.0c01002

Civciristov, S., Ellisdon, A. M., Suderman, R., Pon, C. K., Evans, B. A., Kleifeld, O., . . . Halls, M. L. (2018). Preassembled GPCR signaling complexes mediate distinct cellular responses to ultralow ligand concentrations. *Science Signaling, 11*(551). doi:10.1126/scisignal.aan1188

Cool, A. M., & Lindert, S. (2022). Umbrella Sampling Simulations Measure Switch Peptide Binding and Hydrophobic Patch Opening Free Energies in Cardiac Troponin. *Journal of Chemical Information and Modeling, 62*(22), 5666-5674. doi:10.1021/acs.jcim.2c00508

Davenport, A. P., Scully, C. C. G., de Graaf, C., Brown, A. J. H., & Maguire, J. J. (2020). Advances in therapeutic peptides targeting G protein-coupled receptors. *Nature Reviews Drug Discovery, 19*(6), 389-413. doi:10.1038/s41573-020-0062-z

de Castro, R. J. S., & Sato, H. H. (2015). Biologically active peptides: Processes for their generation, purification and identification and applications as natural additives in the food and pharmaceutical industries. *Food Research International, 74*, 185-198. doi:10.1016/j.foodres.2015.05.013

Del Favero, G., Aichinger, G., Hohenbichler, J., Marko, D., Dall'Asta, C., & Dellafiora, L. (2022). A target fishing study to spot possible biological targets of fusaric acid: Inhibition of protein kinase-A and insights on the underpinning mechanisms. *Food and Chemical Toxicology, 159*. doi:10.1016/j.fct.2021.112663

Dellafiora, L., Magnaghi, F., Galaverna, G., & Dall'Asta, C. (2022). A mechanistic investigation on kokumi-active γ-Glutamyl tripeptides - A computational study to understand molecular basis of their activity and to identify novel potential kokumi-tasting sequences. *Food Research International, 162*. doi:10.1016/j.foodres.2022.111932

Dellafiora, L., Oswald, I. P., Dorne, J. L., Galaverna, G., Battilani, P., & Dall'Asta, C. (2020). An in silico structural approach to characterize human and rainbow trout estrogenicity of mycotoxins: Proof of concept study using zearalenone and alternariol. *Food Chemistry, 312*. doi:10.1016/j.foodchem.2019.126088

Endow, J. K., Agostinho, G. R. A., Baldwin, A. J., Roston, R. L., Yamaguchi, T., Kamikubo, H., & Inoue, K. (2016). Polyglycine Acts as a Rejection Signal for Protein Transport at the Chloroplast Envelope. *Plos One, 11*(12). doi:10.1371/journal.pone.0167802

Galindo, M. M., Voigt, N., Stein, J., van Lengerich, J., Raguse, J. D., Hofmann, T., . . . Behrens, M. (2012). G Protein-Coupled Receptors in Human Fat Taste Perception. *Chemical Senses, 37*(2), 123-139. doi:10.1093/chemse/bjr069

Genchi, G. (2017). An overview on D-amino acids. *Amino Acids, 49*(9), 1521-1533. doi:10.1007/s00726-017-2459-5

Guo, S. M., Zhao, T. T., Yun, Y., & Xie, X. (2022). Recent progress in assays for GPCR drug discovery. *American Journal of Physiology-Cell Physiology, 323*(2), C583-C594. doi:10.1152/ajpcell.00464.2021

Hirasawa, A., Tsumaya, K., Awaji, T., Katsuma, S., Adachi, T., Yamada, M., . . . Tsujimoto, G. (2005). Free fatty acids regulate gut incretin glucagon-like peptide-1 secretion through GPR120. *Nature Medicine, 11*(1), 90-94. doi:10.1038/nm1168

Huang, J., & MacKerell, A. D. (2013). CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of Computational Chemistry, 34*(25), 2135-2145. doi:10.1002/jcc.23354

Hudson, B. D., Shimpukade, B., Mackenzie, A. E., Butcher, A. J., Pediani, J. D., Christiansen, E., . . . Milligan, G. (2013). The Pharmacology of TUG-891, a Potent and Selective Agonist of the Free Fatty Acid Receptor 4 (FFA4/GPR120), Demonstrates Both Potential Opportunity and Possible Challenges to Therapeutic Agonism. *Molecular Pharmacology, 84*(5), 710-725. doi:10.1124/mol.113.087783

Iwasaki, N., Sakamoto, K., Kitajima, S., Maruyama, Y., & Kuroda, M. (2021). GPR120 agonists enhance the fatty orosensation when added to fat-containing system, but do not evoke it by themselves in humans. *Physiology & Behavior, 234*. doi:10.1016/j.physbeh.2021.113383

Janjusevic, L., Karaman, M., Sibul, F., Tommonaro, G., Iodice, C., Jakovljevic, D., & Pejin, B. (2017). The lignicolous fungus Trametes versicolor (L.) Lloyd (1920): a promising natural source of antiradical and AChE inhibitory agents. *Journal of Enzyme Inhibition and Medicinal Chemistry, 32*(1), 355-362. doi:10.1080/14756366.2016.1252759

Janjusevic, L., Pejin, B., Kaisarevic, S., Gorjanovic, S., Pastor, F., Tesanovic, K., & Karaman, M. (2018). Trametes versicolor ethanol extract, a promising candidate for health-promoting food supplement. *Natural Product Research, 32*(8), 963-967. doi:10.1080/14786419.2017.1366484

Jo, S., Kim, T., Iyer, V. G., & Im, W. (2008). Software news and updates - CHARNIM-GUI: A web-based grraphical user interface for CHARMM. *Journal of Computational Chemistry, 29*(11), 1859-1865. doi:10.1002/jcc.20945

Jonsson, L., & Nyman, P. O. (1992). CHARACTERIZATION OF A LIGNIN PEROXIDASE GENE FROM THE WHITE-ROT FUNGUS TRAMETES-VERSICOLOR. *Biochimie, 74*(2), 177-182. doi:10.1016/0300-9084(92)90043-e

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., . . . Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583-+. doi:10.1038/s41586-021-03819-2

Kim, S., Chen, J., Cheng, T. J., Gindulyte, A., He, J., He, S. Q., . . . Bolton, E. E. (2023). PubChem 2023 update. *Nucleic Acids Research, 51*(D1), D1373-D1380. doi:10.1093/nar/gkac956

Kimura, I., Ichimura, A., Ohue-Kitano, R., & Igarashi, M. (2020). FREE FATTY ACID RECEPTORS IN HEALTH AND DISEASE. *Physiological Reviews, 100*(1), 171-210. doi:10.1152/physrev.00041.2018

Krumm, B. E., & Grisshammer, R. (2015). Peptide ligand recognition by G protein-coupled receptors. *Frontiers in Pharmacology, 6*. doi:10.3389/fphar.2015.00048

Lammi, C., Boschin, G., Bollati, C., Arnoldi, A., Galaverna, G., & Dellafiora, L. (2021). A heuristic, computer-driven and top-down approach to identify novel bioactive peptides: A proof-of-principle on angiotensin I converting enzyme inhibitory peptides. *Food Research International, 150*. doi:10.1016/j.foodres.2021.110753

Lee, J., Cheng, X., Swails, J. M., Yeom, M. S., Eastman, P. K., Lemkul, J. A., . . . Im, W. (2016). CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of Chemical Theory and Computation, 12*(1), 405-413. doi:10.1021/acs.jctc.5b00935

Legnik, S., Stular, T., Brus, B., Knez, D., Gobec, S., Janezic, D., & Konc, J. (2015). LiSiCA: A Software for Ligand-Based Virtual Screening and Its Application for the Discovery of Butyrylcholinesterase Inhibitors. *Journal of Chemical Information and Modeling, 55*(8), 1521-1528. doi:10.1021/acs.jcim.5b00136

Lu, V., Gribble, F. M., & Reimann, F. (2018). Free Fatty Acid Receptors in Enteroendocrine Cells. *Endocrinology, 159*(7), 2826-2835. doi:10.1210/en.2018-00261

Mao, C. Y., Xiao, P., Tao, X. N., Qin, J., He, Q. T., Zhang, C., . . . Sun, J. P. (2023). Unsaturated bond recognition leads to biased signal in a fatty acid receptor. *Science, 380*(6640). doi:10.1126/science.add6220

Marcone, G. L., Rosini, E., Crespi, E., & Pollegioni, L. (2020). D-amino acids in foods. *Applied Microbiology and Biotechnology, 104*(2), 555-574. doi:10.1007/s00253-019-10264-9

McKillop, A. M., Miskelly, M. G., Moran, B. M., & Flatt, P. R. (2023). Incretins play an important role in FFA4/GPR120 regulation of glucose metabolism by GW-9508. *Life Sciences, 318*. doi:10.1016/j.lfs.2023.121475

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. doi:10.1093/toxsci/56.1.8

Milligan, G., ShimpukadE, B., Ulven, T., & Hudson, B. D. (2017). Complex Pharmacology of Free Fatty Acid Receptors. *Chemical Reviews, 117*(1), 67-110. doi:10.1021/acs.chemrev.6b00056

Miyamoto, J., Hasegawa, S., Kasubuchi, M., Ichimura, A., Nakajima, A., & Kimura, I. (2016). Nutritional Signaling via Free Fatty Acid Receptors. *International Journal of Molecular Sciences, 17*(4). doi:10.3390/ijms17040450

O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. *Journal of Cheminformatics, 3*. doi:10.1186/1758-2946-3-33

Oh, D. Y., Talukdar, S., Bae, E. J., Imamura, T., Morinaga, H., Fan, W. Q., . . . Olefsky, J. M. (2010). GPR120 Is an Omega-3 Fatty Acid Receptor Mediating Potent Anti-inflammatory and Insulin-Sensitizing Effects. *Cell, 142*(5), 687-698. doi:10.1016/j.cell.2010.07.041

Pal, A., Curtin, J. F., & Kinsella, G. K. (2021). Structure based prediction of a novel GPR120 antagonist based on pharmacophore screening and molecular dynamics simulations. *Computational and Structural Biotechnology Journal, 19*, 6050-6063. doi:10.1016/j.csbj.2021.11.005

Pedroni, L., Perugino, F., Kurtaga, A., Galaverna, G., Dall'Asta, C., & Dellafiora, L. (2023). The bitter side of toxicity: A big data analysis spotted the interaction between trichothecenes and bitter receptors. *Food Research International, 173*. doi:10.1016/j.foodres.2023.113284

Thibeault, P. E., LeSarge, J. C., Arends, D., Fernandes, M., Chidiac, P., Stathopulos, P. B., . . . Ramachandran, R. (2020). Molecular basis for activation and biased signaling at the thrombin-activated GPCR proteinase activated receptor-4 (PAR4). *Journal of Biological Chemistry, 295*(8), 2520-2540. doi:10.1074/jbc.RA119.011461

Udenigwe, C. C., & Aluko, R. E. (2012). Food protein-derived bioactive peptides: production, processing, and potential health benefits. *J Food Sci, 77*(1), R11-24. doi:10.1111/j.1750-3841.2011.02455.x

Vidal-Limon, A., Aguilar-Toalá, J. E., & Liceaga, A. M. (2022). Integration of Molecular Docking Analysis and Molecular Dynamics Simulations for Studying Food Proteins and Bioactive Peptides. *Journal of Agricultural and Food Chemistry, 70*(4), 934-943. doi:10.1021/acs.jafc.1c06110

Wang, W. L., Cui, Z. Y., Ning, M. H., Zhou, T. X., & Liu, Y. (2022). In-silico investigation of umami peptides with receptor T1R1/T1R3 for the discovering potential targets: A combined modeling approach. *Biomaterials, 281*. doi:10.1016/j.biomaterials.2021.121338

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., . . . Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research, 46*(W1), W296-W303. doi:10.1093/nar/gky427

Wu, Y. R., Li, X. X., Hua, T., Liu, Z. J., Liu, H. G., & Zhao, S. W. (2022). MD Simulations Revealing Special Activation Mechanism of Cannabinoid Receptor 1. *Frontiers in Molecular Biosciences, 9*. doi:10.3389/fmolb.2022.860035

Yan, L., Ke, Y. Q., Kan, Y. H., Lin, D., Yang, J. K., He, Y. J., & Wu, L. (2020). New insight into enzymatic hydrolysis of peptides with site-specific amino acid D-isomerization. *Bioorganic Chemistry, 105*. doi:10.1016/j.bioorg.2020.104389

Yu, Z. P., Wang, Y. X., Zhao, W. Z., Li, J. R., Shuian, D., & Liu, J. B. (2022). Identification of Oncorhynchus mykiss nebulin-derived peptides as bitter taste receptor TAS2R14 blockers by in silico screening and molecular docking. *Food Chemistry, 368*. doi:10.1016/j.foodchem.2021.130839

Zoete, V., Cuendet, M. A., Grosdidier, A., & Michielin, O. (2011). SwissParam: A Fast Force Field Generation Tool for Small Organic Molecules. *Journal of Computational Chemistry, 32*(11), 2359-2368. doi:10.1002/jcc.21816

Chapter VIII

A Machine Learning based prefiltering approach to screen GPR120 putative binders

This chapter is based on a manuscript in preparation having the PhD candidate as first author.

Abstract

GPR120, also known as free fatty acid receptor 4 (FFAR4), is a G-protein-coupled receptor (GPCR) activated by long-chain fatty acids. Due to its involvement in metabolic regulation, inflammatory process, as well as in the debated perception of fat taste, GPR120 has gained increasing attention as an attractive target from both a pharmaceutical and a food science perspective. Considering his physiological and pharmacological relevance, the identification of novel GPR120 ligands represents a pivotal step toward the development of new therapeutic agents, nutraceuticals, and functional foods. Structure-based virtual screening relying on molecular docking represents a reliable and efficient strategy for identifying novel GPR120 ligands. However, the application of structure-based virtual screening to ultra-large libraries remains highly demanding in terms of time and computational resources. In this scenario, we developed a machine learning (ML) model capable of predicting potential GPR120 binders, thus reducing the number of molecules requiring molecular docking. In this study, a Gradient Boosting on Decision Trees (GB on DT) algorithm was trained on docking scores of 5 million molecules randomly sampled from the Enamine REAL database encoded as Extendended Connectivity Fingerprints (ECFPs). Two virtual screening strategies were evaluated and compared. Stratification of the molecules according to molecular size was found to be crucial to overcoming inherent docking bias. The refined model successfully reduced by nearly 70% the number of molecules to be further docked. A subset of the predicted molecules was then sorted from each probability range and further docked within GPR120, corroborating the model's ability to correctly classify compounds distinguishing between optimal and non-optimal docking scores. Overall, the developed model provided a reliable and efficient approach for prefiltering molecules from ultra-large libraries before running docking simulation. Notably, the *in silico* pipeline here refined could be readily adapted to other GPCRs or to other molecular targets relevant from a pharmaceutical and/or from a food science perspective.

Chapter VIII

1. Introduction

G-protein-coupled receptors represent (GPCRs) the largest class of cell-surface signaling proteins. They act interacting with the heterotrimeric G proteins with their intracellular domain. Upon agonist binding, GPCRs undergo conformational changes that stabilize their active state, leading to the dissociation of the G protein into Gα and Gβγ subunits. Once activated, these subunits interact with other specific intracellular effectors initiating a cascade of downstream signaling events (Civciristov et al., 2018). Recently, several deorphanized GPCRs were identified as free fatty acid receptors (FFARs), including GPR120 (Figure 1), known as free fatty acid receptor 4 (FFAR4). This receptor is activated by long chain free fatty acids, mainly 14-22 carbons in length including the family of omega-3 (n-3) polyunsaturated fatty acids (PUFA) (Moniri, 2016). Dietary fatty acids play a pivotal role in human physiology, representing not only a primary energy source but also serving as fundamental structural components of cell membranes and modulating the immune and the inflammatory response (Hasan et al., 2017). From a molecular perspective, two GPR120 splice isoforms have been described: a long one (377 amino acids) and a short one (361 amino acids), differing for the insertion of 16 amino acids between position 231 and 247 in the third intracellular loop. However, no evidence pointing out a functional difference between the two isoforms has been reported to date (Pal et al., 2021). GPR120 is expressed in various tissue and cell types, including stomach and duodenum, adipose tissue, macrophages, lung (Cornall et al., 2014), and pancreas (Zhao et al., 2013). It has also been detected in the human lingual gustatory epithelium. For this reason, over the last decade, it has also been described as involved in the still debated human gustatory fatty acid perception, namely “fatty taste” (Galindo et al., 2012). Beyond sensory functions, GPR120 was suggested to be involved in the regulation of hormone secretions including glucagon-like peptide 1 (GLP-1), gastric inhibitory polypeptide (GIP) or cholecystokinin (CCK) in intestinal endocrine cells (Hirasawa et al., 2005; Liu et al., 2015). Consequently, it may play a role in disorders related to obesity, satiety regulations, and type 2 diabetes mellitus (T2DM) (Hirasawa et al., 2005). For all these reasons, the interest in GPR120 as a novel pharmaceutical target is deserving increasing attention (Milligan et al., 2017). Free fatty acids are the natural ligands of GPR120, but over the years, several GPR120 agonists have been identified, including natural and synthetic compounds (e.g. the known synthetic agonist TUG-891) (Hudson et al., 2013; Pedroni et al., 2024). The identification of additional new agonists would be relevant to expanding the current knowledge about chemical space of GPR120 ligands, a fundamental feature in guiding drug discovery (Carlsson & Luttens, 2024).

Considering his physiological and pharmacological relevance of GPR120, a pipeline for ligand identification represents a pivotal step toward the development of new therapeutic agents, nutraceuticals, and functional foods. Structure-based virtual screens of ultra-large chemical libraries are widespread and promising methods used in drug discovery process to identify novel ligands for a therapeutic target (Carullo et al., 2021). Typically, they rely on molecular docking to estimate binding affinities and to predict plausible protein-ligand binding architectures (Adeshina et al., 2020; Pedroni, Perugino, et al., 2023). Molecular docking has already demonstrated significant success also in the discovery of GPCRs ligands, pointing to the suitability of GPCRs binding pocket to structure-based methods (Carlsson & Luttens, 2024). However, the rapid and exponential growth of the make-on-demand databases, now containing billions of synthetical available compounds, makes it unfeasible to screen all the accessible molecules even with the fastest structure-based docking algorithms (Luttens et al., 2025) or high-performance computer resource. In this context, the need for a more efficient virtual screening approach would be essential to reduce both time and computational efforts and algorithms such as DeepDock have been developed (Z. Liao, 2019). Also Machine learning (ML) has already proven to be a valuable strategy to improve, enhance, and accelerate structure-based virtual screening (Bai et al., 2025; Ghislat et al., 2021). Recently, Luttens at al. proved the potential of combining molecular docking and molecular descriptors to screen vast chemical space with ML-based models (Luttens et al., 2025). Indeed, ML models can learn complex non-linear relationships such as between molecular descriptors (e.g., fingerprints) and docking scores, thus predicting the likelihood of molecular binding avoid docking of unpromising molecules. In brief, the aim of this work was to combine the already known reliability of molecular docking (Di Pizio & Nicoli, 2020; Perugino et al., 2024b) with the efficiency of artificial intelligence to empower and reinforce the discovery of bioactive compounds. Specifically, we refined a ML model trained on docking score and Morgan fingerprints (ECFP4) able to classify potential GPR120 binders and non-binders without performing docking calculations. Overall, the model here presented: i) is not intended to replace molecular docking but to prefilter compounds and reduce the number requiring further docking; ii) overcomes key docking limitations (i.e. bias toward larger molecules); iii) minimizes time and computational efforts; iv) provides a flexible and generalizable framework adaptable to other GPCRs or to other relevant molecular targets. By integrating structure-based validation with ML, this approach established a robust and scalable virtual screening pipeline, well framed in the contest of the New Approaching Methodologies (NAMs), which encompasses a variety of non-testing approaches, including computational methods, increasingly applied to predict human outcomes,

assess chemical safety and expedite drug development (Edwards et al., 2025; Pedroni, Louisse, et al., 2023).

2. Material and Methods

2.1 Data source and retrieval

All the known GPR120 ligands available in the ChEMBL database (Gaulton et al., 2012) were retrieved in SMILES format and filtered to remove duplicates obtaining unique compound identifiers. After this curation step, the dataset was further refined by applying Lipinski's Rule of Five (Ro5) (Lipinski et al., 1997) using RDKit (RDKit: Open-source cheminformatics; http://www.rdkit.org) in line with the criteria to be adopted for the prospective screening of the Enamine REAL compounds database. For model development and the subsequent virtual screening campaign, 5 million molecules were randomly sampled from the Enamine REAL database in CXSMILES format and converted to canonical SMILES using RDKit. The corresponding 3D structures were generated using the LigPrep tool implemented in Schrödinger Maestro (Schrödinger 2025-1) setting the pH to 7.0 and allowing the generation of at most one single conformation per molecule. The target protein structure was retrieved from the Protein Data Bank (PDB) with the PDB code 8G59 (Mao et al., 2023) due to its good resolution (2.64 Å). The structural continuity and refinement of the protein were obtained with ColabFold (v. 1.5.5), an AlphaFold2-based notebook using the 8G59 structure as a template and the FASTA sequence of the human GPR120 (Uniprot AC: Q5NUL3) as input. ColabFold is an easy-to-use software for protein structure and complex prediction using AlphaFold2 and Alphafold2-multimer notebook (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb) (Yang et al., 2023).

2.2 Molecular Docking

Molecular docking analyses were initially performed to assess the binding capability of known GPR120 ligands retrieved from ChEMBL and to explore their chemical space. More in detail, molecular docking simulations were performed using PLANTS 1.2 (O. Korb et al., 2009), an algorithm based on ant colony optimization (ACO) (Oliver Korb et al., 2007). The $PLANTS_{PLP}$ internal score was chosen as the scoring function, where a very negative score corresponds to a strong binding while a less negative or even positive score corresponds to a weak or unlikely binding (Gorgulla et al., 2021). All the curated compounds were docked within GPR120 binding site. Ligand 3D structures were prepared using LigPrep as implemented in the Schrödinger

Small-Molecule Drug Discovery (Schrödinger 2025-1) setting pH environment at 7.0 and allowing the generation of at most one conformation per molecule. LigPrep output molecules were obtained in Structured Data File (.sdf) and subsequentially converted into the Tripol Mol2 format (.mol2) using OpenBabel (v. 3.3.1) (O'Boyle et al., 2011). Protein preparation and grid generation were performed using Maestro (Schrödinger 2025-1). The docking grid was generated by selecting the ligand co-crystalized within 8G59 structure (namely, TUG-891). To validate the docking protocol, molecular docking simulation on TUG-891 was performed and the pose recording the best docking score was compared with the one from structural studies. 3D structure of TUG-891 was retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov/ (Kim et al., 2025); CID 57522038) in the sdf format and then converted into mol2 with OpenBabel. 10 poses for each ligand were generated. The chemical space of the docked compounds was analysed using the DataWarrior software (v. 06.01.00) (Lopez-Lopez et al., 2019). Based on this analysis, model development was restricted to molecules from the Enamine REAL database with a heavy atom count (HAC) between 15 and 36. The Enamine REAL database counts approximately 10.1 billion molecules (last database access July 2025), organized in different packages according to their HAC values. A total of 5 million molecules from the Enamine REAL database were proportionally sampled from the different packages in CXSMILES format. After that, the 3D structures of these molecules were generated using the LigPrep tool of Maestro in .sdf format and then converted into .mol2 using OpenBabel before performing docking analysis. Canonical SMILES were used to generate molecular descriptors. Specifically, Morgan fingerprint descriptors were generated using the open-source framework RDKit with 1024 bits and a radius=2.

2.3 The Algorithm

Supervised ML enables the automatic construction of regression and classification models from labeled datasets (Hancock & Khoshgoftaar, 2020). CatBoost is an open-source gradient boosting on decision trees (GB on DTs) algorithm developed by Prokhorenkova et al. (Prokhorenkova, 2018) representing a powerful tool for use in supervised ML applications (Hancock & Khoshgoftaar, 2020). This algorithm introduces a novel and advanced gradient boosting scheme that reduces overfitting using multiple random permutations of the dataset and averaging the statistics across them (Ahn et al., 2023). Another key advantage of CatBoost is its native GPUs support, which significantly accelerates training compared to CPU-based algorithms. This is particularly relevant in the context of our work where million molecules and

their relative molecular descriptors are processed. Furthermore, CatBoost is intended to process data also in small and ordered chunks. This capability minimizes memory issues allowing the training and the application of the model on millions of compounds and relative Morgan fingerprints. These features combined with the already proven success of CatBoost when paired with Morgan fingerprints for GPCRs ligands virtual screening (Luttens et al., 2025) make it the most suitable choice for both accelerating virtual screening process while ensuring high predictive performance.

2.4 The Prefilter Model

For model development, the best scored docking poses of the 5 million molecules sampled from the Enamine REAL database were used. Firstly, compounds were ranked by docking score in ascending order, from the most negative value (best computed binding) to the less negative value (worst computed binding). Two supervised classification models were trained and then compared to determine the most performing one. The label “1” was assigned to the positive dataset (molecules predicted to have good interactions to the receptor) while the label “0” was assigned to the negative dataset (molecules with suboptimal docking score). To set the threshold to determine whether a molecule is class 1 or class 0, two different approaches were tested. In the first approach, the entire dataset was split into two parts and the docking score recorded by the molecule at position 2,500,000 (i.e. -88.0) was set as the threshold. Molecules with scores lower than this threshold were assigned to the positive dataset, while those with higher scores were assigned to the negative dataset. The dataset was split into training (80%) and test (20%) subsets by stratifying over the class label. However, it is well known that docking score carries some intrinsic limitations related to the size of the molecules (Borchers et al., 2025). Specifically, larger molecules often get higher docking scores due to the capability to establish more interactions. To minimize this bias, a second, refined model was developed. In this second approach, only the top 1 million molecules and bottom 1 million molecules were considered, excluding the potential ambiguity of the molecules in the middle range. From these two groups, two balanced and stratified in terms of HAC distribution subsets were selected using RDKit. The aim of this procedure was to both exclude the molecules with an ambiguous docking score from the training of the model and to minimize the risk of overrepresenting very large molecules mitigating docking-related bias. For the sake of the study, the performances of both models were evaluated using Area Under the ROC Curve (AUC), precision and recall metrics. All together this metrics provide a comprehensive overview over model classification capability.

Furthermore, the second model underwent an additional validation step. Specifically, the approximately 770,000 molecules remaining from the stratified top 1 million, along with the approximately 770,000 molecules from the bottom 1 million, were re-evaluated via docking simulations and tested using the second model.

2.5 Virtual Screening

The virtual screening campaign was performed using the in-house-built ML model described above to explore an ultra-large chemical space. Specifically, 10 million molecules with HAC values between 15 and 36 were randomly sampled from the Enamine REAL database. For the sake of the analysis, all the 5 million molecules used during model training were excluded from the sampling procedure. The sampled molecules were retrieved in CXSMILES format and subsequently converted into the SMILES format using RDKit. Molecular descriptors were generated in the form of Morgan fingerprints using the python package RDKit with a vector length of 1024 bits and a radius of 2. Both models introduced in section 2.4 were applied to the dataset to compare their predictive performances. For the refined model (i.e. model 2), two different probability thresholds were tested: the default threshold of 0.5 and a stricter threshold of 0.8 to evaluate the effect of increased classification stringency on screening performance.

2.6 Molecular Docking as Validation

To further validate model prediction, 50.000 molecules were randomly sampled from each predicted probability interval of belonging to class 1 (from 0 to 1 with a step of 0.1) for a total of 550.000 molecules, and subsequentially docked within GPR120 binging site. The 3D structures of the molecules under investigation were generated using LigPrep tool implemented in Maestro generating molecules in .sdf format. The molecules were subsequently converted into .mol2 format using OpenBabel (O'Boyle et al., 2011) (version 3.1.1) before performing molecular docking with PLANTS 1.2, applying the same docking protocol described in paragraph 2.2 to ensure consistency between the analysis and the validation steps. The distribution of the docking scores within each predicted probability range was visualized using a violin plot.

3. Results

3.1 Model development and validation

For the sake of this study, the structural continuity of the human GPR120 was obtained through an AlphaFold2-based notebook using the structure available on PDB with code 8G59 as a template for the model. Specifically, AlphaFold2 is a modelling system based on ML combining neural networks and homology modelling that provides highly reliable 3D protein models (Yang et al., 2023). In this framework, the model scored a high confidence value with a pLDDT=80.8 which indicates a good backbone resolution as per manufacture declaration (https://alphafold.ebi.ac.uk/). Being our model trained on docking score over Morgan fingerprints, the success of virtual screening will considerably depend on the accuracy of the docking program used to predict the docking score and the binding architecture. Therefore, the reliability of the molecular docking procedure was accessed by checking the capability to reproduce the binding architecture of TUG-891 (i.e. a potent GPR120 agonist) comparing its docking pose with the architecture from a cryo-EM study of GPR120-TUG-891 complex within PDB structure 8G59. As shown in Figure 1, the docking procedure was able to reproduce the binding architecture of TUG-891 recording an average RMSD value of 0.58 Å and a $PLANTS_{PLP}$ of -109.14, confirming the soundness of the docking protocol. Once validated the docking procedure, the 3044 GPR120 ligands retrieved from ChEMBL were filtered to remove duplicates obtaining 1161 unique IDs. To prepare the dataset for model training and subsequent virtual screening of Enamine REAL compounds, Lipinski's Ro5 was applied, obtaining at least 731 molecules to further investigate through molecular docking. All 731 molecules were docked within the designed binding site of GPR120, and most of these recorded a very negative docking score, consistent with favorable predicted binding affinities. To develop a classification model able to discriminate reliably true negative from true positive compounds, a chemical space analysis of the hits recording a docking score lower than -80 was performed using DataWarrior. Interestingly, the analysis revealed that all the molecules recording a docking score lower than -80 were in the HAC range 15-36. Based on this observation, the dataset for model development was restricted to compounds within this HAC range. A total of 5 million were randomly selected from the Enamine REAL database only considering the packages including molecules with an HAC between 15 and 36 (see section 2.4). These molecules were subsequently docked within GPR120 binding site, the best docking score among the 10

generated poses was retained for each molecule, and the results were ranked from the lowest (most favorable) to the highest (less favorable). Once all the docking scores were collected and sorted, two ML models were trained and evaluated and the best performing one was optimized by adjusting threshold parameters.

Model 1. As described in Section 2.4, the first model was trained by splitting the dataset at the docking score corresponding to the molecule at position 2,500,000 (i.e. -88.00). Molecules with docking scores lower than -88.00 were assigned to the positive class, while all the molecules recording a docking score higher than -88.00 were assigned to the negative dataset. The model achieved a mean AUC of 0.906 and an average precision (AP) of 0.955 (Figure 3A) demonstrating strong predictive performance. However, this model was affected by one of the intrinsic limitations of the docking methods, the bias toward larger molecules (i.e., formed by a large number of atoms). Indeed, larger molecules often yield more negative scores regardless of their actual binding affinity but because they have more atoms and functional groups capable of forming multiple interaction with the binding site. According to this, the positive dataset may be unbalanced toward molecules with high HAC values. Although calculating ligand efficiency could normalize for size, it may not represent a solution prone to process smaller molecules and resulting in a positive dataset unbalanced toward molecules with low HAC values. These limitations motivated the design of a refined model.

Model 2. To overcome these issues, a second model was built using a stratified sampling approach according to HAC. More in detail, a stratified dataset from the first 1 million molecules of the dataset (the ones with the lowest docking scores) and a stratified dataset from the last 1 million molecules of the dataset (the ones with the highest docking score) were taken. This means that the model was trained purely on molecules belonging to the two opposite extremes in terms of docking scores, thus excluding all the possibly ambiguous intermediate values. This approach may ensure the classification of the molecules in a more rigorous way. According to the stratification criteria applied, 227,270 molecules were sorted from the first 1 million molecules of the dataset. A perfectly balanced dataset was then obtained by selecting an equal number (i.e. 227,270 molecules) of compounds from the last 1 million molecules (worst docking scores), again stratified by HAC. The procedure applied for this second model ensured that the positive and the negative datasets were balanced in terms of size as well as in terms of chemical space coverage. The model achieved a mean AUC of 0.989 and an AP of 0.989 (Figure 3B) proving its high predictive performance. To further validate this model, an additional test was performed using the approximately 770,000 molecules remaining from the stratification of the first 1 million molecules and the approximately 770,000 remaining from

the stratification of the last 1 million molecules (Figure 2). The model was able to classify as class 1 the 84% of the molecules of the first subset (molecules recording good docking scores) and to classify as class 0 the 90% of the molecules of the second subset (molecules recording bad docking scores). Overall, this model succeeded in discriminating between positive and negative ligands as well as in addressing docking methods internal biases.

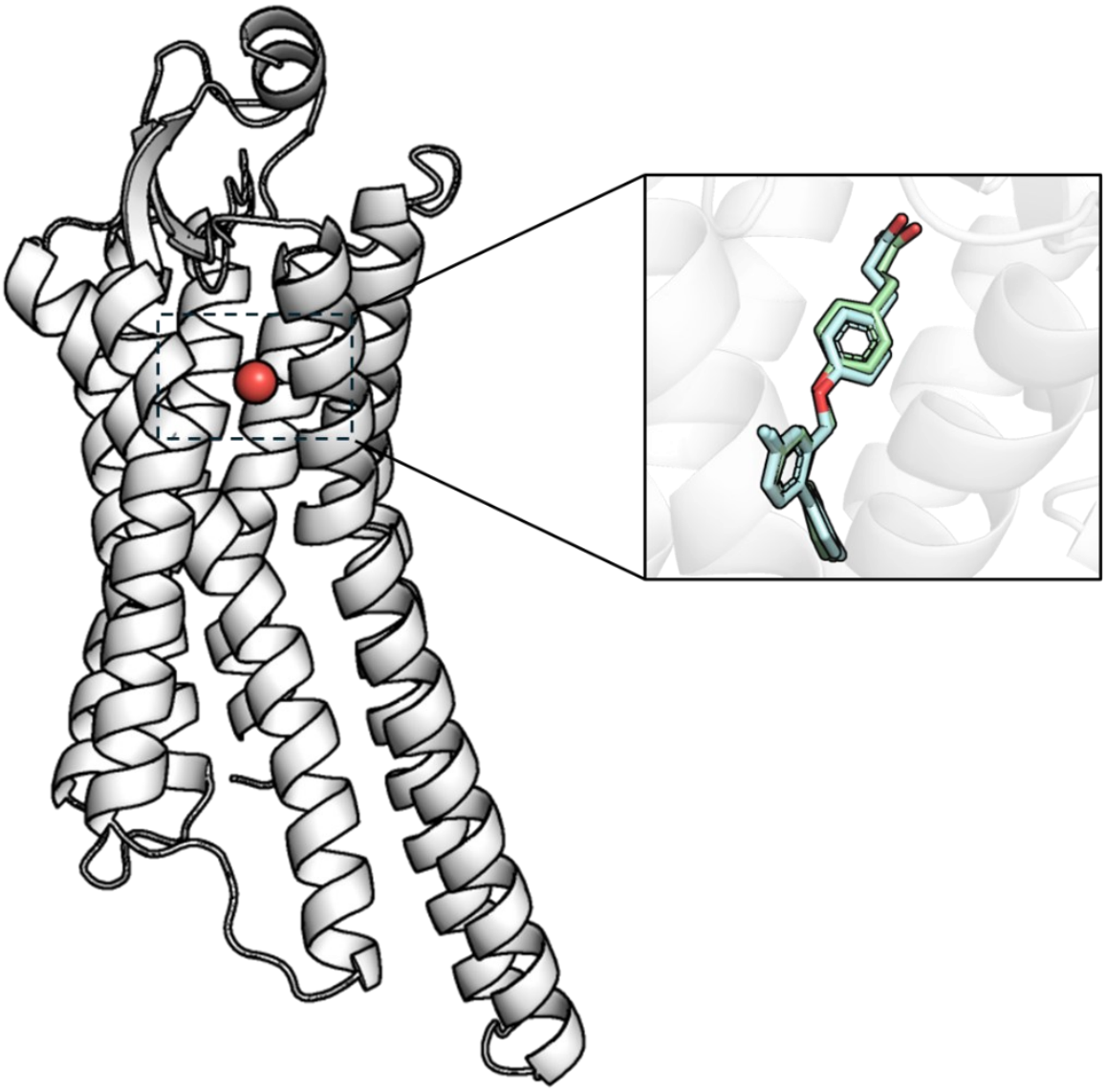


Figure 1. Validation of the docking procedure. GPR120 3D structure is represented in white cartoon, while the red sphere represents the centroid used in docking simulations. The box on the right represents TUG-891 (know GPR120 agonist) best-scored docking pose (cyan sticks) superimposed to the cryo-EM pose of TUG-891 retrieved from the PDB ID 8G59 (green sticks).

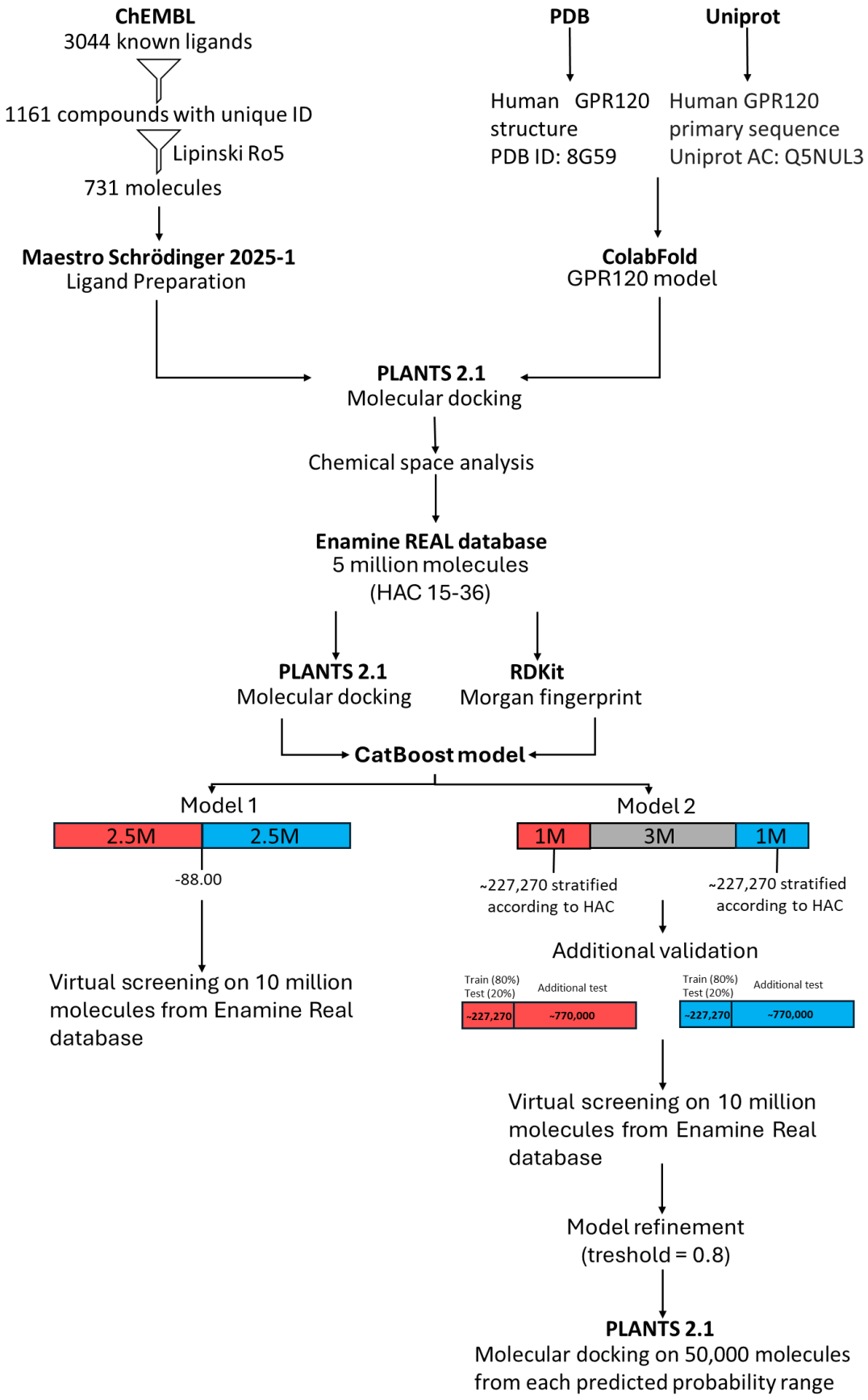


Figure 2. Workflow. Key steps of the pipeline developed and applied.

