UNIVERSITÀ DI PARMA

UNIVERSITA' DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN

"SCIENZE DEGLI ALIMENTI"

CICLO XXXVII

Computational approaches to investigate biological and chemical foodborne threats: exploring mechanisms of action and bioremediation strategies

Coordinatore:

Chiar.mo Prof. Massimiliano Rinaldi

Tutore:

Chiar.mo Prof. Luca Dellafiora

Co-tutore:

Chiar.mo Prof. Sergio Ghidini

Dottorando: **Lorenzo Pedroni**

Anni Accademici 2021/2022 – 2023/2024

Table of Contents

Outline

This PhD thesis frames into the field of Food Science and revolves around computational approaches, and their fit-for-purpose integration with experimental analysis, to tackle biological and chemical foodborne threats, as the title suggests, with two main objectives. The first one is to elucidate the mechanism of action of certain chemicals relevant to food safety and/or ecotoxicology, gaining more structural insights through molecular modelling-based studies, being critical for their risk assessment and subsequent management. The pipelines applied led to the unveiling of new possible targets for several compounds relevant to food safety and environmental health or to further clarify their mechanisms of action, metabolism, pharmaco/toxico-dynamic and -kinetic. The second objective aimed at identifying natural compounds possibly able to mitigate foodborne pathogens growth or action, not only in humans but also in animals and crops.

The thesis is organized in four Sections, as detailed below.

Section I reports a general introduction on the state of art of computational approaches in food science with a special focus on food chemistry and microbiology, and those associated aspects that are relevant to food hygiene. Moreover, it introduces briefly the methodologies used within the several case-studies reported in this PhD thesis. In the second part, the focus was moved to foodborne pathogens and chemical foodborne threats and on how computational approaches can deal with them also considering a One Health framework.

Section II focuses on foodborne pathogens, specifically Hepatitis E Virus, *Listeria monocytogenes* and *Campylobacter jejuni*. Here, in the context of developing anti-viral/bacterial strategies, computational approaches were applied to study virulence factors and proposing natural compounds possibly acting against them.

Section III reports a case-study regarding bioremediation. Here a hybrid *in silico* – *in vitro* approach was applied to find cytochromes able to hydroxylate an emerging mycotoxin, fusaric acid, whose hydroxylation is known to lower its phytotoxic action.

Section IV contains five chapters proposing new mechanistic insights on xenobiotics, either regarding their mechanism of action, possible structural insights on their metabolism and bioactivation or lastly new possible targets underpinning their action in living organisms.

Eventually, **section V** contains a conclusion based on the insights obtained during this three year-long PhD journey and a discussion of possible outlooks.

Section I

General Introduction

1.1. Computational Approaches in Food Science

In the past decade, the number of scientific papers directly associating the term *computational* to the term *food science* exponentially increased, nearly quadrupling in volume with an exceptionally rapid growth during the last five years. This trend is expected to remain stable or most likely to keep growing due to the advent and the rapid advancements of computational sciences, especially those related to artificial intelligence (AI), which nowadays are permeating every field of research [1-4].

Focusing on food science, *in silico* methods are getting more and more integrated into *traditional* wet lab studies, sometimes replacing them [5], for two main reasons: they allow the reduction of the experimental costs, both moneywise and timewise, and the attainment of previously inconceivable results [6-8]. What does the term *computational* refer to?

Computational, as described by one of the most used and famous large language models today available, refers to "anything related to the process of using computers or algorithms to perform calculations, solve problems, or simulate systems. In a broader context, it involves the use of computational techniques to model, analyze, or predict phenomena based on data and mathematical models. In scientific disciplines, computational often pertains to the use of computer-based methods to study complex systems that would be difficult or impossible to analyze purely through experimentation or theory" [9]. Keeping this in mind, the applications of computational approaches in food science are ubiquitous, ranging from food chemistry to food microbiology, and many more.

Food Chemistry. This is one of the widest and most crucial branches of food science focusing on food chemical composition, structure and overall properties along with the associated effects on human health. Although the main methodologies spread within this field are still confined to wet laboratories, *in silico* tools and computer-driven analysis are taking ground enhancing research

capabilities and promoting new discoveries [10, 11]. Recently, Jie Qian and co-workers developed an AI-driven platform, namely ChemSweet, to quickly detect sweet molecules while considering the biosafety of such newly detected sweeteners [6]. This represents a breakthrough in this field since this platform can screen a huge chemical space. This could not be easily achieved in a reasonable timeframe applying conventional techniques, occasionally affected by serendipitous discoveries or trial-and-errors achievements not always matching the needed safety criteria [12, 13]. Moreover, computational approaches allowed researchers to obtain more insights into less understood chemical aspects of food, including taste perceptions, such as kokumi [14-16]. As an example, this complex, and characterized by several facets, taste perception was (partially) explained through the interaction of kokumi compounds with the calcium sensing receptor [14, 17]. Specifically, Dellafiora and colleagues successfully applied an *in silico* pipeline to mechanistically investigate kokumi-active γ-glutamyl peptides to elucidate the molecular basis of their activity also paving the way for the possible identification of novel kokumi-like compounds [18]. These are just two of possible examples among many others investigating bitterness, sourness and many others taste perceptions currently on the process of being recognized as a taste *sensu stricto* (e.g. “fatty” and “starchy” tastes) [19-23]. Concerning another crucial aspect related to food chemistry – the bioactivity assessment of food compounds and/or xenobiotics – computational methodologies allowed to deepen a previous scarce understanding of certain compounds’ bioactivity and/or to pose solid basis to plan further dedicated and targeted investigations [10]. Regarding food compounds bioactivity assessment, there are many examples of fruitful applications of such approaches. They have been effectively exploited to assess the bioactivity of peptides, including the inhibition of the angiotensin-converting enzymes for the sake of modulating the blood pressure [24-27]. Moreover, the so called *dry* (*in silico*) methodologies can be used to investigate the xenoestrogenic potential of food components improving their understanding by taking into account the consequences of transformation upon food processing

and/or metabolism [28-31]. In addition to the assessment of bioactive compounds, the toxicological investigation of certain food-related molecules has been properly carried out through *in silico* tools possibly reducing – or making more targeted – *in vitro*/*in vivo* testing, lastly expanding our knowledge of overlooked compounds or newly detected toxicants and eventually improving the risk assessment process [32-34]. This last aspect is a key factor to offer a more comprehensive insight into the safety profiles of such compounds, lastly helping the competing foods safety authorities to update their guidance. To name one, Cozzini and colleagues performed a full *in silico* investigation, based on a combination of highly curated databases and consensus molecular docking, to evaluate most known food contact chemicals against the nuclear receptor family [5]. As they report, this is a crystal-clear example of how *in silico* driven studies might help in accelerating hazard identification.

Food Microbiology. It focuses on the study of food-related microorganisms including beneficial and harmful species. Regarding the beneficial side, this includes microorganisms responsible for food/beverage fermentation or those with probiotics properties. On the other hand, it focuses on both microorganisms accelerating food spoilage and on the so-called foodborne pathogens posing health risks either through the production of toxins or being pathogenic per-se. Once again, how computational approaches might help in this multi-faceted field? Also in this case, food microbiology may benefit of *in silico* analysis for many reasons and from different perspectives. To cite but a few, in 2023, Zhang and co-workers developed a data-driven machine learning-based platform, called FFExplorer [35]. This *in silico* tool enables to infer bioactive compounds production and/or food contaminants detoxification after fermentation exploring the potential application of certain microorganisms. Regarding foodborne pathogens, one interesting application is related to the analysis of their pathogenic factors, such as key proteins and/or toxins, ultimately identifying ways to disrupt or mitigate their harmful effects [36-39]. This can be achieved with computational methods, and eventually confirmed by targeted *in vitro* testing, providing structural insights into

how the pathogenic proteins might interact with molecules possibly inhibiting/interfering with their pathogen action. In this sense, some researchers started from a 16 000 compounds library to find potential *Fusarium verticillioides* aldehyde dehydrogenase inhibitors and reduced the number of compounds meant to be tested by two orders of magnitude preliminary screening them using molecular docking simulations [40].

Regarding this PhD thesis, the broad-meaning term *computational approaches* is mainly related to bioinformatics and computational biology/chemistry, primarily related to molecular modelling techniques. These include legacy homology modelling or the groundbreaking – and Noble prize winner for chemistry – AI-driven modelling AlphaFold [41], molecular docking and molecular dynamics simulations. Being the previously referred methodologies not so commonly used in the food science field, and to give the readers a brief introduction about them, here are some key points to keep in mind.

3D structure. The application of 3D *in silico* molecular models requires the 3D structure of the molecules under analysis. Indeed, the starting point is a reliable 3D structure, either of a protein or of a small molecule. Regarding protein 3D structures, the main way to collect them is searching the RCSB Protein Data Bank (PDB; https://www.rcsb.org/) [42], the gold standard for the retrieval of biological macromolecules 3D structures. It contains more than 200 000 3D structures mainly obtained through X-ray crystallography, nuclear magnetic resonance spectroscopy and cryo-electron microscopy. When the 3D structure needed is not present researchers might apply so-called homology/comparative modelling. This technique is based on the principle that 3D protein structures are more evolutionary conserved rather than the sequences themselves being the former more strongly associated to the protein function than the latter [43]. Following this principle, if the 3D structure of a related homologous protein is available (i.e. a protein that has a phylogenetic relationship with another), the latter might be used as a template to thread the primary sequence

of the protein of interest on that specific available structure. Nowadays, an enormous breakthrough has exploded thanks to the advancement of AI in the field of computational biology. This breakthrough is called AlphaFold and basically it allows researchers to feed this AI-tool with a protein primary sequence obtaining, in a really short time, the related 3D protein structure along with some metrics measuring the reliability of such result [44]. Despite these advances, both homology modelling and AlphaFold require the primary sequence of the query protein which is typically obtained from the UniProt database (https://www.uniprot.org/) [45]. On the other hand, small molecules are mainly stored in databases like PubChem (https://pubchem.ncbi.nlm.nih.gov/) [46], ChEMBL (https://www.ebi.ac.uk/chembl/) [47], Cambridge Structural Database (https://www.ccdc.cam.ac.uk/solutions/software/csd/) and many more. However, not all the small molecules come along with their 3D structure posing the problem of occasionally obtaining it with a certain degree of reliability. There are several tools available to achieve this such as a newly released one, namely dMXP [48], or a more traditional, less specific and precise, like Open Babel [49].

Molecular docking simulation. Once the 3D structure of the protein and of the small compounds of interest are collected, we can proceed with molecular docking simulations. This is an *in silico* methodology applied to investigate the binding event between the protein, usually referred to as receptor, and the molecules under investigation, typically called ligands. There are several available tools to perform such simulations, even for protein-protein or nucleic acid-protein interaction [50-52]. In principle, molecular docking aims to find the best way of interaction between two molecules. There are three main docking types:

1. *Rigid body docking*. It is the most simplistic and old-fashioned model: both receptor and ligand are evaluated as two rigid solid bodies. It is fast but the least accurate since it

neglects many aspects of the binding event, including a too approximate accounting of intra-molecular mobility (we don't have all the degrees of freedom of torsions).

2. *Semi-Flexible docking*. The receptor is rigid, but the ligand is fully flexible. It is usually exploited when investigating protein-small molecules interactions.
3. *Flexible docking*. Both molecules are flexible although with some limitations to make the accomplishment computationally affordable. It is slower since we are considering a higher number of degrees of freedom but therefore it is more accurate.

Despite the presence of differences among these models, they all apply two main steps to achieve a result. The first one is the conformational space search, where the ligand will be docked within the chosen binding site of the receptor, followed by a scoring procedure meant to score and rank the large amount of the docked poses. The main docking program adopted along this PhD thesis is GOLD (CCDC; https://www.ccdc.cam.ac.uk/). This software allowed to perform semi-flexible docking procedures, usually keeping fully flexible ligands and protein hydrogens free to rotate.

Molecular dynamics simulation. This technique offers a significant complement to docking simulations, enhancing the prediction reliability and giving deeper insights into molecular interactions [53]. Indeed, molecular dynamics (MD) simulations overcome the static nature of the docking procedure allowing to analyze the evolution of the obtained protein-ligand complex over time. This "dynamic" perspective provides further information about the stability or flexibility, i.e. the overall conformational changes, of the complex over time refining the docking outcomes. MD simulations are based on Newtonian mechanics, and they can be divided into several steps. The first step consists in defining the forcefield used to describe the complex under analysis. There are several sets of available forcefields, like CHARMM, GAFF or GROMOS forcefields [54-56], whose proper definition is fundamental to describe the interaction between atoms. Subsequently, we need to place the complex within a virtual box, solvate it with water molecules (using a water model like SPC,

SPC/E or TIP3P among many others [57]), and add an ion concentration useful to neutralize and simulate a as close as possible real-world system. This must be energetically minimized, equilibrated both in terms of temperature (NVT) and pressure (NPT). Eventually, the equilibrated system can proceed to the production phase where all the data regarding the complex under analysis will be collected and interpreted both quantitatively through several statistics, and qualitatively via an expert visual inspection of the resulting MD trajectory using dedicated software [58]. Apart from traditional MD simulations, another applied technique is umbrella sampling. Specifically, an element of the complex under analysis is pulled from the other along a certain reaction coordinate applying steered MD simulations. This improves the sampling of a system, eventually gaining information regarding the free energy of different conformations, as well as to study the inward/outward pathway of ligands [59].

1.2. Foodborne Pathogens and Xenobiotics: *In Silico* tools for One Health

Foodborne threats represent a significant global challenge, both in terms of public health and economic impact. As per the World Health Organization (WHO), there are an estimated 600 million people falling ill after consumption of contaminated food with more than 400 000 deaths every year. Dramatically, one third of these deaths occurred among children under five years of age.

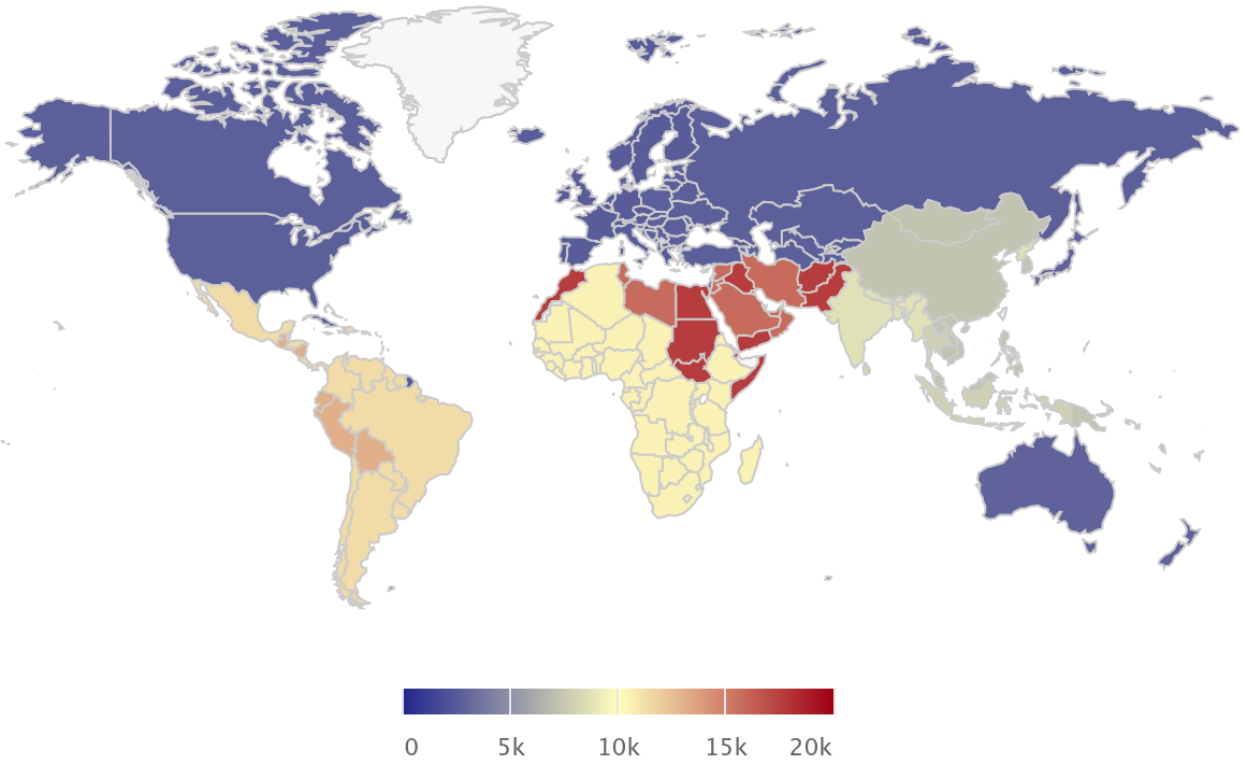


***Figure 1**. Estimated foodborne illnesses per 100 000 people (source Foodborne Disease Burden Epidemiology Reference Group; https://www.foodbornediseaseburden.org/ferg/estimates).*

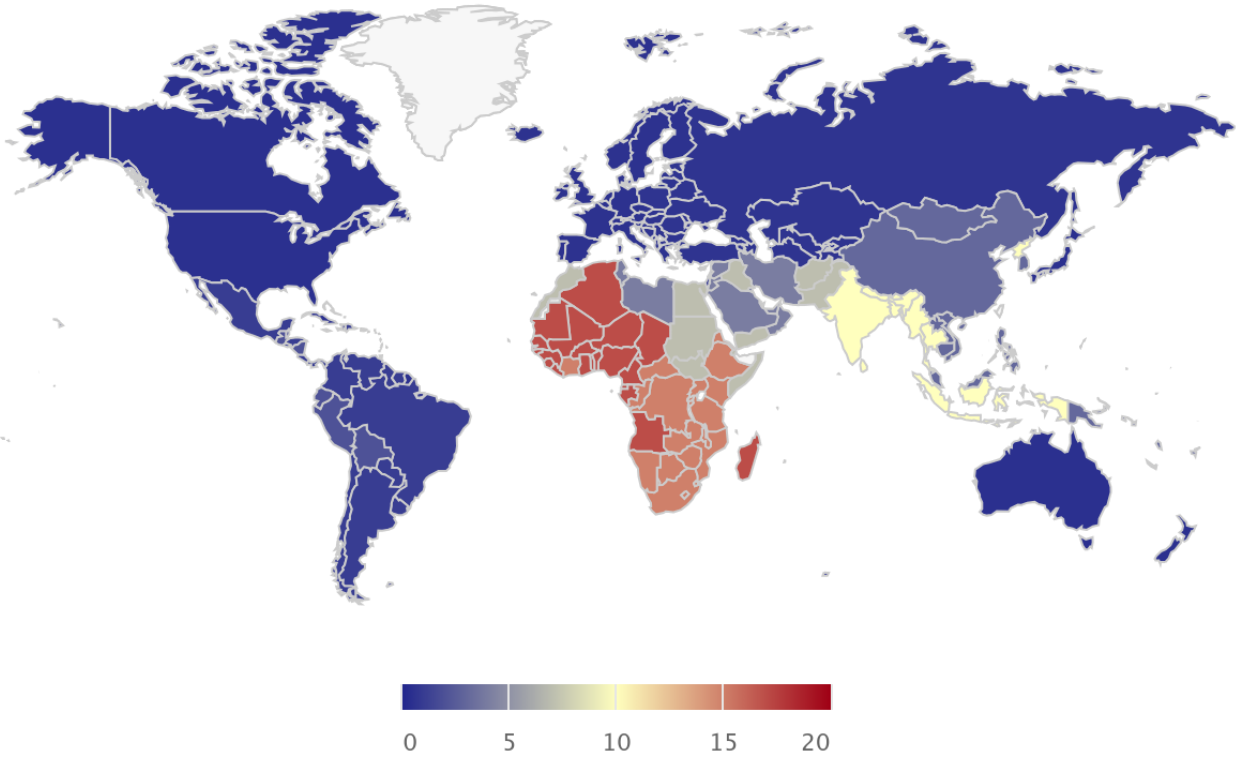


***Figure 2**. Estimated foodborne deaths per 100 000 people (source Foodborne Disease Burden Epidemiology Reference Group; https://www.foodbornediseaseburden.org/ferg/estimates).*

As shown in Figures 1 and 2, foodborne related illnesses/deaths are not equally distributed across the globe. They mostly strike Africa, South/Central America, Arabic peninsula, Iran, Afghanistan, Pakistan, India, China and Southeast Asian countries. Even if the reported data are based on a relatively old report published by WHO almost ten years ago [60] – with an updated version expected for 2025 – this is probably an underestimation of what is really happening. This is especially true for low- and middle-income countries. In fact, food safety is not prioritized as it should be even if, as per the World Bank 2019 report, the economic burden of foodborne diseases in such countries was estimated over 90 billion $/year with an annual cost of roughly 15 billion dollars associated to foodborne illnesses treatment [61]. There are several sources of foodborne threats including pathogens such as bacteria, viruses and parasites, or chemicals like xenobiotics as plant, fungal or

bacterial secondary metabolites, and anthropogenic contaminants as pesticides, dioxins and many others.

Bacteria. Among pathogens, bacteria are the most spread in food and water, giving rise to a series of common symptoms such as nausea, diarrhoea and fever [62]. According to a WHO factsheet released in October 2024, species belonging to *Salmonella*, *Campylobacter*, and *Escherichia coli* are some of the most common foodborne pathogens possibly leading to severe or fatal outcomes. However, there are many more concerning bacteria such as *Vibrio cholerae*, *Clostridium botulinum* or *Listeria monocytogenes*. These pathogens are commonly introduced within the food chain due to improper storage, handling or processing [63, 64]. Specifically, *Campylobacter jejuni* (see Section II, Chapter II) is one of the pathogenic species belonging to the *Campylobacter* genus. It is usually transmitted through contaminated and undercooked poultry or from raw and unpasteurised milk, however there can be other ways of infection such as meat, seafood and vegetables [65-67]. It leads to campylobacteriosis, an infection including gastroenteritis, fever, and diarrhoea. There are more than 1 million cases registered every year bringing to a significant burden of disease [68]. Apart from typical symptoms, there might be significant complications such as the Guillain-Barrè syndrome, leading to numbness, paralysis and eventually death [69]. Regarding *Listeria monocytogenes* (see Section II, Chapter III), it is a ubiquitous bacterium that can be found and persist everywhere even at refrigeration temperatures [70]. For this reason, contamination might occur during every step of food processing being especially true for dairy products, vegetables or generally ready-to-eat products meant to be consumed without cooking [71, 72]. Although the number of *Listeria*-related infection is way lower compared to more common foodborne pathogens, listeriosis' global burden remains significative. This is mainly due to the high mortality rate of this disease – sometimes up to 30% – especially in people belonging to the so-called YOPI group (young, old, pregnant, immunosuppressed) [73, 74].

Viruses. In addition to bacteria there are viruses such as *Norovirus*, Hepatitis A and E Viruses playing a relevant role in foodborne illnesses. These viruses might follow waterborne or foodborne/zoonotic transmission since they can remain environmentally stable even without their host [75-77]. However, depending on the socio-economic status of the considered country certain routes might be more probable than some others [78, 79]. Regarding Hepatitis E Virus (see Section II, Chapter I), it is the agent causing Hepatitis E and it accounts for an estimated 20 million infections per year with more than 3 million symptomatic cases worldwide [80]. Its overall fatality is relatively low, but it increases significantly for pregnant women [81]. Depending on the considered genotype, ranging from 1 to 8 so far, there might be preferential hosts leading to diverse routes of transmission [82, 83].

Xenobiotics. Beyond pathogens, xenobiotics represent another major issue. Theoretically, xenobiotics are chemical substances not naturally produced and not expected to be found within the biological system under analysis. They include anthropogenic substances, such as pesticides or more generally pollutants, and secondary plant, fungal or bacterial metabolites, possibly acting as beneficial bioactive compounds or toxicants [84, 85]. Regarding this PhD thesis, only harmful xenobiotics, namely mycotoxins – ochratoxin A, zearalenone, fusaric acid, trichothecenes – and alkenylbenzenes will be explored. Mycotoxins are secondary fungal metabolites embedding a huge variety of chemicals. They are environmental contaminants, and they are especially concerning for the food and feed chains. Their harmful effects might vary from nephrotoxicity, to hepatotoxicity, to cyto- or phytotoxicity [86, 87]. On the other hand, alkenylbenzenes are low-molecular weight secondary plant metabolites. Some of them, e.g. safrole, have been banned as food/feed additives since researchers proved their possible carcinogenic effects and overall adverse health outcomes [88, 89]. However, they are still naturally occurring in certain food commodities such as essential oils, basil along with related-products, and many others [90].

But how all the above might fit within the One Health framework? As per WHO, One Health is an "integrated, unifying approach that aims to sustainably balance and optimize the health of people, animals and ecosystems". In this framework the tight interconnection among humans, animals and environment, is emphasized and seek mainly through interdisciplinary strategies. Falling under the One Health umbrella, foodborne pathogens and toxicants are a clear example of threats with high priority since they can easily jump from the environment to both humans and animals. In this sense, the computational approaches applied in the case studies reported in this thesis might help in achieving such an enormous and really challenging objective. Starting from hepatitis E virus, it is especially concerning with respect to the swine food chain being pigs its main reservoir [91]. It has been reported most pigs within farms are seropositive for Hepatitis E Virus, with peaks over 90% depending on the region considered [92, 93]. This is true not only for low- to middle-income countries but also for Europe and especially for Italy. In fact, in some regions like Emilia-Romagna, one of the main sources of food-related income is linked to the production of hams and related meat products [94, 95]. For these reasons, as reported in Section II – Chapter I, finding possible feed additives or other food- and feed-grade molecules to lower the rate of infection within pig farms would be a great step to pursue the One Health perspective. Moving to foodborne toxicants, fusaric acid is a mycotoxin mainly produced by fungus belonging to the *Fusarium* genus having both phytotoxic and cytotoxic action [96-98]. It is considered an *emerging* mycotoxin since it has been poorly characterized and, most importantly, it has not been regulated yet, putting consumers and livestock at a major risk of exposure [99, 100]. Interestingly, some researchers found that its hydroxylation lower significantly its phytotoxicity [101]. As reported within Section III – Chapter I, we applied a mixed *in vitro* and computational pipeline finding an enzyme able to biotransform fusaric acid into its hydroxylated and less phytotoxic derivatives. The use of enzymatic inactivation represents a promising strategy to safeguard and bioremediate crops against the impact of fusaric

acid, mycotoxins and toxicants of various nature [102]. In this sense, this is crucial within the One Health framework since this can be considered a valid solution benefitting the ecosystem and consequently humans and animals. Indeed, toxicant mitigation from an agricultural standpoint contributes to both food/feed safety and environmental protection, reducing overall health risks [103-105]. Finally, obtaining structural and species-specific insights into concerning and environmentally spread chemicals, as mycotoxins or alkenylbenzenes with which both animals and humans might be exposed to, could significantly contribute for the achievement of One Health objectives. All these insights can pave the way for targeted interventions to promote environmental, animal and human health.

References

1. Esmaeily, R., M.A. Razavi, and S.H. Razavi, *A step forward in food science, technology and industry using artificial intelligence.* Trends in Food Science & Technology, 2024. **143**.
2. Miyazawa, T., et al., *Artificial intelligence in food science and nutrition: a narrative review.* Nutrition Reviews, 2022. **80**(12): p. 2288-2300.
3. Lui, R.Y., *Deus Ex Machina? The Rise of Artificial Intelligence in Toxicology.* Chemical Research in Toxicology, 2024. **37**(4): p. 525-527.
4. Singh, A.V., et al., *Integrative toxicogenomics: Advancing precision medicine and toxicology through artificial intelligence and OMICs technology.* Biomedicine & Pharmacotherapy, 2023. **163**.
5. Cozzini, P., et al., *Computational methods on food contact chemicals: Big data and in silico screening on nuclear receptors family.* Chemosphere, 2022. **292**.
6. Qian, J., et al., *ChemSweet: An AI-driven computational platform for next-gen sweetener discovery.* Food Chemistry, 2025. **463**.
7. Mangiatordi, G.F., et al., *Mind the Gap! A Journey towards Computational Toxicology.* Molecular Informatics, 2016. **35**(8-9): p. 294-308.
8. Zhang, Y., A.N.A. Aryee, and B.K. Simpson, *Current role of in silico approaches for food enzymes.* Current Opinion in Food Science, 2020. **31**: p. 63-70.
9. OpenAI, *ChatGPT.* 2024.
10. Wang, N., et al., *Recent advances in computational prediction of molecular properties in food chemistry.* Food Research International, 2024. **192**.
11. Pallante, L., et al., *On the human taste perception: Molecular-level understanding empowered by computational methods.* Trends in Food Science & Technology, 2021. **116**: p. 445-459.

12. Anjum, M.R., et al., *Lead acetate induced reproductive and paternal mediated developmental toxicity in rats.* Ecotoxicology and Environmental Safety, 2011. **74**(4): p. 793-799.

13. Juarez, G., et al., *Rotational Spectroscopy as a Tool to Characterize Sweet Taste: The Study of Dulcin.* Chemistryopen, 2024.

14. Maruyama, Y., et al., *Kokumi Substances, Enhancers of Basic Tastes, Induce Responses in Calcium-Sensing Receptor Expressing Taste Cells.* Plos One, 2012. **7**(4).

15. Toelstede, S., A. Dunkel, and T. Hofmann, *A Series of Kokumi Peptides Impart the Long-Lasting Mouthfulness of Matured Gouda Cheese.* Journal of Agricultural and Food Chemistry, 2009. **57**(4): p. 1440-1448.

16. Zhao, C.J., A. Schieber, and M.G. Gänzle, *Formation of taste-active amino acids, amino acid derivatives and peptides in food fermentations - A review.* Food Research International, 2016. **89**: p. 39-47.

17. Guha, S. and K. Majumder, *Comprehensive Review of γ-Glutamyl Peptides (γ-GPs) and Their Effect on Inflammation Concerning Cardiovascular Health.* Journal of Agricultural and Food Chemistry, 2022.

18. Dellafiora, L., et al., *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, 2022. **162**.

19. Goel, M., et al., *FlavorDB2: An updated database of flavor molecules.* Journal of Food Science, 2024.

20. Gu, Y.X., et al., *High-throughput discovery of umami peptides from pork bone and elucidation of their molecular mechanism for umami taste perception.* Food & Function, 2024. **15**(19).

21. Liu, S.H., et al., *Research on Bitter Peptides in the Field of Bioinformatics: A Comprehensive Review.* International Journal of Molecular Sciences, 2024. **25**(18).

22. Su, J.H., et al., *Integrating Computational and Experimental Methods to Identify Novel Sweet Peptides from Egg and Soy Proteins.* International Journal of Molecular Sciences, 2024. **25**(10).

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

24. Dellafiora, L., et al., *Hybrid in Silico/in Vitro Approach for the Identification of Angiotensin I Converting Enzyme Inhibitory Peptides from Parma Dry-Cured Ham.* Journal of Agricultural and Food Chemistry, 2015. **63**(28): p. 6366-6375.

25. Dellafiora, L., et al., *"Bottom-Up" Strategy for the Identification of Novel Soybean Peptides with Angiotensin-Converting Enzyme Inhibitory Activity.* Journal of Agricultural and Food Chemistry, 2020. **68**(7): p. 2082-2090.

26. Lammi, C., et al., *Mechanistic Insights into Angiotensin I-Converting Enzyme Inhibitory Tripeptides to Decipher the Chemical Basis of Their Activity.* Journal of Agricultural and Food Chemistry, 2022.

27. Lammi, C., et al., *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, 2021. **150**.

28. Cozzini, P. and L. Dellafiora, *In silico approach to evaluate molecular interaction between mycotoxins and the estrogen receptors ligand binding domain: A case study on zearalenone and its metabolites.* Toxicology Letters, 2012. **214**(1): p. 81-85.

29. Dellafiora, L., et al., *Modelling the possible bioactivity of ellagitannin-derived metabolites. In silico tools to evaluate their potential xenoestrogenic behavior.* Food & Function, 2013. **4**(10): p. 1442-1451.

30. Dellafiora, L., et al., *An in silico structural approach to characterize human and rainbow trout estrogenicity of mycotoxins: Proof of concept study using zearalenone and alternariol.* Food Chemistry, 2020. **312**.

31. Dellafiora, L., et al., *Molecular insights on xenoestrogenic potential of zearalenone-14-glucoside through a mixed in vitro/in silico approach.* Food and Chemical Toxicology, 2017. **108**: p. 257-266.

32. Benfenati, E., et al., *Integrating in silico models and read-across methods for predicting toxicity of chemicals: A step-wise strategy.* Environment International, 2019. **131**.

33. Rathore, A.S., et al., *ToxinPred 3.0: An improved method for predicting the toxicity of peptides.* Computers in Biology and Medicine, 2024. **179**: p. 108926.

34. Palazzolo, L., et al., *Development of in silico methodologies to predict the toxicity of novel proteins in the context of food and feed risk assessment.* EFSA Support. Publ., 2024. **21**(10).

35. Zhang, D.C., et al., *Data-Driven Prediction of Molecular Biotransformations in Food Fermentation.* Journal of Agricultural and Food Chemistry, 2023. **71**(22): p. 8488-8496.

36. Camargo, A., et al., *Unveiling the pathogenic mechanisms of Clostridium perfringens toxins and virulence factors.* Emerging Microbes & Infections, 2024. **13**(1).

37. Molina, R.D.I., et al., *Inhibition of bacterial virulence factors of foodborne pathogens by paprika (Capsicum annuum L.) extracts.* Food Control, 2022. **133**.

38. Abdelhamid, A.G. and N.K. El-Dougdoug, *Controlling foodborne pathogens with natural antimicrobials by biological control and antivirulence strategies.* Heliyon, 2020. **6**(9).

39. Martinovic, T., et al., *Foodborne pathogens and their toxins.* Journal of Proteomics, 2016. **147**: p. 226-235.

40. Ren, Z.G., et al., *Discovery of Aldehyde Dehydrogenase as a Potential Fungicide Target and Screening of its Natural Inhibitors against Fusarium verticillioides.* Journal of Agricultural and Food Chemistry, 2024. **72**(35): p. 19424-19435.

41. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.

42. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Research, 2000. **28**(1): p. 235-242.

43. Rost, B., *Twilight zone of protein sequence alignments.* Protein Engineering, 1999. **12**(2): p. 85-94.

44. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-+.

45. Bateman, A., et al., *UniProt: the Universal Protein Knowledgebase in 2023.* Nucleic Acids Research, 2023. **51**(D1): p. D523-D531.

46. Kim, S., et al., *PubChem 2023 update.* Nucleic Acids Research, 2023. **51**(D1): p. D1373-D1380.

47. Mendez, D., et al., *ChEMBL: towards direct deposition of bioassay data.* Nucleic Acids Research, 2019. **47**(D1): p. D930-D940.

48. Ai, H.P., et al., *dMXP: A De Novo Small-Molecule 3D Structure Predictor with Graph Attention Networks.* Journal of Chemical Information and Modeling, 2024. **64**(9): p. 3744-3755.

49. O'Boyle, N.M., et al., *Open Babel: An open chemical toolbox.* Journal of Cheminformatics, 2011. **3**.

50. He, J.H., et al., *HNADOCK: a nucleic acid docking server for modeling RNA/DNA-RNA/DNA 3D complex structures.* Nucleic Acids Research, 2019. **47**(W1): p. W35-W42.

51. Kozakov, D., et al., *The ClusPro web server for protein-protein docking.* Nature Protocols, 2017. **12**(2): p. 255-278.

52. Dominguez, C., R. Boelens, and A. Bonvin, *HADDOCK: A protein-protein docking approach based on biochemical or biophysical information.* Journal of the American Chemical Society, 2003. **125**(7): p. 1731-1737.

53. Santos, L.H.S., R.S. Ferreira, and E.R. Caffarena, *Integrating Molecular Docking and Molecular Dynamics Simulations*, in *Docking Screens for Drug Discovery*, W.F. DeAzevedo, Editor. 2019. p. 13-34.

54. Schmid, N., et al., *Definition and testing of the GROMOS force-field versions 54A7 and 54B7.* European Biophysics Journal with Biophysics Letters, 2011. **40**(7): p. 843-856.

55. Vanommeslaeghe, K., et al., *CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields.* Journal of Computational Chemistry, 2010. **31**(4): p. 671-690.

56. Wang, J.M., et al., *Development and testing of a general amber force field.* Journal of Computational Chemistry, 2004. **25**(9): p. 1157-1174.

57. Mark, P. and L. Nilsson, *Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K.* Journal of Physical Chemistry A, 2001. **105**(43): p. 9954-9960.

58. Humphrey, W., A. Dalke, and K. Schulten, *VMD: Visual molecular dynamics.* Journal of Molecular Graphics & Modelling, 1996. **14**(1): p. 33-38.

59. You, W.L., Z.Y. Tang, and C.E.A. Chang, *Potential Mean Force from Umbrella Sampling Simulations: What Can We Learn and What Is Missed?* Journal of Chemical Theory and Computation, 2019. **15**(4): p. 2433-2443.

60. World Health, O., *WHO estimates of the global burden of foodborne diseases*. 2016, Genève, Switzerland: World Health Organization. 265.

61. World, B. and S. Jaffee, *The safe food imperative*. Agriculture and food series. 2019, Washington, D.C., DC: World Bank Publications. 205.

62. Bintsis, T., *Foodborne pathogens.* Aims Microbiology, 2017. **3**(3): p. 529-563.

63. Sheng, L.N. and M.J. Zhu, *Practical in-storage interventions to control foodborne pathogens on fresh produce.* Comprehensive Reviews in Food Science and Food Safety, 2021. **20**(5): p. 4584-4611.

64. Augustin, J.C., et al., *Contribution of Foods and Poor Food-Handling Practices to the Burden of Foodborne Infectious Diseases in France.* Foods, 2020. **9**(11).

65. Kenyon, J., et al., *Campylobacter outbreak associated with raw drinking milk, North West England, 2016.* Epidemiology & Infection, 2020. **148**.

66. Jurinovic, L., et al., *First Data on Campylobacter spp. Presence in Shellfish in Croatia.* Pathogens, 2022. **11**(8).

67. Mohammadpour, H., et al., *The prevalence of Campylobacter spp. in vegetables, fruits, and fresh produce: a systematic review and meta-analysis.* Gut Pathogens, 2018. **10**.

68. Devleesschauwer, B., et al., *Health and economic burden of Campylobacter*. Campylobacter: Features, Detection, and Prevention of Foodborne Disease, ed. G. Klein. 2017. 27-40.

69. Nachamkin, I., B.M. Allos, and T. Ho, *Campylobacter species and Guillain-Barre syndrome.* Clinical Microbiology Reviews, 1998. **11**(3): p. 555-+.

70. Ferreira, V., et al., *Listeria monocytogenes Persistence in Food-Associated Environments: Epidemiology, Strain Characteristics, and Implications for Public Health.* Journal of Food Protection, 2014. **77**(1): p. 150-170.

71. Li, X., et al., *Prevalence of Listeria monocytogenes in Milk and Dairy Product Supply Chains: A Global Systematic Review and Meta-analysis.* Foodborne Pathogens and Disease, 2024. **21**(9): p. 526-535.

72. Alegbeleye, O. and A.S. Sant'Ana, *Survival and growth behaviour of Listeria monocytogenes in ready-to-eat vegetable salads.* Food Control, 2022. **138**.

73. Gambarin, P., et al., *Listeria monocytogenes in Ready-to-Eat seafood and potential hazards for the consumers.* Int. J. Microbiol., 2012. **2012**: p. 497635.

74. de Noordhout, C.M., et al., *The global burden of listeriosis: a systematic review and meta-analysis.* Lancet Infectious Diseases, 2014. **14**(11): p. 1073-1082.

75. Hofmeister, M.G., M.A. Foster, and E.H. Teshale, *Epidemiology and Transmission of Hepatitis A Virus and Hepatitis E Virus Infections in the United States.* Cold Spring Harbor Perspectives in Medicine, 2019. **9**(4).

76. de Graaf, M., J. van Beek, and M.P.G. Koopmans, *Human norovirus transmission and evolution in a changing world.* Nature Reviews Microbiology, 2016. **14**(7): p. 421-433.

77. Kumar, S., et al., *Hepatitis E virus: the current scenario.* International Journal of Infectious Diseases, 2013. **17**(4): p. E228-E233.

78. Waage, J., et al., *Changing food systems and infectious disease risks in low-income and middle-income countries.* Lancet Planet. Health, 2022. **6**(9): p. e760-e768.

79. Olaimat, A.N., et al., *Common and Potential Emerging Foodborne Viruses: A Comprehensive Review.* Life-Basel, 2024. **14**(2).

80. World Health Organization. *Hepatitis E*. 2021; Available from: https://www.who.int/news-room/fact-sheets/detail/hepatitis-e.

81. Webb, G.W. and H.R. Dalton, *Hepatitis E: an underestimated emerging threat.* Therapeutic Advances in Infectious Disease, 2019. **6**.

82. Dalton, H.R., N. Kamar, and J. Izopet, *Hepatitis E in developed countries: current status and future perspectives.* Future Microbiology, 2014. **9**(12): p. 1361-1372.

83. Okamoto, H., *Genetic variability and evolution of hepatitis E virus.* Virus Research, 2007. **127**(2): p. 216-228.

84. Filimonov, D., et al., *Assessment of the Xenobiotics Toxicity Taking into Account Their Metabolism*, in *Machine Learning and Deep Learning in Computational Toxicology*, H. Hong, Editor. 2023. p. 21-51.

85. Rana, A., et al., *Health benefits of polyphenols: A concise review.* Journal of Food Biochemistry, 2022. **46**(10).

86. Hussein, H.S. and J.M. Brasel, *Toxicity, metabolism, and impact of mycotoxins on humans and animals.* Toxicology, 2001. **167**(2): p. 101-134.

87. EFSA, *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, 2014. **12**(12): p. 3916.

88. IARC, *IARC monographs on the evaluation of the carcinogenic risks to humans. Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs* 1987.

89. Gotz, M.E., et al., *Myristicin and Elemicin: Potentially Toxic Alkenylbenzenes in Food.* Foods, 2022. **11**(13).

90. Eisenreich, A., et al., *Alkenylbenzenes in Foods: Aspects Impeding the Evaluation of Adverse Health Effects.* Foods, 2021. **10**(9).

91. Meester, M., et al., *Infection dynamics and persistence of hepatitis E virus on pig farms - a review.* Porcine Health Management, 2021. **7**(1).

92. Pavia, G., et al., *Seroprevalence and phylogenetic characterization of hepatitis E virus in pig farms in Southern Italy.* Preventive Veterinary Medicine, 2021. **194**.

93. Chandler, J.D., et al., *Serological evidence for swine hepatitis E virus infection in Australian pig herds.* Veterinary Microbiology, 1999. **68**(1-2): p. 95-105.

94. Mauceri, C., et al., *Hepatitis E in Italy: A silent presence.* Journal of Infection and Public Health, 2018. **11**(1): p. 1-8.

95. La Rosa, G., et al., *Hepatitis E Virus (Genotype 3) in Slurry Samples from Swine Farming Activities in Italy.* Food and Environmental Virology, 2017. **9**(2): p. 219-229.

96. Gruber-Dorninger, C., et al., *Emerging Mycotoxins: Beyond Traditionally Determined Food Contaminants.* Journal of Agricultural and Food Chemistry, 2017. **65**(33): p. 7052-7070.

97. E Selim, M. and N.A. El-Gammal, *Role of fusaric acid mycotoxin in pathogensis process of tomato wilt disease caused by Fusarium oxysporum.* J. Bioprocess. Biotech., 2015. **5**(10).

98. Mamur, S., et al., *Evaluation of the cytotoxic and genotoxic effects of mycotoxin fusaric acid.* Drug and Chemical Toxicology, 2020. **43**(2): p. 149-157.

99. López-Díaz, C., et al., *Fusaric acid contributes to virulence of Fusarium oxysporum on plant and mammalian hosts.* Molecular Plant Pathology, 2018. **19**(2): p. 440-453.

100. Brown, D.W., et al., *Identification of a 12-Gene Fusaric Acid Biosynthetic Gene Cluster in Fusarium Species Through Comparative and Functional Genomics.* Molecular Plant-Microbe Interactions, 2015. **28**(3): p. 319-332.

101. Crutcher, F.K., et al., *Detoxification of Fusaric Acid by the Soil Microbe Mucor rouxii.* Journal of Agricultural and Food Chemistry, 2017. **65**(24): p. 4989-4992.

102. Karlovsky, P., *Biological detoxification of fungal toxins and its use in plant breeding, feed and food production.* Natural Toxins, 1999. **7**(1): p. 1-23.

103. Xu, X.Y., et al., *Biodegradation strategies of veterinary medicines in the environment: Enzymatic degradation.* Science of the Total Environment, 2024. **912**.

104. Wang, F., et al., *Emerging contaminants: A One Health perspective.* Innovation, 2024. **5**(4).

105. Mackenzie, J.S. and M. Jeggo, *The One Health Approach-Why Is It So Important?* Tropical Medicine and Infectious Disease, 2019. **4**(2).

Section II

Foodborne Pathogens: *Orthohepevirus A*, *Campylobacter jejuni*, *Listeria monocytogenes*

This section is focused on the investigation of three known foodborne pathogens: *Orthohepevirus A*, commonly known as Hepatitis E Virus, *Campylobacter jejuni*, and *Listeria monocytogenes*.

The First Chapter investigates Hepatitis E Virus, a virus following zoonotic/foodborne or waterborne transmission. Computational approaches were here applied to model the RNA Helicase domain of a viral polyprotein to investigate the mechanism of inhibition of certain natural compounds.

Next, in the Second Chapter, the focus moved on *Campylobacter jejuni*. This foodborne bacterium is the leading cause of campylobacteriosis, affecting millions of people worldwide. Here, a hybrid *in silico – in vitro* pipeline was exploited to investigate the potential of dipeptides to inhibit a crucial enzyme for the bacterium survivability, i.e. the chorismate synthase, leading to unexpected results.

Last, *Listeria monocytogenes* will be the focus of the Third Chapter. Specifically, molecular modelling studies were used to deeply characterize a key interaction between a bacterial protein – Internalin A – and a human one – E-cadherin. Indeed, the interaction between these two proteins leads to bacterial internalization and infection. For this reason, the study investigated compounds of natural origin possibly capable of preventing this interaction and thus *Listeria monocytogenes* internalization.

Chapter I

In silico study on the Hepatitis E virus RNA Helicase and its inhibition by silvestrol, rocaglamide and other flavagline compounds

This chapter has been published as:

Pedroni Lorenzo, Dellafiora Luca, Varrà M. Olga, Galaverna Gianni, Ghidini Sergio. *In silico* study on the Hepatitis E virus RNA Helicase and its inhibition by silvestrol, rocaglamide and other flavagline compounds. Sci Rep. 2022 Sep 15;12(1):15512. doi: 10.1038/s41598-022-19818-w.

Abstract

Orthohepevirus A, commonly known as Hepatitis E Virus (HEV), follows waterborne or zoonotic/foodborne transmission. Genotype 3 HEV infections are worldwide spread, especially in swine populations, representing an emerging threat for human health, both for farm workers and pork consumers. Unfortunately, HEV *in vitro* culture and analysis are still difficult, resulting in a poor understanding of its biology and hampering the implementation of counteracting strategies. Indeed, HEV encodes for only one non-structural protein (ORF1), a multifunctional, multidomain, non-cleaved polyprotein, which is a potential druggable target, but whose crystallization and related structural inferences result unfeasible due to size-related reasons. In this context, an *in silico* approach based on molecular modelling succeeded to describe ORF1 RNA Helicase domain from a molecular standpoint allowing the identification of potential inhibitory compounds among natural plant-based flavagline-related molecules such as silvestrol, rocaglamide and derivatives thereof. In the context of scouting potential anti-viral compounds and relying on the outcomes presented, in-further dedicated investigations on silvestrol, rocaglamide and a promising oxidized derivative have been suggested.

1. Introduction

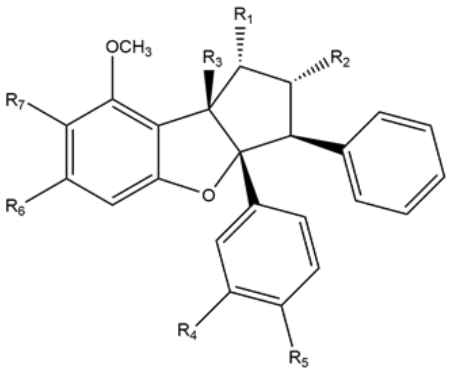
Hepatitis E virus (HEV) is a small positive sense single stranded RNA virus member of the *Hepiviridae* family (*Orthohepevirus* genus). It is the causative agent of the infamous Hepatitis E. As reported by the World Health Organization [1], there are 20 million HEV estimated infections per year and 3.3 million symptomatic cases of Hepatitis E worldwide. While the fatality rate is usually relatively low, ranging from 0.2% to 4%, it significantly increases for pregnant women [2].

The virus spreads following several routes with differences between high- and low-income countries: in the formers it is strictly related to zoonotic and foodborne transmission [3, 4] while in the latter it gives waterborne outbreaks [5]. There are eight genotypes described so far (HEV-1 to -8), differing for host preferences and ways of transmission. Particularly, HEV-3, which is globally spread, and HEV-4, which is mostly limited to Asia, follow zoonotic and foodborne transmission [6, 7]. Reservoirs of HEV include deer, wild boars, cows, sheep, and goats with evidence of human infection reported to be caused by contaminated milk [8-11]. Despite the wide range of potential hosts, the main reservoirs are pigs, and the related meat-based products are a major source of infection [12-14]. HEV was detected also in the berry fruit and leafy green vegetables supply chain due to bad agricultural practices as irrigation with contaminated water [15, 16].

Although this virus is widespread, *in vitro* culture and analysis are still difficult and consequently its molecular biology has not been fully understood yet [17]. One of the critical knowledge gaps concerns the HEV ORF1 gene, which encodes for a 185 kDa poly-protein with no cleavage sites reported [18]. ORF 1 is a multifunctional, multi-domain and non-structural poly-protein with a crucial role in the viral diffusion and replication, which may be a possible druggable target to interfere with the viral mechanisms of infection [19]. However, its huge dimension (185 kDa) and multi-domain architecture make its analysis challenging via canonical molecular biology and structural approaches.

In addition, the lack of available crystallographic and NMR structures hampers the identification and characterization of its druggable sites using canonical medicinal chemistry approaches. This lack of information ultimately prevents the structure-based identification of molecules targeting ORF1 domains for their possible implementation in anti-HEV strategies. In this context, *in silico* approaches can efficiently overcome the lack of structural data either to analyse proteins mechanics, their druggability or to provide a useful tool to study chemical and biological aspects of small molecules, including antiviral compounds [20-23]. For this reason, an *in silico* procedure has been developed and applied to target HEV ORF1. Based on previous studies showing the inhibitory activity of the natural plant secondary metabolite silvestrol (SLV; Figure 1) against HEV-3 [24-26], our study provided a reliable model to: I) investigate the underpinning mechanisms and viral target; and II) provide a predictive framework to estimate the activity of SLV analogues for further dedicated investigations. To do so, the HEV RNA Helicase domain was modelled and refined via a homology modelling approach based on an innovative, hybrid structure- and sequence-based big-data analysis targeting the whole set of crystallographic data available so far in the Protein Data Bank (nearly 190.000 structures; last database access 28th February 2022). Then, the interaction of the model with the already known RNA Helicase inhibitor rocaglamide (RCG) and a set of 9 natural-related compounds [27] was calculated through docking and molecular dynamics (MD) simulations.

A



Ligand	R1	R2	R3	R4	R5	R6	R7
RCG	OH	$CON(CH_3)_2$	OH	H	OCH_3	OCH_3	H
RCG1	OCHO	COOH	OH	H	OCH_3	OCH_3	H
RCG2	OH	$CON(CH_3)_2$	OCH_3	OH	OCH_3	OCH_3	H
RCG3	OH	$COOCH_3$	OCH_3	OCH_2O		OCH_3	H
RCG4	OAc	$CON(CH_3)_2$	OH	H	OCH_3	OCH_2O	
RCG5	OAc	$CON(CH_3)_2$	OH	OH	OCH_3	OCH_2O	
RCG6	OAc	$CON(CH_3)_2$	OH	OCH_3	OCH_3	OCH_2O	

B

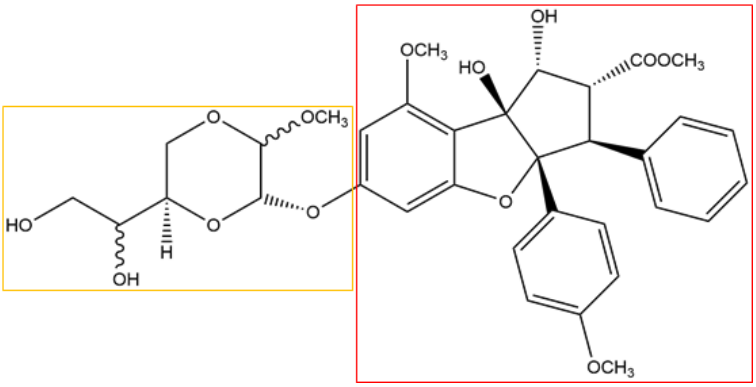


Figure 1. *Chemical structure of RCG, RCG analogues and SLV.* ***A.*** *RCG scaffold with its chemical substituents ranging from R1 to R7.* ***B.*** *SLV chemical structure. The red square surrounds the RCG scaffold while the yellow the typical SLV dioxane portion. The latter is replaced by a hydroxyl group (-OH) in S-Ag.*

2. Materials and Methods

2.1 Homology modelling

2.1.1 Sequence and template selection

The primary protein sequence used to model the 3D structure of HEV RNA Helicase was stored in the UniProt database (release 2021_04) with the Accession Code (AC) Q6J8G2 [28]. As no crystallographic structures of HEV RNA Helicase were available in the Protein DataBank at the time of analysis (last database access January 2022) a HEV RNA Helicase model was obtained via homology modelling. However, the fast-forward search of 3D structures homologous to HEV RNA Helicase to serve as useful template for homology modelling failed. Specifically, the search within the structures available in the Protein Data Bank (last database access January 2022) using the webserver BLAST interface (basic local alignment search tool, https://blast.ncbi.nlm.nih.gov/Blast.cgi) [29], which is a gold standard in this kind of study, gave no results. Then a hybrid structure- and sequence-based strategy was developed, as detailed below, to overcome this lack of results. A set of proteins annotated as RNA Helicases (E.C 3.6.4.13, according to Brenda classification [30]) in UniProt was collected grouping 228 non-eukaryotic sequences with available 3D structure in the Protein Data Bank. To retrieve the 3D structure of this set of proteins, their UniProt ACs were downloaded and mapped towards the Protein Data Bank using the "Retrieve/ID Mapping" tool available on UniProt. The 228 UniProt ACs mapped to 1972 PDB entries, whose 3D structures were iteratively downloaded in the .pdb format using a PyMol script developed in-house (available upon request). Of note, 1940 out of 1972 structures were available for the download while 32 were associated with deprecated codes or no longer accessible file. Each chain belonging to a pdb file was converted to a fasta formatted file using an *ad hoc* script developed in-house (available upon request) and subsequentially concatenated in a unique *multi.fasta* file. This

multi.fasta file was converted into a local database of sequences via the *makeblastdb* command set with default parameters (ncbi-blast+ suite; version 2.11.0+) [29]. Finally, the query sequence (HEV RNA Helicase) was searched within such database using the blast-p algorithm (ncbi-blast+ suite; version 2.11.0+) [29] with default parameters and setting the output in BLASTp tabular output format. The best hit in terms of alignment score (63.5 bits), e-value (1e-12) and identity percentage (nearly 30%) (i.e., PDB code 3WRY) was considered to develop the homology model of HEV RNA Helicase.

2.1.2 Model generation and refinement

Modeller version 10.0 interfaced to Chimera (version 1.15) [31] was used to generate the model of HEV RNA Helicase. The chosen template was the crystallographic structure with PDB code 3WRY (only the residues aligning the RNA Helicase domain were considered). Fifty models were generated, setting the inclusion of non-water heteroatoms, using a thorough optimization and choosing the best scored model according to zDOPE for subsequent analysis.

After computing a first draft of the model and checking its Ramachandran Plot with PROCHECK v3.5 [32], some regions were found with improper dihedrals (namely, residues 30-33; 101-103, 111-113 and 97-99). Therefore, they were stepwise refined with the Modeller loop-refinement tool (version 10.0 interfaced to Chimera [33] version 1.15) using the DOPE modelling protocol, generating 5 models and carrying forth the analysis of the best model according to the zDOPE scoring.

A final assessment of the model was performed by re-building its Ramachandran Plot and checking its Z-score on the ProSa-web Server [34] to verify the proper topology.

2.1.3 Building the model-RNA complex

The HEV RNA Helicase-RNA complex was obtained docking via ClusPro 2.0 [35] set with default parameters the RNA sequence to the previously built RNA Helicase model. The input RNA structure

was chosen based on the following protocol. First, all the structures of RNA Helicase containing RNA were download from Protein Data Bank and aligned to the model previously obtained. The RNA of the structure with the most similar organization (I.e., with the lowest RMSD value calculated in PyMol using the super command) was chosen (PDB ID 6JIM). In the last step, the RNA of 6JIM and the HEV RNA Helicase model were uploaded on ClusPro 2.0 [35]. Out of the obtained complexes the analysis of one of the mostly hydrophobic-favoured poses, showing the RNA molecules arranged in the same area as 6JIM structure, was carried forth. The sequence was than edited to a poly-purine fragment in agreement with previous studies reporting its suitability to interact with SLV-related compounds [36].

2.2 Docking

Docking studies were performed with GOLD (Genetic Optimization for Ligand Docking; version 2021.3) [37] to provide a plausible binding architecture for a set of flavagline compounds. The 3D structure of RCG (Figure 1A), SLV (Figure 1B) and SLV aglycone (S-Ag, Figure 1B) were downloaded from PubChem (https://pubchem.ncbi.nlm.nih.gov/; CID 331783, CID 11787114 and CID 24178739, respectively) [38] in the .sdf format. The other RCG analogues (Figure 1 A) were generated editing the RCG structure using the PyMol Builder tool (version 2.3.0) and further optimized using Chimera (version 1.15) [33] with the Minimize Structure tool (5000 steepest descent steps and 100 conjugate gradient steps).

The RNA Helicase model was used as input structure and the space to arrange ligands was set based on the architecture of binding of the RCG in the human eIF4A1 ATP-dependent RNA helicase having PDB code 5CZ9. The structure was visually aligned to the HEV RNA Helicase model, and the binding site was defined in a 10 Å-radius sphere around the centroid of the inter-bases space occupied by RCG in 5CZ9 structure (Figure 2). RCG was docked first generating 100 poses with no positioning

constraints, setting the ligand fully flexible and allowing polar protein hydrogens free to rotate. The best scored pose according to PLP Scoring function (256 units; the higher the score, the more probable and favoured the ligand interaction) showed a comparable binding architecture to the crystallographic binding pose of RCG. Such pose was then used as a position restraint setting the similarity option with shape overlap (weight constraint 200). This helped docking SLV and other analogues facilitating their proper arrangement into the binding site.

Virtual decoys were also generated to test procedure performances via the DUD-E database Generate tool (http://dude.docking.org/) [39]. The 50 decoys generated were ranked according to chemical similarities to SLV using LiSiCA algorithm [40] and the two extremes (the most similar and the most dissimilar compound; ZINC ID ZINC8584442 and ZINC8387186, respectively) (Figure 3) were chosen for subsequent analysis (see section 3.2 for further detail).

