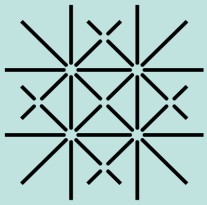
In Silico Prediction of Tartrazine Toxicity for Safer Foods

COMPUTER MODELING OF ADVERSE EFFECTS

28935-01

Author:

Anastasiia Popova



Universität Basel

Contents

Introduction

This report[1] examines the toxicity of one of the most commonly used synthetic food colorants in the industry – yellow dye Tartrazine (also known as E102, Yellow 5 Lake, Acid Yellow 23, Food Yellow 4, and trisodium) [1].

Tartrazine is most frequently found in so-called ultra-processed food, which is food that significantly altered from its original state and usually contains a lot of added salt, sugar, fat, as well as other additives such as preservatives, emulsifiers, flavorings, and colorings [2]. These foods have a long shelf life and are often ready to eat or heat, but they are generally low in nutrients. Consumption of these foods is linked to higher risks of mortality [3], obesity [4], excessive sugar intake [5], cardiovascular diseases [6], and certain types of cancer [7].

Motivation

Computer modeling of interactions of Tartrazine with off-target proteins can help identify potential adverse effects for humans on a molecular level, provide insights into whether such effects require further experimental studies, and guide regulatory decisions of safer food additives.

Goal

This study aims to explore the main concepts of computer modeling in small molecule toxicity research and apply this knowledge to estimate the toxic potential of Tartrazine. In this work, only free online software was used.

1 Chemical Properties of Tartrazine

Tartrazine is an organic sodium salt (Na^+) – a synthetic lemon yellow azo dye (contains –N=N- bounds) - soluble in water, hygroscopic; when heated to decomposition it emits toxic fumes, but it is stable in normal conditions. Its molecular formula is $C_{16}H_9N_4Na_3O_9S_2$, 2D and 3D structures are shown in Figure 1, the summary of main chemical properties is in Table 1.

1.1 Health Hazards

Tartrazine itself is generally considered non-toxic; it is legally permitted as an additive in food and cosmetics in many countries [8], and only a small percentage of people are allergic to it [9]. Despite its widespread use, Tartrazine has also been associated with

[1]The following AI tools were used to assist in report writing: Perplexity AI and ChatGPT. Perplexity AI was used for more effective information searches, while ChatGPT – to improve draft texts. The generated output underwent critical evaluation and fact-checking. The author is solely responsible for the content and academic integrity of this document.

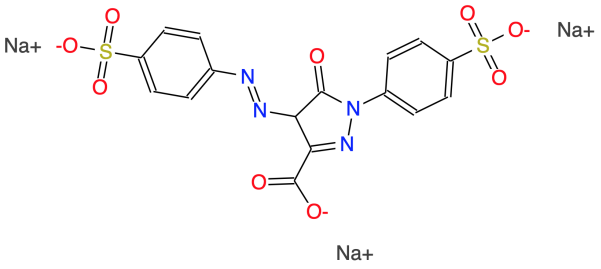
more serious adverse effects, such as neurotoxicity (animal study [10]), genotoxicity and carcinogenic potential (human lymphocytes [11] and eukaryotic cells [12] in vitro).

Thus, in addition to the risks of allergies and other individual reactions, there is also the concern that this compound might have chronic toxicity for humans with long-term exposure to Tartrazine even at low doses.

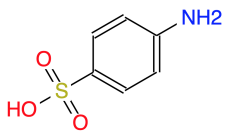
1.2 Metabolism

In order to assess toxicity of a substance more accurately, it is common practice to consider metabolites (chemically active substances that can be produced during the evolution inside of an organism). In this work we analyze not only Tartrazine itself, but also its metabolite. [2]

After being taken orally, Tartrazine undergoes significant metabolism by the gut bacteria, resulting in the formation of sulfanilic acid [1]. It is a white crystalline solid organic compound, soluble in water; its molecular formula is $C_6H_7NO_3S$ (see also Figure 1, Table 1), and it is not considered highly toxic [15]. Studies or data related to the chronic toxicity of sulphanilic acid are limited.