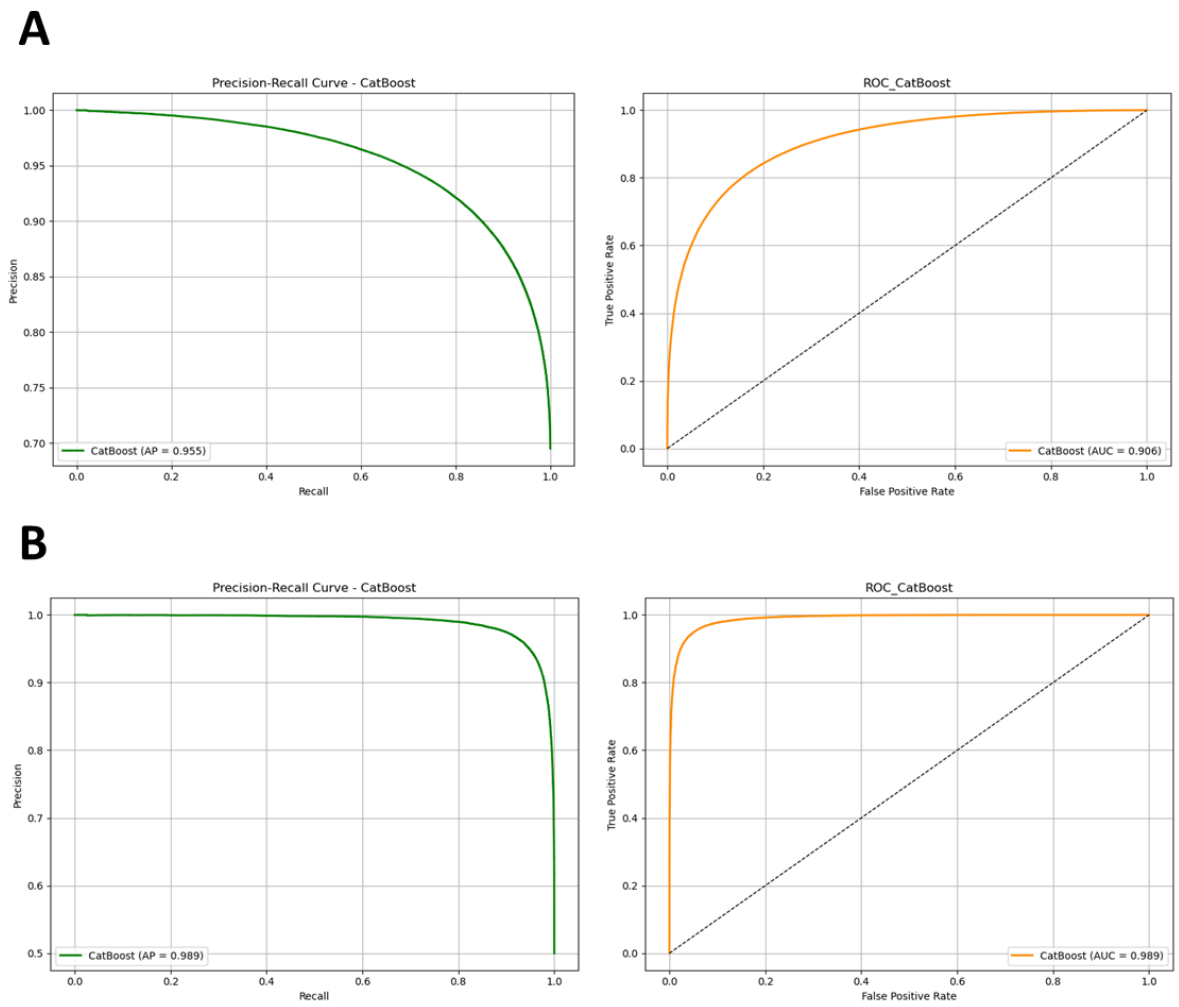


Figure 3. Model performances. **A.** Precision recall curve (left) and ROC (right) of the first model. **B.** Precision recall curve (left) and ROC (right) of the second model.

3.2 Virtual screening outcomes

As discussed above, the main goal of our work was to reduce the number of the molecules requiring docking. For the virtual screening analysis, 10 million molecules were retrieved from the Enamine real database in the 15-36 HAC range. Virtual screening was performed using both the models (Model 1 and 2; see Section 3.1 above) to check which was the best performing one. The first model predicted as class 1 the 63% of the molecules and as class 0 the remaining 37%. These results pointed to the fact that most of the molecules are identified as possible binders, maybe including also molecules that have not a so strong binding affinity. This may be because the model is trained on the whole dataset of docked molecules and because may be biased by molecular size. On the other hand, the second model predicted as class 1 the 47% of the molecules, while as class 0 the 53% (Figure 4). These outcomes pointed to the capability of the second model to reduce the number of molecules having a lower probability to be positive. Stratifying by HAC in fact helps to overcome the correlation between molecular size and docking score, allowing the model to learn true chemical features associated with binding, rather than features related to compound size. Considering the better and more reliable results obtained applying the second model, the following step was to improve the second model setting the probability threshold at 0.8 instead of the default one (i.e. 0.5). Applying this stricter threshold the model identified as class 0 the 68% and as class 1 the 32% of the molecules, reducing by almost 70% the number of the molecules to be further docked (Figure 4). Reducing the number of molecules to dock represents an advantage both in terms of time and costs.

Furthermore, the probability distribution of the ligands to belong to class 1 were calculated using the function predict_proba in scikit-learn. To further check the precision of the model in distinguishing between molecules able to dock and molecules not favourably docked, a randomly sampled subset of molecules from each probability range was selected and further docked. More in detail, 50.000 molecules were randomly sampled from each probability range (i.e. 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0) for a total of 550.00 molecules. Once docked, their docking scores were grouped into four representative ranges (−129 to −97, −97 to −88, −88 to −70, and −70 to 45) and the results were visualized using a violin plot which provides an overview of the relationship between the docking score and the predicted probabilities (Figure 5). As shown in Figure 5, docking analysis revealed the capability of the model to distinguish between molecules potentially able to bind the receptor to those unlikely to bind. The violin plot (Figure 5) highlighted that most of the molecules recording a docking score between -88 and -129 are in the highest part of the probability range, whereas the

molecules recording a docking score between -88 and 45 are predominantly distributed in the lowest probability ranges. More in detail, compounds exhibiting more favourable docking scores (−129 to −97) showed high predicted probabilities, mostly concentrated between 0.8 and 1.0, indicating strong confidence in their classification as active ligands. In the second range (−97 to −88), the probability distribution trend was toward higher values, though with a broader spread, reflecting moderate binding potential. The third interval (−88 to −70) exhibited a shift toward lower probability values, pointing to the increased uncertainty and suggesting a gradual transition between likely binders and non-binders. Finally, compounds belonging to the least favourable docking range (−70 to 45) showed very low predicted probabilities, with the majority clustering below 0.1, in agree with their low calculated binding affinity. Overall, results highlighted the fact that most of the molecules recording a docking score between -88 and -129 are in the highest part of the probability range, whereas the molecules recording a docking score between -88 and 45 are predominantly distributed in the lowest probability ranges. In fact, the wider sections indicate a higher concentration of data while the narrower sections indicate a lower concentration.

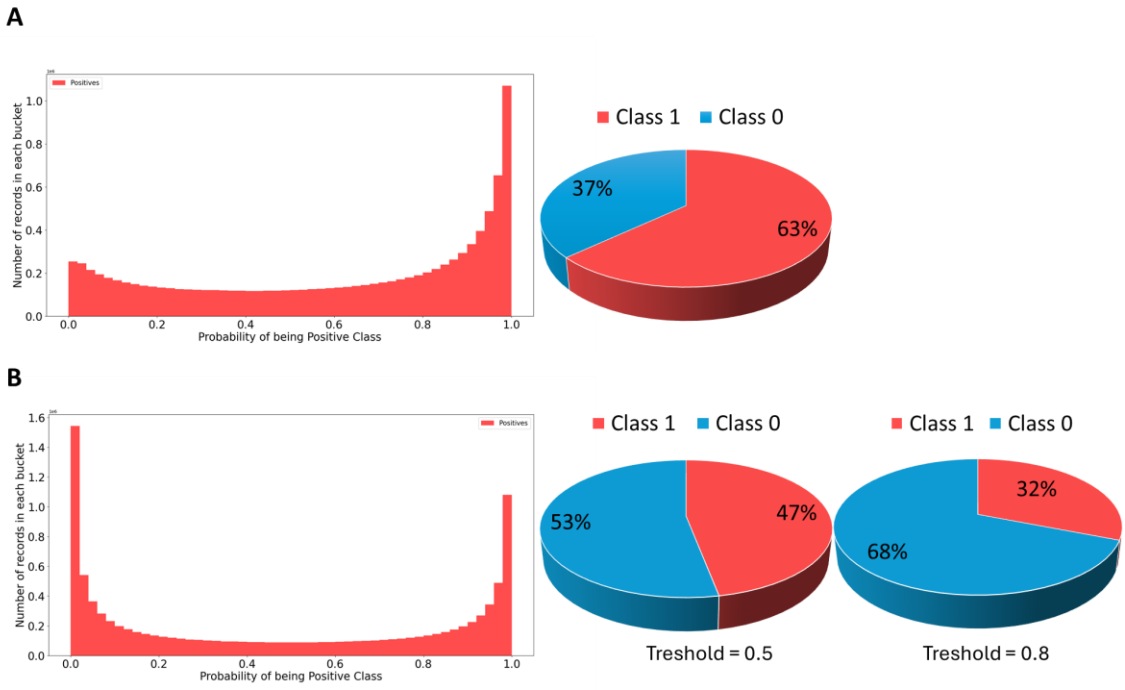


Figure 4. Probability distribution and virtual screening results. **A.** Probability distribution plot (left) of the virtual screening results obtained using model 1 and the corresponding pie chart showing the percentage of the molecules classified as positive and negative (right). **B.** Probability distribution plot (left) of the virtual screening results obtained using model 2 and

the corresponding pie chart showing the percentage of the molecules classified as positive and negative applying the default threshold (i.e. 0.5) and a stricter threshold (i.e. 0.8) (right).

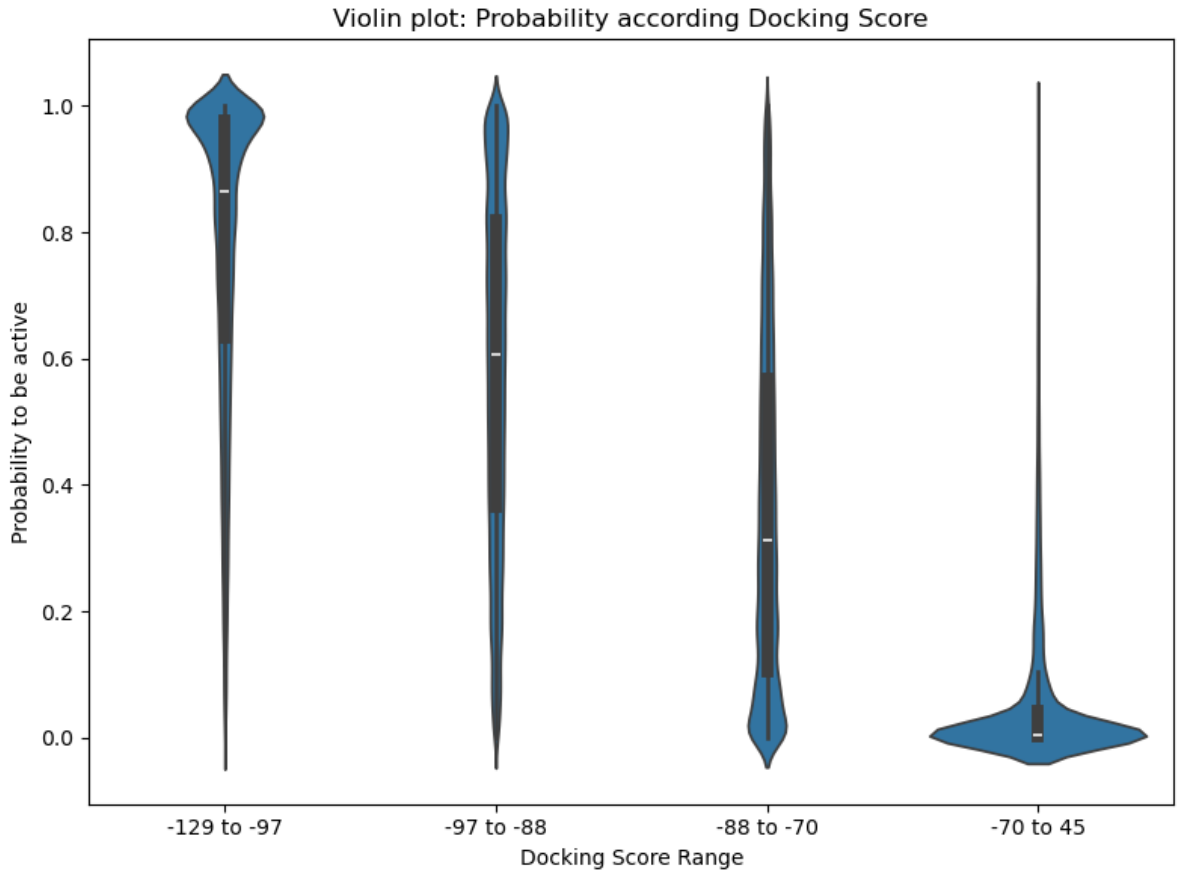


Figure 5. Violin plot showing the distribution of docking scores.

4. Discussion

In the context of empowering structure-based virtual screening with artificial intelligence techniques to identify new GPR120 ligands, ML models represent a valuable strategy to enhance efficiency and scalability (Adeshina et al., 2020) by combining the already proved reliability of molecular docking (Nicoli et al., 2023; Perugino et al., 2024a) with molecular descriptors such as Morgan fingerprints. The implementation of ML in the pipeline allowed the virtual screening of 10 million of compounds to prioritize molecules for further docking calculations. The comparison between the two models developed highlighted how dataset stratification according to molecular size (HAC) significantly improved the predictive capability of the model reducing the number of molecules labelled as class 1 (i.e. 32 %). Furthermore, the exclusion of intermediate docking score values considering only the first and the last 1 million molecules for the stratification reduced uncertainty in classification. On the other hand, the first model trained on unstratified dataset considering the docking score all the 5 million molecules docked tended to classify a larger percentage of molecules as class 1 (i.e. 63%) likely because of the intrinsic bias of docking scores toward compounds with a high number of atoms. Conversely, the model stratified according to HAC successfully overcome this inherent bias allowing the model to only focus on true structure-activity relationship rather than on molecular size of the molecules. This refinement together with the setting of a stricter probability threshold (i.e. 0.8 instead of 0.5 as per default setting) lead to a more accurate prediction of the true positive compounds, reducing the number of molecules that are unlikely to interact satisfactorily with GPR120. In fact, setting the threshold at 0.8 the model was able to discard almost 70% of the molecules before performing molecular docking analysis. This results in a substantial reduction in computational costs and calculation time. The analysis of the predicted probability distributions performed on 550.000 molecules corroborated the model's capability in ranking compounds according to their potential binding affinity. Notably, the violin plot showed that the distribution of the predicted probabilities aligns with docking score calculation. In fact, the investigation about the distribution of the compounds across different predicted probability ranges, based on their respective docking score ranges, revealed a correlation between these two variables. More in detail, the lower the docking score, the higher the distribution in the highest predicted probability ranges, while the higher the docking score, the higher the distribution in the lowest predicted probability ranges. The observed capability of the model in reliably prioritizing promising candidates for further docking

simulation highlighted the pivotal role of ML in representing an advantage when implemented in virtual screening pipelines serving as prefilter. Interestingly, this approach fully aligns with the NAMs, which promote computational approaches to increase the throughput in chemical assessment and drug discovery (Edwards et al., 2025). Beyond its application to GPR120, the developed workflow might be generalized and applied to other GPCRs or other targets requiring the virtual screening of ultra-large chemical libraries. Notably, the model was not trained using an arbitrary docking score cutoff defined based on prior knowledge or target-specific threshold. In fact, the classification threshold was established as the highest docking score within the stratified positive dataset. This data-driven approach makes the ML model independent of target-specific score values, increasing its adaptability to other molecular targets. The aim of this work was to screen a large library of compounds (i.e. Enamine REAL databased) including all molecules falling within the chemical space covered by molecules known to interact with GPR120. Consequently, the main limitation of this approach lies in the need to consider molecules covering a very broad chemical space both for model building and screening. This requirement resulted in a still high computational demand, as a large number of molecules had to be docked to build the model. Specifically, the dataset was designed to include all the molecules falling within the HAC range of those experimentally known to interact with GPR120. In future applications, this limitation could be overcome by restricting the chemical space to the range where the most active compounds are found, thereby reducing the number of molecules to be docked for model development. Moreover, this aspect is particularly target-dependent. In fact, GPR120 is known to interact with a wide variety of molecules in terms of MW, LogP and HAC, while other proteins may show a narrower selectivity toward ligands allowing the construction of the model with a more confined dataset requiring less molecules to be initially docked.

5. Conclusion

GPR120 is a relevant target both from food science and pharmaceutical point of view being involved in the free fatty acid perception as well as in a plethora of physiological outcomes. The present work developed and validated a ML based model specifically designed to empower the prediction of potential GPR120 binders by reducing the number of molecules to analyze via docking simulations. First, the refinement of 3D structure of human GPR120 was ensured by integrating structural information with AlphaFold to obtain a reliable and complete structure. The powerful predictive model CatBoost was trained on docking score and Morgan fingerprints computed from 5 million compounds sampled by the Enamine REAL database. Two models were proposed to overcome docking inherent bias related to molecular size. The second model developed using HAC-stratified balanced datasets succeeded in properly classifying 84% of high-affinity molecules and 90% of low-affinity molecules. The virtual screening campaign performed on 10 million molecules from the Enamine database proved the benefits of the model. Applying the second model setting a stricter probability threshold (0.8), the number of the molecules to be further docked was reduced by approximately 70%. This reduction translates into substantial savings in computational resources and time, highlighting the model's efficiency as a prefiltering tool in large-scale drug discovery campaigns. Finally, a validation on 550.000 compounds belonging to different probability ranges corroborated the capability of the model in correlating predicted probability with actual binding affinities. The outcomes collected in this work highlighted how ML may empower and accelerate virtual screening, reducing both time and computational efforts. Overall, this work may provide a generalizable pipeline that can be extended to other GPCRs or to other molecular targets of interest. Indeed, the developed model might be adapted to screen combinatorial libraries as well as large commercially available collections and in-house collections of compounds relevant both from a pharmacological and/or a food science perspective.

References

Adeshina, Y. O., Deeds, E. J., & Karanicolas, J. (2020). Machine learning classification can reduce false positives in structure-based virtual screening. *Proc Natl Acad Sci U S A, 117*(31), 18477-18488. doi:10.1073/pnas.2000585117

Ahn, J. M., Kim, J., & Kim, K. (2023). Ensemble Machine Learning of Gradient Boosting (XGBoost, LightGBM, CatBoost) and Attention-Based CNN-LSTM for Harmful Algal Blooms Forecasting. *Toxins (Basel), 15*(10). doi:10.3390/toxins15100608

Bai, H., Cheng, Y., Jia, S., Wang, X., Jin, R., Guo, H., . . . Wang, Y. (2025). Machine learning-based QSAR and structure-based virtual screening guided discovery of novel mIDH1 inhibitors from natural products. *J Comput Aided Mol Des, 39*(1), 44. doi:10.1007/s10822-025-00624-1

Borchers, J., Perugino, F., Schlosser, A., Lamer, S., Lutz, L., Pedroni, L., . . . Mally, A. (2025). Identification of small GTPases as potential target proteins of the mycotoxin and renal carcinogen ochratoxin A. *Arch Toxicol*. doi:10.1007/s00204-025-04189-8

Carlsson, J., & Luttens, A. (2024). Structure-based virtual screening of vast chemical space as a starting point for drug discovery. *Curr Opin Struct Biol, 87*, 102829. doi:10.1016/j.sbi.2024.102829

Carullo, G., Mazzotta, S., Vega-Holm, M., Iglesias-Guerra, F., Vega-Perez, J. M., Aiello, F., & Brizzi, A. (2021). GPR120/FFAR4 Pharmacology: Focus on Agonists in Type 2 Diabetes Mellitus Drug Discovery. *J Med Chem, 64*(8), 4312-4332. doi:10.1021/acs.jmedchem.0c01002

Civciristov, S., Ellisdon, A. M., Suderman, R., Pon, C. K., Evans, B. A., Kleifeld, O., . . . Halls, M. L. (2018). Preassembled GPCR signaling complexes mediate distinct cellular responses to ultralow ligand concentrations. *Sci Signal, 11*(551). doi:10.1126/scisignal.aan1188

Cornall, L. M., Mathai, M. L., Hryciw, D. H., & McAinch, A. J. (2014). GPR120 agonism as a countermeasure against metabolic diseases. *Drug Discov Today, 19*(5), 670-679. doi:10.1016/j.drudis.2013.11.021

Di Pizio, A., & Nicoli, A. (2020). In Silico Molecular Study of Tryptophan Bitterness. *Molecules, 25*(20). doi:10.3390/molecules25204623

Edwards, M., Blanquie, O., Moriarty, O., Beken, S., & Ehmann, F. (2025). New approach methodologies: EU regulatory horizons. *Nat Rev Drug Discov, 24*(8), 571-572. doi:10.1038/d41573-025-00053-7

Galindo, M. M., Voigt, N., Stein, J., van Lengerich, J., Raguse, J. D., Hofmann, T., . . . Behrens, M. (2012). G protein-coupled receptors in human fat taste perception. *Chem Senses, 37*(2), 123-139. doi:10.1093/chemse/bjr069

Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., . . . Overington, J. P. (2012). ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res, 40*(Database issue), D1100-1107. doi:10.1093/nar/gkr777

Ghislat, G., Rahman, T., & Ballester, P. J. (2021). Recent progress on the prospective application of machine learning to structure-based virtual screening. *Curr Opin Chem Biol, 65*, 28-34. doi:10.1016/j.cbpa.2021.04.009

Gorgulla, C., Cinaroglu, S. S., Fischer, P. D., Fackeldey, K., Wagner, G., & Arthanari, H. (2021). VirtualFlow Ants-Ultra-Large Virtual Screenings with Artificial Intelligence Driven Docking Algorithm Based on Ant Colony Optimization. *Int J Mol Sci, 22*(11). doi:10.3390/ijms22115807

Hancock, J. T., & Khoshgoftaar, T. M. (2020). CatBoost for big data: an interdisciplinary review. *J Big Data, 7*(1), 94. doi:10.1186/s40537-020-00369-8

Hasan, A. U., Ohmori, K., Hashimoto, T., Kamitori, K., Yamaguchi, F., Noma, T., . . . Kohno, M. (2017). GPR120 in adipocytes has differential roles in the production of pro-inflammatory adipocytokines. *Biochem Biophys Res Commun, 486*(1), 76-82. doi:10.1016/j.bbrc.2017.03.001

Hirasawa, A., Tsumaya, K., Awaji, T., Katsuma, S., Adachi, T., Yamada, M., . . . Tsujimoto, G. (2005). Free fatty acids regulate gut incretin glucagon-like peptide-1 secretion through GPR120. *Nat Med, 11*(1), 90-94. doi:10.1038/nm1168

Hudson, B. D., Shimpukade, B., Mackenzie, A. E., Butcher, A. J., Pediani, J. D., Christiansen, E., . . . Milligan, G. (2013). The pharmacology of TUG-891, a potent and selective agonist of the free fatty acid receptor 4 (FFA4/GPR120), demonstrates both potential opportunity and possible challenges to therapeutic agonism. *Mol Pharmacol, 84*(5), 710-725. doi:10.1124/mol.113.087783

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., . . . Bolton, E. E. (2025). PubChem 2025 update. *Nucleic Acids Res, 53*(D1), D1516-D1525. doi:10.1093/nar/gkae1059

Korb, O., Stutzle, T., & Exner, T. E. (2009). Empirical scoring functions for advanced protein-ligand docking with PLANTS. *J Chem Inf Model, 49*(1), 84-96. doi:10.1021/ci800298z

Korb, O., Stützle, T., & Exner, T. E. (2007). An ant colony optimization approach to flexible protein–ligand docking. *Swarm Intelligence, 1*(2), 115-134. doi:10.1007/s11721-007-0006-9

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews, 23*(1-3), 3-25. doi:10.1016/s0169-409x(96)00423-1

Liu, H. D., Wang, W. B., Xu, Z. G., Liu, C. H., He, D. F., Du, L. P., . . . Sun, J. P. (2015). FFA4 receptor (GPR120): A hot target for the development of anti-diabetic therapies. *Eur J Pharmacol, 763*(Pt B), 160-168. doi:10.1016/j.ejphar.2015.06.028

Lopez-Lopez, E., Naveja, J. J., & Medina-Franco, J. L. (2019). DataWarrior: an evaluation of the open-source drug discovery tool. *Expert Opin Drug Discov, 14*(4), 335-341. doi:10.1080/17460441.2019.1581170

Luttens, A., Cabeza de Vaca, I., Sparring, L., Brea, J., Martinez, A. L., Kahlous, N. A., . . . Carlsson, J. (2025). Rapid traversal of vast chemical space using machine learning-guided docking screens. *Nat Comput Sci, 5*(4), 301-312. doi:10.1038/s43588-025-00777-x

Mao, C., Xiao, P., Tao, X. N., Qin, J., He, Q. T., Zhang, C., . . . Sun, J. P. (2023). Unsaturated bond recognition leads to biased signal in a fatty acid receptor. *Science, 380*(6640), eadd6220. doi:10.1126/science.add6220

Milligan, G., Alvarez-Curto, E., Hudson, B. D., Prihandoko, R., & Tobin, A. B. (2017). FFA4/GPR120: Pharmacology and Therapeutic Opportunities. *Trends Pharmacol Sci, 38*(9), 809-821. doi:10.1016/j.tips.2017.06.006

Moniri, N. H. (2016). Free-fatty acid receptor-4 (GPR120): Cellular and molecular function and its role in metabolic disorders. *Biochem Pharmacol, 110-111*, 1-15. doi:10.1016/j.bcp.2016.01.021

Nicoli, A., Haag, F., Marcinek, P., He, R., Kreissl, J., Stein, J., . . . Di Pizio, A. (2023). Modeling the Orthosteric Binding Site of the G Protein-Coupled Odorant Receptor OR5K1. *J Chem Inf Model, 63*(7), 2014-2029. doi:10.1021/acs.jcim.2c00752

O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. *J Cheminform, 3*, 33. doi:10.1186/1758-2946-3-33

Pal, A., Curtin, J. F., & Kinsella, G. K. (2021). Structure based prediction of a novel GPR120 antagonist based on pharmacophore screening and molecular dynamics simulations. *Comput Struct Biotechnol J, 19*, 6050-6063. doi:10.1016/j.csbj.2021.11.005

Pedroni, L., Louisse, J., Punt, A., Dorne, J., Dall'Asta, C., & Dellafiora, L. (2023). A Computational Inter-Species Study on Safrole Phase I Metabolism-Dependent Bioactivation: A Mechanistic Insight into the Study of Possible Differences among Species. *Toxins (Basel), 15*(2). doi:10.3390/toxins15020094

Pedroni, L., Perugino, F., Kurtaga, A., Galaverna, G., Dall'Asta, C., & Dellafiora, L. (2023). The bitter side of toxicity: A big data analysis spotted the interaction between trichothecenes and bitter receptors. *Food Res Int, 173*(Pt 1), 113284. doi:10.1016/j.foodres.2023.113284

Pedroni, L., Perugino, F., Magnaghi, F., Dall'Asta, C., Galaverna, G., & Dellafiora, L. (2024). Free fatty acid receptors beyond fatty acids: A computational journey to explore peptides as possible binders of GPR120. *Curr Res Food Sci, 8*, 100710. doi:10.1016/j.crfs.2024.100710

Perugino, F., Pedroni, L., Galaverna, G., Dall'Asta, C., & Dellafiora, L. (2024a). A mechanistic toxicology study to grasp the mechanics of zearalenone estrogenicity: Spotlighting aromatase and the effects of its genetic variability. *Toxicology, 501*, 153686. doi:10.1016/j.tox.2023.153686

Perugino, F., Pedroni, L., Galaverna, G., Dall'Asta, C., & Dellafiora, L. (2024b). Virtual display of targets: A new level to rise the current understanding of ochratoxin A toxicity from a molecular standpoint. *Toxicology, 503*, 153765. doi:10.1016/j.tox.2024.153765

Prokhorenkova, L., Gusev, A.G., Vorobev, A., Dorogush, A.V., Gulin, A. (2018). *CatBoost: unbiased boosting with categorical features*. Paper presented at the Proceedings of the 32nd International Conference on Neural Information Processing Systems, Montréal, Canada.

Yang, Z., Zeng, X., Zhao, Y., & Chen, R. (2023). AlphaFold2 and its applications in the fields of biology and medicine. *Signal Transduct Target Ther, 8*(1), 115. doi:10.1038/s41392-023-01381-z

Z. Liao, R. Y., X. Huang, X. Yao, T. Huang and S. Zhu. (2019). DeepDock: Enhancing Ligand-protein Interaction Prediction by a Combination of Ligand and Structure Information. *IEEE International Conference on Bioinformatics and Biomedicine (BIBM), San Diego, CA, USA, 2019*, 311-317. doi:10.1109/BIBM47256.2019.8983365

Zhao, Y., Zha, D., Wang, L., Qiao, L., Lu, L., Mei, L., . . . Qiu, J. (2013). Phenotypic characterization of GPR120-expressing cells in the interstitial tissue of pancreas. *Tissue Cell, 45*(6), 421-427. doi:10.1016/j.tice.2013.07005

Chapter IX

General discussion and perspectives

Chapter IX

Main findings of this PhD thesis

This PhD thesis aimed to integrate AI techniques within computational approaches to investigate the interactions between macromolecules (i.e., proteins) and small molecules of food interest. *In silico* approaches based on virtual screening, molecular docking and molecular dynamics simulations have already been proved to be valuable tools to investigate the interaction between food-related compounds and respective molecular targets. Through the development and application of advanced *in silico* pipelines, this work explored a broad spectrum of molecules with a focus on taste-active, bioactive, and toxic compounds (especially mycotoxins). Despite their apparent diversity, these chemicals share a fundamental aspect: their biological activity is determined by molecular recognition events occurring between ligands and specific protein targets. This link led to the development of the integrative approach adopted throughout this thesis. The growing of food-related databases and make-on-demand databases led to the conditions for applying AI-based approaches in this field. ML and DL techniques, in fact, have already revolutionized drug discovery by reducing costs and computational effort, improving prediction accuracy, and accelerating candidate screening. The integration of AI tools into traditional structure-based and ligand-based virtual screening workflows has provided new strategies to accelerate this investigation. The same improvement may be applied also to the study of food-related compounds and macromolecules and associated effects on human health. A major achievement of this thesis lies in demonstrating how computational methods, when coupled and integrated with AI techniques, can elucidate the inner structural and functional bases of complex molecular interactions. Across the different chapters, this approach proved effective in addressing challenges moving from toxicology, bioactivity and taste perception, thereby contributing to a more integrated and comprehensive understanding of food compounds' biological behavior. The diverse computational workflows developed in this thesis combine both structure-based and ligand-based approaches, integrating AI tools for ligands screening and protein structure prediction. This integration enables a multiscale analysis of molecular systems from the identification of potential ligands using ML-assisted virtual screening, to the refinement of ligand–protein complexes through molecular dynamics and umbrella sampling simulations. The resulting framework represents an efficient and widely reproducible strategy for exploring large molecular datasets and predicting functional outcomes relevant to food quality and safety. The studies presented in Chapters 2–6 focused primarily on toxic compounds, particularly mycotoxins, which represent a major food safety concern. By examining their potential interactions with human proteins, these chapters provided novel

insights into the mechanistic aspects of mycotoxin toxicity and their possible adverse biological effects. In Chapter 2 the discovery of some trichothecenes (i.e. DON and 15-Ac-DON) as potential ligands of TAS2R46 suggests an interesting overlap between food toxicity and taste perception pathways, reinforcing the hypothesis that bitter taste receptors may play a broader role in physiological responses beyond gustation. This outcome aligns with the presence of TASRs not only in the lingual gustatory epithelia but also in other organs and tissues. Furthermore, it could provide a partial mechanistic explanation for DON toxicity. In fact, the capability of DON and 15-Ac-DON to interact with TAS2R46 highlighted the need to a broad investigation of the capability of trichothecenes to activate bitter receptors, especially those expressed in the extra-oral tissues, such as in gut where trichothecenes toxicity has been described. These outcomes may represent the starting point for an in-depth investigation of their mechanism of action.

Keeping working on mycotoxins, in Chapter 3 the investigation of ZEN and α-ZEL interactions with aromatase variants unveiled how a single genetic polymorphism can modulate toxic responses, emphasizing the importance of personalized toxicology in food risk assessment. Specifically, T310S variant was predicted to have an enhanced susceptibility to ZEN and αZEL inhibition. On the other hand, D309G and S478F have been described as potentially inactive aromatase variants.

Chapters 4 and 5 were focused on OTA, its thermal degradation product 2-R'-OTA and its metabolites (i.e. OTα, OTB, OTQ and OTHQ). In Chapter 4, the *in silico* pipeline developed pointed out OGFOD1 as potential target of OTA and 2-R'-OTA. Being this enzyme involved in protein biosynthesis, the hypothesized inhibition by OTA and its diasteromer may result in protein biosynthesis dysregulation. OGFOD1 inhibition may fill some gaps of OTA toxicology, and further investigation toward a more informed description of OTA adverse outcome pathway should be prioritized. In Chapter 5 the capability of OTA and its main metabolites (i.e. OTalpha, OTB, OTH and OTHQ) to interact with enzymes involved in Parkinson's AOP was investigated to elucidate the experimentally proposed effect of OTA on Pankinson's onset. Eventually, both lysosomal (i.e. cathepsins) and mitochondrial (i.e. Complex I) enzymes were found potentially inhibited by OTA and congeners thus proving to be connected to neurogenerative diseases such as Parkinson. The enzymes under investigation (i.e. cathepsins and Complex I) revealed as new potential molecular targets and mechanistic links to neurodegenerative diseases such as Parkinson's Disease and protein biosynthesis damage. In fact, OTA and OTQ may inhibit Complex I possibly disrupting mitochondrial function. Concerning cathepsins OTA and congeners show binding to cathepsins D and L, but not B. These findings highlighted how *in*

silico analyses can complement experimental toxicology by identifying previously unrecognized protein targets and serving as a guide for further *in vitro* and *in vivo* investigations. Chapter 6 addressed the dual behavior of fumonisins (i.e. FB1 and HFB1), providing a molecular rationale for their ability to act both as substrates and inhibitors of ceramide synthases. This work highlighted how molecular dynamics simulations can dissect the dynamic complexity of enzyme–toxin interactions and support the understanding of structure–function relationships that drive toxicity. Beyond toxic compounds, in Chapters 7 and 8, the focus shifted toward bioactive compounds keeping the interest on targets involved in test perception (i.e. GPR120). The identification of GPR120-binding peptides (in Chapter 7) opened new perspectives for the discovery of natural bioactive sequences encrypted within food proteins that may contribute to metabolic health through GPCR modulation. Furthermore, this work also investigated the effect of the substitution of L-amino acids with D-amino acids. In the last work, the development of a ML model (CatBoost) to classify GPR120 binders (Chapter 8) demonstrated how AI can significantly enhance the efficiency of virtual screening workflows, reduce computational costs and facilitate high-throughput ligand prioritization. This methodological advancement represents an important step toward data-driven food bioactivity discovery, paving the way for AI-assisted nutraceutical and functional food design. Altogether, the results of this thesis highlight the potential of combining AI methodologies with computational chemistry in the field of food chemistry and toxicology to promote a deeper understanding of molecular interactions at the food–biological systems. Beyond their scientific implications, these findings emphasize the value of *in silico* tools in supporting regulatory toxicology, food safety assessment, and nutritional innovation. The use of AI-enhanced modeling can not only streamline research pipelines but also contribute to predictive risk assessment, compound prioritization, and mechanistic elucidation, thereby aligning with the 3Rs principles (Replacement, Reduction, and Refinement) and NAMs use by minimizing the need for animal testing.