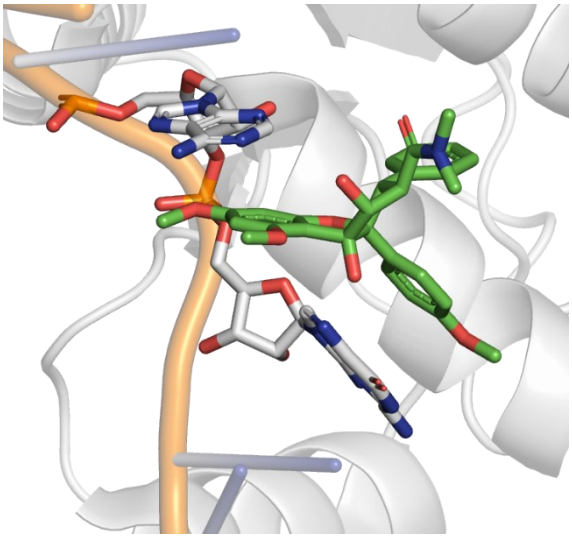


Figure 2. *RCG pose within the PDB 3D structure with code 5ZC9. The RCG is represented as green sticks while the RNA bases are in white sticks.*

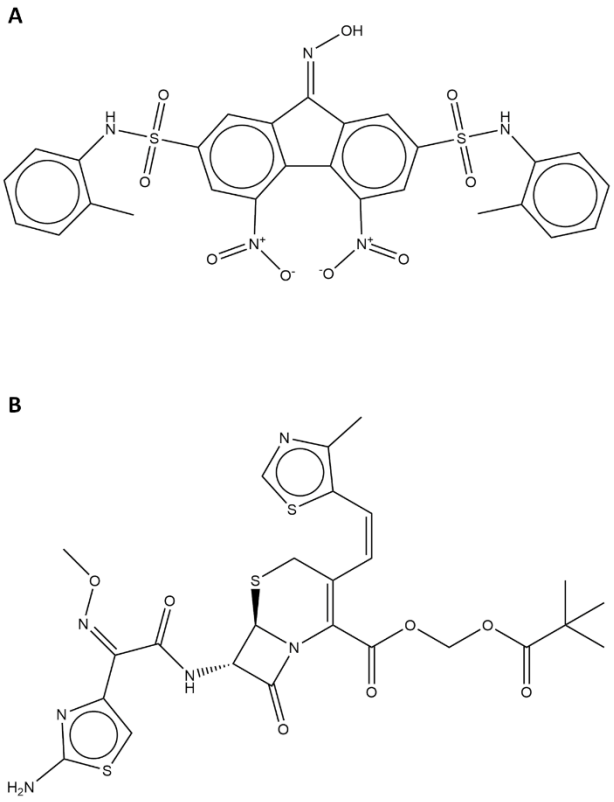


Figure 3. *Structures of the virtual decoys used in this study.* ***A****. ZINC8387186* ***B****. ZINC8584442*

2.3 Molecular Dynamics

Molecular dynamics (MD) simulations were performed to investigate the overall geometrical stability of HEV RNA Helicase-ligands complexes over time. The adopted software was GROMACS (version 2019.4) [41] with CHARMM27 all-atom force field parameters support [42]. All the ligands have been processed and parameterized with CHARMM27 all-atom force field using the SwissParam tool (http://www.swissparam.ch) [43]. Input structures were solvated with SPCE waters in a rhombic dodecahedron periodic boundary condition, and counter ions (Na^+ or Cl^-) were added to neutralize the system. Prior to performing molecular dynamic simulations, the systems were energetically minimized to avoid steric clashes and to correct improper geometries using the steepest descent algorithm with a maximum of 50000 steps. Afterwards, all the systems underwent isothermal (300 K, coupling time 0.1 ps) and isobaric (1 bar, coupling time 2 ps) 100 ps simulations before running a 25 ns MD simulation.

3. Results and Discussion

3.1. Building and refining the HEV RNA Helicase model

The HEV RNA Helicase domain of the HEV ORF1 poly-protein was meant to be modelled via homology modelling (HM), as reported in the 2.1 section, due to the absence of a crystallographic structure in the Protein Data Bank. This method, the relevance of which has been already proved [44], allows the successful modelling of proteins whose 3D structure is missing and it is particularly useful when crystallographic investigations are challenging, as in the case of poly-proteins.

Specifically, the HEV ORF1 domain's coding sequence of HEV RNA Helicase was chosen based on the results reported by Karpe and Lole [45] who succeeded to recombinantly express the putative HEV1 ORF1 RNA Helicase region (from amino acid 960 to 1204) proving its activity. The focus on HEV3 (UniProt AC Q6J8G2) was due to its prevalent foodborne/zoonotic transmission and its worldwide spread [46]. The localization of the RNA Helicase domain on the HEV3 ORF1 was achieved based on the alignment with the HEV1 RNA Helicase sequence proved as active by Karpe and Lole. Particularly, the HEV3 ORF1 region ranging from amino acid 975 to 1219 shared 90.6% identity and 95.5% similarity with the HEV1 RNA Helicase domain (Figure 4), becoming our target sequence (Tar-Seq).

```
HEV3_Helicase    975 GCTISPGIVHYQFTAGVPGSGKSRSIQQGDVDVVVVPTRELRNSWRRRGF   1024
                     ||.::||:|.|||||||||||||.|.|||||||||||||:||||||
HEV1_Helicase    960 GCRVTPGVVQYQFTAGVPGSGKSRSITQADVDVVVVPTRELRNAWRRRGF   1009

HEV3_Helicase   1025 AAFTPHTAARVTIGRRVVIDEAPSLPPHLLLLHMQRASSVHLLGDPNQIP   1074
                     ||||||||||||.|||||||||||||||||||||::||||||||||||
HEV1_Helicase   1010 AAFTPHTAARVTQGRRVVIDEAPSLPPHLLLLHMQRAATVHLLGDPNQIP   1059

HEV3_Helicase   1075 AIDFEHAGLVPAIRPELAPTSWWHVTHRCPADVCELIRGAYPKIQTTSRV   1124
                     ||||||||||||||:|.||||||||||.|||||||||||||.|||||||
HEV1_Helicase   1060 AIDFEHAGLVPAIRPDLGPTSWWHVTHRWPADVCELIRGAYPMIQTTSRV   1109

HEV3_Helicase   1125 LRSLFWNEPAIGQKLVFTQAAKAANPGAITVHEAQGATFTETTIIATADA   1174
                     ||||||.|||:||||||||||.||||::|||||||||:||||||||||
HEV1_Helicase   1110 LRSLFWGEPAVGQKLVFTQAAKPANPGSVTVHEAQGATYTETTIIATADA   1159

HEV3_Helicase   1175 RGLIQSSRAHAIVALTRHTEKCVILDAPGLLREVGISDVIVNNFF   1219
                     |||||||||||||||||||||||:|||||||||||||.||||||
HEV1_Helicase   1160 RGLIQSSRAHAIVALTRHTEKCVIIDAPGLLREVGISDAIVNNFF   1204

Length: 245
Identity:     222/245 (90.6%)
Similarity:   234/245 (95.5%)
Gaps:           0/245 ( 0.0%)
Score: 1179.0
```

Figure 4. *Sequence alignment of the HEV1 RNA Helicase with the HEV3 RNA Helicase.*

Concerning the template selection for HM, it is typically obtained identifying homologous sequences with available 3D structure via the BLAST UniProt web-interface, a gold benchmark standard in this kind of studies [47]. However, using the Tar-Seq as input for a canonical BLAST search gave no results. This was likely due to the nature of the BLAST heuristic algorithm that could not retrieve a best local alignment between a given poly-protein domain and the whole sequence of poly-proteins, which was rather required for the sake of the project. Indeed, the BLAST UniProt web-interface searches target sequences within the whole primary protein sequences associated with 3D data and not limiting the search to the actual sequence resolved in the crystal structure. However, in most of the medium/large proteins, such as poly-proteins, the crystals related to the UniProt ACs partially cover the protein primary structure causing BLAST to fail in identifying homologs with known structures. To overcome the intrinsic weakness of a canonical BLAST-based approach in this kind of study, a successful workflow, with all the steps reported in the 2.1.1 section, was setup. The change in the search space used in this approach made it robust, reusable, and successful in finding a template to build a model for a poly-protein domain. Downloading all the PDBs having a 3D structure belonging to the Prokaryotic reign with E.C 3.6.4.13 (RNA Helicase activity, according to Brenda classification [30]) and converting them to FASTA files via an in-house script allowed to build a database made of the “crystallome” of the prokaryotic RNA Helicase. Afterward, searching it to find the template, as reported in the 2.1.1 section, led to obtain various significant hits.

The best aligning protein was chosen as template (PDB ID 3WRY) [48]. Such protein belongs to the *Tomato mosaic virus* (Tmv), a positive sense single stranded RNA virus belonging to the same HEV class. Before using it as template to model the HEV RNA Helicase domain, the last 160 N-terminal residues, which were organized in a self-standing sub-domain, were removed being not covered by our Tar-Seq. The model and its refinement process, reported in the 2.1.2 section, allowed to obtain a reliable and stable model. Indeed, at the end of the topological refinement steps, 90% of the

residues were within the mostly favoured regions, 10% in allowed regions and there were no residues in generously or disallowed regions according to the Ramachandran plot (Figure 5).

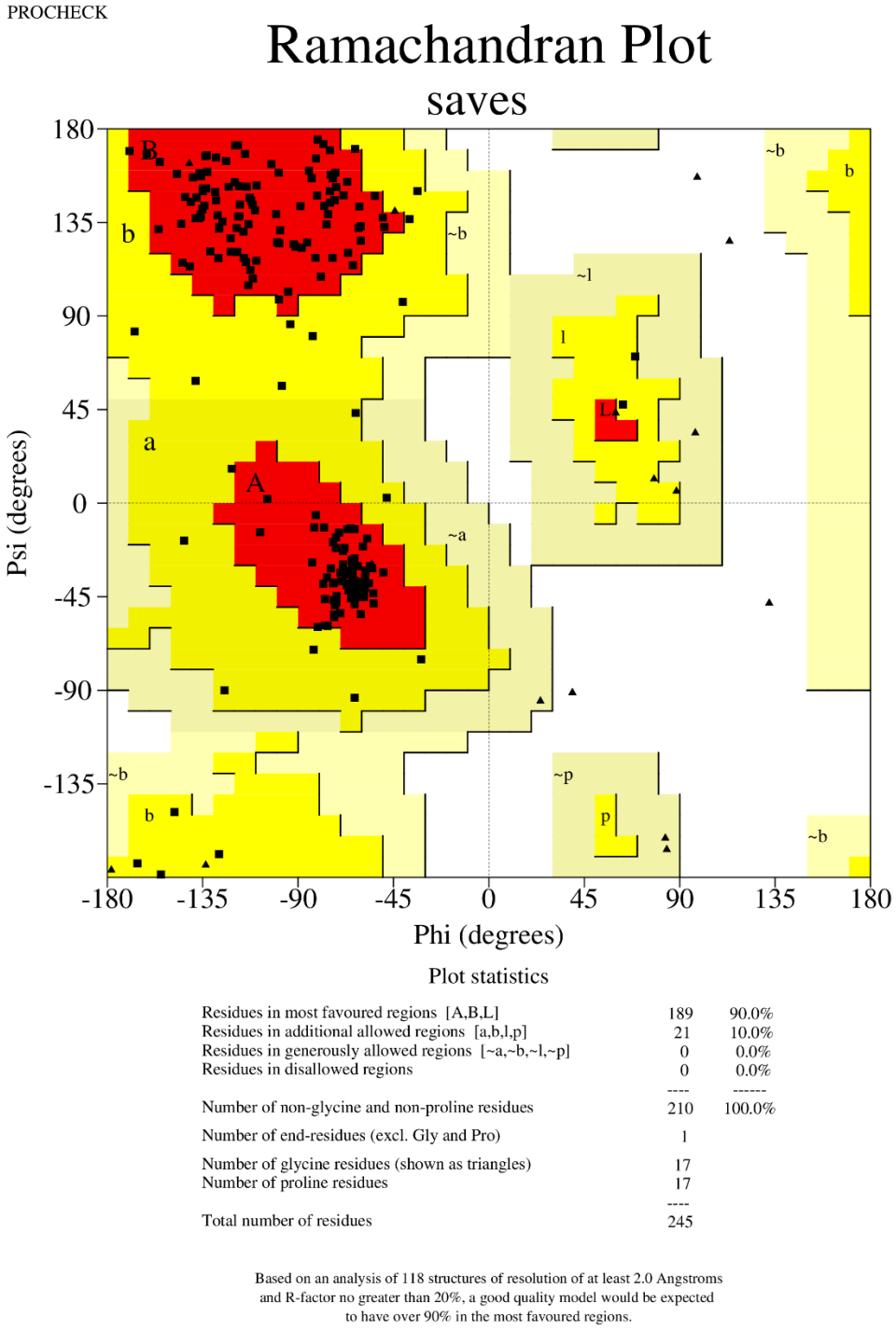


Figure 5. *Ramachandran Plot of the HEV RNA Helicase model computed via PROCHECK v3.5.*

Furthermore, the model was also checked on the ProSA-web Server obtaining a Z-score of 6.14 which is within the range of scores typically found for native proteins of comparable size [34] further confirming the model reliability. The last check was the comparison of the model with the one produced by the blind web resource trRosetta [49]: their structural alignment resulted in an RMSD lower than 2.6 Å, confirming the reliability of the entire procedure described above (Figure 6).

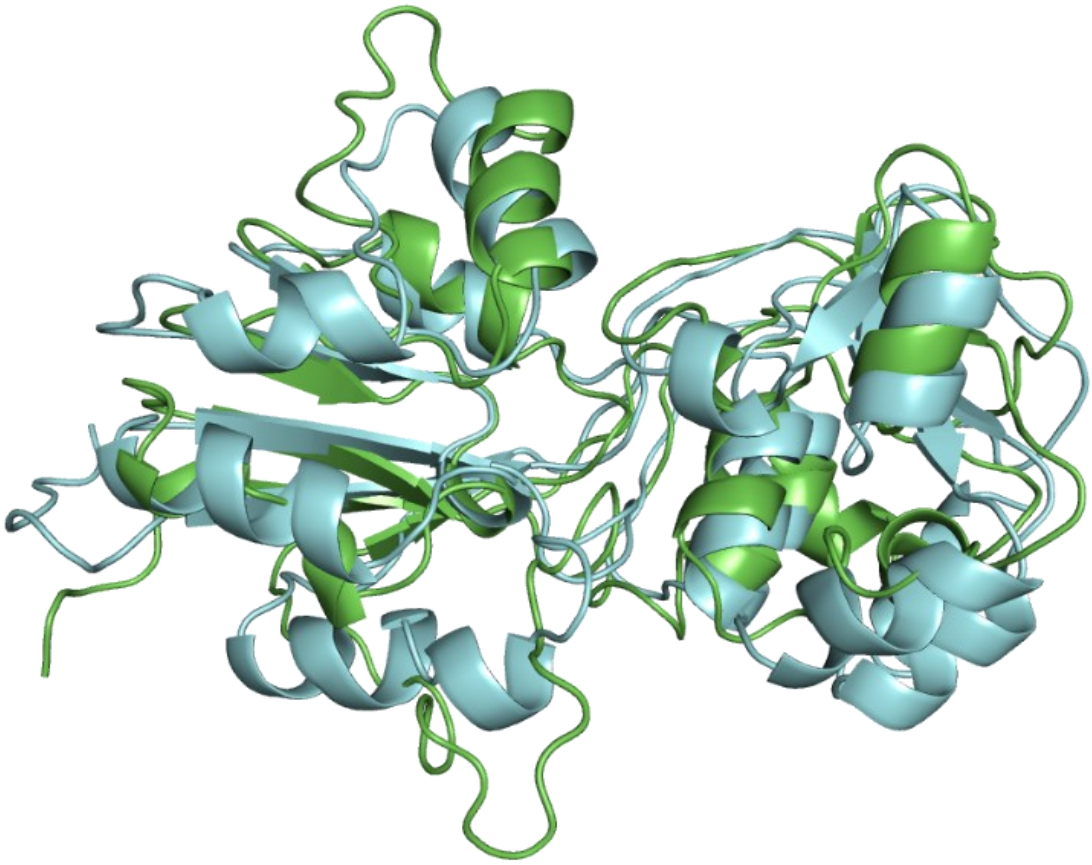


Figure 6. *Structural alignment of the HEV RNA Helicase model (green cartoon) with the model built via the trRosetta web-server (pale-blue cartoon).*

The RNA was added to the modelled protein to complete the HEV RNA Helicase-RNA model as reported in the 2.1.3 section. The protein from which the RNA was obtained is stored in the PDB with 6JIM code [50]. It derives from the *Chikungunya virus*, a species belonging to the *Alsuviricetes* class, as well as HEV and Tmv. The Tar-Seq and the RNA-donor structure 6JIM were structurally

comparable (RMSD ~2.5 Å), also at the level of the RNA-binding pocket. Then, the 6JIM RNA chain was extracted and docked to our model via ClusPro 2.0 [35] obtaining a hydrophobic-favoured complex with a 6JIM-like arrangement.

Finally, the poly-uracil RNA fragment was edited to a poly-purine one (5'-GAGAGAG-3') as previous studies reported the suitability of SLV-related compounds to interact with a poly-purine RNA sequence [51].

3.2. Docking ligands and analyzing complex stability

For the sake of identifying compounds with potential anti-viral properties, the model was targeted with natural bioactive compounds belonging to the cyclopenta[*b*]benzofuran/flavagline class to identify promising candidates to test in further dedicated investigations. The focus on RCG (Figure 1A) and some of its characterized analogues (RCG1 to 6 Figure 1A) was due to the known activity against RNA Helicases [27]. SLV (Figure 1B) was also included in the study as it was an already known inhibitor of HEV replication, but whose underpinning mechanisms and molecular targets were still to be discovered [24-26]. Moreover, S-Ag (CID 24178739, Figure 1B) was analysed due to the chemical similarity with both RCG and SLV, along with two virtual decoys to assess the procedural performances (ZINC ID ZINC8584442 and ZINC8387186; further details are reported in the 2.2 section).

The position to dock the ligands within the ligand binding site was defined based on the PDB structure 5ZC9 showing RCG engaged in a well-defined base-base stacking interaction (Figure 2). The set of ligands was docked as reported in the 2.1.2 section, the results are reported in Table 1 (the higher the score, the better the expected ligand interaction).

Table 1. *Docking PLP score obtained running GOLD. The highest the score the more likely the docking pose is optimal.*

Ligand	Docking score (PLP)
RCG	256.075
RCG1	224.929
RCG2	246.042
RCG3	224.077
RCG4	231.621
RCG5	228.042
RCG6	169.175
SLV	184.600
S-Ag	236.039
ZINC8584442	178.539
ZINC8387186	150.527

As reported in Table 1, RCG showed the highest score followed by all its analogues. All of them clearly showed a stacking interaction except for RCG6 and the two decoys. On the other hand, SLV obtained a higher score than virtual decoys but sensibly lower than both RCG and S-Ag. Of note, the inactive decoys served to validate the model. In particular, the lack of activity was imputable to their incapability to keep a stable stacking interaction resulting in an enzyme inhibition. Therefore, the fact that they both showed no stacking interactions corroborated the model reliability.

Once obtained the docking pose for each ligand, 25 ns long MD simulations were run to investigate the evolution of the built complexes. Based both on the 5ZC9 crystal structure and other evidence from the literature claiming the RCG stabilization of the protein-RNA complex [52], the following geometrical rules were associated to ligands theoretically able to interact with and inhibit HEV RNA

Helicase. Specifically, they should: I) keep the stacking interaction with the RNA stably; and II) preserve the stability of the RNA-protein complex avoiding RNA detachment.

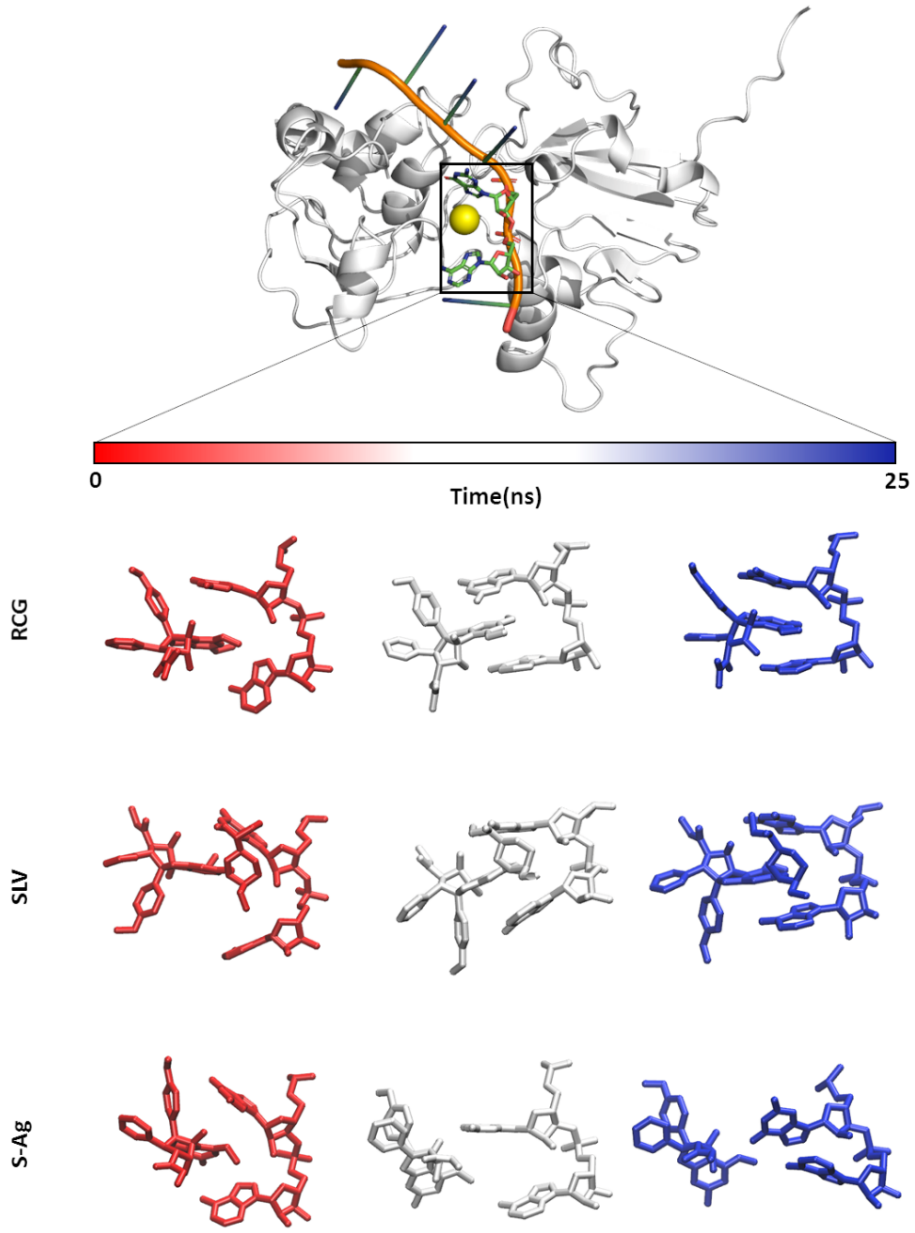


Figure 7. *Protein-RNA-ligand complex and evolution over time of RCG, SLV and S-Ag with respect to the RNA bases. The protein is represented as a white cartoon, the RNA as an orange/green/blue cartoon while the two bases involved in the interaction with the ligand as green sticks. The yellow sphere represents the ligand docking-site. Under the time bar, starting from the top and moving to the bottom we can alternatively see the time-step sticks representation of both the RNA bases and RCG, SLV and S-Ag trajectories.*

RCG and SLV were both engaged in a base-base stacking interaction, also stabilizing the protein-RNA complex with no appreciable RNA detachment (Figure 7). RCG1 did not show a proper stacking interaction and promoted the RNA detachment, pointing to its limited theoretical inhibitory activity against the viral RNA Helicase. Interestingly, this result is in line with the Pan et al. study reporting more than 500-fold activity decrease with respect to RCG against the human eIF4A1 ATP-dependent RNA helicase [27]. RCG2 caused a self-collapse of the RNA, promoting its detachment from the protein, and no stacking interactions were observed. This also suggested a low inhibitory potential, in line with the inactivity reported by Pan et al. [27] against human RNA Helicase eIF4A1. RCG3, a less cytotoxic RCG analogue [27], showed a single base interaction, rather than a base-base stacking interaction, and promoted the RNA detachment. Therefore, it was not deemed an efficient inhibitor. RCG4 and RCG6 both favoured the RNA detachment, with the latter reported as less active than the former, although RCG4 was found in an RCG-like stacking interaction with the RNA. They were not considered able to appreciably inhibit the HEV RNA Helicase activity.

S-Ag did not interact at all with the RNA. Shortly after the beginning of the MD simulation it slipped out the RNA chain and kept a likely unspecific surface interaction with a near protein portion (Figure 7). This was probably due to the presence of the hydrophilic hydroxyl group in position R6 (Figure 1), which prevented a proper a stacking interaction, and to the absence of the SLV dioxane portion. Indeed, this part participates in the SLV stacking interaction confirming the Cencic and co-workers' hypothesis claiming its crucial role for the SLV activity [53]. Concerning the two decoys, they both promoted RNA detaching, particularly ZINC8584442. This, along the lack of base-base stacking, was expected for inactive compounds and confirmed the procedural performances.

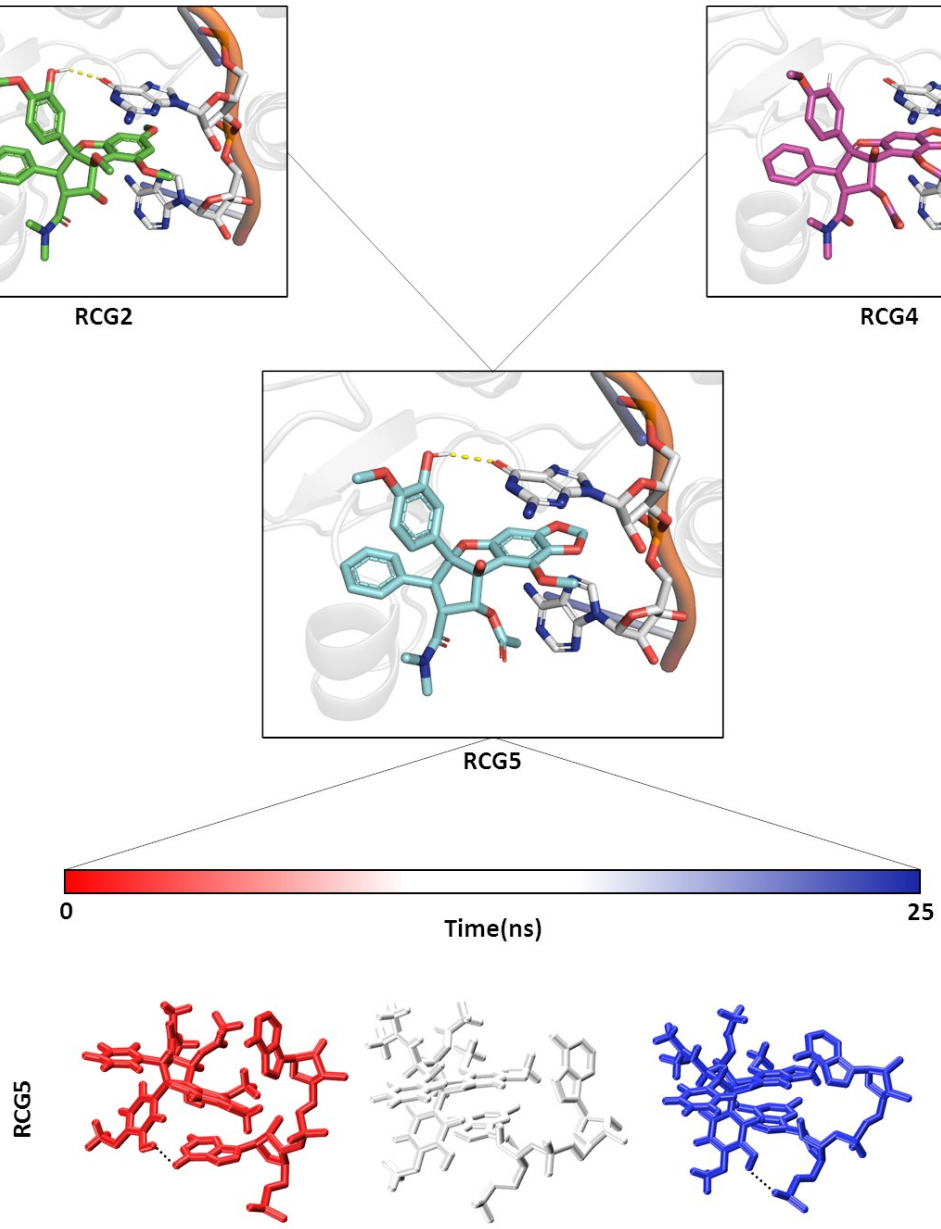


Figure 8. *RCG2 is represented as green sticks, RCG4 as magenta sticks, RCG5 as pale-blue sticks. RNA bases involved in the interaction are represented as white sticks while the RNA backbone as orange cartoon. The yellow dashed lines represent the Hydrogen bond occurring between the R4 of both RCG5 and RCG2, and the RNA base. The time-step representation of the RCG5/RNA-bases trajectories is reported under the time bar. The black dashed lines represent both the starting Hydrogen bond and the final one.*

Lastly, a novel RCG analogue, i.e. RCG5, was rationally designed and evaluated based on the calculated pose of other RCG analogues included in this study. Particularly, starting from docking

poses it was noticed there was a hydrogen bond between the hydroxyl group in position R4 of RCG2 and an RNA base (Figure 8). Moreover, looking at the progression of RCG2 MD simulation, the starting H-bond was not kept over time but a new one between the hydroxyl group in R4 and the RNA backbone was formed. This interaction was thought as favourably stabilizing the complex and theoretically able to enhance the activity of the compound. Therefore, in RCG5, the R4 of RCG4 (i.e., -H) was replaced with a hydroxyl group. RCG5 showed a stable stacking interaction while keeping the RNA-protein complex stable, steadily forming an H-bond with a phosphate group of the RNA backbone (Figure 8). This newly designed compound needs further investigation and experimental evaluation to validate its HEV RNA Helicase inhibitory activity.

3. Conclusion

Overall, this study: i) presented mechanistic insights on HEV RNA Helicases and its SLV/RCG-dependent inhibition; ii) expanded the current understanding of the structure-activity relationship for SLV and RCG-related compounds; iii) provided a blueprint for further analysis targeting HEV RNA Helicases. Moreover, this study described for the first time the inhibition of HEV RNA Helicase as a plausible mechanism of action of SLV, which was already described reducing HEV capability to penetrate cells. This ancillary mechanism, which deserves further investigation with high priority, was found plausible also for other RCG/SLV-like compounds and may complement the already reported activity of this HEV impairing agents on host RNA Helicases. Furthermore, there are evidence reporting SLV as well-tolerated in animals [53] making it an interesting candidate as a feed additive. The described methodology, starting from the template selection moving to the actual modelling and testing, succeeded to build a reliable model able to qualitatively discriminate several ligands. In addition to this, the pipeline is highly versatile, flexible and it can be replicated on other HEV ORF1 domains, such as the RNA Dependent RNA Polymerase which has already been proved as a suitable target to inhibit virus replication [54].

References

1. World Health Organization. *Hepatitis E*. 2021; Available from: https://www.who.int/news-room/fact-sheets/detail/hepatitis-e.

2. Webb, G.W. and H.R. Dalton, *Hepatitis E: an underestimated emerging threat.* Therapeutic Advances in Infectious Disease, 2019. **6**.

3. Meng, X.J., *Zoonotic and Foodborne Transmission of Hepatitis E Virus.* Seminars in Liver Disease, 2013. **33**(1): p. 41-49.

4. Yugo, D.M. and X.J. Meng, *Hepatitis E Virus: Foodborne, Waterborne and Zoonotic Transmission.* International Journal of Environmental Research and Public Health, 2013. **10**(10): p. 4507-4533.

5. Carratala, A. and S. Joost, *Population density and water balance influence the global occurrence of hepatitis E epidemics.* Scientific Reports, 2019. **9**.

6. Dalton, H.R., N. Kamar, and J. Izopet, *Hepatitis E in developed countries: current status and future perspectives.* Future Microbiology, 2014. **9**(12): p. 1361-1372.

7. Okamoto, H., *Genetic variability and evolution of hepatitis E virus.* Virus Research, 2007. **127**(2): p. 216-228.

8. Dziedzinska, R., et al., *Evidence of Hepatitis E Virus in Goat and Sheep Milk.* Viruses-Basel, 2020. **12**(12).

9. Huang, F., et al., *Excretion of Infectious Hepatitis E Virus Into Milk in Cows Imposes High Risks of Zoonosis.* Hepatology, 2016. **64**(2): p. 350-359.

10. Obaidat, M.M. and A.A. Roess, *Individual animal and herd level seroprevalence and risk factors of Hepatitis E in ruminants in Jordan.* Infection Genetics and Evolution, 2020. **81**.

11. Takahashi, K., et al., *Complete or near-complete nucleotide sequences of hepatitis E virus genome recovered from a wild boar, a deer, and four patients who ate the deer.* Virology, 2004. **330**(2): p. 501-505.

12. Boxman, I.L.A., et al., *Porcine blood used as ingredient in meat productions may serve as a vehicle for hepatitis E virus transmission.* International Journal of Food Microbiology, 2017. **257**: p. 225-231.

13. Lopez-Lopez, P., et al., *Risk factors associated with hepatitis E virus in pigs from different production systems.* Veterinary Microbiology, 2018. **224**: p. 88-92.

14. Pallerla, S.R., et al., *Hepatitis E virus genome detection in commercial pork livers and pork meat products in Germany.* Journal of Viral Hepatitis, 2021. **28**(1): p. 196-204.

15. Kokkinos, P., et al., *Virological Quality of Irrigation Water in Leafy Green Vegetables and Berry Fruits Production Chains.* Food and Environmental Virology, 2017. **9**(1): p. 72-78.

16. Maunula, L., et al., *Tracing enteric viruses in the European berry fruit supply chain.* International Journal of Food Microbiology, 2013. **167**(2): p. 177-185.

17. Wang, B. and X.J. Meng, *Structural and molecular biology of hepatitis E virus.* Computational and Structural Biotechnology Journal, 2021. **19**: p. 1907-1916.

18. Ropp, S.L., et al., *Expression of the hepatitis E virus ORF1.* Archives of Virology, 2000. **145**(7): p. 1321-1337.

19. Purdy, M.A., et al., *ICTV Virus Taxonomy Profile: &ITHepeviridae&IT.* Journal of General Virology, 2017. **98**(11): p. 2645-2646.

20. Seyedi, S.S., et al., *Computational Approach Towards Exploring Potential Anti-Chikungunya Activity of Selected Flavonoids.* Scientific Reports, 2016. **6**(1): p. 24027.

21. Patel, C.N., et al., *Identification of antiviral phytochemicals as a potential SARS-CoV-2 main protease (Mpro) inhibitor using docking and molecular dynamics simulations.* Scientific Reports, 2021. **11**(1): p. 20295.

22. Dwivedi, V.D., et al., *Mechanistic insights into the Japanese encephalitis virus RNA dependent RNA polymerase protein inhibition by bioflavonoids from Azadirachta indica.* Scientific Reports, 2021. **11**(1): p. 18125.

23. Alencar, W.L.M., et al., *Interactions of Co, Cu, and non-metal phthalocyanines with external structures of SARS-CoV-2 using docking and molecular dynamics.* Scientific Reports, 2022. **12**(1): p. 3316.

24. Glitscher, M., et al., *Inhibition of Hepatitis E Virus Spread by the Natural Compound Silvestrol.* Viruses-Basel, 2018. **10**(6).

25. Parvez, M.K. and N. Subbarao, *Molecular Analysis and Modeling of Hepatitis E Virus Helicase and Identification of Novel Inhibitors by Virtual Screening.* Biomed Research International, 2018. **2018**.

26. Todt, D., et al., *The natural compound silvestrol inhibits hepatitis E virus (HEV) replication in vitro and in vivo.* Antiviral Research, 2018. **157**: p. 151-158.

27. Pan, L., et al., *Rocaglamide, silvestrol and structurally related bioactive compounds from Aglaia species.* Natural Product Reports, 2014. **31**(7): p. 924-939.

28. Bateman, A., et al., *UniProt: a worldwide hub of protein knowledge.* Nucleic Acids Research, 2019. **47**(D1): p. D506-D515.

29. Altschul, S.F., et al., *BASIC LOCAL ALIGNMENT SEARCH TOOL.* Journal of Molecular Biology, 1990. **215**(3): p. 403-410.

30. Jeske, L., et al., *BRENDA in 2019: a European ELIXIR core data resource.* Nucleic Acids Research, 2019. **47**(D1): p. D542-D549.

31. Webb, B. and A. Sali, *Protein Structure Modeling with MODELLER*, in *Protein Structure Prediction, 3rd Edition*, D. Kihara, Editor. 2014. p. 1-15.

32. Laskowski, R.A., et al., *PROCHECK - A PROGRAM TO CHECK THE STEREOCHEMICAL QUALITY OF PROTEIN STRUCTURES.* Journal of Applied Crystallography, 1993. **26**: p. 283-291.

33. Pettersen, E.F., et al., *UCSF chimera - A visualization system for exploratory research and analysis.* Journal of Computational Chemistry, 2004. **25**(13): p. 1605-1612.

34. Wiederstein, M. and M.J. Sippl, *ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins.* Nucleic Acids Research, 2007. **35**: p. W407-W410.

35. Desta, I.T., et al., *Performance and Its Limits in Rigid Body Protein-Protein Docking.* Structure, 2020. **28**(9): p. 1071-+.

36. Muller, C., et al., *Comparison of broad-spectrum antiviral activities of the synthetic rocaglate CR-31-B (-) and the eIF4A-inhibitor Silvestrol.* Antiviral Research, 2020. **175**.

37. Jones, G., et al., *Development and validation of a genetic algorithm for flexible docking.* Journal of Molecular Biology, 1997. **267**(3): p. 727-748.

38. Kim, S., et al., *PubChem in 2021: new data content and improved web interfaces.* Nucleic Acids Research, 2021. **49**(D1): p. D1388-D1395.

39. Mysinger, M.M., et al., *Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking.* Journal of Medicinal Chemistry, 2012. **55**(14): p. 6582-6594.

40. Legnik, S., et al., *LiSiCA: A Software for Ligand-Based Virtual Screening and Its Application for the Discovery of Butyrylcholinesterase Inhibitors.* Journal of Chemical Information and Modeling, 2015. **55**(8): p. 1521-1528.

41. Pall, S., et al. *Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS*. in *2nd International Conference on Exascale Applications and Software (EASC)*. 2014. Stockholm, SWEDEN.

42. Best, R.B., et al., *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, 2012. **8**(9): p. 3257-3273.

43. Zoete, V., et al., *SwissParam: A Fast Force Field Generation Tool for Small Organic Molecules.* Journal of Computational Chemistry, 2011. **32**(11): p. 2359-2368.

44. Withana-Gamage, T.S., et al., *In Silico Homology Modeling To Predict Functional Properties of Cruciferin.* Journal of Agricultural and Food Chemistry, 2011. **59**(24): p. 12925-12938.

45. Karpe, Y.A. and K.S. Lole, *NTPase and 5 ' to 3 ' RNA Duplex-Unwinding Activities of the Hepatitis E Virus Helicase Domain.* Journal of Virology, 2010. **84**(7): p. 3595-3602.

46. Li, P.F., et al., *The global epidemiology of hepatitis E virus infection: A systematic review and meta-analysis.* Liver International, 2020. **40**(7): p. 1516-1528.

47. Waterhouse, A., et al., *SWISS-MODEL: homology modelling of protein structures and complexes.* Nucleic Acids Research, 2018. **46**(W1): p. W296-W303.

48. Ishibashi, K., et al., *Structural basis for the recognition-evasion arms race between Tomato mosaic virus and the resistance gene Tm-1.* Proceedings of the National Academy of Sciences of the United States of America, 2014. **111**(33): p. E3486-E3495.

49. Du, Z.Y., et al., *The trRosetta server for fast and accurate protein structure prediction.* Nature Protocols, 2021. **16**(12): p. 5634-5651.

50. Law, Y.S., et al., *Structural insights into RNA recognition by the Chikungunya virus nsP2 helicase.* Proceedings of the National Academy of Sciences of the United States of America, 2019. **116**(19): p. 9558-9567.

51. Lwasaki, S., et al., *The Translation Inhibitor Rocaglamide Targets a Bimolecular Cavity between eIF4A and Polypurine RNA.* Molecular Cell, 2019. **73**(4): p. 738-+.

52. Chu, J., et al., *Rocaglates Induce Gain-of-Function Alterations to eIF4A and eIF4F.* Cell Reports, 2020. **30**(8): p. 2481-+.

53. Cencic, R., et al., *Antitumor Activity and Mechanism of Action of the Cyclopenta b benzofuran, Silvestrol.* Plos One, 2009. **4**(4).

54. Parvez, M.S.A., et al., *Prediction of potential inhibitors for RNA-dependent RNA polymerase of SARS-CoV-2 using comprehensive drug repurposing and molecular docking approach.* International Journal of Biological Macromolecules, 2020. **163**: p. 1787-1797.

Chapter II

Dipeptides as possible inhibitors of Chorismate synthase: a case study on *Campylobacter jejuni*

This chapter presents still unpublished results. It is part of a collaboration with Professor Peter Macheraux (University of Graz, Austria) whose laboratory carried out the *in vitro* analysis.

Abstract

Campylobacter jejuni is one of the most common foodborne pathogens and the causative agent of campilobacteriosis, a disease leading to diarrhoea, fever and gastroenteritis. It can form biofilms and survive at refrigeration temperature and thus food industry needs novel antibacterial approaches. Here we applied an *in silico* pipeline to predict possible bioactive dipeptides inhibiting the *C. jejuni* chorismate synthase. This enzyme is part of the shikimate pathway, a crucial biosynthetic route for bacteria and not present in animals, making it an interesting enzyme to target for selective inhibition. A library of 400 dipeptides was screened through molecular docking simulations. Among them, GluAsp stood out as it mimicked CS natural substrate and it maintained a similar interaction pattern during the whole molecular dynamics simulation. However, the *in vitro* assay resulted with ED as a CS activity enhancer raising concerns considering the diffusion of *C. jejuni* in protein rich matrices. This study provides a blueprint for further dedicated investigations, broadening the CS active site chemistry knowledge and opening new directions in designing peptide-based inhibitors, thereby advancing food safety interventions against *C. jejuni*.

1. Introduction

Campylobacter jejuni is a concerning foodborne pathogen and one of the main causes of foodborne illnesses worldwide, as per World Health Organization [1]. It is usually transmitted through contaminated and undercooked poultry or from raw and unpasteurised milk, however there can be other ways of infection such as meat, seafood and vegetables [2-4]. It is a gram-negative bacterium causing campylobacteriosis, an infection including gastroenteritis, fever, and diarrhoea. There are more than 1 million cases registered every year bringing to a significant burden of disease [5]. Apart from its typical symptoms, this illness could lead to significant complications such as the Guillain-Barrè syndrome, causing numbness, paralysis and eventually death [6]. *C. jejuni* preferentially grows at microaerophilic conditions and it is usually inhibited by refrigeration temperatures [7]. However, it can form biofilm which enhances its resistance and persistence; additionally, recent studies identified some cold-resistant *C. jejuni* strains surviving under aerobic conditions, posing more challenges from a food safety standpoint [8, 9]. For these reasons, developing novel strategies to deal with this foodborne pathogen has become essential nowadays.

C. jejuni, as every other prokaryote, plants, and fungi, needs to biosynthesize aromatic amino acids (tryptophan, tyrosine and phenylalanine). In fact, differently from animals and humans, these kind of organisms cannot uptake aromatic amino acids from diet but they evolved a specific biosynthetic route to produce them, the so-called shikimate pathway [10]. This pathway is characterized by seven enzymes, and thus seven steps, that starting from phosphoenolpyruvate leads to the production of chorismate, i.e. the last precursor of folates and aromatic amino acids [11]. This pathway is a suitable and interesting target for developing antibacterial strategies since it is essential in bacterial growth and consequent virulence but without a counterpart in animals. Indeed, researchers are already

trying to target such pathway investigating also the possibility of developing broad-spectrum antibacterial compounds due to the high conservation within the bacterial phyla [12-14].

At the same time, the identification of food-grade antibacterial compounds of natural origin could bring added value as moving steps toward the labels cleaning and meeting the high acceptance of consumers for additives of natural origin, to mention but a few. For these reasons, we decided to identify bioactive peptides, molecules inherently natural and theoretically food grade when derived from food proteins, possibly interacting and interfering with *C. jejuni* chorismate synthase (CS), the last of the seven enzymes found in the shikimate pathway and whose crystallographic structure is available [15]. This could help in ultimately mitigating the bacterium pathogenicity, lowering its survivability. Regarding the focus on peptides they were chosen since CS is a crucial enzyme catalysing the 1,4-trans elimination of the phosphate group from 5-enolpyruvylshikimate-3-phosphate (EPSP), i.e. its natural substrate, through a reaction exploiting flavin mononucleotide (FMN) as a cofactor [16]. Strikingly, certain short acidic peptides might mirror some of the chemical motifs of EPSP, corroborating their possible role as competitive inhibitors. In particular, we applied an *in silico* pipeline based on molecular docking and dynamics simulations since it already proved success in discovering possible bioactive peptides [17]. We built a library containing all the 400 possible dipeptides given by the combination all the 20 proteinogenic amino acids. This was used as input for the subsequent molecular docking simulation while the most promising candidates were also investigated via molecular dynamics simulations. The considered best hit (GluAsp) and worst hit (ValAla) were then analysed through *in vitro* CS activity's assay based on previous studies [18]. The choice of seeking for bioactive peptides, and dipeptides specifically, was in line with previous research proving the intrinsic antimicrobial properties of some of these compounds and their possible multi-purpose use as food/feed additives, also trying to avoid the issue of antibiotic resistance [19-21]. Finally, investigating the behaviour of dipeptides in their interaction with *C. jejuni*

CS is fundamental since this bacterium is commonly found within protein-rich food matrices, such as poultry [7]. Insights in this sense could contribute to better improve the food safety awareness and to design more effective strategies counteracting bacterial growth, spread and pathogenesis.

2. Materials and Methods

2.1. Data source

The 3D structure of *C. jejuni* CS is stored on the Protein Data Bank (PDB; https://www.rcsb.org/) [22] under the PDB ID 1SQ1 [15]. However, since the structure is not fully resolved, it was modelled through ColabFold (v 1.5.5) [23], a publicly available notebook implemented with AlphaFold2 and AlphaFold2-multimer [24, 25], in agreement with previous studies [26]. The primary sequence of the *C. jejuni* CS was collected from UniProt (https://www.uniprot.org; UniProt ID Q9PM41) [27] and given as input to ColabFold along with 1SQ1, the latter used as a custom template. Once obtained a reliable model, the cofactor FMN was added and optimized within CS's binding site through AlphaFill (https://alphafill.eu/) [28] setting default parameters.

The 3D structure of Gallion (CID 135406014), and EPSP (439463) were downloaded in the SDF format from PubChem (https://pubchem.ncbi.nlm.nih.gov/) [29]. They were converted to the MOL2 format through Open Babel (v. 3.1.1) [30] before proceeding with the next steps.

2.2. Construction of the dipeptides library

The library of dipeptides analysed in this study was generated using an *ad hoc* python script executed through PyMol (v.2.3), in line with a previous study [26]. All possible combinations of the 20 proteinogenic amino acids were created, obtaining 400 dipeptides in the MOL2 format. These were protonated at pH 7, each with deprotonated C-terminal and protonated N-terminal, as per physiological conditions.

2.3. Molecular docking simulations

Molecular docking simulations were performed through the software GOLD (Genetic Optimization for Ligand Docking; v. 2022) [31]. A semi-flexible docking approach was applied allowing only the

protein's polar hydrogens free to rotate while keeping the ligands fully flexibles. The binding site was set as 10 Å sphere centred within the CS's active site, keeping the FMN cofactor. 10 poses per ligand were generated and scored using the internal PLPScore score function (the higher the score, the more likely the predicted architecture of binding, as per manufacturer declarations; https://www.ccdc.cam.ac.uk/), in line with previous studies [26].

2.4. Pharmacophoric analysis of CS pocket

A pharmacophoric analysis of the *C. jejuni* CS binding site was conducted using the IsoMif algorithm [32]. This was done to obtain precise insights on the CS's binding pocket chemical fingerprint specifically focusing on regions accepting negatively charged groups. Default parameters were used while setting the maximum distance value between the grid and residues atoms at 3 Å with a grid resolution of 1 Å.

2.5. Molecular dynamics simulations

Molecular dynamics (MD) simulations were executed using GROMACS (v. 2021.4) [33] to verify the evolution over time of the complexes obtained from the molecular docking simulations. Of note, the tested complexes were CS-EPSP, CS-Gallion and CS with 2 different dipeptides, namely GluAsp (ED) and ValAla (VA).

Both the protein and peptides were parametrized with the CHARMM27 all-atom forcefield [34] while the non-protein ligands (i.e. gallion and EPSP) on the SwissParam webserver (https://old.swissparam.ch/) [35]. Each system was placed in an octahedral box, solvated with SPCE waters, and neutralized adding Na^+ and Cl^- ions. Every system was then energetically minimized to avoid steric clashes or to correct improper geometries. This was achieved using a steepest descent algorithm with a maximum of 5 000 optimization steps. Each minimized system underwent isothermal (300 K, 2 ps coupling time) and isobaric (1 bar, 2 ps coupling time) simulations for 100 ps

to allow the system to reach the equilibrium. Once completed these steps, 30 ns long MD simulations were run.

2.6. Chorismate synthase activity's assay

To evaluate the inhibitory effects of gallion, ED and VA a coupled assay incorporating EPSP, CS (from *Paracoccidioides brasiliensis*) and anthralinate synthase was conducted as reported in a previous study [18]. Likewise, the expression and purification of CS was performed as previously described [18]. All the chemicals and peptides were purchased from Merck (peptides at 95% purity), handled and stored prior usage according to supplier's specification.

In brief, the assay setup enabled the forward coupling of the CS reaction, resulting in anthranilate production. This is a stable compound with a fluorescence emission peak at 390 nm that can be detected via spectrofluorometry. Reaction mixtures were prepared in 50 mM MOPS (pH 7.5) containing 3 mM $MgSO_4$, 7.5 mM L-Gln, 22.5 mM $(NH_4)_2SO_4$, 1 mM dithiothreitol (DTT), 10 µM FMN, 80 µM EPSP, 4 µM CS, ~20 µM anthranilate synthase, and various concentrations of gallion, ED and VA (0.1–500 µM), and these mixtures were pipetted into a 96-well plate in triplicate. 500 µM NADPH were added to the plate which was subsequently incubated at 37 °C for one minute. The fluorescence emission changes at 390 nm (λ_{ex} = 340 nm) were recorded every 12 seconds for 5 minutes using a Spectramax Gemini XS Microplate Spectrofluorometer (Molecular Devices, CA, USA). All measurements were performed in triplicate (n = 3). Data are expressed as the mean ± standard error mean (SEM). Statistical analysis was performed using IBM SPSS Statistics for Windows, version 29 (IBM Corp., Armonk, NY). The data was analysed by one-way ANOVA (α = 0.05), followed by *post hoc* Fisher's LSD test (α = 0.05).

3. Results and Discussion

3.1 3D model validation

C. jejuni CS was modelled through ColabFold (v 1.5.5) [23] (details in section 2.1) obtaining a reliable structure with a pLDDT score higher than 90, indicating well predicted backbone and side chains. Since CS is a homotetramer with four identical binding sites, all subsequent analysis was performed on a single monomer also to reduce the computational load. Subsequently, the architecture of binding of the FMN cofactor, added through AlphaFill (https://alphafill.eu/) [28] within the CS binding site, was comparable to the *Streptococcus pneumoniae* CS crystallographic structure (PDB ID 1QXO [36]) which has this co-crystallized cofactor along with EPSP (Figure 1).

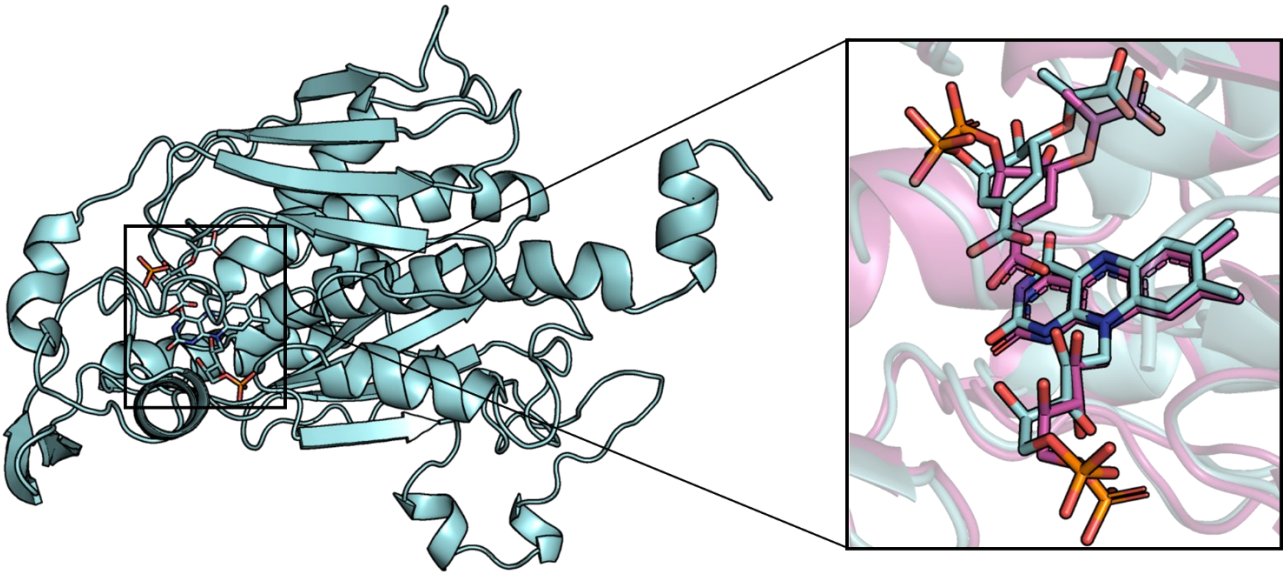


Figure 1. *On the left, 3D structure of the* C. jejuni *CS monomer represented as a cyan cartoon while FMN and EPSP as cyan sticks. On the right, close-up on the binding site with FMN and EPSP within* C. jejuni *CS (cyan) compared to the co-crystallized ligands within* S. pneuomoniae *CS (magenta; PDB ID 1QXO [36]).*

Once assessed the reliability of the predicted CS-FMN complex (Figure 1) we proceeded with molecular docking simulations. The accuracy of the applied docking protocol, prior to using it for gallion and the whole dipeptide library, was assessed on EPSP. Specifically, the architecture of

binding of EPSP (PLPScore 70.03) obtained after the docking procedure was almost superimposed to the one's co-crystallized within *S. pneumoniae* CS crystallographic structure (PDB ID 1QXO [36]), confirming both the model and procedural reliability.

3.2 Dipeptides library docking-based screening

The dipeptides library (400 peptides in total) was constructed using an *ad hoc* python script implemented with PyMol (v. 2.3), in line with a previous study [26]. Specifically, the script takes as input the structure of a C-terminal proteinogenic amino acid and iteratively bonds a N-terminal proteinogenic amino acid to it, maintaining a proper protonation state under physiological conditions (pH 7). Eventually, it stores the 3D coordinates of the generated dipeptide in the MOL2 format – proper format for the docking procedure performed with GOLD. This 400 dipeptides library was screened exploiting the docking protocol set up using EPSP (see section 3.1). It was chosen to investigate peptides since they already proved antibacterial properties and they have a potential as multi-use food/feed additives, helping to possibly avoid issues related to the antibiotic resistance [19-21]. Indeed, they are naturally present in food matrices and food-grade when derived from them, aligning with the cleaner labelling trend which is appealing to consumer preferences. Last, only dipeptides were chosen avoiding longer oligopeptides due to the limited size of the CS binding site.

The obtained scores were all positive, roughly ranging from 50 to 95 PLPScore, indicating that all the dipeptides could theoretically interact with the CS binding site (the higher the score, the higher the physico-chemical ligand-pocket match). However, a visual expert inspection of the whole set of obtained poses was performed through PyMol (v. 2.3) to seek for the best possible hit to analyse through MD simulations and to eventually test *in vitro*. Focusing on the dipeptides 3D structures, it was noticed that certain short acidic peptides, as expected, might convincingly emulate some of the

chemical motifs of CS substrate, suggesting a possible role as competitive inhibitors. Among the 400 tested, ED (PLPScore 84.44) was the most promising since it was the one best mimicking the pattern of interaction of EPSP, the natural CS substrate (Figure 2).

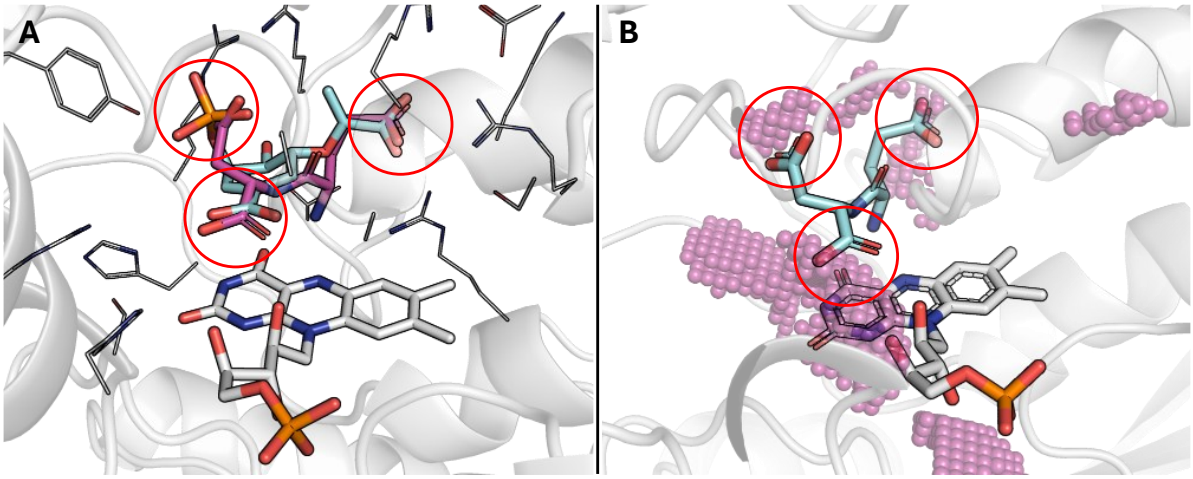


Figure 2. *Focus on CS binding site with the protein represented as a white transparent cartoon, residues as white lines, FMN as white sticks. Negatively charged moieties are embedded by red circles. A) Superimposition of EPSP (magenta) and ED (cyan). B) Regions accepting negatively charged groups are represented as magenta sphere while ED is represented as cyan sticks.*

As shown in Figure 2, the C-terminal carboxylate moiety was almost superimposed to the EPSP's carboxylate moiety next to FMN. Similarly, the carboxylate moiety of the Glu side chain occupied the regions corresponding to the other EPSP's carboxylate while the carboxylate moiety of the Asp side chain occupied the region of the EPSP's phosphate group. Additionally, the analysis of the CS binding pocket with IsoMif [32], performed to find regions likely to accommodate negatively charged groups, confirmed the goodness of such interaction (Figure 2). On the other hand, as a negative hit to test both with MD simulation and *in vitro* assay, we decided for VA (PLPScore 58.29). This dipeptide is significantly different from ED due to its hydrophobic nature. Specifically, it lacks negatively charged

side chains, and it is engaged in a different pattern of interaction than ED and EPSP which are likely to results in a way less favoured interaction with the enzyme (Figure 3).

Finally, gallion (PLPScore 43.75) was also included in the docking simulations. This is a promising CS inhibitor, similar to those already characterized and documented in the literature, on which colleagues from the University of Graz are working on [18]. As reported in figure 3, the pattern of interaction was comparable, and even more dense, to those of ED and EPSP.