(a) 2D chemical structure of Tartrazine



(b) 2D chemical structure of Sulfanilic acid (Tartrazine's metabolite)

Figure 1: Chemical structures of studied compounds [16].

2 Methods

2.1 Metrics

Let us introduce the terms and metrics used in the following sections.

In general, **toxicity** is the degree to which a chemical substance can damage an organism [17]. To assess the toxic effect on the body, factors such as dose, time of

[2] To search for possible metabolites, the NERDD software can be used. We don't use the NERDD for Tartrazine because GLORYx (FAME 3 model) cannot make predictions for molecules containing Na atoms [13], and GLORY (FAME 2) works if all molecules are neutral [14].

Property Name	Tartrazine	Sulfanilic acid
Molecular Weight	534.4 g/mol	173.19 g/mol
Hydrogen Bond Donor Count	0	2
Hydrogen Bond Acceptor Count	12	4
Rotatable Bond Count	3	1
Topological Polar Surface Area	229 Å^2	88.8 Å^2
Heavy Atom Count	34	11
Formal Charge	0	0
Complexity	949	211
Isotope Atom Count	0	0
Defined Atom Stereocenter Count	0	0
Undefined Atom Stereocenter Count	1	0
Defined Bond Stereocenter Count	0	0
Undefined Bond Stereocenter Count	0	0
Covalently-Bonded Unit Count	4	1

Table 1: Chemical properties of the Tartrazine [1] and Sulfanilic acid [15].

exposure, and route of entry are usually considered. For dose-response relationships, acute toxicity (high concentrations over a short period of time) and chronic toxicity (low concentrations over a long period of time) are also distinguished. Toxicity threshold is a level of concentration that does not have adverse effects (or these effects cannot be measured).

To assess the toxicity of a substance in silico, the molecular interaction strength between the substance (referred to as a ligand or small molecule) and potential targets (proteins that play an important role in the organism) is usually predicted. Additionally, an empirical analysis is conducted to determine how similar the substance under investigation is to other known toxic compounds [18],[19]. The compound is considered not toxic if the molecules do not interact or the interaction is weak for all targets. It should be noted that the final assessment of toxicity needs to be supported by its **ADME** (adsorption, distribution, metabolism, and excretion) modeling [20].

The **binding affinity** is a commonly used measure of the interaction strength between a ligand and the target molecule. It is often expressed in terms of the dissociation constant K_d, which is the concentration of ligand at which half of the binding sites are occupied at equilibrium [21]; a lower K_d value indicates a higher binding affinity. In terms of the Gibbs free energy of dissociation ΔG_d (the energy needed to break a chemical bond) it is

$$\Delta G_d = \Delta H_d - T\Delta S_d = -RT \ln \frac{K_d}{c_0} \tag{1}$$

where c_0 is the concentration the standard state, R is the gas constant (8.31 J K^{-1} mol^{-1}), T is the absolute temperature (K), ΔH_d and ΔS_d denote the changes in enthalpy and entropy respectively [21].

The **toxic potential** can be estimated from the normalized binding affinities towards all targets. Low toxicity is associated with toxic potential less than 0.3 [20]. While there is

a strong correlation between potency (strong biological effect at low concentration), off-target affinity, and toxicity, it is not absolute: for example, drugs with high tissue selectivity may achieve therapeutic effects with minimal off-target toxicity [22].

In practice, to categorize substances based on their toxicity, the **LD50 value** (the dose required to kill 50% of a test population in animal/cells studies) is used. Affinity is also often expressed as pKi, which is the negative logarithm of the inhibition constant (K_i, used when the kinetic mechanism can be identified and the ligand lowers the activity of proteins involved in catalyzing reactions). Another possible measure of potency is IC50 (concentration required to inhibit a specific biochemical function by 50%), it is related to receptor affinity through the Cheng-Prusoff equation; to see the detailed description of the dependencies between different measures used in biochemistry, the reader can visit [23].