Chapter IX

Conclusion and Future Perspectives

This PhD thesis aimed to integrate AI into computational approaches to investigate the interactions between macromolecules (i.e. proteins) and small molecules of food interest. AI-based tools, in fact, represent an improving strategy for gold-standard computational methods such as virtual screening, molecular docking and molecular dynamics simulations. By leveraging on the integrated use of AI, this work tried to find a connection between taste-active compounds, toxic compounds and bioactive compounds and their respective target, aiming to shed light on the potential crosstalk between molecular mechanisms apparently unrelated (e.g. the interaction between mycotoxins and TASRs). This overview highlighted how the boundaries between sensory perception, health-promoting effects, and toxicity are often labile and how computational tools can elucidate this complex relation. From a methodological point of view, this research tried to integrate AI in the traditional computational workflows starting from ligand based virtual screening, moving to protein structure prediction till the development of ML model for structure-based virtual screening. Looking ahead, the approaches developed in this thesis could be further extended and generalized to other targets relevant from a food and/or a pharmaceutical perspective. Furthermore, the integration of increasingly advanced AI tools within the proposed computational framework will make this strategy even more powerful and adaptable. In this context, the continuous evolution of AI-driven computational tools would provide new opportunities for the discovery of bioactive compounds, the assessment of food safety, and the design of novel functional ingredients, eventually and significantly contributing to the progress of a safer and more sustainable food system.

Supplementary material

Chapter II

Table S1. List of PubChem CIDs of molecules included in true positive or true negative set

True positive set
100187; 10052; 10133; 101684; 101799; 10251; 10331; 10403; 1048; 10511; 10512; 10518; 10535; 10680; 10709; 10724; 107905; 107971; 10983; 10994; 11020; 11047; 11095; 11107; 11128; 11175902; 1119; 1130; 115149; 115161; 115221; 11820; 119093; 11920271; 11992452; 120678; 12309272; 124021; 12560; 1269845; 128861; 130796; 134906; 1349907; 136419; 13993178; 14109402; 14217; 14362; 145659; 149096; 14969; 14989; 15063; 151004; 1549517; 159055; 160467; 161120; 165208; 165521; 179651; 1803558; 181128; 182362; 185635; 18735; 188289; 1889; 19094; 19649; 196915; 197754; 1983; 19908724; 20835504; 2153; 2164; 2180; 21852; 2194; 2206; 221071; 221493; 222156; 222994; 2230; 2236; 226036; 2265; 22833654; 229016; 2294; 23205; 2331; 235284; 2353; 238782; 246330; 2481; 24832100; 2483; 25021769; 2519; 25200812; 25210; 2576; 259327; 259803; 26937; 2707; 2719; 2723593; 2723650; 2723704; 2725; 275196; 2757; 2764; 27686; 27868; 283906; 28718; 2883177; 29029; 2955; 2972; 3032338; 3033; 3034034; 3034683; 3055; 3082543; 3085401; 3100; 31072; 3108; 313325; 31340; 31401; 31553; 3163205; 3182; 3229; 3242; 326240; 329275; 3321199; 3325; 3371; 338777; 3476; 3516; 354616; 3559; 36145; 3625947; 36573; 36688181; 370; 371257; 373677; 3778; 3819217; 3828; 3938245; 3948; 4004; 4044; 4059; 4064; 4091; 4095; 4114; 413112; 413117; 4162; 4189; 433737; 435597; 439246; 439503; 439533; 440936; 441071; 441293; 441734; 442009; 442011; 44201359; 442138; 442195; 442291; 442292; 442369; 442428; 442435; 442506; 442538; 4427709; 445154; 4472; 4488; 455264; 4583; 4594; 4601; 4619; 4679; 4680; 471002; 4740; 4754; 4770; 4780; 4788; 4806; 4831; 4848; 4913; 494227; 4998; 5088211; 509245; 513197; 51397980; 5147; 5161; 5270542; 5273568; 5280343; 5280373; 5280378; 5280443; 5280445; 5280544; 5280647; 5280666; 5280759; 5280863; 5280906; 5280961; 5281222; 5281377; 5281404; 5281426; 5281428; 5281435; 5281544; 5281607; 5281610; 5281611; 5281614; 5281617; 5281654; 5281670; 5281672; 5281680; 5281697; 5281701; 5281707; 5281708; 5281804; 5281955; 5282073; 5282136; 5282197; 5282362; 5284627; 5284639; 5284648; 5284649; 5315831; 5317284; 5317309; 5317750; 5318517; 5322064; 5322065; 5326; 5329; 5330; 5344; 5353545; 5353547; 5354883; 5355853; 5357070; 5359422; 5360696; 5362830; 5362; 5370611; 5376891; 5377381; 5388962; 5429; 5447130; 5462438; 5463863; 5464155; 5466139; 54676038; 5497; 5531; 5576; 5578; 566062; 5702160; 5743; 5754; 5755; 57697; 57965774; 5811533; 5827; 5838; 5842; 5865; 5917; 5959; 5980988; 6041; 6072; 6106; 6140; 6167; 6197; 624708; 6274; 6275271; 6305; 6306; 636760; 637760; 637776; 638023; 638278; 639665; 6426944; 6427001; 6433476; 6433; 6436032; 6436239; 6436532; 6437604; 6452133; 6454; 6463; 6464; 6465; 6473881; 6473; 6489; 65030; 65064; 65080; 65084; 65243; 6526396; 653; 65571; 656566; 657298; 657311; 6603838; 66160; 6651; 667450; 667477; 667493; 6674; 667581; 667687; 67111; 67145; 676294; 676306; 676307; 676454; 6781; 68051; 68066; 68071; 68077; 68112; 68214; 68415; 68437; 6847; 6849; 68843; 688660; 6911854; 6954; 69794; 700999; 7020; 7028; 7098673; 71105; 71113; 7144; 71542; 7196; 72276; 72277; 72281; 72304; 72320; 72344; 72625; 72964; 7330; 73404; 73415; 73467; 73635; 736366; 7605278; 76808; 768; 784; 7896; 79730; 79881; 8085; 8105; 820777; 8274; 8275; 84029; 8486; 8549; 8593; 8631; 88881; 89421; 89594; 8969; 9015; 90454; 9064; 91461; 91469; 9210; 92141; 92794; 92843; 92953; 92972; 92990; 932; 9475; 9552079; 9568512; 9576240; 9601691; 9603; 9649; 969463; 98046; 99924; 99967
True negative set
10000891; 100081134; 100096; 10012435; 10012587; 10012788; 10012825; 10014246; 10015662; 10019127; 10019324; 10020702; 10021651; 10022591; 10024256; 10024783; 100299; 10035228; 10036635; 10040569; 10041192; 10043095; 10044040; 10045262; 10045406; 10058450; 10060228; 10063024; 10063131; 10063853; 10064084; 10064819; 10064930; 10065723; 10068255; 10070482; 10081060; 10082256; 10082400; 10083634; 10084048; 10084837; 10085909; 10086485; 10086837; 10088795; 10090177; 10090304; 100922253; 100924013; 100924512; 100924513; 10092855; 100935579; 100938761; 100938762; 100938763; 100938764; 100938765; 100938966; 100943801; 100944092; 100944093; 100954261; 100956296; 100956297; 100958904; 100958911; 100964344; 100966758; 100974136; 100974137; 100974138; 100974139; 100974140; 100974141; 100974142; 100974143; 100974147; 100974148; 100974149; 100974150; 100974151; 100974152; 100974153; 100984390; 100984391; 100984392; 100984394; 100984395; 100984396; 100986593; 100987892; 100987893; 100990071; 100993160; 100994798; 100997096; 101001971; 101002408; 101003; 101009545; 101011400; 101015849; 101021685; 101029521; 101033892; 10103794; 10104030; 101041457; 101041458; 101041460; 101041461; 10104370; 101044994; 10104537; 101046110; 101047272; 101053275; 101053276; 101053277; 101053408; 10105599; 101056970; 101060730; 101060808; 10106330; 101065574; 101065575; 101067596; 101072096; 101072271; 101072516; 101072517; 101072518; 101072519; 101072520; 101072521; 101072522; 101072523; 10107920; 101088404; 101088405; 101088853; 101090530; 101090990; 101090991; 101091381; 101091382; 101094129; 101094442; 101094443; 101100886; 101100887; 101100888; 101100889; 101100890; 101100891; 101104340; 101115127; 101115128; 101115129; 101116747; 101117002; 101117003; 101117929; 101117930; 101117931; 101117932; 101121603; 101121604; 101121605; 101122762; 101128134; 101129024; 101129025; 101129026; 101129027; 101130014; 101131699; 101134371; 101135147; 101140985; 101152994; 101159096; 101165429; 101165430; 101165433; 101165434; 101168080; 101169526; 101169882; 101169883; 101171802; 101171803; 101171804; 101171805; 101172687; 101173366; 101177019; 101177677; 101177678; 101178328; 101178392; 101181250; 101183729; 101183730; 101187319; 101188260; 101188860; 101188952; 101188953; 101189813; 101190002; 101190003; 101190004; 101196421; 101196422; 101199388; 101199389; 101199390; 101199930; 101204785; 101206115; 101206116; 101206117; 101206118; 101206119; 101206122; 101206142; 101207250; 101210431; 101212746; 101213958; 101215970; 101220675; 101220676; 101220677; 101220678; 101220679; 101220680; 101220683; 101220685; 101220687; 101220688; 101220689; 101220690; 101220691; 101220692; 101220693; 101244915; 101244916; 101244917; 101244918; 101244919; 101244920; 101244921; 101246778; 101248172; 101248181; 101248182; 101248541; 101249545; 101254308; 101254311; 101255054; 101256204; 101258175; 10125936; 101260555; 101261456; 10126241; 101267615; 101272708; 101272709; 101274624; 101274625; 101274626; 101274628; 101277327; 101297586; 10130209; 10130220; 10130221; 10130228; 10130518; 101316894; 10131751; 10132142; 101332553; 101333507; 101334563; 101334564; 101335350; 101336199; 101337950; 101337951; 101337952; 101341898; 101341900; 101341902; 101341903; 101341906; 101341908; 101343628; 101351588; 101351589; 101352201; 101352202; 101352203; 101352987; 101354421; 101358537; 101359398; 101361460; 101361461; 101363286; 101363287; 101365; 10137403; 101374280; 101393321; 101396024; 101396025; 101396026; 101396027; 101411516; 101414856; 101414857; 101414858; 101415131; 101416579; 101416580; 101416581; 101416582; 101432331; 101432359; 101432361; 101446816; 101450393; 101450394; 101469918; 101479388; 101490301; 101490302;

101496819; 101500794; 101503039; 10150542; 101510116; 101513786; 101524185; 101524186; 101524187; 101527245; 101527246; 101527247; 101527248; 101527249; 101527250; 101527251; 101527252; 101536942; 101537570; 101541823; 10154314; 10154423; 10154424; 101549; 101552; 101561237; 101561238; 101561239; 101563617; 10156843; 10157468; 101576759; 101576760; 10158253; 10158347; 101592001; 101599384; 101601471; 101601472; 101609934; 101612205; 101613883; 101613884; 101613885; 101613886; 101613888; 101616853; 101623194; 101623195; 101623196; 101623271; 101623274; 101623277; 101623278; 101638834; 101639401; 101639402; 101639403; 101639408; 101639409; 101639410; 101639411; 101639514; 101643640; 101643826; 101655828; 101656173; 101656176; 101657985; 101657986; 101660207; 101660954; 101670043; 101670044; 101671697; 101674474; 101675298; 101680880; 101681439; 101682797; 101683441; 101683446; 101683969; 101687593; 101692822; 101704221; 101704222; 101704223; 101707173; 101707929; 101713931; 101713932; 101713933; 101713934; 101713935; 101713936; 101713937; 101713938; 101713939; 101713940; 101713941; 101713942; 101713943; 101713944; 101713945; 101713946; 101713947; 101713948; 101713949; 101714427; 101714783; 101716382; 101716; 101719625; 101736876; 101736877; 101743026; 101748964; 101748; 101749879; 101750; 101754109; 101754110; 101754111; 101754112; 101754113; 101759565; 101761047; 101763212; 101763240; 101763542; 101763543; 10176372; 10176493; 101768715; 101769034; 10177298; 101775360; 101781421; 101787337; 101788996; 101788997; 101788998; 101788999; 101789000; 101791884; 101796201; 10179672; 10179833; 101798; 10180447; 10180589; 101812; 101831361; 101834100; 101850045; 101850046; 101850047; 101850048; 101856242; 101856243; 101856244; 101856245; 101856246; 101856247; 101856248; 101856249; 101856250; 101857765; 101857766; 101857767; 101857768; 101868029; 101884974; 101889223; 101889225; 101889226; 101917632; 101920824; 101921445; 101921446; 101921447; 101929137; 101929138; 101929509; 101931961; 101937950; 101938889; 101938891; 101939804; 10194146; 101941697; 101941700; 101946799; 101950445; 101954256; 101954257; 101956741; 101973271; 101973273; 10197954; 10198045; 101989; 10199055; 101992127; 101992128; 101993908; 101993909; 101995249; 10200582; 102006214; 10200660; 102008722; 102009592; 10201377; 102014817; 102014818; 102017151; 10201852; 102021384; 102021385; 102024; 102025581; 102025582; 102025583; 102025584; 102026061; 102039233; 102039305; 102039306; 102041052; 102041674; 102041675; 102041676; 102041680; 102055047; 102055048; 102061466; 102061477; 102061478; 102061479; 102064125; 102065005; 102065007; 102065640; 102066544; 102074146; 102077710; 102078341; 102089288; 102094139; 102115584; 102129382; 102137583; 102137584; 102140975; 102144861; 102144874; 10215500; 102156372; 102183607; 102183608; 102183897; 102184082; 102185119; 102190; 102191069; 102191387; 102195238; 102195290; 10219674; 10219763; 10219851; 102198; 10219933; 10220444; 102204794; 10220660; 102208394; 10221379; 102214945; 102216879; 102217409; 102217410; 102217411; 102217412; 102217413; 102217415; 10222129; 102241827; 102247923; 10224; 10225; 102272195; 102275215; 102281442; 102284722; 102288; 102301372; 102305954; 102310126; 102311519; 102312079; 102312080; 102315858; 102315893; 102315894; 102315895; 102315896; 102315897; 102315898; 102315900; 102315901; 102318793; 102318799; 10231; 102323123; 102337192; 102340928; 102343888; 102344175; 102348511; 102354558; 102354559; 102354560; 102354561; 102358021; 102359514; 10236867; 102375328; 102375329; 10237567; 102377218; 102379976; 102379979; 102382390; 102382391; 102382392; 102382393; 102385423; 102388301; 10238; 102390120; 102393908; 102406044; 102406103; 102406104; 102406105; 102406106; 102408023; 102420123; 10242501; 102425861; 102429099; 102434321; 102438594; 102452699; 102456668; 10245813; 10246743; 10246748; 10246749; 102476143; 102484158; 102486295; 102489199; 10248; 102498555; 102498556; 102498557; 102504684; 102504685; 102504686; 102504687; 102505156; 102505157; 102505158; 102505287; 102505288; 102505289; 102505290; 102505291; 102505292; 102505293; 102505545; 10250569; 102518679; 102518683; 10251868; 102518926; 102525471; 102531943; 102531948; 102531949; 102531950; 102533390; 102533391; 102533393; 102533750; 102565436; 102565437; 102585774; 102594811; 102596265; 102601177; 102601198; 102601265; 102601816; 102602060; 102602086; 102602138; 102602453; 102602594; 102604; 10261973; 10262957; 10264675; 10266546; 10268056; 10272716; 10272864; 102861; 10286734; 10286828; 10291085; 10291131; 10291132; 102926; 102980360; 103005; 10313382; 10313507; 10315314; 10316099; 10316395; 103176788; 10327727; 10329946; 10330546; 10331746; 10331806; 10333678; 10333679; 10334720; 10335539; 10335668; 10337208; 10337719; 103400907; 10352585; 10352692; 10352742; 10353134; 10353980; 10356391; 10356392; 10357811; 10357; 10362816; 10375257; 10377413; 10378118; 10379675; 10380668; 10380669; 10382018; 10382202; 10382476; 10385025; 10397757; 10398656; 10399290; 10405207; 10405378; 10407; 10408; 10420029; 1042025; 1042026; 10420631; 10420749; 10421380; 10423065; 10425762; 10428933; 10432218; 10432591; 10442556; 10444197; 10444790; 10445127; 10450027; 10450108; 10450164; 10451345; 10451346; 10453; 10464871; 10465358; 10467862; 10468817; 10469683; 10471605; 10472035; 10475946; 10476759; 10488210; 10488243; 104901990; 10491000; 10492175; 10492695; 10492712; 10493415; 10496005; 10497333; 10497334; 10497840; 104983338; 10498455; 10501351; 10502137; 10512044; 10513421; 10517733; 10522051; 10522220; 10522299; 10523503; 10525452; 10525951; 10526231; 10534; 10537873; 10537875; 10537876; 10539954; 10540625; 10541329; 10541816; 10542811; 10542812; 10543948; 10545089; 10546035; 10559749; 10559752; 10560187; 10562744; 105652; 10567972; 10569286; 10570227; 10570307; 10572072; 10572420; 10573150; 10573892; 105755; 10585873; 10586892; 10586; 10593074; 10594543; 10595320; 10595553; 10608036; 10608335; 10610796; 10611433; 10611781; 10611929; 10613577; 10615808; 10616841; 10618693; 106199634; 10620358; 10620635; 10621051; 10621285; 10637535; 10638669; 10639564; 10640869; 10640870; 10641574; 106446717; 10645741; 10646380; 106530; 10655819; 10656750; 10657543; 10657544; 10657; 10658324; 10659; 10661302; 10663864; 10663892; 10665415; 10665479; 106664862; 10666832; 10669869; 10670013; 10670251; 106756; 106766; 10685040; 10685185; 10689221; 10689952; 10690500; 10690501; 10690; 106997; 10702958; 10705814; 10710079; 10710355; 10712450; 10712; 10713322; 10713608; 107148; 107152; 107153; 10716040; 10716568; 107166; 10716945; 10717058; 107217; 10726508; 10728742; 10730036; 10730679; 10733467; 10733587; 10734482; 10735234; 10735452; 10735721; 10736131; 10736204; 10736573; 10737557; 107381; 10738435; 10739848; 10740394; 107415; 107428; 10748; 10750355; 10751385; 107526; 107531; 10754209; 10754484; 10755458; 10757174; 10762834; 10764455; 10764477; 10774220; 10774652; 10776176; 10778501; 10778968; 10780109; 10783330; 10783457; 10786023; 10786234; 10787641; 10797810; 10799776; 10799923; 10803581; 10804709; 10807050; 10807081; 10808024; 10808238; 10809722; 10810866; 10811932; 10822901; 10824559; 10828482; 10828483; 10828595; 10829672; 10831587; 10831701; 10831951; 10832216; 10834398; 10848285; 10850593; 108518; 10852189; 10854476; 10856008; 10856009; 10856614; 10857428; 10857465; 10858800; 10860373; 10861057; 10861500; 10866852; 10867438; 10867736; 10868700; 10870059; 10871088; 10871527; 10871589; 10871590; 10871647; 10871924; 10872923; 10875801; 10877695; 10877712; 10878214; 10878730; 10878761; 10880407; 10881934; 10882351; 10882580; 10883845; 108865888; 10886; 10888696; 10889367; 10889846; 108910161; 108910232; 10892749; 10893330; 10895273; 10899282; 10899502; 10899810; 10900196; 10900197; 10900550; 10901910; 10903976; 10904041; 10910141; 10910359; 10910683; 10911004; 10911015; 10911122; 10911178; 10914184; 10914307; 10914762; 10914822; 10915407; 10915413; 10915634; 10921067; 10921394; 10921395; 10921661; 10922510; 10923602; 10923890; 10924111; 10924935; 10925369; 109257; 10925912; 10926361; 10927432; 10927615; 10927952; 10932578; 10932900; 10936097; 10936727; 10937184; 10938531; 10943594; 10943988; 10944427; 10945502; 109455; 109457; 109467; 10946824; 10946883; 10947739; 10947; 10948020; 10948021; 10948022; 10948259; 10949340; 10954115;

10954241; 10954800; 10957294; 10958608; 10959366; 10965080; 10965116; 10965117; 10965664; 10965754; 10965982; 10976419; 10976576; 10977693; 10978491; 10979132; 10979133; 10979870; 10980308; 10980962; 10980989; 10981151; 10981878; 10982819; 10987888; 10989578; 10991489; 11001447; 11003731; 11003760; 11005; 110061; 11006; 11008545; 11009821; 11009939; 11010303; 11011823; 11012820; 11013287; 11013385; 11013486; 11019447; 11019448; 11019677; 1101; 11020081; 11022828; 11024312; 11024313; 11024690; 11030343; 11030410; 11030934; 11030980; 11032733; 11032861; 11033660; 11035136; 11035450; 1103; 11041412; 11041669; 11042198; 11044462; 11044463; 11045418; 11045419; 11053559; 110543; 11054999; 11057392; 11061531; 11062926; 11062992; 11063897; 11065760; 110664; 11067474; 11073962; 110774501; 110774506; 110774508; 110801; 11080218; 11080340; 11085706; 11085769; 11087087; 11088079; 11088249; 11089343; 11095734; 11095756; 11095817; 11096585; 11096986; 110975; 11097894; 11099875; 11099946; 11100123; 11101912; 111037; 11105941; 11106485; 111066; 111068; 11107340; 11108287; 11108609; 11108647; 111112; 11111355; 111123; 11112690; 111162; 11116905; 11117256; 11118; 11120942; 11121559; 11121591; 11127487; 111276; 11128649; 11129405; 11129623; 11130568; 11130997; 11131493; 111317; 11131836; 11132129; 11133629; 11137679; 11137711; 11138574; 11138698; 11140408; 11140811; 11140872; 11141474; 11143357; 11144038; 111440; 11149189; 11152467; 11153318; 11159858; 11159887; 1115; 11160553; 11160699; 11160700; 11161320; 11161383; 11163251; 11163665; 11164126; 11165767; 11166685; 11172404; 11178325; 11186917; 11187028; 11189328; 11194651; 11197106; 11197537; 11199796; 112045; 11209669; 11210019; 11210041; 11210480; 11212308; 11219061; 11221120; 11229130; 11229319; 11230718; 11231136; 11231669; 11232903; 11235268; 11240926; 11242951; 11245841; 112499931; 11253924; 11254395; 11256606; 11257846; 11258605; 112587029; 11264589; 11264633; 11265551; 11265594; 11265793; 11267428; 112683664; 11268594; 112733; 11275354; 11277776; 11278765; 11279303; 11279546; 11280355; 11282103; 11287450; 11287704; 11288533; 11290467; 11293112; 11300413; 11302723; 11309871; 11309884; 11326713; 11330952; 11332852; 11333; 11334208; 11334467; 11336802; 11337214; 11339365; 11339; 113426744; 11344362; 11345108; 11345644; 11346715; 11349616; 11350578; 113514; 113555; 113556; 113557; 11355843; 11355844; 11355845; 113558; 11356861; 11360256; 11367383; 11367738; 11367770; 11368070; 11368078; 11370367; 11370731; 11371704; 113798150; 11380413; 11382278; 11390118; 11390178; 11394931; 11395358; 11401514; 11402060; 11405597; 11406057; 11412693; 11412863; 11413111; 11413787; 11415532; 11417566; 11417938; 11418100; 114242; 11425306; 11425530; 11428058; 11428910; 11435344; 11439335; 11439443; 1143; 114416; 11446864; 11447324; 11448816; 11449685; 11458041; 11458; 11459635; 11460996; 11462690; 11464650; 11469444; 11469606; 11471482; 11472702; 11472836; 11473920; 11474759; 11476180; 11479470; 11480803; 11480819; 11480827; 11480996; 11481187; 11482534; 11482768; 11482938; 114840; 11484198; 11485060; 114917551; 11492034; 11493885; 11499542; 11505884; 11506979; 11513915; 11517; 11521526; 11521527; 11521869; 115296; 11530949; 11535892; 11535896; 11537196; 11543024; 11543901; 11546576; 11550267; 11553047; 11557526; 11557604; 11565360; 11567875; 11571906; 11571917; 11572003; 11573738; 11579278; 11579977; 11581224; 11586487; 11586716; 11586787; 11587282; 11587557; 115880410; 115881882; 11600783; 11600832; 11602584; 11607321; 11608027; 11608251; 11608359; 11609865; 11615984; 11622665; 11622; 11629953; 11631979; 11635; 11638838; 11651921; 11652469; 11652903; 11658435; 11658617; 11658901; 116625; 11666434; 11667300; 11672764; 11672998; 11673181; 116740; 11688487; 11688608; 11688997; 11695553; 117064633; 117065485; 117065522; 117071910; 117072200; 117072211; 11709065; 11711550; 11715628; 11722740; 11723884; 11723885; 11724912; 11724913; 11726200; 11727528; 11727586; 11727633; 11729495; 117336; 11736191; 11737575; 11739408; 11739958; 11745149; 11745183; 11745248; 11745625; 11745761; 11745797; 11746218; 11748462; 11750118; 11759509; 11759599; 11759824; 11759867; 117600270; 11760194; 11760621; 117611212; 117611218; 117614134; 117614266; 117616350; 117627717; 117630657; 117630658; 117637232; 117637248; 117637283; 117638753; 117655275; 117660649; 117660650; 117660673; 117660679; 117676065; 117695343; 117701887; 11770452; 117704746; 117710040; 117710087; 117710153; 117713264; 11771430; 117717938; 117717939; 117718721; 11772676; 11774580; 117748621; 117748622; 117753; 117754; 117758023; 117758921; 117758922; 117767346; 117767347; 117768413; 117770884; 11777164; 117775985; 117783790; 117783796; 11778392; 117802471; 117804851; 117812831; 117813967; 117814005; 11782179; 11782294; 117827648; 117828125; 117828231; 11782878; 117829171; 11783601; 117838737; 11784200; 117846090; 117846436; 117846460; 117846599; 117847498; 117847499; 117879819; 11788604; 11788789; 117888349; 117891021; 11789843; 117904274; 11790776; 11791189; 11792647; 117927398; 117932049; 11793217; 117932481; 11793500; 117938207; 11793851; 11794053; 117947517; 117956535; 117956699; 117956764; 117962430; 11796727; 117971366; 117983464; 117986631; 117999540; 118002230; 118014527; 118014533; 118016972; 118016987; 118017465; 118017508; 118026843; 118036177; 118049141; 118052396; 118052411; 11805319; 11805588; 11806200; 11807136; 11807209; 11807674; 11808195; 118082090; 118083704; 118085104; 118085720; 118095399; 118098295; 11810357; 118104178; 118120887; 118120888; 118133312; 118134; 118137522; 118146605; 118149201; 11816899; 118180349; 118198083; 118213810; 118225429; 118225467; 11823039; 11823089; 118230914; 11823898; 118243157; 11824562; 118262473; 118262517; 118274795; 118288994; 118290; 11830220; 11830551; 11830590; 11830914; 118311318; 11832737; 118332684; 11833785; 11834367; 11834368; 118357123; 118360319; 118360420; 118362244; 118362245; 118362246; 118362247; 118362248; 118362252; 118362253; 118362256; 118362299; 118362305; 118362307; 118362308; 118362311; 118362312; 118362313; 118362314; 118362315; 118362318; 118372947; 118391655; 118391741; 118398343; 118398855; 118424040; 118428901; 118429239; 118437699; 118443208; 118457258; 118457310; 118459248; 118460906; 118469570; 118469679; 118472759; 118489617; 118503764; 118508; 11850; 118520353; 118525346; 118525384; 118529038; 118530; 118536398; 118543688; 118543701; 118544661; 118546377; 118546382; 118549274; 118550378; 118550457; 118550487; 118550724; 118561842; 11857496; 11857498; 11857500; 11857717; 11857723; 118579370; 11858221; 11858646; 118630063; 11865435; 11865436; 11865437; 11865438; 118661532; 118663275; 118687354; 11869260; 118704897; 118705308; 118705318; 118705681; 118711818; 118796425; 118797328; 118797416; 118797420; 118797422; 118797598; 118797610; 118797621; 118855887; 118855889; 118855898; 118855901; 118855904; 118855910; 118855920; 118855927; 11886768; 11886769; 118870586; 118873056; 11887; 118887899; 118888882; 118891706; 118892456; 118907360; 118908285; 118908291; 118908294; 118908313; 118908314; 118908315; 118908316; 118908318; 118908320; 118908324; 118908329; 118908330; 118908332; 118908340; 118924468; 118961526; 118968871; 118968872; 118968873; 118968874; 118968875; 118968876; 118973953; 118973954; 118973955; 118973957; 118981422; 118991; 119025350; 119057692; 11908385; 11908386; 11908449; 11908450; 11908451; 11908483; 11908488; 11908489; 11908490; 11947891; 11947892; 11947893; 11947894; 11947895; 11947902; 11947903; 11947904; 11949368; 11949370; 11949485; 119508; 11954203; 11954275; 11957474; 11959770; 11959966; 11959967; 11961382; 11961810; 11963500; 11963549; 11968638; 11968999; 11969000; 11969488; 11970126; 11970252; 11974512; 11975273; 11975469; 119831; 11987845; 119900; 11990295; 11996452; 12000178; 12003173; 12003287; 12009644; 1201442; 12038; 12051505; 12054511; 12065179; 120673; 120700; 12071879; 12082513; 12091582; 121222434; 121225574; 121226874; 121234041; 121254746; 12125606; 121260612; 121263949; 121269198; 121269337; 121274592; 12128377; 121289047; 121290475; 121290477; 121290492; 121297517; 121329665; 12133246; 12134740; 121354405; 121354406; 121354407; 121354414; 121354420; 121354836; 121354838; 121356770; 121357112; 121360446; 121372592; 121372594; 121381026; 121391436; 121391565;

1213; 121419897; 121426842; 121430411; 121434193; 121434245; 121435209; 121435219; 121443518; 121443528; 121446423; 121448521; 121452036; 121452037; 121459517; 121461316; 121484514; 12163610; 12163807; 121667; 12193653; 121942; 12225100; 122305; 12237817; 12245515; 12245518; 12245522; 12245526; 12245536; 12245537; 12285853; 12285856; 12285861; 12285862; 12285863; 12285870; 12285871; 12285877; 12285878; 12285879; 12285885; 12285886; 12285889; 12285890; 12285891; 12285892; 12302012; 12302019; 12302243; 12302626; 12302628; 12302658; 12302663; 12302671; 12302672; 12302675; 12302681; 12302683; 12302691; 12302694; 12303301; 12303303; 12304257; 12304900; 12304904; 12304907; 12304908; 12304961; 12304964; 12305017; 12305019; 12305247; 12305733; 12305793; 12305795; 12305796; 12305797; 12305799; 12305800; 12305811; 12306047; 12306048; 12308888; 12308889; 12309449; 12310779; 12313029; 12313177; 12313180; 12313181; 12313320; 12313605; 12314098; 12314099; 12314100; 12314116; 12314994; 12314995; 12315259; 12315386; 12315387; 12327; 12329; 12355590; 12358318; 12358319; 12358320; 12358335; 12358344; 12358354; 12358355; 12366; 12378870; 12378871; 12378873; 12386; 12388; 12389; 12391; 12395; 12397; 12398; 12401; 12403; 12405; 12406; 124448; 12444995; 12444996; 124726; 124766; 124789; 125020; 125207; 12523; 125299; 12530; 12533; 12534; 125414; 125507; 125608; 125691; 12571389; 12592; 12599; 12600; 126056; 12635860; 12635862; 126561; 12667626; 12678469; 1268109; 1268142; 126987; 127027; 12719461; 127259; 12729; 12762300; 12765315; 12775774; 12775775; 12775778; 12775781; 12775782; 12775784; 12775787; 12789962; 127955; 12800343; 12800344; 12800345; 128003; 128004; 128033; 128287; 12831083; 12831084; 12844; 12848726; 12858404; 12866362; 12866365; 12883686; 12883687; 128869; 12897372; 12897381; 12897873; 12897874; 12898223; 12898660; 12900623; 12900888; 12900889; 12900890; 12900892; 12900894; 12900895; 12900896; 12900897; 12900898; 12903136; 12903171; 12903190; 12903215; 12903236; 12904912; 12904914; 12904917; 12904922; 12904923; 12905080; 12905081; 12914607; 12914615; 12917773; 12920857; 12920878; 129231; 129679; 12968428; 12978129; 12980878; 12980879; 12986433; 129885; 12988705; 129887; 130008; 13020135; 13035200; 13035201; 13036549; 13036554; 130442; 13059510; 13059914; 13059935; 13059936; 13059940; 13059943; 13059947; 13059951; 13093580; 13093581; 13093585; 13093706; 13093710; 13093712; 13097725; 131122; 13122553; 13131822; 13174983; 13174986; 13186786; 13186787; 13186790; 13186791; 13186794; 13187263; 13213649; 13217259; 13224598; 13229822; 13231946; 13231947; 13233; 132485; 13253013; 13254459; 13255083; 13255084; 132578; 13270883; 13273272; 132795; 13289406; 13317985; 13319174; 13319175; 13319176; 13321318; 13334675; 13334676; 13340696; 13340697; 133473; 13362146; 13365325; 13365327; 13365329; 13365341; 13365345; 13365350; 13365355; 13371686; 13371687; 13385085; 13392894; 134038; 134081; 134200; 13436; 134512; 13455605; 134601; 13465780; 134815; 13537818; 13537819; 13575093; 135941; 13606766; 13606768; 13640738; 13640739; 13641754; 136460; 13672003; 13688019; 13688020; 137065; 13743924; 13746805; 13751942; 13751943; 13751946; 13752534; 13752536; 13752538; 13767387; 13771704; 13771705; 13783679; 1379956; 1381794; 13823553; 13829314; 13829320; 13829321; 13830884; 13832522; 13832523; 13832658; 13832659; 13833211; 13844864; 13844865; 13844867; 13845841; 13845844; 13845845; 13845846; 13845847; 13846039; 13847795; 13849; 13855185; 13855186; 13855187; 13871654; 13871655; 13871992; 13871996; 13873432; 13875145; 138824; 13883773; 13886369; 13886371; 13886608; 13886746; 13887784; 13889162; 13894893; 139200; 13922705; 13922706; 13932299; 13932300; 13932301; 13932302; 13939150; 13939151; 13944845; 13944847; 13944848; 13958183; 13958185; 13962219; 13962248; 13962978; 13985537; 13985538; 13989841; 13991149; 13991152; 13996640; 14017900; 14018116; 14029492; 14034359; 14037246; 14037247; 14037250; 14052123; 14052124; 14054189; 14054190; 14055602; 14060872; 14077987; 14078720; 14080177; 14101679; 14101757; 14102402; 14102449; 14102450; 14102495; 14104214; 14104215; 14104219; 14104220; 14104222; 14123183; 14123184; 14123185; 14127007; 14127008; 14130767; 14130768; 14130780; 14130782; 14131758; 14131778; 14132412; 14138808; 14151452; 14151453; 14151454; 14151455; 14151456; 14151460; 14151461; 14151462; 14151463; 14151464; 14151465; 14151466; 14151467; 14151468; 14151470; 14151484; 14159378; 14159382; 14159383; 14188133; 14189035; 14194046; 14206185; 14211991; 14211993; 14212063; 14212831; 14214709; 142266; 14228; 14235; 14236845; 14237331; 14237950; 14239330; 14257492; 14258186; 14258187; 14259293; 14259437; 14259438; 14259439; 14281967; 14281968; 14281993; 14298900; 14298901; 14298903; 14298905; 14307947; 14308478; 14309979; 14311531; 14311887; 14311888; 14313004; 14313005; 14325798; 14328; 14334452; 14334453; 14334456; 14334457; 14337663; 14353155; 14369164; 14393376; 14393378; 14393381; 14393383; 14393409; 14408225; 14411656; 14413875; 14413876; 14427684; 14428217; 14428218; 14428854; 14429924; 14429925; 14454353; 14456283; 14456285; 14462207; 14462211; 14462212; 14462253; 14462261; 14462263; 14464116; 14464444; 14464994; 144746; 14481266; 14487703; 14493338; 14493339; 14493340; 14493341; 14493; 14504742; 14506107; 14506108; 14515637; 14522746; 14522747; 14524540; 14535450; 145386; 14542060; 14542061; 14558364; 14558751; 14558757; 14561089; 14564496; 14564514; 14576077; 14580688; 14580689; 14585418; 14585977; 14587816; 14614417; 14625414; 14629480; 14629481; 14631135; 14631136; 146374; 14651051; 14655094; 14655096; 14655652; 14656371; 14656373; 14661162; 146694; 146759; 14679078; 1467; 14680929; 14681008; 14682822; 14682824; 14701841; 147108; 147194; 14730853; 14752585; 14803267; 14803269; 14803271; 14803280; 14803281; 14803283; 14803285; 14803287; 14803296; 14803337; 14803339; 14820031; 14823978; 14842875; 14865241; 14875887; 14885840; 14889443; 14889445; 14889447; 14922; 14928; 14947697; 15016409; 15042978; 15043184; 15081097; 15081520; 15088993; 15093640; 15094; 151009; 151050; 151149; 15118715; 151199; 151200; 151263; 15126540; 1513138; 15135477; 15138379; 15139364; 151407; 151410; 15143807; 151464; 15156515; 15159804; 151728; 15173104; 15178973; 15178974; 15206720; 152109; 152205; 15225519; 152304; 15268; 15269; 152867; 15385692; 15403645; 15408887; 15408938; 15461361; 15487548; 1548879; 1548880; 1548943; 1549025; 1549026; 1549045; 1549080; 1549089; 1549370; 1549371; 1549778; 1549943; 1549944; 1549945; 1549946; 1549947; 1549948; 1549949; 1549950; 1550884; 1550888; 1550890; 155152; 15516023; 15540144; 155448; 15546721; 15558051; 15558280; 15558282; 15558283; 15559359; 15559364; 15559365; 15559366; 15559367; 15559548; 15559908; 15559911; 15559912; 15559964; 15559974; 15560011; 15560016; 15560222; 15560223; 15560224; 15560225; 15560226; 15560227; 15560228; 15560229; 15560230; 15560285; 15560349; 15560358; 15560361; 15567214; 15567215; 15571350; 15571; 15580525; 15594137; 15594139; 15596425; 15596677; 15601861; 15601862; 15604509; 15606053; 15609908; 15613266; 15652290; 15655013; 15694154; 15694155; 15714465; 15714826; 15717; 157355; 15743318; 15755615; 15772871; 15814211; 158410; 15929663; 159296; 15931060; 15931793; 15938965; 15938966; 1593918; 15942877; 15946243; 15946245; 15949153; 15949155; 15949156; 15949159; 15949160; 15949162; 15949244; 15949245; 15949248; 15949530; 15949531; 15949532; 15949533; 15949534; 15949535; 15949537; 15949623; 15949625; 15949627; 15949628; 15949629; 15949631; 15949632; 15949633; 15949716; 15949717; 15949719; 15949720; 15949721; 15949723; 15949724; 15949726; 15949727; 15949802; 15949804; 15949805; 15949806; 15949813; 15949816; 15949817; 15949821; 15954437; 15954438; 159630; 15975980; 15977259; 15979; 15984857; 15991568; 16007308; 16019992; 16035194; 16037931; 16038696; 16046428; 16048618; 160514; 16057040; 16057976; 16058096; 16058296; 16058491; 16064706; 16066486; 16066491; 16066506; 16066519; 16066531; 16066533; 16066536; 16066555; 16066574; 16066595; 16066600; 16066605; 16066620; 16066642; 16066656; 16066657; 16067095; 16067133; 16071271; 160739; 160799; 16088454; 16095604; 16095605; 16095607; 16100574; 16100870; 16104182; 16104183; 16104184; 16104186; 16104197; 16112216; 16112225; 16112232; 16112265; 16112269; 16112272; 16112276; 16112278; 16112279; 16113593;