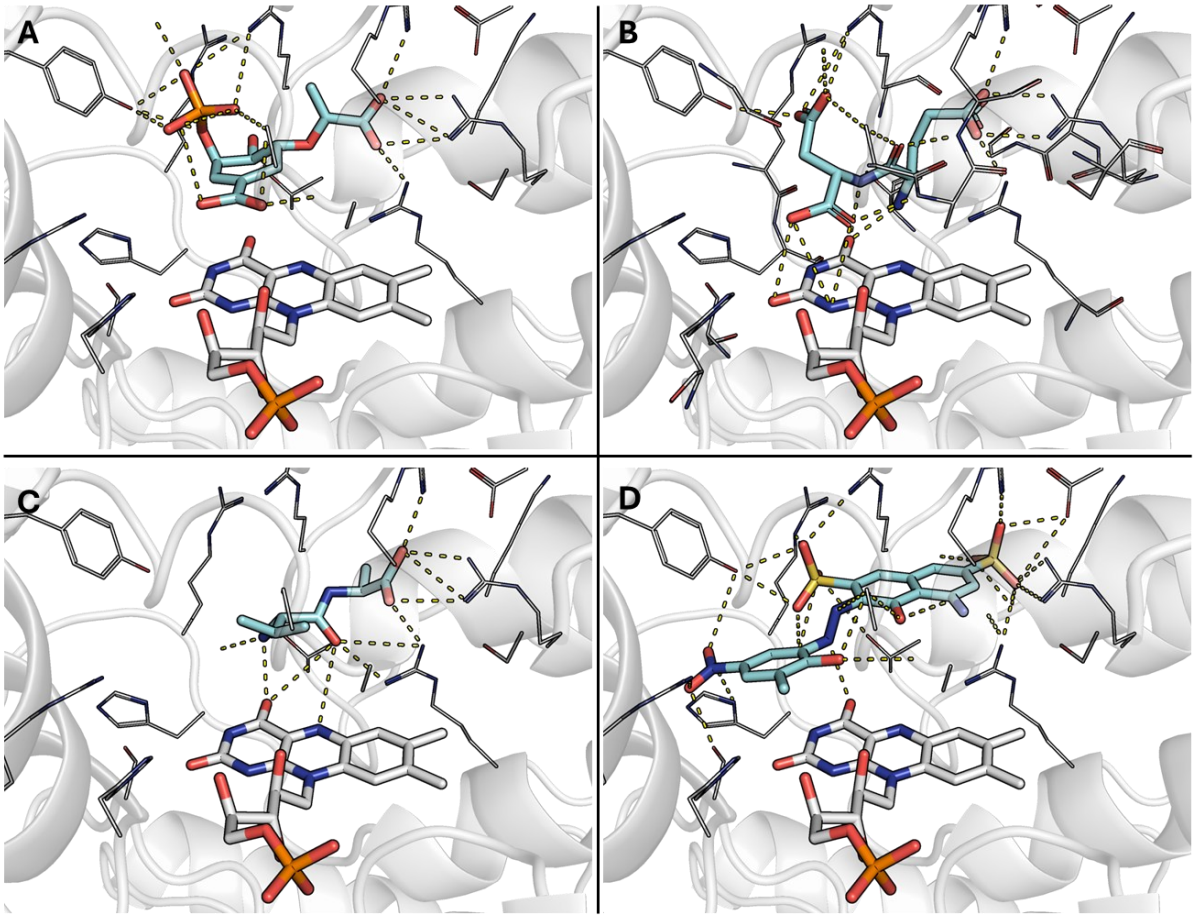


Figure 3. *Docking poses of the four chosen compounds for MD simulation and* in vitro *testing. Ligands, namely EPSP (A), ED (B), VA (C) and Gallion (D), are represented as cyan sticks. Protein is represented as white transparent cartoon while residues around ligands as white lines. Polar interactions are represented as dashed lines.*

3.2. MD simulations and *in vitro* assay

MD simulations. The complexes obtained from the docking simulations, namely CS-EPSP, CS-Gallion, CS-ED and CS-VA, were used as input for 30 ns long MD simulations. As shown in figure 4, with respect to the overall protein stability – monitored via root mean squared deviation (RMSD) trends analysis) –, all the complexes were comparable from the second half of the MD simulation with VA considered the worst based on its higher RMSD average value. However, this was not considered a critical factor to discriminate between good and bad CS interactors.

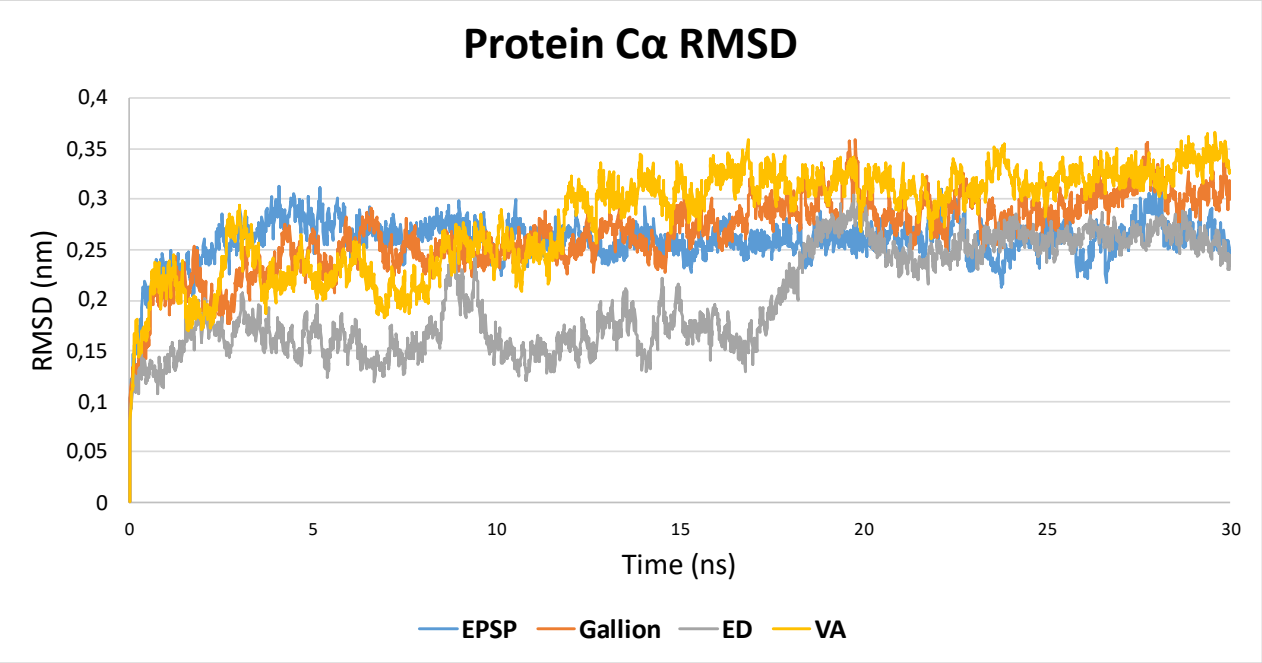


Figure 4. *CS Cα RMSD computed over the 30 ns long MD simulations. EPSP is reported in blue (average 0.259 ± 0.021 nm), Gallion in orange (average 0.267 ± 0,0369 nm), ED in grey (average 0.202 ± 0.051 nm) and VA in yellow (average 0.281 ± 0.050 nm).*

Subsequently, an H-bond analysis monitoring the evolution of h-bond numbers over time was conducted on all the four tested complexes. Specifically, it was observed that CS-EPSP maintained a good amount of H-bonds, more than 9 on average (9.48 ± 1.01), over the 30 ns MD simulation.

Moving to gallion, it kept a similar but slightly lower amount of H-bonds during the simulation (9.20 ± 1.14) while ED was the best maintaining more than 10 H-bonds on average (10.65 ± 1.30). On the other hand, VA was the worst, as expected from the docking pose, keeping less than 7 hydrogen bonds on average (6.57 ± 1.11) due to the lack of side chains useful for interacting with the polar environment of the pocket. These data suggested that, in line with its similarity with EPSP and gallion, ED could be a promising competitive CS inhibitor while VA should not obtain any result of this kind.

Activity assay. After computational analysis, an *in vitro* activity assay was performed to evaluate the inhibitory effects of gallion, ED and VA (see section 2.6). The ligands were tested in different concentrations, starting from 0.1 µM reaching 500 µM, and each measurement was done in triplicates.

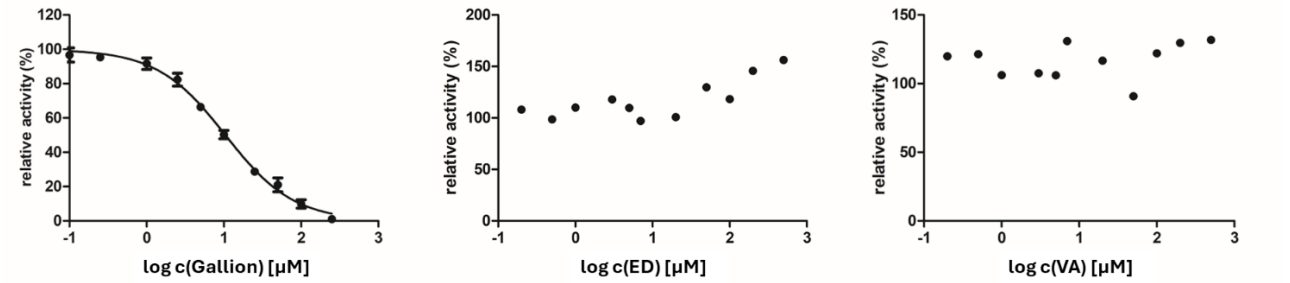


Figure 5*. Graphs reporting the relative activity of CS with gallion, ED and VA, at different concentrations. The log scale ranges from -1 (0.1 µM) to 3 (500 µM)*

As shown in figure 5, gallion acted as expected as a good CS inhibitor while on the other hand VA did not show any significant effect on the CS activity. Serendipitously ED, expected to be a promising inhibitor for the CS activity for the reasons stated above, enhanced significantly ($p < 0.05$) the activity of CS for the two highest concentrations tested (Figure 6).

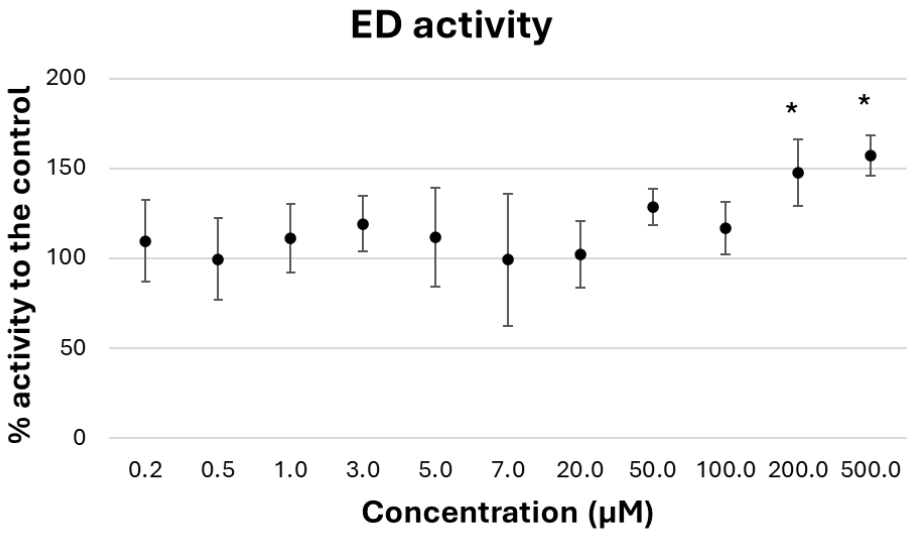


Figure 6. *Focus on the ED activity with each datapoint represented as mean ± SEM. As reported in the graph, the activity of CS is significantly enhanced at 200 µM and 500 µM, indicated with stars above the respective values.*

This behavior, unexpected based on the computational analysis though interesting, could be due to several reasons, which deserve future dedicated investigations. First, the computational analysis was performed on the *C. jejuni* CS while the *in vitro* activity assay was done on the *P. brasiliensis* one due to technical reasons like the high production yield and consolidated production protocol for the latter (see section 2.6). Despite the active site and overall folding are conserved some differences must be expected, with a possible impact on ligands recruitment. Therefore, the small differences in terms of primary and ternary protein structure may have had an impact on the capability of the *P. brasiliensis* enzyme to recruit the peptide as describe for *C. jejuni* CS. Second, peptides are intrinsically very flexible – way more flexible than both the inhibitor and substrate tested here. The higher flexibility might have had a role either in preventing the peptide from accessing the binding site, or from arranging it there as described by the *in silico* outcome. In this respect, chemical modifications on the peptide scaffold aimed at reducing its flexibility (e.g. exploiting staple peptides) might result both in a proper recruitment by the enzyme and in a proper arrangement at the catalytic site. Regarding the enhancement of the CS activity, which lies unexplained yet, it could be reasonably

due to a preferential binding to some other regions of CS, external to the active site, worth of further dedicated investigations. Indeed, CS is a homotetrameric enzyme and ED could either bind to floppy regions stabilizing them or to the interface among monomers, acting as a glue and enforcing their interactions to maintain the most active folding. The analysis of such mechanisms is important to extend the mechanistic understanding of a critical point of bacterial metabolism, which is fundamental to drive its “druggability” and design new counteracting anti-bacterial strategies. On the other hand, the observed peptide-dependent enhancement of CS activity described here draws attention to the matrix effects on bacterial growth. Indeed, if confirmed by future experiments, the effects of certain peptides on the activity of CS might have a role on the bacterial fitness when growing on specific protein substrates. This would extend the current understanding of the ecology of food-related bacteria with possible consequences on food technology, while pursuing the growth of desired strains, or food safety, while exploiting food composition to prevent the growth, spread and pathogenicity of undesired microorganisms.

4. Conclusion

The CS from *C. jejuni* was investigated through an *in silico* pipeline as it is an interesting target for antibacterial intervention. The molecular docking-based screening of the dipeptides library identified ED as a promising bioactive inhibitory peptide against CS. This is due to the mimicking of the natural substrate EPSP, especially regarding ED's carboxylates which superimposed with EPSP's carboxylate and phosphate – all negatively charged groups. Moreover, the pattern of interaction between CS and EPSP/ED was kept stable along the MD simulation and a comparable number of hydrogen bonds were maintained. However, the *in vitro* activity unexpectedly reported how ED acted as a CS enhancer at the two highest concentrations tested rather than inhibiting it. This is also concerning from a food safety perspective since typical matrices where *C. jejuni* is found in (e.g. poultry) are protein-rich matrices likely containing this kind of dipeptides.

The reasons behind this serendipitous discovery likely lie within the nature of one or both the two involved actors, CS and ED. Regarding the latter, it is a dipeptide and thus it is naturally very flexible. This flexibility could have prevented it from reaching CS active site with the orientation obtained with the docking procedure, and suitable to inhibit the enzyme. For this reason, follow-ups should be done and a stabilized version of ED – for instance exploiting peptidomimetics strategies – should be pursued to verify its ability to inhibit CS. Moreover, related to both ED flexibility and CS homotetrameric nature, the peptide could act as a complex stabilizer binding in region far from the binding site. This last part could be investigated through computational approaches seeking for possible pockets on the CS homotetramer surface.

Overall, this study provided a solid blueprint for further dedicated analysis of *C. jejuni* CS broadening the knowledge regarding its binding site chemical characteristics and general activity. Last, it gave

some clues on a possible scaffold to be used to develop newly effective and peptide-based CS inhibitors.

References

1. World Health, O., *WHO estimates of the global burden of foodborne diseases*. 2016, Genève, Switzerland: World Health Organization. 265.
2. Kenyon, J., et al., *Campylobacter outbreak associated with raw drinking milk, North West England, 2016.* Epidemiology & Infection, 2020. **148**.
3. Jurinovic, L., et al., *First Data on Campylobacter spp. Presence in Shellfish in Croatia.* Pathogens, 2022. **11**(8).
4. Mohammadpour, H., et al., *The prevalence of Campylobacter spp. in vegetables, fruits, and fresh produce: a systematic review and meta-analysis.* Gut Pathogens, 2018. **10**.
5. Devleesschauwer, B., et al., *Health and economic burden of Campylobacter*. Campylobacter: Features, Detection, and Prevention of Foodborne Disease, ed. G. Klein. 2017. 27-40.
6. Nachamkin, I., B.M. Allos, and T. Ho, *Campylobacter species and Guillain-Barre syndrome.* Clinical Microbiology Reviews, 1998. **11**(3): p. 555-+.
7. European Food Safety, A., et al., *The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2016.* Efsa Journal, 2017. **15**(12).
8. Hur, J.I., et al., *Cold tolerance in Campylobacter jejuni and its impact on food safety.* Food Research International, 2024. **175**.
9. Zhong, X., et al., *Campylobacter jejuni Biofilm Formation Under Aerobic Conditions and Inhibition by ZnO Nanoparticles.* Frontiers in Microbiology, 2020. **11**.
10. Herrmann, K.M., *The shikimate pathway as an entry to aromatic secondary metabolism.* Plant Physiol., 1995. **107**(1): p. 7-12.
11. Shende, V.V., K.D. Bauman, and B.S. Moore, *The shikimate pathway: gateway to metabolic diversity.* Natural Product Reports, 2024. **41**(4): p. 604-648.

12. Frlan, R., *An Evolutionary Conservation and Druggability Analysis of Enzymes Belonging to the Bacterial Shikimate Pathway.* Antibiotics-Basel, 2022. **11**(5).

13. Nunes, J.E.S., et al., *Mycobacterium tuberculosis Shikimate Pathway Enzymes as Targets for the Rational Design of Anti-Tuberculosis Drugs.* Molecules, 2020. **25**(6).

14. McConkey, G.A., *Targeting the shikimate pathway in the malaria parasite Plasmodium falciparum.* Antimicrobial Agents and Chemotherapy, 1999. **43**(1): p. 175-177.

15. Forouhar, F., et al., *Crystal Structure of the Chorismate Synthase from Campylobacter jejuni, Northeast Structural Genomics Target BR19.* 2004.

16. Macheroux, P., et al., *A unique reaction in a common pathway: mechanism and function of chorismate synthase in the shikimate pathway.* Planta, 1999. **207**(3): p. 325-334.

17. Pedroni, L., et al., *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**(21).

18. Bueno, P.S.A., et al., *New inhibitors of chorismate synthase present antifungal activity against Paracoccidioides brasiliensis.* Future Microbiology, 2019. **14**(11): p. 969-980.

19. Xuan, J.Q., et al., *Antimicrobial peptides for combating drug-resistant bacterial infections.* Drug Resistance Updates, 2023. **68**.

20. Adaro, M., et al., *Biosynthesis of a Novel Antibacterial Dipeptide, Using Proteases From South American Native Fruits, Useful as a Food Preservative.* Frontiers in Nutrition, 2021. **8**.

21. Raheem, N. and S.K. Straus, *Mechanisms of Action for Antimicrobial Peptides With Antibacterial and Antibiofilm Functions.* Frontiers in Microbiology, 2019. **10**.

22. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Research, 2000. **28**(1): p. 235-242.

23. Mirdita, M., et al., *ColabFold: making protein folding accessible to all.* Nature Methods, 2022. **19**(6): p. 679-+.

24. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-+.

25. Evans, R., et al., *Protein complex prediction with AlphaFold-Multimer.* 2021.

26. Pedroni, L., et al., *Free fatty acid receptors beyond fatty acids: A computational journey to explore peptides as possible binders of GPR120.* Current Research in Food Science, 2024. **8**.

27. Bateman, A., et al., *UniProt: the Universal Protein Knowledgebase in 2023.* Nucleic Acids Research, 2023. **51**(D1): p. D523-D531.

28. Hekkelman, M.L., et al., *AlphaFill: enriching AlphaFold models with ligands and cofactors.* Nature Methods, 2023. **20**(2): p. 205-+.

29. Kim, S., et al., *PubChem 2023 update.* Nucleic Acids Research, 2023. **51**(D1): p. D1373-D1380.

30. O'Boyle, N.M., et al., *Open Babel: An open chemical toolbox.* Journal of Cheminformatics, 2011. **3**.

31. Jones, G., et al., *Development and validation of a genetic algorithm for flexible docking.* Journal of Molecular Biology, 1997. **267**(3): p. 727-748.

32. Chartier, M. and R. Najmanovich, *Detection of binding site molecular interaction field similarities.* J. Chem. Inf. Model., 2015. **55**(8): p. 1600-1615.

33. Abraham, M.J., et al., *GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers.* SoftwareX, 2015. **1-2**: p. 19-25.

34. Best, R.B., et al., *Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone φ, ψ and Side-Chain χ1 and χ2 Dihedral Angles.* Journal of Chemical Theory and Computation, 2012. **8**(9): p. 3257-3273.

35. Bugnon, M., et al., *SwissParam 2023: A Modern Web-Based Tool for Efficient Small Molecule Parametrization.* Journal of Chemical Information and Modeling, 2023. **63**(21): p. 6469-6475.

36. Maclean, J. and S. Ali, *The structure of chorismate synthase reveals a novel flavin binding to a unique chemical reaction.* Structure, 2003. **11**(12): p. 1499-1511.

Chapter III

Structural insights on Internalin A – Human E-cadherin interaction to prevent *Listeria monocytogenes* internalization

This chapter presents still unpublished results as outcome of the work carried out during my period abroad at the University of Barcelona, under the supervision of Professor Francisco Javier Luque Garriga and colleagues.

Abstract

Listeria monocytogenes poses significant health risks, especially due to its resilience in various environments and potential to cause severe infections, especially pregnant women and generally in vulnerable populations. Here we explored the potential of natural compounds to prevent *L. monocytogenes* internalization, focusing on the prevention of the interaction between the bacterial protein internalin A (InlA) and the human E-cadherin (Ecad). An *in silico* pipeline consisting of molecular docking and dynamics simulations was applied to identify key clusters of interaction between InlA and Ecad and natural compounds possibly preventing such interaction. Two compounds, CNP0101638 and CNP0154950, demonstrated stable interactions with InlA, potentially shielding key residues and preventing the InlA-Ecad complex formation. Overall, this research provided structural and mechanistic insights to better define InlA as a possible target to mitigate *L. monocytogenes* internalization. These findings highlighted the potential for natural compounds to act as food/feed additives or food-contact materials coatings, proposing novel antimicrobial strategies to reduce *L. monocytogenes* pathogenicity. Further dedicated *in vitro* and *in vivo* studies should be carried out to properly pave the way for applications in a food safety framework.

1. Introduction

Listeria monocytogenes is a widespread and ubiquitous gram-positive and facultative anaerobic bacterium. It can be found everywhere, ranging from soil to water to food or feed (REF), and it can persist even at refrigeration temperatures (from 0 °C) and under a broad pH range (roughly from 4 to 9) [1, 2]. For this reason, the contamination might occur during many steps of food processing, and it is especially concerning for dairy products, vegetables or ready-to-eat products [3, 4]. *L. monocytogenes* is the infamous causative agent of listeriosis, a concerning foodborne disease affecting animals and humans. Although the number of *Listeria*-related infections is way lower compared to more common foodborne pathogens the listeriosis' global burden remains significative, both in terms of public health and food production [5]. This is mainly due to the high mortality rate of listeriosis, up to 30%, especially in people belonging to the so-called YOPI group (young, old, pregnant, immunosuppressed) [5, 6].

L. monocytogenes infection begins with the invasion of human intestinal cells where it crosses their membrane by inducing its own endocytosis [7, 8]. Once internalized, the bacterium starts replicating and eventually kills the host cell and spreads to the next ones, potentially entering the host's bloodstream and causing more severe outcomes because of systemic and multi-organ infections [9]. The internalization of *L. monocytogenes* is primarily mediated by a class of proteins exposed on its outer membrane, known as internalins (Inl), with 5 main representatives, namely InlA/B/C/F/P [7]. Among these, InlA/B/F facilitate the entry of the bacterium within the host cells by interacting with the human (or mammalian) E-cadherin (Ecad) [7]. Of note, InlA-driven internalization is the one promoting *L. monocytogenes* quick transcytosis across the intestinal epithelium, possibly leading to a systemic infection and severe outcomes such as meningoencephalitis or, in pregnant women, foetal-placental infection [8-10]. Interestingly, the InlA-Ecad interaction has been extensively

investigated through crystallographic studies providing a useful background for modelling approaches [11-13]. Regarding InlA, it is a 460 residues-long protein with an α/β folding and characterized by a leucine-rich repeats (LRR) domain adjacent to an inter-repeat domain resembling an immunoglobulin fold [7, 13] (Figure 1).

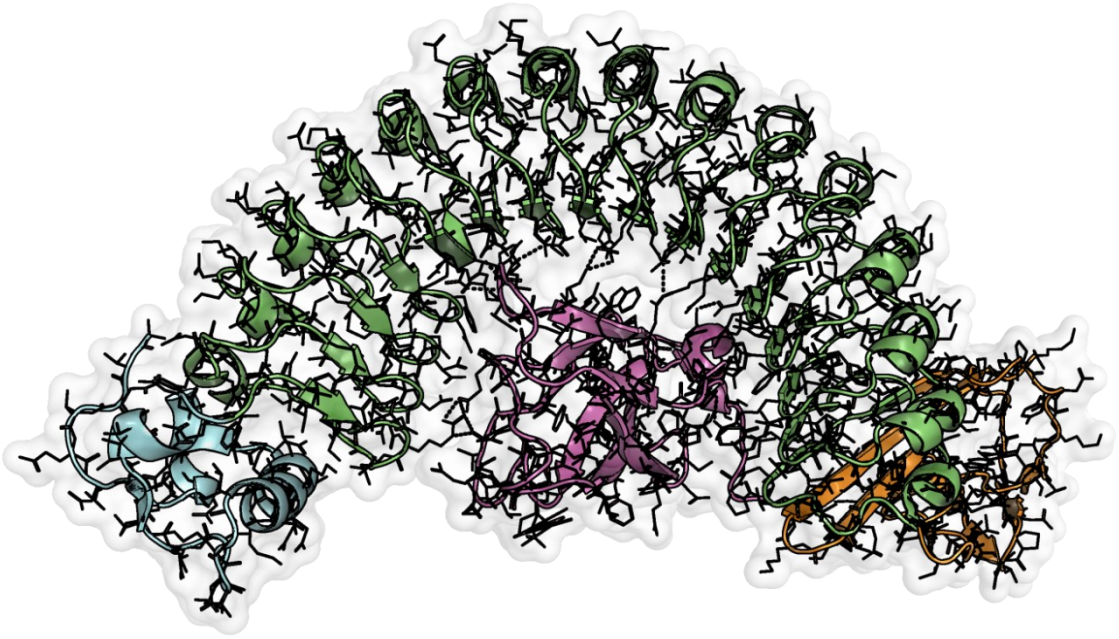


Figure 1. *Internalin A – human E-cadherin crystallographic structure (PDB ID 1O6S [13]). Both InlA and Ecad are represented as cartoons, with residues represented as lines and with white transparent surface. The Ecad is magenta while InlA has a cyan C-terminal, a green LRR domain and an orange IR domain.*

It is the LRR domain which binds to the human Ecad extracellular domain (EC1) leading to *L. monocytogenes* internalization [14]. In this sense, as already reported in other studies [15, 16], finding strategies effectively preventing the Ecad-InlA interaction could serve as a promising strategy to hamper the bacterial invasion possibly mitigating *L. monocytogenes* virulence. Of note, some *L. monocytogenes* strains with InlA mutations are not harmful being unable to efficiently recognize cadherins [17, 18]. This was corroborated through *in silico* studies which proved how some key InlA mutations led to a less favoured interaction with Ecad possibly disfavouring *L. monocytogenes* internalization [19]. Moreover, pathogenic *Listeria* strains for humans were not able to infect mice

largely due to differences in some residues on the mouse Ecad preventing the interaction with InlA [20-22].

In this framework, we exploited computational approaches to seek for natural compounds able to prevent the formation of the InlA-Ecad complex. Briefly, we collected natural compounds from three publicly available databases, namely COCONUT (https://coconut.naturalproducts.net/) [23], LOTUS (https://lotus.naturalproducts.net/documentation) [24] and FOODB (https://foodb.ca/), filtered them from more than 600 000 to roughly 4 000. These compounds were investigated using a docking-based screening to identify candidates capable of interacting with InlA to inhibit the InlA-Ecad complex formation. For the most promising compounds, we performed molecular dynamics simulations to verify their stability at the InlA interface ideally masking it from Ecad, consequently hampering InlA-Ecad interaction. This computational pipeline offers a promising strategy for identifying possible food-grade natural compounds inhibiting or attenuating *L. monocytogenes* pathogenic action and it could pave the way for the development of possible novel food/feed additives or “functionalized” food contact materials.

2. Materials and Methods

2.1. Data source

2.1.1. Protein structure

The 3D structure of the *Listeria monocytogenes* InlA – human Ecad complex was obtained from the Protein Data Bank (PDB; https://www.rcsb.org) (REF), under the PDB ID 1O6S [13]. The structure was resolved at 1.80 Å through X-ray diffraction and every residue was fully resolved. Starting from 1O6S the Inla – murine Ecad complex was obtained mutating human Ecad residues to the murine's ones. This was done exploiting the PyMol (v. 2.3) *Wizard mutagenesis* tool in agreement with previous studies (REF).

2.1.2. Natural compounds library

The chemical structures of natural compounds were retrieved from three publicly available databases: COCONUT (https://coconut.naturalproducts.net/) [23], LOTUS (https://lotus.naturalproducts.net/) [24] and FOODB (https://foodb.ca/). Both the COCONUT and LOTUS databases were downloaded in the SMILE format while FOODB in the JSON format. From the FOODB JSON file, only the information related to molecule name and SMILE were kept, as per the other two databases.

The three databases in SMILE format were merged, and the duplicate entries were removed. Subsequently, they were filtered using an *ad hoc* python script mainly based on the RDKit toolkit implemented for python. The following criteria were followed: logP ($-1 \leq \log P \leq 3$), rotatable bonds (excluding hydrogens, rotatable bonds ≤ 5), molecular weight ($550 \leq$ molecular weight ≤ 700) and "molecular length", defined as the highest distance between two atoms belonging to the same molecule (molecular length ≤ 15 Å). The first two parameters, logP and rotatable bonds, were set to

obtain molecules possibly acting in a water-based environment without excessive flexibility and thus with reduced computational costs. The last two size-driven parameters were set to filter only molecules able to interact with the InlA region to be targeted (see section 3.2).

The filtering procedure led to a reduction by two orders of magnitude of the considered dataset, from over 600 000 starting compounds to 4 086 compounds. Eventually, the *Ligprep* module of MAESTRO [25] was used to prepare the 3D structure library of these 4 086 compounds in the SDF format. Specifically, it was set to generate up to 8 stereoisomers and 6 tautomers per compound, considering a protonation state at a pH of 7.4. This generated a final library of more than 20 000 conformers.

2.3. Molecular dynamics simulations

Molecular dynamics (MD) simulations were conducted using the AMBER20 package [26]. InlA and Ecad were parametrized adopting the ff19SB forcefield [27] while the ligands with the GAFF forcefield [28]. All the simulated systems were neutralized (adding Na^+ and Cl^- at 0.15 M) and embedded within an octahedral box of OPC3 water molecules (roughly 45 000 water molecules/system) [29]. Each system was energetically minimized in more steps starting from the hydrogens minimization, moving to the solvent (water and ions) and eventually to all the atoms. After minimization, each system was equilibrated via thermalization to 300 K (within 600 ps, 700 ps or 800 ps, depending on the replica) and pressurization to 1 bar (within 1 ns). Prior to the MD production, an equilibration in NVT ensemble was run for 1 ns. Eventually for the MD production, 3 replicas of 1 µs each were run for the InlA – E-cad complex (alternatively with human or murine Ecad) while for the InlA – ligand systems they were shortened to 400 ns. Regarding the latter, a restraint of 5.0 kcal/mol/Å^2 for the ligand was kept during the equilibration phase lowered to 2.5 kcal/mol/Å^2 during the first 10 ns of the MD production and eventually removed until the end (400

ns). Replicas for InlA – ligand were run only for the ones showing an interesting behaviour during the first replica. Of note, prior to running InlA – ligand MD simulations the partial atomic charges for the ligands were derived using the RESP protocol [30] at the B3LYP/6-31G(d) level of theory through Gaussian16 [31].

A per residue free energy decomposition analysis was conducted exploiting the MMPBSA.py [32], available in the AmberTools package [26], to study the differences in terms of energy contribution to the InlA-Ecad complex of the eleven distinct residues between the murine and the human Ecad.

2.4. Docking-based screening

The virtual screening of the filtered dataset of natural compounds was completed exploiting molecular docking simulations. It was conducted using the Glide docking program from Schrödinger [33-35]. It was used a semi-flexible procedure, keeping the receptor (InlA) rigid while allowing the ligands to be fully flexible. The box considered by Glide to perform the docking was set to embed from cluster 1 to cluster 3 of the InlA – Ecad complex, previously identified through MD simulations (see section 3.1). The first round of docking-based screening was performed using HTVS mode, the most computationally efficient. Subsequently, the top 4 000 conformers were docked using the SP mode, more precise, allowing final torsional refinement and sampling. Apart from the mode, the other parameters were kept as per default setting allowing the generation of 10 poses per ligand (*POSES_PER_LIG 10*) and the ring conformation energy below 2.5 kcal/mol (*RINGCONFCUT 2.500000*).

3. Results and Discussion

3.1. Characterization of the InlA-Ecad interface

We applied a computational pipeline based on MD simulations as they already proved high reliability in gaining insights on the InlA-Ecad interaction [19]. Three replicas of 1 μs long MD simulations were run to investigate the InlA – E-cad complex stability (PDB ID 1O6S [13]) and to identify key residues contributing to the inter-molecular interaction.

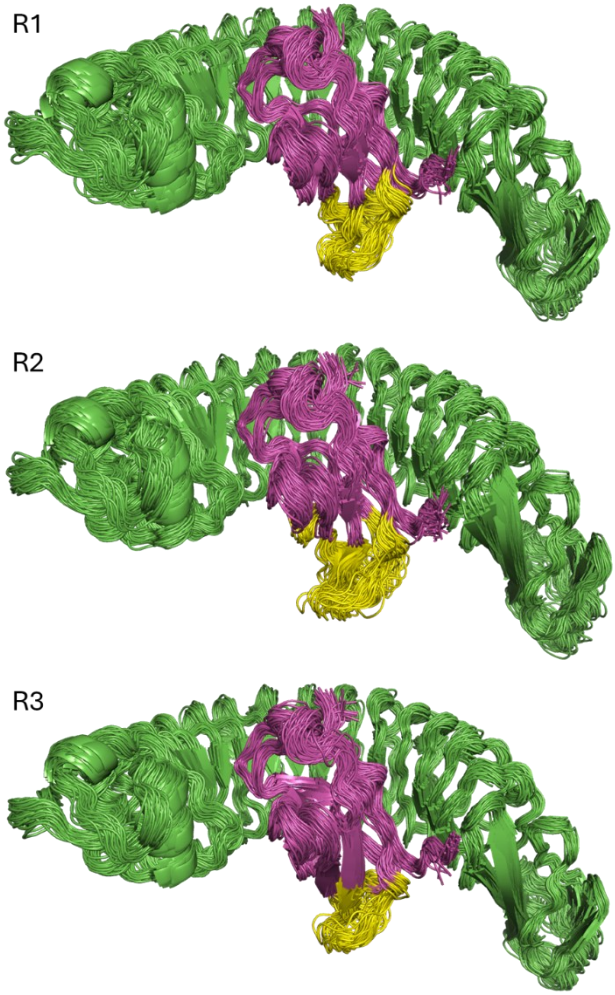


Figure 2. *Collection of 50 snapshots (1 every 20 ns) of the 1 μs long MD simulations replicas (R1, R2, R3). InlA is represented as green cartoon, Ecad as magenta cartoon with the mobile loop in yellow.*

As shown in figure 2, apart from the N- and C-terminal regions expected to be more mobile in both the considered proteins, the overall complex was stable especially on the InlA side. On the other hand, human Ecad showed a loop (reported in yellow in Figure 2) which was consistently less stable, especially in the second replica. This was expected since this region is the most exposed to the solvent and it is not interacting with InlA. Moreover, the root mean squared deviation (RMSD) of both InlA and Ecad was computed confirming what visually observed, i.e. a good stability of the complex over the three replicas (Figure S1). Once established the stability of the complex, the presence of possible conserved patterns of interaction between InlA and Ecad was analysed.

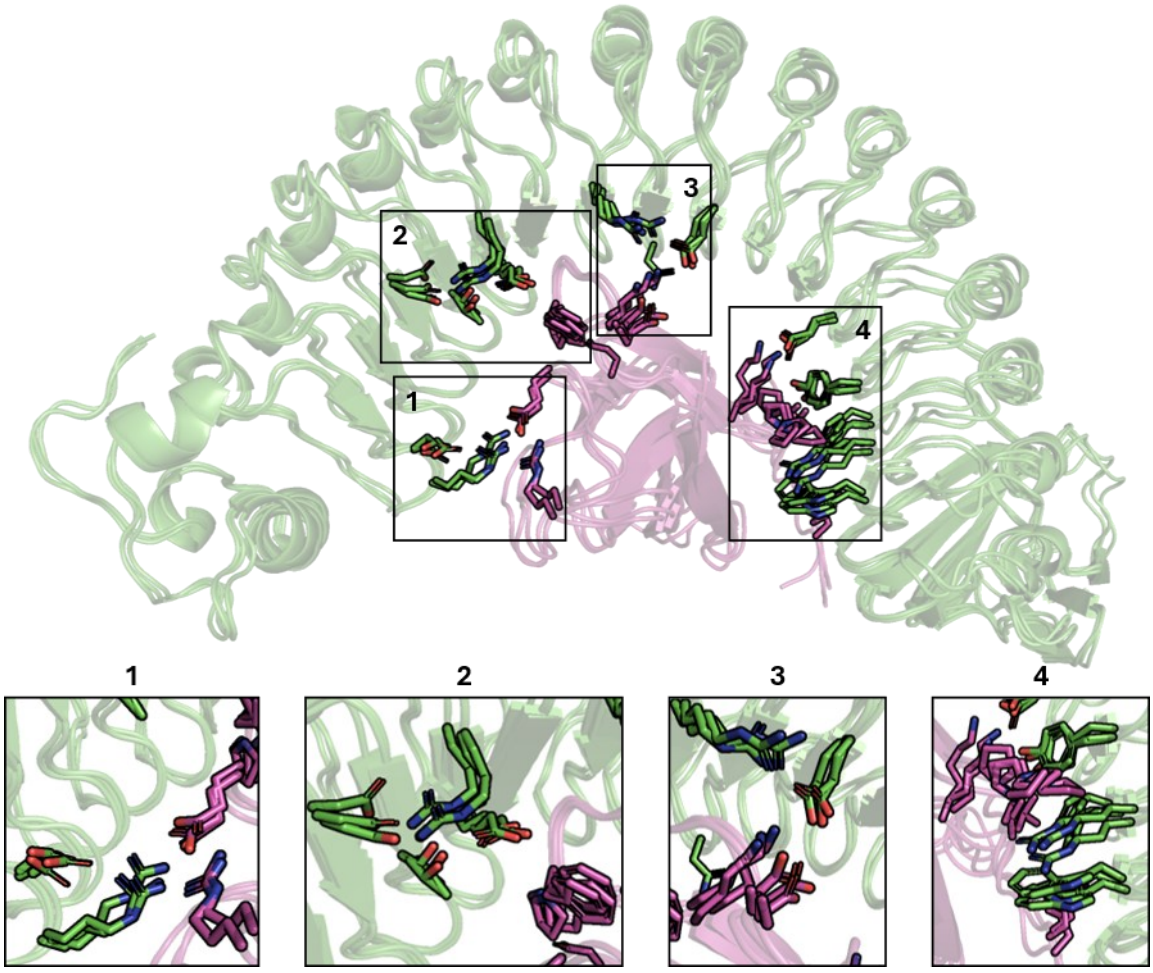


Figure 3. *Inla (green) – Ecad (magenta) representation of the three replicas with most represented interacting residues reported as sticks. The four identified clusters of interaction are numbered (1 to 4) and embedded within black squares. Of note, the visible gap between cluster 3 and 4 there is occupied by a conserved cavity on the InlA-Ecad interface.*

As reported in Figure 3, four conserved interaction clusters were observed along with a cavity, possibly suitable to receive ligands, between the third and fourth cluster. The identification of those interaction hot-spots was crucial to define the regions to target for the sake of preventing/impairing the InlA-Ecad interaction. Indeed, the aim of this analysis was finding the best InlA region to target through the following docking-based screening. This was done to possibly find compounds masking the interaction hot-spots and preventing the interaction of InlA with Ecad. In this sense, while the first three clusters were mainly characterized by the presence of arginine and aspartates residues, conferring a typical polar environment, the fourth one was mainly hydrophobic-driven and considered less suitable and more challenging to be possibly targeted by natural compounds. Moreover, the first three clusters cover more than a half of the InlA-Ecad interaction surface (Figure 3). Therefore, identifying natural compounds able to steady interact with these three InlA-side clusters – shielding them from the Ecad surface thus preventing their interaction – was considered more promising rather than working on the most isolated fourth one. The last reason to prefer targeting the region covering clusters from 1 to 3 is related to the differences between human and murine Ecad. As proven in previous studies, mice are resistant to *L. monocytogenes* infection since their Ecad brings some different residues exposed to the interface with InlA [20, 21]. For this reason, we performed three replicas of 1 μs long MD simulations of the mouse Ecad-InlA complex. Despite we observed a geometrical stability comparable to the ones with human Ecad, we analysed the energetic contribution of the eleven distinct residues between the murine and the human Ecad. We performed a per-residue free energy decomposition analysis and observed how the major differences in terms of free energy were due to two main different residues: a proline in the human Ecad replaced by a glutamic acid in the murine Ecad and a glutamic acid in the human Ecad replaced by a glutamine in the murine one (Table 1). These results confirmed the importance of the proline residue on the human Ecad, embedded between the second and third cluster, in driving a proper

interaction with InlA [16, 21, 22]. Therefore, shielding this interaction along with the one with glutamic acid – contained in the first cluster – could be a promising strategy for preventing *L. monocytogenes* internalization.

Table 1. *Per-residue free energy decomposition of the 11 distinct residues between human and murine Ecad. The position reported here is relative to the crystal structure 1O6S (REF).*

E-cadherin	Residue	Position	Average (kcal/mol)
Human	**PRO**	**480**	**-6,969487259**
	LYS	492	0,627559344
	GLY	496	0,11704
	THR	509	0,251426667
	GLU	**528**	**-3,888673168**
	ARG	534	1,272199163
	THR	537	0,184073333
	THR	539	0,181073333
	PHE	541	0,087384763
	ASN	550	0,025193333
	LEU	559	0,181039328
	Total contribution		-7,93117
Murine	**GLU**	**480**	**-5,177202827**
	ARG	492	0,298311616
	THR	496	0,149826667
	LYS	509	0,346166667
	GLN	**528**	**-1,889393504**
	ALA	534	0,058553333
	LYS	537	0,46956
	ILE	539	0,135506667
	TYR	541	0,120046667
	GLU	550	0,052313333
	VAL	559	0,335286912
	Total contribution		-5,101024469

3.2. Docking-based screening and analysis of InlA-ligands stability

Prior to running the docking-based screening we investigated the InlA stability by removing the Ecad from the complex. We performed three replicas of 1 μs long MD simulations and observed that the overall structure of InlA remained stable, as observed when in complex with Ecad (Figure S2).

The library of natural compounds was built collecting more than 600 000 chemicals present in three publicly available databases, namely COCONUT (https://coconut.naturalproducts.net/) [23], LOTUS (https://lotus.naturalproducts.net/) [24] and FOODB (https://foodb.ca/). We decided to filter the whole collection based on molecular weight, between 550 and 700 Da, and on molecular length, over 15 Å, to select only molecules possibly able to cover the region spanning from cluster 1 to 3. Regarding the logP, it was set between -1 and 3 to select only compounds with a moderate hydrophobicity but keeping a good balance with solubility, fundamental for making them soluble in water-based environments [36]. Finally, the rotatable bonds were set to be less than 6 to have molecules with a limited conformational flexibility. This was also crucial to keep a good computational efficiency while building the final compounds library (4 096 different natural compounds with more than 20 000 conformers) and to test it via the docking-based screening procedure.

The first step consisted in screening the whole 20 000 conformers library using a less precise but more efficient mode (HTVS). The top 4 000 obtained poses were re-docked using a more precise mode (SP) (see section 2.4). The first 500 poses were visually inspected on PyMol (v. 2.3) to identify the most promising compounds to proceed with further MD simulations.

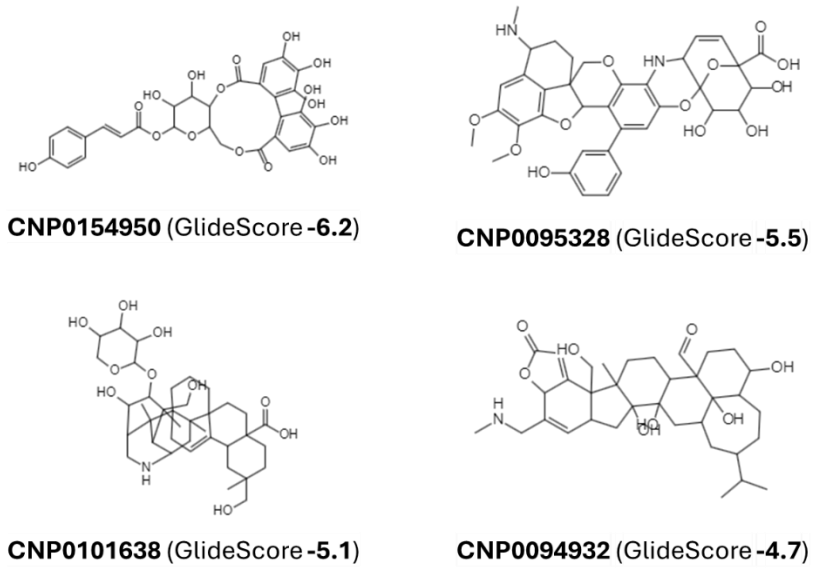


Figure 4. *Chemical structures of the selected from the docking-based screening procedure along with their database ID and docking score.*

As reported in figure 4, the four chosen molecules were all from COCONUT database entries (REF), and they obtained comparable GlideScores (from -6.2 to -4.7 kcal/mol), a value approximating the ligand binding free energy [33]. Regarding CNP0154950, it was chosen since it was the first ranked pose, and 5 conformers appeared among the top ten poses obtained from the final docking simulation (pose 1, GlideScore -6.2). The other three compounds (CNP0095328 pose 13, GlideScore -5.5; CNP0101638 pose 26, GlideScore -5.1; CNP0094932 pose 40; score -4.7) were chosen because their visual inspection showed interesting patterns of interactions that, if maintained stable, can be expected to prevent the InlA-Ecad complex formation.

All four ligands were tested via a 400 ns long MD simulation and only for stable compounds two more replicas would have been run. Regarding CNP0094932, it showed a clear instability along the MD simulation eventually detaching from InlA. CNP0095328 lost the first pattern of interaction whose justified its selection, partially detaching from InlA (Figure 5). For these reasons they were both discarded.

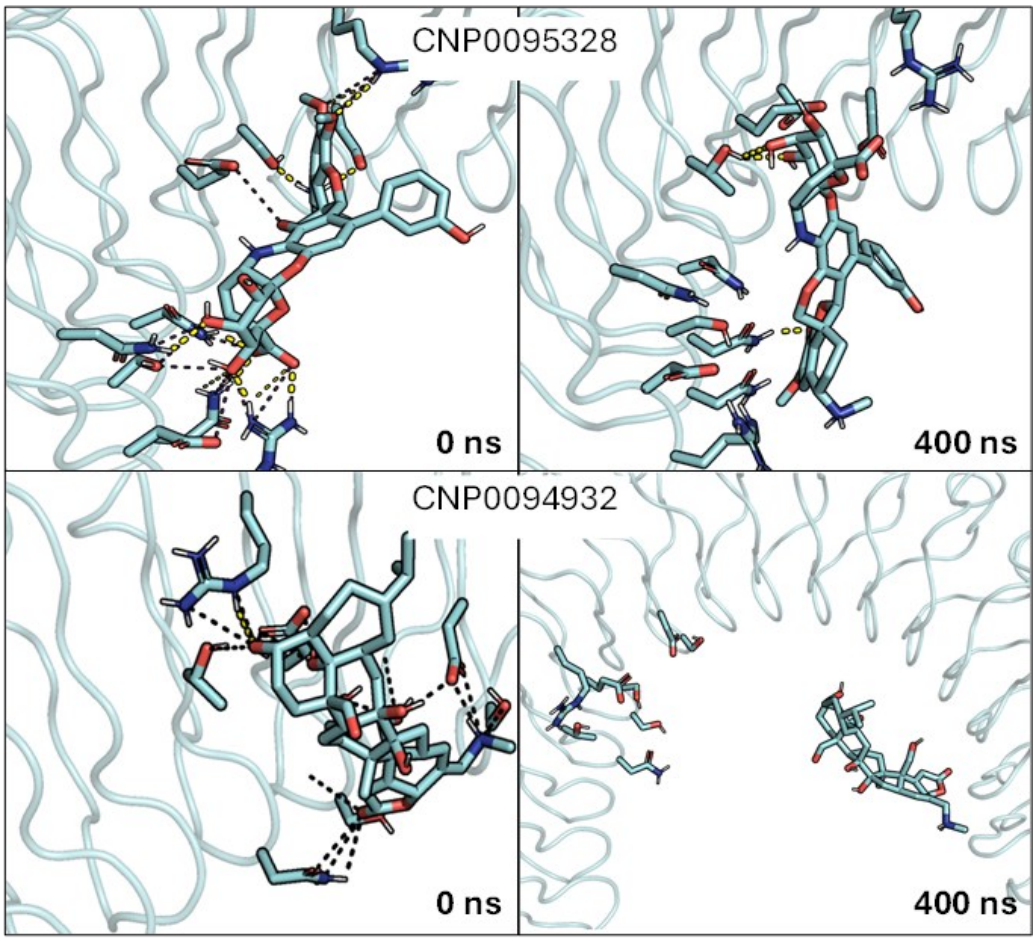


Figure 5. *Focus on the first (0 ns) and last (400 ns) frame of the MD simulations between InlA and the two discarded ligands. Both ligands, CNP0095328 and CNP0094932, are represented as cyan sticks with polar interactions reported as dashed lines. In the upper part it can be appreciated how the initial pattern of interaction between InlA and CNP0095328 is not maintained while in the lower part the complete detachment of CNP0094932 from InlA can be observed.*

On the other hand, compounds CNP0154950 and CNP0101638 showed a good stability and a comparable behaviour in the first, second and third replica (Figure S3). Moreover, as reported in Figure 6 they kept a similar pattern of interactions starting from the first frame (0 ns) to the last frame (400 ns) of the MD simulation pointing to their overall stability over that specific InlA region. As shown in Figure S3, CNP0154950 RMSD was stable over 400 ns for two distinct MD simulations replicas. Of note, in one replica it showed some small short-lasting readjustments within the first half of the simulation, but it reached a comparable steady state keeping a high amount of hydrogen bonds till the end of simulation. Interestingly, briefly after the beginning of the MD simulations, in

all the replicas, we noticed the formation of a “hole” next to the ligand tyrosine-moiety. This was due to water molecules acting as bridge to maintain a hydrogen bond between a hydroxyl group from CNP0154950 and an InlA glutamate (Figure 7).

CNP0101638 behaved similarly maintaining a comparably stable trend as shown in the RMSD in figure S3. Specifically, in addition to a good amount of hydrogen bonds it kept a salt bridge between its carboxylate moiety and an InlA arginine and another one between its charged nitrogen and an InlA glutamate.

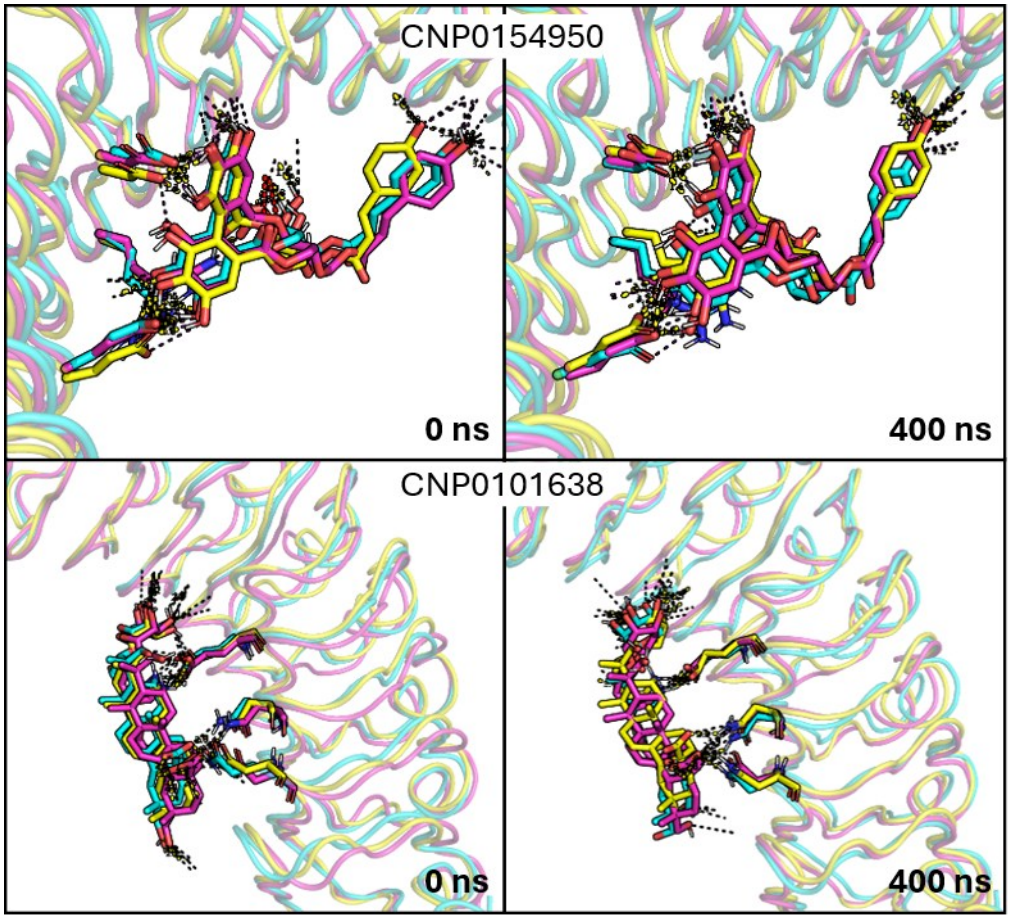


Figure 6. *Focus on the first (0 ns) and last (400 ns) frame of the three MD simulations replicas between InlA and the two promising ligands. In the upper part, CNP0154950 and the interacting residues are represented as sticks and differently coloured depending on the replica (magenta, cyan, yellow) with polar interactions reported as dashed lines. In the lower part, CNP0101638 and the interacting residues are represented as sticks and differently coloured depending on the replica (magenta, cyan, yellow) with polar interactions reported as dashed lines.*

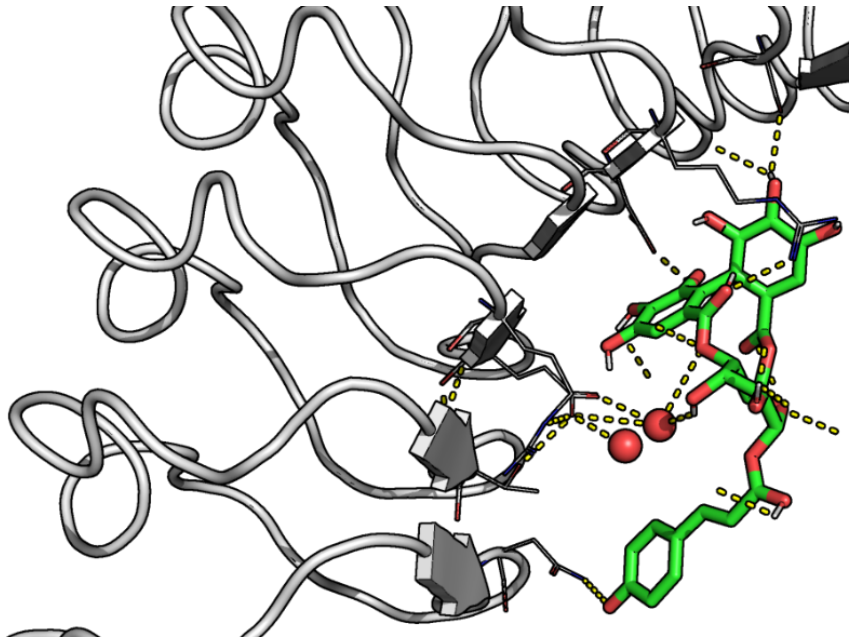


Figure 7. *Focus on the "hole" formed between CNP0154950 and InlA. InlA is reported as a white cartoon, the ligand as green sticks while the interacting InlA residues as white lines with polar interaction reported as yellow dashed lines. Of note, water is reported as red sphere, and it can be noticed how water molecules act as bridges to keep a hydrogen-bonding interaction between an hydroxylate group belonging to the ligand and an InlA glutamate.*

3.3 Usage of compounds of natural origin to enhance food safety

The usage of compounds of natural origin with antimicrobial properties in food gained attention from both consumers and the food industry itself [37]. The first reason is due to the abuse of antibiotics which selected more and more resistant pathogenic strains, pushing researchers to find alternatives to the so-called conventional methods [38]. The second reason is related to the consumers' growing trend to accept products with additives from natural sources over those produced by synthesis [39]. Apart from these reasons, a wealth of natural compounds, besides the well-known "natural" antibiotics described so far, already proved good antibacterial properties against a broad range of foodborne pathogens, including *L. monocytogenes* [40-42]. Regarding the natural compounds identified in this study, CNP0154950 and CNP0101638, they belong to two

different class of molecules of natural origin: gallotannins and oleanane triterpenoids. Both classes proved already antibacterial properties such as bacteriostatic activity, inhibition of biofilm formation, growth inhibition or bacterial toxins neutralization [42-44]. Moreover, there are studies showing the antibacterial activity of these classes of compounds against *L. monocytogenes* [45-47] pointing to the likeliness of a real-world application for the outcomes collected. A wealth of studies described the extractability of gallotannins from many different sources, including some considered for or deriving from the food production chain [48]. Concerning oleanane triterpenoids, oleanolic acids are widely diffused in plants with many high-content species that entered the food production chain a long time ago, like *Olea europaea L.*, and for whom the extraction of oleanolic acids fractions is reasonably to consider [49]. Indeed, moving to a real-world scenario, the applicability of the two identified compounds to enhance the food safety of certain products could exploit the pythocomplexes included in plant's extracts, considering the presence of congeners and analogues having a comparable function as likely. This aligns with many of the previously reported studies, as the isolation individual compounds would be certainly more challenging both time- and money-wise. Furthermore, phytocomplexes might better meet "clean label" trends, not underestimating the consumer's preferences [50]. However, their toxicity, their possible combined effect with other compounds, and more in general their safety profile should be thoroughly investigated prior to their usage as food safety enhancers. Last, it should always be considered how these compounds when added to certain food matrices as additives, rather than being used as sanitizers to clean out the producing environment, could possibly lead to unwanted taste/aroma. For this reason, further research should be also conducted to find the best threshold to achieve the desired antibacterial activity without altering the sensory characteristics.

4. Conclusion

This study deeply investigated through molecular modelling approaches the InlA – human Ecad complex and ways to prevent the complex formation by using natural compounds. This was done since this interaction is known to led to *L. monocytogenes* internalization and infection giving rise to listeriosis. In the first part, through MD simulations, the InlA-Ecad interface was thoroughly characterized identifying four promising clusters of interaction with three of them (from 1 to 3) having a high interaction potential. Moreover, within the region covered by these 3 clusters, there are two Ecad key-residues (480PRO and 528GLU as per PDB ID 1O6S [13]) whose mutation in mice appears to avoid InlA proper interaction and *L. monocytogenes* internalization. For these reasons, these 3 clusters might be targeted on the InlA-side to prevent the complex formation.

In the second part, through a docking-based screening strategy we tested natural compounds to investigate their capability of interacting with InlA and shielding the previously identified crucial region from Ecad, preventing the interaction. Specifically, out of the four chosen compounds only CNP0101638 and CNP0154950 demonstrated stable interactions with the targeted clusters over 400 ns MD simulations. CNP0154950 maintained consistent binding through hydrogen bonds and two salt bridges while CNP0101638 formed interestingly a water-mediated hydrogen bond with an InlA glutamate. Concluding, the identification of these two natural compounds as stable InlA binders on the main Ecad interface highlights their potential as inhibitors of *L. monocytogenes* internalization.

Overall, this study provided a solid foundation for further dedicated *in vitro* and *in vivo* analysis that could confirm the effectiveness of the identified compounds, along testing their safe use in food production. Moreover, provided they prove to be safe, the identification of natural compounds can pave the way to naturally derived antimicrobial strategies possibly using such compounds, or the pythocomplexes they are included in, as feed/food additives or as chemicals functionalizing food-contact materials or food-processing surfaces.

References

1. te Giffel, M.C. and M.H. Zwietering, *Validation of predictive models describing the growth of Listeria monocytogenes.* International Journal of Food Microbiology, 1999. **46**(2): p. 135-149.
2. Walker, S.J., P. Archer, and J.G. Banks, *GROWTH OF LISTERIA-MONOCYTOGENES AT REFRIGERATION TEMPERATURES.* Journal of Applied Bacteriology, 1990. **68**(2): p. 157-162.
3. Li, X., et al., *Prevalence of Listeria monocytogenes in Milk and Dairy Product Supply Chains: A Global Systematic Review and Meta-analysis.* Foodborne Pathogens and Disease, 2024. **21**(9): p. 526-535.
4. Alegbeleye, O. and A.S. Sant'Ana, *Survival and growth behaviour of Listeria monocytogenes in ready-to-eat vegetable salads.* Food Control, 2022. **138**.
5. de Noordhout, C.M., et al., *The global burden of listeriosis: a systematic review and meta-analysis.* Lancet Infectious Diseases, 2014. **14**(11): p. 1073-1082.
6. Khalil, A., et al., *Listeria outbreaks cause maternal and perinatal mortality and morbidity: we must do better.* Lancet Microbe, 2023. **4**(4): p. e206-e207.
7. Ireton, K., et al., *Role of internalin proteins in the pathogenesis of Listeria monocytogenes.* Molecular Microbiology, 2021. **116**(6): p. 1407-1419.
8. Drolia, R. and A.K. Bhunia, *Crossing the Intestinal Barrier via Listeria Adhesion Protein and Internalin A.* Trends in Microbiology, 2019. **27**(5): p. 408-425.
9. Nikitas, G., et al., *Transcytosis of Listeria monocytogenes across the intestinal barrier upon specific targeting of goblet cell accessible E-cadherin.* Journal of Experimental Medicine, 2011. **208**(11): p. 2263-2277.
10. Lecuit, M., *Listeria monocytogenes, a model in infection biology.* Cellular Microbiology, 2020. **22**(4).