2.2 Modeling

The Structure-Activity Relationship (SAR) shows how different chemical modifications to a molecule can affect its interaction with proteins. The idea of using them in toxicology studies can be expressed as follows: If we systematically modify the chemical structure, the resulting changes in the molecule's activity can be correlated with these modifications. QSAR (Quantitative Structure-Activity Relationship) models leverage statistical methods to predict the toxic effects of new compounds based on their chemical structure [24]. There are several types of algorithms used for QSAR modeling, which can be classified as linear (linear regression-based methods), non-linear (neural networks, support vector machines), and data-based (clustering methods and decision trees) [19].

The prediction of QSAR models can be enhanced by the application of structure-based methods, which give a more detailed mechanistic interpretation of molecular interactions. These methods include:

- **Molecular docking** is a computational technique that predicts the preferred orientation of a ligand when it binds to a target protein and its binding affinity. The docking process involves *sampling* and *scoring*: sampling explores the space of possible states to generate binding poses while scoring evaluates these poses to predict the most likely binding mode and affinity [25].

- **Molecular dynamics simulations** are computational methods that model the evolution of the physical movements of atoms and molecules over time. Knowledge about the stability of ligand-protein complexes can be valuable to assess long-term toxic effects [26].

- **The 3D binding mode modeling** refers to the spatial arrangement of a ligand when it is bound to a protein. A mechanistic interpretation of how a substance interacts with its target can be linked to adverse outcome pathways (AOPs) [27].

To examine the flowchart of in silico prediction models in detail, refer to Figure 9 in [19].

2.3 Software

In this work, we applied several free online tools to study Tartrazine and Sulfanilic acid toxicity: VirtualToxLab, PanScreen, Molinspiration, ProTox-3.0, and SwissADME. Below, we provide a brief description of each tool and highlight its most relevant features for this report.

- **VirtualToxLab** simulates and quantifies the binding of small molecules to 16 protein targets known or suspected to trigger adverse effects. This platform combines an automated flexible docking model with a multi-dimensional QSAR (mQSAR) model to provide predictions of toxic potential and activity [28]. VirtualToxLab takes 3D structure in SDF format of a compound from a user.

- **PanScreen** provides both qualitative (binding poses) and quantitative (binding affinities) predictions of toxicity. It uses a consensus model, based on an ensemble of molecular docking models and po-sco interaction analysis (uses algorithms to identify various types of interactions, such as hydrogen bonds, pi-pi stacking, and hydrophobic contacts, with high precision), unsaturated functional groups analysis (based on qualitative tests of bounds) and attention-based models to generate interaction fingerprints which gives a high predictive performance [29].

  PanScreen also accepts 3D structure in SDF format of a compound. The output of Panscreen is able to generate flags to warn the user about the reliability of the modeling. Examples of flags obtained for the studied compounds include: “Compound not binding”, “Compound could not be docked by all programs”, “Compound has very bad docking scores” and “Calculated applicability score is very low”. It is also able to render 3D poses in the web application, which significantly increases the speed of assessing results.

- **Molinspiration** predicts the bioactivity scores for various targets, including GPCR ligands, ion channel modulators, kinase inhibitors, nuclear receptor ligands, protease inhibitors, and other enzyme targets [30]. As input, it uses SMILES, by drawing the structure, or SDfile of the molecule of interest and usesa Bayesian statistical model trained on the fragment-based table of active and non-active molecules [31]. Bioactivity score is a universal drug-likeness measure to compare activity within a class of targets. A bioactivity score greater than 0.0 suggests that the compound is likely to be active against the target, values between 0.0 and -5.0 indicate moderate activity and compounds with scores less than -5.0 are considered inactive.

  ProTox-3.0 provides predictions of the median lethal dose (LD50), toxicity class, and accuracy. It also yields information on organ toxicity, toxicological endpoints, toxicological pathways, molecular initiating events, and metabolism. Users can input the chemical structure either by specifying the PubChem name, canonical SMILES string, or by drawing the structure. To predict toxicity, ProTox-3.0 uses machine learning models (random forest and deep neural networks), which are validated using 10-fold cross-validation and test datasets [32].