16116680; 16116681; 16119738; 16120724; 16120726; 16120750; 16122159; 16127315; 16127316; 16131851; 16131893; 161426; 161605; 1616260; 161658; 161660; 16204499; 16204529; 16204532; 16205756; 16211032; 16211337; 16211884; 16211940; 16211941; 16211942; 16211984; 16211985; 16211986; 16212017; 16212688; 16212773; 16212959; 16212960; 16212965; 16212966; 16213096; 16213406; 16213544; 16213545; 16213546; 16213640; 16213872; 16215116; 16215295; 16217112; 16217535; 16217536; 16217605; 16217663; 16217703; 16218587; 16219440; 16219567; 16219580; 16219668; 16219991; 16220012; 16220080; 16222281; 1623625; 16237; 162458; 162464; 16255; 16296; 16299; 16315; 16330; 16385908; 16396482; 16397049; 16401316; 16401533; 16401564; 164615; 16467934; 16467935; 16467937; 16467941; 16467942; 16467944; 16467945; 16467949; 164929; 165139; 165171; 165224; 16537; 16582; 165863; 16617; 16654696; 16662633; 16663020; 16663874; 16664000; 16664002; 16678017; 16682922; 16683120; 16683462; 16683610; 16683615; 16683848; 16683858; 16685680; 16686633; 16687863; 16688152; 16689313; 166991; 16726322; 16727226; 16727374; 16728279; 16728280; 16728929; 16734862; 167350; 16738719; 16742819; 16746177; 16750077; 16752556; 16752557; 16752558; 16752559; 16752560; 16754230; 16757696; 16757726; 16758060; 16758093; 16758204; 16759282; 16759610; 16760025; 16760035; 16760165; 16760260; 16760266; 16760534; 16760663; 167720; 167792; 16792990; 168037; 168120; 16821; 168973; 169185; 169509; 17029; 17047057; 17093839; 17093873; 171147; 1711944; 1711945; 1712057; 1713001; 171355; 1714417; 1715136; 171554; 17170430; 173312; 173534; 17356; 1736; 1737; 1738122; 173843; 1742210; 1742211; 174618; 175553; 1757003; 17617; 17623; 176439; 17697; 17747541; 17749715; 17750876; 17753098; 17753099; 17753100; 17753101; 17753215; 17753216; 17753217; 17753218; 17753338; 17753339; 17753340; 17753341; 17753460; 17753463; 17753586; 17753587; 17753692; 17753808; 17753821; 17753866; 17753954; 17753975; 17754245; 17757223; 17757224; 17757248; 17757268; 17786297; 17786298; 1778770; 1778771; 1779141; 1779142; 17801046; 17801047; 17804309; 1780944; 17850364; 17865563; 17876128; 17879449; 17879454; 17898023; 17899; 17907921; 17920225; 17941481; 179451; 17967519; 17975430; 17975431; 17975434; 17978147; 17989646; 17989647; 17990888; 17996413; 17998692; 18000520; 18000521; 18003269; 18003831; 18004373; 18008843; 18008844; 1801285; 1801286; 1801289; 1801290; 1801291; 1801292; 1801293; 1801294; 181215; 181487; 18172328; 18172963; 18185507; 18218184; 18218210; 18218266; 18218299; 18218318; 18218373; 18218435; 18218453; 18218480; 18218481; 18218518; 18218589; 18219020; 18219039; 18219099; 18219118; 18219158; 18219178; 18219198; 18219227; 18219228; 18219234; 18219235; 18219236; 18219276; 18219354; 18219374; 18219411; 18219449; 18219468; 18219488; 18219506; 18219546; 18219566; 18219603; 18219610; 18219612; 18219613; 18219615; 18219616; 18219617; 18219619; 18219620; 18219621; 18219623; 18219625; 18219628; 18219641; 18219659; 18219678; 18219738; 18219816; 18220203; 18220512; 18220552; 18220572; 18220937; 18220957; 18221051; 18221103; 18221104; 18221208; 18221651; 18221671; 18221745; 18221848; 18221849; 18221971; 18222022; 18222041; 18222206; 18222207; 18222336; 18222373; 18222391; 18222556; 18222743; 18223063; 18223099; 18223100; 18223106; 18223107; 18223116; 18223118; 18223119; 18223121; 18223122; 18223123; 18223125; 18223126; 18223127; 18223128; 18223130; 18223134; 18223193; 18223194; 18223228; 18223229; 18223244; 18223245; 18223262; 18223404; 18223405; 18223487; 18223798; 18223836; 18223854; 18223932; 18224086; 1822408; 18232225; 18232263; 18232283; 18232398; 18232416; 18232463; 18232464; 18232635; 18232655; 18232735; 18232843; 18232844; 18233388; 18233527; 18233787; 18233927; 18234034; 18234041; 18235120; 18235380; 18235420; 18235496; 18235515; 18235594; 18235623; 18235624; 18237093; 18237740; 18237741; 18237797; 18237804; 18237879; 18237880; 18246520; 18246660; 18246766; 18246773; 18249311; 18249450; 18249557; 18249564; 18249610; 18251690; 18251697; 18251830; 18251837; 18252642; 18302478; 18305099; 18315103; 18315104; 18315107; 18322814; 18322815; 18329476; 18336619; 18340876; 18341314; 18341322; 18341357; 18341358; 18342218; 18343463; 18343955; 18345552; 18345553; 18345556; 18350600; 18363761; 18372047; 18374253; 18380328; 18380350; 18380383; 18380433; 18387200; 18392540; 18393063; 18393072; 18394841; 18394842; 18395993; 18396184; 18396276; 18398597; 18406734; 18410144; 18410145; 18411140; 18411141; 18411204; 18411231; 18417383; 18428334; 18431280; 18439495; 18459124; 18465617; 18471261; 18475307; 18479573; 18484765; 18484805; 18484825; 18484884; 18484901; 18484979; 18485008; 18485009; 18485550; 18485689; 18485849; 18485949; 18486089; 18486196; 18486203; 18486249; 18487140; 18487533; 18487571; 18487590; 18487661; 18487763; 18489085; 18489861; 18489917; 18489924; 18489998; 18490690; 18490710; 18506303; 18506371; 18513146; 18519671; 18520874; 18530284; 18533869; 18533910; 18533955; 18533964; 18533970; 18534040; 18534054; 18534055; 18534082; 18534116; 18534147; 18534154; 18534227; 18534243; 18534247; 18534317; 18534336; 18534350; 18534396; 18534427; 18539356; 18539357; 18544364; 18544366; 18544368; 18544369; 18546721; 18546752; 18547325; 18551495; 185515; 185698; 18595379; 18598341; 18598879; 18598981; 18599330; 18600476; 18601090; 18601091; 18601215; 18601800; 18602030; 18602123; 18602519; 18603795; 18604375; 18604492; 18604493; 18604494; 18604495; 18604925; 18604996; 18604997; 18606405; 18606872; 18607776; 18607777; 18607778; 18607955; 18607956; 18607958; 18608245; 18608252; 18608253; 18610684; 18611133; 18611585; 18616104; 18617867; 18620506; 18632628; 18633348; 18635982; 18635983; 18637152; 18637693; 18637778; 18637780; 18637781; 18638133; 18639325; 18639457; 18639622; 18639745; 18640098; 18640100; 18640998; 18641319; 18641494; 18641667; 18642055; 18643544; 18643545; 18644376; 18644774; 18645216; 18648928; 18648929; 18648931; 18649049; 18649154; 18651106; 18652383; 18652485; 18652523; 18652729; 18653296; 18653354; 18654465; 18654542; 18654745; 18656773; 18657710; 18658606; 18658629; 18658630; 18658631; 18660336; 18661166; 18661236; 18662833; 18663854; 18663855; 18664161; 18664599; 18664602; 18664777; 18665577; 18667086; 18667087; 18667088; 18667390; 18667392; 18669076; 18671255; 18683842; 18683843; 18686698; 18699538; 18723938; 18740061; 18755257; 18755964; 18755966; 18755968; 18755975; 18755979; 18757201; 18759588; 18780185; 18781703; 18798221; 18815; 18826; 18845; 188966; 188; 189091; 18913674; 18946123; 18950; 18951434; 18951435; 1895423; 1895424; 1895425; 1895426; 18956941; 18956943; 18963706; 18972208; 18974248; 18974249; 18974250; 18974251; 18980652; 18985125; 18985882; 18988210; 18988211; 18988266; 18990784; 18999282; 18999286; 18999287; 19001072; 19003895; 19009567; 19011297; 19011298; 19018346; 19021974; 19025158; 19036787; 19040011; 19044925; 19044926; 19045976; 19066599; 19067470; 19068859; 19068862; 19074062; 19082593; 19082594; 19088191; 19088192; 19091660; 19093231; 19094548; 19094646; 19099178; 191008; 19108248; 1912; 19137; 1913; 191554; 191596; 19169216; 191759; 191804; 191; 192435; 192877; 192904; 193184; 193307; 193325; 193385; 193395; 193527; 193541; 193557; 19362607; 19362608; 19365643; 19365658; 19365677; 19365683; 19365684; 193692; 19374170; 19389577; 19389578; 19418688; 19418689; 19422583; 19422584; 194237; 194279; 19428469; 19428470; 194780; 195189; 195270; 195592; 195678; 195829; 195887; 19591441; 19602160; 19604878; 19607576; 19607577; 19609335; 19611869; 19611870; 19611977; 19611978; 19611980; 19611981; 196128; 196333; 196415; 196585; 196657; 196658; 196694; 196695; 19693366; 19695046; 19710830; 197415; 19742652; 19758224; 19760719; 19761661; 19761829; 19761830; 19769110; 19776208; 19776216; 19777718; 19777719; 19777721; 19780791; 19780793; 19780794; 19780795; 19780796; 19780797; 19780798; 19784027; 19787716; 197; 19800018; 19809340; 19812632; 19814647; 19818370; 19818375; 19821516; 198223; 19824025; 19827232; 19830873; 19836284; 19839186; 19839187; 19842619; 19844; 19846620; 19850728; 19850729; 19856954; 19856955; 19857768; 19863272; 19866392; 19866394; 19866396; 19867605; 19872333; 19874037; 19880722; 19891141; 19893198; 19895964; 19899983; 19905508; 19905517; 19911211; 19915339; 19915600; 19917387; 19928914; 19933555; 19939983; 19956016; 19956018; 19956022; 19956026; 19956925;

19956935; 19962237; 19962242; 19966371; 19987162; 19988414; 19988415; 19988418; 19988419; 19991269; 19995093; 19995094; 20001519; 20001520; 20001577; 20001584; 20001659; 20001660; 20002714; 20002721; 20002854; 20002861; 20004307; 20004308; 20004364; 20004371; 20004444; 20018362; 20018761; 20020259; 20020279; 20055035; 20058190; 20059702; 20064926; 20067220; 20076427; 20076428; 20076448; 20076459; 20076460; 20076463; 20076464; 20076474; 20078710; 20081214; 20094688; 20094689; 20095000; 20111798; 20129705; 20129901; 20129902; 20130713; 20132139; 20139821; 20139867; 20142969; 20193634; 20196978; 20196979; 20196980; 20196981; 20200164; 20200167; 20201341; 20205203; 20207806; 20215803; 20218158; 20219; 20221918; 20230607; 20230823; 20230824; 20231052; 20233464; 20233465; 20233467; 20245554; 2024; 20250536; 20250537; 20257219; 20259181; 20268659; 20268666; 20279586; 2027; 20296829; 20310205; 20317688; 20323513; 20345185; 20345186; 20345187; 20345189; 20345191; 20345192; 20381858; 20388518; 20444942; 20445823; 20445958; 20447780; 20449346; 20458168; 20463797; 20463956; 20463957; 20467050; 20472106; 20475249; 20485499; 20487690; 20487692; 20487697; 20487698; 20487699; 20487702; 20494258; 20494259; 20494260; 20494261; 20494262; 20494263; 20494264; 20499; 20505803; 20505804; 20510305; 20512830; 20523693; 20528817; 20532257; 20536059; 20536060; 20536798; 2054178; 20553998; 20559684; 20559685; 20559807; 20559811; 20565517; 20573903; 20576123; 20584080; 20584086; 20584689; 20588384; 20588386; 20596792; 20596793; 20596798; 20596799; 20596800; 20598613; 20598616; 20598630; 20598844; 20599422; 20603426; 20603763; 206037; 20603874; 20608403; 20611025; 20617222; 20637823; 20638152; 20638461; 20638747; 20639623; 20639661; 20639662; 206439; 20648246; 20652498; 20654318; 20661184; 20662494; 2066644; 20669851; 20669855; 20670004; 20675958; 20682092; 20682839; 20686103; 20686759; 20686760; 20692296; 206; 20702672; 20706023; 20707644; 20708704; 20710692; 20711570; 20716954; 20717004; 20721392; 20721397; 20722697; 20739025; 20739030; 20739031; 20739047; 20739055; 20745934; 20751289; 20752070; 20759811; 20764115; 20764116; 20788567; 20793104; 20797; 20798417; 2080; 20812345; 20812346; 20812347; 20812348; 20813434; 20815780; 20816007; 20818499; 20820696; 20822752; 20823422; 20827823; 20831806; 20831812; 20833507; 20835508; 20835708; 20836326; 20836489; 20836757; 20838599; 20841988; 20843154; 20846226; 20846248; 20846931; 20849148; 20849203; 20978568; 20978569; 20980904; 21057649; 21057651; 21063707; 2108; 21107893; 21116717; 21116845; 21117001; 21118466; 21119048; 21119049; 21119829; 21120114; 21120400; 21120522; 21120523; 21120721; 21120793; 21120827; 21120926; 21120928; 21121201; 21121202; 21121203; 21121205; 21121206; 21121207; 21121284; 21121761; 21121859; 21121860; 21122725; 21123006; 21123007; 21123012; 21123028; 21123714; 21123715; 21124359; 21124407; 21124633; 21124634; 21124737; 21124738; 21124739; 21125434; 21125569; 21126041; 21126276; 21126419; 21126420; 21126429; 21126430; 21126431; 21126472; 21126473; 21126491; 21126498; 21126534; 21126535; 21126537; 21126538; 21127109; 21127636; 21127637; 21134396; 21139461; 21139462; 21139463; 21140823; 21141025; 21143471; 21144857; 21145293; 21145305; 21145314; 21145315; 21145377; 21145378; 21145384; 21145993; 21155384; 21155407; 21155408; 21155412; 21155426; 21155664; 21155775; 21157973; 21157974; 21158794; 21158795; 21158796; 21158797; 21158809; 21158835; 21159022; 21159616; 21189366; 21197433; 21217625; 21225084; 21225091; 21225462; 21225463; 21240766; 21242031; 21253900; 21254459; 21260249; 21263301; 21263316; 21265673; 21265; 21267571; 21267572; 21275180; 21282860; 21285384; 21287615; 21287616; 21287617; 21287626; 21299923; 21302164; 21302165; 21302166; 21302172; 21304745; 21305777; 21306037; 21308126; 21309178; 21309328; 21309382; 21310266; 21311073; 21311708; 21334747; 21339731; 21340501; 21340540; 21342001; 21351659; 21352015; 21355827; 21356457; 21356504; 21362286; 21365250; 21437101; 21442684; 21452266; 21456135; 21480381; 21488125; 21494654; 21503; 21513959; 21515276; 21533040; 21538309; 21549111; 21550415; 21550669; 21550670; 21554668; 21563826; 21572943; 21572944; 21572945; 21572946; 21572947; 21572973; 21575216; 21576867; 21577020; 21577021; 21577768; 21577769; 21579143; 21579144; 21580289; 21581130; 21581143; 21581151; 21581152; 21581900; 21581901; 21585138; 21585140; 21585141; 21587486; 21587581; 21587582; 21589113; 21589114; 21589115; 21589116; 21589215; 21589749; 21589941; 21590186; 21590187; 21590848; 21590849; 21590890; 21590894; 21590992; 21590993; 21593675; 21593777; 21594015; 21594061; 21594062; 21595764; 21596276; 21596277; 21596; 21597299; 21597300; 21597301; 21597302; 21597311; 21597312; 21597313; 21597314; 21599455; 21599456; 21599461; 21599462; 21599611; 21600128; 21600129; 21600159; 21600436; 21600441; 21600442; 21600443; 21600444; 21600445; 21600446; 21600447; 21601151; 21602447; 21602448; 21602449; 21602450; 21603812; 21604003; 21604004; 21604418; 21607273; 21607385; 21607386; 21607387; 21607651; 21607652; 21607653; 21607654; 21607655; 21607656; 21607907; 21608590; 21608591; 21611372; 21611406; 21619889; 21620060; 21626034; 21626035; 21626036; 21626037; 21626038; 21626317; 21627865; 21627866; 21627867; 21627868; 21627869; 21627874; 21629947; 21629948; 21630033; 21630034; 21630534; 21630536; 21631268; 21632987; 21633007; 21633008; 21633013; 21633014; 21633015; 21633016; 21633018; 21633019; 21633020; 21633021; 21633043; 21633044; 21633294; 21634696; 21635049; 21635051; 21636127; 21636128; 21636382; 21637607; 21637608; 21637717; 21647622; 21649216; 21654400; 21659245; 21666305; 21666920; 21667556; 21668361; 21668984; 21669636; 21669637; 21669638; 21669639; 21669640; 21669641; 21670482; 21670483; 21670484; 21670842; 21670843; 21671815; 21671828; 21671829; 21671858; 21671859; 21671860; 21671861; 21671862; 21671863; 21671864; 21675005; 21676472; 21676631; 21676632; 21676638; 21676639; 21680747; 21680748; 21680749; 21681705; 21681706; 21681707; 21682080; 21682385; 21684606; 216878; 21687904; 21691892; 21692112; 21699729; 21704; 21711896; 21711897; 21714625; 21720362; 21721799; 21721800; 21721801; 21721802; 21721803; 21721804; 21721805; 21722576; 21722577; 21722580; 21722581; 21722747; 21724899; 21724900; 21724901; 21724902; 21724903; 21726858; 21726859; 21726861; 21726862; 21726863; 21726864; 21726865; 21726866; 21730847; 21731179; 21732422; 21732423; 21738345; 21739635; 21752099; 21764240; 21767238; 21767241; 21771330; 21771919; 21771920; 21775202; 21775237; 21775658; 21777192; 21778437; 21786111; 21786112; 21786116; 21811227; 21829704; 21829705; 21829710; 21829711; 21835962; 21836311; 21863272; 218640; 21871772; 21871773; 21871782; 21871787; 21871788; 21872029; 21881183; 21881396; 21882146; 21885600; 21889598; 21895; 21896263; 21896270; 21902519; 21903980; 21909112; 21910566; 21910567; 21912482; 21918499; 21924743; 21924836; 21924838; 21924839; 21924840; 21924841; 21924842; 21932710; 21941319; 21947040; 21951661; 21951662; 21954815; 21954816; 219658; 219659; 219662; 21966338; 219672; 21969130; 21969158; 21969691; 21972054; 219722; 21977604; 21982957; 21983273; 219889; 219894; 219936; 21998341; 21999768; 22011100; 22011445; 22015065; 22015066; 22015297; 22020807; 22027040; 22027042; 22027571; 22034852; 22046106; 22058820; 220674; 22071584; 22073254; 22079972; 22081773; 22082179; 22082627; 22090799; 22090829; 22094514; 22094516; 22094536; 22099225; 22099227; 22104957; 22114555; 22114556; 22114594; 22114632; 22114633; 22114635; 22114636; 22131951; 22132196; 22132197; 22138711; 22138715; 221417; 22145309; 221641; 22164467; 22164469; 221691; 22173764; 221771; 22182699; 222076; 22212315; 22212323; 22212824; 22213407; 22213632; 22213799; 22214151; 22214152; 22214409; 22214413; 22214415; 22214796; 22214832; 22214857; 22215476; 22215481; 22215483; 22215510; 22215511; 22215512; 22215514; 22219099; 222284; 22235151; 22242693; 22247539; 22256299; 222626; 22266944; 22269604; 22274889; 22274892; 22274914; 22278476; 22280833; 22281164; 22286958; 22286962; 22297699; 22308819; 22321367; 22327942; 22327943; 22328204; 223328; 22341999; 22342384; 22355568; 22378532; 22378533; 22383; 22395552; 22408696; 22415298; 22416292; 22416293; 22419490; 2242; 22438595; 22457493; 22461935; 22472078; 22491096; 22491930; 22492011; 22492023; 22492025;

22492086; 22492130; 22492244; 22492260; 22492262; 22492267; 22492341; 22492359; 22492658; 22496660; 224; 2251948; 22524185; 22524215; 22524269; 22524270; 22524501; 22524534; 22524649; 22556; 22569128; 22569137; 22569138; 22571156; 22590381; 22595621; 22603872; 22605245; 22606971; 22611009; 22611010; 22619178; 22619179; 22624075; 22625306; 226311; 22639402; 22639564; 22639619; 226486; 22652094; 22654886; 22655024; 22655184; 22658214; 22660042; 22660043; 22660044; 22662186; 22667524; 22673025; 22673026; 22692100; 22692293; 22700722; 22716776; 22719441; 22721381; 22728848; 22735106; 22750782; 22753029; 22766689; 227829; 22787944; 22788155; 22788156; 22788194; 22788226; 22788822; 22789067; 22789068; 22790541; 22790904; 22792327; 22795532; 22795604; 22796170; 22796221; 22796852; 22797248; 22797710; 22797711; 22797713; 22798618; 22799153; 22803478; 22803479; 22803664; 22803666; 22804138; 22804140; 22804361; 22805053; 22806614; 22806693; 22807550; 22808396; 22808894; 22811263; 22811428; 22812602; 22814147; 22814148; 22814150; 22814360; 22817663; 22817850; 22819942; 22821197; 22822788; 22822793; 22822988; 22823186; 22823460; 22823478; 22824305; 22824370; 22825318; 22826161; 22826239; 22826240; 22826301; 22826600; 22827006; 22827299; 22827486; 22827612; 22827616; 22827630; 22827725; 22828754; 22831877; 22836365; 22836366; 22836602; 22837929; 22838233; 22838243; 22838260; 22838602; 22838706; 22839137; 22839230; 22840260; 22840266; 22840267; 22840268; 22841402; 22842016; 22842017; 22842857; 22843672; 22843785; 22843878; 22844325; 22844798; 22847930; 22847931; 22848108; 22848422; 22848660; 22850735; 22851127; 22851371; 22851372; 22851637; 22851688; 22851690; 22851898; 22852473; 22853398; 22854257; 22854268; 22855130; 22855182; 22855277; 22857103; 22857645; 22857916; 22858199; 22858417; 22858436; 22858659; 22859275; 22859547; 22859674; 22860172; 22860173; 22860174; 22860493; 22861193; 22861208; 22865093; 22866008; 22866745; 22866772; 22866779; 22866900; 22867231; 22868285; 22868286; 22868665; 22868666; 22870925; 22871761; 22872022; 22872023; 22872777; 22873; 22874142; 22874730; 22874946; 22875104; 22875305; 22875306; 22875308; 22875309; 22875312; 22875315; 22875316; 22875623; 22876962; 228769; 22878123; 22878208; 22878463; 22878465; 22882169; 22882906; 22886035; 22886932; 22887028; 22887029; 22887034; 22887095; 22887103; 22887125; 22887158; 22887159; 22887195; 22887213; 22887221; 22887238; 22887244; 22887245; 22887246; 22887271; 22887279; 22887295; 22887296; 22887297; 22887298; 22887299; 22887326; 22887330; 22889059; 22890125; 22890129; 22892881; 22892887; 22894967; 22894968; 22895032; 22895034; 22895047; 22895048; 22895050; 22899510; 228; 22901015; 22906672; 22910955; 22910986; 22911059; 22911063; 22911073; 22911074; 22911078; 22911079; 22911092; 22911094; 22911102; 22911124; 22911127; 22911135; 22911158; 22912570; 229153; 22923111; 229385; 229387; 22941602; 22947329; 22956664; 22967427; 22969211; 22969570; 22974872; 22974874; 22979809; 22979958; 22983843; 22986441; 22986447; 229888; 22992890; 22994808; 22995433; 22996547; 22999064; 2302435; 23037255; 23052563; 230525; 23052640; 23052641; 23063599; 23080781; 23081672; 23115740; 23131151; 23154558; 23164327; 23166968; 23170527; 23171546; 23182249; 23182251; 23185605; 23185606; 23195546; 23204986; 23210698; 23210757; 23210780; 23215705; 23215706; 23234527; 23235; 23237123; 23237624; 23237625; 23238443; 23241243; 23241244; 23242840; 23244449; 23244450; 23244520; 23246789; 23246979; 23248124; 23249329; 23249420; 23249440; 23252085; 23252086; 23256531; 23257716; 23258880; 23258881; 23261859; 23261860; 23263392; 23263393; 23272613; 23272614; 23274369; 23274603; 23278838; 23278980; 23279550; 23287375; 23290838; 23295571; 23304512; 23304513; 23304514; 23304668; 23304723; 23314141; 23331601; 23335965; 23335966; 23336346; 23336765; 23336829; 23341028; 23347; 23360998; 23361000; 23362578; 23364408; 23364409; 23364674; 23367540; 23367; 23368; 23374202; 23382597; 23384219; 23398518; 23398520; 23404592; 23412508; 23412755; 23412774; 23421209; 23421440; 23421441; 23421442; 23422869; 234237; 23424028; 23424086; 23424823; 23448603; 2345; 23496757; 23498234; 23498235; 23504570; 23504572; 23506334; 23506672; 23513241; 23513268; 23513273; 23513291; 23517449; 23519593; 23519877; 23520138; 23523623; 23524485; 23525642; 23535; 235414; 23546782; 23548331; 235483; 23558689; 23558721; 23561657; 23563582; 23573691; 23574384; 23574387; 23574390; 23574393; 23574396; 23574399; 23574406; 23583495; 23591264; 23591357; 23615194; 23615231; 23615303; 23615550; 23615841; 23615842; 23615843; 23616132; 23616133; 23616146; 23616171; 23618712; 23619277; 23623651; 23623724; 23623868; 23624080; 23638293; 23640046; 23643895; 23644417; 23644578; 23644587; 236537; 23656132; 23657810; 23657840; 23657846; 23657860; 236687; 23723022; 23723023; 23724695; 23815418; 23851385; 238666; 23902375; 23928115; 23988592; 24003359; 240119; 240121; 24031035; 24114; 24116; 24117; 24178880; 24197; 24207195; 24209769; 24212124; 24212151; 242517; 242530; 24310; 243675; 243691; 24433; 24510; 24513; 24515; 24516; 245572; 245622; 246089; 246929; 24699; 247323; 24743972; 24744895; 24748222; 24749; 24752885; 24755546; 247573; 24759316; 24762201; 24773269; 24778044; 24778045; 24778046; 24778047; 24778176; 24778177; 24778178; 24778179; 24778180; 24778181; 24778182; 24778310; 24778312; 24778313; 24778314; 24778315; 24778316; 24778437; 24778438; 24787636; 24794799; 24799590; 24799591; 24799753; 24802025; 24802149; 24802156; 24802159; 24802163; 24802189; 24802191; 24802240; 24802260; 24802277; 24802279; 24802281; 24802287; 24802290; 24802291; 24802292; 24802293; 24802294; 24802302; 24802417; 24802436; 24802445; 24802515; 24802526; 24802549; 24803871; 24809553; 24812250; 24820766; 24821053; 24822258; 24824402; 24824547; 24824601; 24828119; 24832089; 24835; 24838361; 24838; 248488; 24849373; 24852164; 24858512; 24858513; 24858934; 24858935; 24858936; 24858939; 24858940; 24859094; 24867478; 24872453; 24879712; 24881058; 24883457; 24883458; 24884162; 24884264; 24884265; 24884380; 24886703; 24892722; 24892744; 24898638; 24906306; 24915; 24938378; 249494; 24976875; 24982156; 24982158; 25008017; 250096; 25012069; 25014878; 25021438; 25021704; 25022345; 25023363; 25027315; 25028803; 25045615; 25050060; 25050459; 25052621; 25052622; 25052623; 25052643; 25052645; 25052718; 25058196; 25060182; 25060183; 25060184; 25060426; 25061232; 25061462; 25061463; 25061697; 25061698; 25061699; 25061700; 25061717; 25061959; 25061960; 25074416; 25076066; 25085166; 25088425; 25092346; 25092347; 25094253; 25101143; 25101144; 25101145; 25101146; 25109610; 25109673; 25123330; 25128690; 25151757; 251529; 251531; 25163992; 25164059; 25173432; 25178780; 25183088; 25183212; 25183216; 25185694; 25185729; 25186193; 25187070; 25188450; 25188720; 25192153; 25192154; 25194279; 25194280; 25198463; 25200839; 25201135; 25201228; 25201877; 25202439; 25202954; 25203498; 25203575; 25206501; 25209274; 25209457; 25210762; 25228280; 25231871; 25232491; 25232872; 25232874; 25233220; 25233383; 25234544; 25235746; 25235747; 25235749; 25240274; 25240443; 25240618; 25240942; 25241087; 25241109; 25241114; 25241361; 25243851; 25244377; 25244885; 25245485; 25245486; 25245519; 25245652; 25245814; 25245946; 25256616; 25311; 25322904; 25322905; 25322906; 25322907; 25322945; 25322946; 25322948; 25323005; 25323032; 25323033; 25323034; 25323035; 25323039; 25323112; 25323113; 25323114; 25323115; 25323116; 25323117; 25323127; 25323128; 25323129; 25323130; 253724; 256105; 256106; 25630328; 25630332; 25630334; 25630337; 25632541; 25632546; 25632550; 25749205; 25767094; 25789443; 25789445; 25789447; 25789449; 25791357; 25791358; 25791360; 25791362; 25791639; 25791640; 25791641; 25791642; 25791726; 25791727; 25791728; 25791729; 25795681; 25796437; 25796491; 25796492; 25796493; 25796494; 25796499; 25796500; 25796501; 25796502; 25796533; 25796534; 25796555; 25880661; 25880662; 259325; 259627; 259812; 259815; 26080931; 26087903; 26087904; 26087905; 26087906; 261549; 26176673; 26176674; 26176682; 26176683; 26202329; 26202331; 26202333; 26202335; 26202354; 26202355; 26202357; 26202359; 26202363; 26202365; 26202367; 26202369; 26296498; 26296499; 263034; 263620; 26437430; 26437431; 26437432; 26437433; 26470411; 26470417; 26470418; 26470514; 26470523; 26470525; 26470586; 26470587;

26470590; 26470591; 26470594; 26470602; 26470603; 26470604; 26470605; 26470608; 26470610; 26470624; 26470626; 26470628; 26470631; 26470634; 26470635; 26470636; 26470637; 26470801; 26470802; 26471086; 26495719; 26495723; 26495725; 26495727; 266056; 26722743; 26790551; 26791003; 26791083; 26793290; 2682; 26940520; 26967480; 26967481; 26985977; 26985978; 26985980; 26985981; 27191380; 27208; 27209; 27233725; 27233726; 27233728; 2723610; 2723872; 2724333; 2724334; 2724488; 2724551; 2724552; 27290; 2729834; 2731519; 2733453; 2733977; 2734150; 2734151; 2734157; 2735269; 2736963; 2737298; 2737299; 2737300; 2737301; 2737302; 2737303; 2737304; 274161; 274194; 27470369; 274; 275049; 275052; 27507142; 27507143; 27507144; 27507145; 275228; 2753347; 2754210; 2755077; 27551520; 27551524; 2758977; 2758978; 277452; 2781296; 2783883; 2783884; 2783885; 27867; 27876079; 27876080; 27876081; 27876082; 27876084; 27876085; 27876087; 27892; 278933; 281254; 28125550; 28125551; 2817448; 2827465; 282789; 2828526; 285097; 287439; 287687; 288227; 289151; 290479; 291344; 29251; 294564; 294; 295257; 297137; 29746; 298225; 298261; 298278; 298284; 298300; 29921593; 29933357; 29971343; 29975730; 29975731; 29975733; 29975734; 29976985; 29976994; 29979916; 29979920; 29979924; 29980565; 29980568; 29980572; 29980576; 29980621; 29980623; 29980627; 29980663; 29980665; 29980667; 29980671; 29980868; 29982418; 29982710; 29982711; 29982714; 29982715; 29982717; 29982718; 29982720; 29982721; 29982723; 29982725; 29982850; 29982854; 29982856; 29982859; 29982861; 29982862; 29983518; 29983520; 29983522; 29983524; 29983526; 29983529; 29983531; 29983830; 29983834; 29983838; 29983842; 29983845; 29983849; 29983852; 29983856; 29984439; 29987092; 29987094; 29987096; 29987099; 29987102; 29987104; 29987106; 29987108; 29987110; 29987112; 29987114; 29987116; 29987262; 29987265; 29987268; 29987272; 2998; 3000450; 3002158; 3002863; 3010374; 3012046; 3013082; 301637; 3018619; 3018771; 3023273; 3023610; 3023838; 3023839; 30248; 3025939; 3026803; 3026805; 3026833; 3026835; 3026837; 3027253; 303133; 3032092; 3032299; 3033866; 3034742; 3034828; 3035534; 3035590; 3036036; 3036169; 3036743; 3037063; 3037086; 3037556; 3050419; 307254; 30751; 3080770; 3080780; 3080835; 3080859; 3080922; 3080923; 3081390; 3081460; 3081923; 3081991; 3082029; 3082374; 3082398; 3082524; 3083134; 3084331; 3085020; 3085213; 3085365; 3085666; 3086428; 3086538; 309286; 3094792; 3095457; 309; 3102; 3103012; 31209; 31210; 31211; 31225; 31226; 31237; 31249; 31251; 31271; 31278; 31283; 31284; 31291; 3131005; 313121; 3131586; 313244; 313247; 313299; 31353; 31404; 314397; 314398; 314420; 314440; 315332; 315411; 316167; 316641; 316771; 32014; 321651; 321652; 322626; 322686; 3232583; 323498; 323587; 324113; 324224; 3246006; 3246058; 3246347; 3248861; 325954; 326265; 3291269; 3292082; 32953741; 33022; 3310512; 3317082; 3323164; 3347410; 335051; 3354132; 3355893; 3363163; 3368; 337268; 3372845; 3373861; 3373862; 3388732; 33934; 34173590; 3426; 3442834; 3450688; 3451358; 345715; 345716; 345717; 345721; 345967; 3459978; 3465607; 3465608; 3467658; 3471143; 3473; 34747; 34768; 3476988; 348011; 3481845; 3482394; 3494325; 349482; 349897; 3504251; 351383; 351658; 351821; 352020; 3525955; 352800; 3531744; 353794; 35408; 355056; 3554458; 3571964; 35821; 358296; 3591866; 359614; 3600220; 36081; 360951; 3617170; 3635972; 3637388; 3640407; 3648225; 3650195; 3661074; 3665684; 36688187; 36688188; 36688888; 36689003; 36689004; 36689005; 36689006; 36689007; 36689008; 36689009; 36689042; 36689043; 36689044; 36689045; 36689046; 36690486; 3707845; 3713156; 3719039; 3726022; 3745830; 3749604; 3754400; 3778793; 37888170; 3793749; 3796232; 3804937; 381430; 3832553; 38355134; 38362747; 38362755; 38363039; 38363043; 3837287; 3871; 3872; 387977; 3893; 38988608; 38988609; 38989258; 39222484; 4007; 40152688; 4033947; 403919; 40426991; 40428042; 40428622; 40428623; 40428789; 40429747; 40430063; 40459107; 40459108; 40459109; 40459110; 40463684; 40463685; 40467048; 40467049; 40467851; 40467852; 40467853; 40467874; 40467875; 40467876; 40467877; 40468191; 40468192; 40469155; 40469467; 40469468; 4047281; 40475156; 40475323; 40484567; 40484568; 40484569; 40488034; 40488036; 40488845; 40488846; 40488847; 40488848; 40488850; 40488851; 40488852; 40488865; 40488866; 40488867; 40488868; 40488869; 40488870; 40489049; 40489121; 40489122; 40489141; 40491308; 40492284; 40492285; 40493553; 40498267; 40498268; 40498269; 40498270; 40510104; 40521639; 40521640; 40522813; 40530756; 40530757; 40565718; 40565727; 40565751; 40565755; 40565921; 40565922; 40566223; 40571596; 40571610; 40571756; 40571768; 40579784; 40579793; 40599843; 40599861; 40609154; 40618672; 40622550; 40628235; 40628236; 40628237; 40628238; 40628239; 40628240; 40628241; 40628242; 40653939; 40768247; 40773639; 40773640; 40773641; 40773661; 40773662; 40773663; 40776052; 40805799; 40805800; 40805801; 40805802; 40805803; 40805849; 40805852; 40805853; 40815181; 40815182; 40823008; 4084218; 40953; 410079; 4108321; 4114916; 41211; 4127941; 4131; 4132952; 4134252; 4138758; 4149213; 41570; 4160231; 4168454; 4180364; 418610; 4188678; 4195243; 4197495; 4200772; 42013937; 42013938; 42051884; 42051885; 42051886; 42051887; 42051891; 42051892; 42051893; 42051894; 42051983; 42051984; 42051985; 42051986; 42066856; 42066894; 42066898; 42105547; 42105548; 42105550; 42105552; 42149278; 42149495; 4223808; 42311552; 42312550; 4239068; 42552636; 42552662; 42552694; 42552695; 42552696; 42552697; 42552711; 42552712; 42552713; 42552714; 42552736; 42552737; 42552738; 42552739; 42552748; 42560788; 42560789; 42560790; 42580057; 42580058; 42580059; 42580060; 42580141; 42580142; 42580143; 42597021; 42597285; 426001; 42603421; 42603423; 42609785; 42609819; 42612216; 42626680; 42626695; 42626713; 42626774; 42626816; 42627143; 42627493; 42627551; 42627552; 42627693; 42627719; 42631466; 42631467; 42631468; 42631469; 42631470; 42631471; 42637647; 42639268; 4276; 428040; 428899; 4302177; 430252; 43045; 43046; 43054191; 430654; 430899; 4310120; 43170129; 43170132; 43242758; 4330796; 43353222; 43369484; 43369833; 43399806; 4349664; 435434; 435728; 43579; 4359730; 43598; 4366627; 4369530; 43707618; 43709427; 43713034; 4373101; 437600; 437619; 437627; 437682; 4386393; 439178; 439186; 439192; 439193; 439304; 439312; 439315; 439341; 439353; 439357; 439451; 439494; 439507; 439512; 439537; 439538; 439541; 439557; 439559; 439583; 439637; 439645; 439660; 439680; 439709; 439867; 440141; 440142; 440240; 440299; 440315; 4403; 440446; 440545; 440655; 440656; 440658; 440995; 441005; 441014; 441032; 441033; 441034; 441035; 441036; 44119776; 44119910; 44119912; 44119950; 44120059; 44120115; 44120163; 441214; 44123306; 44123569; 44128626; 44128990; 44128991; 44130948; 44134557; 44135592; 44135834; 44137749; 441418; 441419; 441422; 441426; 441427; 441428; 441429; 441432; 441435; 441436; 441439; 44147198; 44147433; 441483; 441484; 441487; 44149764; 44153679; 44153717; 44153778; 4415668; 44158049; 44176417; 44182035; 44183763; 44184143; 44184144; 44191427; 44199360; 44199361; 44202883; 44207096; 44225826; 44226857; 44228788; 44228859; 44229138; 44229140; 44230153; 44232355; 442343; 442353; 442355; 44237111; 44237138; 44237281; 442384; 442393; 44239561; 44241312; 44241502; 44244950; 44246477; 44246478; 44246480; 44253454; 44257260; 44257339; 44258064; 44258071; 44258406; 44258413; 44258415; 44258420; 44258567; 44263858; 44268421; 44269524; 44277004; 44277176; 44278642; 44278644; 44278656; 44278923; 44282660; 44283360; 44283376; 44283391; 44283392; 44283502; 44283803; 44283828; 44283919; 44283944; 44283945; 44284170; 44287562; 44288860; 44302187; 44302221; 44302364; 44302433; 44302663; 44318407; 443248; 443265; 44328781; 44328785; 44335037; 44336256; 44338426; 4433883; 4434125; 44347898; 44348067; 443502; 443504; 44352348; 44352507; 44354166; 44354842; 44355779; 44355780; 44355849; 44355850; 44355851; 44357706; 44366290; 44366344; 44366547; 44366599; 44366616; 44369299; 44369384; 44369403; 44369406; 44369506; 44369578; 44373026; 44373664; 44379455; 44379476; 44379701; 44379702; 44379983; 44383136; 44383236; 44391931; 44398601; 44399922; 44399923; 44400560; 44402217; 44405732; 44408268; 44408385; 44408489; 44408498; 444176; 44420311; 44420320; 44420335; 44420336; 44420340; 44420352; 444212; 444314; 44442323; 44442325; 44457530; 44462511; 44462514; 44462515; 44462516; 44465655; 444685;