11. Yamazaki, T., et al., *Anti-InlA single-domain antibodies that inhibit the cell invasion of Listeria monocytogenes.* Journal of Biological Chemistry, 2023. **299**(10).

12. Wollert, T., D.W. Heinz, and W.D. Schubert, *Thermodynamically reengineering the listerial invasion complex InIA/E-cadherin.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(35): p. 13960-13965.

13. Schubert, W.D., et al., *Structure of internalin, a major invasion protein of Listeria monocytogenes, in complex with its human receptor E-cadherin.* Cell, 2002. **111**(6): p. 825-836.

14. Mengaud, J., et al., *E-cadherin is the receptor for internalin, a surface protein required for entry of L-monocytogenes into epithelial cells.* Cell, 1996. **84**(6): p. 923-932.

15. Shivaee, A., et al., *Investigating the effect of the inhibitory peptide on L.monocytogenes cell invasion: an in silico and in vitro study.* Gut Pathogens, 2023. **15**(1).

16. Sahu, S.C., et al., *A synthetic polypeptide based on human E-cadherin inhibits invasion of human intestinal and liver cell lines by Listeria monocytogenes.* Journal of Medical Microbiology, 2007. **56**(8): p. 1011-1016.

17. Van Stelten, A., et al., *Listeria monocytogenes Isolates Carrying Virulence-Attenuating Mutations in Internalin A Are Commonly Isolated from Ready-to-Eat Food Processing Plant and Retail Environments.* Journal of Food Protection, 2016. **79**(10): p. 1733-1740.

18. Nightingale, K.K., et al., *inlA Premature Stop Codons Are Common among Listeria monocytogenes Isolates from Foods and Yield Virulence-Attenuated Strains That Confer Protection against Fully Virulent Strains.* Applied and Environmental Microbiology, 2008. **74**(21): p. 6570-6583.

19. Dellafiora, L., et al., *A Structural Study on the Listeria Monocytogenes Internalin A-Human E-cadherin Interaction: A Molecular Tool to Investigate the Effects of Missense Mutations.* Toxins, 2020. **12**(1).

20. Genheden, S. and L.A. Eriksson, *OF MICE AND MEN: DISSECTING THE INTERACTION BETWEEN LISTERIA MONOCYTOGENES INTERNALIN A AND E-CADHERIN.* Computational and Structural Biotechnology Journal, 2013. **6**(7): p. e201303022.

21. Wollert, T., et al., *Extending the host range of Listeria monocytogenes by rational protein design.* Cell, 2007. **129**(5): p. 891-902.

22. Lecuit, M., et al., *A single amino acid in E-cadherin responsible for host specificity towards the human pathogen Listeria monocytogenes.* Embo Journal, 1999. **18**(14): p. 3956-3963.

23. Sorokina, M., et al., *COCONUT online: Collection of Open Natural Products database.* Journal of Cheminformatics, 2021. **13**(1).

24. Rutz, A., et al., *The LOTUS initiative for open knowledge management in natural products research.* Elife, 2022. **11**.

25. Ligprep, S., *Schrödinger Release 2021-1: LigPrep. Schrödinger, LLC*. 2021, New York, NY.

26. Case, D.A., et al., *The AmberTools.* Journal of Chemical Information and Modeling, 2023. **63**(20): p. 6183-6191.

27. Tian, C., et al., *ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution.* Journal of Chemical Theory and Computation, 2020. **16**(1): p. 528-552.

28. Wang, J.M., et al., *Development and testing of a general amber force field.* Journal of Computational Chemistry, 2004. **25**(9): p. 1157-1174.

29. Xiong, Y.Y., P.S. Shabane, and A.V. Onufriev, *Melting Points of OPC and OPC3 Water Models.* Acs Omega, 2020. **5**(39): p. 25087-25094.

30. Wang, J., P. Cieplak, and P.A. Kollman, *How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?* J. Comput. Chem., 2000. **21**(12): p. 1049-1074.

31. Frisch, M.J., et al., *Gaussian 16 Rev. C.01*. 2016: Wallingford, CT.

32. Miller, B.R., et al., *MMPBSA.py: An Efficient Program for End-State Free Energy Calculations.* Journal of Chemical Theory and Computation, 2012. **8**(9): p. 3314-3321.

33. Friesner, R.A., et al., *Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes.* Journal of Medicinal Chemistry, 2006. **49**(21): p. 6177-6196.

34. Halgren, T.A., et al., *Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening.* Journal of Medicinal Chemistry, 2004. **47**(7): p. 1750-1759.

35. Friesner, R.A., et al., *Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy.* Journal of Medicinal Chemistry, 2004. **47**(7): p. 1739-1749.

36. Grabowski, T. and J.J. Jaroszewski, *Bioavailability of veterinary drugs in vivo and in silico.* Journal of Veterinary Pharmacology and Therapeutics, 2009. **32**(3): p. 249-257.

37. Gyawali, R. and S.A. Ibrahim, *Natural products as antimicrobial agents.* Food Control, 2014. **46**: p. 412-429.

38. Pancu, D.F., et al., *Antibiotics: Conventional Therapy and Natural Compounds with Antibacterial Activity-A Pharmaco-Toxicological Screening.* Antibiotics-Basel, 2021. **10**(4).

39. Quinto, E.J., et al., *Food Safety through Natural Antimicrobials.* Antibiotics-Basel, 2019. **8**(4).

40. Stan, D., et al., *Natural Compounds With Antimicrobial and Antiviral Effect and Nanocarriers Used for Their Transportation.* Frontiers in Pharmacology, 2021. **12**.

41. Kawacka, I., et al., *Natural Plant-Derived Chemical Compounds as Listeria monocytogenes Inhibitors In Vitro and in Food Model Systems.* Pathogens, 2021. **10**(1).

42. Chen, J.C., et al., *Identification of Escherichia coli enterotoxin inhibitors from traditional medicinal herbs by in silico, in vitro, and in vivo analyses.* Journal of Ethnopharmacology, 2009. **121**(3): p. 372-378.

43. Wang, J., et al., *Antibacterial oleanane-type triterpenoids from pericarps of Akebia trifoliata.* Food Chemistry, 2015. **168**: p. 623-629.

44. Engels, C., A. Schieber, and M.G. Gänzle, *Inhibitory Spectra and Modes of Antimicrobial Action of Gallotannins from Mango Kernels (Mangifera indica L.).* Applied and Environmental Microbiology, 2011. **77**(7): p. 2215-2223.

45. Kurek, A., et al., *The effect of oleanolic and ursolic acids on the hemolytic properties and biofilm formation of Listeria monocytogenes.* Pol. J. Microbiol., 2014. **63**(1): p. 21-25.

46. Wu, J., et al., *Pomegranate peel (Punica granatum L) extract and Chinese gall (Galla chinensis) extract inhibit Vibrio parahaemolyticus and Listeria monocytogenes on cooked shrimp and raw tuna.* Food Control, 2016. **59**: p. 695-699.

47. Penduka, D., et al., *Evaluation of the anti-Listeria potentials of some plant-derived triterpenes.* Annals of Clinical Microbiology and Antimicrobials, 2014. **13**.

48. Suntar, I., F. Labanca, and L. Milella, *Gallotannins in Food*, in *Handbook of Dietary Phytochemicals*. 2021, Springer Singapore: Singapore. p. 1173-1200.

49. Pollier, J. and A. Goossens, *Oleanolic acid.* Phytochemistry, 2012. **77**: p. 10-15.

50. Kurćubić, V.S., et al., *Multitarget phytocomplex: Focus on antibacterial profiles of Grape Pomace and Sambucus ebulus L. lyophilisates against extensively drug-resistant (XDR) bacteria and in vitro antioxidative power.* Antibiotics (Basel), 2024. **13**(10).

Supplementary Materials

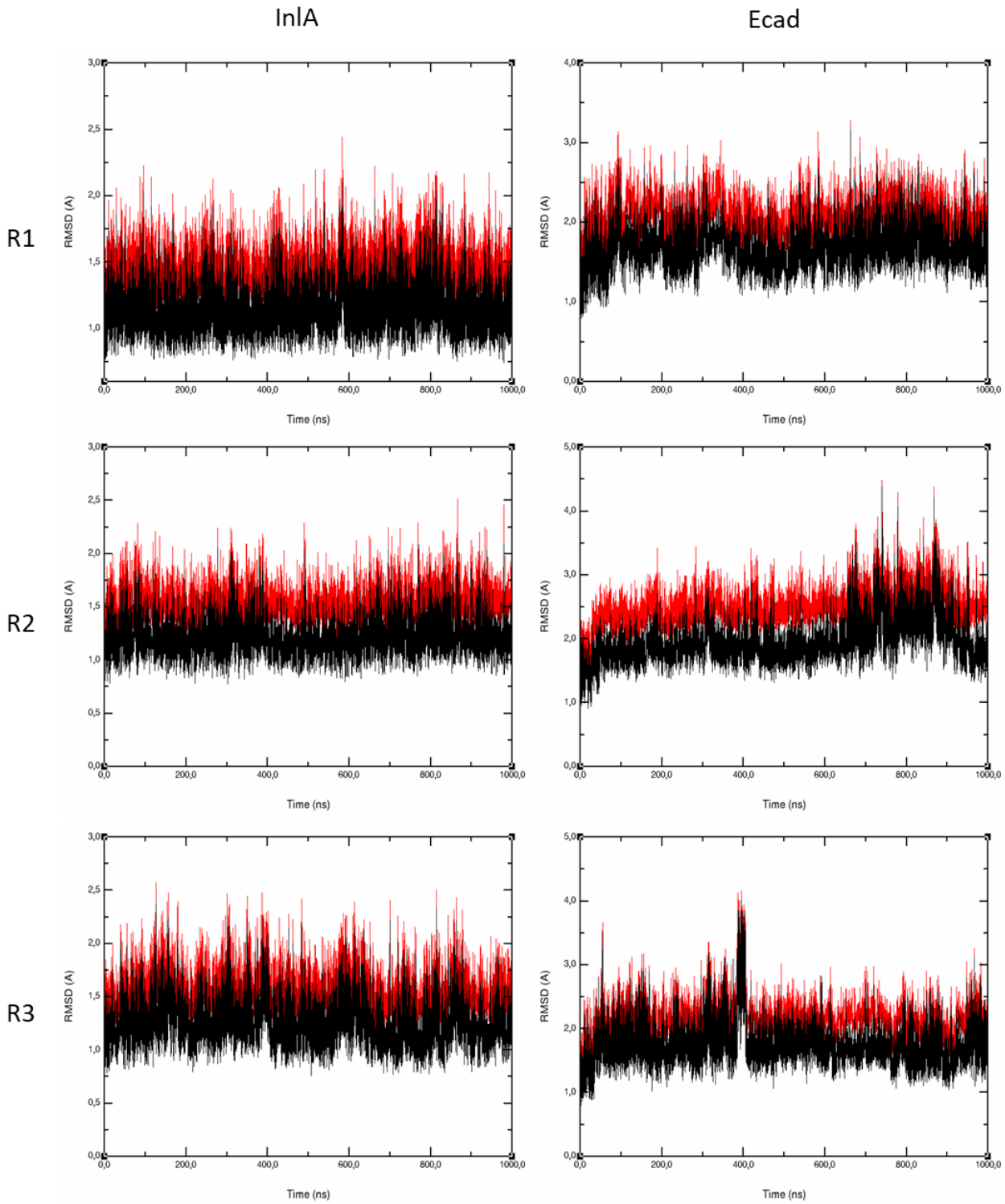


Figure S1. *RMSD of InlA, on the right, and Ecad, on the left, over the three 1 µs long MD simulations replicas. The red line represents the RMSD computed over all the protein's heavy atoms while the black one considering only the backbone.*

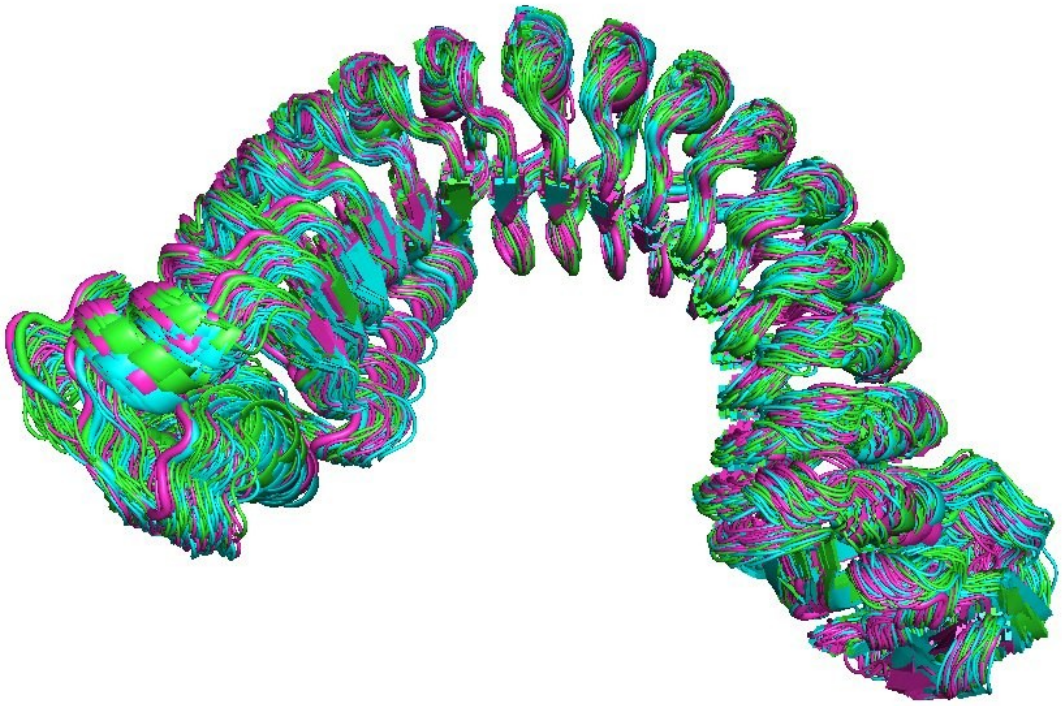


Figure S2. *Collection of 50 snapshots (1 every 20 ns) of the 1 μs long MD simulations replicas of InlA. The protein is represented as a cartoon differently coloured depending on the replica (green, magenta and cyan). It can be noticed how most of the mobility is given by the N- and C-terminal region while the clamp-alike structure over the LRR domain is maintained.*

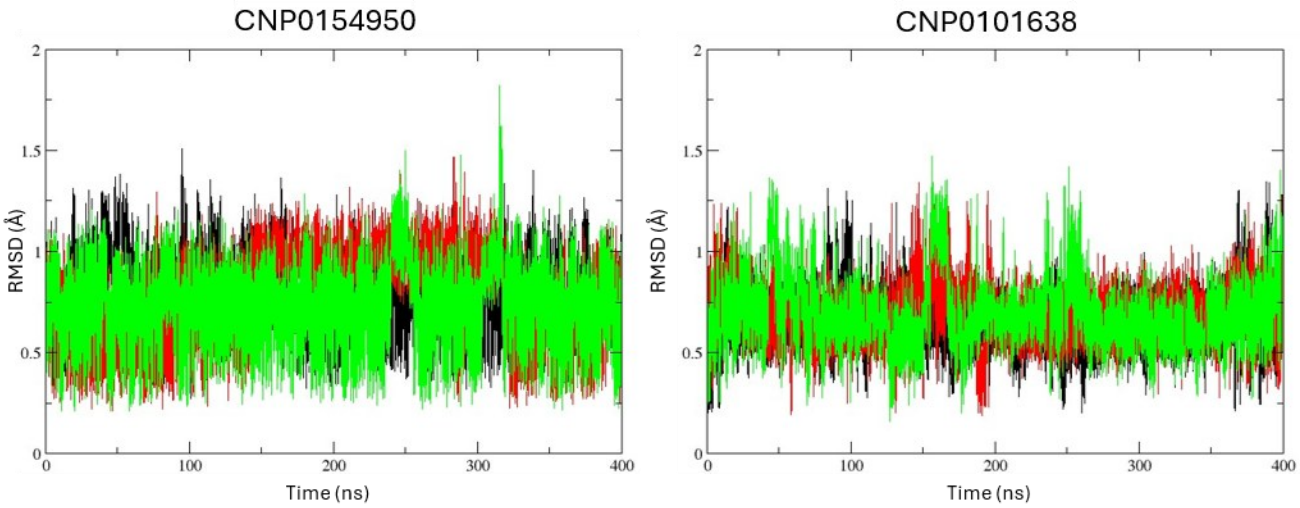


Figure S3. *RMSD of CNP0154950 (on the left) and CNP0101638 (on the right) over the three replicas (red, green and black). It can be seen how the trends are similar in all the three replicas apart from slight readjustments.*

Section III

In silico approaches to pave the way for toxins' bioremediation

Chapter I

Computational methods meet *in vitro* techniques: A case study on fusaric acid and its possible detoxification through cytochrome P450 enzymes

This chapter has been published as:

Lorenzo Pedroni, Daniel Zocchi Doherty, Chiara Dall'Asta, Gianni Galaverna, Stephen G. Bell, Luca Dellafiora. Computational methods meet *in vitro* techniques: A case study on fusaric acid and its possible detoxification through cytochrome P450 enzymes. Ecotoxicol Environ Saf. 2024 Mar 15;273:116167. doi: 10.1016/j.ecoenv.2024.116167.

Abstract

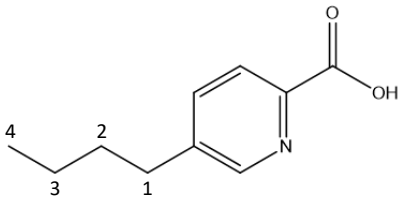
Mycotoxins are known environmental pollutants that may contaminate food and feed chains. Some mycotoxins are regulated in many countries to limit the trading of contaminated and harmful commodities. However, the so-called emerging mycotoxins are poorly understood and need to be investigated further. Fusaric acid is an emerging mycotoxin, noxious to plants and animals, but is known to be less toxic to plants when hydroxylated. The detoxification routes effective in animals have not been elucidated yet. In this context, this study integrated *in silico* and *in vitro* techniques to discover potential bioremediation routes to turn fusaric acid to its less toxic metabolites. The toxicodynamics of these forms in humans have also been addressed. An *in silico* screening process, followed by molecular docking and dynamics studies, identified CYP199A4 from the bacterium *Rhodopseudomonas palustris* HaA2 as a potential fusaric acid biotransforming enzyme. Its activity was confirmed *in vitro.* However, the effect of hydroxylation seemed to have a limited impact on the modelled toxicodynamics against human targets. This study represents a starting point to develop a hybrid *in silico*/*in vitro* pipeline to find bioremediation agents for other food, feed and environmental contaminants.

1. Introduction

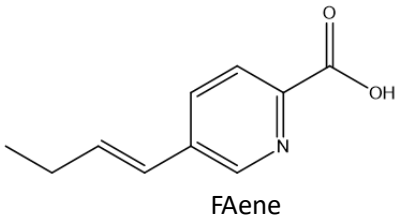
Mycotoxins are toxic secondary metabolites produced by several fungi that can contaminate food, feed and the environment. This can occur worldwide causing a variety of adverse effects to living organisms [1, 2]. Besides a short list of mycotoxins regulated to limit their spread through the food and feed production chain, the majority of those released into the environment and carried through the food chain are not yet regulated. This puts consumers and livestock at risk of exposure to mycotoxin mixtures whose composition and toxicological profile are hard to predict. In particular, the term “emerging mycotoxins” refers to those who are poorly characterized and thus prioritized for study to fill the current data gaps. In this respect, the identification of mitigation strategies to reduce the level of contamination (e.g. turning mycotoxins to detoxified metabolites), is as desirable as their toxicological and exposure profiling. In this context, this work focused on the "emerging mycotoxin" fusaric acid (FA; Figure 1; CID 3442) [3] and applied a hybrid *in silico/in vitro* pipeline to discover novel enzymatic means either for its detoxification for animals or environmental bioremediation.

FA is a low molecular weight (179.22 g/mol) mycotoxin produced by several *Fusarium* species [4]. *F. oxysporum* (affecting tomato), *F. verticillioides* (affecting maize) and *F. fujikuroi* (affecting rice) are known and widespread FA-producing fungi making this mycotoxin relatively common in food and environment, despite this is a poorly checked and yet unregulated food and feed contaminant [5, 6]. Interestingly, FA gained prominence in the seventies due to its reported activity as a hypotensive agent acting as a dopamine-β-hydroxylase (DBH) inhibitor [7]. Nevertheless, this emerging mycotoxin showed both phytotoxicity and cytotoxicity [3, 8, 9]. Its phytotoxic effects were primarily associated with, though not restricted to, the development of tomato wilt symptoms, which may be attributed to its ion chelating capacity, disruption of membrane permeability, and a more general

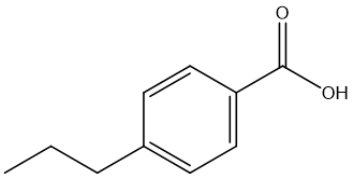
inhibition of plant growth [10]. Regarding its cytotoxicity and toxic activity against animals, it was demonstrated that it might act against mitochondria, impair cell growth *in vitro* and inhibit protein kinase-A (PKA) [9, 11, 12]. Furthermore, FA deserves more attention due to its co-occurrence with other mycotoxins, mainly deoxynivalenol (DON), which might cause a combined response, as described in pigs fed with contaminated corn [11].



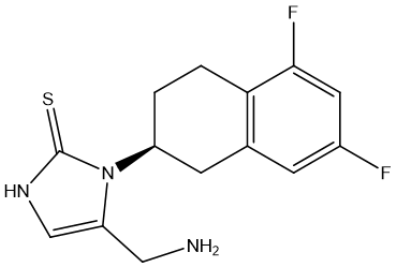
Ligand	C position 1(α)	2(β)	3(γ)	4(δ)
FA	H	H	H	H
FA1	OH	H	H	H
FA2	H	OH	H	H
FA3	H	H	OH	H
FA4	H	H	H	OH



FAene



8ZU



Nepicastat

***Figure 1**. Chemical structures of the analyzed compounds. On the left, FA scaffold (CID 3442) with its chemical substituent in position R1, R2, R3 and R4. On the right, FAene, 8ZU (i.e. 4-n-propylbenzoic acid; CID 137601) and nepicastat (CID 9796181). In bold are represented the metabolites identified by Crutcher and coworkers [14].*

To cope with FA and mycotoxins in general, the use of enzymatic inactivation presents an appealing approach for bioremediating agricultural products and safeguarding crops against the harmful impacts of these fungal byproducts [13]. In this respect, it was recently reported that a fungus belonging to the *Mucor* genus can convert FA to a series of hydroxy-FA derivatives (Figure 1), which are less phytotoxic compounds [14]. Therefore, an *in silico* target fishing approach was employed, as

previously described [15], utilizing a reverse screening methodology targeting the “crystallome” available at the time of analysis for the cytochrome P450 (CYP) family of monooxygenase enzymes (last database access 22nd December 2022) [16]. The choice of focusing the analysis on CYPs relied on the assumption that many members of this class of enzymes can efficiently hydroxylate FA. The *in silico* analysis identified a possible FA biotransforming enzyme from *Rhodopseudomonas palustris* HaA2 [17]. Its activity for FA was confirmed *in vitro*.

Finally, the capability of FA and its metabolites (Figure 1) to target the human enzymes DBH and PKA was studied via molecular docking and dynamics simulations, as these techniques have already been demonstrated to be useful for the investigation of toxicodynamics (TD) of mycotoxins and related metabolites [18]. The importance of the method presented and the relevance of related outcomes to environmental bioremediation and food safety have been also discussed.

2. Materials and Methods

2.1. Computational analysis

2.1.1. Ligands database construction

The identifier for the CYP protein family was retrieved from InterPro (https://www.ebi.ac.uk/interpro/) [19] with ID IPR036396. This was used in the *Advanced Search* tool on PDB (https://www.rcsb.org) [20] to collect all the CYP protein crystal structures and consequently download all the non-redundant ligands co-crystallized within them (658 non-redundant ligands; last database access 22nd December 2022). The 3D structure of FA was retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov; PubChem CID: 3442) [21]. FA and all the 658 ligands were converted from SDF to MOL2 format using Open Babel (v. 3.1.1) [22]. This was done to enable the reverse screening step.

The set of FA metabolites was derived by modifying the structure of FA through UCSF Chimera (version 1.16) [23], while the 3D structure of nepicastat was retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov; PubChem CID: 9796181) [21].

2.1.2. Reverse screening

The ligands database construction enabled target fishing based on reverse screening, as described previously [15]. FA was used as bait to screen the whole ligands database by means of LiSiCA (Ligand Similarity using Clique Algorithm) [24] setting default parameters. This tool allows the quantification of chemical similarities between a reference compound (i.e. FA), and a set of target compounds through the Tanimoto coefficient, the gold standard to measure chemical analogies among chemicals [25]. The PDB entries related to the most FA-alike ligands per LiSiCA were thoroughly checked and considered for subsequent analysis.

2.1.3. Protein model preparation

The 3D model of CYP199A4 from *Rhodopseudomonas palustris* HaA2 was obtained from the crystallographic structure of this enzyme in complex with the ligand 4-n-propylbenzoic acid (PDB ID 7R8S) [26]. The structure is complete (no unresolved parts are present), and the model used for the following analysis was built removing water molecules, ions and the co-crystallized ligand, apart from heme.

The human PKA protein model was derived by humanizing the crystallographic PKA murine structure with PDB code 4DH1 [27], following the procedure reported in previous work [12].

The model of the human DBH was derived from AlphaFold (https://alphafold.ebi.ac.uk/) [28] browsing for the UniProt P09172 code (https://www.uniprot.org/) [29]. This was done since the unique crystallographic structure recorded in the PDB at the time of the analysis (last database access 16th February 2023; PDB ID 4ZEL [30]), contained some unresolved portions preventing its use for molecular dynamics simulations (see below). Once the AlphaFold structure was obtained, the copper was added and optimized within the DBH's binding site through AlphaFill (https://alphafill.eu/) [31] using default settings.

2.1.4. Molecular docking simulation

This procedure provides the most likely architectures of binding for the ligands of interest within the binding site of a given protein model. The docking simulations were performed through GOLD (Genetic Optimization for Ligand Docking; version 2022) [32] using the internal GOLDScore scoring function, as it has a proven high reliability for studying protein–ligand interactions [18]. Specifically, positive scores indicate favored interactions, as per manufacturer declarations (), and the higher the score, the better the fitting into the pocket [18]. For each ligand, ten poses were generated, but considering only the best scored for each ligand for further analysis. A semi-flexible docking protocol

was applied, allowing the protein's polar hydrogens to rotate freely while keeping the ligands fully flexible. The binding site was set as a 10 Å radius sphere around the pocket centroid for each protein model defined as following: for CYP199A4 model the centroid was set at the center of mass of the co-crystallized ligand-binding site of 7R8S PDB structure; for PKA model the centroid was set in agreement with a previous work targeting the interaction between FA and PKA [12]; for DBH model the centroid was set at the copper coordinates, as per Vendelboe and co-workers [30].

2.1.5. Molecular dynamics simulation

Molecular dynamics (MD) simulations were performed using GROMACS (version 2019.4) [33] to verify the potential of FA to be biotransformed by CYP199A4 or the theoretical capability of FA metabolites to inhibit DBH and/or PKA.

The whole system was parametrized with the CHARMM27 all-atom force field [34]. The ligands were parametrized on the SwissParam webserver (https://www.swissparam.ch/) [35]. The system was solvated with SPCE waters in a cubic periodic boundary condition and neutralized adding counter ions, namely Na^+ and Cl^-. Prior to the MD production phase, a series of preparatory steps were carried out for each system. These steps involved the energetic minimization of the system to eliminate steric clashes and rectify any improper geometries. This minimization was achieved by means of a steepest descent algorithm with a maximum of 5000 optimization steps. Subsequently, each system underwent an isothermal simulation conducted at 300 K with a coupling time of 2 ps, and an isobaric simulation maintained at 1 bar with a coupling time of 2 ps. These initial simulations lasted for 100 ps, ensuring that the systems reached equilibrium under the specified conditions. Eventually, the systems were subjected to 40 ns long MD simulations.

2.2. Experimental analysis

2.2.1. Substrate binding assays

Purified CYP199A4 [36] was eluted through a PD-10 column in 50 mM Tris (pH 7.4) buffer to remove glycerol. The protein concentration was quantitated using a Varian CARY 60 spectrophotometer (ε418nm = 119 $mM^{-1}cm^{-1}$) and diluted to approximately 1 µM. Spin-state assays were then carried out at 30 ± 5°C with UV-Vis spectra collected between 250 and 750 nm at a scan rate of 600 nm/min. Spectral data (after 50 mM Tris baseline subtraction) was first recorded for the substrate free CYP199A4 spectrum, and then for 1 µL additions of FA (100 mM in DMSO) until no further spectral perturbations of the heme Soret absorbance bands were detected. The percentage of high-spin ferric heme (~ 397 nm peak), indicative of substrate binding, was estimated by comparison to a calibrated set of spectra of camphor-bound P450cam [37].

2.2.2. *In vitro* NADH activity assay

Purified CYP199A4, and its ferredoxin redox partner HaPux were eluted through a PD-10 size exclusion column into 50 mM Tris (pH 7.4) buffer to remove glycerol. The in vitro reactions were performed in a total volume of 1.2 mL in 50 mM Tris (pH 7.4) buffer, containing CYP199A4 (1 µM), ferredoxin HaPux (10 µM), ferredoxin reductase HaPuR (0.5 µM), and catalase (0.1 mg/mL). The Tris buffer was oxygenated prior to use. After addition of these components, NADH (320 µM) was added and the UV-Vis absorbance of NADH at 340 nm (ε340¬ = 6.22 $mM^{-1}cm^{-1}$) was monitored for several minutes. FA (1 mM) was then added and the rate of NADH oxidation was monitored by the drop in 340 nm absorbance over time. Upon the exhaustion of all NADH, the NADH oxidation rate was calculated (rate of decrease of NADH ($mMmin^{-1}$)) per concentration of P450 enzyme (mM). Reactions were then carried out with increased NADH concentrations (1 mM) and varying concentrations of FA (0.5 to 1 mM) to maximise product yield.

2.2.3. HPLC analysis of metabolites

The reaction mixture (100 µL) was acidified with trifluoroacetic acid (TFA, 0.2%) upon completion. Internal standard (2 µL of 20 mM 9-hydroxyfluorene in EtOH) was added, followed by HPLC-grade acetonitrile (98 µL). Particulate matter was then removed by centrifugation at 13,000 rpm for 3 min at room temperature. The sample was then used directly for HPLC analysis.

Analytical HPLC was performed using a Schimadzu Prominence LC-20AD Pump equipped with an SIL-20A autosampler connected to a Phenomenex Kinetex XB-C18 column (250 × 4.6 mm, 5µm) inside a CTO-20A column oven. Gradients of 20−95% or 0−50% acetonitrile to water (both with 0.1% TFA) at a flowrate of 1 mL/min over 30 min were used. An SPD-20A UV-Vis detector was used with the detection wavelength set to 254 nm.

2.2.4. GC-MS analysis of metabolites

Concentrated HCl (4 µL) and internal standard (5 µL of a 10 mM octanoic acid in EtOH stock) were added to the reaction mixture (1 mL). The metabolites were then extracted into ethyl acetate (1 x 400 µL, 2 x 300 µL), dried with $MgSO_4$ and the solvent was removed under a stream of N_2. The dried extract was subsequently dissolved in anhydrous acetonitrile (150 µL) and derivatised with BSTFA:TMCS (99:1, 15 µL). The mixture was left at 37 °C for 2 hrs and was then directly used for GC-MS analysis.

GC-MS was conducted using a Schimadzu QP2010S GC-MS coupled with an AOC-20S autosampler, an AOC-20i Plus autoinjector and a DB-5 MS fused silica column (30 m x 0.25 mm x 0.25 µm). The carrier gas was helium, and the column flow rate was 1.40 mL/min. The injection temperature and interface temperature were both held at 280 °C. The sample was then injected (1 µL). The column temperature was set to 100 °C and held for 2 min. The temperature was increased at 9°C/min to a final temperature of 280 °C (20 min). The column was then held at 280 °C for 10 min. Products were identified by their mass spectra and relative retention times.

3. Results and Discussion

3.1. Reverse screening

The reverse screening applied in this study has been used previously to discover new ligands for selected proteins [12, 38]. Briefly, the screening is based on the principle that molecules with a sufficient degree of chemical similarity may target the same proteins and be substrates when they bind to the enzymes. As a result, the screening estimates the similarity between FA and the whole set of CYP ligands and substrates stored in the Protein Data Bank (PDB) at the time of analysis to identify CYPs capable of oxidizing FA. To do so, the first step consisted of the construction of a CYP ligand database (658 non-redundant ligands; last database access 22nd December 2022). This was built collecting all the 3D structures of the non-redundant ligands co-crystallized within the available CYP structures in the PDB (1 285 structures, last database access 22nd December 2022; https://www.rcsb.org). Specifically, the CYPs were selected by filtering the PDB as per the InterPro ID IPR036396 (https://www.ebi.ac.uk/interpro/) [19] while the co-crystallized ligands were processed as reported in the *Materials and Methods* (section 2.1.1). Subsequently, in line with previous studies [15], a reverse screening procedure based on target fishing was applied to find possible FA biotransforming enzymes. Indeed, FA was used as bait to screen all the 658 molecules belonging to the previously built ligands database by means of LiSiCA [24], setting default parameters. This tool ranked the set of molecules according to the similarity to FA using the Tanimoto coefficient's metric (from 0 to 1, with 1 meaning identical molecules). All the ligands obtaining a score equal to or higher than 0.7 were analyzed along with the related CYP structure that they are co-crystallized with (Table 1). This arbitrary threshold was set to focus only on the most FA-alike ligands. The top hit obtained was 4-n-propylbenzoic acid, which was co-crystallized with 3 structures of the same CYP, namely CYP199A4 (PDB IDs 7R8S, 5YQA and 7UDF) from

Rhodopseudomonas palustris HaA2 [26, 39, 40]. Interestingly, this CYP is known to catalyze the oxidation of some alkyl-benzoic acids both through hydroxylation and desaturation [26, 36]. It was then considered using molecular modelling analysis to further investigate its capability to bind and transform FA.

***Table 1**. Target fishing based on reverse screening results.*

Ligand ID	Score*	PDB ID
8ZU	0.8	7R8S 7UDF 5YQA
8QM	0.8	5UVB
4IA	0.8	5KDB 6OOX
QZP	0.8	6VJX
IZP	0.7	3VM4
EH1	0.7	6C2D 7TNU
EGM	0.7	4EGM 6OOW 7R8T
EJV	0.7	6C3H
OW7	0.6	6PRS
NCT	0.6	1P2Y 1P7R 4EJG 4EJJ

Note: * Scores computed by LiSiCA [24] based on the Tanimoto coefficient (the closer the score to 1, the greater the chemical similarity between the test and reference compound).

3.2. Defining the reference scenario for protein models

It has been previously demonstrated that the combination of docking simulations and molecular dynamics can estimate whether a compound is efficiently transformed by a given CYP [41, 42]. This

analysis monitored the CYP-ligand's geometry and interactions over time by measuring the distances between the CYP catalytic core, i.e. the heme's Fe, and the region of the molecule which undergoes the oxidation reaction. Initially, the interaction between CYP199A4 and a reference substrate (PDB ID 8ZU; 4-n-propylbenzoic acid) was calculated to retrieve a reference set of data to compare with FA. This would allow us to estimate the capability of the latter to arrange itself into the CYP binding site in a substrate. To do so, the 3D model for CYP199A4 from *Rhodopseudomonas palustris* HaA2 was derived from the crystallographic structure deposited in PDB (7R8S) [26], as detailed in section 2.1.3. The analysis pipelined molecular docking and MD simulations, as reported in previous studies [41, 42]. The first step consisted of molecular docking (see section 2.1.4) of the CYP199A4's reference substrate co-crystallized within the catalytic site, i.e. 4-n-propylbenzoic acid from the PDB ID 7R8S (8ZU; Figure 1).

***Table 2**. Molecular docking results.*

Protein	Ligand	Score
CYP199A4	FA	59.43
	8ZU	55.32
DBH	FA	60.81
	FA1 [a]	61.43
	FA2 [b]	65.68
	FA3 [b]	62.02
	FA4 [b]	64.81
	FAene [a]	60.03
	Nepicastat	52.99
PKA	FA	42.15
	FA1 [a]	42.64
	FA2 [b]	48.76
	FA3 [b]	43.12
	FA4 [b]	44.15
	FAene [a]	43.05

Note: [a] identified in this work; [b] identified by Crutcher and co-workers [14].

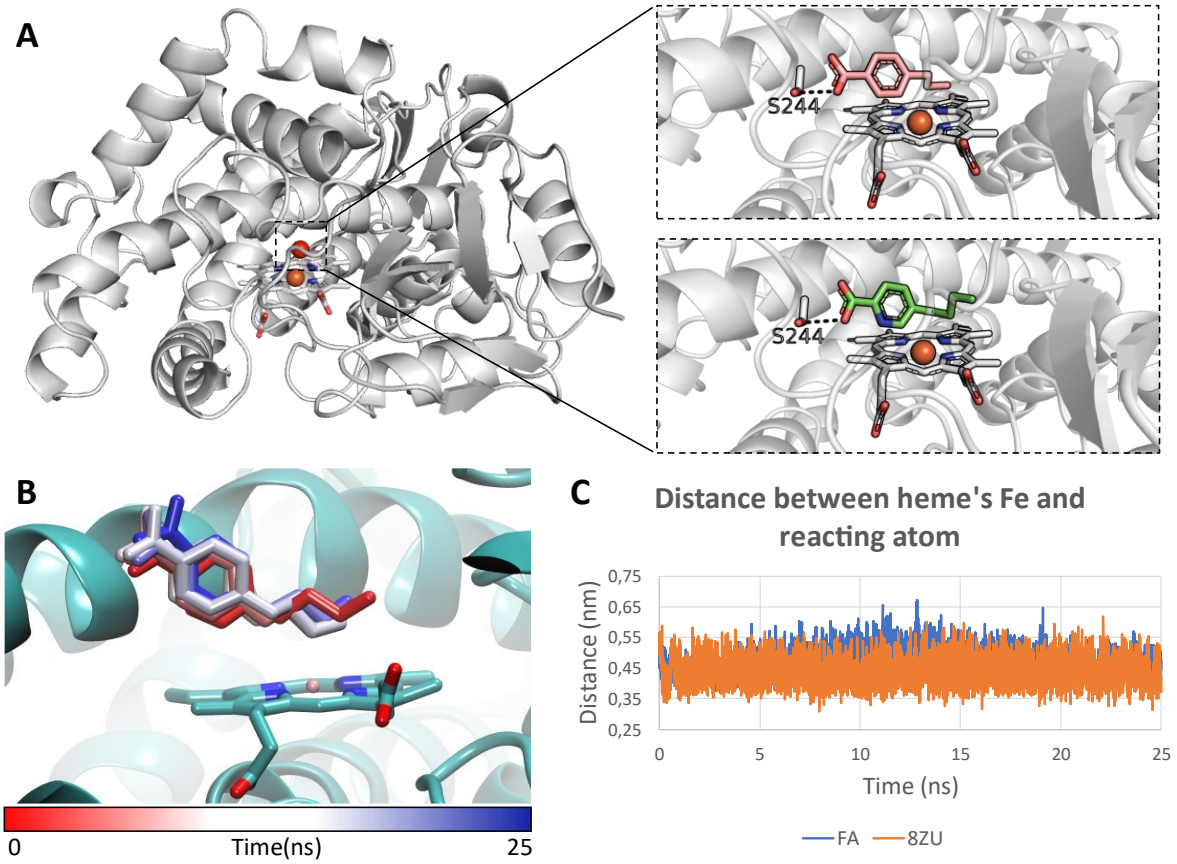


***Figure 2**. Molecular modelling results of 8ZU and FA. **A**. On the left, the 3D model of CYP199A4 from Rhodopseudomonas palustris HaA2 (PDB ID 7R8S). Protein is represented by white cartoon, heme by white sticks and the centroid as a red sphere. On the right, focus on the binding site with the 244Ser represented by white sticks. On top, 8ZU is represented by pink sticks and the hydrogen bond with 244Ser is represented by dashed lines. On bottom, FA is represented by green sticks and the hydrogen bond with 244Ser is represented by dashed lines. **B**. Trajectory of FA within CYP's binding site. The protein is represented by cyan cartoon while FA by sticks. The red-to-blue transition reports the stepwise changes of coordinates over the 25 ns MD simulation. **C**. Interatomic distance between the heme's Fe and putative FA reacting atom (blue; average value 0.406 ± 0.041) and 8ZU reacting atom (orange; average value 0.435 ± 0.047).*

The binding pose obtained a positive score (55.32 units; Table 2), indicating a favorable interaction (the higher the score, the better the interaction with the pocket), with a geometry of interaction almost identical to the crystallographic pose of the 8ZU ligand. This pointed to the reliability of the method to provide reliable CYP199A4-ligand binding architectures for FA-alike compounds. Starting from this docking pose, a 25 ns long MD simulation was run to analyze the geometrical evolution of

the CYP199A4-8ZU complex. Specifically, the interatomic distance between the region of 8ZU known to undergo the reaction and the heme Fe was monitored along the whole MD simulation. The average distance between the region undergoing the reaction and the heme's Fe of less than 0.53 nm indicate that a biotransformation is likely to occur [41, 42]. As shown in Figure 2, the Cα of the aliphatic chain of 8ZU was kept stably below such distance threshold (0.406 ± 0.041 nm) pointing to the reliability of the CYP199A4 model and providing a reference scenario to estimate the substrate likeliness of FA.

3.3. *In silico* analysis of FA binding to CYP199A4

Once the reliability of the CYP199A4 model to predict possible substrates was ascertained (see section 3.1.2), the interaction with FA was calculated to estimate its likelihood of being a good substrate. Notably, it was recently proved that FA hydroxylation might lead to less phytotoxic derivatives [14]. Therefore, the identification of enzymes able to perform such biotransformation might be used and optimized for environmental/food bioremediation strategies to mitigate the presence of FA in the environment and/or through the food chain, as demonstrated for other contaminants [43, 44].

The binding architecture of FA within the CYP199A4 catalytic site was obtained via docking simulations (Figure 2). The pose obtained a positive score, in line with that recorded by the reference compound 8ZU (59.43 and 55.32 units, respectively; Table 2), also showing a comparable geometry of binding (section 3.2 and Figure 2A). This could suggest a proper arrangement at the CYP binding site to undergo biotransformation. Subsequently, a 25 ns long MD simulation was run to monitor the CYP199A4-FA stability over time, recording a steady-state interaction of FA within the CYP catalytic site (Figure 2B). In particular, the distance between the heme's Fe and the portion of the aliphatic chain of FA made of Cα and Cβ was monitored as deemed the most probable to undergo

the reaction. The FA's alkyl chain was kept substantially close to the heme's Fe during the whole simulation, where the average distance of FA's Cα-Cβ portion to heme's Fe was 0.476 ± 0.040 nm. This distance was comparable to the average distances kept by the atom undergoing the reaction of the reference substrate 8ZU (0.435 ± 0.047 nm) and below the 0.53 nm threshold previously identified (8ZU; Figure 2C). This pointed to the potential ability of CYP199A4 from *Rhodopseudomonas palustris* HaA2 to biotransform (namely, hydroxylate) FA. Based on the distances observed, the reaction should take place preferentially at Cα and/or Cβ. The actual capability of CYP199A4 from *Rhodopseudomonas palustris* HaA2 to biotransform FA was then assessed *in vitro.*

3.4. FA binding to and biotransformation by CYP199A4

We first assessed if FA could bind within the CYP199A4 active site using UV-Vis absorbance spectroscopy. Addition of FA to CYP199A4 resulted in an almost complete shift in the Soret band of CYP199A4 from 419 nm to 397 nm (Figure 3A), indicative of a change in the coordination of the heme from a 6-coordinate low-spin ferric species to 5-coordinate high-spin one. The spectral change is similar to those reported for 4-n-propyl- and 4-n-butyl-benzoic acids with CYP199A4 [26]. This demonstrates that FA can bind to CYP199A4 and that it binds close enough to the heme to displace the 6th distal aqua ligand. This suggests that FA is likely to be bound in a good position relative to the heme to enable oxidation by the enzyme.

Next, we assessed the catalytic activity of the CYP199A4 enzyme for FA by conducting *in vitro* oxidation reactions with the fully reconstituted CYP199A4 system from *Rhodopseudomonas palustris* HaA2 (using the electron transfer partners HaPur and HaPux) [26]. The rate of NADH oxidation in the fully reconstituted system in presence of FA was high (1180 ± 40 min^{-1}) approaching

that observed with 4-methoxybenzoic acid and exceeding that of the alkyl substituted benzoic acids (Figure 4) [26].

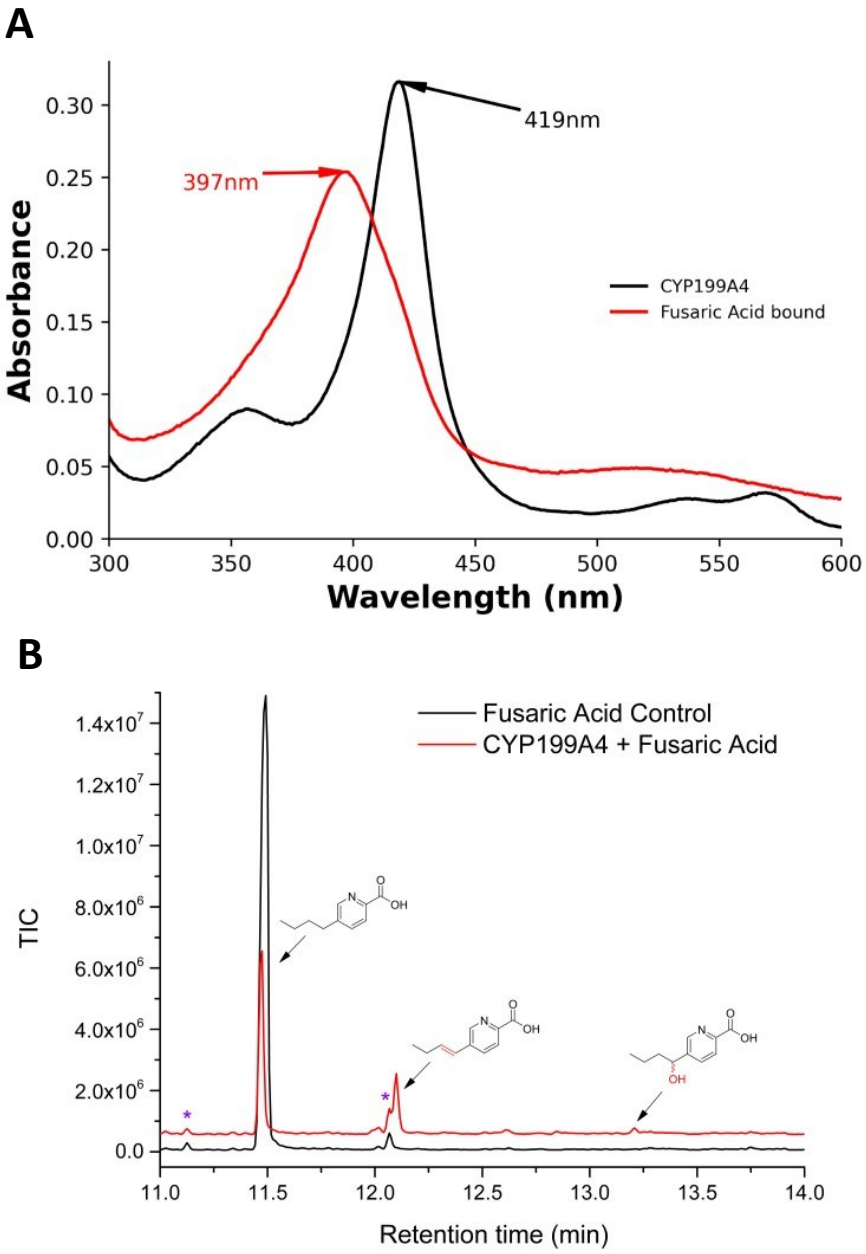


Figure 3. *In vitro analysis results.* ***A****. UV-vis spectra demonstrating the spin-state shift of CYP199A4 from low spin (419 nm) to high spin (397 nm) upon the addition of FA (0.33 mM). This shift is indicative of displacement of the heme iron's axial water ligand, a feature characteristic of successful binding of many CYP substrates.* ***B****. GC-MS chromatogram of the oxidation of fusaric acid (0.5 mM) using an in vitro mixture of CYP199A4 (1 µM) and its redox partner proteins HaPux (10 µM) and HaPuR (1 µM), along with NADH (1 mM) as the source of electrons. The oxidation is shown in red and the no CYP199A4 control in black. The alkene product was characterized by MS.*

Both HPLC and GC-MS analysis of the reaction mixture indicated that a proportion of the substrate had been consumed and that product metabolites had been formed (Figure 5). When the amount of substrate and NADH was the same or an excess of NADH was used (double) a significant amount of the substrate remained. This demonstrates that not all the reducing equivalents from the nicotinamide cofactor are being used to generate oxidized FA metabolites and that uncoupling reactions must be occurring [26].

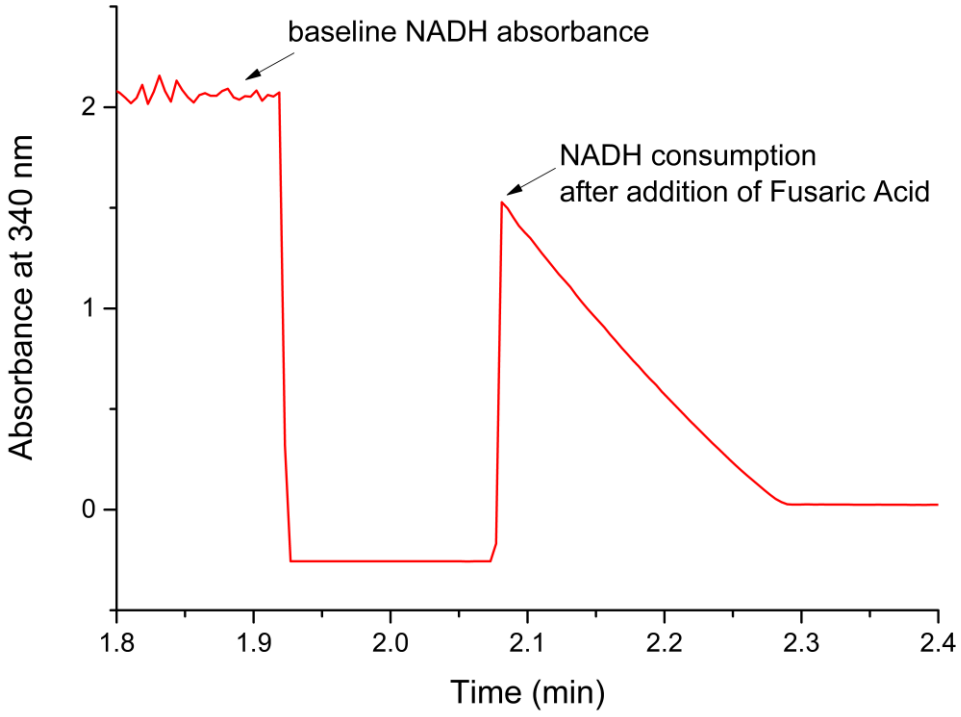


Figure 4. *UV-Vis time-course (red) of the NADH consumption resulting from the reaction of CYP199A4 (1 µM) and its redox partner proteins HaPux (10 µM) and HaPuR (1 µM) with fusaric acid (1 mM). The initial NADH concentration was 320 µM.*

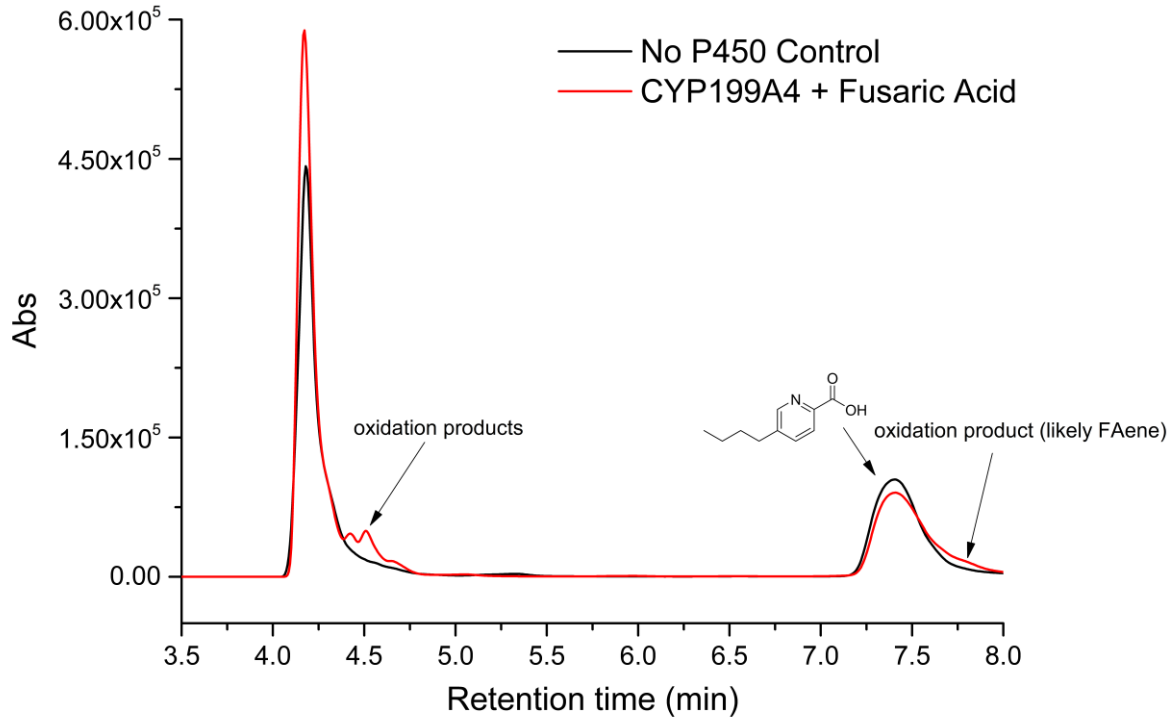
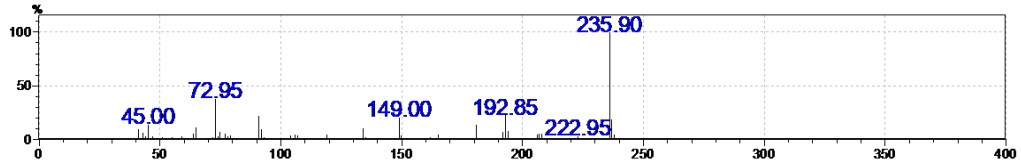


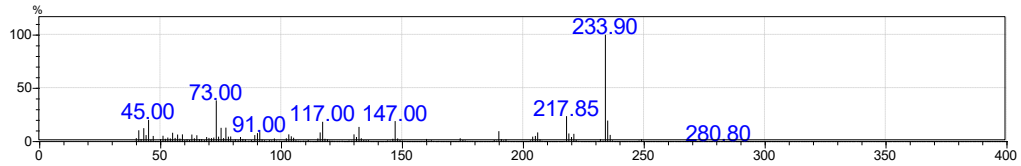
Figure 5. *Reverse-phase HPLC chromatogram (20-95% AcCN) of the oxidation of fusaric acid (1 mM) using an in vitro mixture of CYP199A4 (1 µM) and its redox partner proteins HaPux (10 µM) and HaPuR (1 µM), along with NADH (320 µM) as the source of electrons. The oxidation is shown in red and the no P450 control in black.*

Two of these metabolites were identified using GC-MS from the MS fragmentation patterns of the BSTFA/TMS derivatized metabolites (Figure 3B and Figure 6). One of these arose from hydroxylation at the benzylic Cα carbon of the n-butyl alkyl side chain (4-(1-hydroxybutyl)picolinic acid). The other was from a desaturation reaction to generate an alkene which based on the identity of the hydroxylation metabolite and previous work with this enzyme we propose has the double bond between Cα and Cβ (4-(3-hydroxylbut-1-en-1-yl)-picolinic acid; Figure 3B) [26]. Importantly, the identity of the observed metabolites is consistent with those that would be expected based on the molecular docking experiments in section 3.2 and previous experiments with 4-n-butylbenzoic acid - the one described undergoing the reaction [26]. This is a strong insight supporting the reliability of the model.

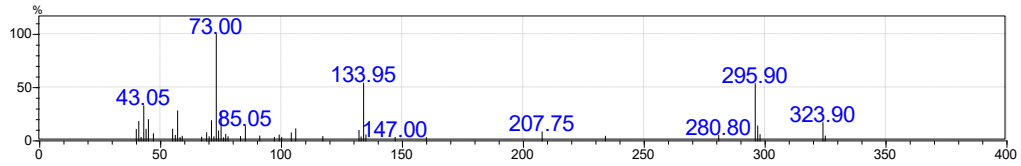
Fusaric acid: RT: 11.45 – 11.52 min (expected TMS-derivatised parent ion = 251 m/z, most abundant = 236 m/z (-15))

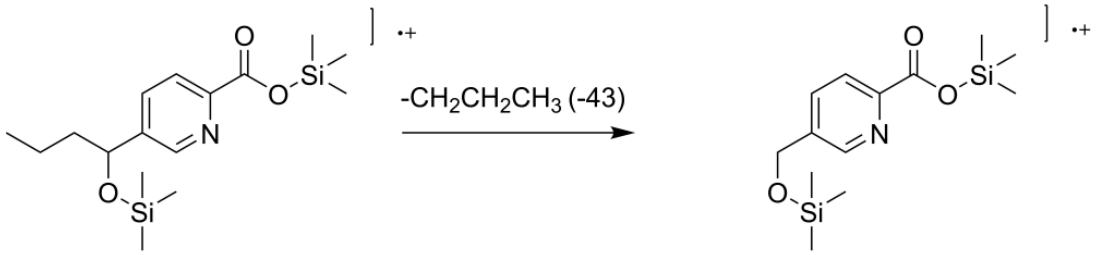


5-but-1-en-yl-pyridine-2-carboxylic acid major product: RT: 12.08 – 12.13 min (expected TMS-derivatised parent ion = 249 m/z, most abundant = 234 m/z (-15))



5-butan-1-ol-pyridine-2-carboxylic acid minor product RT: 13.19 – 13.23 min. (expected TMS-derivatised parent ion = 339 m/z, most abundant = 296 m/z (-43))





Doubly derivatised FA1 product
expected m/z = 339

fragmented doubly derivatised FA1 product
expected m/z = 296

Figure 6. *Top: Mass spectra of trimethylsilyl-derivatised fusaric acid and reaction products, separated by GC. Bottom: Rationalisation of the 5-butan-1-ol-pyridine-2-carboxylic acid product mass spectrum, with the expected m/z of the parent ion (339 m/z) fragmenting to a more stable derivative through loss of the propyl side-chain(m/z = 296).*

3.5. Analysis of FA and derivatives interaction with PKA and DBH

Concerning the fit-for-purpose validation of DBH model, the interaction with the inhibitor nepicastat was taken as reference to estimate the capability of FA hydroxylated derivatives to inhibit the enzyme. In this respect, the arrangement of DBH inhibitors, including nepicastat and FA, at the

enzyme binding site was described in previous studies, which also reported interferences with the copper at the catalytic site among the underpinning mechanisms of inhibition [45-47].