44472831; 44473138; 44473139; 444792; 445070; 445093; 44517915; 445187; 445226; 445238; 445316; 445318; 44544430; 44544536; 445463; 4454759; 44548660; 44551308; 44551309; 44551689; 44551690; 44551691; 44551692; 44558885; 44560080; 44560081; 445638; 445639; 44567462; 44567463; 44574756; 44575502; 4457602; 445801; 445826; 445864; 445902; 445913; 44591695; 44591721; 44591722; 445996; 44602412; 44602452; 446048; 446090; 446103; 44611138; 4462283; 44629624; 44629828; 44629916; 44629928; 44629931; 44629937; 44630092; 44630390; 446397; 446398; 446488; 446578; 446584; 446585; 446588; 446673; 446882; 4469125; 4469582; 4469584; 4469809; 447027; 447084; 4470982; 447139; 44715199; 44717407; 44717408; 44717409; 44717410; 44717414; 44717415; 44717416; 44717420; 44717700; 44717759; 447198; 44719912; 447231; 44723594; 447270; 4473084; 4474230; 4474425; 4474578; 44754938; 44754939; 44754940; 44755000; 44755048; 44755069; 44778202; 4477868; 447832; 4479103; 447916; 4479241; 448165; 448170; 448173; 448208; 448210; 448244; 44828743; 448294; 448323; 448324; 448325; 448374; 448378; 448388; 448615; 448683; 448684; 4487010; 448702; 448789; 448895; 44889986; 448925; 4489683; 449225; 449370; 449469; 449599; 45006121; 450201; 45033102; 45038102; 45038103; 45038817; 45038818; 45038819; 45038951; 45039313; 45039356; 45039357; 45039595; 45039641; 45039680; 45039681; 45039682; 45039826; 45039910; 45040427; 45040428; 45040429; 45040430; 45040431; 45040476; 450503; 45050701; 45051848; 45051849; 45051850; 45052211; 45052261; 45071976; 45071983; 45073472; 45076917; 45076930; 45076931; 45076933; 450776; 45093622; 45093626; 45093627; 45093628; 45093641; 45093699; 45093719; 45093720; 45100758; 45100759; 45109759; 45109765; 45109775; 45109778; 45109790; 45109797; 45109800; 45109807; 45109809; 45109814; 45109825; 45110040; 45110168; 45110169; 451187; 451188; 451189; 451195; 451196; 451197; 451198; 451200; 45157509; 451597; 452245; 4526533; 45266829; 45266831; 45266868; 45281236; 453548; 45356796; 45357252; 45358029; 45360385; 453; 45479484; 45480566; 45480573; 45480574; 45480582; 45480584; 45480590; 45480625; 45480626; 45482654; 455147; 4582981; 4586058; 45923771; 45933847; 459494; 46173743; 46173764; 46187997; 46190835; 46202857; 46202858; 46202859; 46202860; 46202935; 46203185; 46217193; 46222220; 46224569; 46224570; 46224590; 46224603; 4623424; 4629011; 4629153; 4630703; 4631010; 466397; 4673050; 46780090; 46780091; 46780441; 46780447; 46780483; 46781088; 46781665; 46781666; 46781992; 46782101; 46782468; 46782906; 46782953; 46782954; 46849869; 46863234; 46873926; 46874321; 46874390; 46874396; 46874427; 46874455; 46874558; 46874702; 46875018; 46875160; 46875232; 46875281; 46875597; 46876125; 46876203; 46876227; 46876229; 46876275; 46876278; 46876694; 46876753; 46876893; 46876894; 46876921; 46876922; 46876985; 46877005; 46877073; 46877074; 46877121; 46877217; 46877231; 46878518; 46878519; 46878532; 46878549; 46878556; 46897877; 46905263; 46905267; 46912129; 46916361; 46916362; 46916363; 46916401; 46916402; 46916403; 46916404; 46916405; 46916406; 46916407; 46931125; 46936190; 46936216; 46936223; 46936294; 46936364; 46936389; 46936458; 46936484; 46936495; 46936584; 46936616; 46936629; 46936648; 46936762; 46936810; 46936820; 46936892; 46936904; 46936958; 46937000; 4693886; 46941523; 471580; 471581; 474; 4798; 493591; 4939316; 4943005; 494311; 494348; 4947; 4953775; 4958619; 49764355; 49766505; 49766620; 49766621; 4979092; 49792020; 498201; 49820613; 49822387; 49822388; 49824738; 49825415; 49837892; 49851295; 49852317; 49866464; 49866558; 49866637; 49871294; 49873679; 49873685; 49873686; 49873695; 49873697; 4995412; 5018391; 5044262; 5050297; 5059891; 506590; 506591; 5066234; 5076194; 5082332; 5084146; 50896765; 50908570; 50909874; 50909884; 50916477; 50916478; 50916646; 50916647; 50916648; 50930791; 50936912; 50939543; 50988013; 50989701; 50990442; 50990598; 50996743; 51003418; 51003661; 5104362; 51051214; 51064083; 51136478; 5113708; 5118446; 5118447; 511916; 511918; 51340644; 51340649; 51340650; 51340651; 51340669; 51340682; 51351764; 51351765; 51351766; 51351815; 51351816; 51351817; 51352722; 51380928; 51396505; 51397474; 51397475; 51397495; 51398009; 51398010; 51398011; 51398012; 51398013; 51404081; 51404248; 51404249; 51404250; 51404251; 51404257; 51404258; 5140436; 51407476; 51407477; 5143; 51488153; 51488154; 51536919; 51536920; 51536921; 51536926; 51536927; 51536928; 51536929; 515377; 515384; 51546380; 51548332; 51548334; 51573356; 51573357; 51576595; 51576810; 51579139; 51579421; 516810; 516811; 51683431; 517637; 51777301; 51777307; 51777308; 51777325; 51777337; 51777338; 51777342; 517827; 517970; 5179954; 518472; 5184980; 518810; 51921893; 51921894; 51921895; 51921896; 519361; 519382; 5194599; 519564; 519662; 519764; 519857; 520148; 520296; 520455; 52083986; 52083988; 52083990; 520895; 521185; 521253; 521326; 521332; 521334; 521380; 522171; 522266; 522673; 522676; 5228854; 5229910; 5230271; 5232275; 5232276; 5232277; 5232278; 5232295; 52325763; 52325764; 5242923; 524461; 5248464; 524975; 525032; 5258034; 526152; 526616; 526624; 526993; 5270647; 5271246; 5271249; 5271250; 5271252; 527144; 527268; 5273464; 5273465; 5273467; 5273469; 5273513; 5273515; 527498; 5275726; 5280341; 5280435; 5280450; 5280462; 5280598; 5280934; 5281162; 5281515; 5281516; 5281517; 5281519; 5281520; 5281522; 5281531; 5281534; 5281535; 5281659; 5281929; 5281; 5282108; 5282109; 5282110; 5282184; 5282231; 5282800; 5283361; 5283363; 5283366; 5283368; 5283381; 5283383; 5284369; 5284421; 5284499; 5284507; 5284656; 528708; 5287473; 528763; 5287649; 5287770; 5287784; 5287809; 5287852; 5287883; 5288074; 5288136; 5288150; 5288166; 5288211; 5288275; 5288324; 5288340; 5288341; 5288350; 5288399; 5288400; 5288423; 5288433; 5288436; 5288438; 5288439; 5288729; 5288739; 5288746; 5288749; 5288760; 5288763; 5288768; 5288769; 5288771; 5288773; 5289595; 5289596; 5289598; 5289601; 5289603; 5289606; 5289608; 529052; 529056; 529068; 52914458; 52914495; 52916942; 52916943; 52921598; 52921627; 52921646; 52921662; 529388; 52939419; 52939789; 529496; 5311443; 531518; 531527; 5315346; 531538; 5315594; 5315912; 5315941; 5316068; 5316444; 5316713; 5316845; 5317246; 5317259; 5317319; 5317407; 5317570; 5318103; 5318557; 5318734; 5319249; 5319264; 5319595; 5319691; 5319706; 5319879; 5320128; 5320530; 5321399; 5321888; 5321889; 5321995; 5321996; 5321997; 5322111; 53239715; 53239717; 53239719; 53239720; 5326987; 5327005; 53298384; 53301851; 5331187; 53319148; 53338773; 533390; 53346513; 53381393; 53381395; 53384462; 53384471; 53384542; 5338463; 53384900; 53384914; 53384915; 53384916; 53384918; 53384930; 53392809; 53393646; 53393661; 53393674; 53393774; 53413358; 53413479; 53420613; 53422473; 53423642; 53424808; 534429; 53461945; 53462054; 53463499; 53463521; 53463645; 53463776; 53469479; 53470622; 53477592; 53477593; 53477647; 53477733; 53477771; 53477780; 53481006; 53481024; 5352140; 5352162; 5352428; 5352435; 5352436; 5352437; 5352459; 5352460; 5352481; 5352543; 5352830; 5352845; 5352905; 5352973; 5352974; 5353011; 5354238; 5354280; 5354882; 5355293; 5355850; 5355855; 5355856; 5355858; 5356544; 5357002; 5357153; 5362566; 5362567; 5362579; 5362581; 5362583; 5362591; 5362592; 5362595; 5363120; 5363141; 5363269; 5363289; 5363426; 5363741; 5363817; 5363876; 5363896; 5364430; 5364438; 53645503; 53646005; 53646750; 53650133; 5365069; 5365136; 5365703; 5365794; 5365850; 5365910; 5365916; 5365976; 5365982; 53659831; 53659832; 5365991; 5365992; 5366044; 5366074; 5366077; 5366244; 53672722; 5367369; 5367530; 53675565; 5367657; 5367665; 5367679; 5367681; 5367682; 5367698; 5367706; 5367719; 5367752; 5367767; 5367785; 5367848; 53679620; 5368064; 53681435; 53681436; 53681437; 53681439; 53682947; 53688435; 5369371; 53693898; 53693900; 5369459; 53698760; 536987; 53699899; 5370052; 53701918; 5370599; 5370602; 53709419; 5371002; 5371084; 53712816; 5371723; 53718249; 5372174; 5372195; 53722046; 5372353; 53724373; 53733134; 5373603; 5373729; 5373780; 53743681; 5374527; 53748486; 53748487; 53748489; 53748490; 53748491; 53748492; 5375218; 53766604; 53769462; 537754; 53779437; 53780324; 53782679; 53782680; 53782685; 53782686; 53782687; 53782688; 53782690; 53782691; 53782694; 53782699; 53782700; 53782701; 53782702; 53782703; 53782704; 53782706; 53782707; 53785320; 53788624; 537892; 537903; 53790668; 53790670;

537918; 537931; 53796906; 53799656; 53799657; 53802214; 53802215; 53802216; 53823656; 5382490; 53825654; 53830004; 53831058; 53836010; 5383634; 53839849; 53839850; 53849308; 53850676; 53850677; 538506; 53852252; 53853553; 5385498; 5385520; 53857340; 53862083; 53865173; 53866836; 53868634; 53868635; 53868636; 53868637; 53868638; 53868639; 53869986; 53871970; 53871972; 53871973; 53871974; 53871975; 53871976; 53871977; 53871978; 53871979; 53871980; 538756; 53877958; 53877959; 53878077; 53879659; 53880293; 53880294; 53880295; 53880296; 53881496; 53883286; 53883512; 53883513; 53883544; 5388696; 53893119; 53893132; 53893133; 53897792; 53900994; 53906336; 53906337; 539111; 53913978; 53913979; 53914603; 53921953; 53923841; 53927285; 53933732; 53942697; 53942698; 53949092; 53950108; 53951057; 539829; 53985909; 53985910; 53989233; 539902; 539909; 54002986; 54006130; 540226; 54033044; 54045554; 54048924; 54051293; 54055604; 54056768; 54063114; 54065548; 54065549; 54075063; 54087631; 54087632; 54089947; 54090800; 54097206; 54108221; 54108785; 54108786; 54111948; 54112424; 54114556; 54121424; 54129978; 54130119; 54133627; 54145635; 54146274; 54146531; 54148725; 54149509; 54152451; 54159177; 54160479; 54160480; 54162515; 54170885; 54173895; 54179255; 54179999; 54194100; 54208; 54212763; 54222646; 54229533; 54230212; 54230213; 54230563; 542369; 54237266; 54249633; 54249634; 54253685; 54259910; 54259911; 54263984; 54267583; 54270338; 54274762; 54281757; 54283243; 54292281; 54293242; 54297884; 54297885; 54301727; 54301728; 54301978; 54305663; 54310221; 54321691; 54322905; 54347226; 54351452; 54351900; 54355510; 543579; 54358386; 54358387; 54361181; 54361182; 54362057; 54362427; 54362428; 54362893; 54363528; 54364205; 54374854; 54384642; 54389263; 543912; 54395770; 54402619; 54405023; 54405359; 54405360; 54405361; 54405362; 54407177; 54415622; 54419765; 54420529; 54421096; 54424873; 54431436; 54434891; 54434931; 54436675; 54441515; 54441516; 54441517; 54445182; 54453238; 54453901; 54456433; 54456844; 54456846; 54467119; 54473728; 54478248; 54486861; 54488379; 54498735; 54503417; 54503633; 54503634; 54504194; 54524082; 54527287; 54527288; 54527504; 54540528; 54540529; 54541830; 54544180; 54544695; 54547403; 54549170; 54550870; 54561130; 54561131; 54561132; 54564397; 54564621; 54567100; 54567101; 54572099; 54572698; 54577411; 545865; 5459355; 5459879; 5459956; 5459986; 5460024; 5460026; 5460044; 5460054; 5460126; 5460226; 5460248; 54603882; 5461127; 54611954; 546235; 5462988; 5463277; 5463485; 5463840; 5463904; 5463905; 5463906; 5463909; 5463911; 5463913; 5463914; 5463954; 5464164; 546577; 546617; 54670067; 54675810; 54676825; 547071; 54732806; 54751688; 54759162; 5479772; 5479783; 548228; 548230; 5484202; 5486682; 5486715; 5487627; 549546; 55079885; 551449; 551496; 552029; 552038; 552058; 552214; 552226; 552303; 55250368; 55295881; 552964; 55297539; 553354; 554076; 554158; 554180; 5541; 554202; 556940; 557446; 558123; 5590; 559389; 559859; 560074; 560109; 560145; 560150; 560156; 560175; 560470; 560516; 560575; 5634; 564550; 565690; 56589766; 56589851; 56596546; 56603822; 56604911; 56605751; 56606153; 56608803; 56609660; 56609717; 56611045; 56611277; 56613364; 56613399; 56614780; 56616753; 56618775; 56618939; 56618957; 56621519; 56622076; 56623053; 56624369; 56625777; 56627161; 56628147; 56630667; 56632570; 56632603; 56637006; 56637337; 56650093; 56650094; 56657680; 56666128; 56683010; 568093; 568193; 568208; 568210; 56832479; 56842518; 56843050; 56845432; 56845689; 56845990; 56845995; 56846079; 56846177; 56846218; 56846237; 56846239; 56846514; 56846519; 56846584; 56846618; 568507; 568522; 568523; 568530; 568666; 569214; 56924143; 56928211; 56961467; 56971649; 56973539; 56973688; 56973689; 56973690; 56978075; 56980840; 56983870; 56988235; 56992506; 56992563; 56994182; 56997393; 57000527; 57002308; 57004472; 57004639; 57008544; 57010447; 57012832; 57013856; 57015141; 5702004; 57024401; 57025446; 5702665; 57031297; 57046032; 57046035; 57046140; 5705112; 570574; 57057775; 57065080; 57066580; 57067244; 57074459; 57077522; 57077765; 57078006; 57080069; 57089454; 57090722; 57090934; 57095774; 57097467; 57106630; 57111376; 57119450; 57122031; 57122883; 57123435; 57128667; 57131490; 57133406; 57138681; 57142780; 57143133; 57144513; 57151025; 57151100; 57157186; 57159962; 57169063; 57169337; 5717479; 57175863; 57180332; 57181456; 57183118; 57187433; 57190132; 57193742; 57194632; 57195169; 57195345; 57197748; 57198574; 57198737; 57199213; 57203161; 57207201; 57231876; 57234304; 57238170; 57242813; 57245434; 57253673; 57263958; 57264900; 57265344; 57268732; 57272854; 57274391; 57278218; 57278611; 57280352; 57288387; 57294184; 57294977; 57297101; 57298281; 57306402; 57310038; 57310776; 57311216; 57317456; 57319120; 57320331; 57320554; 57321363; 57324057; 57324405; 57331769; 57338422; 57338423; 57339194; 57345928; 57346909; 57347734; 57348607; 57348610; 57349502; 57349564; 57350398; 57352382; 57353021; 57353502; 57355543; 57357963; 57358045; 57358095; 57358374; 57358602; 57369524; 57369599; 57370084; 57375809; 57376074; 57384727; 57384973; 57389496; 57399808; 57403324; 57411650; 5742864; 57449163; 57454431; 57459281; 57461573; 57464725; 57464814; 57464831; 57465144; 57467888; 57468051; 57469225; 57470296; 57474150; 57474151; 57483300; 57483528; 57483670; 57483671; 5748368; 57483912; 57484151; 57484152; 57488047; 57488722; 57493519; 57494380; 57496803; 57500615; 57503973; 57503980; 57503984; 57504852; 5750612; 57506881; 5750805; 57509241; 57511168; 57514608; 57515112; 57540272; 57557846; 57598586; 57614969; 57614970; 57621284; 57621285; 57621288; 57621291; 576240; 576263; 57642148; 57642149; 57642164; 57642168; 57642176; 57642182; 57677216; 57679915; 57691826; 57691841; 57693892; 57695700; 57695701; 57695702; 57695703; 57695704; 57701550; 57701551; 57702372; 57704059; 577126; 5771642; 5772036; 57735986; 57745769; 57751761; 57775944; 57806008; 57806013; 5780; 57823250; 57823254; 57824481; 57824588; 57824589; 57824592; 57828600; 57828601; 57832405; 57834566; 57834567; 57834577; 578423; 5784388; 57864639; 57879890; 57879893; 57879900; 57879902; 57879908; 57885068; 57893250; 57895777; 57912333; 579336; 57939995; 5793; 57940286; 57972066; 57973135; 5797328; 57975703; 57994455; 57997981; 58013936; 58013940; 58019724; 58030924; 5803450; 58040644; 58069220; 58069229; 58090235; 58132950; 58132951; 58132958; 58165917; 58165919; 58165920; 58183850; 58183855; 58183865; 58183867; 58183871; 58213413; 58213449; 58213454; 58213457; 58213459; 58265153; 58265160; 58265166; 58265178; 58265180; 58265190; 58265196; 58265369; 58265414; 58283927; 58283928; 58300638; 58301384; 58307996; 58307999; 58312001; 58312006; 58312008; 58332685; 58351750; 5839423; 5842730; 58444029; 58463494; 58468411; 58481202; 58482138; 58488956; 58498438; 5853596; 58538088; 5853952; 5857493; 58590346; 585911; 58594768; 58595959; 58606984; 58618581; 5864137; 58645386; 58645387; 58654615; 58666267; 58669405; 58687444; 58687451; 5869600; 58699843; 58700677; 58701490; 58761794; 58761797; 58761800; 58761802; 58776116; 58798222; 58805971; 58814650; 58831415; 58833558; 58862400; 58863022; 58863068; 58905859; 58922336; 589365; 58958914; 58958922; 58958924; 58969552; 58971923; 58971924; 58981618; 58991510; 59002381; 59015080; 59022366; 59024526; 59034276; 59034985; 59036328; 59039607; 59040622; 59048761; 59049257; 59057550; 59059868; 59059876; 59059877; 59059978; 59074633; 59083882; 59096512; 59097773; 59099798; 59101134; 59103604; 59104448; 59105062; 59105109; 59110138; 59111037; 59112131; 59119178; 59119179; 59125088; 59140762; 59146659; 59149367; 59149376; 59149384; 59149387; 59153838; 59153839; 59154466; 59163159; 59169495; 59172547; 59185624; 59202444; 59202804; 59284384; 5934745; 59348397; 59349386; 59350540; 59350541; 59360925; 5936462; 59378726; 59383280; 59391721; 59400491; 59441449; 59441477; 59445439; 59451317; 59460532; 5947009; 59476548; 59483380; 59489508; 59491123; 59491272; 59503407; 59503411; 59515148; 595479; 595480; 595518; 59561892; 5956805; 59574095; 59574110; 59574111; 59574112; 59574114; 59575442; 59575448; 59629099; 59646259; 59646263; 59646264; 59660913; 59661875; 59664699; 59709179; 59722278; 59760383; 59760917; 59779296; 59782654; 59805006; 59805009; 59806223; 59816556; 59819893; 59848323; 59848324; 59848327; 59848329; 59848331;

59848338; 5984; 59859430; 59862514; 59875236; 59878724; 59883195; 59886072; 5988; 59893355; 59894669; 59904077; 59904081; 59904158; 59906316; 59907918; 59909322; 59909325; 59912513; 59912514; 59920722; 59921549; 59927511; 59932066; 59932088; 59935106; 59936830; 59946029; 59957382; 59960729; 59960750; 59965103; 59966141; 59967020; 59967032; 59968218; 59968988; 59969557; 59979821; 60000525; 60000688; 60000689; 60018944; 60041168; 60042689; 60042747; 60046036; 60052896; 6005821; 60066215; 60070412; 60070413; 60076984; 60078498; 60078499; 60078505; 60078646; 60078648; 60078967; 60078972; 6008212; 60087033; 60093952; 60101813; 60109619; 60122943; 60122944; 60122947; 60122951; 60156292; 60161240; 60162291; 60196286; 60209635; 60209767; 6022; 602788; 6036; 604; 6050; 605629; 6076971; 6076; 6083; 6086514; 609480; 60961; 60968; 60979; 60985; 60997; 60998; 60999; 61005; 61007; 61014; 61015; 6101770; 6101927; 61024; 6102530; 6102790; 61027; 6102839; 61030; 61031; 61035; 61040; 61048; 61050; 61051; 61052; 61079; 61085; 61092; 61093; 61096; 61098; 61125; 61129; 6114390; 61156; 61159; 61185; 61192; 61204; 61206; 61229; 61262; 61270; 61290; 61294; 61303; 61304; 61331; 61342; 6134; 61361; 61384; 61386; 61401; 61402; 61403; 61404; 61405; 61406; 61407; 61410; 61412; 61415; 61416; 61417; 61431; 61434; 61435; 61446; 61447; 61451; 61452; 61455; 61457; 61464; 61466; 61516; 61517; 61521; 61524; 61585; 61586; 6165108; 61658; 61659; 61712; 61746; 61780; 61781; 61834; 61835; 61892; 61953; 61954; 61955; 61979; 61980; 61982; 61984; 61985; 61989; 61990; 61997; 62007; 62013; 62016; 62017; 62080; 62090; 62092; 62098; 620; 62105; 62106; 62126; 62127; 62151; 62182; 62183; 621; 62217; 62240; 62247; 62321; 6232217; 62328; 62329; 62332; 62333; 62336; 62349236; 62362; 6240234; 62469; 62473244; 62475040; 62475765; 62490; 62506157; 62506158; 62506160; 62506324; 62506983; 62507315; 62507811; 62507976; 62508308; 62508309; 62508473; 62508474; 62508637; 6251; 62553; 6255; 62566; 62578; 62580; 62592; 6259976; 62653; 62654; 62661; 6284401; 62852; 62876; 62901; 62907; 6321330; 6323336; 6323395; 6325163; 6325432; 6325503; 6326183; 6328565; 6336472; 6337812; 6338568; 63456625; 6364671; 6366072; 6366434; 636822; 63686330; 6370843; 637195; 637312; 637517; 6375937; 637616; 637758; 6378927; 6379211; 638014; 638024; 638034; 638036; 638042; 638500; 6386037; 6400264; 6409631; 6409820; 6409821; 641423; 641622; 641623; 641624; 641625; 641626; 641627; 641628; 641681; 6419581; 6419706; 6420052; 6420200; 642117; 6421197; 6421199; 6421200; 6421281; 6421426; 6421445; 6421450; 6421459; 6421834; 642376; 6426625; 6426626; 6426631; 6426632; 6427012; 6428089; 6428535; 6428538; 6428936; 6428986; 6429039; 6429195; 6429289; 6429302; 6430180; 6430188; 6430738; 6430754; 6430759; 6430833; 6430879; 6431015; 6432005; 6432404; 6432703; 6432903; 6432934; 6433150; 6433214; 6435825; 6435834; 6435835; 6435837; 6435839; 6435861; 6435876; 6435882; 6435903; 6435914; 6436336; 6436374; 6436375; 6436439; 6436624; 6437428; 6437467; 643757; 6437648; 6437720; 644104; 644174; 644237; 644264; 6445771; 6448; 6450812; 6450832; 6451569; 6452874; 6453284; 64537000; 6454902; 6454927; 6455074; 64574070; 64689; 64731; 64763; 64791; 6481827; 64947; 6501; 65041; 6505; 6506; 6508206; 6532; 6536998; 6537093; 6537492; 65428; 6552007; 6552084; 65575; 655982; 6560132; 6560213; 6560357; 6560765; 656504; 656512; 656941; 656985; 657030; 657035; 657055; 657130; 6589427; 6602110; 6602503; 6603718; 6604063; 6604194; 6604825; 6604826; 6610262; 6610315; 66233564; 66277787; 6631; 6636316; 66509015; 66509130; 66509209; 66563684; 66571414; 66595; 66599114; 66599367; 66600964; 66601034; 66622690; 66622691; 66629241; 66629908; 66634210; 66647886; 66654; 66687020; 66688498; 66696082; 66696125; 66696263; 6669; 66726720; 66728469; 66747170; 66751; 66764679; 66764715; 66764998; 66782500; 66788661; 66793680; 66793752; 66794013; 66795785; 66795817; 66799368; 66799626; 66810357; 66814858; 66822116; 66828; 66831164; 66831165; 66835956; 66835988; 66846131; 66846132; 66846134; 66874017; 66880466; 66880468; 66880470; 66887108; 66899979; 66916523; 66921550; 66959484; 66971997; 66985221; 66985403; 67005; 67018288; 67043323; 67062982; 67063005; 67063007; 67063008; 67063097; 67072633; 67072634; 67111368; 67116; 67128679; 6713130; 6713792; 6713976; 6714006; 67153144; 67158675; 67159212; 67172434; 67176901; 67224923; 67225311; 67233004; 67233006; 67235049; 67243529; 67245415; 67247301; 67257954; 67258870; 67265595; 673160; 67320492; 67327847; 67329001; 67329046; 67330028; 67349037; 67350816; 67350; 67351818; 67351891; 67366665; 67375559; 67409435; 67410024; 67410034; 67410584; 67410585; 67410603; 67410605; 67410620; 67410629; 67410630; 67410681; 67410683; 67410760; 67410919; 67410942; 67410946; 67410953; 67410978; 67411005; 67411006; 67411125; 67411650; 67411725; 67411764; 67411765; 67412095; 67412096; 67412100; 67412128; 67412168; 67412170; 67412175; 67412341; 67412488; 67412599; 67412812; 67413029; 67413033; 67413049; 67413139; 67413170; 67413211; 67413670; 67413707; 67413730; 67414353; 67424980; 67441279; 67450830; 67472765; 67476019; 67476020; 67477525; 67518639; 67519198; 67521; 67533353; 67533916; 67536612; 67536616; 67553847; 67553850; 67555473; 67567238; 67577048; 67578402; 67592695; 67592697; 67594286; 67612895; 67613034; 67613262; 67615000; 67615455; 67632183; 67633216; 67633278; 67633301; 67633620; 67636320; 67636345; 67640490; 67640900; 67641695; 67641738; 67641933; 67642211; 67645588; 67649153; 67666554; 67673287; 67683356; 67687200; 67687725; 67689648; 67689652; 67699046; 67699946; 67701638; 67702969; 67706075; 67706638; 67706683; 67711343; 67720038; 67723193; 67723200; 67724956; 67726479; 67742010; 67744739; 67749698; 67749700; 67766483; 67770886; 67780106; 67780588; 67784620; 67785483; 67786945; 67787076; 67787236; 67787489; 67789572; 67795854; 67797725; 67810472; 67814152; 67814325; 67814862; 67814867; 67817666; 67819105; 67821681; 67825991; 67826002; 67826707; 67829356; 6782; 67831081; 67831440; 67839156; 67856350; 67862674; 67877528; 67889287; 67894176; 67902260; 67909796; 67918867; 67918907; 67918967; 67919076; 67919082; 67920059; 67921484; 67921485; 67921489; 67938791; 67939114; 67944215; 67944290; 67950444; 67952436; 67974753; 67979606; 67987016; 67987943; 67996277; 68006605; 68009482; 68009554; 68009582; 68012968; 68013637; 68015385; 68016903; 68017365; 68018077; 68023710; 68024271; 68024276; 68026771; 68026943; 68028651; 68029373; 68030084; 68030194; 68030465; 68047758; 68047764; 68053980; 68057; 68074721; 68075241; 68080051; 68080207; 68097423; 68110212; 68112028; 68114367; 68119541; 68145569; 68149497; 68151892; 68151893; 68152163; 68152165; 68152679; 68164441; 68167579; 68178188; 68186041; 68196072; 68196650; 68197081; 68200247; 68201113; 68206404; 68208485; 68214784; 68216895; 68216896; 68219335; 68219336; 68227641; 68229136; 68229180; 68233016; 68240371; 68243417; 68250129; 68253465; 68253539; 68255877; 68256528; 68259401; 68259849; 68265923; 68274; 68300528; 68304742; 68316879; 68317013; 68317863; 68317864; 68317865; 68321388; 68321398; 68324546; 68324547; 68324676; 68324677; 68326572; 68334110; 68335909; 68336367; 68338374; 68340058; 68340059; 68342374; 68356185; 68356186; 68356187; 68360668; 68361009; 68361010; 68361013; 68361014; 68366113; 68366897; 68366899; 68366976; 68366980; 68372215; 68372829; 68375596; 68375600; 68375602; 68375608; 68376099; 68387821; 68393076; 68397078; 68397083; 68397084; 68397232; 68397264; 68397312; 68403210; 68415506; 68417406; 68417675; 68438564; 68438565; 68443554; 68445272; 68482; 68503264; 6850716; 6850753; 6851304; 68519390; 68519886; 68524108; 68526191; 68534174; 68536504; 68544316; 6857367; 6857369; 6857417; 6857563; 6857564; 6857680; 6857681; 68578936; 6858240; 68592972; 68621102; 68628058; 68661802; 68663294; 68668347; 68730400; 6873; 6874; 68753715; 68762295; 68762298; 68794529; 68808470; 68812359; 68821824; 68821849; 68822540; 68825009; 68825010; 68825027; 68829936; 68841466; 68853745; 68854491; 68855383; 68867749; 68875807; 68881485; 68901676; 68903797; 68905554; 68913217; 68968644; 68973174; 68976688; 68977337; 68977340; 68977341; 68977346; 68977357; 68977360; 68977393; 68977777; 68977795; 68978165; 68978206; 68978419; 68978425; 68979094; 68979673; 68980122; 68980305; 68980307; 68980308; 68980316; 68980803; 68981109; 68982979; 68984637; 68984672; 68985879; 68986673; 68986734; 68986737; 68998565; 69013070; 69013072;

69013279; 69017244; 69017249; 69028286; 69029979; 69030580; 69030583; 69031034; 69032256; 69032917; 6904; 69052094; 69052440; 69059987; 69077688; 69079620; 69099845; 69099846; 69107323; 69112276; 69117899; 69131502; 6913230; 69135463; 6915737; 69164048; 69173362; 69173730; 69175355; 6918391; 69188119; 69196037; 69196040; 69205326; 69206104; 69220633; 69279541; 69301022; 69301023; 69303687; 6930; 69332325; 69350; 69390674; 69471726; 69474952; 69475300; 69475959; 69477460; 69490632; 69497010; 6949; 69501288; 69525402; 69525405; 69525406; 69528681; 69534922; 69537435; 69537438; 69567545; 69567546; 69569193; 69569196; 69572815; 69574979; 69577820; 69582370; 69605895; 69615520; 69620801; 69620814; 69620815; 69622414; 69640475; 69643591; 69647610; 69649038; 69686048; 6971003; 6971007; 6971016; 6971020; 6971021; 6971096; 6971097; 6971098; 6971099; 6971278; 69714643; 69715361; 69742873; 69746937; 69747942; 69758213; 69767294; 6976821; 69772419; 69772579; 69772586; 69772587; 69825678; 69825730; 6985535; 69860867; 69874022; 69874582; 69877289; 69880231; 69889658; 69890236; 69903934; 69904223; 69905368; 69905492; 69905556; 69910058; 69910059; 69911057; 69911115; 69911116; 69912472; 69912473; 69912670; 69912672; 6992021; 6992060; 6992061; 6992062; 6992063; 6992064; 6992065; 6992066; 6992084; 6992107; 69921592; 6992262; 6992309; 6992310; 6992311; 6992312; 6992313; 6992314; 6992315; 6992316; 6992394; 6992395; 6992396; 6992397; 6992398; 6992399; 6992400; 6992493; 6992494; 6992642; 6992643; 6993119; 6993120; 6993372; 6993373; 6993580; 69937083; 69942121; 69945324; 6994673; 6994674; 6994675; 6994676; 6994677; 6994678; 6994679; 6994680; 69967075; 6997328; 69978571; 69988856; 6999110; 7000063; 7000064; 7000182; 70004140; 70004295; 70004838; 70005072; 70005431; 70005453; 70006106; 70006157; 70008427; 70010706; 70014676; 70023099; 7009595; 7009596; 7009597; 7009598; 7010136; 7010137; 70104268; 70104270; 70117899; 70119455; 70123921; 70124491; 70140441; 70141988; 70148970; 7014910; 7014911; 7015111; 7015112; 70152886; 7015293; 7016046; 7016047; 7016061; 7016101; 70172077; 70179573; 70179816; 70180134; 70180468; 70180535; 7018164; 70183479; 7018876; 70189621; 7019841; 7020089; 7020090; 7020240; 7020241; 7020414; 7020415; 7020495; 7020636; 7020637; 70211039; 70213063; 7021781; 7021823; 7021824; 70226107; 70237473; 70261299; 70261302; 70269959; 70272868; 70273202; 70273203; 70276722; 70288909; 70291272; 70305770; 70307233; 7031900; 7031901; 7031902; 7031903; 70321050; 70322858; 70331340; 70331648; 70363151; 70363270; 70385390; 70386981; 70386984; 70387533; 70387563; 70389207; 70389398; 7039483; 7039484; 7039485; 7039486; 70395784; 70396989; 70400129; 704002; 704003; 7040251; 70403533; 70405797; 70422664; 70433391; 70434780; 7043897; 70439872; 7044038; 70452145; 7045801; 70459479; 70459579; 70459695; 70459753; 70460000; 70460192; 70460212; 70460460; 70460974; 70464362; 70466709; 70476280; 70476596; 70478070; 7048607; 70486134; 70489348; 70519217; 70519348; 70519515; 70523329; 70524924; 70525127; 70543179; 70543261; 70554896; 70563644; 70564972; 70564973; 7057263; 7057264; 7057265; 7057266; 70576554; 70576573; 70578623; 7058055; 7058175; 7058176; 70582949; 70583822; 70584129; 70584563; 7059469; 7059470; 7059571; 70621337; 70640512; 70651226; 70660211; 70671022; 70671024; 70671195; 70671469; 7067544; 7067545; 7067546; 7067854; 7067855; 70681492; 70683051; 70698306; 7076609; 7077319; 7077320; 70789225; 70793475; 70795740; 70796534; 70800569; 70800737; 70801708; 70801711; 70806232; 70813850; 70822531; 70822967; 70822971; 70822972; 70822986; 70822988; 70823002; 70824171; 70825865; 70828496; 70829558; 70832729; 70837736; 70839899; 70839901; 70841419; 70841423; 70841899; 70845657; 70846687; 70846688; 70846691; 70846723; 70846753; 70849134; 70863586; 70864328; 70864857; 70865245; 70865567; 70872441; 70872953; 70873588; 70875643; 70876119; 7087742; 7087743; 70878456; 70878457; 70879692; 70884982; 70885145; 70889269; 70896299; 70896300; 70898408; 70899487; 70899488; 70899489; 70902169; 70904694; 70905922; 70905923; 70905927; 70905929; 70905939; 70905940; 70905942; 70905945; 70908313; 70912444; 70912546; 70912730; 70916039; 70917665; 70919851; 70920815; 70920852; 70923257; 70923258; 70924503; 70927147; 70930734; 70943460; 70943489; 70943532; 70943569; 70947409; 70960076; 70971406; 70971407; 70971830; 70971908; 70974146; 70974147; 70974157; 70974385; 70978041; 70978083; 7098617; 7098630; 7098663; 7098664; 70988912; 70995958; 71003464; 71009079; 71010344; 71010531; 71010739; 71010740; 71013333; 71013334; 71017103; 71019926; 71021213; 71021512; 71021552; 71022303; 71024598; 71024599; 71024600; 71024602; 71024603; 71024604; 71024610; 71026502; 71026504; 71026721; 71026767; 71026769; 71026770; 71026779; 71026783; 71026785; 71026901; 71031015; 71034150; 71034440; 71038157; 71039422; 71039470; 71039551; 71046037; 71052800; 71056489; 71068117; 71068228; 71072839; 71076901; 71078848; 71094775; 71096673; 71096885; 71096995; 71098378; 71098; 71102198; 71104487; 71122057; 71122742; 71125059; 71126249; 71126470; 71132378; 71143139; 71143140; 71153395; 71153540; 71153553; 71154273; 71160516; 71163818; 71168400; 71170840; 71171065; 71171249; 71171400; 71179324; 71179347; 71179348; 71182837; 71182838; 71182839; 71182840; 71189237; 71198285; 71204667; 71207412; 71210615; 71210617; 71217107; 71221169; 71223763; 71229078; 71229164; 71230487; 71231247; 71243125; 71243126; 71245564; 71245774; 71248371; 71248372; 71248605; 71250382; 71255014; 71267248; 71268059; 71268060; 71273370; 7127; 71287921; 71295147; 71298075; 71298167; 71298168; 71298169; 71299613; 71299614; 71299615; 71299616; 71299649; 71299754; 7130189; 71306906; 71306938; 71308202; 71308275; 71308285; 71308848; 71308893; 71309127; 71309128; 71309129; 71309139; 71309140; 71309142; 71309249; 71309251; 71309329; 71309338; 71309394; 71309492; 71309493; 71309503; 71309513; 71309514; 71309671; 71309810; 71309828; 71309852; 71309883; 71309905; 71309908; 71309918; 71309926; 71309927; 71310055; 71310073; 71310254; 71310259; 71310301; 71311469; 71313037; 71313040; 71313042; 71315458; 71315459; 71315460; 71315461; 71315462; 71315463; 71315464; 71316999; 71317000; 71317026; 71317027; 71317028; 71317085; 71317086; 71317094; 71317097; 71317125; 71317126; 71317131; 71317182; 71317242; 71321270; 71321271; 71321276; 71331643; 71339417; 71346174; 71346179; 71359088; 71360836; 71365640; 7136; 71403382; 71430872; 71431203; 71434121; 71434145; 71434189; 71434190; 71434223; 71434605; 71444487; 71444956; 71450313; 71464515; 71469257; 71476884; 71485; 71493251; 71497050; 71544; 7156899; 7156900; 7157007; 71571511; 71586808; 71587038; 71587638; 71587751; 71587802; 71587848; 71587930; 71620555; 71627239; 71651389; 71656069; 71656070; 71656071; 71656072; 71665536; 7167919; 7167921; 7167923; 7167925; 7170; 71716017; 71716605; 71716695; 71717298; 71717818; 71719661; 71720259; 71738110; 7173834; 7173835; 71744466; 71749567; 71749896; 71749899; 71749900; 71749901; 71750413; 71751872; 71752268; 71752285; 71752710; 7175; 71771808; 71771896; 71771927; 71774002; 71774137; 71774138; 71774139; 71774261; 71774293; 71774861; 71776184; 71777431; 71777455; 71777457; 71777654; 7184; 7187; 7191874; 7191875; 7191876; 7191877; 7191878; 7191879; 7193; 7194; 719782; 7202581; 7204907; 7204908; 7204909; 7204910; 7205332; 7205334; 72192941; 72196693; 72196696; 72199363; 72199473; 72200289; 72200; 72203026; 72231158; 72232122; 72560466; 72577947; 72577957; 72644633; 7264502; 7264503; 7264506; 7264507; 7264509; 7264510; 7264512; 7264513; 7269409; 72710558; 72715171; 7271821; 7271822; 72943164; 7311; 7325239; 7325240; 7326191; 7326192; 7326193; 7326194; 7326195; 7326196; 7326197; 7326216; 7326217; 7326218; 7326219; 7326220; 7326221; 7326222; 7326223; 7326224; 7326225; 7326229; 7326230; 73265279; 73325031; 7332594; 7332596; 7332597; 7332598; 73346478; 73346479; 73353118; 7335; 7336; 73386824; 73409720; 73425469; 73456809; 73557493; 73557527; 736172; 73661; 73662; 7370969; 7370970; 7370971; 7370972; 7370973; 7370974; 73755491; 73759827; 73796489; 73896887; 73940646; 73943104; 73951817; 73952988; 73; 74008; 74050978; 7408080; 7408081; 7427693; 7427696; 7427697; 7427700; 7427; 744412; 74764917; 74892754; 75320; 7533; 75357255; 75357256; 75411985; 75552; 7566553; 7566558; 7566942; 7566944; 7566987; 7566990; 7567181; 7567208;

7568072; 7568467; 7568469; 7568471; 7568532; 7568542; 7568545; 7568941; 7568945; 7568949; 7568953; 7593; 7599; 7600; 7601; 76069839; 76091866; 76286767; 76319594; 76320714; 7633; 764139; 7646; 7650; 7651; 76521977; 76530107; 765490; 7655; 7657; 7658; 76599; 76607; 7662; 7663; 7664; 7685; 76935; 76958155; 76959719; 7695; 76966087; 76971851; 76973265; 76973373; 770919; 7714; 77191561; 7733; 77477206; 77611128; 7778; 7781; 778574; 7790; 77933344; 7800; 78066606; 78144563; 78172108; 78224430; 78224436; 78225275; 78225644; 78350453; 78759339; 78894584; 78896428; 78941; 79025; 792941; 79427476; 79457; 7986; 80046; 80127; 802; 80354; 80390539; 8042; 8048; 8049; 8050; 80584; 8063362; 80796; 81045; 8122; 81264; 8138; 8139; 8167; 81696; 8180; 8181; 8193; 8194; 81964; 8201; 8205; 8207; 8209; 82170; 8221; 8222; 82308; 82336; 82491371; 82492671; 82494036; 82530; 82635; 82670420; 82679912; 82788327; 8294; 82956349; 83036; 83136; 83246; 8347; 8361; 8363; 83862; 8400; 8403; 84170; 8418; 84224052; 8432; 84559; 8456; 84571; 84695; 8469; 84740; 84996; 85090; 85520; 85521; 85570370; 85665113; 85741; 8615; 86278404; 86291136; 86301313; 86306394; 86306395; 86308127; 86308128; 86308792; 86308793; 86346640; 86559; 8655; 86575080; 86597; 86608; 86609; 86623592; 86742497; 86818; 8698; 8702; 87056795; 87060661; 87076217; 87079111; 87086528; 87090113; 87090114; 87095087; 87104154; 87107558; 87109007; 87110794; 87110805; 87110806; 87110811; 87110812; 87110826; 87110840; 87110841; 87110924; 87111033; 87129298; 87131217; 87139785; 87139859; 87166135; 87166396; 87203092; 87215252; 87225033; 87228929; 872363; 87262507; 87262570; 87288514; 87297824; 87299047; 872; 87304078; 87307864; 87317807; 87319225; 87331000; 87331002; 87336732; 87337393; 87356223; 87367794; 87374053; 87390328; 87393634; 87396752; 87438652; 87444302; 87452171; 87492103; 87523115; 87526544; 87562557; 87568891; 87576224; 87577602; 87590188; 87599894; 876160; 87616196; 8763; 87642889; 87645899; 87650794; 87650802; 87671148; 8767; 8768887; 8768888; 87708397; 87711378; 87724; 87732154; 87754686; 87767135; 87770206; 87770436; 87773300; 87775570; 87789691; 87813; 87816284; 87829687; 87830984; 87839202; 87845441; 87845444; 87877619; 87883498; 87893564; 87903642; 87918040; 87919553; 87929779; 8792; 87932701; 87994332; 88004027; 88016334; 88016626; 88034483; 88042761; 88049597; 8804; 88099966; 88102110; 88102112; 88103721; 88153078; 88185247; 88185248; 88204766; 88259209; 88299131; 88315503; 88318617; 88318618; 8831; 88326921; 8833; 8834; 8835; 88363299; 88369144; 88376505; 88388338; 88388339; 88391753; 88423175; 88447213; 88472051; 88473051; 88500158; 88504031; 88547603; 88579948; 88593574; 88620291; 88635004; 88639344; 8863; 88640192; 88689528; 88697864; 88708160; 88719; 88726758; 88726946; 88735; 88747143; 88755072; 88792560; 88798302; 88801967; 88909360; 88909362; 88921607; 88933144; 88938648; 88969992; 88971157; 88974141; 88980388; 88990100; 89000581; 89015893; 89020782; 89026460; 89026497; 89040318; 89040320; 89042466; 89042470; 8907; 89086645; 89094786; 89112838; 89112896; 8911609; 8911610; 8911616; 8911617; 89119439; 89126758; 89126762; 89126801; 89126802; 89136694; 89138282; 89140855; 8914859; 8914860; 8914869; 8914870; 89150867; 89166647; 89174364; 8918; 89190251; 89199537; 89200515; 89213821; 89213822; 89220688; 89226510; 89240705; 89242058; 89242343; 89244051; 89249780; 89255866; 89266219; 89289782; 89289848; 89306; 89307; 89309343; 89309984; 89310001; 89315322; 89315824; 89316; 89317890; 89323130; 89327884; 89327885; 89332529; 89332981; 89333506; 89337510; 89345843; 89349860; 89355142; 89357936; 89359565; 89360325; 89360759; 89363316; 89373960; 89374440; 89380297; 89383664; 89383666; 89402503; 89404282; 89409889; 89424182; 89428930; 89440; 89443239; 89449179; 89472723; 89488955; 89489561; 89502493; 89507839; 89512946; 89513347; 89516125; 89517129; 89518751; 89518758; 89535031; 89545930; 89546211; 89548557; 89548631; 89548634; 89548635; 89551425; 89552046; 89552558; 89558981; 89565170; 89565550; 89583253; 89583266; 89584157; 89593341; 89595591; 89605339; 89609533; 89611496; 89641306; 89647283; 89659018; 89662080; 89666589; 89669330; 89669499; 89681650; 89682175; 89694893; 89697507; 89705920; 89707354; 89717031; 8973; 89742272; 89742280; 89742285; 89756746; 89762380; 89762396; 89766756; 89770228; 89770229; 89770231; 89770232; 89770233; 89770509; 89770915; 89771266; 89771759; 89771762; 89772093; 89792263; 89797755; 89797805; 89797812; 89798663; 89805324; 89810855; 89818556; 89821241; 89822457; 89822465; 89822685; 89822687; 89824769; 89855666; 89858513; 89858596; 89858605; 89858707; 89858725; 89858800; 89859068; 89862; 89865599; 89869873; 89888626; 89888627; 89888870; 89901022; 89906833; 89906834; 89928693; 89928735; 89951765; 89951991; 89981081; 89988126; 89988896; 89989618; 89989620; 89989962; 89990199; 89990200; 89994930; 89994936; 90008; 90046345; 90057933; 90082222; 90086045; 90086046; 90086055; 90087353; 90108770; 90108771; 90115605; 90115606; 9012; 90132269; 90132725; 90136590; 9013; 90146798; 90146799; 90159939; 90162006; 90165009; 90165975; 90173; 90175644; 9017; 90181506; 90182081; 90188133; 90194848; 90228352; 90228373; 90228378; 90229004; 90229005; 9024; 90256488; 90260243; 90277732; 90277741; 90305152; 90305156; 90305157; 90305173; 90305308; 90320248; 90327874; 90327879; 90334378; 90334384; 90334386; 90334387; 90338074; 90338078; 90338104; 90339263; 90339325; 90346100; 90346255; 90346392; 90346504; 90347071; 90352177; 90352730; 90360112; 90360141; 90360523; 90375056; 90436597; 90443626; 90445228; 90448890; 90462020; 90470579; 90470917; 90471261; 90471562; 90472355; 90472363; 90472459; 90472720; 90472721; 90472751; 90472752; 90472753; 90472760; 90472761; 90472762; 90472763; 90473145; 90473857; 90474691; 90474830; 90475611; 90476898; 90479210; 90481571; 90481572; 90481573; 90540; 9060; 90657320; 90657618; 90658077; 90658939; 90664134; 90672358; 90673512; 90674662; 90674663; 90674664; 90674665; 90674666; 90674669; 90677558; 90677559; 90678741; 90678742; 90678749; 90678750; 90693018; 90698384; 90709690; 90728273; 90728282; 90753032; 90758804; 90766007; 90771119; 90781811; 90785; 90787552; 90790175; 90808165; 90825154; 90826456; 90840509; 90844641; 90855986; 90857639; 90876488; 90878489; 90882790; 90888514; 90895196; 90898185; 90898916; 90908346; 90918543; 90926336; 90929920; 90945623; 90947720; 90955620; 90957436; 90957794; 90959679; 90972; 90983566; 90986652; 90987644; 91016471; 91017827; 91037596; 91056676; 91057721; 91061602; 91071931; 91082956; 91112615; 91114576; 91121618; 91132899; 91145095; 91146595; 91151945; 91168237; 91170404; 91177337; 91184277; 91188557; 91194955; 91205183; 91213526; 91216621; 91217274; 91219843; 91225200; 91230442; 91249181; 91273349; 91274936; 91276820; 91282883; 91292920; 91295965; 91303323; 91313799; 91325874; 91329420; 91330433; 91332636; 91338702; 91342565; 91354; 91364718; 91365133; 91373772; 91379669; 91380274; 91383278; 91384988; 91389762; 91390759; 91392448; 91397843; 91398481; 91415912; 91417366; 91430781; 91432920; 91433979; 91439947; 91452818; 91457; 91459214; 91459638; 91459680; 91476; 91477103; 91486604; 91489367; 91496029; 91504241; 91516452; 91516637; 91516969; 91520522; 91524007; 91525512; 91538020; 91541906; 91549213; 91552003; 91557427; 91576; 91587; 91588637; 91591124; 91600064; 91605017; 91606114; 91606305; 91606477; 91608018; 91608141; 91612026; 91617326; 91654167; 91658939; 91666332; 91691114; 91691897; 91692850; 91694274; 91696341; 91697149; 91697152; 91697156; 91697164; 91746259; 91746261; 91746262; 91746263; 91820740; 91820744; 91844452; 91844950; 91844957; 91845017; 91845077; 91845082; 91845089; 91845130; 91845250; 91845272; 91845299; 91845440; 91845480; 91845496; 91845572; 91845600; 91845614; 91845620; 91845631; 91845639; 91845719; 91845818; 91845916; 91845934; 91845961; 91845976; 91846073; 91846074; 91846101; 91846154; 91846265; 91846325; 91846422; 91846446; 91846498; 91846511; 91846522; 91846543; 91846610; 91846618; 91846689; 91846802; 91846810; 91846820; 91846826; 91846874; 91846899; 91846921; 91846994; 91847008; 91847012; 91847038; 91847079; 91847098; 91847133; 91847163; 91847170; 91847171; 91847258; 91847313; 91847425; 91847441; 91847452; 91847513; 91847529; 91847586; 91847597; 91847674; 91847820; 91847828; 91847854; 91847875; 91847912; 91847929; 91847971; 91848000; 91848006; 91848034; 91848093; 91848181; 91848192; 91848204;

91848207; 91848213; 91848228; 91848231; 91848256; 91848260; 91848285; 91848311; 91848435; 91848466; 91848471; 91848498; 91848517; 91848530; 91848613; 91848691; 91848700; 91848704; 91848734; 91848742; 91848757; 91848802; 91848814; 91848825; 91848826; 91848940; 91848955; 91848965; 91848991; 91849015; 91849019; 91849022; 91849026; 91849037; 91849119; 91849129; 91849135; 91849166; 91849191; 91849208; 91849217; 91849281; 91849291; 91849295; 91849302; 91849304; 91849310; 91849315; 91849327; 91849368; 91849372; 91849375; 91849397; 91849442; 91849443; 91849602; 91849641; 91849645; 91849703; 91849727; 91849732; 91849787; 91849805; 91849807; 91849857; 91849861; 91849868; 91849887; 91849942; 91849958; 91850030; 91850082; 91850104; 91850119; 91850127; 91850171; 91850189; 91850197; 91850224; 91850288; 91850360; 91850365; 91850594; 91850627; 91850649; 91850656; 91850729; 91850775; 91850784; 91850827; 91850840; 91850930; 91850940; 91850964; 91850971; 91850980; 91850990; 91851080; 91851089; 91851108; 91851134; 91851190; 91851217; 91851221; 91851243; 91851260; 91851265; 91851341; 91851426; 91851492; 91851496; 91851539; 91851581; 91851612; 91851627; 91851652; 91851660; 91851705; 91851757; 91851925; 91851944; 91852024; 91852038; 91852058; 91852106; 91852107; 91852120; 91852206; 91852315; 91852403; 91852415; 91852433; 91852434; 91852464; 91852545; 91852554; 91852609; 91852610; 91852644; 91852646; 91852649; 91852667; 91852675; 91852676; 91852760; 91852777; 91852847; 91852871; 91852874; 91852889; 91852891; 91852892; 91852923; 91852926; 91852987; 91853001; 91853022; 91853044; 91853053; 91853079; 91853108; 91853111; 91853124; 91853203; 91853214; 91853229; 91853257; 91853282; 91853291; 91853326; 91853333; 91853358; 91853375; 91853406; 91853446; 91853482; 91853508; 91853519; 91853653; 91853704; 91853778; 91854031; 91854245; 91854267; 91854314; 91854336; 91854353; 91854363; 91854364; 91854380; 91854416; 91854437; 91854438; 91854471; 91854515; 91854560; 91854605; 91854626; 91854706; 91854726; 91854727; 91854741; 91854769; 91854786; 91854787; 91854857; 91854863; 91854991; 91855002; 91855013; 91855080; 91855142; 91855144; 91855216; 91855250; 91855274; 91855296; 91855325; 91855470; 91855550; 91855557; 91855584; 91855653; 91855725; 91855758; 91855876; 91855937; 91855947; 91855961; 91856054; 91856110; 91856178; 91856185; 91856190; 91856204; 91856221; 91856309; 91856390; 91856437; 91856527; 91856591; 91856615; 91856620; 91856632; 91856790; 91856806; 91856813; 91856920; 91856956; 91856976; 91856978; 91856992; 91857024; 91857055; 91857064; 91857070; 91857074; 91857077; 91857084; 91857090; 91857104; 91857212; 91857267; 91857279; 91857306; 91857355; 91857430; 91857448; 91857496; 91857502; 91857519; 91857566; 91857682; 91857744; 91857779; 91857813; 91857854; 91857874; 91857985; 91858046; 91858156; 91858170; 91858235; 91858256; 91858267; 91858296; 91858341; 91858379; 91858392; 91858443; 91858478; 91858569; 91858803; 91858813; 91858898; 91858985; 91859003; 91859019; 91859052; 91859055; 91859061; 91859080; 91859090; 91859111; 91859136; 91859168; 91859252; 91859292; 91859299; 91859350; 91859352; 91859370; 91859395; 91859538; 91859568; 91859571; 91859577; 91859593; 91859613; 91859717; 91859947; 91859954; 91860012; 91860074; 91860084; 91860100; 91860122; 91860132; 91860138; 91860408; 91860473; 91860478; 91860523; 91860527; 91860622; 91860639; 91860652; 91860687; 91860792; 91860813; 91860899; 91860907; 91860988; 91861030; 91861064; 91861077; 91861149; 91861175; 91861230; 91861257; 91861282; 91861289; 91861291; 91861305; 91861329; 91861338; 91861375; 91861390; 91861472; 91861497; 91861502; 91861539; 91861593; 91861681; 91861703; 91861738; 91861747; 91861823; 91861827; 91861836; 91861878; 91861917; 91861936; 91861988; 91862095; 91862116; 91862160; 91862215; 91862242; 91862332; 91862340; 91862372; 91862411; 91862423; 91862445; 91862466; 91862557; 91862565; 91862601; 91862688; 91862716; 91862739; 91862745; 91862772; 91862849; 91862858; 91862871; 91862933; 91865027; 91871300; 91871303; 91871304; 91884551; 91928504; 91928505; 91929660; 91944491; 91975487; 91978; 91980681; 92001949; 92021378; 92021550; 92021586; 92021587; 92021753; 92021812; 92021836; 92021997; 92022033; 92022058; 92022078; 92022082; 92022138; 92022148; 92022194; 92022239; 92022242; 92022250; 92022263; 92022288; 92022293; 92022294; 92022406; 92022528; 92022588; 92022593; 92022728; 92022808; 92022824; 92022826; 92022852; 92022904; 92023101; 92023398; 92023642; 92023837; 92023967; 92024069; 92024258; 92024268; 92024284; 92024333; 92033962; 92033963; 92033966; 92034723; 92034727; 92034729; 92042748; 92043446; 92043651; 92043770; 92092; 92116; 92131872; 92131873; 92132076; 92132096; 92132143; 92132156; 92132178; 92132338; 92132348; 92132372; 92132416; 92132419; 92138; 92139; 92144436; 92144936; 92147214; 92147461; 92160327; 92170307; 92171766; 92176858; 92177211; 92178856; 92179531; 92199; 92204555; 92209276; 92209277; 92209278; 92209279; 92209388; 92209389; 92209390; 92209391; 92209423; 92209424; 92209425; 92209426; 92209589; 92209590; 92209591; 92209690; 92209718; 92209719; 92209720; 92209721; 92209722; 92209723; 92209724; 92209725; 92209726; 92209727; 92209888; 92209889; 92209894; 92209895; 92209896; 92211307; 92211308; 92220880; 92220881; 92221095; 92221096; 92221097; 92221098; 92221117; 92221118; 92221119; 92221120; 92221121; 92221122; 92221123; 92221124; 92221125; 92221438; 92222362; 92222363; 92222364; 92222365; 92222756; 92222757; 92222758; 92222759; 92223100; 92223101; 92223147; 92223148; 92223149; 92223150; 92223463; 92223514; 92223515; 92223516; 92223517; 92223617; 92223618; 92223619; 92223620; 92223627; 92223628; 92223629; 92223630; 92223671; 92223672; 92223673; 92223674; 92229747; 92229748; 92229749; 92229750; 92231; 92236291; 92236292; 92236293; 92236372; 92236373; 92236374; 92237260; 92237261; 92251111; 92251112; 92251113; 92251114; 92256235; 92256236; 92256237; 92256238; 92257; 92258963; 92258964; 92258965; 92259095; 92259096; 92259097; 92259098; 92259141; 92259142; 92259143; 92259144; 92260816; 92260817; 92260818; 92262603; 92262604; 92262605; 92265623; 92265624; 92265625; 92271525; 92271526; 92271527; 92276331; 92276332; 92276333; 92276334; 92276335; 92276336; 92276337; 92276338; 92282144; 92282145; 92282147; 92286301; 92286510; 92291844; 92291845; 92291846; 92291847; 92295558; 92295560; 92295561; 92295565; 92295567; 92295568; 92298418; 92298422; 92298478; 92298482; 92308786; 92308787; 92308791; 92308796; 92308799; 92308800; 92308804; 92308811; 92313; 92323595; 92323600; 92323710; 92323711; 92333; 92354764; 92354765; 92354766; 92354767; 9240203; 9240211; 92403968; 92403969; 92436381; 92436439; 92436440; 92436442; 92436485; 92436526; 92436527; 92436597; 92436598; 92436612; 92436642; 92436643; 92436691; 92436692; 92436702; 92436703; 92436722; 92436726; 92436727; 92449115; 92449116; 92449117; 92449118; 92449121; 92449122; 92449123; 92449669; 92449670; 92449671; 92449672; 92450077; 92450259; 92460667; 92460668; 92467509; 92467510; 92467511; 92467512; 92468684; 92468685; 92472035; 92472; 92526854; 92649; 92762; 9276579; 9276580; 9276581; 9276582; 92776; 92843727; 92843728; 92843780; 92843796; 92854535; 92854979; 92855694; 92856405; 929262; 92954120; 92954121; 92961758; 92961759; 92961760; 92961761; 92965357; 92965358; 92965359; 92965360; 92969702; 92973609; 93019029; 93019031; 93019033; 93019035; 93019379; 93019380; 93075; 93078; 93081; 93090040; 93091914; 93199; 93233208; 93233209; 93233210; 93233211; 93234246; 93234247; 93286457; 93286458; 93481170; 93481369; 93481370; 93481372; 93481373; 93482988; 93482989; 93483189; 93484960; 93484961; 93485152; 93485153; 93485154; 93485155; 93485689; 93485756; 93485769; 93485770; 93485835; 93485836; 93493140; 93493141; 93504016; 93504017; 93520065; 93520072; 93532282; 93532283; 93532285; 93532286; 93532287; 93532288; 93532320; 93532321; 93532858; 93532870; 93532871; 93532872; 93838; 93970200; 93971; 93981192; 94001675; 94001676; 94001677; 94001678; 94005023; 94008900; 94011385; 94011386; 94011387; 94136; 94164; 94169; 94174; 94214; 94254; 94314; 94378; 9443734; 9443735; 9443738; 9443739; 94579; 9473; 94780; 94827838; 94886; 95048713; 95048714; 95048715; 95048716; 95048717; 95048718; 95048752; 95048753; 95048767; 95049172; 95086; 95162202; 95162203; 95162390; 95162391; 95162392; 95162393; 95162394; 95162395; 95162396; 95162406;

95162407; 95162408; 95162409; 95162410; 95162411; 95162412; 95162413; 95162414; 95162415; 95162416; 95162421; 95162422; 95162423; 95162424; 95162425; 95162426; 95162427; 95162428; 95162429; 95162430; 95162434; 95162435; 95162563; 95162564; 95162565; 95162566; 95162567; 95162568; 95162569; 95162570; 95162571; 95162572; 95162573; 95162574; 95162575; 95289; 95337; 95392; 95397; 95479525; 95479526; 95479937; 95479938; 95479939; 95479940; 9548577; 9548706; 9548799; 9548800; 9548801; 9548802; 9549179; 9549264; 9549265; 95495; 9549809; 9549810; 9549940; 9549951; 9549953; 9554554; 95566325; 95566326; 95566327; 95566328; 95566337; 95566421; 95566482; 95572; 95731630; 95731631; 95731632; 95731633; 95858802; 95858803; 95858804; 95858805; 9595223; 9601439; 96122502; 9623123; 96241; 9626261; 96361113; 96361114; 96361115; 96361116; 96361126; 96361127; 96361128; 96361129; 96361162; 96361163; 96361164; 96361165; 96365186; 96365188; 96365189; 96365191; 96365193; 96365196; 96365683; 96367908; 96367909; 96367911; 96367913; 96368; 96538813; 96559682; 9663048; 96814; 96851; 96866127; 96866363; 96866364; 96866369; 96873724; 96879480; 96885403; 96923; 97012111; 97012112; 97012113; 97041847; 97041848; 97043248; 97043249; 97043250; 97043251; 97043714; 97049941; 97049942; 97050062; 97050063; 97050064; 97050065; 97050066; 97050130; 97050131; 97050132; 97050408; 97050409; 97050410; 97050411; 97050428; 97052224; 97054139; 97054443; 97054444; 97054445; 97054446; 97054473; 97054474; 97054475; 97054476; 97054497; 97054509; 97054510; 97054511; 97054512; 97054524; 97054525; 97054526; 97054530; 97054532; 97054534; 97054535; 97054540; 97054541; 97267886; 97267887; 97289621; 97289622; 97289623; 97289787; 97299238; 97302378; 97302379; 97302383; 97302384; 97398; 97545874; 97546477; 97546499; 97546500; 97549014; 97549042; 97549043; 97549121; 97549122; 97549123; 97549190; 97549191; 97550280; 97550281; 97550373; 97550374; 97805; 97897; 9794056; 9794176; 9795796; 97960688; 97960689; 97960691; 97960692; 97960693; 97960694; 97961389; 97961390; 9796227; 9796699; 9797219; 9798107; 9798108; 9798109; 9798201; 9799725; 9799930; 9800420; 9801256; 98042586; 98043; 98045367; 98047596; 98047597; 98047598; 98092727; 98092728; 9810996; 9813700; 98137963; 98138983; 98138984; 98138985; 98138986; 98145185; 98145186; 98145696; 98145697; 98148765; 98148766; 98148767; 98148768; 9815418; 9815668; 9815696; 9815697; 98158547; 98158548; 98158549; 9815895; 9816473; 98170437; 9817091; 9817271; 98175187; 98175188; 98175189; 98175190; 98175191; 9818005; 9818041; 9818067; 9818810; 9819399; 9819400; 9819401; 9819497; 98207156; 9821051; 9823052; 9834129; 9837332; 9840209; 9840291; 9840964; 9840965; 9840966; 9840967; 9841053; 9841054; 9841288; 9841289; 9841827; 9841841; 9843444; 9843487; 98467040; 98467041; 98467042; 98467044; 98492; 984; 9855644; 9855656; 9855659; 9855795; 9857914; 9858617; 9858975; 9859136; 9859243; 985; 9860847; 9860975; 98615029; 98615030; 98617437; 98617438; 98617439; 98617440; 98617441; 98617443; 98618675; 98618676; 98618760; 9862315; 9862857; 9862858; 9863584; 9864257; 9864513; 9874786; 9877540; 9880679; 9882133; 9883148; 9883227; 9883997; 9884439; 9884532; 9884533; 9884534; 9884535; 9884536; 9884680; 9885688; 98860466; 98860467; 9888852; 9891231; 9891369; 9899003; 9899007; 9900706; 9905606; 9906026; 9906131; 9907067; 9907121; 9908826; 9909410; 99118573; 99134536; 9914698; 9920525; 9920533; 9920539; 9920559; 9921731; 9923756; 9926179; 9927340; 9927794; 9927987; 9927988; 9928082; 9930516; 9930733; 9942503; 99459; 9949726; 9950236; 99513; 9951889; 9954687; 99643478; 9964407; 9967861; 99721432; 99721433; 99721434; 99721435; 99726334; 9973509; 9973629; 9973885; 9974387; 9974754; 9975088; 9976525; 9979387; 9989796; 9990201; 9990240; 99906173; 99906174; 9992125; 9992463; 99933010; 99933011; 99933012; 99933013; 99933019; 99933020; 9993896; 99940264; 99947932; 99947934; 99957894; 9997457; 99975954; 9998008; 999888

Table S2. List of T3DB's molecules predicted as bitter by PyRMD model

9153 (T3D0010); 40470 (T3D0013); 3589 (T3D0034); 3118 (T3D0038); 3048 (T3D0041); 13940 (T3D0055); 3017 (T3D0056); 4115 (T3D0061); 6434141 (T3D0067); 9146 (T3D0070); 8343 (T3D0076); 1493 (T3D0087); 8461 (T3D0094); 2730 (T3D0127); 768 (T3D0133); 2268 (T3D0135); 8268 (T3D0144); 991 (T3D0145); 38199 (T3D0166); 42128 (T3D0168); 9131 (T3D0174); 3082 (T3D0229); 6433155 (T3D0233); 4420 (T3D0235); 1486 (T3D0248); 5104 (T3D0311); 8480 (T3D0313); 66826 (T3D0315); 23670523 (T3D0316); 27959 (T3D0394); 33100 (T3D0395); 36399 (T3D0396); 36982 (T3D0397); 36980 (T3D0398); 38032 (T3D0405); 37804 (T3D0406); 37803 (T3D0407); 38034 (T3D0409); 41541 (T3D0410); 38035 (T3D0411); 41555 (T3D0412); 41540 (T3D0413); 41551 (T3D0414); 38033 (T3D0415); 38030 (T3D0416); 27514 (T3D0418); 27960 (T3D0420); 38036 (T3D0422); 37805 (T3D0423); 37806 (T3D0424); 40467 (T3D0430); 63103 (T3D0431); 38875 (T3D0432); 38875 (T3D0433); 63108 (T3D0434); 521055 (T3D0435); 51041 (T3D0437); 38876 (T3D0438); 44163 (T3D0439); 63106 (T3D0440); 63102 (T3D0444); 38879 (T3D0445); 91705 (T3D0446); 63109 (T3D0447); 91718 (T3D0448); 36304 (T3D0449); 36401 (T3D0450); 63105 (T3D0451); 53035 (T3D0452); 63110 (T3D0453); 36402 (T3D0454); 36185 (T3D0455); 91717 (T3D0456); 51821 (T3D0457); 91674 (T3D0458); 36186 (T3D0459); 38878 (T3D0461); 36218 (T3D0463); 51042 (T3D0465); 63101 (T3D0467); 63093 (T3D0472); 40468 (T3D0473); 47651 (T3D0474); 41404 (T3D0475); 38014 (T3D0476); 41362 (T3D0477); 51822 (T3D0478); 50100 (T3D0479); 50099 (T3D0480); 40469 (T3D0481); 63084 (T3D0482); 63095 (T3D0483); 38012 (T3D0484); 38881 (T3D0486); 43238 (T3D0487); 38013 (T3D0488); 37807 (T3D0490); 36188 (T3D0494); 51075 (T3D0495); 51074 (T3D0496); 91704 (T3D0497); 93440 (T3D0498); 38015 (T3D0499); 91719 (T3D0501); 63087 (T3D0502); 53036 (T3D0503); 91720 (T3D0504); 35823 (T3D0507); 63083 (T3D0508); 50101 (T3D0509); 91728 (T3D0511); 47650 (T3D0512); 63088 (T3D0513); 41363 (T3D0518); 40473 (T3D0519); 91680 (T3D0520); 38017 (T3D0521); 40496 (T3D0523); 40517 (T3D0524); 37250 (T3D0526); 37035 (T3D0527); 63080 (T3D0528); 43004 (T3D0529); 40234 (T3D0530); 91635 (T3D0531); 63076 (T3D0532); 63071 (T3D0533); 93442 (T3D0534); 40172 (T3D0535); 63081 (T3D0536); 92391 (T3D0537); 38016 (T3D0538); 63075 (T3D0539); 40471 (T3D0540); 38019 (T3D0545); 50891 (T3D0546); 63072 (T3D0547); 91631 (T3D0548); 63079 (T3D0549); 93443 (T3D0550); 63074 (T3D0551); 63079 (T3D0552); 63073 (T3D0553); 40479 (T3D0556); 37037 (T3D0559); 40478 (T3D0560); 40481 (T3D0561); 50102 (T3D0562); 38025 (T3D0563); 91633 (T3D0564); 63097 (T3D0565); 40477 (T3D0566); 37036 (T3D0569); 53038 (T3D0570); 63098 (T3D0571); 40476 (T3D0572); 92392 (T3D0573); 40500 (T3D0574); 38857 (T3D0579); 53039 (T3D0580); 40485 (T3D0584); 39253 (T3D0585); 50103 (T3D0587); 38410 (T3D0590); 40483 (T3D0592); 9152 (T3D0622); 7047 (T3D0634); 14048876 (T3D0663); 161350 (T3D0664); 11012797 (T3D0680); 86418 (T3D0797); 3352 (T3D0798); 942 (T3D0799); 6758 (T3D0800); 8427 (T3D0801); 1480 (T3D0803); 7204 (T3D0804); 7207 (T3D0805); 7153 (T3D0806); 3030 (T3D0807); 2078 (T3D0810); 1988 (T3D0811); 2180 (T3D0813); 10658 (T3D0815); 68082 (T3D0816); 108104 (T3D0819); 181312 (T3D0821); 4825 (T3D0822); 6199 (T3D0825); 68081 (T3D0831); 7251175 (T3D0834); 4114 (T3D0836); 10212 (T3D0838); 6440498 (T3D0842); 68079 (T3D0843); 3083726 (T3D0845); 98608 (T3D0846); 11128 (T3D0859); 4626626 (T3D0860); 38479 (T3D0890); 82537 (T3D0898); 17563 (T3D0902); 27975 (T3D0903); 62097 (T3D0904); 27999 (T3D0905); 9570071 (T3D0919); 2314 (T3D0926); 28780 (T3D0928); 6129 (T3D0940); 25429 (T3D0942); 2566 (T3D0944); 34766 (T3D0964); 51605 (T3D0969); 17517 (T3D0977); 5353758 (T3D0984); 4944 (T3D1003); 11442 (T3D1024); 37261 (T3D1025); 4767 (T3D1026); 5053 (T3D1027); 83975 (T3D1028); 6436606 (T3D1029); 16722131 (T3D1030); 104926 (T3D1031); 5281873 (T3D1032); 38283 (T3D1033); 2912 (T3D1034); 40585 (T3D1035); 5361514 (T3D1036); 3347 (T3D1037); 50516 (T3D1039); 42586 (T3D1088); 4931 (T3D1090); 95019 (T3D1104); 5468405 (T3D1128); 160036 (T3D1137); 25113527 (T3D1147); 92244 (T3D1186); 10692 (T3D1193); 659840 (T3D1221); 25021466 (T3D1228); 22833302 (T3D1230); 22833337 (T3D1280); 22833324 (T3D1281); 16682738 (T3D1283); 91634 (T3D1285); 16683004 (T3D1286); 16682828 (T3D1287); 10749 (T3D1295); 23749 (T3D1303); 6537679 (T3D1313); 25113478 (T3D1349); 30556 (T3D1357); 441373 (T3D1361); 17023 (T3D1371); 8696 (T3D1403); 43957 (T3D1404); 151932 (T3D1445); 161975 (T3D1461); 6436086 (T3D1462); 31391 (T3D1463); 16683111 (T3D1537); 6433210 (T3D1548); 18730 (T3D1649); 169132 (T3D1662); 11443 (T3D1665); 11443 (T3D1667); 7607 (T3D1669); 35409 (T3D1670); 7069 (T3D1672); 220601 (T3D1677); 7764 (T3D1680); 23087 (T3D1695); 2603 (T3D1697); 3330 (T3D1698); 81530 (T3D1702); 10994 (T3D1703); 23167 (T3D1706); 389129 (T3D1771); 61828 (T3D1775); 114810 (T3D1780); 6035 (T3D1782); 5802 (T3D1790); 2453 (T3D1791); 11974 (T3D1798); 6618 (T3D1803); 5281872 (T3D1843); 162381 (T3D1844); 47326 (T3D1846); 12797557 (T3D1847); 6033664 (T3D1848); 91768 (T3D1849); 40326 (T3D1850); 5281874 (T3D1851); 48132 (T3D1852); 6433896 (T3D1853); 5281548 (T3D1854); 5371268 (T3D1855); 12304690 (T3D1856); 6433154 (T3D1857); 6433155 (T3D1858); 16685126 (T3D1894); 189283 (T3D1919); 189285 (T3D1921); 153879 (T3D1922); 153886 (T3D1923); 153887 (T3D1924); 153888 (T3D1925); 153889 (T3D1926); 153890 (T3D1928); 153891 (T3D1929); 180001 (T3D1937); 153899 (T3D1938); 3034399 (T3D1949); 38385 (T3D1953); 154302 (T3D1954); 154303 (T3D1955); 182662 (T3D1956); 185748 (T3D1957); 180583 (T3D1961); 554868 (T3D1964); 181211 (T3D1965); 181213 (T3D1968); 591320 (T3D1974); 49972 (T3D1983); 49996 (T3D1984); 107888 (T3D1985); 119124 (T3D1986); 108023 (T3D1987); 155277 (T3D1989); 155278 (T3D1990); 155279 (T3D1991); 154405 (T3D1995); 119134 (T3D1997); 154409 (T3D1998); 154465 (T3D2000); 154475 (T3D2001); 154478 (T3D2002); 157767 (T3D2003); 154483 (T3D2006); 154484 (T3D2007); 158272 (T3D2008); 158627 (T3D2009); 158629 (T3D2011); 154485 (T3D2014); 181948 (T3D2015); 181949 (T3D2016); 154496 (T3D2020); 154499 (T3D2021); 154509 (T3D2022); 154516 (T3D2026); 21123676 (T3D2032); 21582035 (T3D2033); 22023651 (T3D2034); 39506 (T3D2038); 38386 (T3D2039); 36159 (T3D2040); 22833472 (T3D2041); 3034400 (T3D2042); 45472 (T3D2044); 39727 (T3D2048); 41117 (T3D2054); 158065 (T3D2056); 158064 (T3D2057); 38253 (T3D2058); 36614 (T3D2059); 48888 (T3D2060); 115206 (T3D2061); 158074 (T3D2062); 158075 (T3D2063); 93472 (T3D2064); 93471 (T3D2065); 93470 (T3D2066); 158068 (T3D2067); 158067 (T3D2068); 158066 (T3D2069); 50627 (T3D2070); 91654 (T3D2071); 51376 (T3D2073); 48889 (T3D2074); 40811 (T3D2075); 51377 (T3D2076); 51378 (T3D2077); 51379 (T3D2078); 51380 (T3D2079); 51373 (T3D2080); 49803 (T3D2082); 38525 (T3D2083); 36934 (T3D2084); 51375 (T3D2087); 39732 (T3D2088); 44069 (T3D2089); 38254 (T3D2092); 38439 (T3D2093); 42836 (T3D2094); 51457 (T3D2095); 54848 (T3D2096); 54846 (T3D2097); 54845 (T3D2098); 51456 (T3D2099); 54844 (T3D2100); 51455 (T3D2101); 48890 (T3D2102); 54847 (T3D2103); 51466 (T3D2104); 51454 (T3D2105); 47367 (T3D2107); 42668 (T3D2110); 42671 (T3D2112); 42669 (T3D2115); 37270 (T3D2116); 42672 (T3D2117); 35733 (T3D2132); 176052 (T3D2134); 51129 (T3D2136); 153906 (T3D2137); 54337 (T3D2138); 50611 (T3D2139); 50613 (T3D2140); 56261 (T3D2141); 55127 (T3D2142); 44152416 (T3D2143); 41510 (T3D2144); 119363 (T3D2145); 51130 (T3D2146); 56262 (T3D2147); 55128 (T3D2148); 41522 (T3D2149); 49923 (T3D2150); 55129 (T3D2151); 35331 (T3D2152); 176051 (T3D2154); 42140 (T3D2155); 56400 (T3D2156); 42139 (T3D2157); 53176 (T3D2158); 55131 (T3D2159); 55132 (T3D2160); 55103 (T3D2161); 51720 (T3D2163); 86281 (T3D2164); 42138 (T3D2165); 55133 (T3D2166); 55104 (T3D2167); 55134 (T3D2168); 527707 (T3D2169); 55105 (T3D2170); 49931 (T3D2171); 53291 (T3D2172); 55130 (T3D2173); 50610 (T3D2175); 35332 (T3D2176); 51463 (T3D2177); 42839 (T3D2178); 51464 (T3D2179); 55137 (T3D2180); 51128 (T3D2181); 47300 (T3D2182); 50544 (T3D2184); 51195 (T3D2185); 41558 (T3D2186); 55136 (T3D2187); 55290 (T3D2188); 42844 (T3D2189); 55107 (T3D2190); 47298 (T3D2191); 51465 (T3D2192); 55116 (T3D2194); 55135 (T3D2195); 51121 (T3D2196); 55108 (T3D2197); 42840 (T3D2198); 51125 (T3D2199); 51122 (T3D2200); 56398 (T3D2201); 51126 (T3D2203); 42133 (T3D2204); 51462 (T3D2205); 55102 (T3D2206); 42132 (T3D2207); 47426