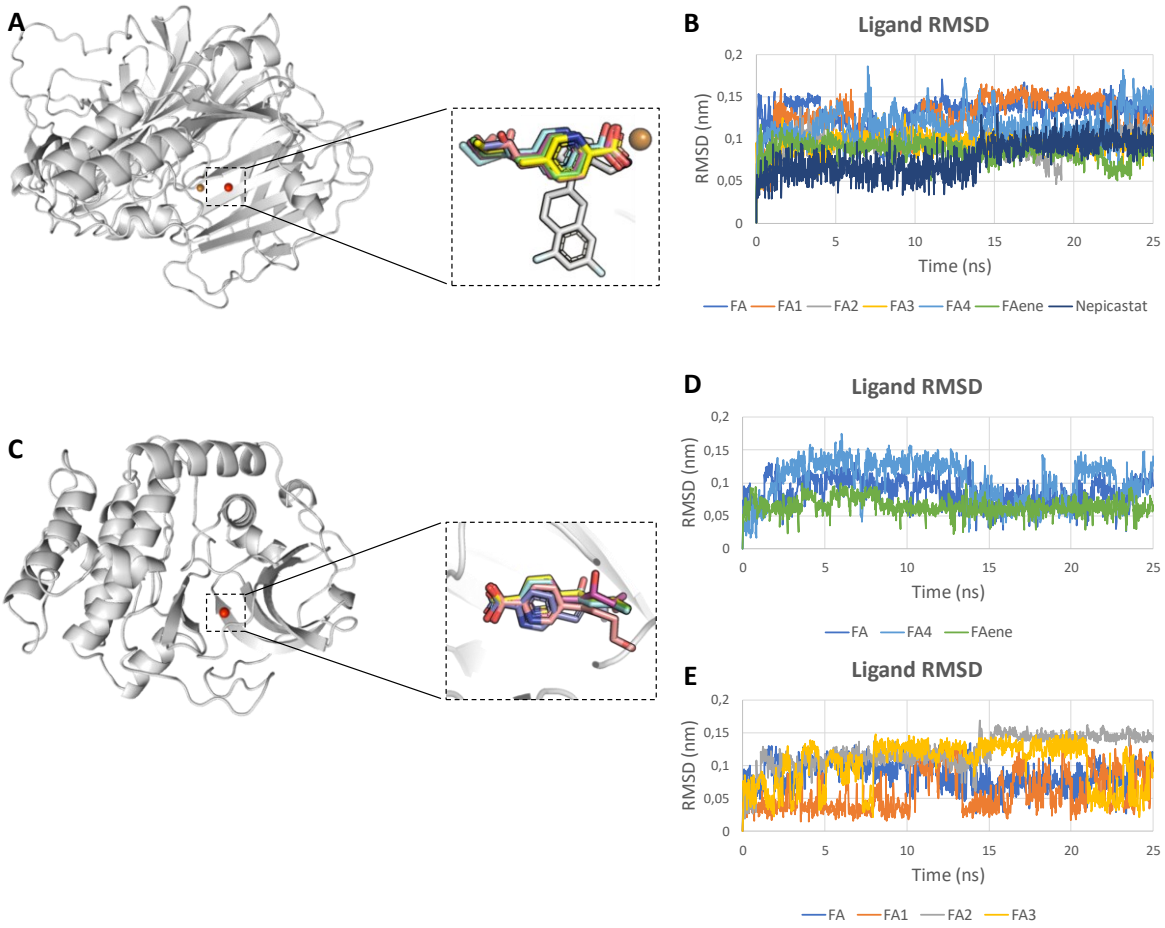


***Figure 7**. Molecular modelling results of DBH and PKA. **A**. DBH 3D model represented by white cartoon with copper represented by a brown sphere and the centroid represented by a red sphere. On the right, close-up of the best docking pose for each tested ligand represented by sticks. **B**. Protein C-alpha RMSD of DBH in complex with nepicastat, FA and all FA putative metabolites. **C**. PKA 3D model represented by white cartoon and the centroid represented by a red sphere. On the right, close-up of the best docking pose for each tested ligand represented by sticks. **D**. Protein C-alpha RMSD of PKA in complex with FA and all FA putative metabolites.*

As shown in Figure 7, nepicastat and FA kept a stable interaction with the copper ion and showed a similar RMSD trend pointing to their comparable steady interaction at the DBH active site. This could validate the model and provide a reference scenario to compare the FA metabolites. Of note, docking and MD simulations were performed either with or without the copper atom at the DBH binding

site. Interestingly, neither of the two ligands could stably interact at the catalytic site without the copper ion (data not shown) pointing to its role to stabilize the enzyme-inhibitor interaction, in agreement with previous evidence [45-47].

The human PKA model, has already been validated in a previous study [12] and the previously published workflow of analysis was applied here to estimate whether the FA metabolites might have an altered capability to interact with PKA compared to their parental compound.

As for the CYP199A4 model, once defined the reference scenario for PKA and DBH models, the capability of FA metabolites (FA1, FA2, FA3, FA4, FAene; Figure 1) to interact with both protein targets PKA and DBH was calculated to investigate whether those modifications would alter the TD of FA. In this respect, it has been previously demonstrated that geometrical criteria such as the ligand RMSD could provide a sound rational to distinguish good inhibitors from non-inhibitor molecules [48]. Specifically, a stable RMSD trend over time indicates that the ligand is likely to have strong interactions at the substrate binding site and competitively inhibit the enzyme. Conversely, variable RMSD trends may indicate an unstable interaction, which is associated with a limited or null capability to inhibit the enzyme. Of note, FA2, FA3 and FA4 were included in the list as their formation was previously demonstrated by Crutcher and coworkers [14], while the formation of FA1 and FAene was observed in the present work. First, the docking pose of nearly all the considered metabolites on both PKA and DBH scored higher than FA (Table 2), possibly suggesting a better fit into the enzyme pockets. This was expected since, apart from FAene, the other metabolites had an additional hydroxyl group able to form additional hydrogen bonds theoretically strengthening the ligand-pocket interaction. As a matter of fact, the only metabolite which obtained a comparable score was FAene when docked to DBH (60.81 and 60.03 units for FA and FAene, respectively; Table 2). The MD simulation of DBH with the FA metabolites showed an interaction trend in terms of RMSD quite comparable to FA (Figure 7), pointing to their possible capability to stably stay at the enzyme binding

site. This could point to a certain capability to inhibit the enzyme. The MD simulations of PKA with FA4 and FAene demonstrated the most stable interaction well retracing the RMSD trend of FA. Conversely, FA1, FA2 and FA3 showed a less stable interaction which might suggest a weaker capability to interact and perhaps inhibit PKA compared to FA. Taken together these results may suggest that biotransformation of FA by means of CYP199A4 may have a relatively limited impact on its TD, although it provided a promising mean for environmental bioremediation. Nonetheless, FA metabolites need further dedicated investigations, specifically targeting the toxicokinetics (TK) in living organisms, given that hydroxylation/unsaturation may prelude FA conjugations and subsequent faster clearance as a part of an efficient detoxifying metabolic route.

4. Conclusion

FA is relatively widespread though overlooked food, feed and environmental contaminant. It may threaten humans and animals for its toxicity, the mechanisms of which include PKA and DBH inhibition, and plants due to its phytotoxicity. Nevertheless, recent evidence has demonstrated that its hydroxylation can substantially reduce its phytotoxicity, suggesting a promising route for bioremediation [13, 14]. In this context, this study successfully discovered a CYP enzyme, namely CYP199A4 from *Rhodopseudomonas palustris* HaA2, as a promising enzymatic mean to catalyze the hydroxylating reaction. Therefore, this enzyme might be considered for developing future bioremediation strategies. Moreover, two FA metabolites, FA1 and FAene, were identified.

Furthermore, the effects of FA hydroxylation and unsaturation on the capability to interact with PKA and DBH were calculated. The 3D modelling studies revealed that all the FA metabolites maintained the ability to interact with DBH like FA. With PKA, FA4 and FAene showed an interaction comparable to that of FA, while the interaction of FA1, FA2 and FA3 seemed less stable. This could suggest that the hydroxylation/unsaturation of FA may have a limited impact on the capability to interact with DBH, while the interaction with PKA seemed more sensitive to the chemical modifications of FA. Nonetheless, besides TD, the TK of FA in living organisms needs further investigation to evaluate whether the FA metabolites analyzed here may prelude conjugation and excretion. If so, the use of CYP199A4 from *Rhodopseudomonas palustris* HaA2 might be considered also for the future development of mitigation strategies through the food and feed production chain, in case its usage proves to be food grade. In summary, a hybrid *in silico*/*in vitro* and highly reproducible pipeline was successfully developed and applied to find possible bioremediation agents to cope with the emerging mycotoxin FA. This same workflow could be applied to find new bioremediation agents for other mycotoxins, or more generally food, feed, and environmental contaminants.

References

1. Hussein, H.S. and J.M. Brasel, *Toxicity, metabolism, and impact of mycotoxins on humans and animals.* Toxicology, 2001. **167**(2): p. 101-134.

2. EFSA, *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, 2014. **12**(12): p. 3916.

3. Gruber-Dorninger, C., et al., *Emerging Mycotoxins: Beyond Traditionally Determined Food Contaminants.* Journal of Agricultural and Food Chemistry, 2017. **65**(33): p. 7052-7070.

4. Munkvold, G.P., *Fusarium Species and Their Associated Mycotoxins*, in *Mycotoxigenic Fungi: Methods and Protocols*, A. Moretti and A. Susca, Editors. 2017. p. 51-106.

5. López-Díaz, C., et al., *Fusaric acid contributes to virulence of Fusarium oxysporum on plant and mammalian hosts.* Molecular Plant Pathology, 2018. **19**(2): p. 440-453.

6. Brown, D.W., et al., *Identification of a 12-Gene Fusaric Acid Biosynthetic Gene Cluster in Fusarium Species Through Comparative and Functional Genomics.* Molecular Plant-Microbe Interactions, 2015. **28**(3): p. 319-332.

7. Hidaka, H., et al., *Fusaric acid, a hypotensive agent produced by fungi.* J Antibiot (Tokyo), 1969. **22**(5): p. 228-30.

8. E Selim, M. and N.A. El-Gammal, *Role of fusaric acid mycotoxin in pathogensis process of tomato wilt disease caused by Fusarium oxysporum.* J. Bioprocess. Biotech., 2015. **5**(10).

9. Mamur, S., et al., *Evaluation of the cytotoxic and genotoxic effects of mycotoxin fusaric acid.* Drug and Chemical Toxicology, 2020. **43**(2): p. 149-157.

10. Ismaiel, A.A. and J. Papenbrock, *Mycotoxins: Producing Fungi and Mechanisms of Phytotoxicity.* Agriculture-Basel, 2015. **5**(3): p. 492-537.

11. Hooft, J.M. and D.P. Bureau, *Deoxynivalenol: Mechanisms of action and its effects on various terrestrial and aquatic species.* Food and Chemical Toxicology, 2021. **157**.

12. Del Favero, G., et al., *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, 2022. **159**.

13. Karlovsky, P., *Biological detoxification of fungal toxins and its use in plant breeding, feed and food production.* Natural Toxins, 1999. **7**(1): p. 1-23.

14. Crutcher, F.K., et al., *Detoxification of Fusaric Acid by the Soil Microbe Mucor rouxii.* Journal of Agricultural and Food Chemistry, 2017. **65**(24): p. 4989-4992.

15. Dellafiora, L., et al., *An In Silico Target Fishing Approach to Identify Novel Ochratoxin A Hydrolyzing Enzyme.* Toxins, 2020. **12**(4).

16. Poulos, T.L., *Heme Enzyme Structure and Function.* Chemical Reviews, 2014. **114**(7): p. 3919-3962.

17. Coleman, T., et al., *Cytochrome P450 CYP199A4 from Rhodopseudomonas palustris Catalyzes Heteroatom Dealkylations, Sulfoxidation, and Amide and Cyclic Hemiacetal Formation.* Acs Catalysis, 2018. **8**(7): p. 5915-5927.

18. Pedroni, L., et al., *The bitter side of toxicity: A big data analysis spotted the interaction between trichothecenes and bitter receptors.* Food Research International, 2023. **173**.

19. Paysan-Lafosse, T., et al., *InterPro in 2022.* Nucleic Acids Research, 2023. **51**(D1): p. D418-D427.

20. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Research, 2000. **28**(1): p. 235-242.

21. Kim, S., et al., *PubChem 2023 update.* Nucleic Acids Research, 2023. **51**(D1): p. D1373-D1380.

22. O'Boyle, N.M., et al., *Open Babel: An open chemical toolbox.* Journal of Cheminformatics, 2011. **3**.

23. Pettersen, E.F., et al., *UCSF chimera - A visualization system for exploratory research and analysis.* Journal of Computational Chemistry, 2004. **25**(13): p. 1605-1612.

24. Legnik, S., et al., *LiSiCA: A Software for Ligand-Based Virtual Screening and Its Application for the Discovery of Butyrylcholinesterase Inhibitors.* Journal of Chemical Information and Modeling, 2015. **55**(8): p. 1521-1528.

25. Dellafiora, L., et al., *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, 2022. **162**.

26. Coleman, T., et al., *Exploring the Factors which Result in Cytochrome P450 Catalyzed Desaturation Versus Hydroxylation.* Chemistry-an Asian Journal, 2022. **17**(24).

27. Kovalevsky, A.Y., et al., *Low- and room-temperature X-ray structures of protein kinase A ternary complexes shed new light on its activity.* Acta Crystallographica Section D-Biological Crystallography, 2012. **68**: p. 854-860.

28. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-+.

29. Bateman, A., et al., *UniProt: the Universal Protein Knowledgebase in 2023.* Nucleic Acids Research, 2023. **51**(D1): p. D523-D531.

30. Vendelboe, T.V., et al., *The crystal structure of human dopamine β-hydroxylase at 2.9 Å resolution.* Science Advances, 2016. **2**(4).

31. Hekkelman, M.L., et al., *AlphaFill: enriching AlphaFold models with ligands and cofactors.* Nature Methods, 2023. **20**(2): p. 205-+.

32. Jones, G., et al., *Development and validation of a genetic algorithm for flexible docking.* Journal of Molecular Biology, 1997. **267**(3): p. 727-748.

33. Abraham, M.J., et al., *GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers.* SoftwareX, 2015. **1-2**: p. 19-25.

34. Best, R.B., et al., *Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone φ, ψ and Side-Chain χ1 and χ2 Dihedral Angles.* Journal of Chemical Theory and Computation, 2012. **8**(9): p. 3257-3273.

35. Bugnon, M., et al., *SwissParam 2023: A Modern Web-Based Tool for Efficient Small Molecule Parametrization.* Journal of Chemical Information and Modeling, 2023. **63**(21): p. 6469-6475.

36. Podgorski, M.N., et al., *Investigation of the requirements for efficient and selective cytochrome P450 monooxygenase catalysis across different reactions.* Journal of Inorganic Biochemistry, 2020. **203**.

37. Maddigan, N.K. and S.G. Bell, *The self-sufficient CYP102 family enzyme, Krac9955, from Ktedonobacter racemifer DSM44963 acts as an alkyl- and alkyloxy-benzoic acid hydroxylase.* Archives of Biochemistry and Biophysics, 2017. **615**: p. 15-21.

38. Pedroni, L., et al., *An in silico insight on the mechanistic aspects of gelsenicine toxicity: A reverse screening study pointing to the possible involvement of acetylcholine binding receptor.* Toxicology Letters, 2023. **386**: p. 1-8.

39. Bell, S.G., et al., *Selective oxidative demethylation of veratric acid to vanillic acid by CYP199A4 from Rhodopseudomonas palustris HaA2.* Molecular Biosystems, 2010. **6**(1): p. 206-214.

40. Bell, S.G., et al., *Investigation of the Substrate Range of CYP199A4: Modification of the Partition between Hydroxylation and Desaturation Activities by Substrate and Protein Engineering.* Chemistry-a European Journal, 2012. **18**(52): p. 16677-16688.

41. Dorne, J., et al., *A Computational Understanding of Inter-Individual Variability in CYP2D6 Activity to Investigate the Impact of Missense Mutations on Ochratoxin A Metabolism.* Toxins, 2022. **14**(3).

42. Pedroni, L., et al., *A computational study on the biotransformation of alkenylbenzenes by a selection of CYPs: Reflections on their possible bioactivation.* Toxicology, 2023. **488**.

43. Narayanan, M., S.S. Ali, and M. El-Sheekh, *A comprehensive review on the potential of microbial enzymes in multipollutant bioremediation: Mechanisms, challenges, and future prospects.* Journal of Environmental Management, 2023. **334**.

44. Palacios-Mateo, C., et al., *Enzymes for microplastic-free agricultural soils.* Ecotoxicology and Environmental Safety, 2023. **258**.

45. Zaman, A., *Docking studies and network analyses reveal capacity of compounds from Kandelia rheedii to strengthen cellular immunity by interacting with host proteins during tuberculosis infection.* Bioinformation, 2012. **8**(21): p. 1012-1020.

46. Kapoor, A., M. Shandilya, and S. Kundu, *Structural Insight of Dopamine β-Hydroxylase, a Drug Target for Complex Traits, and Functional Significance of Exonic Single Nucleotide Polymorphisms.* Plos One, 2011. **6**(10).

47. Dey, S.K., et al., *Dopamine β hydroxylase as a potential drug target to combat hypertension.* Expert Opinion on Investigational Drugs, 2020. **29**(9): p. 1043-1057.

48. Weng, Y.L., et al., *Molecular dynamics and in silico mutagenesis on the reversible inhibitor-bound SARS-CoV-2 main protease complexes reveal the role of lateral pocket in enhancing the ligand affinity.* Scientific Reports, 2021. **11**(1): p. 7429.

Section IV

Mechanistic insights into xenobiotics metabolism and bioactivation

Section IV contains five chapters containing mechanistic insights on a series of foodborne xenobiotics, starting from alkenylbenzenes and moving to mycotoxins.

In the First and Second Chapter, the focus is on alkenylbenzenes bioactivation mediated by cytochrome P450 metabolism. Indeed, the hydroxylation mediated by cytochromes is known to produce the proximate carcinogen whose subsequent sulfation would lead to the ultimate one. Here, computational approaches were applied to obtain insights on both the species-specific metabolism of safrole – Chapter I – and on the differences occurring among specific human cytochromes when interacting with some members of this class of compounds – Chapter II.

Starting from the Third Chapter, until the Fifth Chapter, the focus was moved on mycotoxins. Regarding both the Third and the Fifth Chapter, computational approaches were applied to find human proteins that had not previously been documented as target for specific mycotoxins. In the third one, trichothecenes were found to possibly interact with a bitter receptor, namely TAS2R46. In the fifth one, ochratoxin A was investigated for its possible interaction with a key player protein in protein translation, i.e. OGFOD1.

In the Fourth Chapter, the focus was put on zearalenone a mycotoxin known to act as a xenoestrogen and thus exerting endocrine-disrupting effects. A computational investigation allowed to obtain more insights on the interaction between aromatase and this mycotoxin, specifically investigating the influences of genetic variability among individuals.

Chapter I

A Computational Inter-Species Study on Safrole Phase I Metabolism-Dependent Bioactivation: A Mechanistic Insight into the Study of Possible Differences among Species

This chapter has been published as:

Pedroni Lorenzo, Louisse Jochem, Punt Ans, Dorne Jean Lou C.M., Dall'Asta Chiara, Dellafiora Luca. 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). 2023 Jan 18;15(2):94. doi: 10.3390/toxins15020094.

Abstract

Safrole, a 162.2 Da natural compound belonging to the alkenylbenzenes class, is classified as a possible carcinogen to humans by IARC (group IIB) and has proven to be genotoxic and carcinogenic to rodents. Despite its use as a food or feed additive, it is forbidden in many countries due to its documented toxicity; yet, it is still broadly present within food and feed and is particularly abundant in spices, herbs and essential oils. Specifically, safrole may exert its toxicity upon bioactivation to its proximate carcinogen 1'-hydroxy-safrole via specific members of the cytochrome P450 protein family with a certain inter/intra-species variability. To investigate this variability, an in-silico workflow based on molecular modelling, docking and molecular dynamics has been successfully applied. This work highlighted the mechanistic basis underpinning differences among humans, cats, chickens, goats, sheep, dogs, mice, pigs, rats and rabbits. The chosen metric to estimate the likeliness of formation of 1'-hydroxy-safrole by the species-specific cytochrome P450 under investigation allowed for the provision of a knowledge-based ground to rationally design and prioritise further experiments and deepen the current understanding of alkenylbenzenes bioactivation and CYPs mechanics. Both are crucial for a more informed framework of analysis for safrole toxicity.

1. Introduction

Safrole (Figure 1) is a natural low-molecular-weight molecule (162.2 Da) belonging to alkenylbenzenes [1]. Alkenylbenzenes are secondary metabolites of herbs and spices (e.g., basil, fennel and parsley) that are commonly abundant in essential oils, with many members of interest for their potential adverse health effects, including safrole [2]. It has been reported that some alkenylbenzenes are carcinogenic via a genotoxic mode of action, through their highly reactive phase II 1'-sulfooxy metabolites [3]. Specifically, safrole may exert its toxicity upon its conversion to 1'-hydroxy-safrole (a proximate carcinogen) due to the cytochrome P450 (CYP)-dependent hydroxylation formed in humans mainly by CYP2A6 and subsequent sulfation to form 1'sulfooxy-safrole (the ultimate carcinogen). The 1'-sulfooxy metabolite is responsible for the formation of DNA adducts causing the genotoxic insult [4]. Moreover, animal studies have suggested that safrole may cause tumors in liver and other tissues in vivo, and IARC classified safrole as possibly carcinogenic to humans (group IIB) [1,5].

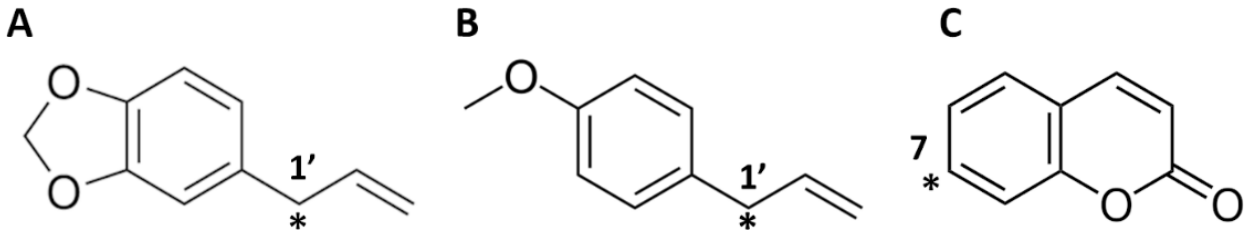


Figure 1. *Chemical structure of molecules under analysis. The asterisk indicates the atom undergoing the reaction. (**a**) Safrole; (**b**) Estragole; (**c**) Coumarin.*

The use of safrole as a food or feed additive is no longer allowed in many countries due to its documented toxicity (e.g., its use has been banned in the US since the 1960s). However, it is still present in food and feed due to its abundance in spices, herbs and essential oils used as ingredients, giving rise to multiple safety issues [6–8]. Safrole has been found in a variety of food products including cola-tasting beverages and Bologna and Vienna sausages to cite but a few, and it can also

be found in plant-based foods, food supplements and feed/food products with safrole-containing essential oils added as ingredients [2,6,9,10].

Safrole toxicity has been proven in animal models, although a certain inter-species variability has been documented. This discrepancy in the data raised the question of whether the genotoxic effects observed in experimental animals are also expectable in humans [11] or other animal species, including farm and companion animals. In line with this interpretation, a certain degree of variability is likely expected among various animal species. As the toxicity of safrole and other alkenylbenzenes depends on its bioactivation of 1'-hydroxy-safrole, inter- and intra-species differences in biotransformation are of interest. In fact, differences in terms of 1'-hydroxy-safrole formation are expected to diversify inter- and intra-species sensitivity to safrole. Indeed, many in vitro studies have been performed with human and rat liver fractions to study the bioactivation and detoxification kinetics of various alkenylbenzenes [12]. Besides these in vitro studies, it would be of interest to obtain further insight into inter- and intra-species differences in bioactivation via molecular modelling studies. Furthermore, there are limited data available on the biotransformation of alkenlylbenzenes in animals other than mice and rats, and the existing in vitro data are only for rats and humans. This indicates a knowledge gap regarding the bioactivation potential of alkenylbenzenes in other species, including farm and companion animals, which may be at risk from consuming feed with safrole-containing ingredients. Nonetheless, such knowledge is important both to ensure better feed safety and animal welfare and to drive a more informed selection of animal models for data extrapolation to humans. However, the systematic study of livestock animals and pets is hampered by the high costs of such studies and related ethical issues. In this sense, the use of new approach methodologies (NAMs) may provide alternative and effective methods of analysis. NAMs include in vitro methods to study the biotransformation of chemicals, but these generally require primary animal material and can be expensive for large-scale analysis due to, for example,

the extensive chemical analytical measurements required. NAMs also include a variety of non-testing approaches, including computational methodologies, with increasing use in the evaluation of chemical safety and, more generally, within the framework of next-generation risk assessments [13]. Regarding in silico biotransformation studies, 3D molecular modelling techniques have already been applied to study the biotransformation of small molecules by CYPs, providing a useful means of distinguishing substrates from non-substrates [14–16]. Specifically, molecular dynamics at a low nanosecond scale (<50 nanoseconds) applying the same method used in this study has already succeeded in predicting the likeliness of small molecules to be substrates of CYPs [15]. Therefore, this previously validated 3D modelling framework based on docking studies and molecular dynamics simulations is applied in the present study to assess the CYP-dependent capability of nine animal species including farm and companion animals (i.e., rat, mouse, dog, cat, rabbit, pig, goat, sheep and chicken) to form the safrole proximate carcinogen 1'-hydroxy-safrole. Such information could provide insights into possible species differences in bioactivation and related susceptibility to safrole-dependent genotoxicity, supplying a blueprint to (i) informedly design further experimental trials; (ii) refine other knowledge-based modelling techniques; and (iii) expand the current understanding of safrole toxicity and bioactivation and CYPs mechanics. More specifically, the present study assessed the dynamics of the interactions of the alkenylbenzenes safrole and estragole with human CYP1A2 and CYP2A6. These CYPs are reported to be involved in the 1'-hydroxylation of various alkenylbenzenes [4] (safrole by CYP2A6, estragole by CYP1A2 and CYP2A6) and coumarin, which is reported to be a substrate that is detoxified (to 7-hydroxycoumarin) by CYP2A6 [17]. Subsequently, molecular dynamics were used to study the interactions between the alkenylbenzenes and homologs of CYP1A2 and CYP2A6 in other species (rat, mouse, dog, cat, rabbit, pig, goat, sheep and chicken), estimating the capability of these species to bioactivate safrole to its proximate carcinogen 1'-hydroxy-safrole.

2. Materials and Methods

2.1. Data Source

The 3D structures of safrole, estragole and coumarin were retrieved in .sdf format from the PubChem database (https://pubchem.ncbi.nlm.nih.gov accessed on the 12th of September 2022) [27] with CID 8815, 5144 and 323, respectively (CAS codes 140-67-0, 94-59-7 and 91-64-5, respectively). The models of human CYP1A2 and CYP2A6 were derived from the crystallographic structures recorded in the Protein Databank (PDB; https://www.rcsb.org accessed on the 20th of September 2022) [28] with PDB codes 2HI4 [29] and 2PG6 [30], respectively. In particular, the latter structure had two mutations, L240C and N297Q, which had reverted to the wild-type sequence, duly replacing the respective amino acid side chain using the Structure Editing/Rotamer tool of UCSF Chimera software (version 1.15) [31], as previously described [32]. The search for sequences of animal homologs to human CYP2A6 was conducted using the blastp tool of Web BLAST (Basic Local Alignment Search Tool) (https://blast.ncbi.nlm.nih.gov/Blast.cgi accessed on the 20th of _September 2022) [21]. Default parameters were used with the sequence of human CYP2A6 as the input and searching in the non-redundant protein sequence database for each species under analysis. The sequence of animal homologs to human CYP1A2 and CYP2A6 were retrieved from the NCBI Entrez database (https://www.ncbi.nlm.nih.gov/search accessed on the 20th of September 2022) [25] and stored in FASTA format for further analysis (see below for further details). The accession numbers are reported in the Supplementary Materials (Table S3). The actual expression in the liver of the considered CYPs was checked by browsing references databases of protein existence (i.e., UniProt; https://www.uniprot.org/ accessed on the 10th of January 2023) and gene transcription and expression (i.e., Expression Atlas (https://www.ebi.ac.uk/gxa/home accessed on the 10th of January 2023) [18] and Bgee (https://bgee.org accessed on the 10th of January 2023) [19]). Specifically, UniProt provides information on protein in terms of protein/transcript existence and whether the

protein has been inferred from homology. Expression Atlas and Bgee provide quantitative expression data at a protein or transcript level, with the latter specifically meant to provide data for intra-experiments quantitative comparisons [19].

2.2. Homology Modelling

The structures of animal homologs of rat, mouse, dog, cat, rabbit, pig, goat, sheep and chicken to human CYP1A2 or CYP2A6 were not available in PDB at the time of analysis (last database accessed 20th September 2022). Therefore, in agreement with previous studies [20,33], the respective models were derived via homology modelling using the human CYP1A2 or CYP2A6 structure as a template with the software Modeller (version 10.2) [34] interfaced in the UCSF Chimera software (version 1.16) [31]. Specifically, default settings were used apart from including non-water HETATM residues from templates to ensure the inclusion of the heme group in each model. The best-scored homolog model for each human CYP according to the ZDope scoring was carried forth to the analysis. Of note, we performed homology modelling via Modeller for continuity of the analysis with our previous study [15]. However, we compared the results obtained with those provided by AlphaFold2 [35], which has provided a significant breakthrough in structural biology. As shown in the Supplementary Materials, the legacy software Modeller and AlphaFold2 performed comparably, confirming the reliability of Modeller outputs. Indeed, the two sets of structures superimposed similarly to the human homolog with low RMSDs (see Supplementary Materials, Tables S4 and S5, and Figure S4). The percentage identity matrix and multiple sequence alignment proved the high identity among sequences and at the level of human residues interacting with safrole, and suggested the usability of human CYP as a reliable template for a canonical homology model procedure (see Figures 4 and S5 and Supplementary Materials). The percentage identity matrices (PIMs; Supporting Materials, Figure S5) were conducted using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/ accessed on the 18th of November 2022) with default parameters [36], while the multiple sequence

alignment was obtained using CLC Sequence Viewer version 7.7 (QIAGEN, Aarhus, Denmark; https://www.qiagenbioinformatics.com).

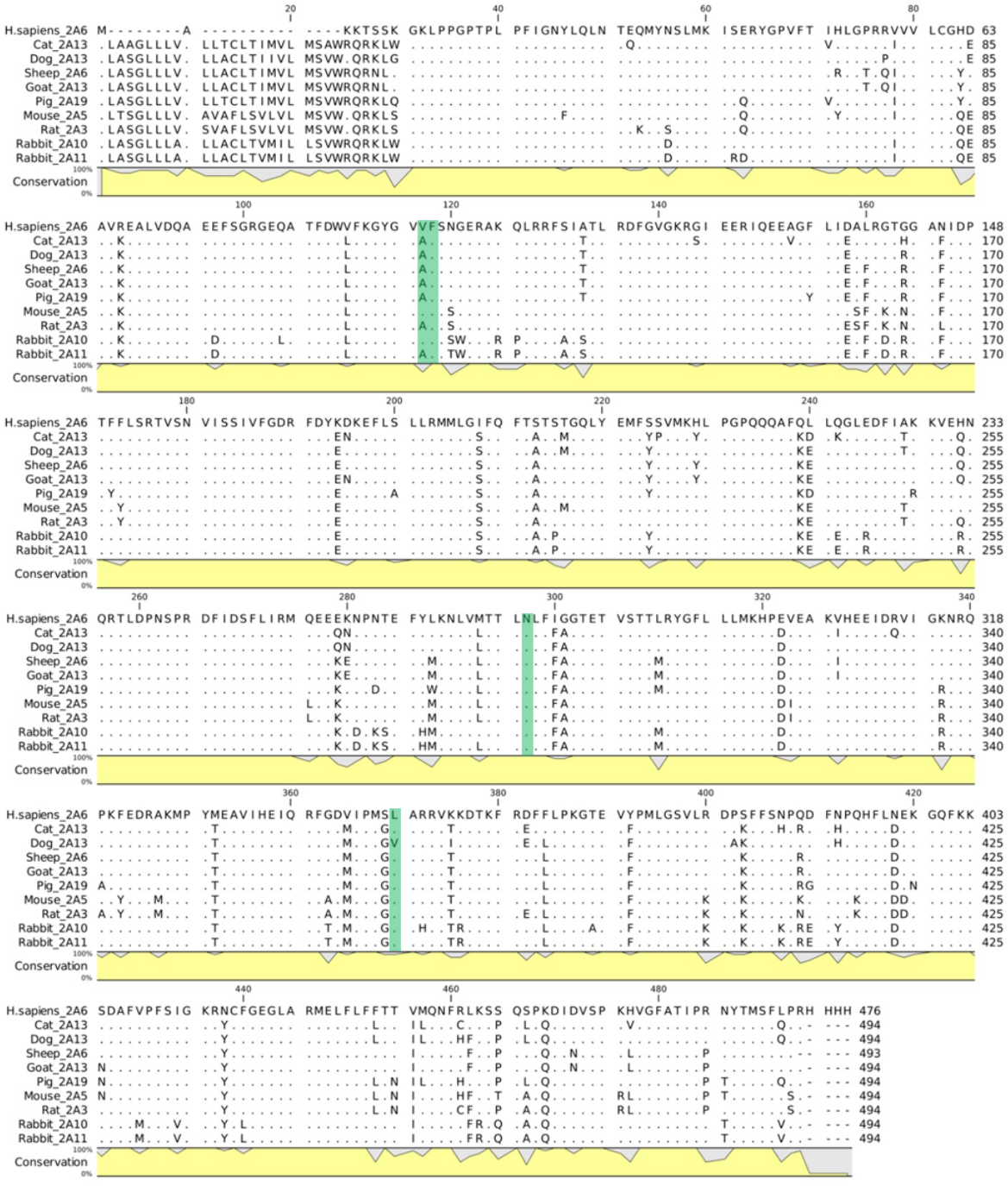


Figure 4. *Multiple sequence alignment of animal homologs to human CYP2A6. Dots represent matching residues while dashes indicate gaps. The conservation line plot is reported in yellow and shows the percentage of conservation of each residue among the species with respect to the human sequence. The residues interacting with safrole in the human complexes are highlighted in green.*

2.3. Docking Simulations

The docking analysis aimed to provide a plausible binding architecture for the molecules under analysis within the catalytic site of CYPs. This was performed with the GOLD software (version 2021), which has already shown high reliability for the evaluation of protein–ligand interactions, including in studies of the interactions with CYPs [15,33,37]. The binding site was defined within a 10-Å radius sphere around the centroid of the substrate binding site. The docking protocol was set according to previous studies, for which ligands were kept fully flexible and proteins semi-flexible, allowing polar hydrogens to rotate freely [15]. As a minor modification, the internal scoring function GOLDScore was used as it has been optimised for the prediction of binding architectures, according to the manufacturer declaration (https://www.ccdc.cam.ac.uk; accessed 15 July 2022). In addition, the binding architecture of styrene—which has a structure closely related to that of estragole and safrole—and coumarin derived from a cytochrome's crystallographic structures (PDB ID 4HGF and 1Z10, respectively) were used as a position restraint (constraint weight of 100 units) to facilitate the arrangement of safrole and estragole or coumarin, respectively. The pictures of binding poses (see Supporting Materials) were obtained through PLIP2021 (https://plip-tool.biotec.tu-dresden.de/ accessed on the 18th of November 2022) [38] and PyMol (version 2.3.0 Open-Source; https://pymol.org).

2.4. Molecular Dynamics

Molecular dynamics aimed to investigate the overall geometrical stability of CYP-ligand complexes over time to discriminate CYP substrates from non-substrate molecules. This was performed using GROMACS (version 2019.4) [39], while ligands were parametrised with the CHARMM27 all-atom force field [40]. The hydrogen database was modified according to previous works [41,42] to

parametrise the heme group. Input structures were solvated with SPCE waters in a cubic periodic boundary condition, and counter ions (Na^+ and Cl^-) were added to neutralise the system. Prior to running simulations, each system was energetically minimised to avoid steric clashes and correct improper geometries using the steepest descent algorithm with a maximum of 5000 steps. Subsequently, each system underwent isothermal (300 K, coupling time 2 psec) and isobaric (1 bar, coupling time 2 psec) 100 psec simulations before running 25 nsec simulations (300 K with a coupling time of 0.1 psec and 1 bar with a coupling time of 2.0 psec).

2.5. Statistical Analysis

The statistical analysis of interatomic distances between the atom undergoing the reaction and the Fe atom of the heme group was conducted using SPSS IBM (v. 27.0, SPSS Inc., Chicago, IL, USA). For each complex, distance values of 5000 frames were considered, expressed as means ± standard deviation (SD) and compared to each other using one-way ANOVA ($\alpha = 0.05$) with Bonferroni as a posthoc test ($\alpha = 0.05$).

2.6. Cluster and Binding Affinity Analysis of Protein-Ligand Complex Trajectories

Each complex underwent a cluster analysis to identify geometries of binding representative of the whole simulation which were used to estimate the binding free energy. This was conducted using GROMACS (version 2019.4) [39] through the cluster command, using the gromos method and setting the cutoff at 0.3 nm, in agreement with previous studies [43]. The binding affinity was obtained by analysing the retrieved geometry of binding for each complex via the PRODIGY web application (https://wenmr.science.uu.nl/prodigy/lig accessed on the 18th of November 2022) [26].

3. Results and Discussion

In a previous study, we demonstrated that computing the distance between the atom undergoing the reaction of small molecules and the Fe atom of the heme group over time can be used as a reliable geometrical criterion to distinguish CYP substrates from non-substrate molecules [15]. The same workflow has been applied here to study the substrate likeliness of safrole for CYP1A2 and CYP2A6 in a set of animal species including rats, mice, dogs, cats, rabbits, pigs, goats, sheep and chickens. For this purpose, the primary sequence of each CYP under analysis was downloaded in the FASTA format to generate 3D models via homology modelling (see Materials and Methods and Supplementary Materials for further details). Before the analysis, the actual expression of the considered CYPs in the liver of animals under analysis has been checked by browsing reference databases of protein existence (i.e., UniProt; https://www.uniprot.org accessed on the 10th of January 2023) and gene transcription and expression (i.e., Expression Atlas, https://www.ebi.ac.uk/gxa/home [18] accessed on the 10th of January 2023; Bgee, https://bgee.org [19] accessed on the 10th of January 2023). For all of them, at least one piece of evidence for existence and expression in the liver, which is the metabolizing organ relevant to safrole toxicity, has been found (Table S1; Supporting Material).

Once the actual hepatic expression of CYPs under analysis had been ascertained, the binding architecture of ligands was calculated via docking simulations and the geometry of binding was analysed over time with molecular dynamics. The binding poses produced by docking analysis describing the interactions of safrole with amino acids forming the binding site and its orientation to the Fe heme are reported in Supplementary Materials (Figures S1–S3). Specifically, safrole showed similar binding poses in all the calculated complexes with CYP2A6 or CYP1A2 homologs, where the atom undergoing the reaction was directed to the heme Fe. The contribution of

hydrophobic/hydrophobic interactions was found to be predominant in all the safrole-CYP complexes, as shown in Figures S1–S3 (Supporting Materials). The formation of hydrogen bonds was observed in certain complexes (i.e., human, cat, pig, rabbit, chicken and rat homologs of CYP1A2 and in human, cat, pig and sheep homologs of CYP2A6). However, the hydrophobic/hydrophobic interactions were expected to be crucial based on the marked hydrophobicity of safrole (LogP 3.45, as per PubChem database; PubChem CID 5144, last accessed 24 November 2022). Although safrole was found to be comparable among the various complexes generated by docking, the arrangement of the atom undergoing the reaction with the heme Fe was found to vary over time in certain complexes. Specifically, the complexes where the distance to the heme Fe increased over a certain threshold were predicted not likely to lead to 1'-hydroxy-safrole formation (see below).

3.1. Fit-for-Purpose Validation

Once the CYP models had been derived, a fit-for-purpose validation was carried out to assure the applicability to CYP1A2 and CYP2A6 of the procedure and related threshold that previously succeeded with CYP2D6 [15]. The validation relied on data retrieved from the literature describing the capability of safrole to be biotransformed to 1'hydroxy-safrole by human CYP2A6, while its conversion by human CYP1A2 was found to be negligible [4]. The safrole congener estragole (Figure 1) was also included, being a substrate of human CYP1A2 and CYP2A6 [4], along with coumarin (Figure 1) as an additional positive control. In fact, coumarin is a substrate of human CYP2A6 and rat CYP2A3 (the proximate homolog of human CYP2A6) [17]. Therefore, the interaction of safrole and estragole was computed in human CYP1A2 and CYP2A6, while the interaction of coumarin was computed in human CYP2A6 and rat CYP2A3 to challenge the geometrical criteria and threshold previously described to infer the substrate likeliness of ligands [15]. In more detail, in agreement with a previous study [15], the distance between the atom undergoing the reaction and the heme's Fe can be monitored to qualitatively distinguish CYP substrates from non-substrate molecules. Based

on previous evidence [15], an average interatomic distance of 0.53 nm was set as the threshold and molecules whose atoms undergoing the reaction show a wider distance could be considered to not be efficiently biotransformed at that specific position.

Therefore, the distance between the heme's Fe and the atom undergoing the reaction of estragole in CYP1A2 and CYP2A6, that of safrole in CYP2A6, and that of coumarin in CYP2A6 and rat CYP2A3 were expected to be lower compared to that of the safrole-CYP1A2 complex, for which a lack of biotransformation has been demonstrated [4]. The initial architectures of binding were computed via docking simulations. As shown in Table 1, docking scores were all positive and relatively high, suggesting the theoretical capability of the molecules under analysis to appreciably interact with the CYP's catalytic site, in agreement with previous studies (the higher the score, the better the fitting into the model's pocket). Therefore, substrates could not be distinguished from non-substrate molecules based on docking scores alone. However, we previously showed that dynamic simulations should be combined with docking studies to better estimate the likeliness of CYPs substrates [15]. As shown in Figure 2, both the interatomic distance analysis and the inspection of trajectories over time supported this assumption as the atom undergoing the reaction was kept close to the heme's Fe only in complexes with an appreciable biotransformation capacity that had been previously demonstrated in vitro.

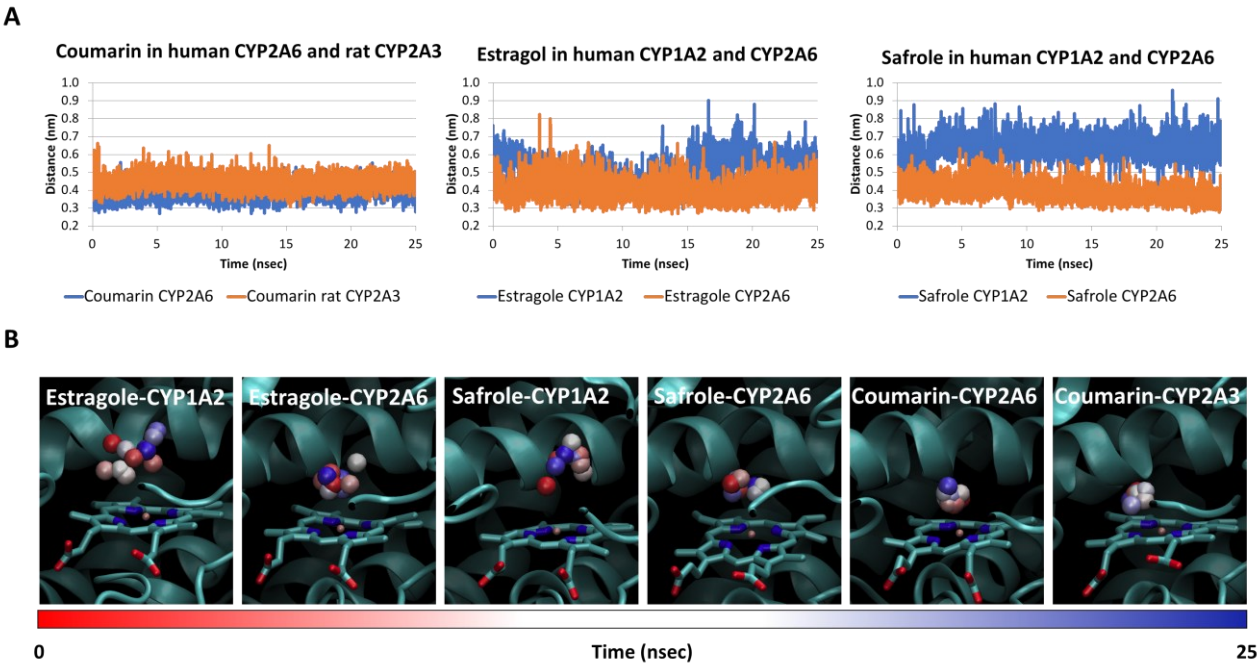


Figure 2. *Molecular dynamics results of estragole and safrole in complexes with human CYP1A2 and CYP2D6 and coumarin in complexes with human CYP2A6 or rat CYP2A3. (**A**) Analysis of the distance between heme's Fe and the atom undergoing the reaction and of coumarin in human CYP2A6 and rat CYP2A3, estragole in human CYP1A2 and CYP2A6, and safrole in human CYP1A2 and CY2A6. (**B**) Trajectory analysis of the complexes mentioned above. The proteins are represented by cartoons, while heme is represented by sticks. The atom undergoing the reaction is represented by spheres and a red-to-blue colour change indicates the stepwise changes of coordinates over time.*

Table 1. *Docking scores of estragole, coumarin and safrole within human CYP1A2 and CYP2A6, and their animal homologs*

Molecule	Species	CYP2A6 and its homologues	CYP1A2
Estragole	Human	162	154
Coumarin	Rat	144	n.p.
	Human	146	n.p.
Safrole	Human	162	159
	Cat	174	153
	Chicken	n.p.	150
	Dog	164	152
	Goat	173	156
	Mouse	157	171
	Pig	172	155
	Rat	167	155
	Sheep	153	159
	Rabbit	174 169	161

Note: "n.p." stands for "not performed".

Jeurissen and co-workers reported that estragole is a better substrate for human CYP2A6 than for CYP1A2, with Kcat/Km values of 341 and 59, respectively [4]. As shown in Figure 2A, the interatomic distance between the atom undergoing the reaction of estragole and the heme's Fe was similar for CYP1A2 and CYP2A6 up to 15 nsec while moving away, with a significantly larger mean distance for CYP1A2 compared to CYP2A6 (0.51 ± 0.09 and 0.40 ± 0.06 nm, respectively; $p < 0.001$). Considering that the proximity between the atom undergoing the reaction and heme's Fe is fundamental to ensure the reaction (see above), the transformation by CYP1A2 was expected to be less effective compared to CYP2A6, in agreement with in vitro experimental results. Additionally, the analysis of trajectories (Figure 2B) revealed a spread of the atom undergoing the reaction that was wider when

estragole was in a complex with CYP1A2 than when it was in a complex with CYP2A6. In line with this interpretation, safrole was described as a substrate of CYP2A6 (with Kcat/Km of 160) but not of CYP1A2 (no biotransformation), as reported by Jeurissen et al. [4]. Both the computed distances and trajectories are in line with this evidence as the interatomic distances were significantly different in complexes with CYP1A2 or CYP2A6 (0.63 ± 0.06 and 0.40 ± 0.04 nm, respectively; $p < 0.001$). Furthermore, the analysis of trajectories showed a wider spread of the atom undergoing the reaction when in a complex with CYP1A2 than when in a complex with CYP2A6 (Figure 2B). As an additional control, coumarin was also included in the analysis and its interaction was calculated with human CYP2A6 and rat CYP2A3, which showed a comparable in vitro reaction yield, as previously reported [17]. As shown in Figure 2, the interatomic distance (0.38 ± 0.04 and 0.44 ± 0.04 nm with CYP2A6 and rat CYP2A3, respectively) and the spread of the atom undergoing the reaction to the heme's Fe in the two coumarin-CYP complexes were similar. Taken together, these results further supported the reliability of the modelling framework to predict whether chemicals are substrates or not for human CYP1A2 and CYP2A6 and animal homologues. The analysis of distances over time confirmed the reliability of the mean distance threshold that was previously identified and set at 0.53 nm. Of note, these data might suggest that a (semi)quantitative analysis is possible considering that the computed distances of estragole were consistent with the rank of the two CYPs in terms of biotransformation yield (the better the CYP to transform estragole, the shorter the computed interatomic distance and the lower the spread of the atom undergoing the reaction within the catalytic site). However, a wide set of positive and negative controls structurally related to the molecules under investigation are needed to properly validate the model for quantitative purposes, which could not be found at the time of analysis. Therefore, the model was meant to provide qualitative predictions.

3.2. Interaction of Safrole with Animal Homologs of Human CYP1A2 and CYP2A6

The 3D models of animal homologs of human CYP1A2 and CYP2A6 were derived via homology modelling in agreement with a previous study [20] and carried forth to docking analysis and molecular dynamic simulations to investigate the interaction with safrole over time and estimate its substrate likeliness to form 1'-hydroxy-safrole [15].

As per CYP2A6, only sheep had a cytochrome entry referred to as CYP2A6. For the other species, the proximate homolog to human CYP2A6, i.e., the animal's CYP with the highest identity percentage to human CYP2A6, was searched for in the respective proteome to identify the CYP theoretically responsible for the safrole biotransformation in the animal species under investigation. Otherwise stated, it was assumed that the CYP that was most likely responsible for the potential biotransformation of safrole in each animal was the one that was most similar in terms of % identity to human CYP2A6. The search was conducted using the blastp tool of Web BLAST (Basic Local Alignment Search Tool) (https://blast.ncbi.nlm.nih.gov/Blast.cgi accessed on the 20th of September 2022; see Materials and Methods for further details) [21]. As shown in Table S2 (Supplementary Materials), according to the blastp algorithm, at least one CYP with a very high percentage of sequence identity (above 80%) to human CYP2A6 was found in each species. As an exception, in the chicken proteome, the CYP most similar to the human homolog showed less than 50% of identity and was therefore deemed too divergent to preserve significant functional conservation in terms of its capability to form 1'-hydroxy-safrole. Therefore, it was not further analysed. Once the accession numbers of all the proximate homologs had been identified, their primary sequence was collected in FASTA format and used to derive the respective 3D model via homology modelling (see Materials and Methods for further details). For rabbits, two homologs (CYP2A10 and CYP2A11) were carried forth to the analysis as they showed the same identity percentage to human CYP2A6.

Each safrole-CYP complex showed a high docking score, suggesting the theoretical capability of safrole to favourably interact with all the animal homologs under analysis. Molecular dynamics simulations were then analysed to verify this assumption. As shown in Figure 3 and Table 2, the interaction of safrole with the rat and mouse homologs showed a dynamic of interaction similar to that found for the human CYP2A6, although in both species, the distance between the atom undergoing the reaction and the heme's Fe was larger compared to the human case. This could suggest that safrole might be less efficiently biotransformed to the 1'-hydroxymetabolite by rat and mouse homologs compared to human CYP2A6.

Table 2. *Interatomic distance between the atom undergoing the reaction of estragole, coumarin or safrole and the heme's Fe atom.*

Molecule	Species	CYP2A6 and its Homologues		CYP1A2	
		Distance (nm)	**Biotransformation [1]**	**Distance (nm)**	**Biotransformation [1]**
Estragole	Human	0.40 ± 0.06	Yes (experimental) [2]	0.51 ± 0.09	Yes (experimental) [2]
Coumarin	Rat	0.45 ± 0.04	Yes (experimental) [3]	n.p.	-
	Human	0.39 ± 0.04	Yes (experimental) [3]	n.p.	-
Safrole	Human	0.39 ± 0.04	Yes (experimental) [2]	0.64 ± 0.09	No (experimental) [2, 4]
	Cat	0.74 ± 0.08	No (calculated)	0.52 ± 0.07	Yes (calculated)
	Dog	0.63 ± 0.14	No (calculated)	0.47 ± 0.06	Yes (calculated)
	Pig	0.65 ± 0.15	No (calculated)	0.43 ± 0.07	Yes (calculated)
	Goat	0.43 ± 0.07	Yes (calculated)	0.47 ± 0.06	Yes (calculated)
	Rabbit	0.72 ± 0.09	No (calculated)	0.47 ± 0.06	Yes (calculated)
		0.47 ± 0.08	Yes (calculated)		
	Chicken	n.p.	n.p.	0.47 ± 0.09	Yes (calculated)
	Sheep	0.43 ± 0.06	Yes (calculated)	0.44 ± 0.06	Yes (calculated)
	Mouse	0.50 ± 0.09	Yes (calculated)	0.48 ± 0.06	Yes (calculated)
	Rat	0.52 ± 0.08	Yes (calculated)	0.47 ± 0.06	Yes (calculated)

Note: Distances are expressed as mean values ± standard deviation. "n.p." stands for "not performed". [1] Refers to the capability to form 1'-hydroxy-metabolite of safrole and estragole or 7-hydroxy-coumarin from coumarin; [2] experimental evidence according to [4]; [3] experimental evidence according to [17]; [4] experimental evidence according to [22].

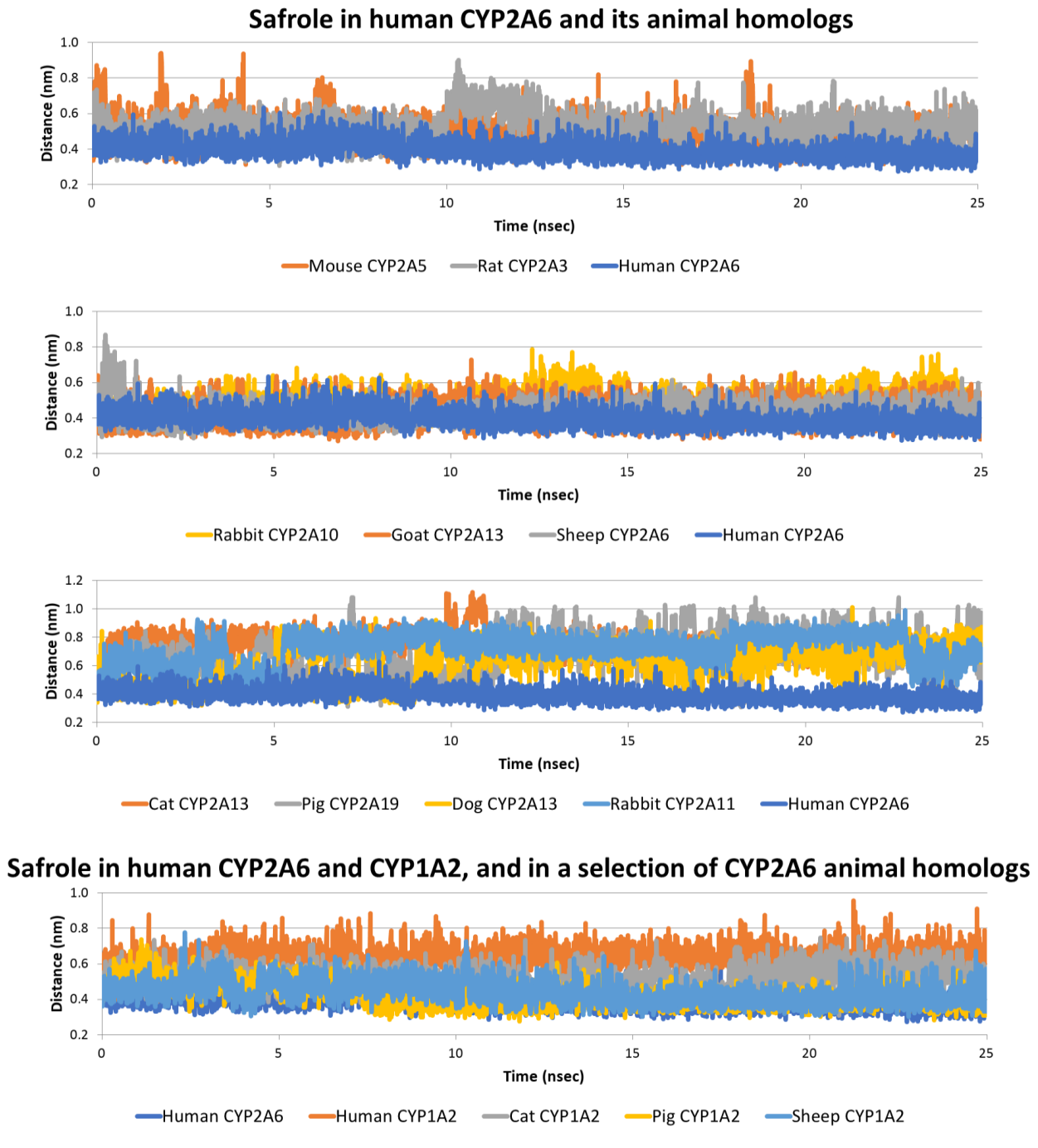


Figure 3. *Molecular dynamics results of safrole in complexes with human CYP1A2 and CYP2A6 and a selection of CYP1A2 and CYP2A6 animal homologs.*

This result is in line with previous in vitro evidence describing the lower activity of rat hepatic microsomes compared to human liver microsomes to form 1'-hydroxy-safrole [12,23] and may propose a plausible mechanistic explanation (which could be also expected for mice). However, a diverse level of expression of human CYP2A6 and rat CYP2A3 in liver tissues could have a role in

diversifying the formation of 1'-hydroxy-safrole in the two species as the hepatic expression of CYP2A3 in rats has been documented but not quantitatively compared to that of humans [24]. Nonetheless, the agreement between the computational results and the evidence previously collected (see above), along with the data collected for coumarin (see above; Section 2.1), further support the reliability of the procedure and suggest that these in silico analyses can provide important insights for other animal homologs. As shown in Figure 3 and Table 2, safrole in sheep CYP2A6, goat CYP2A13 and rabbit 2A10 showed dynamics of interactions, with interatomic distances between its atoms undergoing the reaction and the heme's Fe, which were similar to that observed for human CYP2A6. This suggested that sheep CYP2A6, goat CYP2A13 and rabbit CYP2A10 may appreciably form 1'-hydroxy-safrole. It must be noted that, although the in silico analysis suggested safrole bioactivation, data for protein expression, which is, of course, required for bioactivation to occur, are still fragmentary and require further dedicated investigations. Conversely, the CYP2A6 homologs identified in cats, dogs and pigs and the 2nd homolog identified in rabbits (CYP2A11) showed a dynamic of interaction not in line with an appreciable formation of 1'-hydroxy-safrole. Indeed, the interatomic distances between the atom undergoing the reaction and the heme's Fe were similar to that observed for the human safrole-CYP1A2 complex (i.e., not showing biotransformation in vitro [4]) and above the 0.53 nm threshold. Although the existence of other CYP2A6 homologs able to form 1'hydroxy-safrole cannot be excluded completely, our results suggest that for cats, dogs and pigs, a much less efficient formation of the proximate carcinogen can be expected, which may possibly translate to lower degrees of safrole toxicity. This cannot be concluded from the present in silico analysis as a CYP2A6 homolog-independent biotransformation may exist, but this study provides clear directions for further dedicated (in vitro) investigations.

Concerning the animal homologs of human CYP1A2, all the species under investigation (i.e., rat, mouse, dog, cat, rabbit, pig, goat, sheep and chicken) had in the respective proteome recorded in

the NCBI Entrez database (https://www.ncbi.nlm.nih.gov/search accessed on the 20th of September 2022) [25] a cytochrome annotated as CYP1A2 (Table S2; Supplementary Materials). Each sequence was downloaded in the FASTA format and used to derive a 3D model via homology modelling (see Materials and Methods for further details). As shown in Table 1, all the docking scores were again relatively high, suggesting the theoretical capability of safrole to interact with all of them. Each complex was then analysed via molecular dynamics to monitor the geometrical stability and respective orientation of the atom undergoing the reaction to the heme's Fe. Surprisingly, as shown in Table 2 and Figure 3, all the animal homologs showed a better arrangement of safrole compared to the complexes with human CYP1A2 (i.e., the atom undergoing the reaction was closer to the heme's Fe compared to safrole within human CYP1A2 and below the threshold of 0.53 nm). This may suggest that a certain degree of safrole biotransformation to 1'hydroxy-safrole could be expected for all the species analysed, even if the respective CYP2A6 homolog is inactive or with low biotransformation activity. It is of interest that the cat homolog showed the widest mean interatomic distance and an interatomic distance trend resembling that observed for human CYP1A2. This could suggest a low capability of the cat CYP1A2 homolog to form 1'-hydroxy-safrole compared to the other species considered. This result, along with the evidence observed for the cat homolog of CYP2A6 (see above), might point to its possible low production of 1'hydroxy-safrole, which should be assessed in further experiments, along with the role of phase II reactions to better assess the associated sensitivity to the genotoxic effects of safrole.

3.3. Estimate of Binding Energy

The Prodigy Web Server [26] (https://wenmr.science.uu.nl/prodigy accessed on the 18th of November 2022) was used to estimate the binding free energy of each complex to study the safrole-CYP interaction beyond the capability to form the proximate carcinogen of safrole (i.e., 1'-hydroxy-safrole). To do so, a cluster analysis of each complex trajectory was performed first to obtain

geometries of binding that are representative of the whole dynamics and using such poses as input for Prodigy (see Materials and Methods for further details). Of note, this approach was chosen for its actual applicability to an “in-bulk” analysis (24 complexes analysed), which is often not affordable in terms of computational efforts using other MD-based approaches. Moreover, this would have been beyond the scope of our work considering that the binding-free energy is not related directly to the likelihood of a substrate being converted into a specific product.

From the analysis, all the protein-ligand complexes were found to be energetically favoured (the calculated free energy ranged from –7.4 to –7.0 Kcal/mol; Supplementary Materials, Table S6) suggesting that all the CYPs analysed may favourably interact with safrole (in line with docking analysis; see above). However, the geometry of binding was not always compatible with the formation of the 1’-hydroxy-safrole, although the formation of other metabolites could not be excluded. Therefore, the interaction between safrole and the CYPs under analysis was hypothesised as likely to occur with all the CYPs analysed, but the formation of safrole proximate carcinogen (i.e., 1’-hydroxy-safrole) may occur on a species-specific basis.

4. Conclusion

This study demonstrated the effectiveness of the proposed modelling framework to predict the likelihood of small molecules being metabolised by members of the CYP enzyme family. The distance computed over time between the atom undergoing the reaction and the heme's Fe could be used as a geometric rationale to estimate the likeliness of the formation of a given metabolite.

Concerning the case studies presented, the evidence collected provides a knowledge-based ground that is useful to rationally design further experiments, such as a focused analysis of the hepatic expression of CYP enzymes in different species and the kinetics of biotransformation for specific CYP isoforms, to cite but a few. This is an essential piece of information to characterise species-specific differences in terms of biotransformation, bioactivation and related toxicity. In this respect, specific CYP isoforms have been mined from animal proteomes and prioritised for future assessment. Specifically, all the species considered (rat, mouse, dog, cat, rabbit, pig, goat, sheep and chicken) seemed prone to bioactivate safrole, having at least one homolog of CYP1A2 or CYP2A6 theoretically able to form 1'hydroxy-safrole. In addition, all the complexes analysed were found to be energetically stable, suggesting the possible capability of the CYPs under analysis to interact with safrole and biotransform it. However, the production of safrole's proximate carcinogen (i.e., 1'-hydroy-safrole) has been found to differ between species. Cats seem to be the species less likely to bioactivate safrole, considering the calculated lack of formation of 1'hydroxy-safrole by its CYP2A6 homolog (CYP2A13) and the low biotransformation yield expected for its CYP1A2. Nonetheless, the existence in cat proteome of other CYPs able to significantly form 1'hydroxy-safrole cannot be excluded. Additionally, the relatively high hepatic expression of cat CYPs reported in the Bgee database (Table S1; Supporting Material), which is in line with the expression reported for human homologs, may eventually lead to significant biotransformation of safrole. Therefore, in safrole biotransformation

studies using cat liver tissue fractions, the measurement of CYPs expression and activity, and the study of phase II reactions would be informative and aid in a better characterisation of cats' susceptibility to safrole toxicity. This would be important from a feed safety perspective as cats have the lowest calculated safe concentrations for the presence of safrole-containing essential oils (22 mg/Kg of complete feed) in feed reported thus far for farm and companion animals [10]. Conversely, sheep and goats, for which the highest calculated safe concentrations of safrole-containing essential oils in feed have been proposed (50 mg/Kg; EFSA, 2022), appeared to have the highest performing CYP homologs in terms of forming 1'hydroxy-safrole. However, according to the Bgee database, sheep have the lowest hepatic expression scores for the considered CYPs, possibly suggesting reduced biotransformation due to their limited expression at the main metabolizing organ for safrole toxicity (see Table S1; Supporting Material). Therefore, information on the actual hepatic expression of CYP1A2 in both species and CYP2A13 and CYP2A6 in goat and sheep, respectively, along with a precise measurement of enzymatic activities, would be required to obtain better insights into the bioactivation capacities of these species. As the toxic effect depends on the formation of the reactive 1'sulfooxysafrole metabolite, (quantitative) information on other phase I and phase II reactions (integrated into physiologically-based kinetic (PBK) models) would be required to predict kinetic-dependent species differences in sensitivity to safrole toxicity. This will eventually provide a science-based approach to determine the safe use of safrole-containing essential oils for different species. From a general point of view, the results presented expand the current understanding of safrole toxicity and bioactivation and CYPs mechanics, which are both pivotal for a more informed framework of analysis to assess safrole toxicity. In addition, the outcomes described, when integrated either with experimental confirmations or other modelling techniques, may provide a useful line of evidence to support a more informed species extrapolation, helping the identification of animal species close to humans and valuable for the development of effective model systems.

References

1. Gotz, M.E.; Sachse, B.; Schafer, B.; Eisenreich, A. Myristicin and Elemicin: Potentially Toxic Alkenylbenzenes in Food. *Foods* **2022**, *11*, 1988. https://doi.org/10.3390/foods11131988.

2. Eisenreich, A.; Gotz, M.E.; Sachse, B.; Monien, B.H.; Herrmann, K.; Schafer, B. Alkenylbenzenes in Foods: Aspects Impeding the Evaluation of Adverse Health Effects. *Foods* **2021**, *10*, 2139. https://doi.org/10.3390/foods10092139.

3. Atkinson, R.G. Phenylpropenes: Occurrence, Distribution, and Biosynthesis in Fruit. *J. Agric. Food Chem.* **2018**, *66*, 2259–2272. https://doi.org/10.1021/acs.jafc.6b04696.

4. Jeurissen, S.M.F.; Punt, A.; Boersma, M.G.; Bogaards, J.J.P.; Fiamegos, Y.C.; Schilter, B.; van Bladeren, P.J.; Cnubben, N.H.P.; Rietjens, I. Human cytochrome p450 enzyme specificity for the bioactivation of estragole and related alkenylbenzenes. *Chem. Res. Toxicol.* **2007**, *20*, 798–806. https://doi.org/10.1021/tx700012d.

5. IARC. IARC Monographs on the Evaluation of the Carcinogenic Risks to Humans. In *Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs*' IARC: Lyon, France, 1987.

6. Bampidis, V.; Azimonti, G.; Bastos, M.D.; Christensen, H.; Durjava, M.F.; Kouba, M.; Lopez-Alonso, M.; Puente, S.L.; Marcon, F.; Mayo, B.; et al. Safety and efficacy of a feed additive consisting of a tincture from the bark of Cinnamomum verum J. Presl (cinnamon tincture) for use in all animal species (FEFANA asbl). *Efsa J.* **2021**, *19*, e06986. https://doi.org/10.2903/j.efsa.2021.6986.

7. Hausner, E.; Poppenga, R.H. Chapter 26—Hazards Associated with the Use of Herbal and Other Natural Products. In *Small Animal Toxicology (Third Edition)*, Peterson, M.E., Talcott, P.A., Eds.; Elsevier: Amsterdam, The Netherlands, 2013; pp. 335–356.

8. Prinsloo, G.; Nogemane, N.; Street, R. The use of plants containing genotoxic carcinogens as foods and medicine. *Food Chem. Toxicol.* **2018**, *116*, 27–39. https://doi.org/10.1016/j.fct.2018.04.009.

9. Siano, F.; Ghizzoni, C.; Gionfriddo, F.; Colombo, E.; Servillo, L.; Castaldo, D. Determination of estragole, safrole and eugenol methyl ether in food products. *Food Chem.* **2003**, *81*, 469–475. https://doi.org/10.1016/s0308-8146(03)00004-9.

10. Bampidis, V.; Azimonti, G.; Bastos, M.D.; Christensen, H.; Durjava, M.F.; Kouba, M.; Lopez-Alonso, M.; Puente, S.L.; Marcon, F.; Mayo, B.; et al. Safety and efficacy of a feed additive consisting of an essential oil from Cinnamomum camphora (L.) J. Presl (camphor white oil) for use in all animal species (FEFANA asbl). *Efsa J.* **2022**, *20*, e06985. https://doi.org/10.2903/j.efsa.2022.6985.

11. Bode, A.M.; Dong, Z.G. Toxic Phytochemicals and Their Potential Risks for Human Cancer. *Cancer Prev. Res.* **2015**, *8*, 1–8. https://doi.org/10.1158/1940-6207.capr-14-0160.

12. Martati, E.; Boersma, M.G.; Spenkelink, A.; Khadka, D.B.; van Bladeren, P.J.; Rietjens, I.; Punt, A. Physiologically Based Biokinetic (PBBK) Modeling of Safrole Bioactivation and Detoxification in Humans as Compared With Rats. *Toxicol. Sci.* **2012**, *128*, 301–316. https://doi.org/10.1093/toxsci/kfs174.

13. Carmichael, P.L.; Baltazar, M.T.; Cable, S.; Cochrane, S.; Dent, M.; Li, H.Q.; Middleton, A.; Muller, I.; Reynolds, G.; Westmoreland, C.; et al. Ready for Regulatory Use: NAMs and NGRA for Chemical Safety Assurance. *Altex-Altern. Anim. Exp.* **2022**, *39*, 359–366. https://doi.org/10.14573/altex.2204281.

14. Itoh, T.; Takemura, H.; Shimoi, K.; Yamamoto, K. A 3D Model of CYP1B1 Explains the Dominant 4-Hydroxylation of Estradiol. *J. Chem. Inf. Model.* **2010**, *50*, 1173–1178. https://doi.org/10.1021/ci1000554.

15. Dorne, J.; Cirlini, M.; Louisse, J.; Pedroni, L.; Galaverna, G.; Dellafiora, L. A Computational Understanding of Inter-Individual Variability in CYP2D6 Activity to Investigate the Impact of Missense Mutations on Ochratoxin A Metabolism. *Toxins* **2022**, *14*, 207. https://doi.org/10.3390/toxins14030207.

16. Sridhar, J.; Goyal, N.; Liu, J.W.; Foroozesh, M. Review of Ligand Specificity Factors for CYP1A Subfamily Enzymes from Molecular Modeling Studies Reported to-Date. *Molecules* **2017**, *22*, 1143. https://doi.org/10.3390/molecules22071143.

17. von Weymarn, L.B.; Murphy, S.E. Coumarin metabolism by rat esophageal microsomes and cytochrome P450 2A3. *Chem. Res. Toxicol.* **2001**, *14*, 1386–1392. https://doi.org/10.1021/tx010065v.

18. Papatheodorou, I.; Moreno, P.; Manning, J.; Fuentes, A.M.P.; George, N.; Fexova, S.; Fonseca, N.A.; Fullgrabe, A.; Green, M.; Huang, N.; et al. Expression Atlas update: From tissues to single cells. *Nucleic Acids Res.* **2020**, *48*, D77–D83. https://doi.org/10.1093/nar/gkz947.

19. Bastian, F.B.; Roux, J.; Niknejad, A.; Comte, A.; Costa, S.S.F.; de Farias, T.M.; Moretti, S.; Parmentier, G.; de Laval, V.R.; Rosikiewicz, M.; et al. The Bgee suite: Integrated curated expression atlas and comparative transcriptomics in animals. *Nucleic Acids Res.* **2021**, *49*, D831–D847. https://doi.org/10.1093/nar/gkaa793.

20. Righetti, L.; Rolli, E.; Dellafiora, L.; Galaverna, G.; Suman, M.; Bruni, R.; Dall'Asta, C. Thinking Out of the Box: On the Ability of Zea mays L. to Biotrasform Aflatoxin B1 Into Its Modified Forms. *Front. Plant Sci.* **2021**, *11*, 9158. https://doi.org/10.3389/fpls.2020.599158.

21. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. https://doi.org/10.1006/jmbi.1990.9999.

22. Jeurissen, S.M.F.; Bogaards, J.J.P.; Awad, H.M.; Boersma, M.G.; Brand, W.; Fiamegos, Y.C.; van Beek, T.A.; Alink, G.M.; Sudholter, E.J.R.; Cnubben, N.H.P.; et al. Human cytochrome P450 enzyme specificity for bioactivation of safrole to the proximate carcinogen 1 ‘-hydroxysafrole. *Chem. Res. Toxicol.* **2004**, *17*, 1245–1250. https://doi.org/10.1021/tx040001v.

23. Martati, E.; Boersma, M.G.; Spenkelink, A.; Khadka, D.B.; Punt, A.; Vervoort, J.; van Bladeren, P.J.; Rietjens, I. Physiologically Based Biokinetic (PBBK) Model for Safrole Bioactivation and Detoxification in Rats. *Chem. Res. Toxicol.* **2011**, *24*, 818–834. https://doi.org/10.1021/tx200032m.

24. Kishida, T.; Muto, S.I.; Hayashi, M.; Tsutsui, M.; Tanaka, S.; Murakami, M.; Kuroda, J. Strain differences in hepatic cytochrome P450 1A and 3A expression between Sprague-Dawley and Wistar rats. *J. Toxicol. Sci.* **2008**, *33*, 447–457. https://doi.org/10.2131/jts.33.447.

25. Sayers, E.W.; Bolton, E.E.; Brister, J.R.; Canese, K.; Chan, J.; Comeau, D.C.; Connor, R.; Funk, K.; Kelly, C.; Kim, S.; et al. Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **2022**, *50*, D20–D26. https://doi.org/10.1093/nar/gkab1112.

26. Vangone, A.; Schaarschmidt, J.; Koukos, P.; Geng, C.L.; Citro, N.; Trellet, M.E.; Xue, L.C.; Bonvin, A. Large-scale prediction of binding affinity in protein-small ligand complexes: The PRODIGY-LIG web server. *Bioinformatics* **2019**, *35*, 1585–1587. https://doi.org/10.1093/bioinformatics/bty816.

27. 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.; et al. PubChem in 2021: New data content and improved web interfaces. *Nucleic Acids Res.* **2021**, *49*, D1388–D1395. https://doi.org/10.1093/nar/gkaa971.

28. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. https://doi.org/10.1093/nar/28.1.235.

29. Sansen, S.; Yano, J.K.; Reynald, R.L.; Schoch, G.A.; Griffin, K.J.; Stout, C.D.; Johnson, E.F. Adaptations for the oxidation of polycyclic aromatic hydrocarbons exhibited by the structure of human P450 1A2. *J. Biol. Chem.* **2007**, *282*, 14348–14355. https://doi.org/10.1074/jbc.M611692200.

30. Sansen, S.; Hsu, M.H.; Stout, C.D.; Johnson, E.F. Structural insight into the altered substrate specificity of human cytochrome P450 2A6 mutants. *Arch. Biochem. Biophys.* **2007**, *464*, 197–206. https://doi.org/10.1016/j.abb.2007.04.028.

31. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. https://doi.org/10.1002/jcc.20084.

32. Louisse, J.; Dorne, J.L.C.M.; Dellafiora, L. 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. *Toxycology Lett.* **2022**, *355*, 19–30.

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

34. Sali, A.; Blundell, T.L. Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815. https://doi.org/10.1006/jmbi.1993.1626.

35. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–+. https://doi.org/10.1038/s41586-021-03819-2.

36. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948. https://doi.org/10.1093/bioinformatics/btm404.

37. Maldonado-Rojas, W.; Olivero-Verbel, J. Potential interaction of natural dietary bioactive compounds with COX-2. *J. Mol. Graph. Model.* **2011**, *30*, 157–166. https://doi.org/10.1016/j.jmgm.2011.07.002.

38. Adasme, M.F.; Linnemann, K.L.; Bolz, S.N.; Kaiser, F.; Salentin, S.; Haupt, V.J.; Schroeder, M. PLIP 2021: Expanding the scope of the protein-ligand interaction profiler to DNA and RNA. *Nucleic Acids Res.* **2021**, *49*, W530–W534. https://doi.org/10.1093/nar/gkab294.

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

40. Best, R.B.; Zhu, X.; Shim, J.; Lopes, P.E.M.; Mittal, J.; Feig, M.; MacKerell, A.D. 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.* **2012**, *8*, 3257–3273. https://doi.org/10.1021/ct300400x.

41. Zhang, L.; Silva, D.A.; Yan, Y.J.; Huang, X.H. Force field development for cofactors in the photosystem II. *J. Comput. Chem.* **2012**, *33*, 1969–1980. https://doi.org/10.1002/jcc.23016.

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

43. Del Favero, G.; Mayer, R.M.; Dellafiora, L.; Janker, L.; Niederstaetter, L.; Dall'Asta, C.; Gerner, C.; Marko, D. Structural Similarity with Cholesterol Reveals Crucial Insights into Mechanisms Sustaining the Immunomodulatory Activity of the Mycotoxin Alternariol. *Cells* **2020**, *9*, 847. https://doi.org/10.3390/cells9040847.

Supplementary Materials

Table S1. *Expression data for considered CYPs at the liver level in the species under analysis.*

Species	CYP2A6 and its homologues			CYP1A2		
	UniProt [1]	Expression Atlas [2]	Bgee [3]	UniProt [1]	Expression Atlas [2]	Bgee [3]
Human	Evidence at protein level	692010 parts per billion	98.57	Evidence at protein level	519668 parts per billion	95.32
Rat	Evidence at transcript level	n.a.	n.a.	Evidence at protein level	n.a.	98.65
Mouse	Evidence at transcript level	37033 parts per billion	91.96	Evidence at protein level	216750 parts per billion	99.16
Cat	Inferred from homology	n.a.	99.5	Evidence at transcript level	n.a.	98.42
Dog	Evidence at transcript level	n.a.	75.43	Evidence at protein level	n.a.	74.07
Pig	Evidence at protein level	n.a.	96.28	Evidence at protein level	n.a.	97.64
Goat	Inferred from homology	n.a.	96.66	Inferred from homology	n.a.	88.84
Rabbit	Evidence at protein level (CYP2A10)	n.a.	99.49	Evidence at protein level	n.a.	99.46
	Evidence at protein level (CYP2A11)	n.a.	n.a.			
Chicken	---	---	---	Evidence at transcript level	n.a.	94.68
Sheep	Evidence at transcript level	n.a.	53.57	n.a.	n.a.	58.23

Note: [1] Information of protein existence according to UniProt database (https://www.uniprot.org); [2] Data at protein level was considered (if any) and the highest value was reported according to Expression Atlas database (https://www.ebi.ac.uk/gxa/home); [3] Expression scores according to Bgee database (https://bgee.org) are reported; "n.d." stands for "data not available".

Table S2. *Accession code and percentage identity of animal homologs to human CYP2A6.*

Species [1]	Homolog to CYP2A6	Identity % to human CYP2A6
Rat (10116)	P20812 (CYP2A3)	85.38%
Mouse (10090)	P20852 (CYP2A5)	85.59%
Dog (9615)	Q307K8 (CYP2A13)	88.87%
Rabbit (9986)	Q05555 (CYP2A10)	83.69%
	Q05556 (CYP2A11)	83.47%
Pig (9823)	Q8SQ68 (CYP2A19)	86.44%
Goat (9925)	A0A452DNG6 (CYP2A13)	88.35%
Sheep (9940)	F1CGV2 (CYP2A6)	88.54%
Cat (9685)	M3W9T6 (CYP2A13)	86.65%
Chicken (9031)	ND [2]	< 50%

Note: [1] the species taxon ID is reported between brackets; [2] ND stands for data not found, meaning that the identity percentage to the human homolog was lower than 50% at the time of analysis (last database access 8th July 2022).

Table S3. *Accession number for animal homolog to human CYP1A2 or CYP2A6.*

Species [1]	Homolog to CYP1A2	Homolog to CYP2A6
Rat (10116)	P04799	P20812 (CYP2A3)
Mouse (10090)	P00186	P20852 (CYP2A5)
Dog (9615)	P56592	Q307K8 (CYP2A13)
Rabbit (9986)	P00187	Q05555 (CYP2A10)
		Q05556 (CYP2A11)
Pig (9823)	F1SJ26	Q8SQ68 (CYP2A19)
Goat (9925)	A0A452EFF2	A0A452DNG6 (CYP2A13)
Sheep (9940)	XP027812985	F1CGV2 (CYP2A6)
Cat (9685)	Q5KQT6	M3W9T6 (CYP2A13)
Chicken (9031)	Q01741	ND [2]

Note: [1] the species taxon ID is reported between brackets; [2] ND stands for data not found, meaning that the identity percentage to the human homolog was lower than 50% at the time of analysis (last database access 8th July 2022).

Table S4. *Structural comparison between the CYP2A6 homology modelling structures obtained using Modeller and the AlphaFold Protein Structure Database.*

Species [1]	Homolog to CYP2A6	RMSD (Å) [2]	RMSD (Å) [3]	RMSD (Å) [4]
Rat (10116)	P20812 (CYP2A3)	0.27	0.32	0.35
Mouse (10090)	P20852 (CYP2A5)	0.28	0.33	0.37
Dog (9615)	Q307K8 (CYP2A13)	0.28	0.33	0.62
Rabbit (9986)	Q05555 (CYP2A10)	0.27	0.42	0.42
	Q05556 (CYP2A11)	0.27	0.42	0.42
Pig (9823)	Q8SQ68 (CYP2A19)	0.27	0.31	0.33
Goat (9925)	A0A452DNG6 (CYP2A13)	0.28	0.31	0.34
Sheep (9940)	F1CGV2 (CYP2A6)	0.28	0.31	0.33
Cat (9685)	M3W9T6 (CYP2A13)	0.27	0.32	0.34

Note: [1] species taxon ID is reported between brackets; [2] RMSD computed superimposing the structure obtained via homology modelling using Modeller and the human CYP2A6 (PDB ID 2PG6); [3] RMSD computed superimposing the structure obtained from the AlphaFold Protein Structure Database and the human CYP2A6 (PDB ID 2PG6); [4] RMSD computed superimposing the structure obtained via Homology modelling and the one obtained on the AlphaFold Protein Structure Database.

Table S5. *Structural comparison between the CYP1A2 homology modelling structures obtained using Modeller and the AlphaFold Protein Structure Database.*

Species [1]	RMSD (Å) [2]	RMSD (Å) [3]	RMSD (Å) [4]
Rat (10116)	0.33	0.40	0.41
Mouse (10090)	0.30	0.43	0.44
Dog (9615)	0.17	0.38	0.39
Rabbit (9986)	0.17	0.44	0.46
Chicken (9031)	0.25	0.57	0.58
Pig (9823)	0.16	0.39	0.40
Goat (9925)	0.15	0.38	0.38
Sheep (9940)	0.17	0.73*	0.72*
Cat (9685)	0.18	0.48	0.63

Note: [1] species taxon ID is reported between brackets; [2] RMSD computed superimposing the structure obtained via homology modelling using Modeller and the human CYP1A2 (PDB ID 2HI4); [3] RMSD computed superimposing the structure obtained from the AlphaFold Protein Structure Database and the human CYP1A2 (PDB ID 2HI4); [4] RMSD computed superimposing the structure obtained via homology modelling using Modeller and the one obtained on the AlphaFold Protein Structure Database. The sheep CYP1A2 was not available on the AlphaFold Protein Structure Database and was then computed via the AlphaFold2 Colab notebook using default parameter (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb).*

Table S6. *Average predicted binding energy of the complexes under analysis*

Molecule	Species CYP	ΔG (Kcal/mol) [1]	Number of clusters [2]
Estragole	Human 1A2	-7.20	1
	Human 2A6	-7.00	1
Coumarin	Human 2A6	-7.10	1
	Rat 2A3	-7.00	1
Safrole	Human 1A2	-7.30	1
	Human 2A6	-7.20	1
	Cat 1A2	-7.40	1
	Cat 2A13	-7.25 ± 0.05 *	3
	Chicken 1A2	-7.30	1
	Dog 1A2	-7.30	1
	Dog 2A13	-7.30	1
	Goat 1A2	-7.30	1
	Goat 2A13	-7.10	1
	Mouse 1A2	-7.20	1
	Mouse 2A5	-7.20	1
	Pig 1A2	-7.40	1
	Pig 2A19	-7.20	1
	Rat 1A2	-7.20	1
	Rat 2A3	-7.30	1
	Rabbit 1A2	-7.30	1
	Rabbit 2A10	-7.30	1
	Rabbit 2A11	-7.30	1
	Sheep 1A2	-7.20	1
	Sheep 2A6	-7.20	1

*Note: [1] Binding energy computed on the PRODIGY web application on the cluster's representative complex (https://wenmr.science.uu.nl/prodigy/lig); [2] Number of clusters retrieved along the whole molecular dynamicssimulation; * The value reported represents the average value of each representative cluster.*

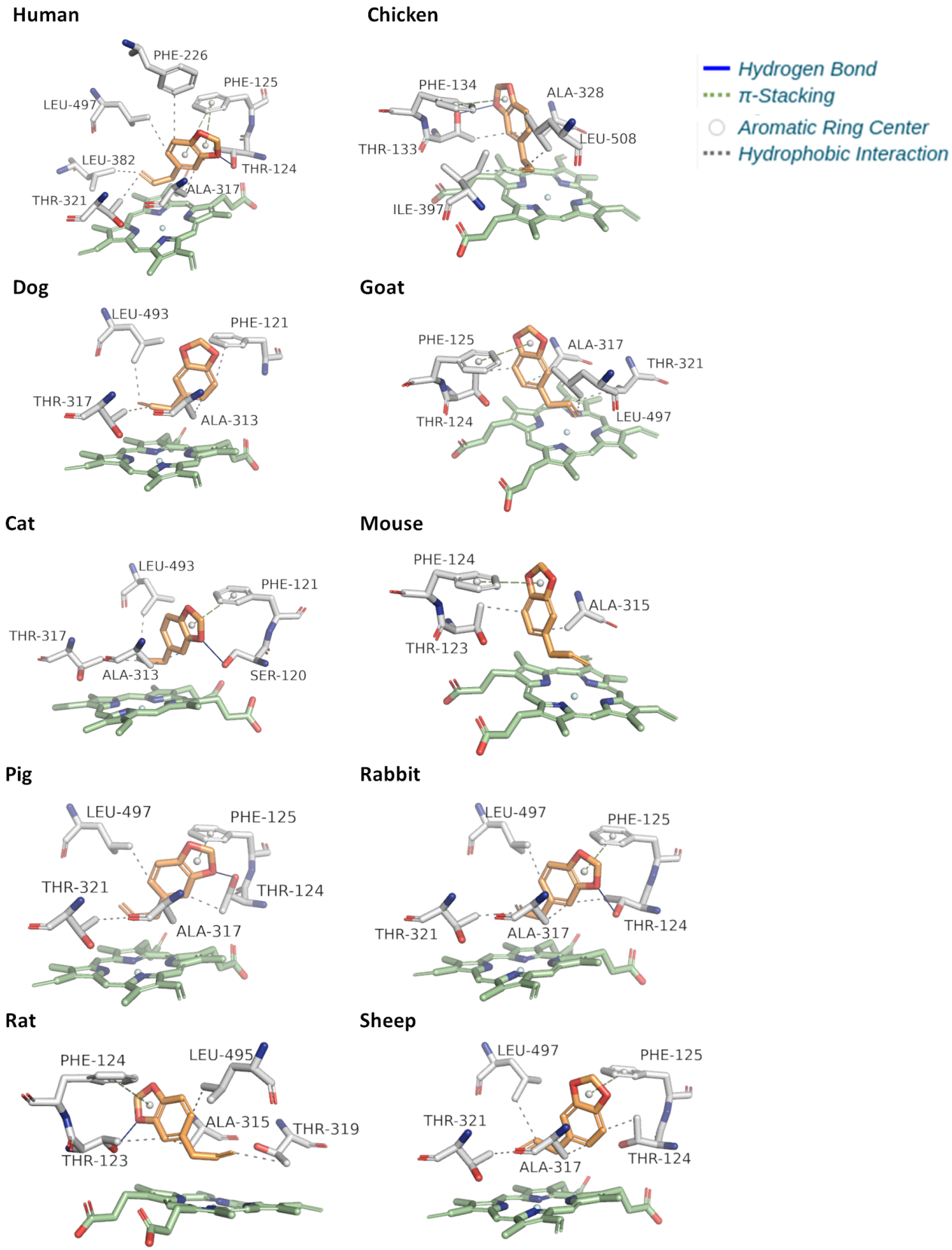


Figure S1. *Binding architecture of safrole in complex with human CYP1A2 and its animal homologs. Protein is shown in white sticks, ligand in yellow and heme in green.*

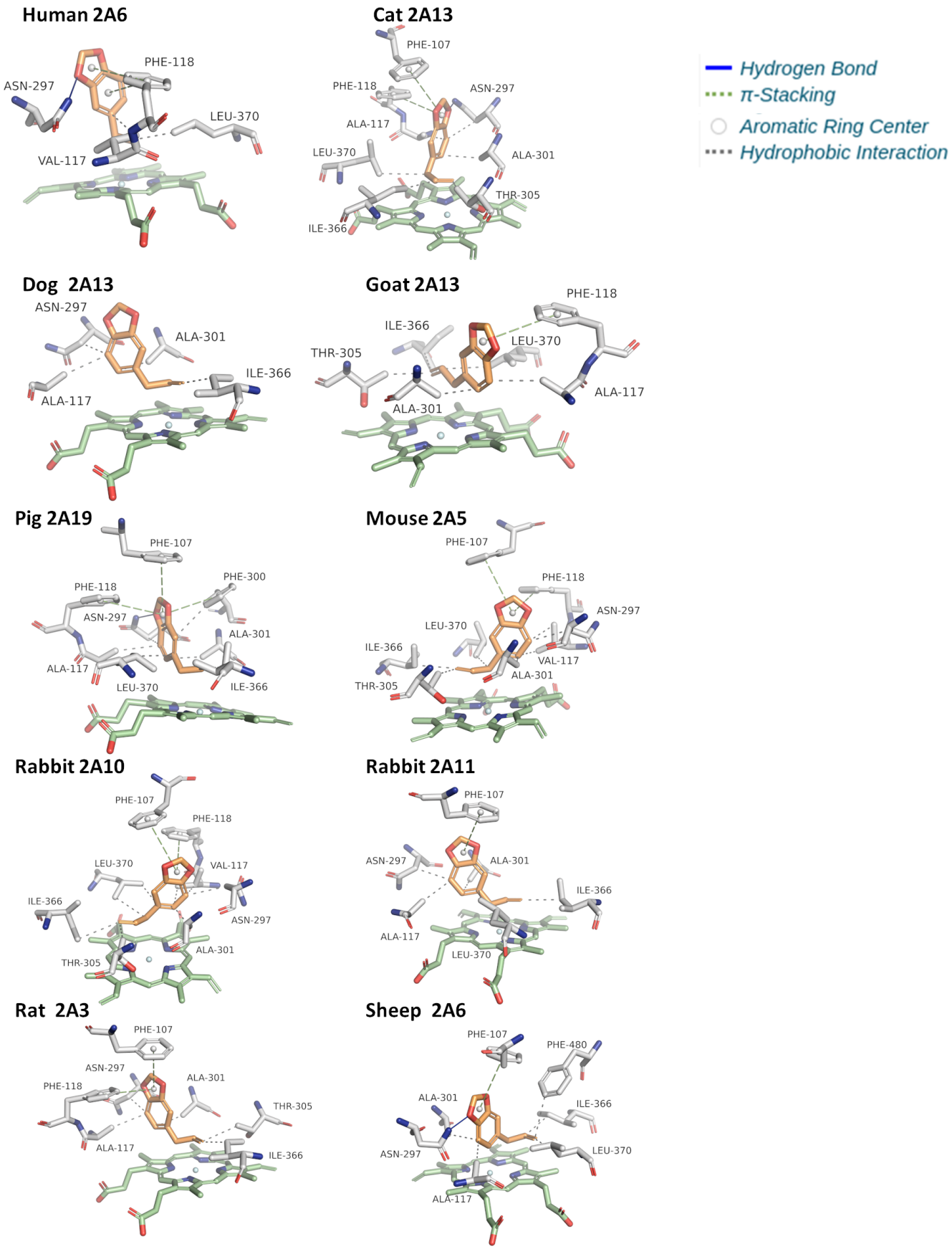


Figure S2. *Binding architecture of safrole in complex with human CYP2A6 and its animal homologs. Protein is shown in white sticks, ligand in yellow and heme in green.*

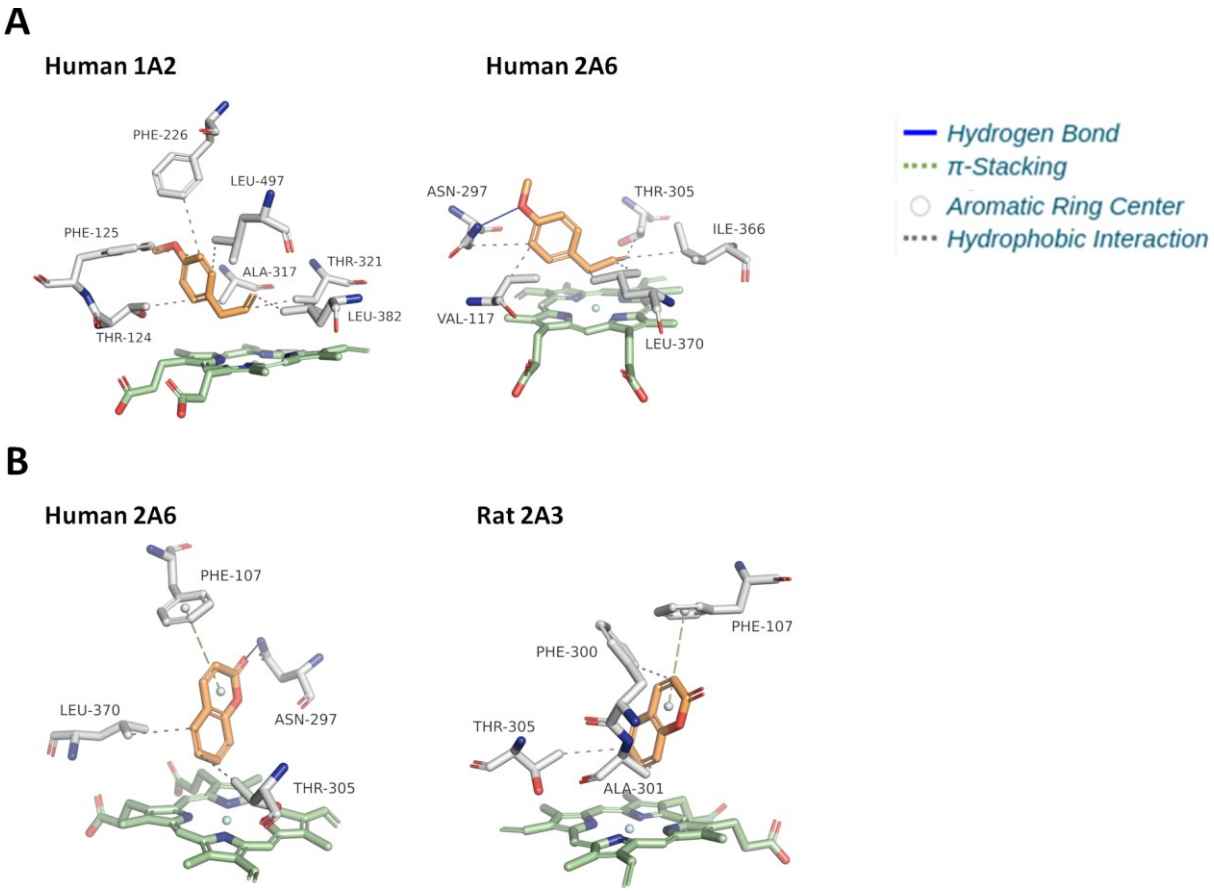


Figure S3. *Binding architecture of estragole within human CYP1A2 and CYP2A6 (**A**) and coumarin in human CYP2A6 and rat CYP2A3 (**B**). Protein is shown in white sticks, ligand in yellow and heme in green.*

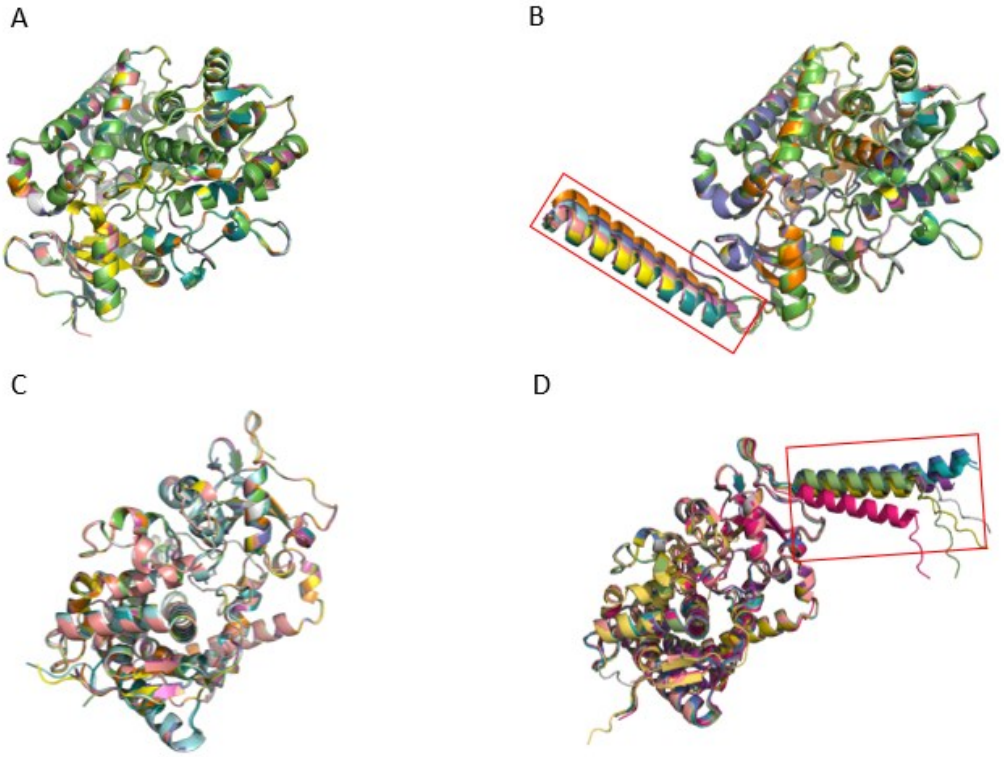


Figure S4. *Structural alignment of the animals' CYPs homologues over the human CYP2A6 (PDB ID 2PG6) and CYP1A2 (PDB ID 2HI4).* ***A****) CYP2A6 animal homologues structures obtained via the homology modelling procedure based on Modeller superimposed to the human CYP2A6 (PDB ID 2PG6).* ***B****) CYP2A6 animal homologues structures retrieved from the AlphaFold Protein Structure Database superimposed to the human CYP2A6 (PDB ID 2PG6). The main difference to structures reported in figure A is the transmembrane helix (within the red box), which was not relevant for the present study.* ***C****) CYP1A2 animal homologues structures obtained via the homology modelling procedure based on Modeller superimposed to the human CYP1A2 (PDB ID 2HI4).* ***D****) CYP1A2 animal homologues structures retrieved from the AlphaFold Protein Structure Database superimposed to the human CYP1A2 (PDB ID 2HI4). The main difference to structures reported in figure C is the transmembrane helix (within the red box), which was not relevant for the present study.*

```
#
#
# CYP2A6 Percent Identity  Matrix - created by Clustal2.1
#
#
```

1: Human-2a6	100.00	83.69	83.47	85.59	85.38	86.65	88.77	86.44	88.54	88.35
2: Rabbit-2a10	83.69	100.00	98.38	86.23	86.03	85.83	87.45	87.85	89.05	88.87
3: Rabbit-2a11	83.47	98.38	100.00	86.44	86.64	86.44	88.06	88.66	89.66	89.47
4: Mouse-2a5	85.59	86.23	86.44	100.00	96.15	86.23	89.07	89.07	89.25	89.27
5: Rat-2a3	85.38	86.03	86.64	96.15	100.00	87.25	90.08	88.87	89.86	89.68
6: Cat-2a13	86.65	85.83	86.44	86.23	87.25	100.00	94.13	91.30	91.89	92.11
7: Dog-2a13	88.77	87.45	88.06	89.07	90.08	94.13	100.00	92.11	93.31	93.12
8: Pig-2a19	86.44	87.85	88.66	89.07	88.87	91.30	92.11	100.00	93.10	93.32
9: Sheep-2a6	88.54	89.05	89.66	89.25	89.86	91.89	93.31	93.10	100.00	99.39
10: Goat-2a13	88.35	88.87	89.47	89.27	89.68	92.11	93.12	93.32	99.39	100.00

Figure S5. *Percent Identity Matrices (PIMs) for CYP2A6 homologs.*

Chapter II

A computational study on the biotransformation of alkenylbenzenes by a selection of CYPs: Reflections on their possible bioactivation

This chapter has been published as:

Pedroni Lorenzo, Louisse Jochem, Dorne Jean-Lou C.M., Dall'Asta Chiara, Dellafiora Luca. A computational study on the biotransformation of alkenylbenzenes by a selection of CYPs: Reflections on their possible bioactivation. Toxicology. 2023 Apr;488:153471. doi: 10.1016/j.tox.2023.153471.

Abstract

Alkenylbenzenes are aromatic compounds found in several vegetable foods that can cause genotoxicity upon bioactivation by members of the cytochrome P450 (CYP) family, forming 1'-hydroxy metabolites. These intermediates act as proximate carcinogens and can be further converted into reactive 1'-sulfooxy metabolites, which are the ultimate carcinogens responsible for genotoxicity. Safrole, a member of this class, has been banned as a food or feed additive in many countries based on its genotoxicity and carcinogenicity. However, it can still enter the food and feed chain. There is limited information about the toxicity of other alkenylbenzenes that may be present in safrole-containing foods, such as myristicin, apiole, and dillapiole. *In vitro* studies showed safrole as mainly bioactivated by CYP2A6 to form its proximate carcinogen, while for myristicin this is mainly done by CYP1A1. However, it is not known whether CYP1A1 and CYP2A6 can activate apiole and dillapiole. The present study uses an *in silico* pipeline to investigate this knowledge gap and determine whether CYP1A1 and CYP2A6 may play a role in the bioactivation of these alkenylbenzenes. The study found that the bioactivation of apiole and dillapiole by CYP1A1 and CYP2A6 is limited, possibly indicating that these compounds may have limited toxicity, while describing a possible role of CYP1A1 in the bioactivation of safrole. The study expands the current understanding of safrole toxicity and bioactivation and helps understand the mechanisms of CYPs involved in the bioactivation of alkenylbenzenes. This information is essential for a more informed analysis of alkenylbenzenes toxicity and risk assessment.

1. Introduction

Alkenylbenzenes are secondary metabolites of herbs and spices, like basil, fennel, and parsley to cite but a few, which are consumed worldwide as food *per se*, used as food and feed ingredients, as well as to extract essential oils widely used in food and feed [1]. Alkenylbenzenes are structurally related chemical analogues, including safrole, apiole, myristicin and dillapiole, sharing a 1,3-benzodioxole scaffold substituted by an allyl group at position 5 and various substitutions by methoxy groups on the aromatic ring (Figure 1). Several members of this class proved marked toxicity to animals, including carcinogenicity, raising food safety concerns due to their occurrence in certain food and feed [1-3]. The carcinogenicity of alkenylbenzenes has a genotoxic mode of action through the cytochromes P450 (CYPs)-dependent formation of 1'-hydroxy metabolites (proximate carcinogens), which can subsequently be converted by phase II metabolism to 1'-sulfooxy metabolites (ultimate carcinogens) [4, 5]. The ultimate carcinogens 1'-sulfooxy metabolites may form DNA adducts causing the genotoxic insult [5]. Hence, the genotoxicity of alkenylbenzenes inherently relies on the biotransformation to 1'-hydroxy metabolites, which is a key factor to investigate when studying their (geno)toxicity [6].

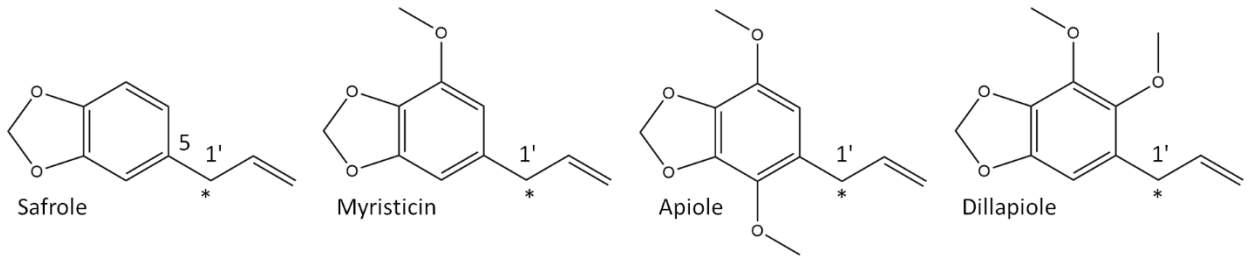


Figure 1. *Chemical structure of alkenylbenzenes under investigation. The asterisk indicates the atom undergoing the reaction.*

Safrole is one of the alkenylbenzenes for which a carcinogenicity study in rodents is available [7]. Such data may be used to derive a point of departure (e.g. a BMDL10) that can be used to determine the margin of exposure (MOE; margin between BMDL10 and exposure estimate) in order to

determine whether the chemical is a priority for risk management actions. Myristicin, apiol and dillapiole have stark structural analogies to safrole sharing the safrole 1,3-benzodioxole scaffold along with the allylic side chain (Figure 1) whose 1'-hydroxylation and subsequent 1'-sulfoxidation may produce the ultimate carcinogen metabolite. However, carcinogenicity studies on these compounds are still missing or inconclusive and toxicity data are generally scarce, although they have been described developmental toxicants, hepatotoxic, nephrotoxic and potentially able to reduce fertility [8]. They have been found at a variable level (from 0.2 to 67.5% of the total content of alkenylbenzenes, depending on the matrix) in essential oils mainly derived from nutmeg (myristicin), dill (dillapiole) and parsley (apiole) (Dosoky and Setzer 2021). However, detected levels may vary based on environmental factors, which may influence their production by plants, and on the analytical methods used to perform the analysis [9]. Besides occurrence analysis, and germane to toxicity investigations to fill the current data gaps for risk assessment, the application of novel approach methodologies (NAMs), including *in vitro* and *in silico* methods, is considered preferred above generation of new *in vivo* data. An *in vitro* study has shown the following order of potency to form DNA adducts of these alkenylbenzenes in human HepG2 cells: safrole ≈ myristicin > dillapiole > apiole [10]. Concerning the bioactivation to proximate carcinogens, available *in vitro* evidence shows that safrole is solely bioactivated by CYP2A6 [5], while myristicin is mainly bioactivated by CYP1A1, which has been also described as crucial for the myristicin-dependent formation of DNA adducts [11, 12]. However, neither the role of CYP1A1 to bioactivate safrole, nor of both CYPs to bioactivate apiole and dillapiole, or of CYP2A6 to bioactivate myristicin, have been duly assessed so far. This scenario provided a compelling rationale to further test CYP1A1 and CYP2A6 over other possible alkenylbenzenes as this incomplete picture may prevent those compounds from being properly evaluated for their toxicity and assessed for related risks.

The use of safrole as pure compound in food and feed production has been prohibited within both the EU and US due to its genotoxic and carcinogenic potential [13, 14]. However, safrole may still be present in food and feed being part of essential oils and ingredients widely used in their production [1]. This makes still advisable to advance its toxicological understanding. In addition, safrole may co-occur along with other congeners [1], including myristicin, apiole and dillapiole, suggesting the need for a broader assessment of safrole congeners – especially for those with strong structural similarities which may show analogies also from a toxicological standpoint. In such a process, toxicokinetic similarities including CYP-mediated formation of toxic metabolites should be addressed thoroughly while closely related members of a class are evaluated [15]. Germane to this work's topic, the CYP-mediated formation of alkenylbenzenes proximate carcinogens may represent an upstream functional event to support such analysis. On this basis, safrole, apiole, dillapiole and myristicin have been assessed for their likeliness to act as substrates of CYP1A1 and CYP2A6, being those CYPs identified as crucial for the activation of the considered alkenylbenzenes (see above).

Based on the above, a previously validated *in silico* protocol [16, 17] has been applied to fill the current data gap for safrole, myristic, apiol and dillapiole bioactivation. Of note, stand-alone *in silico* methods already proved to succeed in studying the toxicity of small molecules (e.g. [18-20] to cite but a few) and molecular modelling approaches like that used here have already been applied to study the safety of chemicals (e.g. [20]). Specific applications of molecular modelling methodologies have a long story of use to study the CYP-mediated biotransformation of small molecules [16, 17, 21-23]. In this work, the interaction of safrole, apiole, dillapiole and myristicin with CYP1A1 and CYP2A6 has been investigated via docking and molecular dynamics simulations to assess the complex stability over time as a mean to predict the substrate likeliness of those compounds, in agreement with previous studies [16, 17]. Briefly, the geometries of interaction between the atoms undergoing the reaction (Figure 1) and the Fe-heme have been monitored being a probing

geometrical feature to discriminate CYP substrates from non-substrate molecules [16]. This approach already succeeded to study specifically the CYP-dependent bioactivation of safrole [17] and it has been applied here to extend the analysis over a set of safrole analogues possibly relevant to food safety.

Overall, this study investigated similarities and differences of bioactivation of safrole, apiole, dillapiole and myristicin by CYP1A1 and CYP2A6 being relevant to: i) extend the current understanding of safrole's congeners toxicity; ii) shed light on the mechanics of CYPs involved in such a process; and iii) to rationally design further dedicated investigations.

2. Materials and Methods

2.1. Data source

The 3D structures of safrole, apiole, dillapiole and myristicin were retrieved in the .sdf format from the PubChem database (https://pubchem.ncbi.nlm.nih.gov) [24] with CID 8815, 10659, 10231 and 4276 respectively (CAS codes 140-67-0, 523-80-8, 484-31-1 and 607-91-0, respectively). The models of human CYP1A1 and CYP2A6 were derived from the crystallographic structures recorded in the Protein Databank (PDB; https://www.rcsb.org) [25] with PDB code 4I8V [26] and 2PG6 [27], respectively. The latter structure had two mutations, L240C and N297Q, which have been reverted to the wild-type sequence replacing the respective amino acid side chain using the Structure Editing/Rotamer tool of UCSF Chimera software (version 1.15) [28], in agreement with previous studies [29].

2.2. Docking simulations

Docking simulations provided plausible binding architecture for the molecules under analysis within the catalytic site of the two CYPs. This was performed, in agreement with previous studies, using the GOLD software (version 2021) as it proved a high reliability to study protein–ligand interactions, including those with CYPs [16, 30, 31]. The binding site was defined within a 10 Å radius sphere around the centroid of the substrate-binding site. The docking protocol was set according to previous studies setting ligands fully flexible and protein semi-flexible, allowing polar hydrogens to rotate freely [16]. The internal scoring function GOLDScore was used being optimised to predict ligand-binding positions as reported in the manufacturer declaration (https://www.ccdc.cam.ac.uk; accessed on 6th December 2022). The best scored pose for each ligand was further analysed through molecular dynamics (see below). In addition, the binding architecture of styrene as per PDB structure 4HGF – which has a structure closely related to that of alkenylbenzenes under analysis –

has been used as a position restrain (constraint weight of 100 units) to facilitate the arrangement of Ligands, in agreement with previous studies [17, 32].

2.3. Molecular dynamics

Molecular dynamics allowed to assess the geometrical stability of CYP-ligand complexes over time as a mean to discriminate the likeliness of molecules under analysis to be substrates of CYP1A1 and CYP2A6. This was performed using GROMACS (version 2019.4) [33], while ligands were parametrised with the CHARMM27 all-atom force field [34]. The hydrogen database was modified according to previous works [16, 35, 36] to parameterise the heme group. Input structures were solvated with SPCE waters in a cubic periodic boundary condition, and Na^+ and Cl^- were added as counter ions to neutralise the system. Prior to running simulations, each system was energetically minimised to avoid steric clashes and to correct improper geometries using the steepest descent algorithm with a maximum of 5000 steps. Subsequently, each system underwent isothermal (300 K, coupling time 2 psec) and isobaric (1 bar, coupling time 2 psec) 100 psec simulations before running 25 nsec simulations (300 K with a coupling time of 0.1 psec and 1 bar with a coupling time of 2.0 psec).

2.4. Statistical analysis

The statistical analysis of interatomic distances between the atom undergoing the reaction and the Fe atom of heme group has been done with SPSS IBM (v. 27.0, SPSS Inc., Chicago, IL, USA). For each complex, distance values of 5000 frames have been considered, expressed as means ± standard deviation (SD) and compared to each other using one-way ANOVA (α=0.05) with Bonferroni as post hoc test (α=0.05).

2.5 Cluster analysis of protein-ligand complexes trajectories

All the complexes were analysed to retrieve geometries of binding representative of the whole simulations, in agreement with previous studies [37]. The GROMACS (version 2019.4) [33] *cluster* command was used setting gromos as method and the cutoff at 0.2 nm.

3. Results and Discussion

3.1. Fit-for-purpose validation

The procedure used here was validated in previous studies. Briefly, measuring the interatomic distance over time between the atom undergoing the reaction and the heme's Fe previously proved to be a probative geometrical parameter to predict the capability of small molecules, including safrole, to be biotransformed by CYPs [16, 17]. Indeed, as previously demonstrated, the likeliness of a ligand to act as a substrate to form a certain metabolite (e.g. hydroxylated compounds) for a given CYP can be inferred when the atom undergoing the reaction is properly arranged to the heme's Fe by docking simulations and stably persist therein during molecular dynamics. In this respect, a distance threshold calculated as the average interatomic distance over time has been previously set at 0.53 nm to distinguish CYPs substrates [16, 17]. On this basis and in order to provide a fit-for-purpose validation of the procedure for the case studies under analysis, myristicin and safrole were considered reference compounds. In fact, they are well-known substrates of CYP1A1 and CYP2A6, respectively, both forming 1'-hydroxy-metabolites [5, 11]. Concerning the arrangement of safrole and myristicin at the respective CYP's catalytic site provided by docking simulations, both compounds had the atom undergoing the reaction properly oriented toward the heme's Fe (Figure 2A). Also, they both collected positive docking scores, pointing to their favourable interaction (the higher the score, the better the interaction [16]) with the respective CYP (GOLDScore for safrole and myristicin of 162 and 172 units in CYP2A6, and CYP1A1, respectively; Table 1). Then, the distances between the atom undergoing the reaction and the heme's Fe was monitored over time through molecular dynamics. As shown in Figure 2B, the atom undergoing the reaction of both safrole and myristicin was kept stably close to heme's Fe along all simulation with a mean distance of 0.39 ± 0.05 and 0.38 ± 0.04 nm for safrole within CYP2A6 and myristicin within CYP1A1, respectively. Of note,

these distances were both below the threshold mentioned above and like those previously described for the atoms of substrates proven to undergo hydroxylation by CYPs [16]. This evidence eventually confirmed the reliability of the procedure used to the case studies under analysis.

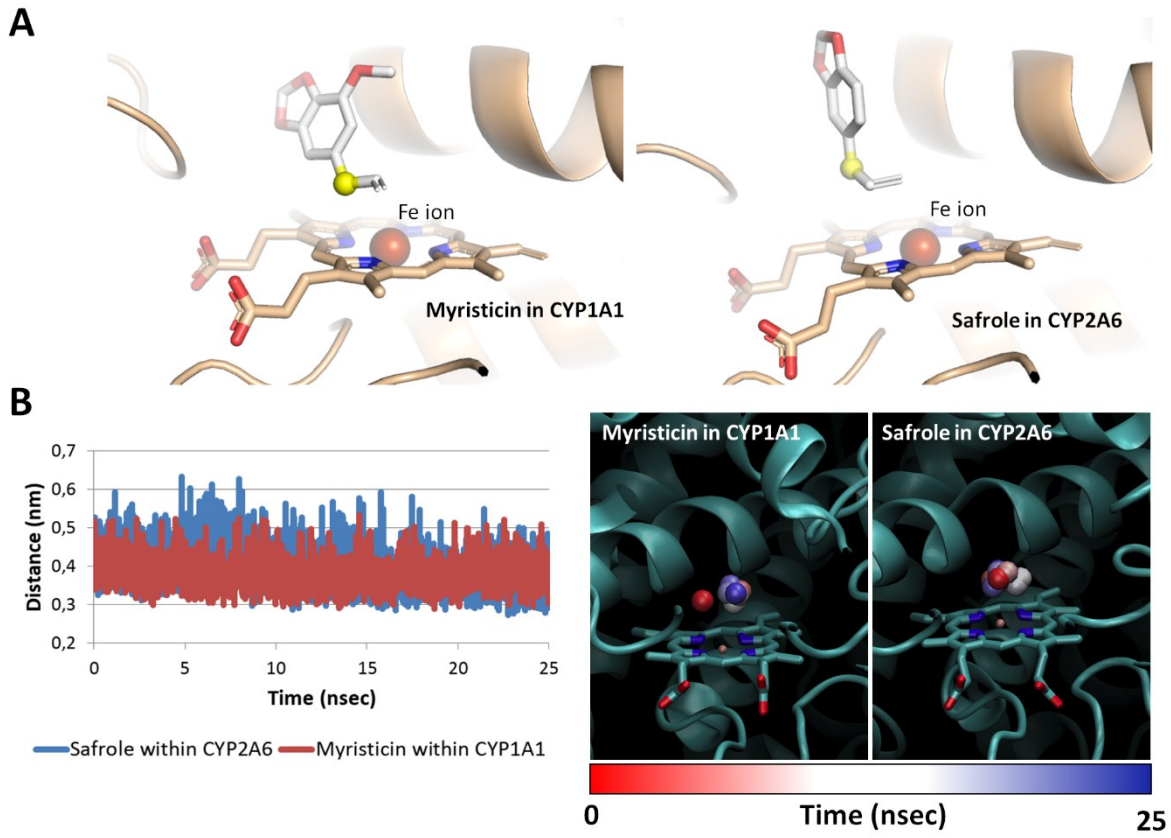


Figure 2. *Docking and molecular dynamics results for myristicin and safrole within CYP1A1 and CYP2A6, respectively.* ***A.*** *Docking pose of safrole and myristicin within CYP2A6 and CYP1A1, respectively. Ligands and heme group are represented in sticks while proteins are represented in cartoon. The heme's Fe atom is represented by red spheres. The atom undergoing the reaction is represented by yellow spheres.* ***B.*** *Molecular dynamics of safrole and myristicin within CYP2A6 and CYP1A1, respectively. The graph on the left shows the interatomic distances between heme's Fe and the atom undergoing reaction of myristicin and safrole within CYP1A1 and CYP2A6, respectively. The figure on the right shows the time step representation of the trajectories of atom undergoing reaction (shown in sphere) of myristicin within CYP1A1 and safrole within CYP2A6. Proteins are represented in cartoon, while heme is represented in sticks. The from-red-to-blue colour fade indicates the stepwise changes of coordinates along the simulation.*

3.2. Extended analysis of alkenylbenzenes interaction with CYP1A1 and CYP2A6

Once ascertained the procedural reliability, the interaction of dillapiole and apiole with CYP1A1 and CYP2A6, as well as that of safrole with CYP1A1 and myristicin with CYP2A6, was calculated to infer their respective likeliness to form 1-hydroxy metabolites by the CYPs considered. Particularly, CYP1A1 was investigated since recent evidence proved its hepatic expression, suggesting to deeper study its role in xenobiotics clearance and possible bioactivation to genotoxic compounds [38]. Moreover, CYP1A1 plays a pivotal role in myristicin 1'-hydroxylation as well as in myristicin-dependent DNA adducts formation, although other CYPs may have an ancillary role in its bioactivation [10, 11]. Therefore, the marked structural analogies between myristicin and the other alkenlybenzenes under analysis provided a sound rationale to test CYP1A1 against safrole, dillapiole and apiole. Adhering to this line of interpretation, CYP2A6, which is the sole described able to bioactivate safrole, has been tested against myristicin, apiol and dillapiole.

Table 1. *Docking scores*

Compound	**GOLDScore***	
	CYP1A1	**CYP2A6**
Safrole	168	162
Myristicin	172	148
Apiole	174	152
Dillapiole	179	158

*Note: * a positive score indicates a favourable interaction with the protein, as per manufacturer declaration (https://www.ccdc.cam.ac.uk; accessed on 6th December 2022)*

As shown in Table 1, all the alkenylbenzenes under analysis recorded positive and relatively high docking scores pointing to their theoretical capability to favourably interact with CYP1A1 and CYP2A6. However, the visual inspection of the docking poses revealed substantial differences in the way the various congeners docked at the catalytic site of CYP1A1 and CYP2A6 (Figure 3). Concerning CYP2A6 (Figure 3A), safrole properly oriented the atom undergoing the reaction to the heme's Fe (in

line with its previously demonstrated biotransformation to 1'hydroxy-safrole, as described above) and a similar arrangement was observed for apiole and myristicin.