(T3D2209); 32437 (T3D2210); 48039 (T3D2211); 54935 (T3D2212); 55110 (T3D2213); 47424 (T3D2214); 55293 (T3D2215); 2757441 (T3D2225); 43495 (T3D2229); 153905 (T3D2230); 527711 (T3D2231); 49888 (T3D2232); 44101 (T3D2233); 107998 (T3D2234); 55111 (T3D2236); 55112 (T3D2237); 42131 (T3D2238); 42136 (T3D2239); 42134 (T3D2240); 42130 (T3D2241); 42841 (T3D2244); 42129 (T3D2245); 42135 (T3D2247); 42838 (T3D2249); 41131 (T3D2250); 527712 (T3D2251); 43255 (T3D2257); 47643 (T3D2258); 6856 (T3D2266); 32781 (T3D2268); 53144 (T3D2281); 32781 (T3D2286); 153880 (T3D2306); 53154 (T3D2307); 130491 (T3D2324); 114900 (T3D2325); 125530 (T3D2326); 130492 (T3D2327); 130490 (T3D2328); 156484 (T3D2329); 175927 (T3D2330); 3016682 (T3D2331); 153863 (T3D2336); 162113 (T3D2338); 131379 (T3D2339); 183885 (T3D2350); 153851 (T3D2358); 173619 (T3D2378); 526349 (T3D2384); 56776 (T3D2415); 53155 (T3D2416); 56777 (T3D2417); 54762 (T3D2418); 47299 (T3D2420); 53156 (T3D2421); 56780 (T3D2422); 55054 (T3D2423); 47425 (T3D2424); 55117 (T3D2425); 55114 (T3D2426); 55120 (T3D2427); 32436 (T3D2428); 55119 (T3D2429); 53524 (T3D2430); 55118 (T3D2431); 54937 (T3D2432); 55121 (T3D2433); 51119 (T3D2434); 42842 (T3D2435); 55122 (T3D2436); 54936 (T3D2437); 55123 (T3D2438); 55124 (T3D2439); 55125 (T3D2440); 55126 (T3D2441); 55106 (T3D2442); 51123 (T3D2443); 55109 (T3D2445); 55113 (T3D2447); 89579 (T3D2451); 5079496 (T3D2452); 4479242 (T3D2454); 1233 (T3D2455); 10624 (T3D2457); 4980 (T3D2458); 149096 (T3D2478); 443368 (T3D2484); 6437364 (T3D2528); 6440480 (T3D2529); 1204 (T3D2530); 5470308 (T3D2531); 10051 (T3D2532); 10257 (T3D2534); 10103 (T3D2535); 11969465 (T3D2536); 6441116 (T3D2550); 37165 (T3D2551); 5002 (T3D2555); 46507594 (T3D2566); 54678486 (T3D2567); 3878 (T3D2570); 1983 (T3D2571); 131007 (T3D2626); 60795 (T3D2677); 171548 (T3D2690); 5280453 (T3D2692); 5283731 (T3D2694); 1130 (T3D2696); 5280793 (T3D2697); 6037 (T3D2698); 445354 (T3D2699); 14985 (T3D2700); 1054 (T3D2701); 6112 (T3D2702); 5280795 (T3D2703); 4055 (T3D2704); 2519 (T3D2712); 2771 (T3D2715); 2812 (T3D2729); 5284627 (T3D2733); 5656 (T3D2736); 5284371 (T3D2749); 4585 (T3D2754); 2678 (T3D2756); 4205 (T3D2767); 71384 (T3D2771); 5732 (T3D2787); 2913 (T3D2790); 3957 (T3D2796); 3696 (T3D2797); 3034034 (T3D2800); 2978 (T3D2801); 5281081 (T3D2808); 5362449 (T3D2815); 10917 (T3D2832); 938 (T3D2841); 54575 (T3D2849); 5358 (T3D2856); 43815 (T3D2868); 4914 (T3D2869); 5073 (T3D2871); 5184 (T3D2877); 444795 (T3D2884); 3016 (T3D2903); 3152 (T3D2910); 4595 (T3D2925); 2244 (T3D2936); 3348 (T3D2938); 3446 (T3D2952); 5284607 (T3D2959); 3672 (T3D2970); 896 (T3D2972); 3100 (T3D2976); 13505 (T3D2978); 2337 (T3D2980); 68617 (T3D2985); 2725 (T3D2989); 444 (T3D2999); 146570 (T3D3007); 66265 (T3D3014); 6047 (T3D3028); 115121 (T3D3074); 442021 (T3D3075); 5281132 (T3D3076); 3503 (T3D3081); 6436464 (T3D3083); 1548943 (T3D3085); 441071 (T3D3086); 51683 (T3D3087); 75729 (T3D3089); 19604097 (T3D3090); 107976 (T3D3091); 73051 (T3D3092); 64692 (T3D3095); 637577 (T3D3096); 6000 (T3D3097); 5360723 (T3D3098); 5281411 (T3D3099); 41736 (T3D3101); 2181 (T3D3108); 4276 (T3D3114); 16666 (T3D3224); 17739 (T3D3230); 4225619 (T3D3233); 7843 (T3D3426); 6360 (T3D3427); 1388 (T3D3474); 2249 (T3D3480); 3346 (T3D3503); 4594 (T3D3524); 128563 (T3D3532); 2520 (T3D3544); 44135439 (T3D3558); 38853 (T3D3577); 6947 (T3D3579); 219203 (T3D3597); 14403 (T3D3598); 6223 (T3D3604); 61913 (T3D3619); 15610387 (T3D3620); 104234 (T3D3621); 7348 (T3D3622); 23687235 (T3D3623); 23668817 (T3D3624); 38258 (T3D3625); 62277 (T3D3626); 22947 (T3D3627); 9028 (T3D3629); 23665884 (T3D3630); 40138 (T3D3633); 174876 (T3D3636); 6781 (T3D3638); 8560 (T3D3642); 6777 (T3D3643); 8970 (T3D3647); 34277 (T3D3648); 8559 (T3D3655); 4696 (T3D3661); 6436831 (T3D3662); 5284457 (T3D3663); 5375083 (T3D3665); 15558498 (T3D3666); 430147 (T3D3668); 23648 (T3D3669); 14421 (T3D3670); 23670 (T3D3671); 10903619 (T3D3672); 440908 (T3D3674); 22283 (T3D3675); 442072 (T3D3684); 443884 (T3D3685); 134553 (T3D3687); 73484 (T3D3692); 15519 (T3D3693); 11051 (T3D3694); 14987 (T3D3695); 160462 (T3D3696); 431810 (T3D3698); 429922 (T3D3699); 104759 (T3D3700); 536551 (T3D3701); 6437832 (T3D3702); 587712 (T3D3703); 429923 (T3D3704); 6451372 (T3D3706); 304599 (T3D3707); 426719 (T3D3708); 14083644 (T3D3713); 599328 (T3D3714); 73049 (T3D3715); 20806 (T3D3716); 114680 (T3D3717); 6436600 (T3D3718); 107974 (T3D3719); 6441510 (T3D3721); 6437363 (T3D3722); 104989 (T3D3723); 5284338 (T3D3725); 5359485 (T3D3727); 5360741 (T3D3728); 104860 (T3D3729); 11969548 (T3D3730); 31181 (T3D3731); 54676038 (T3D3735); 2908 (T3D3736); 6436023 (T3D3737); 5385314 (T3D3738); 105014 (T3D3740); 3442 (T3D3741); 403923 (T3D3744); 114734 (T3D3745); 6438150 (T3D3749); 6441012 (T3D3750); 6438901 (T3D3751); 10255275 (T3D3752); 27790 (T3D3753); 3220 (T3D3755); 16727691 (T3D3756); 16727692 (T3D3757); 164754 (T3D3758); 467499 (T3D3759); 6436022 (T3D3760); 30840 (T3D3768); 3840 (T3D3769); 73431 (T3D3774); 88363 (T3D3775); 99596 (T3D3776); 114932 (T3D3779); 6437365 (T3D3781); 8693 (T3D3943); 1982 (T3D3788); 44073 (T3D3790); 12932 (T3D3797); 213013 (T3D3800); 8606 (T3D3804); 21803 (T3D3808); 8425 (T3D3814); 86367 (T3D3820); 15331 (T3D3821); 7430 (T3D3823); 4284 (T3D3827); 41381 (T3D3832); 17432 (T3D3835); 31070 (T3D3837); 3033674 (T3D3841); 443048 (T3D3846); 92425 (T3D3847); 54745 (T3D3849); 73675 (T3D3850); 8607 (T3D3853); 91763 (T3D3856); 39965 (T3D3857); 125098 (T3D3859); 37175 (T3D3860); 107720 (T3D3864); 37517 (T3D3866); 4004 (T3D3870); 8575 (T3D3873); 8227 (T3D3885); 20392 (T3D3886); 5371562 (T3D3888); 6336 (T3D3889); 39327 (T3D3896); 74483 (T3D3899); 5372405 (T3D3909); 43234 (T3D3910); 91751 (T3D3912); 91753 (T3D3916); 91739 (T3D3918); 62829 (T3D3920); 5216 (T3D3922); 86369 (T3D3924); 5430 (T3D3931); 91776 (T3D3933); 5564 (T3D3939); 6923 (T3D3944); 6943 (T3D3945); 8814 (T3D3946); 6619 (T3D3948); 5280972 (T3D3953); 408 (T3D3956); 92272 (T3D3960); 75318 (T3D3961); 91469 (T3D3962); 5281708 (T3D3964); 89472 (T3D3965); 123917 (T3D3967); 53481631 (T3D3971); 20393 (T3D3973); 7456 (T3D3977); 8434 (T3D3978); 7175 (T3D3979); 104856 (T3D3984); 9554 (T3D3985); 96088 (T3D3990); 9555 (T3D3999); 77222 (T3D4005); 31736 (T3D4006); 4970 (T3D4015); 2833 (T3D4020); 5154 (T3D4022); 104753 (T3D4023); 441451 (T3D4024); 441445 (T3D4025); 264115 (T3D4026); 10143 (T3D4027); 25565 (T3D4032); 46173840 (T3D4038); 11593018 (T3D4039); 54601877 (T3D4040); 5284645 (T3D4052); 13932155 (T3D4053); 5280906 (T3D4059); 10198 (T3D4060); 441581 (T3D4063); 72378 (T3D4064); 441753 (T3D4066); 6391813 (T3D4068); 442070 (T3D4069); 5462429 (T3D4071); 10237 (T3D4073); 442189 (T3D4074); 6321388 (T3D4078); 68085 (T3D4080); 91466 (T3D4081); 442972 (T3D4082); 10098 (T3D4083); 6070 (T3D4084); 2703 (T3D4086); 197810 (T3D4087); 101300 (T3D4092); 197001 (T3D4103); 2236 (T3D4113); 431734 (T3D4124); 114989 (T3D4138); 790 (T3D4150); 802 (T3D4152); 999 (T3D4153); 439260 (T3D4154); 441564 (T3D4157); 10258 (T3D4158); 464 (T3D4159); 3845 (T3D4160); 4455 (T3D4162); 227 (T3D4165); 936 (T3D4168); 123831 (T3D4171); 160437 (T3D4173); 439918 (T3D4174); 73318 (T3D4175); 169148 (T3D4176); 119593 (T3D4177); 69698 (T3D4178); 440810 (T3D4179); 5283344 (T3D4180); 4615423 (T3D4184); 6439956 (T3D4193); 5283334 (T3D4195); 558 (T3D4196); 123979 (T3D4197); 123800 (T3D4198); 836 (T3D4199); 2733785 (T3D4200); 1549111 (T3D4201); 5144 (T3D4212); 27919 (T3D4215); 42628600 (T3D4232); 17754066 (T3D4234); 6128 (T3D4240); 12594 (T3D4242); 6166 (T3D4245); 439939 (T3D4247); 439391 (T3D4251); 193530 (T3D4258); 469 (T3D4259); 7405 (T3D4270); 1021 (T3D4274); 60961 (T3D4275); 5957 (T3D4276); 15993 (T3D4278); 190 (T3D4279); 160666 (T3D4281); 6238 (T3D4282); 6013 (T3D4283); 11222 (T3D4284); 780 (T3D4287); 6322 (T3D4289); 586 (T3D4290); 967 (T3D4292); 16950 (T3D4294); 9750 (T3D4295); 33032 (T3D4296); 144 (T3D4298); 681 (T3D4300); 65065 (T3D4310); 11025495 (T3D4314); 5961 (T3D4319); 72423 (T3D4322); 114935 (T3D4323); 72424 (T3D4324); 6137 (T3D4328); 93556 (T3D4332); 221493 (T3D4343); 439709 (T3D4345); 5280352 (T3D4350); 124886 (T3D4352); 5793 (T3D4356); 6267 (T3D4362); 6274 (T3D4363); 6306 (T3D4364); 6140 (T3D4366); 6305 (T3D4369); 6057 (T3D4371); 173909 (T3D4372); 160556 (T3D4398); 1188 (T3D4409); 47 (T3D4424); 65070 (T3D4430); 115015 (T3D4434); 5610 (T3D4435); 321 (T3D4438); 121893 (T3D4439); 4971 (T3D4440); 26840 (T3D4454); 439155 (T3D4455); 34756 (T3D4456); 21252321 (T3D4467); 5312 (T3D4471); 2256 (T3D4477); 17038 (T3D4482); 13754 (T3D4494); 91757 (T3D4499); 3120

(T3D4500); 62210 (T3D4510); 53395436 (T3D4512); 93379 (T3D4513); 54738 (T3D4516); 15415 (T3D4519); 93541 (T3D4521); 5794 (T3D4524); 34526 (T3D4525); 16760153 (T3D4526); 34469 (T3D4527); 91650 (T3D4532); 39385 (T3D4543); 41428 (T3D4545); 5569 (T3D4548); 3162 (T3D4551); 969472 (T3D4552); 2564 (T3D4553); 6834 (T3D4554); 7028 (T3D4555); 3516 (T3D4556); 1530 (T3D4581); 13314 (T3D4585); 119033 (T3D4618); 91434 (T3D4619); 11665476 (T3D4631); 10667 (T3D4636); 15847 (T3D4657); 9213 (T3D4658); 17604 (T3D4660); 5757 (T3D4668); 2230 (T3D4675); 4173 (T3D4703); 4764 (T3D4708); 2083 (T3D4735); 2154 (T3D4741); 5991 (T3D4743); 448537 (T3D4744); 5870 (T3D4746); 10635 (T3D4751); 5447130 (T3D4752); 5280343 (T3D4758); 5920 (T3D4759); 46936256 (T3D4760); 2662 (T3D4771); 3767 (T3D4777); 2153 (T3D4780); 3639 (T3D4781); 54454 (T3D4788); 5756 (T3D4790); 65211 (T3D4793); 5280863 (T3D4795); 3735 (T3D4798); 4632 (T3D4807); 19241 (T3D4810); 71398 (T3D4813); 338 (T3D4816); 493570 (T3D4817); 57363 (T3D4820); 47936 (T3D4827); 68570 (T3D4832); 6506 (T3D4833); 1752 (T3D4836); 4133 (T3D4848); 1183 (T3D4849); 4947 (T3D4850); 5372174 (T3D4853); 3610 (T3D4855); 21831736 (T3D4857); 164825 (T3D4858); 16043 (T3D4859); 33258 (T3D4860); 17520 (T3D4861); 17730 (T3D4862); 19700 (T3D4864); 23691036 (T3D4865); 6505 (T3D4871); 6950 (T3D4873); 3293 (T3D4874); 323 (T3D4875); 14339290 (T3D4878); 31364 (T3D4885); 5280373 (T3D4886); 5280443 (T3D4888); 444539 (T3D4889); 6998 (T3D4892); 5281607 (T3D4893); 61585 (T3D4895); 8361 (T3D4905); 8400 (T3D4906); 1623625 (T3D4908); 637759 (T3D4909); 10245972 (T3D4910); 8738 (T3D4913); 27872 (T3D4914); 3400691 (T3D4916); 9395 (T3D4920); 37123 (T3D4921); 20087 (T3D4922); 62542 (T3D4923); 10680 (T3D4924); 135191 (T3D4926); 440968 (T3D4930); 6548 (T3D4931); 66597 (T3D4933); 7122 (T3D4934); 204 (T3D4936); 8697 (T3D4938); 32518 (T3D4939); 5282108 (T3D4940); 12101 (T3D4941); 91695 (T3D4942); 121232962 (T3D4954)

Note: the PubChem CID and T3DB ID (between brackets) are reported for each molecule

Table S3. Docking scores of molecules under analysis

Compound	**GOLDScore units***
Strychnine	78.4
Artemisin	59.1
Aristolochic acid	71.3
Trichothecolone	59.9
DON	44.8
DON-3-Sulf	57.9
DON-15-Sulf	59.9
DON-3-GlcA	51.9
DON-15-GlcA	65.0
3-Ac-DON	50.4
15-Ac-DON	56.5
DON-3-Glc	32.5
DON-15-Glc	60.5

Note: * positive scores indicates the theoretical capability of ligands to interact with the protein (the higher the score, the more likely is the precited binding architecture, according to manufacturer declaration; https://www.ccdc.cam.ac.uk; as of 17th April 2023)

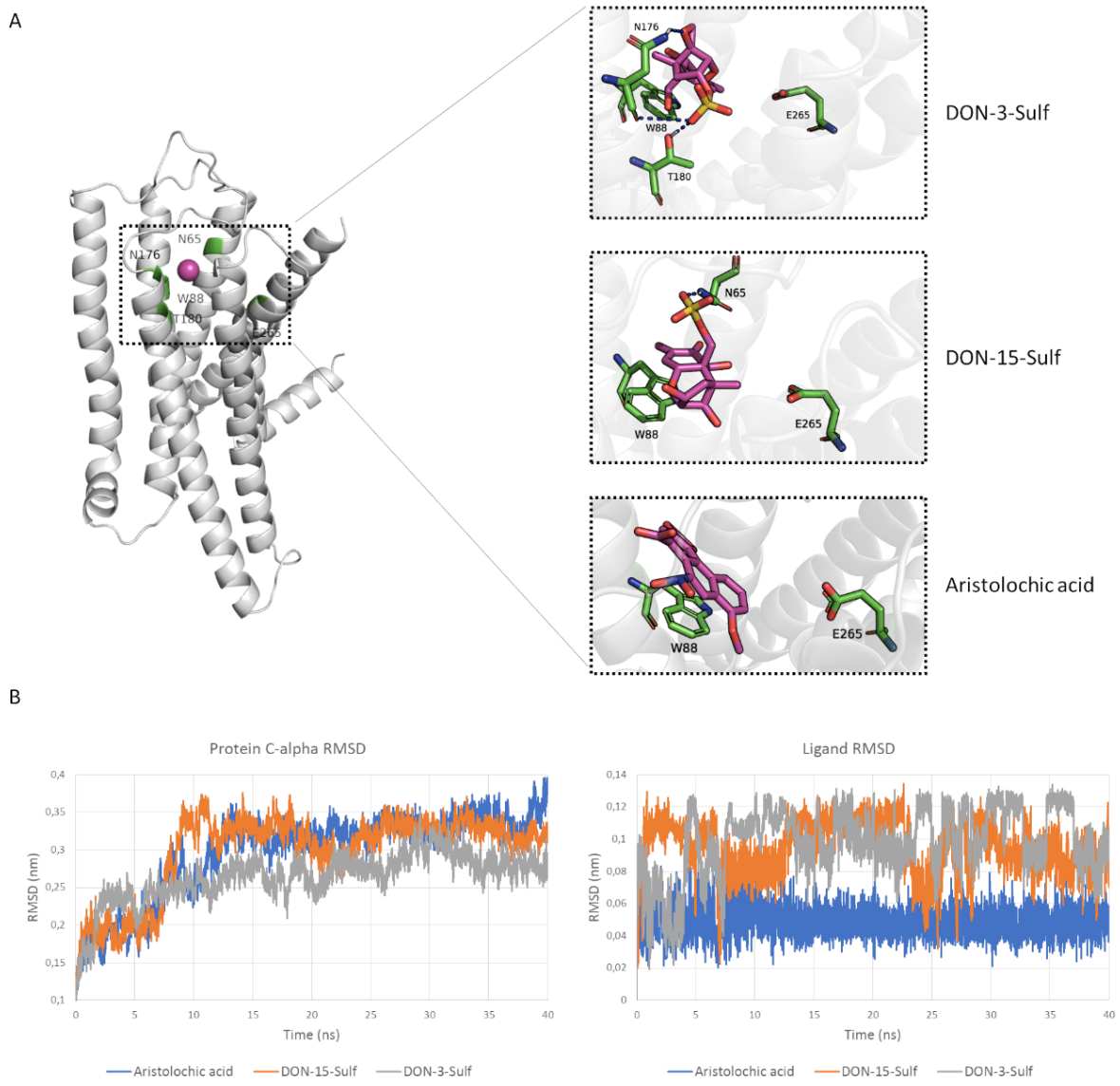


Figure S1. Molecular modelling results of DON sulphates. **A.** Docking pose of DON-3 and –15-Sulf and aristolochic acid. On the left, TAS2R46 is represented in white cartoon. On the right, close-up of the ligand binding site with residues side chains and ligands represented as sticks. Blu dashed lines indicate polar contacts. **B.** On the left, protein C-alpha RMSD of TAS2R46 in complex with aristolochic acid (average value 0.297 ± 0.057 nm), DON-15-Sulf (average value 0.30 ± 0.05 nm) and DON-3-Sulf (average value 0.26 ± 0.03 nm). On the right, ligand RMSD of aristolochic acid (average value 0.049 ± 0.009 nm), DON-15-Sulf (average value 0.095 ± 0.017 nm) or DON-3-Sulf (average value 0.099 ± 0.021 nm).

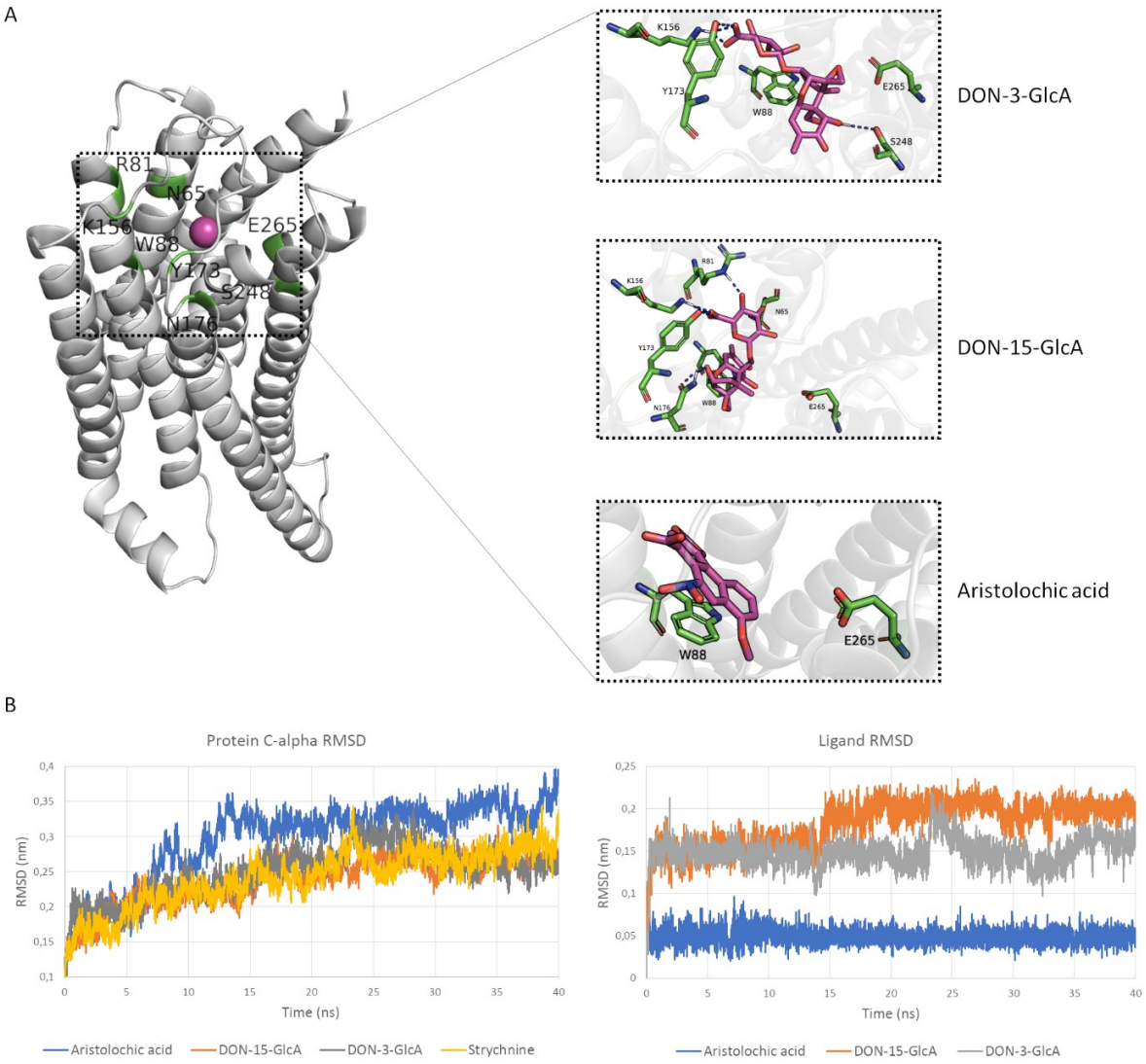


Figure S2. Molecular modelling results of DON glucuronides. **A.** Docking pose of DON-3- and –15-GlcA and aristolochic acid. On the left, TAS2R46 is represented in white cartoon. On the right, close-up of the ligand binding site with amino acid side chains and ligands represented as sticks. Blu dashed lines indicate polar contacts. **B.** On the left, protein C-alpha RMSD of TAS2R46 in complex with aristolochic acid (average value 0.297 ± 0.057 nm), DON-15-GlcA (average value 0.24 ± 0.03 nm), DON-3-GlcA (average value 0.25 ± 0.03 nm) and strychnine (average value 0.241 ± 0.041 nm). On the right, ligand RMSD of aristolochic acid (average value 0.049 ± 0.009 nm), DON-15-GlcA (average value 0.18 ± 0.03 nm) and DON-3-GlcA (average value 0.15 ± 0.02 nm).

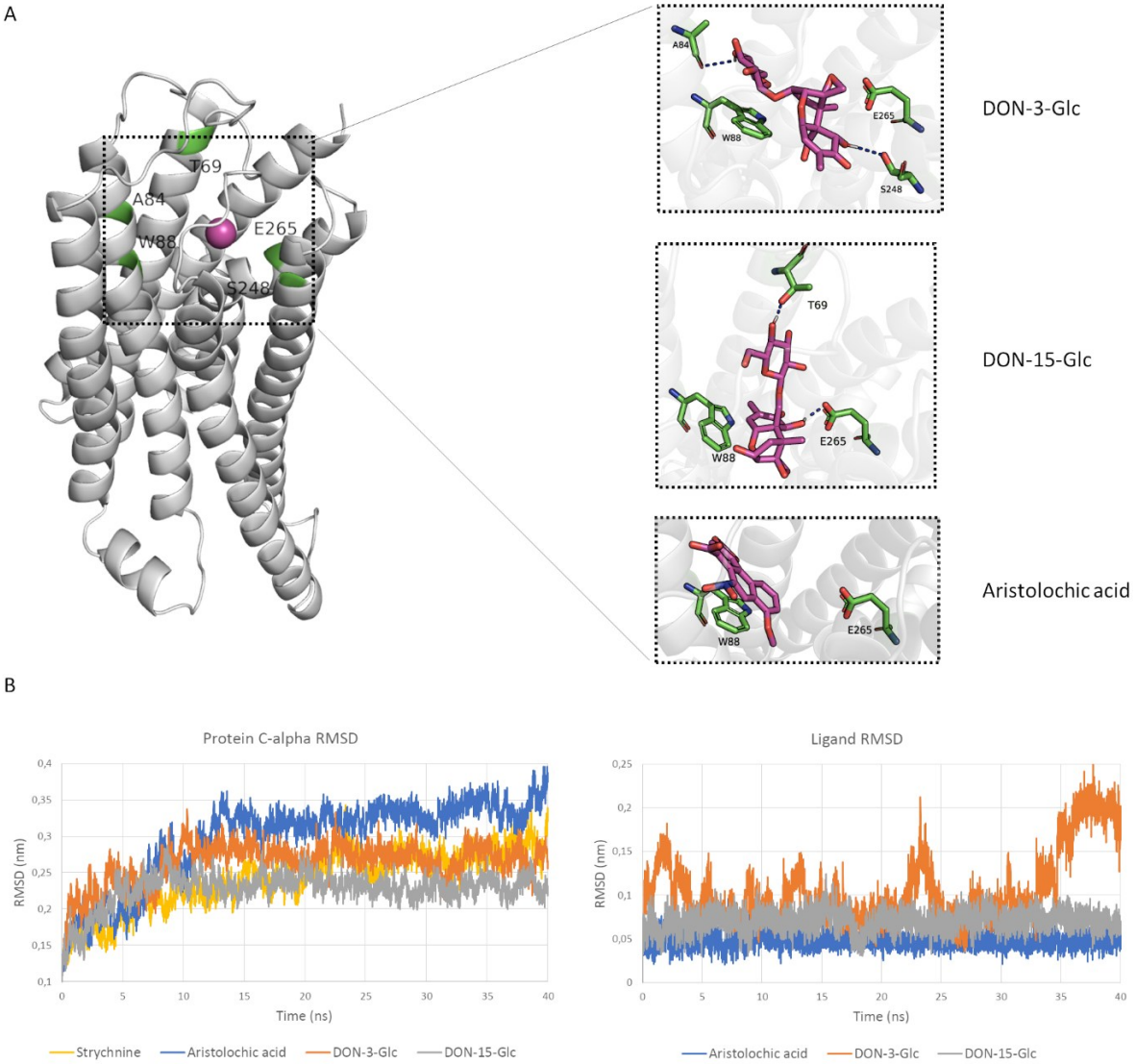


Figure S3. Molecular modelling results of DON glucosides. **A.** Docking pose of DON-3- and –15-Glc and aristolochic acid. On the left, TAS2R46 is represented in white cartoon. On the right, close-up of the ligand binding site with amino acid side chains and ligands represented as sticks. Blu dashed lines indicate polar contacts. **B.** On the left, protein C-alpha RMSD of TAS2R46 in complex with aristolochic acid (average value 0.297 ± 0.057 nm), DON-15-Glc (average value 0.23 ± 0.02 nm), DON-3-Glc (average value 0.26 ± 0.03 nm) and strychnine (average value 0.241 ± 0.041 nm). On the right, ligand RMSD of aristolochic acid (average value 0.049 ± 0.009 nm), DON-15-Glc (average value 0.07 ± 0.01 nm) and DON-3-Glc (average value 0.11 ± 0.04 nm).

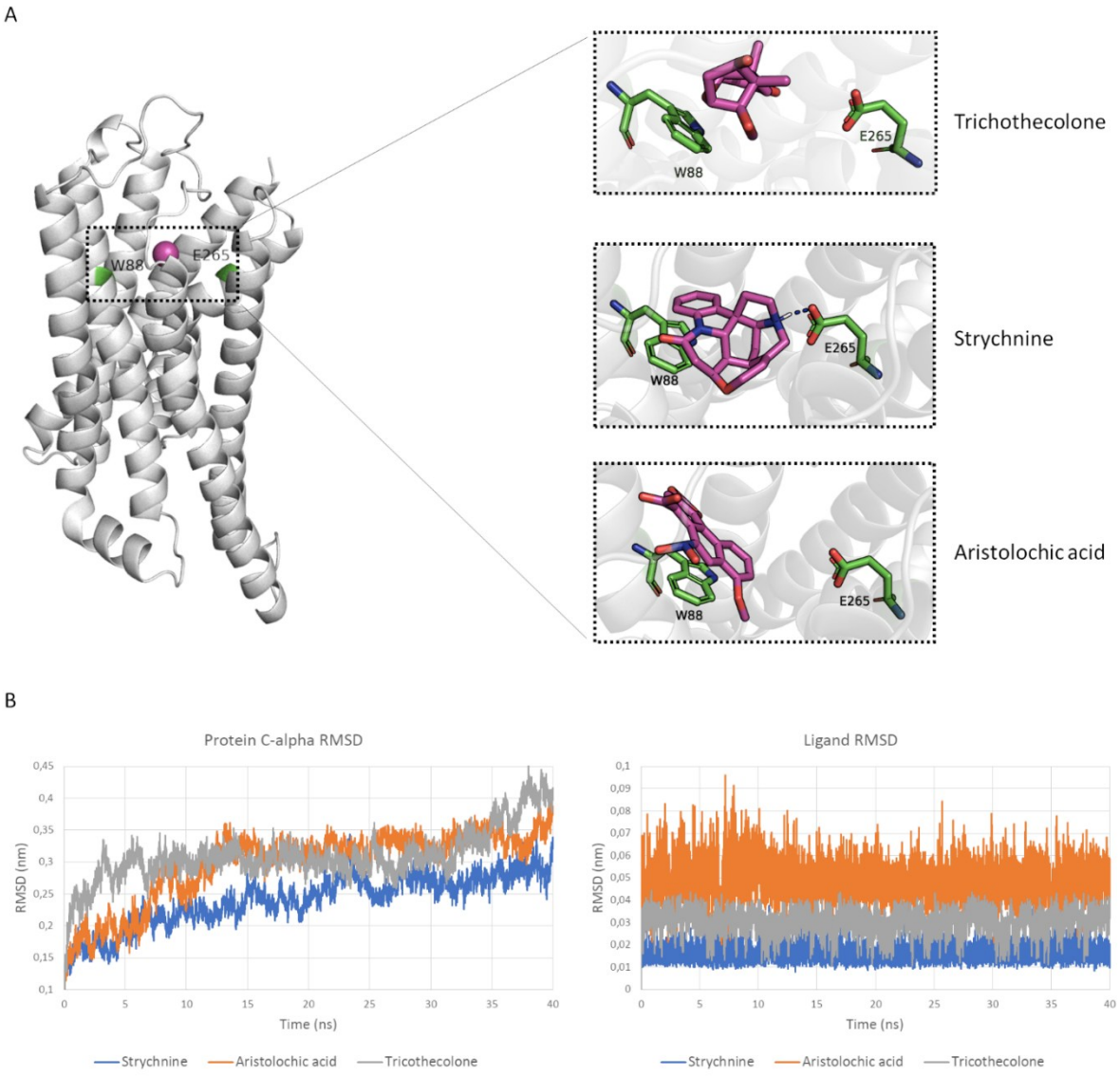


Figure S4. Molecular modelling results of trichothecolone, strychnine and aristolochic acid. **A.** Docking pose of trichothecolone, strychnine and aristolochic acid. On the left, TAS2R46 is represented in white cartoon. On the right, close-up of the ligand binding site with amino acid side chains and ligands represented as sticks. Blu dashed lines indicate polar contacts. **B.** On the left, protein C-alpha RMSD of TAS2R46 in complex with trichothecolone (average value 0.31 ± 0.04 nm), aristolochic acid (average value 0.297 ± 0.057 nm) and strychnine (average value 0.241 ± 0.041 nm). On the right, ligand RMSD of trichothecolone (average value 0.031 ± 0.005 nm), aristolochic acid (average value 0.049 ± 0.009 nm), and strychnine (average value 0.016 ± 0.004 nm).

Chapter V

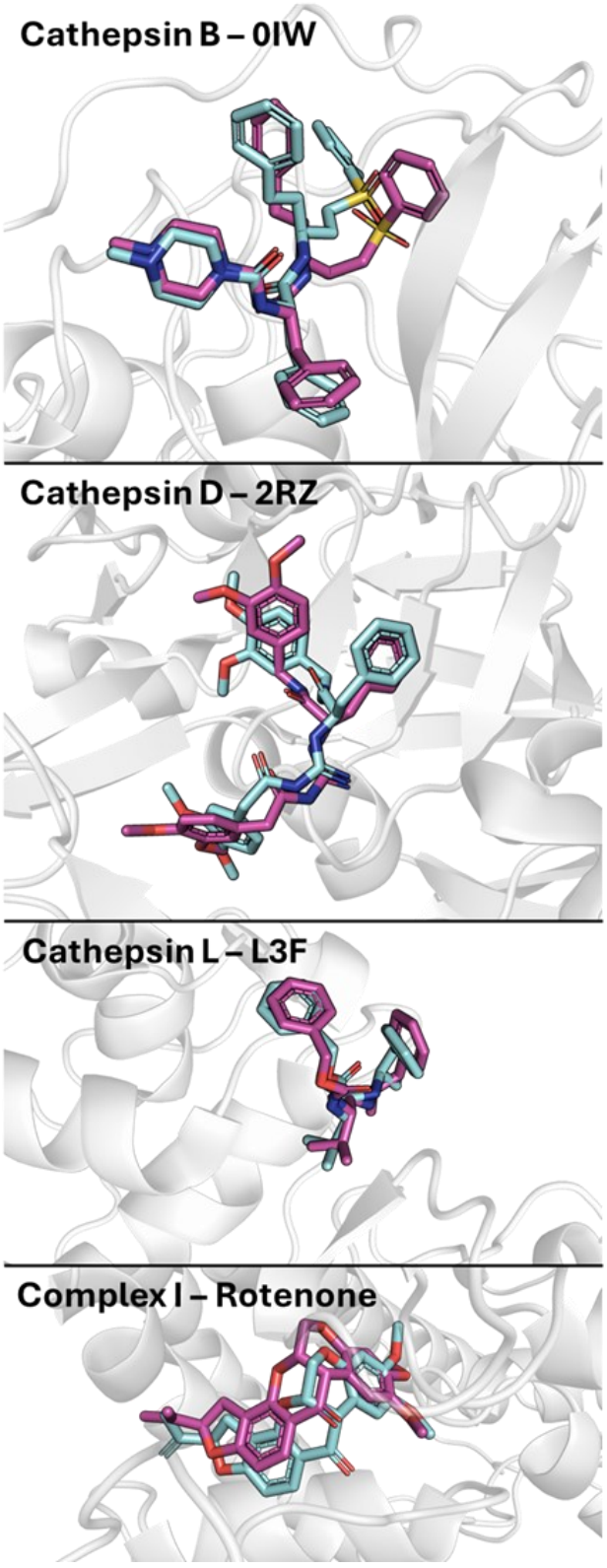


Figure S1. Binding poses of benchmarked ligands. Proteins are represented in cartoons while ligands in sticks (docking poses in cyan, structural data from PDB in magenta).

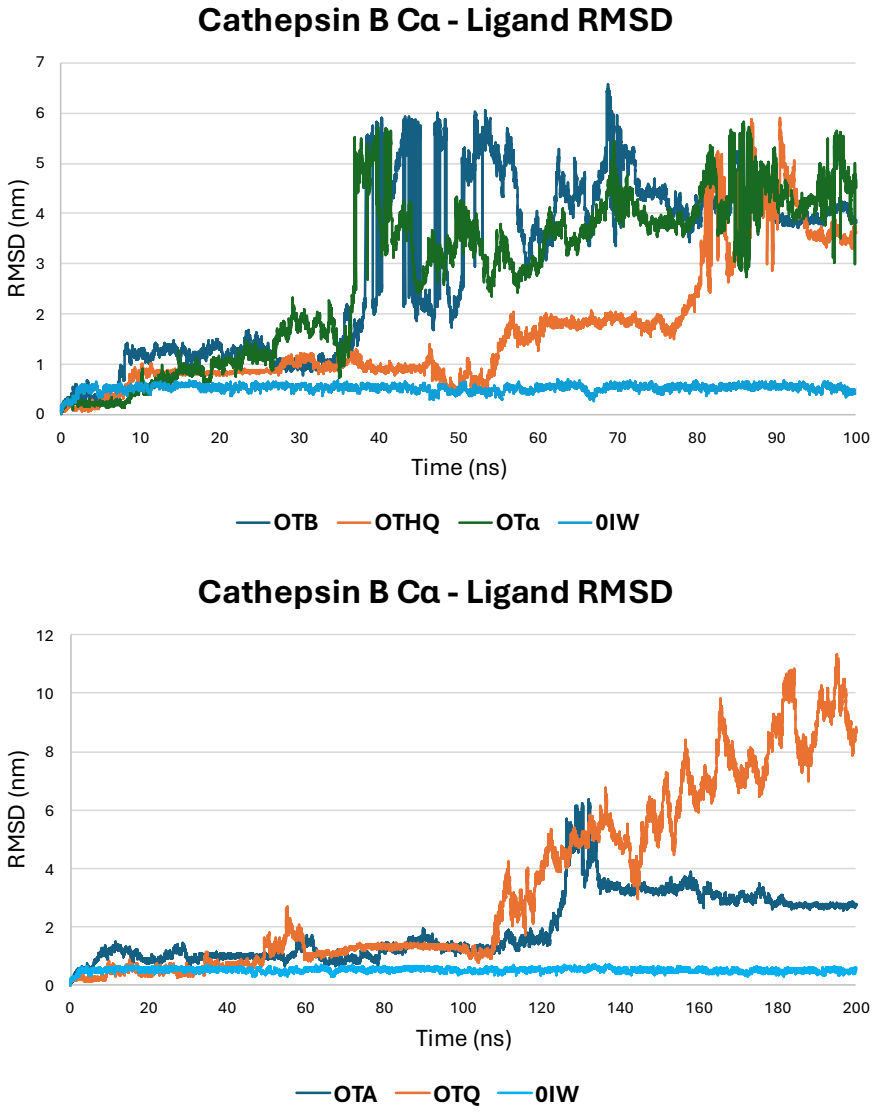


Figure S2. RMSD plot of OTA and congeners within cathepsin B. The increase of RMSD values over time indicates the movement of ligands away from the enzyme's catalytic site until their detachment. On the other hand, the benchmarked ligand 0IW stably contacted Cathepsin B for the whole MD production.

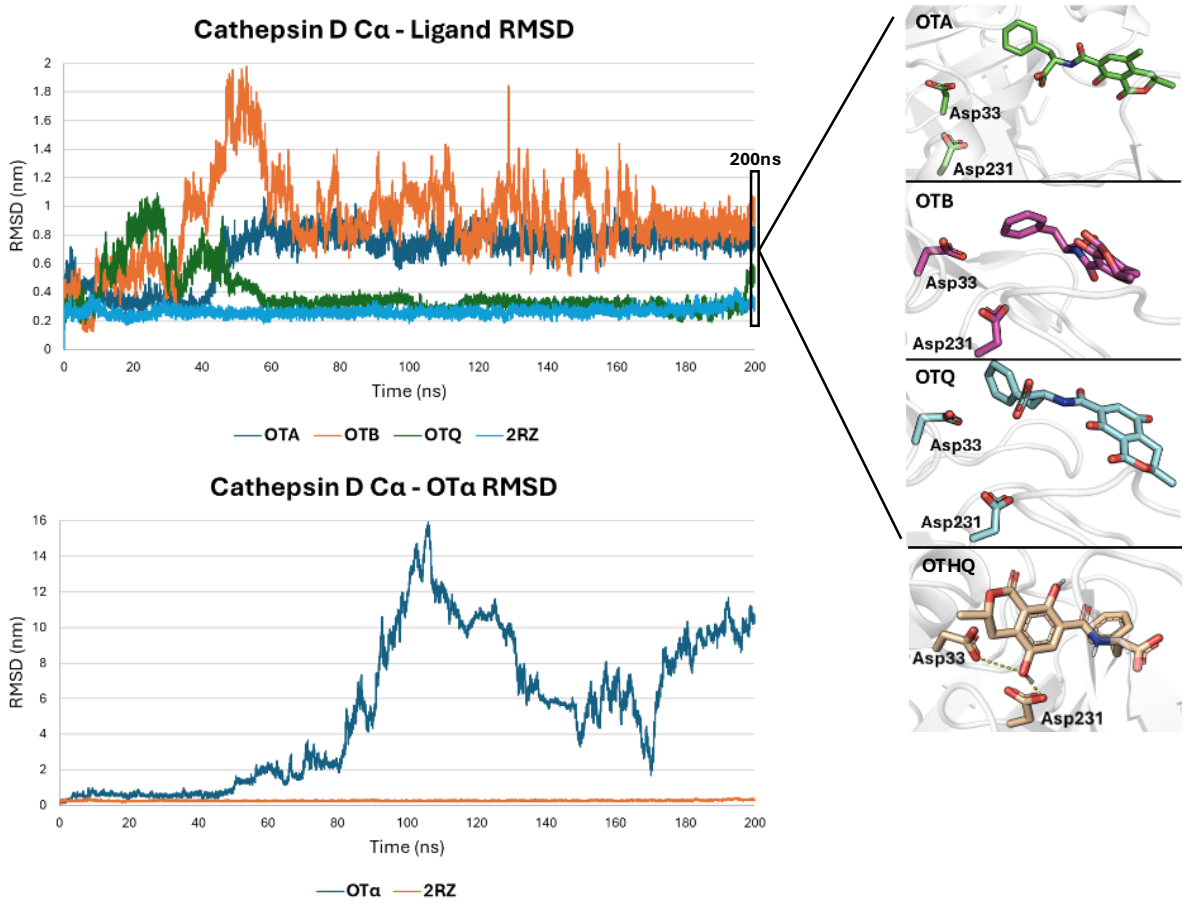


Figure S3. RMSD plot of OTA and congeners within cathepsin D. The increase of RMSD values over time for OTα indicates a clear detachment from the protein, while OTA and OTQ showed stability after an initial rearrangement (from 60 ns onward), though OTQ retraced better the trend of the benchmarked ligand 2RZ. Conversely OTB, although it did not detach from the protein, showed the worst stability. On the top right, a close-up of the last MD production frame (200ns) compared to OTHQ, for OTA (green sticks), OTB (magenta sticks), and OTQ (cyan sticks) showing a not proper orientation with respect to both the catalytic aspartate residues.

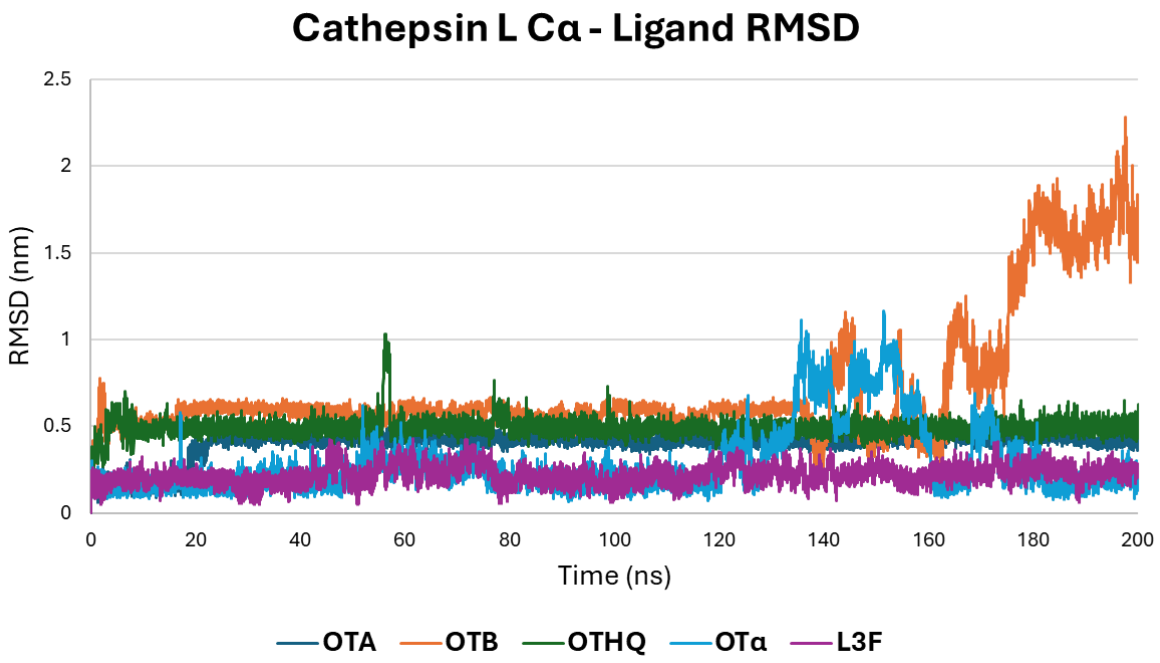


Figure S4. RMSD plot of OTA and congeners within cathepsin L. The increase of RMSD values over time of OTα and OTB indicates the movement of ligands away from the enzyme's catalytic site. OTA and OTHQ showed comparable stability.

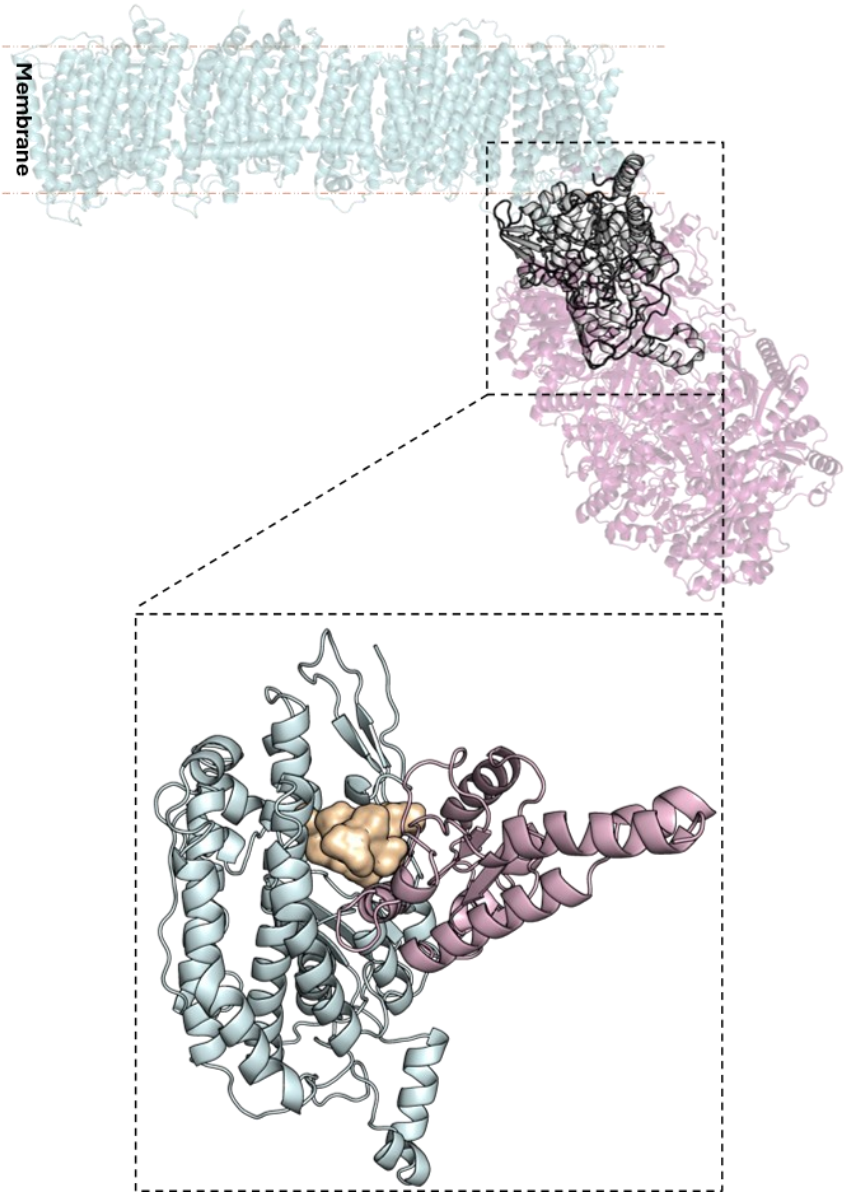


Figure S5. 3D representation of the mitochondrial Complex I. The proteins are represented in cartoon, while the binding pocket is represented in wheat surface. Protein chains not considered in the model are shown in transparency, while those considered (i.e. UniProt AC O75251 and O75306) are highlighted in the close-up.

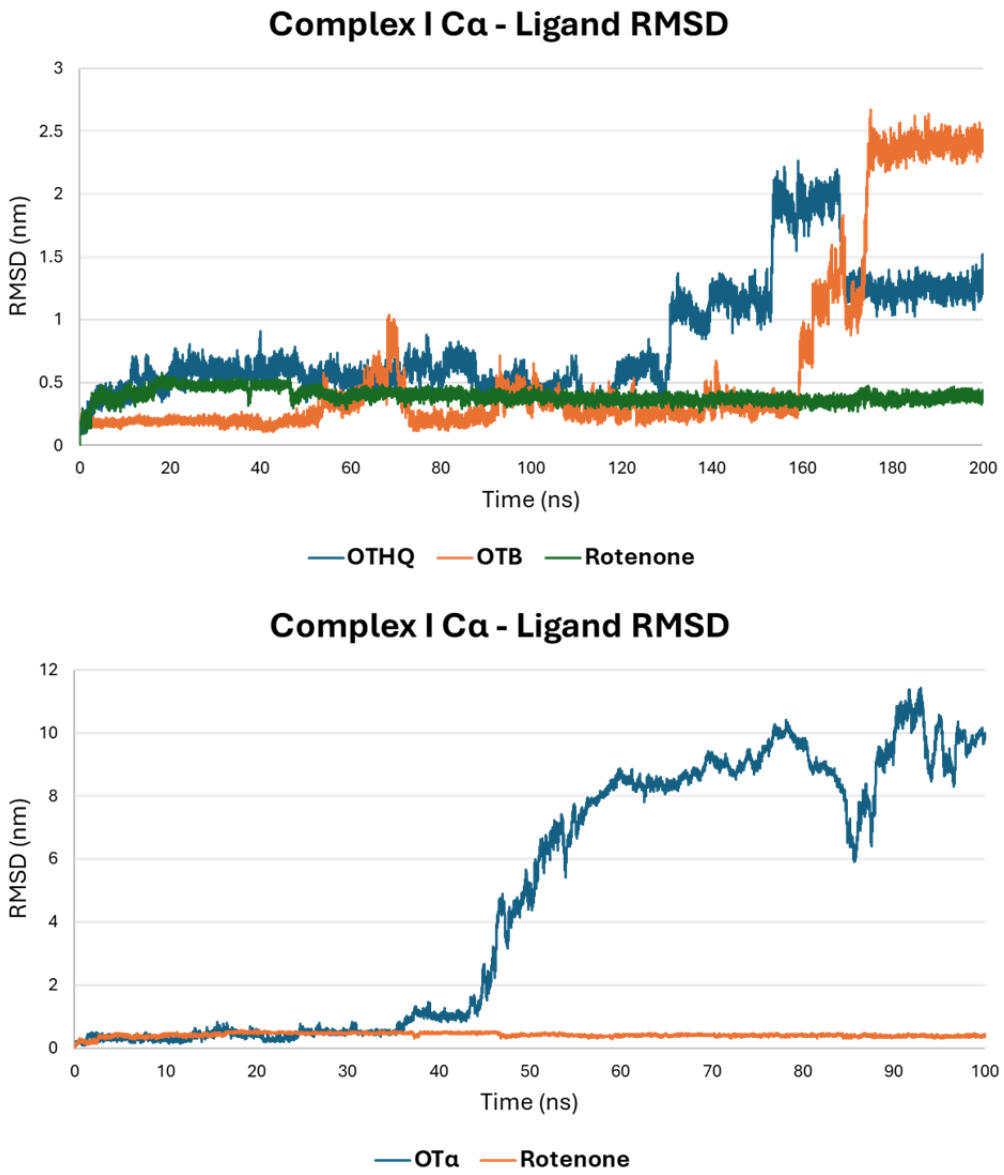


Figure S6. RMSD plot of OTA and congeners within the mitochondrial Complex I. The increase of RMSD values over time of OTα, OTHQ and OTB indicates the movement of ligands away from the enzyme's catalytic site, with a complete detachment for OTα.

Table S1. MMGBSA analysis results reported in kcal/mol**.**

	Cathepsin D	**Cathepsin L**	**Complex I**
OTA	-23.5	-25.55	-28.22
OTB	-17.58	np	np
OTHQ	-18.12	-26.74	-37.97
OTQ	-23.81	-22.77	np
Rotenone	np	np	-40.91
L3F	np	-37.75	np
2RZ	-60.44	np	np

Note: np stands for “calculation not performed”

Chapter VII

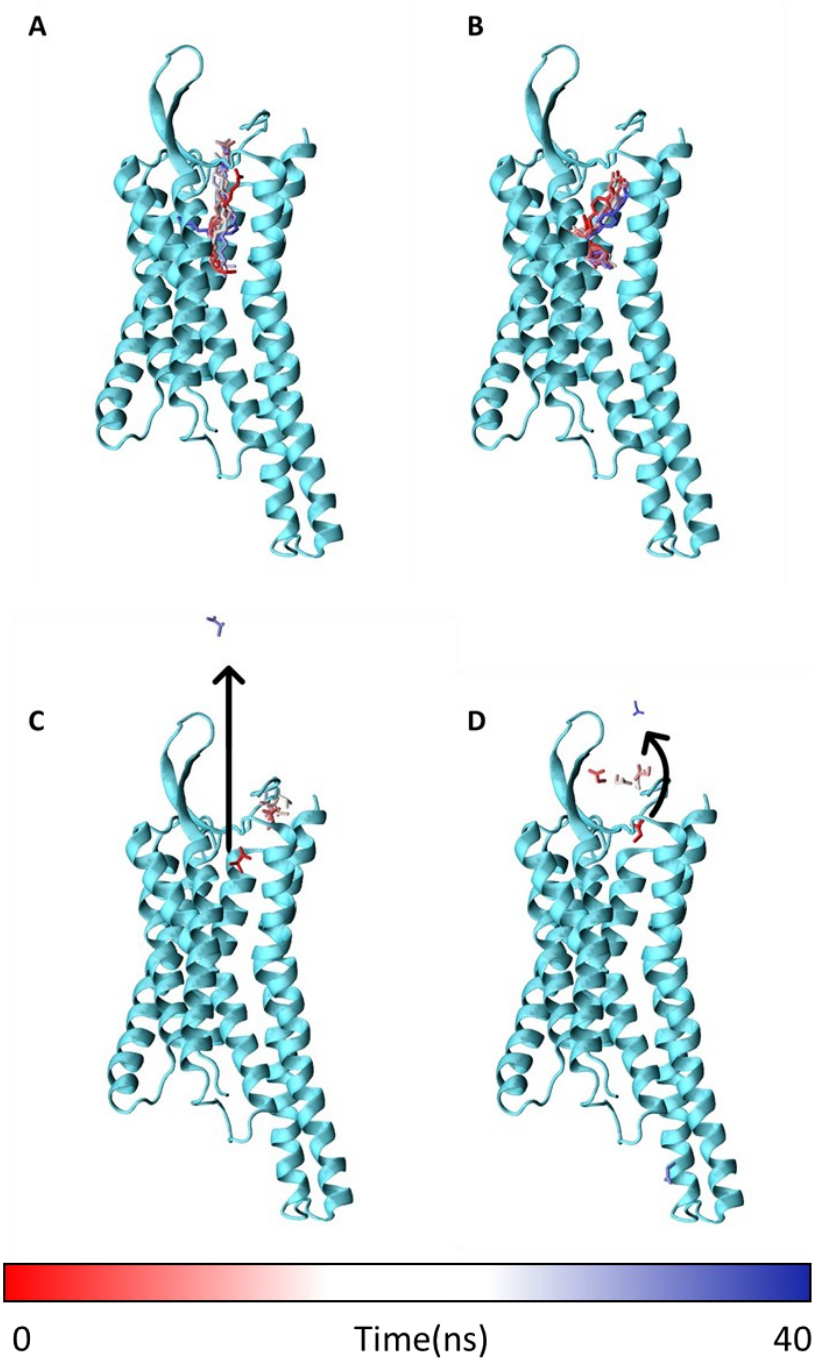


Figure S1. Trajectory of oleic acid (A), TUG-891 (B), isobutyric acid (C) and propionic acid (D) within GPR120 binding site. The protein is represented by cyan cartoon while each ligands by sticks. The red-to-blue transition reports the stepwise changes of coordinates for each ligand, represented as sticks, over the 40 ns CMD simulation. Of note, both isobutyric acid (C) and propionic acid (D) detached from the protein.

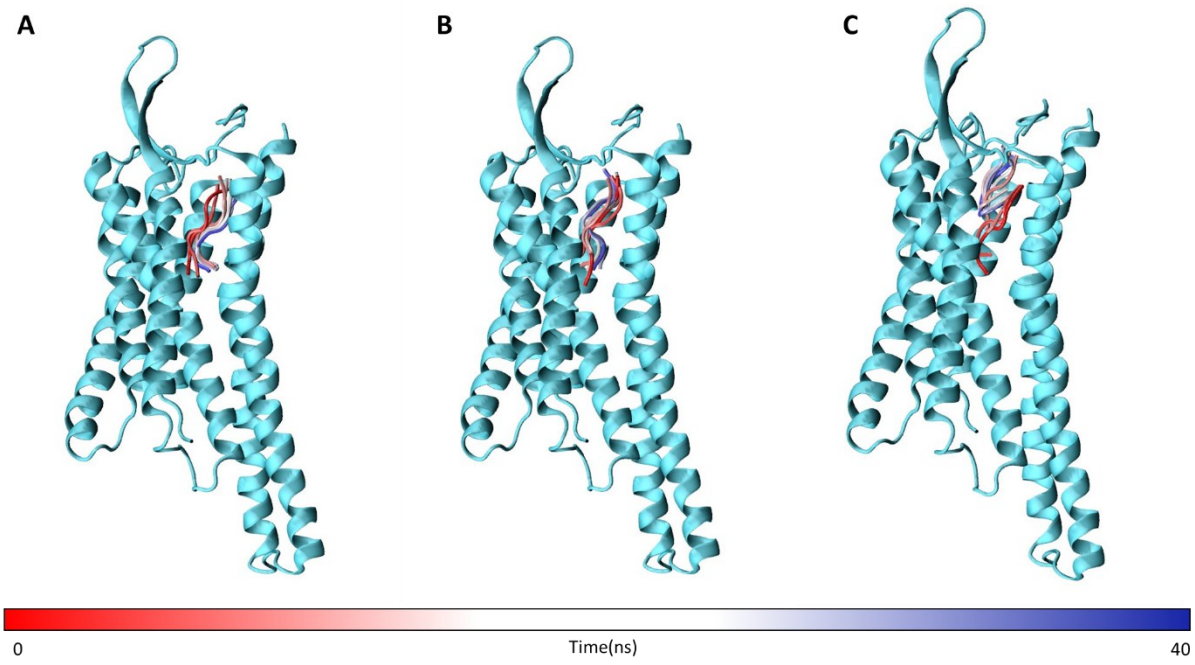


Figure S2. Trajectory of GIFGGG (A), GLFGGG (B) and GdIFGGG (C) within GPR120 binding site. The protein is represented by cyan cartoon. The red-to-blue transition reports the stepwise changes of coordinates for each peptide, represented as cartoon, over the 40 ns CMD simulation.

Acknowledgements

Acknowledgements

At the end of these three years of PhD I would like to take a moment to thank all the people that have been part of this journey.

First and foremost, I would like to thank my co-supervisors, Professor Luca Dellafiora for his guidance, for everything he has taught me over these years, and for the scientifical and personal growth he has encouraged in me.

A very special thanks to my supervisor Professor Gianni Galaverna, I'm grateful for the opportunity I have had.

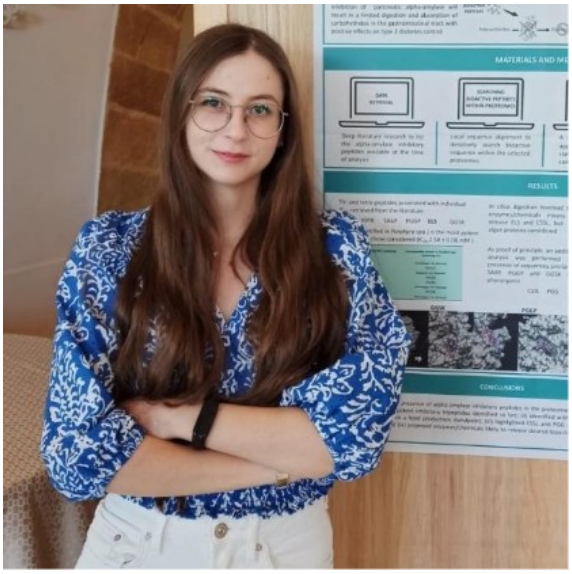
I would like to thank Lorenzo, my (desk)mate from day one.

A special thanks to the whole Molecular Modelling group at The Leibniz Institute for Food System Biology in Fresing. In particular, to Professor Antonella Di Pizio and Francesco Ferri, for their support during my period abroad in Munich and for contributing to my personal and professional development.

I would like to thank all the people who, at some point, have been part of this journey.

Infine, grazie alla fine famiglia. A mamma e papà, per non aver condiviso tutte le mie scelte, ma per avermi sempre e comunque sostenuta. A nonno Nindo e nonno Donato, per aver contribuito più di tutti a rendermi la persona che sono oggi.

About the author



Florinda Perugino was born in Lanciano (Italy) in May 2nd 1997, and raised in Santa Maria Imbaro. She earned her bachelor's degree in Food Science and Technology at The Polytechnic University of Marche (UNIVPM). In October 2022 she obtained her master's degree in Food Science and Technology at University of Parma defending an experimental thesis entitled "A bioinformatic approach to identify and study bioactive peptides focusing on pancreatic alpha-amylase and dipeptidyl peptidase IV inhibitors from microalgae and cyanobacteria". In November 2022 she joined the National PhD program in Artificial Intelligence, specialization area Agrifood and Environment, coordinated by the University of Naples Federico II. She carried out her research activity at the Department of Food and Drug at the University of Parma under the supervision of Professor Gianni Galaverna and Professor Luca Dellafiora. In 2025 she spent 4 months at The Leibniz Institute for Food System Biology at Technical University of Munich as visiting PhD student in the Molecular Modelling group led by Professor Antonella Di Pizio.

She is author of scientific papers published in ranked journals:

Pedroni, L.; Perugino, F.; Galaverna, G.; Dall'Asta, C.; Dellafiora, L. An In Silico Framework to Mine Bioactive Peptides from Annotated Proteomes: A Case Study on Pancreatic Alpha Amylase Inhibitory Peptides from Algae and Cyanobacteria. Nutrients 2022, 14, 4680 https://doi.org/10.3390/nu14214680.

Pedroni, L.; Perugino, F.; Kurtaga, A.; Galaverna, G.; Dall'Asta, C.; Dellafiora, L. The bitter side of toxicity: A big data analysis spotted the interaction between trichothecenes and bitter receptors, Food Research International, 2023, 173, https://doi.org/10.1016/j.foodres.2023.113284.

Perugino, F.; Pedroni, L.; Galaverna, G.; Dall'Asta, C.; Dellafiora, L. A mechanistic toxicology study to grasp the mechanics of zearalenone estrogenicity: spotlighting aromatase and the effects of its genetic variability. *Toxicology*, 501, 2024. doi.org/10.1016/j.tox.2023.153686

About the author

Pedroni, L., Perugino, F., Magnaghi, F., Dall'Asta, C., Galaverna, G., Dellafiora, L. Free fatty acid receptors beyond fatty acids: A computational journey to explore peptides as possible binders of GPR120, Current Research in Food Science, 8, 2024, 100710, doi.org/10.1016/j.crfs.2024.100710.

Perugino, F., Pedroni, L., Galaverna, G., Dall'Asta, C., Dellafiora, L. Virtual display of targets: A new level to rise the current understanding of ochratoxin A toxicity from a molecular standpoint, Toxicology, 503, 2024, 153765, doi.org/10.1016/j.tox.2024.153765.

Perenzoni, D., Dellafiora, L., Perugino, F., Vrhovsek, U., Piombino, P., Pittari, E., Guzzon, R., Moio, L., Galaverna, G., Mattivi, F. *Exploring Putative Kokumi Oligopeptides in Classic Sparkling Wines with a UHPLC-ESI-MS/MS Targeted Protocol (2024). J. Agric. Food Chem.,* 72, 47, 26189–26208, https://doi.org/10.1021/acs.jafc.4c08213.

Pedroni, L., Perugino, F., Dall'Asta, C., Galaverna, G., Buratti, F.M., Testai, E., Dellafiora, L. Computational journey to unveil organophosphorothioate pesticides' metabolism: A focus on chlorpyrifos and CYP2C19 mutational landscape. (2024). *Ecotoxicology and Environmental Safety,* 288*,* 117354, https://doi.org/10.1016/j.ecoenv.2024.117354.

Borgonovi, S.M., Perugino, F., Dellafiora, L., Annunziata, F., Pedroni, L., Galaverna, G., Pinto, A., Dallavalle, S., Iametti, S., Di Nunzio, M. Assessing the impact of food-derived bioactives on digestive proteases by in vitro and in silico approaches. (2025). Food Funct., 2025, 16, 2959, DOI: 10.1039/d4fo04022h.

Perugino, F., Pedroni, L., Galaverna, G., Dall'Asta, C., Dellafiora, L. Grasping the duality of fumonisins versus ceramide synthases: a computational approach to study their role as inhibitors and substrates from a mechanistic standpoint. (2025). Heliyon, 11, 13, e43725, https://doi.org/10.1016/j.heliyon.2025.

Pedroni, L., Perugino, F., Dall'Asta, C., Erviti, L.A., Mally, A., Vettorazzi, A., Dellafiora, L. (2025). Delving into the possible role of Ochratoxin A in Parkinson's disease: a computational study spotlighting the potential mechanisms of action. Food Research International, 221, 117532, https://doi.org/10.1016/j.foodres.2025.117532.

Borchers, J., Perugino, F., Schlosser, A. *et al.* Identification of small GTPases as potential target proteins of the mycotoxin and renal carcinogen ochratoxin A. *Arch Toxicol* (2025). https://doi.org/10.1007/s00204-025-04189-8

About the author

Conference Proceedings

An In Silico Framework to Mine Bioactive Peptides from Annotated Proteomes: A Case Study on Pancreatic Alpha Amylase Inhibitory Peptides from Algae and Cyanobacteria / **Perugino, F.**, Pedroni, L., Galaverna, G., Dall'Asta, C., Dellafiora L. Poster presentation Perugino at Chimali2023, Marsala, Italy.

A Bioinformatic Approach to Explore Zearalenone Estrogenicity: How a Single Nucleotide Variant in the Aromatase Enzyme Could Affect Its Inhibitory Activity / **Perugino, F.**, Pedroni, L., Tavellin, Galaverna, Dall'Asta, C., Dellafiora, L. *Poster presentation* at Eurotox2023, Ljubljana.

Unveiling the mechanisms of toxicity of enniatins – what's beyond their ionophoric characteristics? / **Perugino, F.**, Pedroni, L., Dall'Asta, C., Galaverna, G., Søderstrøm, S., Lie, K.K., Dellafiora, L. *Poster presentation* at 45th Mycotoxin Workshop, Vienna.

A computational approach to investigate ochratoxin A toxicity from a molecular standpoint / **Perugino, F.**, Pedroni, L., Galaverna, G., Dall'Asta, C., Dellafiora, L. *Oral presentation* at 14th International Conference "Mycotoxins and moulds", Bydgoszcz, Poland.

P05-36 May depsipeptide mycotoxins interfere with heme? An in silico case study on the possible impact of enniatin B and beauvericin on Atlantic salmon / **Perugino, F.**, Pedroni, L., Dall'Asta, C., Galaverna, G., Søderstrøm, S., Lie, K.K., Dellafiora, L., *Toxicology Letters*, 399, S146. *Poster presentation* at Eurotox2024, Copenaghen.

In silico framework for investigating the mechanics of kokumi taste: unveiling the chemistry to discover new kokumi compounds / **Perugino, F.**, Pedroni, L., Ferri, F., Dall'Asta, C., Mattivi, F., Di Pizio, A., Galaverna, G., Dellafiora, L. *Oral presentation* at ECRO2025, Bilbao, Spain.