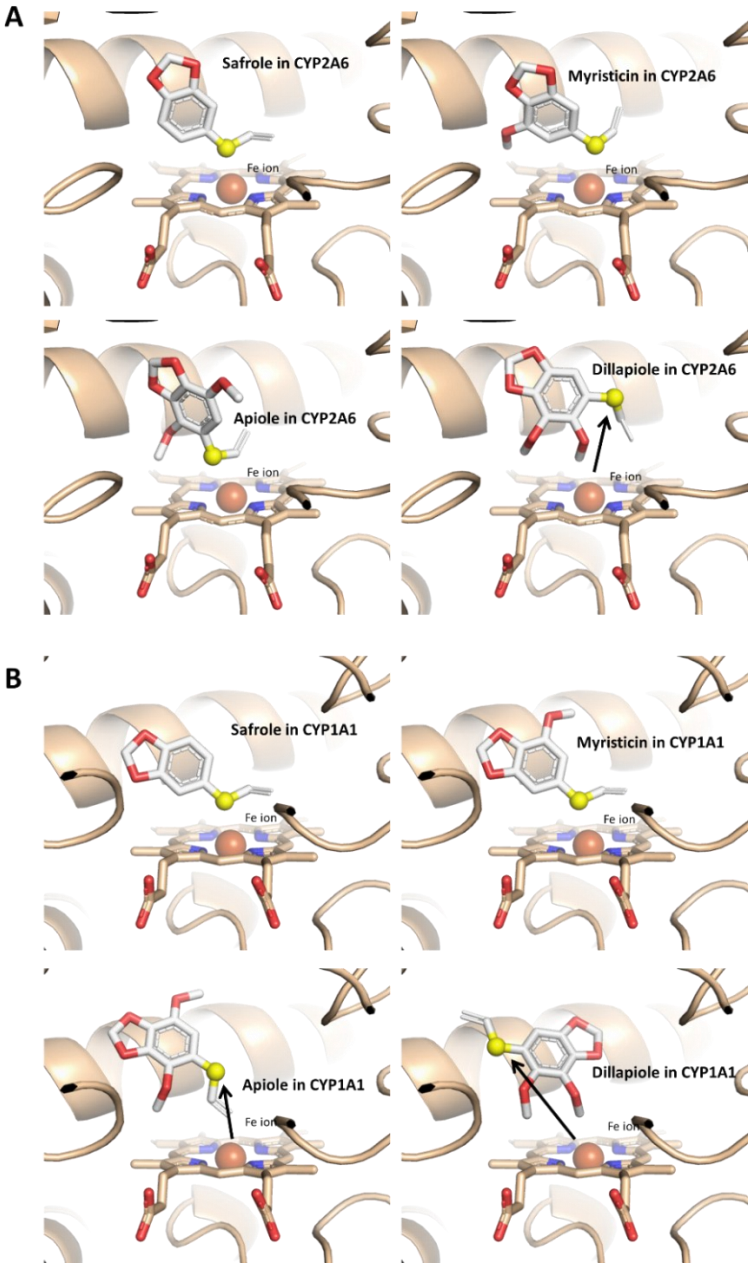


Figure 3. *Docking results for the alkenylbenzenes under investigation within CYP2A6 and CYP1A1. Ligands and heme group are represented in sticks while proteins are represented in cartoon. The heme's Fe atom is represented by red spheres. The atom undergoing the reaction is represented by yellow spheres. The black arrows indicate the displacement of the atom undergoing the reaction to the heme's Fe in certain complexes.* ***A.*** *Safrole, myristicin, apiole and dillapiole in CYP2A6.* ***B.*** *Safrole, myristicin, apiole and dillapiole in CYP1A1.*

Conversely, dillapiole showed a different architecture of binding with the atom undergoing the reaction arranged further away from the heme's Fe. Concerning CYP1A1 (Figure 3B), as per safrole within CYP2A6, myristicin properly oriented the atom undergoing the reaction to the heme's Fe, in line with the previously described biotransformation to 1'-hydroxy-myristicin (see above). Conversely, the atom undergoing the reaction of apiol and dillapiole was displaced far from the heme's Fe. Surprisingly, the atom undergoing the reaction of safrole was found instead arranged close to the heme's Fe retracing the placement of that of myristicin, possibly suggesting its biotransformation to 1'-hydroxy-safrole by CYP1A1. Each complex was then analysed through molecular dynamics to monitor whether the arrangement of the atom undergoing the reaction of the considered alkenylbenzenes to the heme's Fe was kept stable over time.

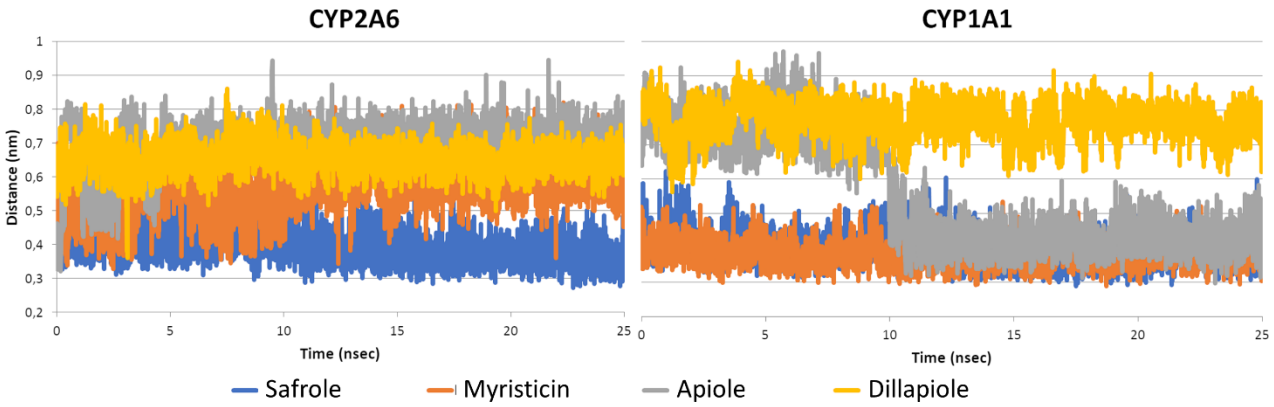


Figure 4. *Interatomic distances between the atom undergoing the reaction of safrole, myristicin, apiole and dillapiole to the heme's Fe within CYP2A6 or CYP1A1.*

As shown in Figure 4, concerning the dynamics within CYP2A6, safrole was the sole of the alkenylbenzenes considered that properly arranged the atom undergoing the reaction close to the heme's Fe (with an average distance of 0.39 ± 0.05 nm, see above). Conversely, myristicin, apiole and dillapiole showed significantly higher distances compared to safrole, and all above the previously established cut-off of 0.53 nm (0.57 ± 0.07, 0.70 ± 0.07 and 0.65 ± 0.04 nm, respectively; $p < 0.001$). This suggests that the biotransformation to their respective 1'-hydroxy-metabolite may be less likely than that of safrole, considering previous evidence describing those distance ranges non-compliant

to undergo an efficient hydroxylation by CYPs [16, 17]. Concerning CYP1A1, dillapiole showed the highest mean distance between the atom undergoing the reaction and the heme's Fe (0.77 ± 0.05 nm) suggesting its biotransformation to 1'-hydroxy-dillapiole as not likely. Conversely, the atom undergoing the reaction of apiole approached the heme's Fe during the simulation (from around 10 nanoseconds onward) while being stable till the end of dynamics. Although the mean distance of 0.57 ± 0.17 nm, apiole might show a certain degree of 1'-hydroxylation by CYP1A1 based on the close interatomic proximity reached during the simulation. Surprisingly, the atom undergoing the reaction of safrole was kept close to the heme's Fe along the whole simulation (Figure 4) and, though the mean interatomic distance was found statistically higher than that of myristicin (0.40 ± 0.04 and 0.38 ± 0.04; $p<0.001$), it was compliant to those previously described as prone to undergo hydroxylation by CYPs [16, 17]. A certain degree of safrole biotransformation to 1-'hydroxy-safrole by CYP1A1 shall be expected accordingly.

Based on the dynamics of interaction of safrole, myristicin, apiole and dillapiole, the latter two appeared to be those less likely prone to receive 1'-hydroxylation considering that the atom undergoing the reaction was not stably oriented toward the heme's Fe over time in neither of the two CYPs considered. However, apiole was considered prone of a limited 1'-hydroxylation by CYP1A1 as it properly arranged the atom undergoing the reaction to the heme's Fe during the simulation.

Taken together, the *in silico* results collected are in line with reported *in vitro* data on formation of the 1'-hydroxy-metabolites of safrole, myristicin and apiole, showing the highest catalytic efficiency for safrole 1'-hydroxylation (4.3 µL/min per mg microsomal protein) [39], followed by myristicin 1'-hydroxylation (0.73 µL/min per mg S9 protein) [40], and the lowest efficiency for 1'-hydroxylation of apiol (0.34 µL/min per mg S9 protein) [41]. Based on the *in silico* results from the present study, one would expect an even lower efficiency of dillapiole 1'-hydroxylation, although it must be noted that possible bioactivation by CYPs other than CYP1A1 and CYP2A6 cannot be excluded. In this respect,

the present work provided a mechanistic explanation for the available *in vitro* evidence detailing the incapability to properly arrange the atom undergoing the reaction to the heme's Fe as a basis for the lower efficiency in the bioactivation of myristicin, apiole and dillapiole compared to safrole, which is consistent with the lower toxicity of apiol and dillapiole described *in vitro* compared to safrole [10, 41]. Therefore, they may deserve a lower priority for future risk assessment evaluations. Although the final amount of formation of the ultimate carcinogens (genotoxicants), i.e., the 1'-sulfooxy-metabolites, do not solely depend on the rate of formation of the proximate carcinogen, i.e., the 1'-hydroxymetabolites, it is of interest to note that, in line with the *in silico* evaluation, in *in vitro* studies with human HepG2 liver cells, the formation of DNA adducts was more pronounced upon exposure to safrole and myristicin compared to apiole and dillapiole [10].

Concerning myristicin, the results collected in this study are in line with the data available so far pointing out that it is preferentially 1'-hydroxylated by CYP1A1, although a certain degree of biotransformation by CYP2A6 can be expected as well. This is in line with previous evidence describing that other CYPs are involved in the bioactivation of myristicin, though with a lower efficiency compared to CYP1A1, and CYP2A6 could be counted in the list of CYPs bioactivating myristicin deserving further dedicated investigations [10].

Concerning safrole, our data highlighted the mechanistic basis of its biotransformation by CYP2A6, showing the proper arrangement of the atom undergoing the reaction to the heme's Fe over the whole dynamic simulation. However, the present work also described that safrole could be efficiently 1'-hydroxylated by CYP1A1 showing a stable and proper arrangement of the atom undergoing the reaction to the heme's Fe. Of note, CYP1A1 has not been considered before as relevant for the biotransformation of safrole to 1'-hydroxy-safrole, though a certain degree of biotransformation was reported, but since its expression has been considered for long mainly extra-hepatic, its relevance for bioactivation of hepatoxicants has been thought limited [42]. However,

recent studies demonstrated its inter-individual and high inducibility also in the liver, suggesting the need to reconsider the role of CYP1A1 in xenobiotics clearance and their bioactivation to genotoxic compounds [38]. Moreover, although limited, previous evidence described the inducibility of CYP1A1 by safrole [43]. These outcomes, along with the data reported in the present work, provided a compelling line of evidence pointing out the relevance to evaluate thoroughly the role of CYP1A1 in the biotransformation of safrole and congeners. This is critical to provide a sound background of knowledge to properly investigate the group of alkenylbenzenes from a risk assessment standpoint and eventually pointing out genotypes potentially with a high yield of formation of safrole's proximate carcinogen (e.g. when CYP1A1 is particularly expressed/induced in the liver).

4. Conclusion

Alkenylbenzenes include several genotoxic compounds that are bioactivated by CYPs to form 1'-hydroxy-metabolites, which act as proximate carcinogens that upon further sulfation form reactive metabolites that can damage DNA. However, for few alkenylbenzenes animal carcinogenicity data are available, hampering their inclusion in a risk assessment, as no margin of exposure (margin between BMDL10 of cancer study and estimated human exposure) can be determined. To fill these data gaps, application of NAMs is considered preferred above the performance of new *in vivo* animal studies, including *in silico* molecular dynamics studies as done in the present work. In this respect, this work addressed safrole, a well-known genotoxic alkenylbenzene for which a cancer study is available, and a series of poorly characterised structural analogues, i.e. apiole, dillapiole and myristicin. Our results described the low capability of apiole and dillapiole to being transformed into their 1'hydroxy metabolites by CYP1A1 and CYP2A6, in line with the limited *in vitro* kinetic and toxicity data available so far describing a lower efficiency of bioactivation (no data available for dillapiole) and limited DNA adducts formation compared to safrole. Taken together, these results may suggest that apiole and dillapiole are of lower concern for their toxicity, with a low priority in further risk assessment studies. In addition, safrole was shown in the present study as potentially being biotransformed to 1'-hydroxy-safrole by CYP1A1. Of note, CYP1A1 has been considered not relevant for safrole bioactivation as it was considered for long expressed mainly in extra-hepatic tissues. However, recent evidence has reported that CYP1A1 can be efficiently expressed and induced also in liver and safrole may act as an effective inducer. Taken together these data, along with the outcome collected in this work, point to the need of duly assessing CYP1A1 for its role in the formation of 1-hydroxy-metabolites of safrole and other alkenylbenzenes to obtain more insights into the roles of different CYPs in their bioactivation.

As a general comment, analogies in the TK of compounds meant to be included in group assessment are pivotal for being actually grouped. Based on the differences described in this work, further analysis to better describe the CYPs involved in the bioactivation of alkenylbenzenes are advisable toward a more informed background for decision-making in that sense.

Regarding the use of NAMs in biotransformation studies, *in silico* 3D molecular modelling approaches, like the one presented here, already succeeded to study the biotransformation of small molecules by CYPs providing useful means to discriminate substrates from non-substrates [16, 17, 21]. Therefore, 3D modelling may provide a self-standing and effective first-line analytical framework for the systematic analysis of CYP-related biotransformation. Assessing the role of specific CYPs in biotransformation reactions is useful to guide further dedicated *in vitro* investigations with respect to their full quantitative kinetic profiles (K_m and V_{max} of most relevant reactions), which are required input parameters for physiologically based kinetic (PBK) modelling.

References

1. Eisenreich, A., et al., *Alkenylbenzenes in Foods: Aspects Impeding the Evaluation of Adverse Health Effects.* Foods, 2021. **10**(9).

2. Bampidis, V., et al., *Safety and efficacy of a feed additive consisting of a tincture from the bark of Cinnamomum verum J. Presl (cinnamon tincture) for use in all animal species (FEFANA asbl).* Efsa Journal, 2021. **19**(12).

3. Groh, I.A.M., et al., *Genotoxic potential of methyleugenol and selected methyleugenol metabolites in cultured Chinese hamster V79 cells.* Food & Function, 2012. **3**(4): p. 428-436.

4. Atkinson, R.G., *Phenylpropenes: Occurrence, Distribution, and Biosynthesis in Fruit.* Journal of Agricultural and Food Chemistry, 2018. **66**(10): p. 2259-2272.

5. Jeurissen, S.M.F., et al., *Human cytochrome p450 enzyme specificity for the bioactivation of estragole and related alkenylbenzenes.* Chemical Research in Toxicology, 2007. **20**(5): p. 798-806.

6. Rietjens, I., et al., *Flavonoids and alkenylbenzenes: Mechanisms of mutagenic action and carcinogenic risk.* Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis, 2005. **574**(1-2): p. 124-138.

7. Jin, M., et al., *Comprehensive toxicity study of safrole using a medium-term animal model with gpt delta rats.* Toxicology, 2011. **290**(2-3): p. 312-321.

8. Dosoky, N.S. and W.N. Setzer, *Maternal Reproductive Toxicity of Some Essential Oils and Their Constituents.* International Journal of Molecular Sciences, 2021. **22**(5).

9. Gotz, M.E., et al., *Myristicin and Elemicin: Potentially Toxic Alkenylbenzenes in Food.* Foods, 2022. **11**(13).

10. Zhou, G.D., et al., *DNA adducts from alkoxyallylbenzene herb and spice constituents in cultured human (HepG2) cells.* Environmental and Molecular Mutagenesis, 2007. **48**(9): p. 715-721.

11. Zhu, X., et al., *Metabolic Activation of Myristicin and Its Role in Cellular Toxicity.* Journal of Agricultural and Food Chemistry, 2019. **67**(15): p. 4328-4336.

12. Seneme, E.F., et al., *Pharmacological and Therapeutic Potential of Myristicin: A Literature Review.* Molecules, 2021. **26**(19).

13. van den Berg, S., et al., *Safety assessment of plant food supplements (PFS).* Food & Function, 2011. **2**(12): p. 760-768.

14. Hausner, E. and R.H. Poppenga, *Hazards Associated with the Use of Herbal and Other Natural Products*, in *Small Animal Toxicology*, M.E. Peterson and P.A. Talcott, Editors. 2013. p. 335-356.

15. More, S.J., et al., *Guidance on harmonised methodologies for human health, animal health and ecological risk assessment of combined exposure to multiple chemicals.* Efsa Journal, 2019. **17**(3).

16. Dorne, J., et al., *A Computational Understanding of Inter-Individual Variability in CYP2D6 Activity to Investigate the Impact of Missense Mutations on Ochratoxin A Metabolism.* Toxins, 2022. **14**(3).

17. Pedroni, L., et al., *A Computational Inter-Species Study on Safrole Phase I Metabolism-Dependent Bioactivation: A Mechanistic Insight into the Study of Possible Differences among Species.* Toxins, 2023. **15**(2): p. 94.

18. Qi, Z., G.W. Miller, and E.O. Voit, *Rotenone and paraquat perturb dopamine metabolism: A computational analysis of pesticide toxicity.* Toxicology, 2014. **315**: p. 92-+.

19. Vinken, M., et al., *Safer chemicals using less animals: kick-off of the European ONTOX project.* Toxicology, 2021. **458**.

20. Rosell-Hidalgo, A., A.L. Moore, and T. Ghafourian, *Prediction of drug-induced mitochondrial dysfunction using succinate-cytochrome c reductase activity, QSAR and molecular docking.* Toxicology, 2022. **28**(485): p. 153412.

21. Itoh, T., et al., *A 3D Model of CYP1B1 Explains the Dominant 4-Hydroxylation of Estradiol.* Journal of Chemical Information and Modeling, 2010. **50**(6): p. 1173-1178.

22. Sridhar, J., et al., *Review of Ligand Specificity Factors for CYP1A Subfamily Enzymes from Molecular Modeling Studies Reported to-Date.* Molecules, 2017. **22**(7).

23. Lewis, D.F.V., et al., *Molecular modelling of CYP1 family enzymes CYP1A1, CYP1A2, CYP1A6 and CYP1B1 based on sequence homology with CYP102.* Toxicology, 1999. **139**(1-2): p. 53-79.

24. Kim, S., et al., *PubChem in 2021: new data content and improved web interfaces.* Nucleic Acids Research, 2021. **49**(D1): p. D1388-D1395.

25. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Research, 2000. **28**(1): p. 235-242.

26. Walsh, A.A., G.D. Szklarz, and E.E. Scott, *Human Cytochrome P450 1A1 Structure and Utility in Understanding Drug and Xenobiotic Metabolism.* Journal of Biological Chemistry, 2013. **288**(18): p. 12932-12943.

27. Sansen, S., et al., *Structural insight into the altered substrate specificity of human cytochrome P450 2A6 mutants.* Archives of Biochemistry and Biophysics, 2007. **464**(2): p. 197-206.

28. Pettersen, E.F., et al., *UCSF chimera - A visualization system for exploratory research and analysis.* Journal of Computational Chemistry, 2004. **25**(13): p. 1605-1612.

29. Louisse, J., J.L.C.M. Dorne, and L. Dellafiora, *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. Toxycology Letters, 2022. **355**: p. 19-30.

30. Dellafiora, L., et al., *An in silico structural approach to characterize human and rainbow trout estrogenicity of mycotoxins: Proof of concept study using zearalenone and alternariol.* Food Chemistry, 2020. **312**.

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

32. Pedroni, L., et al., *In silico study on the Hepatitis E virus RNA Helicase and its inhibition by silvestrol, rocaglamide and other flavagline compounds.* Scientific Reports, 2022. **12**(1).

33. Abraham, M.J., et al., *GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers.* SoftwareX, 2015. **1-2**: p. 19-25.

34. Best, R.B., et al., *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, 2012. **8**(9): p. 3257-3273.

35. Panneerselvam, S., et al., *A Combined Molecular Docking/Dynamics Approach to Probe the Binding Mode of Cancer Drugs with Cytochrome P450 3A4.* Molecules, 2015. **20**(8): p. 14915-14935.

36. Zhang, L., et al., *Force field development for cofactors in the photosystem II.* Journal of Computational Chemistry, 2012. **33**(25): p. 1969-1980.

37. Del Favero, G., et al., *Structural Similarity with Cholesterol Reveals Crucial Insights into Mechanisms Sustaining the Immunomodulatory Activity of the Mycotoxin Alternariol.* Cells, 2020. **9**(4).

38. Lang, D., M. Radtke, and M. Bairlein, *Highly Variable Expression of CYP1A1 in Human Liver and Impact on Pharmacokinetics of Riociguat and Granisetron in Humans.* Chemical Research in Toxicology, 2019. **32**(6): p. 1115-1122.

39. Martati, E., et al., *Physiologically Based Biokinetic (PBBK) Modeling of Safrole Bioactivation and Detoxification in Humans as Compared With Rats.* Toxicological Sciences, 2012. **128**(2): p. 301-316.

40. Al-Malahmeh, A.J., et al., *Physiologically based kinetic modeling of the bioactivation of myristicin.* Archives of Toxicology, 2017. **91**(2): p. 713-734.

41. Alajlouni, A.M., et al., *Mode of action based risk assessment of the botanical food-borne alkenylbenzene apiol from parsley using physiologically based kinetic (PBK) modelling and read-across from safrole.* Food and Chemical Toxicology, 2016. **89**: p. 138-150.

42. Ueng, Y.F., et al., *Identification of the main human cytochrome P450 enzymes involved in safrole 1 '-hydroxylation.* Chemical Research in Toxicology, 2004. **17**(8): p. 1151-1156.

43. Hu, W.Y., et al., *Induction of Cyp1a1 is a nonspecific biomarker of aryl hydrocarbon receptor activation: Results of large scale screening of pharmaceuticals and toxicants in vivo and in vitro.* Molecular Pharmacology, 2007. **71**(6): p. 1475-1486.

Chapter III

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

This chapter has been published as:

Pedroni Lorenzo, Perugino Florinda, Kurtaga Ambra, Galaverna Gianni, Dall'Asta Chiara, Dellafiora Luca. The bitter side of toxicity: A big data analysis spotted the interaction between trichothecenes and bitter receptors. Food Res Int. 2023 Nov;173(Pt 1):113284. doi: 10.1016/j.foodres.2023.113284.

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.

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 [1]. 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 [2]. Germane to humans, about 25 functional TAS2Rs are expressed in many tissues besides the lingual epithelium [3]. 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 [4]. 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 [1, 5, 6].

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 [7] recently made

available in the Protein Data Bank (PDB; https://www.rcsb.org) [8]. Indeed, Xu and co-workers described precisely for the first time the topology of a human TAS2R in the active ligand-bound conformations (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 [9]), 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 [10]. 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) [11] with the support of the VirtualTaste Prediction tool (https://virtualtaste.charite.de/VirtualTaste) [12]. This analysis described trichothecolone, a fungal metabolite primarily produced by *Trichotecium* species [13], 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 [14, 15]. 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 shed more light on their mechanism of actions eventually improving their risk assessment, and more in general its 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 [3, 16, 17]. 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 [18] (Figure 1).

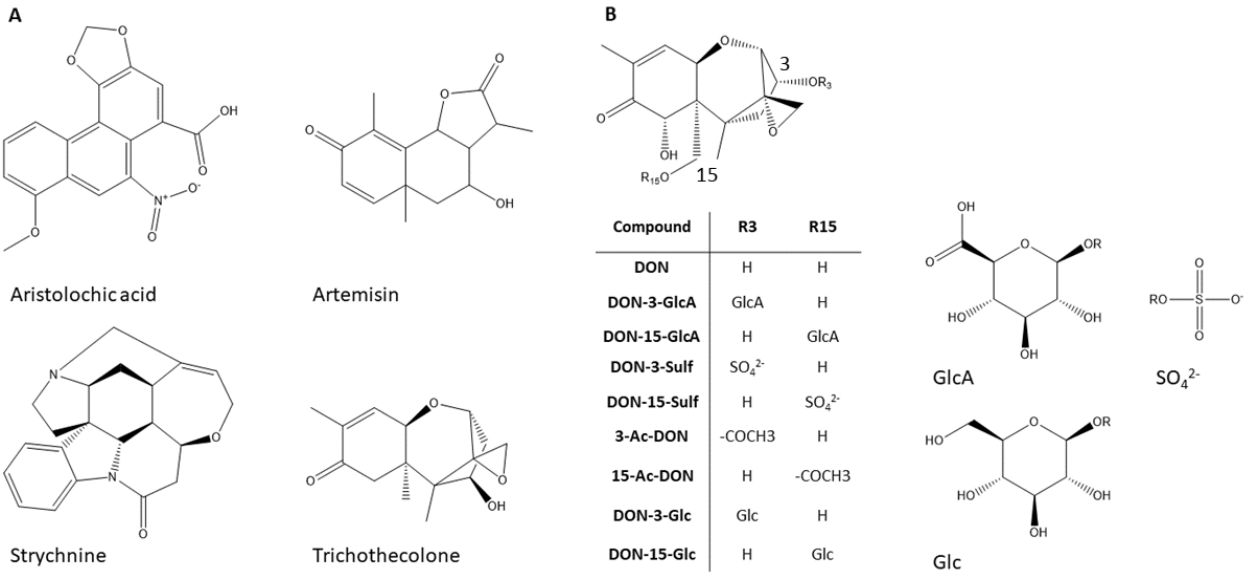


Figure 1*. Chemical structure of compounds analysed via molecular modelling.* ***A.*** *Aristolochic acid (CID 2236), artemisin (CID 65030), strychnine (CID 441071) and trichothecolone (CID 107974).* ***B.*** *DON scaffold with its chemical substituent in position 3 (R3) and 15 (R15).*

This integrated analysis aimed to distinguish TAS2R46's binders from non-binders, in agreement with previous studies [10, 19]. 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.

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 respectively retrieved in the JSON format from the publicly available database of flavour molecules FlavorDB (https://cosylab.iiitd.edu.in/flavordb; 25 595 entries as of 17th April 2023) [20] and in the SMILE format from the T3DB database (http://www.t3db.ca; 3 533 entries as of 17th April 2023) [11]. Both databases were scraped using *ad hoc* python scripts, which may be made available upon request. The 3D structure of molecules meant to undergo docking studies and molecular dynamics was downloaded from the PubChem database (https://pubchem.ncbi.nlm.nih.gov) [21] 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) [22].

2.1.2 Protein structure

The model of Taste Receptors Type 2 Member 46 (TAS2R46) was derived from the strychnine-binding cryo-EM structure deposited in the Protein Data Bank (PDB; https://www.rcsb.org) with PDB code 7XP6 [7]. Specifically, the model was derived from the chain E after completing its non-terminal

missing parts (residues 157-172) through Modeller (version 10.2) [23] implemented in UCSF Chimera (version 1.16) [22], 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 [24].

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 [25]. 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/) [11] as either bitter or not bitter. Moreover, the Prediction tool of VirtualTaste (https://virtualtaste.charite.de/VirtualTaste) [12], a webserver for the prediction of organoleptic properties of chemical compounds, 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 the molecules 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 event [24, 26]. 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 [19]. 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) [27].

2.4 Molecular dynamics simulation

Molecular dynamics (MD) simulations were performed through GROMACS (version 2019.4) [28] to verify the ligands potential to activate the TAS2R46. The whole system was parametrised with the CHARMM27 all-atom force field [29] and specifically each ligand parametrisation was performed on the SwissParam webserver (https://www.swissparam.ch/) [30]. 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/) [20] 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) [25], 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) [11]. 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 [31], 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) [11]. 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/) [12]. 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.

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

Name	Pubchem CID	VirtualTaste Confidence Score [1]	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

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) [32]. [14, 15, 33].

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 [34]. 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 [16, 35, 36]. 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) [16, 36]. 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 [3, 16, 17]. Further analysis are needed to provide final proofs, however a possible functional corelation 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 [37]. 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).

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 [19, 24]. 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) [7]. 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 [32], 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 [19, 24]. In this respect, the molecular dynamics simulation of TAS2R46-strychnine complex served as a reference since it describes the active topology of TASR46 [7]. 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.

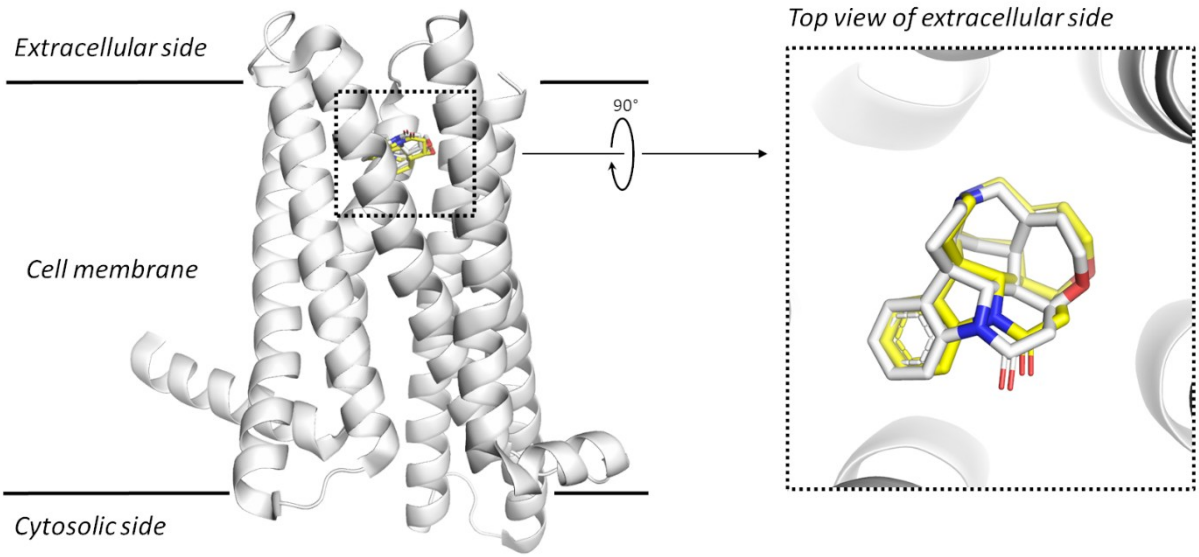


***Figure 2**. Comparison between cryo-EM structure (white coloured) and docking (yellow coloured) pose of strychnine within human TAS2R46. The protein is represented in cartoon, while strychnine in sticks.*

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 [38, 39]. 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 [7], in agreement with the inactivity reported for aristolochic acid [38]. 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 ($p<0.05$) that 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.

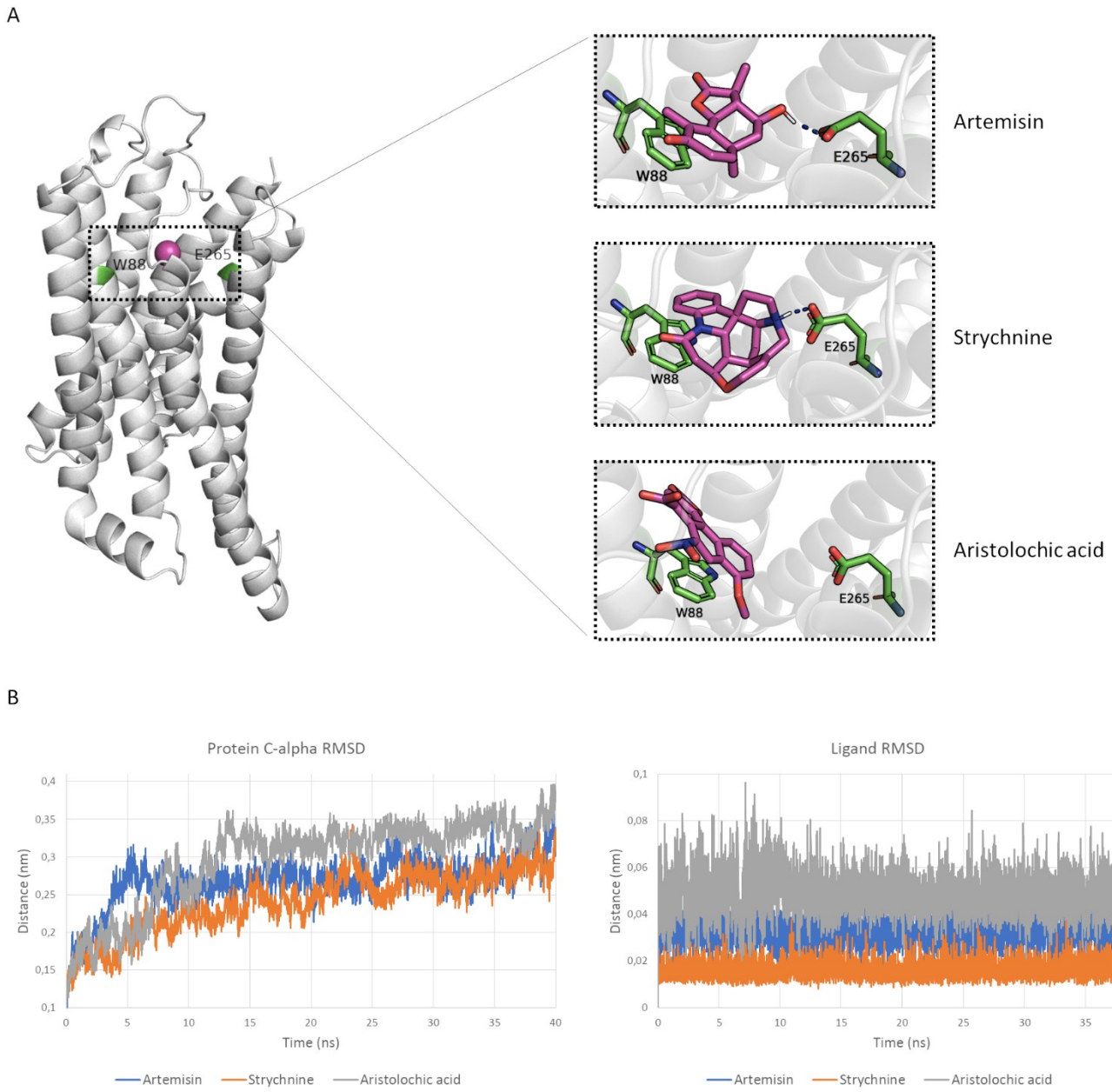


Figure 3. *Molecular modelling results of artemisin, strychnine and aristolochic acid.* ***A.*** *Docking pose of artemisin, strychnine and aristolochic acid. On the left, TAS2R46 is represented in white cartoon with W88 and E265 shown in green stick and the ligand centroid as a magenta sphere. On the 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 [40]. 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 [41]. 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 [42].

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 similar ligand RMSD trend with respect 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 [37] which are present in districts exposed to DON, including but not limited to the gastrointestinal tract [32]. Moreover, the direct involvement of TAS2R46 activation in enteroendocrine cells to release of GLP-1 has been reported [37]. 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, www. https://www.proteinatlas.org, last access 8th May 2023 [9]), and DON proved to permeate the Blood Brain Barrier (BBB) exerting neurotoxicity [43], 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 [44], and TAS2R46 activation has been associated to Bcl-2 decrease in many cell lines [45]. 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 [46] while the 3-Ac-DON ones are likely to be more aggressive towards plants [47]. Indeed, 15-Ac-DON was found with a high incidence in contaminated commodities showing a toxicity comparable to DON [48]. 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* [48]. 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.

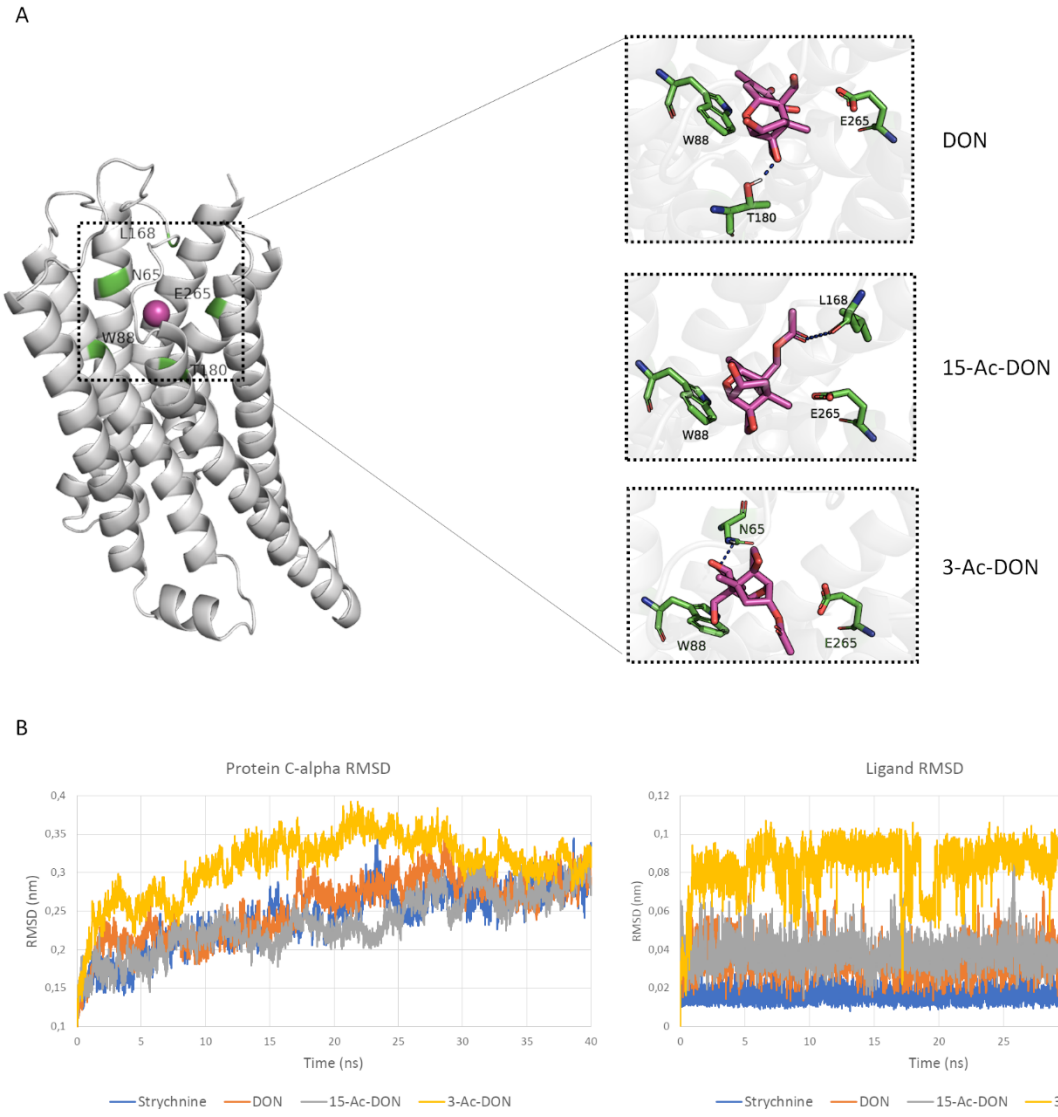


***Figure 4**. Molecular modelling results of DON, 15-Ac-DON and 3-Ac-DON. **A.** Docking pose of DON, 15-Ac-DON and 3-Ac-DON. On the left, TAS2R46 is represented in white cartoon with the ligand binding site centroid as a magenta sphere. On the right, close-up of the ligand binding site with amino acid 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).*

DON Sulphates. The protein RMSD trend of DON3- 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 TAS246 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 [49]. 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 [9]) and the exposure of urogenital tissues to those forms (being the main DON metabolites excreted via urine).

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 of 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 toxicity.

References

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

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

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

4. Lu, P., et al., *Extraoral bitter taste receptors in health and disease.* Journal of General Physiology, 2017. 149(2): p. 181-197.

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

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

7. Xu, W.X., et al., *Structural basis for strychnine activation of human bitter taste receptor TAS2R46.* Science, 2022. 377(6612): p. 1298-1303.

8. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Research, 2000. 28(1): p. 235-242.

9. Uhlen, M., et al., *Tissue-based map of the human proteome.* Science, 2015. 347(6220).

10. Del Favero, G., et al., *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, 2022. 159.

11. Lim, E., et al., *T3DB: a comprehensively annotated database of common toxins and their targets.* Nucleic Acids Research, 2010. 38: p. D781-D786.

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

13. Konishi, K., et al., *Cancer preventive potential of trichothecenes from Trichothecium roseum.* Bioorganic & Medicinal Chemistry, 2003. 11(12): p. 2511-2518.

14. Tolosa, J., et al., *Multi-mycotoxin occurrence in feed, metabolism and carry-over to animal-derived food products: A review.* Food and Chemical Toxicology, 2021. 158.

15. Khaneghah, A.M., et al., *A systematic review and meta-analysis to investigate the concentration and prevalence of trichothecenes in the cereal-based food.* Trends in Food Science & Technology, 2020. 102: p. 193-202.

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

17. Jia, H., et al., *Substance P and Glucagon-like Peptide-1(7-36) Amide Mediate Anorexic Responses to Trichothecene Deoxynivalenol and Its Congeners.* Toxins, 2022. 14(12).

18. Warth, B., et al., *Deoxynivalenol-sulfates: identification and quantification of novel conjugated (masked) mycotoxins in wheat.* Analytical and Bioanalytical Chemistry, 2015. 407(4): p. 1033-1039.

19. Dellafiora, L., et al., *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, 2022. 162.

20. Garg, N., et al., *FlavorDB: a database of flavor molecules.* Nucleic Acids Research, 2018. 46(D1): p. D1210-D1216.

21. Kim, S., et al., *PubChem 2023 update.* Nucleic Acids Research, 2022.

22. Pettersen, E.F., et al., *UCSF chimera - A visualization system for exploratory research and analysis.* Journal of Computational Chemistry, 2004. 25(13): p. 1605-1612.

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

24. Dellafiora, L., et al., *An in silico structural approach to characterize human and rainbow trout estrogenicity of mycotoxins: Proof of concept study using zearalenone and alternariol.* Food Chemistry, 2020. 312.

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

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

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

28. Abraham, M., et al., *GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers.* SoftwareX, 2015. 1–2: p. 19–25.

29. Best, R.B., et al., *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, 2012. 8(9): p. 3257-3273.

30. Zoete, V., et al., *SwissParam: A Fast Force Field Generation Tool for Small Organic Molecules.* Journal of Computational Chemistry, 2011. 32(11): p. 2359-2368.

31. Hosmer, D.W., S. Lemeshow, and R.X. Sturdivant, *Applied Logistic Regression, 3rd Edition*, in *Applied Logistic Regression, 3rd Edition*. 2013. p. 1-500.

32. Sun, Y., et al., *Toxicokinetics and metabolism of deoxynivalenol in animals and humans.* Archives of Toxicology, 2022. 96(10): p. 2639-2654.

33. Gallo, M., et al., *Relationships between food and diseases: What to know to ensure food safety.* Food Research International, 2020. 137.

34. Hou, S.L., et al., *The toxicity mechanisms of DON to humans and animals and potential biological treatment strategies.* Critical Reviews in Food Science and Nutrition, 2023. 63(6): p. 790-812.

35. Patience, J.F., et al., *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, 2014. 92(2): p. 620-626.

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

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

38. Brockhoff, A., et al., *Structural requirements of bitter taste receptor activation.* Proceedings of the National Academy of Sciences of the United States of America, 2010. 107(24): p. 11110-11115.

39. Brockhoff, A., et al., *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, 2007. 55(15): p. 6236-6243.

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

41. Berthiller, F., et al., *Masked mycotoxins: A review.* Molecular Nutrition & Food Research, 2013. 57(1): p. 165-186.

42. Vidal, A., et al., *Mycotoxin Biomarkers of Exposure: A Comprehensive Review.* Comprehensive Reviews in Food Science and Food Safety, 2018. 17(5): p. 1127-1155.

43. Zhang, J.J., et al., *The neurotoxicity of trichothecenes T-2 toxin and deoxynivalenol (DON): Current status and future perspectives.* Food and Chemical Toxicology, 2020. 145.

44. Wang, X.C., et al., *Deoxynivalenol induces apoptosis in PC12 cells via the mitochondrial pathway.* Environmental Toxicology and Pharmacology, 2016. 43: p. 193-202.

45. Costa, A.R., et al., *Bitter taste signaling in cancer.* Life Sciences, 2023. 315.

46. Crippin, T., et al., *Comparing genotype and chemotype of Fusarium graminearum from cereals in Ontario, Canada.* Plos One, 2019. 14(5).

47. Puri, K.D. and S.B. Zhong, *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, 2010. 100(10): p. 1007-1014.

48. Knutsen, H.K., et al., *Risks to human and animal health related to the presence of deoxynivalenol and its acetylated and modified forms in food and feed.* Efsa Journal, 2017. 15(9).

Supplementary Materials

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
Trichotecholone	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)*

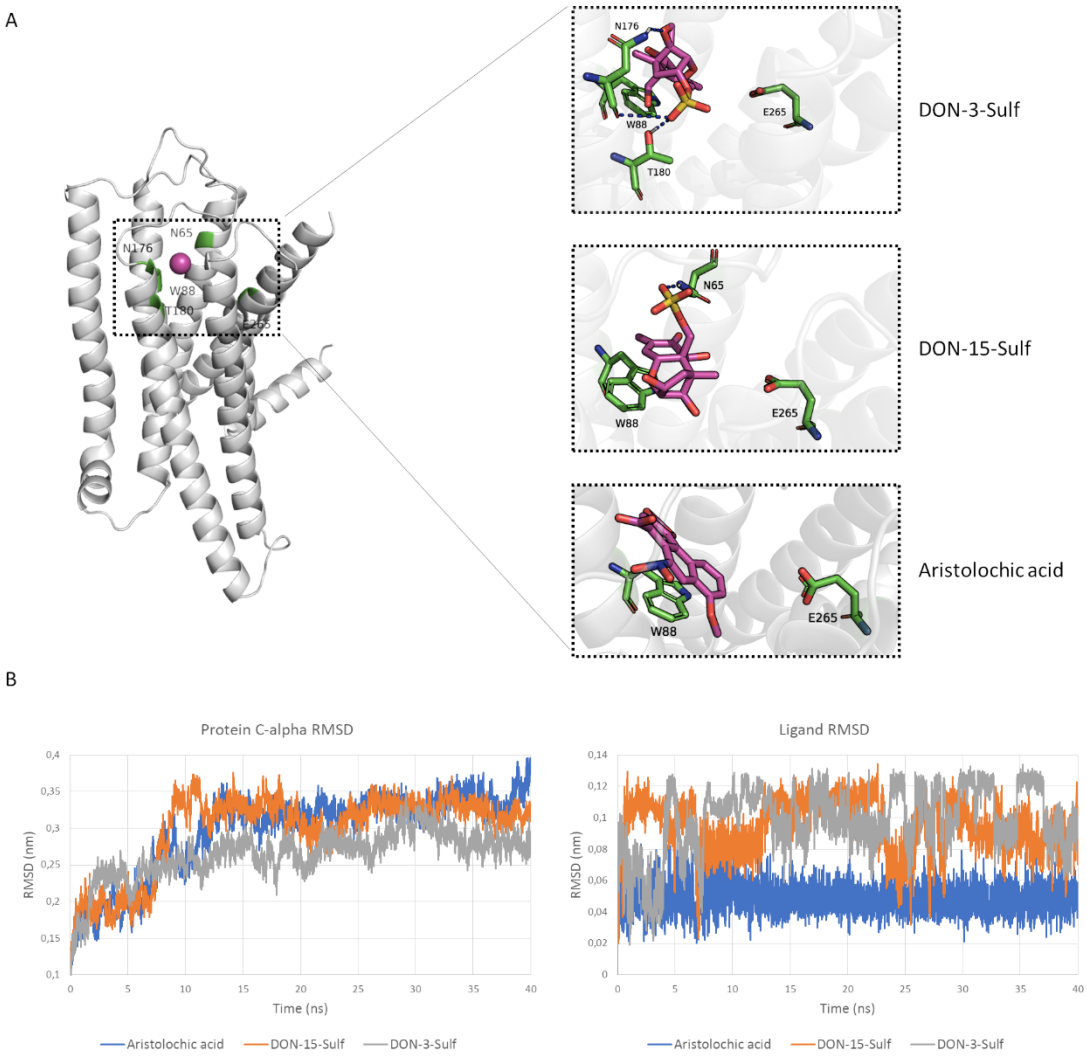


Figure S1. *Molecular modelling results of DON sulphates.* ***A.*** *Docking pose of DON-3 and –15-Sulf and aristolochic acid. On the left, TAS2R46 is represented in white cartoon. On the right, close-up of the ligand binding site with residues side chains and ligands represented as sticks. Blu dashed 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).*

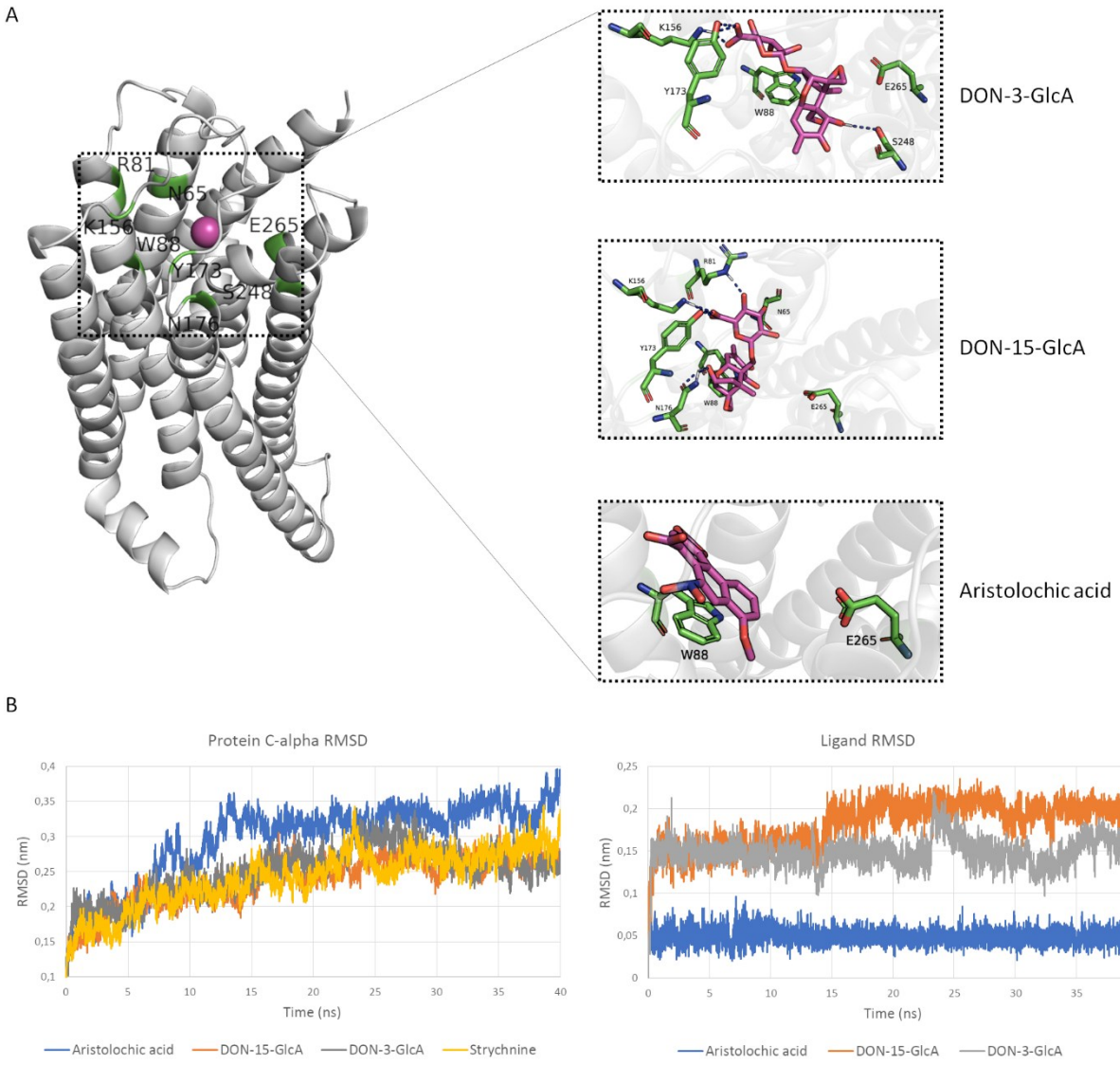


Figure S2. *Molecular modelling results of DON glucuronides.* ***A.*** *Docking pose of DON-3- and –15-GlcA and aristolochic acid. On the left, TAS2R46 is represented in white cartoon. On the right, close-up of the ligand binding site with amino acid side chains and ligands represented as sticks. Blu 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).*

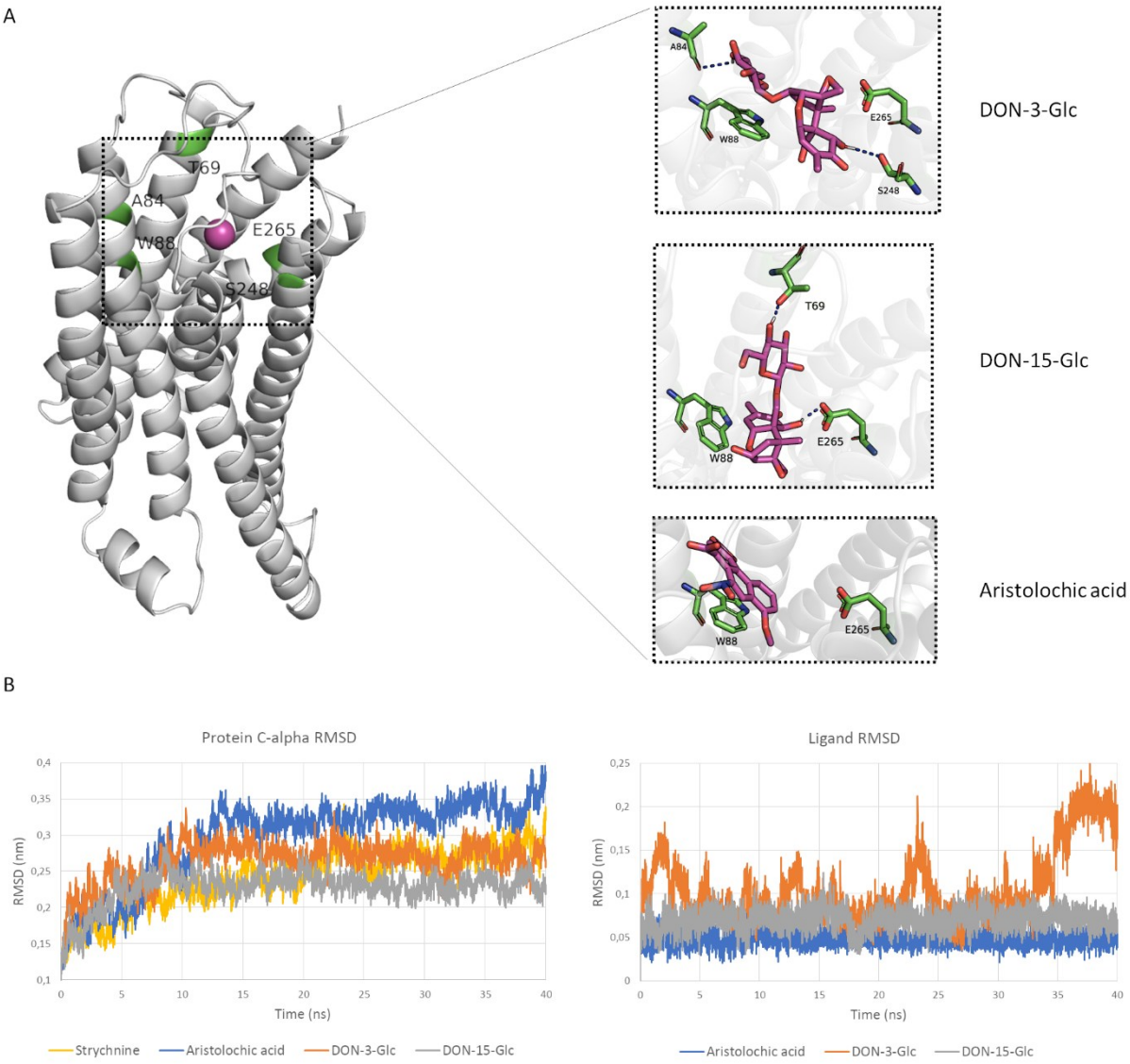


Figure S3. *Molecular modelling results of DON glucosides.* ***A.*** *Docking pose of DON-3- and –15-Glc and aristolochic acid. On the left, TAS2R46 is represented in white cartoon. On the right, close-up of the ligand binding site with amino acid side chains and ligands represented as sticks. Blu dashed 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).*

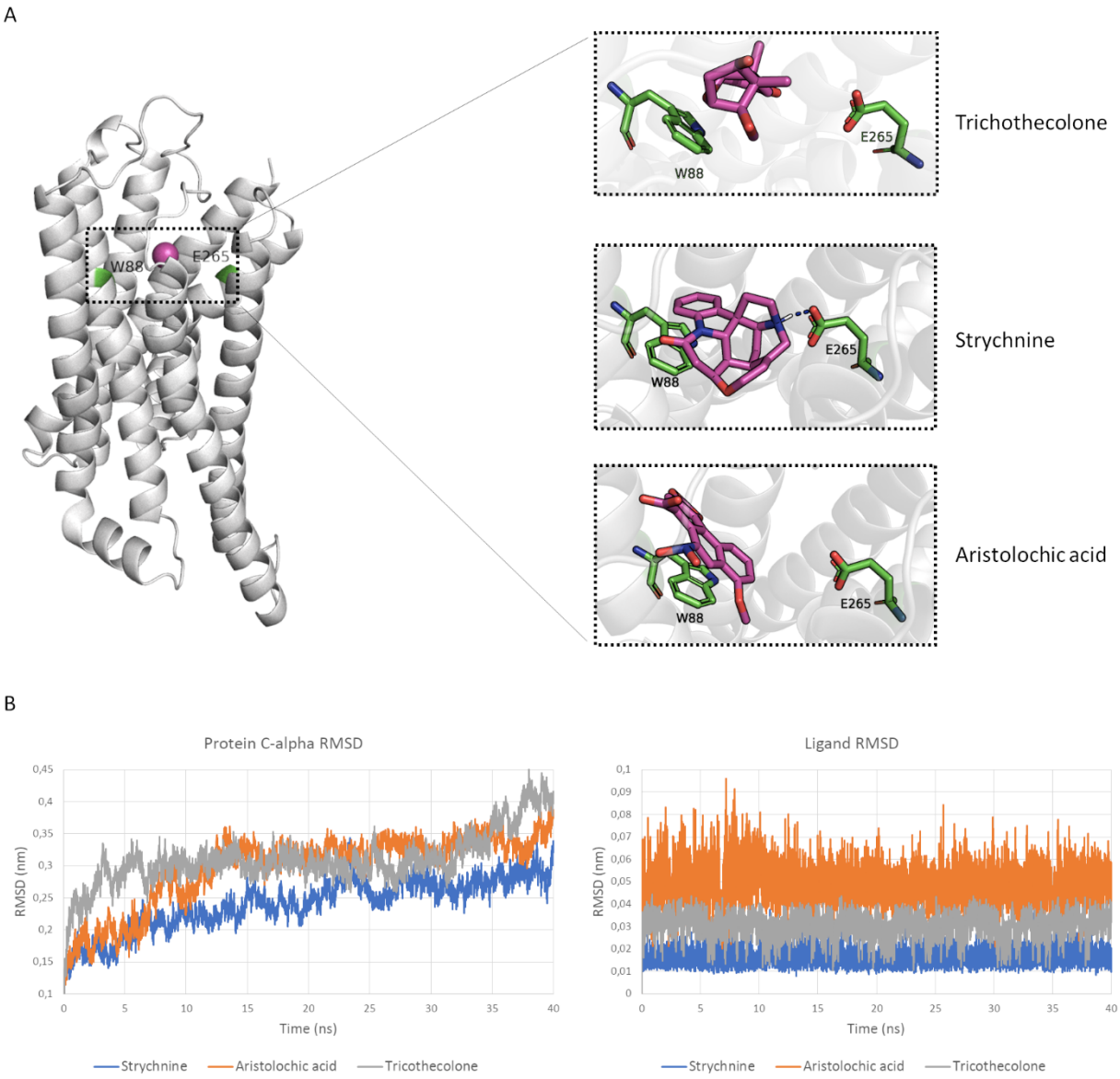


Figure S4. *Molecular modelling results of trichothecolone, strychnine and aristolochic acid.* ***A.*** *Docking pose of trichothecolone, strychnine and aristolochic acid. On the left, TAS2R46 is represented in white cartoon. On the right, close-up of the ligand binding site with amino acid side chains 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 IV

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

This chapter has been published as: Perugino Florinda, **Pedroni Lorenzo**, Galaverna Gianni, Dall'Asta Chiara, Dellafiora Luca. A mechanistic toxicology study to grasp the mechanics of zearalenone estrogenicity: Spotlighting aromatase and the effects of its genetic variability. Toxicology. 2024 Jan;501:153686. doi: 10.1016/j.tox.2023.153686.

Abstract

Zearalenone 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, zearalenone 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 zearalenone. It has been also calculated the inhibition capability of α-zearalenol, the most prominent and estrogenically potent phase I metabolite of zearalenone, 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 zearalenone 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 zearalenone; iii) prioritized aromatase variants for future investigations toward a better comprehension of ZEN xenoestrogenicity at an individual level.

1. Introduction

Zearalenone (ZEN; Figure 1) is a mycotoxin produced by fungi belonging to the *Fusarium* genus (mainly *F. culmorum* and *F. graminearum*) [1, 2] and it is among those most relevant to food safety due to its high incidence in food and feed commodities and toxicity [3, 4]. 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 [5, 6].

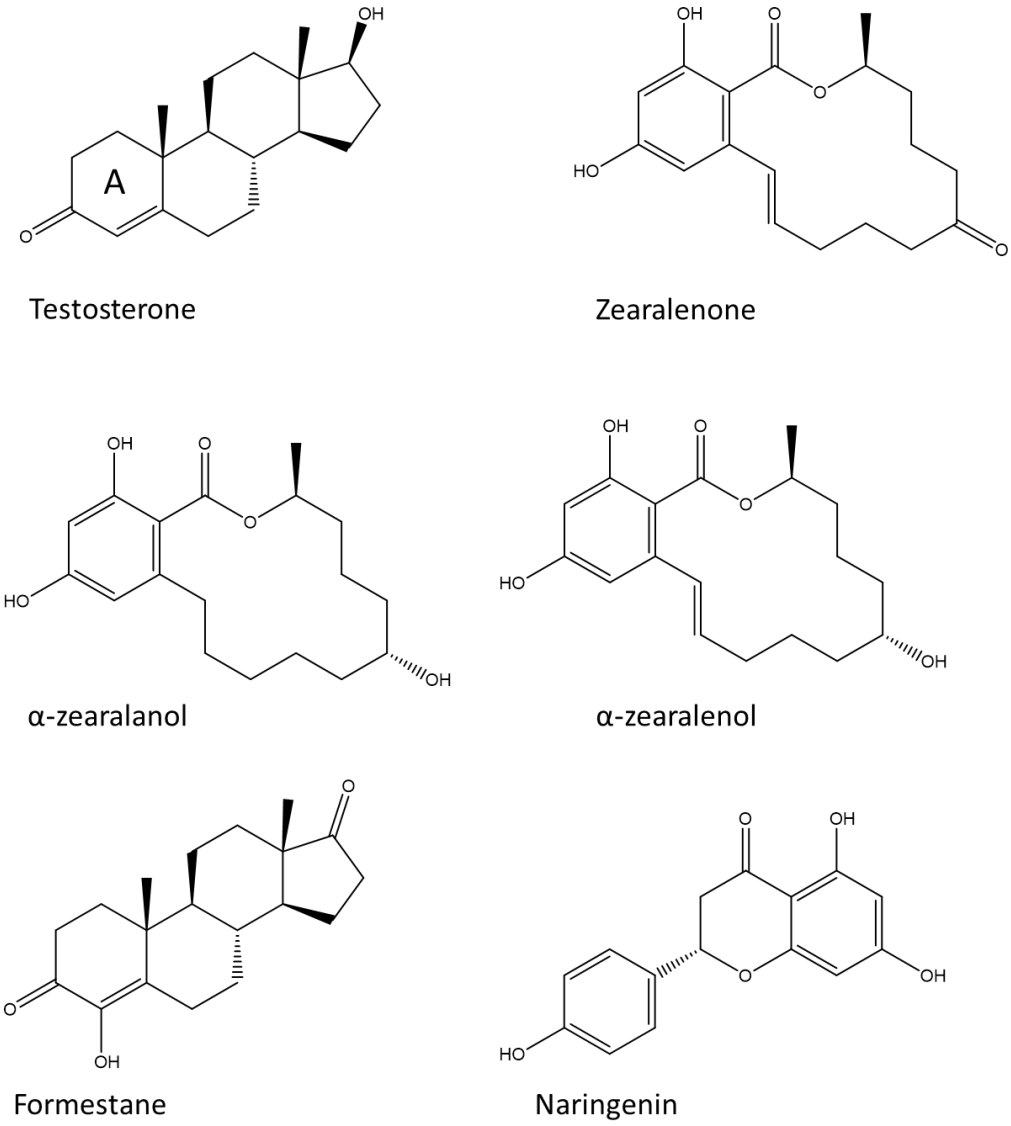


Figure 1. *Chemical structure of molecules under analysis.*

ZEN has shown an endocrine disrupting activity *in vivo* as animal studies reported its capability to impair hormonal balance [7]. 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 [8-11]. It was also hypothesized that ZEN can be associated with cancer development/progression [12-15], though it has been defined as not classifiable as carcinogenic to humans due to the limited evidence in experimental animals [12, 16]. Germane to the mechanisms of toxicity, ZEN estrogenicity may rely on its concerted capability to activate estrogen receptors (ERs) and inhibit aromatase (CYP19A1) [2, 17]. From a structural point of view, ZEN is a resorcyclic acid lactone (molecular formula $C_{18}H_{22}O_5$) 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 [2]. Similarly, the structural analogies to steroids make ZEN suitable to compete also with aromatase substrates/products resulting in its inhibition [17]. This event leads to a ZEN-dependent reduction of E2 production being aromatase responsible for the conversion of androgens to estrogens [18]. The final balance of these two apparently contradictory effects concurs determining 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 [19], but it is also the predominant ZEN phase I metabolite in humans, and it has shown a xenoestrogenic activity way more potent than ZEN [20]. 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 [21]. 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 [17], 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 [22, 23]. 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 [24, 25]. 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) [26] at the time of analysis (i.e. 434 variants; last database access 25th September 2023), focusing on those surrounding the ligand binding pocket (15 variants) 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 [27-29]. 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 [30, 31], 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 to 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.

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) [26] which listed 434 variants at the time of analysis (last database access 25th September 2023). However, the analysis focused on those forming the ligand binding pocket being most likely to affect 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) [32] in the Structure Data File (.sdf) format (CID: 5281576, 5284645, 6013, 439246 and 11273 respectively).

The 3D model of human wild-type aromatase was derived from the crystallographic structure retrieved in the .pdb format from the Protein Data Bank (PDB; https://www.rcsb.org; last database access 24th July 2023) [33] with PDB code 3S79 [34]. 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) [34] was processed with UCSF Chimera software (version 1.16) [35] 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 wild-type structure with Chimera (version 1.16) [35], 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 [31].

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 [36, 37]. 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) [38] and considering non hydrogen atoms only.

2.3 Molecular Dynamics Simulations

Molecular dynamics (MD) simulations were performed through GROMACS (version 2021.4) [39] 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) [40] and the whole system was parametrised with the CHARMM27 all-atom force field [41]. The hydrogen database was modified according to previous works [30, 36, 42, 43] 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 Root Mean Square Deviations (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 [26] 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 who 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).www.uniprot.org) [26] at the time of analysis (last database access 25th September 2023) to make the study technically feasible. Specifically, this investigation targeted the substitutions of amino acids surrounding the ligand binding site, assuming they are most likely to have an impact on aromatase-ZEN interaction. Therefore, the study coped with 15 variants: A306T, D309G, D309N, F221S, M374I, M374L, M374T, R115Q, S478F, T310A, T310I, T310S, V370A, V370M and V373F (Figure 2A).

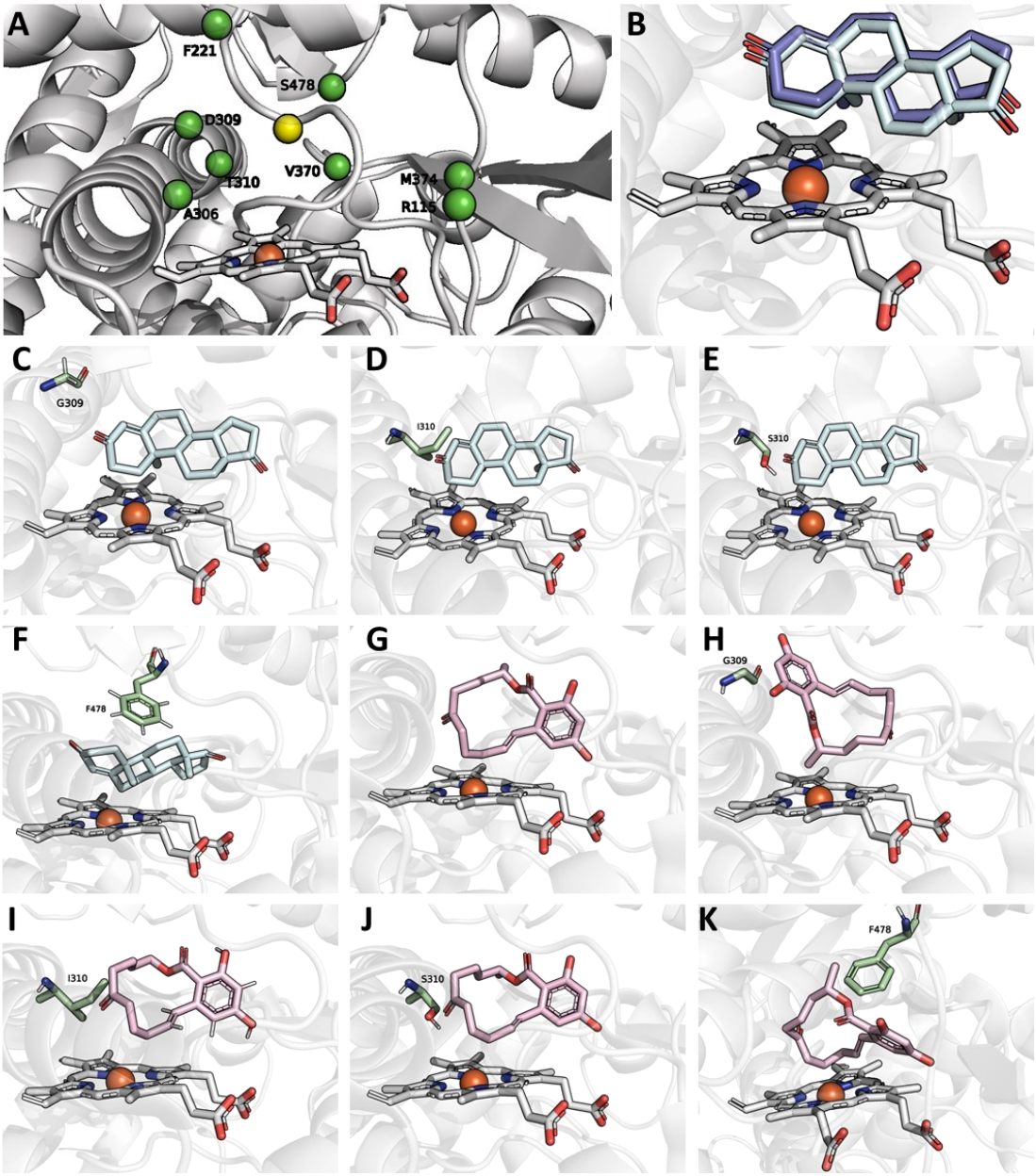


***Figure 2**. Protein binding site and docking poses of TES (pale cyan sticks) and ZEN (pale pink sticks) in WT and in a selection of variants. Heme is represented as white sticks while the protein is represented as transparent white cartoon. The mutated residue in each variant is represented as pale 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.*

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 [30, 31].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 [31, 44]. 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 [24, 45, 46], were analysed combining data from docking and MD simulations. Specifically, it was monitored the geometrical stability of complexes over time using the root-mean-squared-deviations (RMSD) of ligands, proteins and ligands trajectories, and monitoring the number of protein-ligand hydrogen bonds over time, in agreement with previous studies [28, 30].

V370M and D309N have been previously described as inactive aromatase variants [24, 45, 46], while T310S has been described active though more susceptible of being inhibited by NAR compared to the WT enzyme [24]. 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 [47]. To do so and according to previous studies [30], 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, 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 [30], 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, 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 [47, 48], 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 [48, 49]. 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 and the T310S variant, which is a functional variant more susceptible to NAR inhibition compared to the WT enzyme [50]. 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, 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 [28], 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 [24].

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.

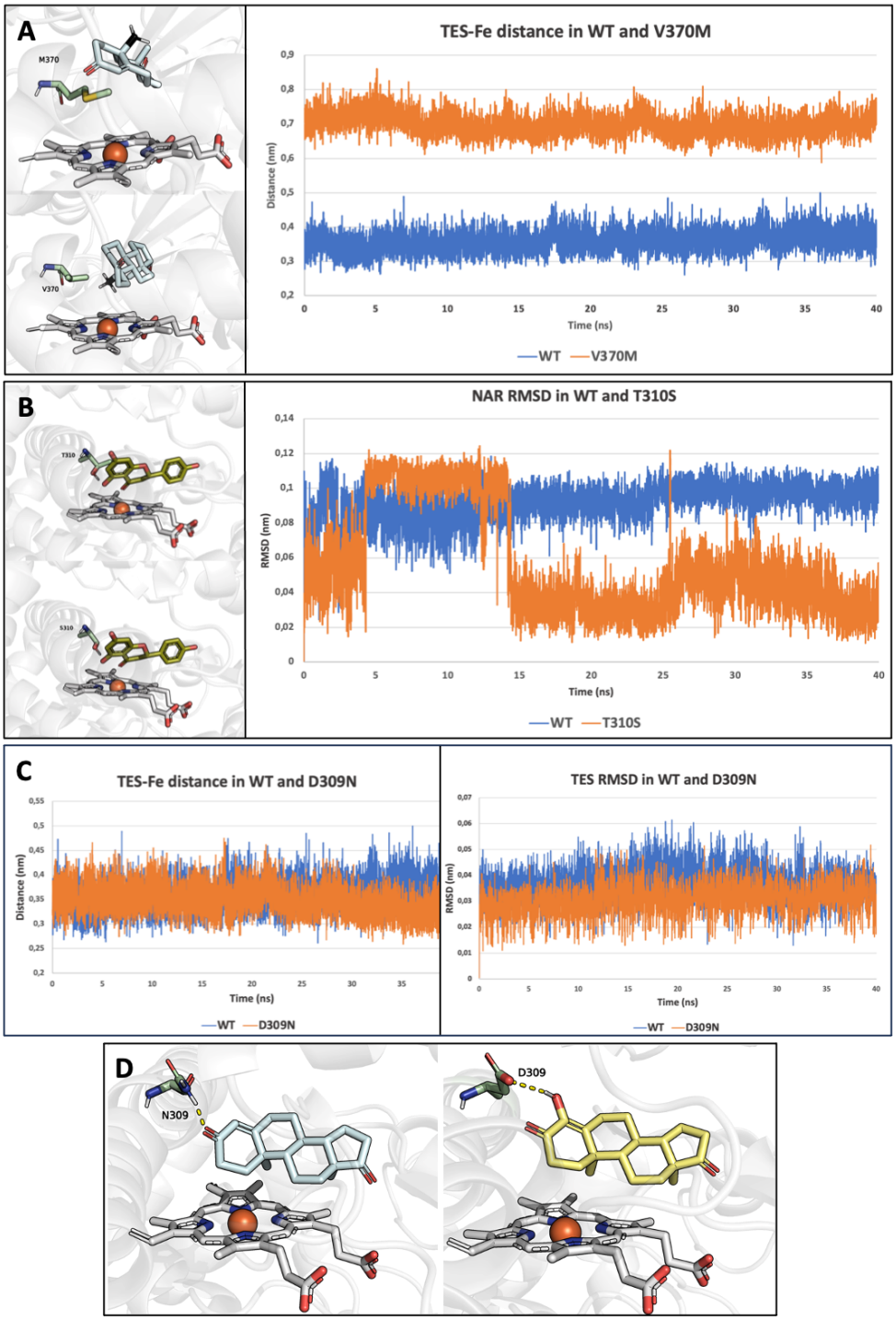


Figure 3. *Model validation. Protein is reported as transparent white cartoon while heme in white sticks. The average distances between the atom undergoing reaction and heme's iron atom ± SD are reported between brackets.* ***A****. Comparison between TES-V370M and TES-WT complex. TES is represented 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 the variations magnitude may reflect the impact of amino acid substitution in the capability to bind ZEN, in agreement with previous studies [30, 31]. 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 [29, 31]. 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 [24].

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.

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.

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 309D 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 D309G this 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.

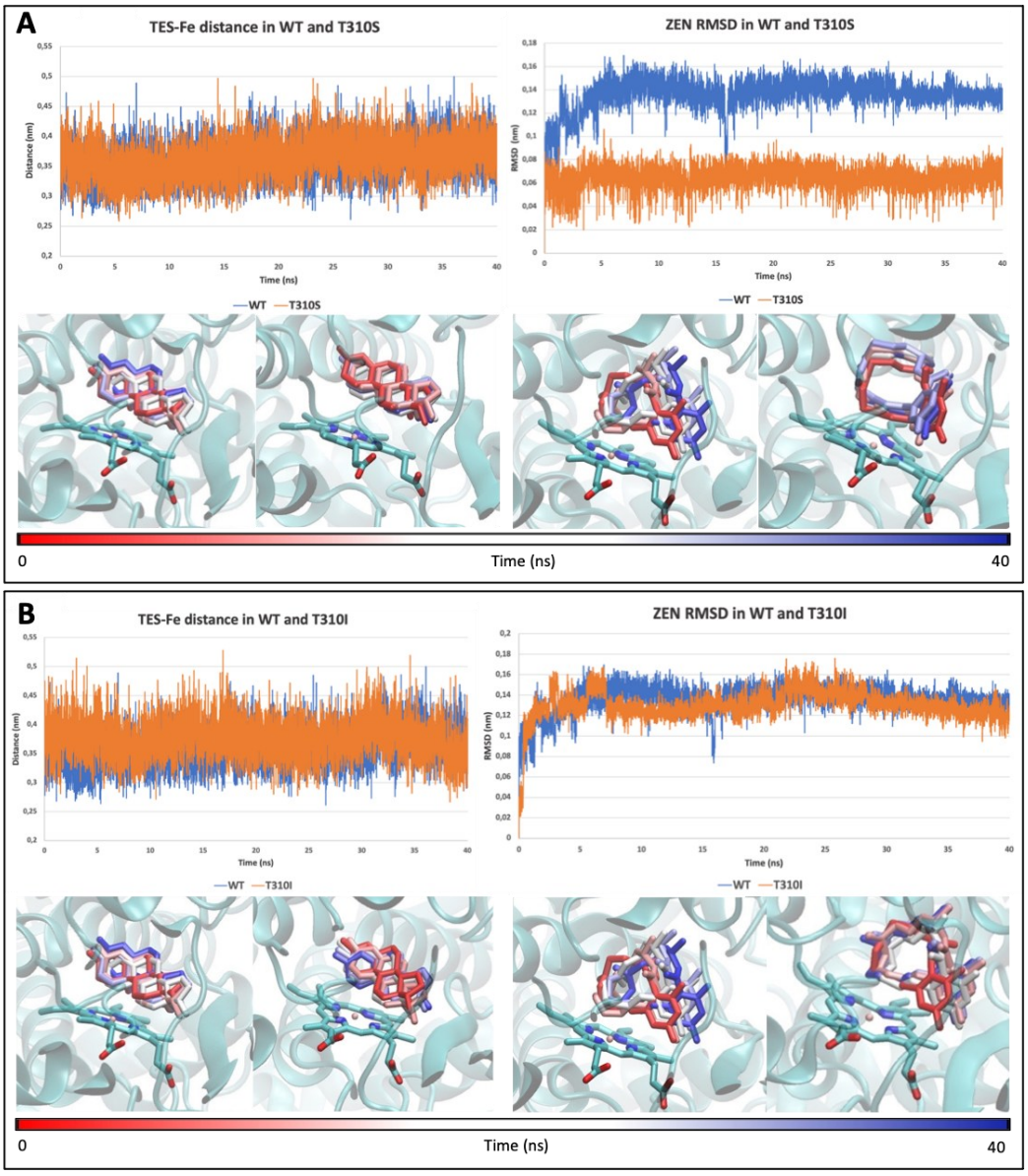


Figure 4. *MD results on T310 variants compared to WT. The average distances between the atom undergoing reaction and heme's iron atom ± SD are reported between brackets.* ***A****. On the top-left, interatomic distances between TES's atom undergoing reaction and heme's iron within WT (blue; 0.359 ± 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.

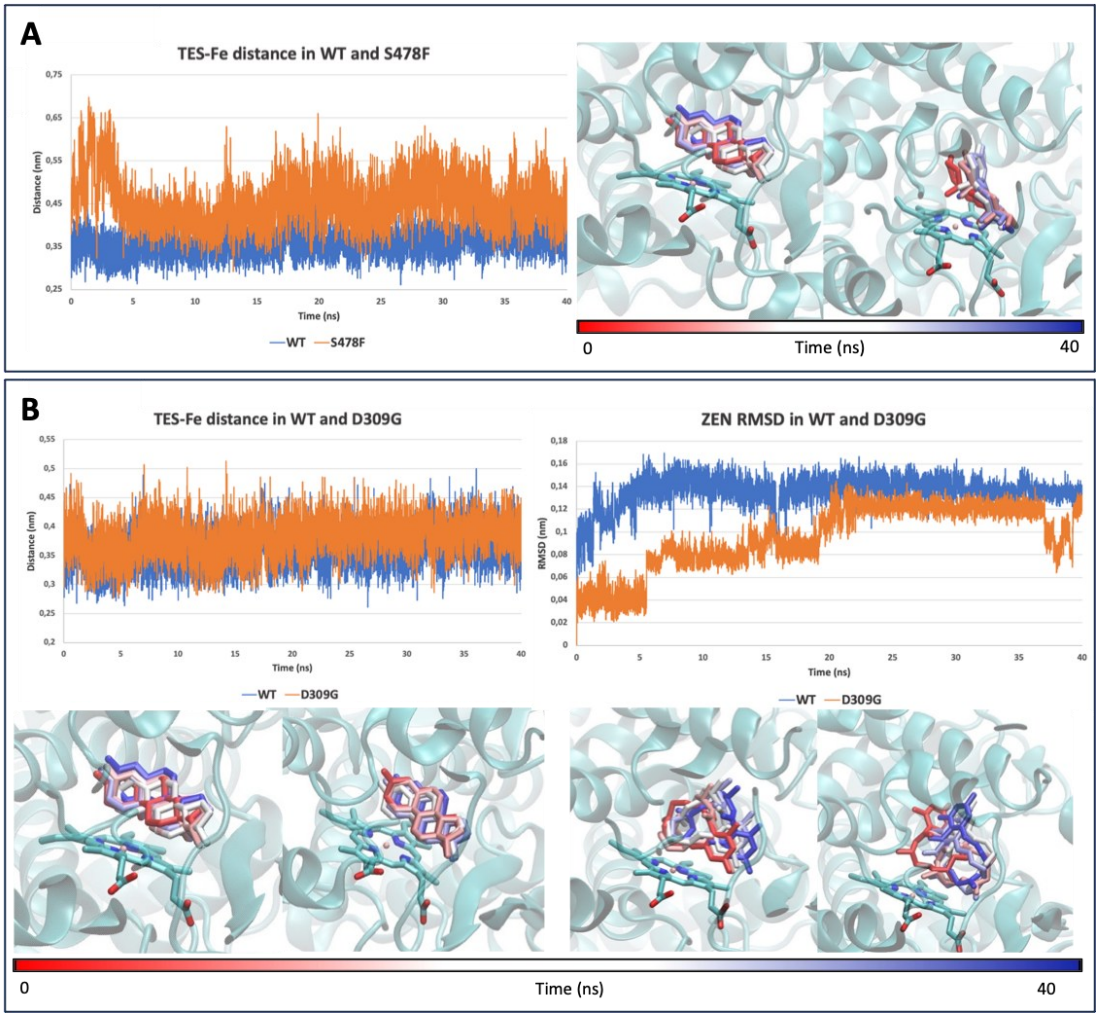


***Figure 5**. MD results on S478F and D309G compared to WT. **A**. On the left, interatomic distances between TES's atom undergoing reaction and heme's iron within WT (blue; average value ± SD of 0.359 ± 0.032 nm) and S478F (orange; average value ± SD of 0.455 ± 0.057 nm). On the right, time-step representation (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 a close analogue of αZEL (αZAL) as able to inhibit aromatase though with a lower efficacy compared to ZEN [17], suggesting a certain degree of inhibitory potential for αZEL as well. Moreover, αZEL is a more potent agonist of ER compared to ZEN [51]. 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 [17]. 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.

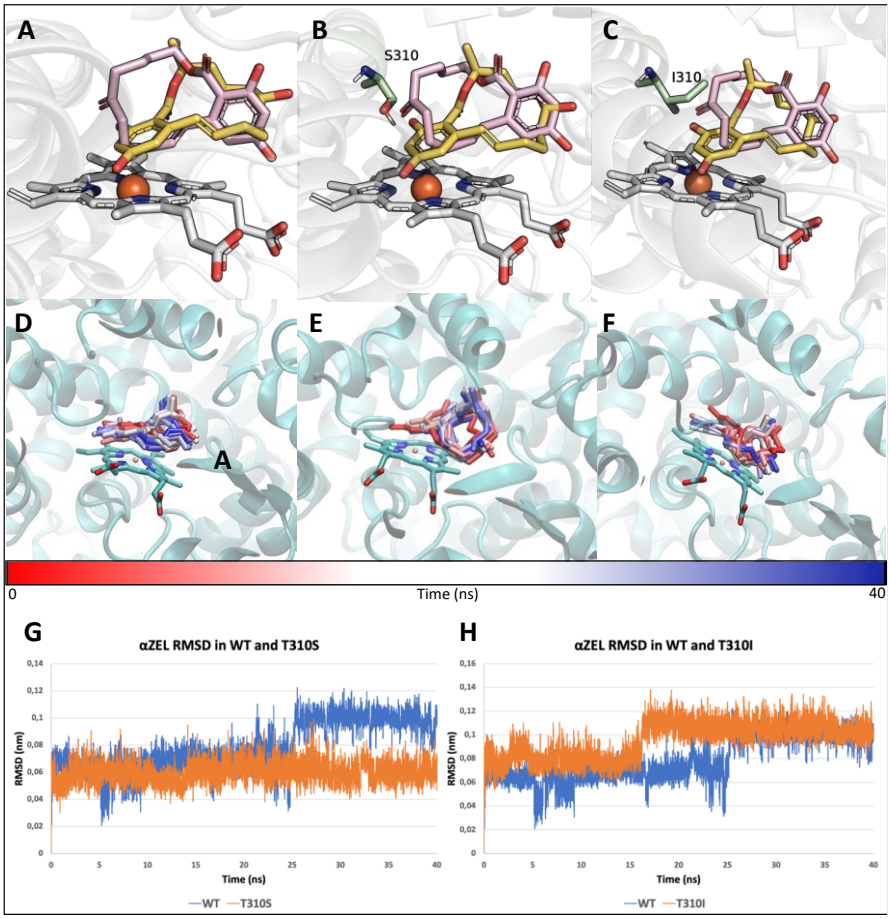


Figure 6. *Molecular docking and MD simulations results of αZEL within T310S and T310I compared to WT.* ***A.*** *Docking pose superimposition of ZEN (pink sticks) and αZEL (yellow sticks) within WT.* ***B.*** *Docking pose superimposition of ZEN (pink sticks) and αZEL (yellow sticks) within T310S.* ***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 [24]. This suggested that those individuals who express 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 [52].

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 prioritized 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

1. Malir, F., et al., Hazard characterisation for significant mycotoxins in food. Mycotoxin Research, 2023.

2. EFSA, Scientific Opinion on the risks for public health related to the presence of zearalenone in food. EFSA Journal 2011. **9**(6): p. 2197.

3. Catteuw, A., et al., 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, 2019. **67**(12): p. 3448-3458.

4. Silva, A.S., et al., UHPLC-ToF-MS method for determination of multi-mycotoxins in maize: Development and validation. Current Research in Food Science, 2019. **1**: p. 1-7.

5. Pfleger, F. and C. Schwake-Anduschus, Relevance of Zearalenone and its modified forms in bakery products. Mycotoxin Research, 2023. **39**(3): p. 153-163.

6. Jing, S.Y., et al., Toxicity of zearalenone and its nutritional intervention by natural products. Food & Function, 2022. **13**(20): p. 10374-10400.

7. Kowalska, K., D.E. Habrowska-Gorczynska, and A.W. Piastowska-Ciesielska, Zearalenone as an endocrine disruptor in humans. Environmental Toxicology and Pharmacology, 2016. **48**: p. 141-149.

8. Massart, F. and G. Saggese, Oestrogenic mycotoxin exposures and precocious pubertal development. International Journal of Andrology, 2010. **33**(2): p. 369-376.

9. Massart, F., et al., High growth rate of girls with precocious puberty exposed to estrogenic mycotoxins. Journal of Pediatrics, 2008. **152**(5): p. 690-695.

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

11. Rivera-Nunez, Z., et al., Urinary mycoestrogens and age and height at menarche in New Jersey girls. Environmental Health, 2019. **18**.

12. Claeys, L., et al., Mycotoxin exposure and human cancer risk: A systematic review of epidemiological studies. Comprehensive Reviews in Food Science and Food Safety, 2020. **19**(4): p. 1449-1464.

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

14. Lo, E.K.K., et al., Mechanistic insights into zearalenone-accelerated colorectal cancer in mice using integrative multi-omics approaches. Computational and Structural Biotechnology Journal, 2023. **21**: p. 1785-1796.

15. Rong, X., et al., Close association between the synergistic toxicity of zearalenone-deoxynivalenol combination and microRNA221-mediated PTEN/PI3K/AKT signaling in HepG2 cells. Toxicology, 2022. **468**.

16. IARC, IARC monographs on the evaluation of carcinogenic risks to humans. IARC monographs, 1993. **56**: p. 599.

17. Wang, Y.F., et al., Assessing the effect of food mycotoxins on aromatase by using a cell-based system. Toxicology in Vitro, 2014. **28**(4): p. 640-646.

18. Rizner, T.L. and A. Romano, Targeting the formation of estrogens for treatment of hormone dependent diseases-current status. Frontiers in Pharmacology, 2023. **14**.

19. De Boevre, M., et al., Human exposure to mycotoxins and their masked forms through cereal-based foods in Belgium. Toxicology Letters, 2013. **218**(3): p. 281-292.

20. Gupta, R.C., et al., Chapter 49 - Trichothecenes and zearalenone, in Reproductive and Developmental Toxicology (Third Edition), R.C. Gupta, Editor. 2022, Academic Press. p. 1003-1016.

21. Knutsen, H.K., et al., Risks for animal health related to the presence of zearalenone and its modified forms in feed. Efsa Journal, 2017. **15**(7).

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

23. Katzenellenbogen, J.A., et al., Structural underpinnings of oestrogen receptor mutations in endocrine therapy resistance. Nature Reviews Cancer, 2018. **18**(6): p. 377-388.

24. Kao, Y.C., et al., Molecular basis of the inhibition of human aromatase (estrogen synthetase) by flavone and isoflavone phytoestrogens: A site-directed mutagenesis study. Environmental Health Perspectives, 1998. **106**(2): p. 85-92.

25. Russell, J.F., J.C. Wong, and M.M. Grumbach, Chapter 3J - Aromatase Deficiency and Aromatase Excess, in Genetic Steroid Disorders, M.I. New, et al., Editors. 2014, Academic Press: San Diego. p. 165-190.

26. Bateman, A., et al., UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Research, 2023. **51**(D1): p. D523-D531.

27. Wang, Y., et al., A review on application of molecular simulation technology in food molecules interaction. Current Research in Food Science, 2022. **5**: p. 1873-1881.

28. Zhao, L., et al., 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, 2021. **4**: p. 603-611.

29. Ji, D.W., et al., Physicochemical characterisation, molecular docking, and drug-likeness evaluation of hypotensive peptides encrypted in flaxseed proteome. Current Research in Food Science, 2020. **3**: p. 41-50.

30. Dorne, J., et al., A Computational Understanding of Inter-Individual Variability in CYP2D6 Activity to Investigate the Impact of Missense Mutations on Ochratoxin A Metabolism. Toxins, 2022. **14**(3).

31. Louisse, J., J. Dorne, and L. Dellafiona, 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, 2022. **355**: p. 19-30.

32. Kim, S., et al., PubChem 2023 update. Nucleic Acids Research, 2023. **51**(D1): p. D1373-D1380.

33. Berman, H.M., et al., The Protein Data Bank. Nucleic Acids Research, 2000. **28**(1): p. 235-242.

34. Ghosh, D., et al., Novel Aromatase Inhibitors by Structure-Guided Design. Journal of Medicinal Chemistry, 2012. **55**(19): p. 8464-8476.

35. Pettersen, E.F., et al., UCSF chimera - A visualization system for exploratory research and analysis. Journal of Computational Chemistry, 2004. **25**(13): p. 1605-1612.

36. Pedroni, L., et al., A computational study on the biotransformation of alkenylbenzenes by a selection of CYPs: Reflections on their possible bioactivation. Toxicology, 2023. **488**.

37. Pedroni, L., et al., A Computational Inter-Species Study on Safrole Phase I Metabolism-Dependent Bioactivation: A Mechanistic Insight into the Study of Possible Differences among Species. Toxins, 2023. **15**(2).

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

39. Abraham, M.J., et al., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX, 2015. **1-2**: p. 19-25.

40. Zoete, V., et al., SwissParam: A Fast Force Field Generation Tool for Small Organic Molecules. Journal of Computational Chemistry, 2011. **32**(11): p. 2359-2368.

41. Brooks, B.R., et al., CHARMM: The Biomolecular Simulation Program. Journal of Computational Chemistry, 2009. **30**(10): p. 1545-1614.

42. Panneerselvam, S., et al., A Combined Molecular Docking/Dynamics Approach to Probe the Binding Mode of Cancer Drugs with Cytochrome P450 3A4. Molecules, 2015. **20**(8): p. 14915-14935.

43. Zhang, L., et al., Force field development for cofactors in the photosystem II. Journal of Computational Chemistry, 2012. **33**(25): p. 1969-1980.

44. Dellafiora, L., et al., 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, 2022. **162**.

45. Ludwig, M., et al., Female Pseudohermaphroditism Associated with a Novel Homozygous G-to-A (V370-to-M) Substitution in the P-450 Aromatase Gene. 1998. **11**(5): p. 657-664.

46. Lo, J., et al., Structural Basis for the Functional Roles of Critical Residues in Human Cytochrome P450 Aromatase. Biochemistry, 2013. **52**(34): p. 5821-5829.

47. Di Nardo, G., et al., Evidence for an Elevated Aspartate pK(a) in the Active Site of Human Aromatase. Journal of Biological Chemistry, 2015. **290**(2): p. 1186-1196.

48. Murthy, J.N., et al., Active site acidic residues and structural analysis of modelled human aromatase: A potential drug target for breast cancer. Journal of Computer-Aided Molecular Design, 2005. **19**(12): p. 857-870.

49. Martin, G.D.A., et al., Biotransformation and molecular docking studies of aromatase inhibitors. Steroids, 2016. **113**: p. 95-102.

50. Lephart, E.D., Modulation of Aromatase by Phytoestrogens. Enzyme Research, 2015. **2015**: p. 594656.

51. Chain, E.P.o.C.i.t.F., 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, 2014. **12**(12): p. 3916.

52. Alexander, J., et al., Appropriateness to set a group health-based guidance value for zearalenone and its modified forms. Efsa Journal, 2016. **14**(4).

Chapter V

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

This chapter has been published as: Perugino Florinda, **Pedroni Lorenzo**, Galaverna Gianni, Dall'Asta Chiara, Dellafiora Luca. Virtual display of targets: A new level to rise the current understanding of ochratoxin A toxicity from a molecular standpoint. Toxicology. 2024 Mar;503:153765. doi: 10.1016/j.tox.2024.153765.

Abstract

Ochratoxin A (OTA) is a worldwide spread mycotoxin contaminating several food and feed commodities, 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 stereoisomer. 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 [1]. Over the last 60 years, OTA toxicity has been widely investigated and thoroughly described. Nowadays, it 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 [2]. Also, the International Agency for Research on Cancer (IARC) classified it as possibly carcinogenic to humans (group 2B) [3]. Along with OTA, certain food commodities can be contaminated also by 2'R-OTA, an OTA diastereoisomer known to be produced upon thermal treatments like roasting and baking of OTA-containing food [4, 5]. Of note, 2'R-OTA seems to be less toxic compared to its diasteroisomer [6] although certain population groups may be highly exposed depending on the diet habits (e.g. coffee drinkers) [4].

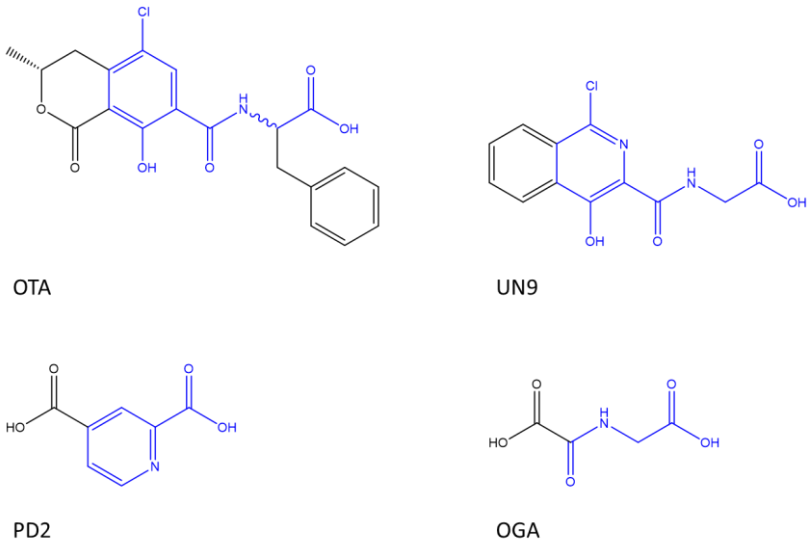


Figure 1. *Chemical structure of molecules under analysis. From the top left, N-Oxalyglycine (OGA), Pyridine-2,4-Dicarboxylic Acid (PD2) and N-[1-Chloro-4-Hydroxysoquinolin-3-yl)carbonyl]glycine (UN9). From the bottom left, Ochratoxin A (OTA) and its stereoisomer (2'R-OTA). The OGA, PD2 and UN9 common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

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 [7]. Specifically, these effects might occur via several pathways such as nucleotide metabolism, siege of RNA synthesis and many more [8]. 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 [9] that identified the strong chemical similarity between OTA and the Protein Data Bank (PDB; https://www.rcsb.org) [10] ligand UN9 (Figure 1), which is an inhibitor of the human OGFOD1 protein [11]. 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 [12]. 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 [13]. 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 [14-16] 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 [17, 18]. 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 [19] (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 [11] 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 as well [11]. The protonation state of OTA when arranged close to the iron cation was calculated using CHARMM36 via the SwissParam webserver (http://www.swissparam.ch) [20] 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.

2.2. Modelling studies

For the sake of a fit-for-purpose validation aimed a 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 [9, 14]. 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 [11]. 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) [21] 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 both 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 [9, 14], 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) [11], were compared to their respective crystallographic binding architectures.

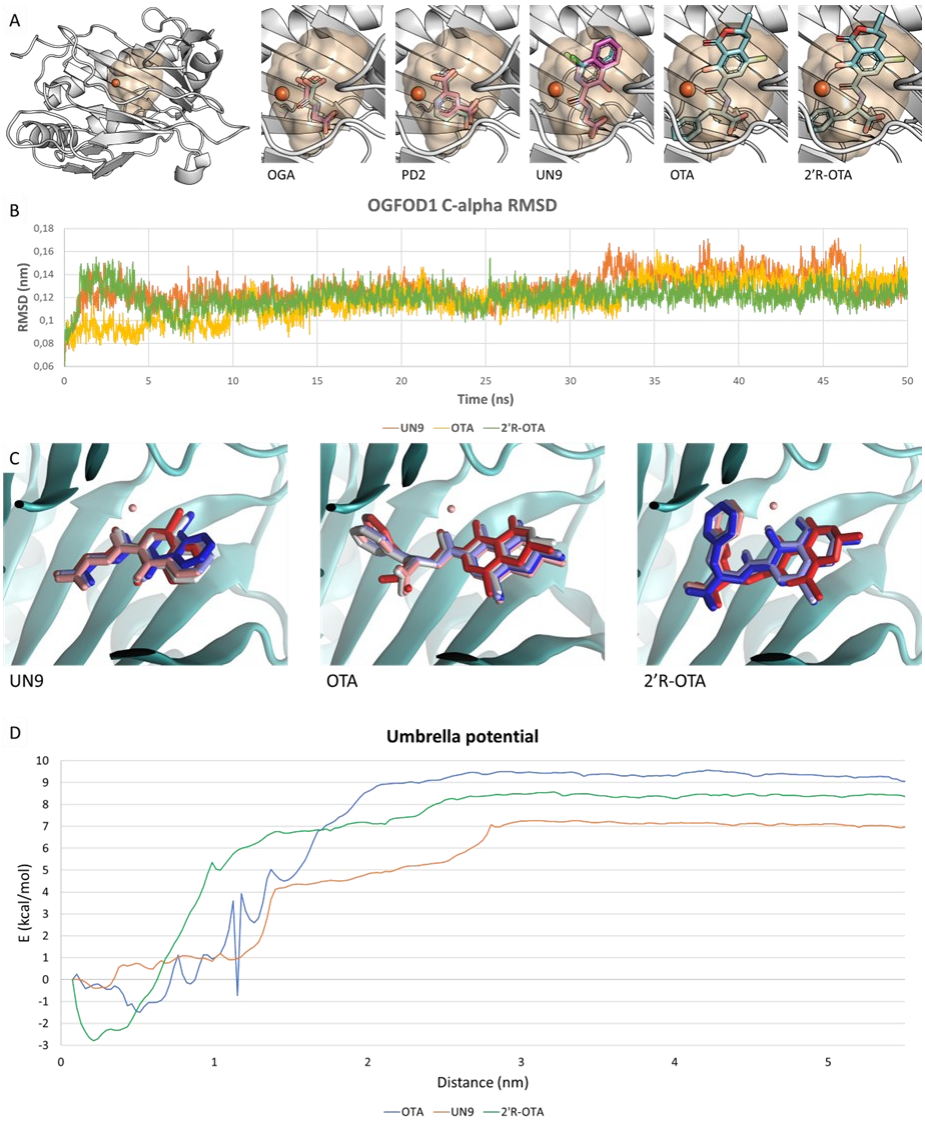


***Figure 2**. Molecular modelling validation and results of the considered ligands. **A**. On the left the 3D OGFOD1 model represented in white cartoon, with Fe represented as an orange sphere and the binding pocket as wheat semi-transparent surface. On the right, the docked poses are represented as 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.*

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 [9] with UN9 obtaining the best score (172 units), followed by PD2 (152 units) and OGA (150 units) (Table 1).

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. [11].

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 comparable inhibitory activity [11] (Table 1). With respect to the 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[14-16].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 [22]. 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 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 pKa switch the 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 the 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 [23]. 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 [23]. 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 on what expected from a salt bridge (3.04 ± 0.031 Å regarding OTA, 2.75 ± 0.016 regarding 2'R-OTA) [24].

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 [25]. 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 [26, 27]. 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 [28], a key element for translation initiation [29]. Notably, the phosphorylation of eIF2α, which depends also on OGFOD1, turns the protein in its inactive form hindering the translation process [30]. This apparently contradicts the hypothesised OTA-dependent inhibition of OGFOD1 as OGFOD1 activity may lead to an increase of eIF2α phosphorylation [25]. 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 [27]. 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 [31] and this may have come with the here hypothesized inhibition of OGFOD1. Following this line of interpretation, it has been also described that OGFOD1 may increase LC3B lipidation [27], which is also observed upon OTA treatment of kidney cells [32]. 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) [26]. Interestingly, CDK2 was proposed as one of the key regulators of the G1 cell-cycle arrest induced by low nanomolar concentrations of OTA [33]. 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 [32]. 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 diastereoisomer 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 [4]. 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.

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. Indeed, the OTA-dependent inhibition of OGFOD1 is soundly consistent with the current molecular understanding of OTA toxicity and may fill some of its mechanistic gaps. 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 actual capability of OTA and its 2'R stereoisomer to inhibit OGFOD1. On this basis, OGFOD1 may be a mechanistic connector interlacing the yet scarcely connected molecular actors and events associated with the OTA-dependent 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 its adverse outcome pathway.

References

1. Vettorazzi, A., E. González-Peñas, and A.L. de Cerain, *Ochratoxin A kinetics: A review of analytical methods and studies in rat model.* Food and Chemical Toxicology, 2014. **72**: p. 273-288.

2. Tao, Y.F., et al., *Ochratoxin A: Toxicity, oxidative stress and metabolism.* Food and Chemical Toxicology, 2018. **112**: p. 320-331.

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

4. Cramer, B., et al., *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, 2015. **59**(9): p. 1837-1843.

5. Bryla, M., et al., *Transformation of ochratoxin A during bread-making processes.* Food Control, 2021. **125**.

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

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

8. Liang, R., et al., *Apoptosis Signal-regulating Kinase 1 promotes Ochratoxin A-induced renal cytotoxicity.* Scientific Reports, 2015. **5**.

9. Dellafiora, L., et al., *An In Silico Target Fishing Approach to Identify Novel Ochratoxin A Hydrolyzing Enzyme.* Toxins, 2020. **12**(4).

10. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Research, 2000. **28**(1): p. 235-242.

11. Horita, S., et al., *Structure of the Ribosomal Oxygenase OGFOD1 Provides Insights into the Regio-and Stereoselectivity of Prolyl Hydroxylases.* Structure, 2015. **23**(4): p. 639-652.

12. Thinnes, C.C., et al., *Selective Inhibitors of a Human Prolyl Hydroxylase (OGFOD1) Involved in Ribosomal Decoding.* Chemistry-a European Journal, 2019. **25**(8): p. 2019-2024.

13. Park, M., et al., *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, 2023. **24**(6).

14. Pedroni, L., et al., *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**(21).

15. Perugino, F., et al., *A mechanistic toxicology study to grasp the mechanics of zearalenone estrogenicity: Spotlighting aromatase and the effects of its genetic variability.* Toxicology, 2024. **501**: p. 153686.

16. Del Favero, G., et al., *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, 2022. **159**.

17. Cournia, Z., B. Allen, and W. Sherman, *Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations.* Journal of Chemical Information and Modeling, 2017. **57**(12): p. 2911-2937.

18. Kuppusamy, A., M. Arumugam, and S. George, *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, 2017. **95**: p. 199-203.

19. Kim, S., et al., *PubChem 2023 update.* Nucleic Acids Research, 2023. **51**(D1): p. D1373-D1380.

20. Bugnon, M., et al., *SwissParam 2023: A Modern Web-Based Tool for Efficient Small Molecule Parametrization.* Journal of Chemical Information and Modeling, 2023. **63**(21): p. 6469-6475.

21. Abraham, M.J., et al., *GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers.* SoftwareX, 2015. **1-2**: p. 19-25.

22. Perry, J.L., et al., *Binding of ochratoxin A derivatives to human serum albumin.* Journal of Physical Chemistry B, 2003. **107**(27): p. 6644-6647.

23. Harding, M.M., *Small revisions to predicted distances around metal sites in proteins.* Acta Crystallographica Section D-Biological Crystallography, 2006. **62**: p. 678-682.

24. Ban, X.F., et al., *Evolutionary Stability of Salt Bridges Hints Its Contribution to Stability of Proteins.* Computational and Structural Biotechnology Journal, 2019. **17**: p. 895-903.

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

26. Fujisaki, T., et al., *The prolyl hydroxylase OGFOD1 promotes cancer cell proliferation by regulating the expression of cell cycle regulators.* Febs Letters, 2023. **597**(8): p. 1073-1085.

27. Singleton, R.S., et al., *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, 2014. **111**(11): p. 4031-4036.

28. Wang, Y., et al., *Ochratoxin A induces cytotoxicity through ROS-mediated endoplasmic reticulum stress pathway in human gastric epithelium cells.* Toxicology, 2022. **479**.

29. Zoll, W.L., et al., *Characterization of mammalian eIF2A and identification of the yeast homolog.* Journal of Biological Chemistry, 2002. **277**(40): p. 37079-37087.

30. Jiang, H.Y. and R.C. Wek, *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, 2005. **280**(14): p. 14189-14202.

31. Yang, S.A., et al., *Ochratoxin A induces endoplasmic reticulum stress and fibrosis in the kidney via the HIF-1α/miR-155-5p link.* Toxicol Rep, 2023. **10**: p. 133-145.

32. Akpinar, H.A., H. Kahraman, and I. Yaman, *Ochratoxin A Sequentially Activates Autophagy and the Ubiquitin-Proteasome System.* Toxins, 2019. **11**(11).

33. Dubourg, V., et al., *Weighted Correlation Network Analysis Reveals CDK2 as a Regulator of a Ubiquitous Environmental Toxin-Induced Cell-Cycle Arrest.* Cells, 2020. **9**(1).

Section V

Conclusions and Outlook

In the current scenario, where the food system is changing and more globalized than ever before, the concern for foodborne pathogens, contaminants and toxicants rose from both industries and consumers. The increased prevalence of such threats highlighted the need for the identification of novel strategies to tackle them and for clarifying their mechanisms of action to develop informed food safety measures. To deal with these challenges, computational approaches might benefit the research in the food sector offering promising alternatives, or better integrations, with traditional *in vitro* – *in vivo* studies. Moreover, in this sense, identifying novel measures addressing these kinds of threats perfectly fits the One Health framework, acknowledging the tight interconnection among humans, animals and the environment.

Regarding this PhD thesis, computational approaches were applied, either alone or combined with *in vitro* methodologies, reaching promising results and solutions aligning with the One Health paradigm. Among the several case studies reported, three macro areas were investigated:

1. Tackling foodborne pathogens possibly with compounds of natural origin.
2. Obtaining more mechanistic insights on known toxicants.
3. Finding promising bioremediation strategies for naturally occurring toxins.

Computational modelling approaches successfully succeeded in better characterizing essential viral and bacterial proteins to target with compounds of natural origin. The latter could hopefully serve as food/feed additives, chemicals functionalizing food contact materials or as sanitizers for food processing surfaces. Regarding their identification, computational approaches enable the screening of huge chemical libraries otherwise unfeasible with traditional methodologies, both timewise and moneywise.

Moving to the characterization of the mechanistic insights of naturally occurring toxicants, *in silico* methodologies allowed to gain information on either their metabolic fate at an inter-individual – and

inter-species – level, or on their interaction with possible biological targets involved in their toxicokinetics and toxicodynamics never considered before. Regarding the former, findings about the differential zearalenone xenoestrogenicity based on the aromatase variants, as well as results on the interindividual, but also species-specific, alkenylbenzenes bioactivation, might help to tailor more precise food safety measures. Moreover, the identification of possible novel targets for already known toxicants might enhance our understanding on their fate in the human body thus broadening their understanding to find more targeted counter measures. Both outcomes may eventually improve the risk assessment process for contaminants of utmost relevance to food safety.

Regarding the possibility of finding novel bioremediation strategies, computational approaches succeeded and could be promising methodologies to identify enzymes able to biotransform toxins lowering their toxicity. In this sense, such approaches could help in the zone interconnecting food safety and environmental health to develop sustainable approaches.

Concluding, this thesis corroborated the efficacy of computational approaches, also when integrated to *in vitro* experiments, to address biological and chemical foodborne threats, aligning with the One Health framework. Indeed, it offered promising tools to support objectives such as prevention and mitigation, starting from their biological and chemical understanding. This PhD project, by moving steps toward a holistic approach to food safety, recognizing the crucial interplay among humans, animals and the environment, highlights the key role of computational sciences in the nowadays food safety research. The integration of computational methods into traditional *in vitro* – *in vivo* studies helps create a comprehensive approach where *in silico* predictions are combined with empirical robust testing. In this sense, the integration of these methodologies would fasten the decision-making process regarding the identification and selection of the most useful compounds to cope with specific foodborne pathogens or, about other threats, it can improve the precision of food safety interventions. Last, based on the inherently rapid evolution of foodborne threats and on the

case studies reported here, this thesis aims for a more extended use of computational science, either per-se or better along with traditional methods, toward the development of a sustainable, precise and more moneywise and timewise efficient, framework within food safety research.

Acknowledgements

My three-year PhD journey is coming to an end, and I would like to express my acknowledgements to those who supported me along the way.

First, I would like to acknowledge Professor Luca Dellafiora, whose guidance allowed me to fully live every moment of this journey. It all started with him being *just* my tutor and it ended up with him being primarily a friend.

I would like to acknowledge Professor Javier Luque, along with all the wonderful people within his research group, since it allowed me to conduct an amazing research work during my abroad period at the University of Barcelona and it has always been available and kind with this (me) Italian intruder.

I would also like to acknowledge Regione Emilia Romagna, which funded the current PhD project and gave me an amazing opportunity to work on such an interesting topic.

Last, but not for importance, I would like to express my gratitude to all the special people in my life whose support have been a source of strength throughout the whole journey, especially during the hardest moments of it. Your presence helped me in overcoming all the encountered challenges.

Now, as some beloved toon's characters once famously said, "*That's all folks!*".

About the author

Lorenzo Pedroni was born on October 20th, 1997, in Fidenza, a city next to Parma (Italy). He obtained a bachelor's degree in Biotechnology at the university of Parma in 2019. Then, he decided to move to a Bioinformatics international master's degree at the University of Bologna where he graduated in 2021. In November 2021 he started his PhD project in Food Science and Technology at the Food and Drug Department at the University of Parma under the supervision of Professor Luca Dellafiora. From March to June 2024, he worked as a visiting PhD candidate at the University of Barcelona under the supervision of Professor Javier Luque. During the PhD he authored several scientific papers, published in peer-reviewed journals, and participated to important scientific congresses as listed below.

List of publications

Dorne JLCM, Cirlini M, Louisse J, Pedroni L, Galaverna G, Dellafiora L. A Computational Understanding of Inter-Individual Variability in CYP2D6 Activity to Investigate the Impact of Missense Mutations on Ochratoxin A Metabolism. Toxins. 2022; 14(3):207. https:// doi.org/10.3390/toxins14030207

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(21):4680. https:// doi.org/10.3390/nu14214680

Pedroni, L., Dellafiora, L., Varrà, M.O. et al. In silico study on the Hepatitis E virus RNA Helicase and its inhibition by silvestrol, rocaglamide and other flavagline compounds. Sci Rep 12, 15512 (2022). https://doi.org/10.1038/s41598-022-19818-w

Pedroni L, Louisse J, Punt A, Dorne JLCM, Dall'Asta C, Dellafiora L. A Computational Inter- Species Study on Safrole Phase I Metabolism-Dependent Bioactivation: A Mechanistic Insight into the Study of Possible Differences among Species. Toxins. 2023;15(2):94. doi: 10.3390/toxins15020094

Pedroni L, Louisse J, Dorne JLCM, Dall'Asta C, Dellafiora L. A computational study on the biotransformation of alkenylbenzenes by a selection of CYPs: Reflections on their possible bioactivation. Toxicology. 2023;488:153471. doi:10.1016/j.tox.2023.153471

Pedroni L, Dorne JLCM, Dall'Asta C, Dellafiora L. An in silico insight on the mechanistic aspects of gelsenicine toxicity: A reverse screening study pointing to the possible involvement of acetylcholine binding receptor. Toxicology Letters. 2023;386:1-8. doi: 10.1016/j.toxlet.2023.09.003

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;173:113284. doi: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. 2024;501:153686. doi:10.1016/j.tox.2023.153686

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 ofGPR120. Current Research in Food Science. 2024;8:100710. doi:10.1016/j.crfs. 2024.100710

Poór M, Dombi Á, Fliszár-Nyúl E, Pedroni L, Dellafiora L. Effects of chrysin and Chrysin-7- sulfate on ochratoxin A-Albumin interactions and on the plasma and kidney levels of the mycotoxin in rats. ACS Omega. April 2024. doi:10.1021/acsomega.4c01738

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. 2024;503:153765. doi:10.1016/j.tox.2024.153765

Pedroni L, Doherty DZ, Dall'Asta C, Galaverna G, Bell SG, Dellafiora L. Computational methods meet in vitro techniques: A case study on fusaric acid and its possible detoxification through cytochrome P450 enzymes. Ecotoxicology and Environmental Safety. 2024;273:116167. doi:10.1016/j.ecoenv.2024.116167

Delfino D, Prandi B, Ridolo E, et al. Allergenicity of tropomyosin variants identified in the edible insect Hermetia illucens (black soldier fly). Food Chemistry. 2024;437:137849. doi: 10.1016/j.foodchem.2023.137849

Delfino D, Prandi B, Calcinai L, et al. Molecular Characterization of the Allergenic Arginine Kinase from the Edible Insect Hermetia illucens (Black Soldier Fly). Molecular Nutrition & Food Research. April 2024. doi:10.1002/mnfr.202300911

Abdallah, M. F., Recote, J. M., Van Camp, C., Van Hassel, W. H. R., Pedroni, L., Dellafiora, L., … Rajkovic, A. (2024). Potential (co-)contamination of dairy milk with AFM1 and MC-LR and their synergistic interaction in inducing mitochondrial dysfunction in HepG2 cells. *Food and Chemical Toxicology: An International Journal Published for the British Industrial Biological Research Association*, *192*(114907), 114907. doi:10.1016/j.fct.2024.114907

Louisse, J., Pedroni, L., van den Heuvel, J. J. M. W., Rijkers, D., Leenders, L., Noorlander, A., … Dellafiora, L. (2024). In vitro and in silico characterization of the transport of selected perfluoroalkyl carboxylic acids and perfluoroalkyl sulfonic acids by human organic anion transporter 1 (OAT1), OAT2 and OAT3. *Toxicology*, *509*(153961), 153961. doi:10.1016/j.tox.2024.153961

List of poster and oral communications

Oral at the 75° Convegno Sisvet, 15th-18th June 2022, Lodi (Italy). Lorenzo Pedroni, Luca Dellafiora and Sergio Ghidini. *In silico* evaluation of silvestrol, rocaglamide and other flavagline compounds as possible hepatitis E virus inhibitors.

Poster at the 26th Workshop on the Developments in the Italian PhD Research on Food Science Technology and Biotechnology, 19th- 21st September 2022, Asti (Italy). Lorenzo Pedroni. Computational approaches integrated to biomolecular technologies to study foodborne viral infections: virulence factors and identification of food-grade components with potential anti-viral activity.

Poster at the 44th Mycotoxin-Workshop, 5th – 7th June 2023, Celle (Germany). Lorenzo Pedroni, Gianni Galaverna, Chiara Dall'Asta and Luca Dellafiora. An *in silico* method to mine biological targets and biotransforming enzymes for emerging mycotoxins – A case study on fusaric acid.

Poster at the EUROTOX - 57th congress of the European societies of toxicology 10th – 13th September 2023, Ljubljana (Slovenia). Lorenzo Pedroni, Florinda Perugino, Gianni Galaverna, Chiara Dall'Asta, Luca Dellafiora. In *Silico* Site-directed Mutagenesis of TAS2R Enzymes Integrated with Molecular Modelling Techniques to Predict Chemical Susceptibility to Toxic Compounds.

Oral at the 45th Mycotoxin-Workshop, 2nd – 5th June 2024, Vienna (Austria). Pedroni, Lorenzo; Perugino, Florinda; Galaverna, Gianni; Dall'Asta, Chiara; Dellafiora, Luca. A new level to rise the current understanding of ochratoxin A toxicity from a molecular standpoint: an *in silico* journey.

Poster at the EUROTOX – 58th congress of the European societies of toxicology 8th – 11th September 2024, Copenhagen (Denmark). Lorenzo Pedroni, F. Perugino, G. Galaverna, C. Dall'Asta, F. M. Buratti, E. Testai, L. Dellafiora. Molecular modelling study on CYP2C19 mediated biotransformation of organophosphorothioate pesticides: insights on a possible mutational landscape.