- **SwissADME** predicts properties of small molecules relevant to drug metabolism and pharmacokinetics and integrates multiple computational tools [33]. As input it takes molecule's SMILES and returns a detailed report about chemical structure, predicted parameters, and SwissADME Bioavailability Radar chart, which provides a rapid assessment of a drug-likeness. Each axis on the radar chart represent one of six physicochemical properties, and the optimal range (for effective drugs) is depicted as a pink area; description of the properties can be found in Additional data.

3 Results

The **VirtualToxLab** results for Tartrazine and its metabolite show the same activity levels for all targets ($>100\,\mu\mathrm{m}$), indicating that only very high concentrations of these substances can lead to the suppression of all modeled receptors. The toxic potentials are low: for Tartrazine, it is **0.053**, and for Sulfanilic acid, it is **0.073**. Thus, the VirtualToxLab results indicate that Tartrazine and its metabolite are non-toxic.

According to the results from **PanScreen** in Figure 2 Tartrazine probably binds to several targets. For the Androgen, Melatonin, and Estrogen beta receptors, three flags were raised, indicating that the prediction might be inaccurate; thus, we excluded these targets from the analysis. A high probability of interaction, probability of being active, and predicted affinity (-11 kcal/mol) with only a single warning (compound could not be docked by all programs) is demonstrated only for the Glucocorticoid receptor. This receptor regulates the expression of numerous genes involved in metabolism, immune response, development, and other physiological processes [34].

For Tartrazine's metabolite, Sulfanilic acid, we cannot confirm binding with the modeled targets. A flag indicates that the prediction is not reliable, as shown in Figure 3). The inspection of 3D poses for the high-affinity receptors supports this conclusion.

The **Molinspiration** results for Tartrazine show that it is bioactive against enzyme inhibitors, while Sulfanilic acid has moderate bioactivity against them (the absolute difference between their bioactivity scores is 0.01). Both Tartrazine and Sulfanilic acid have moderate bioactivity for GPCR ligands, ion channel modulators, kinase inhibitors, nuclear receptor ligands, and proteases, see Figure 5.

The **ProTox-3.0** results for both compounds are shown in Figure 6. The predicted toxicity class for Tartrazine is 3, with an LD50 of 81 mg/kg, while for Sulfanilic acid, the toxicity class is 5, with an LD50 of 3770 mg/kg. There is no binding for any targets, and Tartrazine is active against only a single toxicity endpoint - chemical-induced BBB-barrier toxicity [35] with a probability of **0.64**. Notably, that Sulfanilic acid has similar results: no binding and the same BBB-barrier toxicity endpoint with a higher probability of **0.86**.

In this report, we do not analyze in detail all the metrics provided by **SwissADME**, only the most relevant to the estimation of the toxic potential of Tartrazine. More data from SwissADME can be found in Additional data. [3]

[3] On the SwissADME website, it is specified that it is preferable to input the neutral form of the molecule: "submitting an ionized structure would lead to severe biases in the predictions unless it is a permanent ion

Target name	Probability of interaction	Probability active	Predicted affinity	Flags
Melatonin receptor 1B	High	0.90	-9.5	1; 2; 3
Glucocorticoid receptor	High	1.00	-11.0	2
Dopamine receptor D2	Uncertain	0.83	-8.2	
Estrogen receptor beta	Low	0.52	-8.1	1; 2; 3
Tyrosine-protein kinase JAK2	Uncertain	1.00	-9.0	
Estrogen receptor alpha	Uncertain	0.98	-9.0	2; 3
Beta-2 adrenergic receptor	Low	0.02	-7.3	
Androgen receptor	High	1.00	-10.0	1; 2; 3
Tyrosine-protein kinase JAK1	Uncertain	0.96	-8.5	
Substance-P receptor	Low	0.03	-6.1	
5HT receptor 1B	Low	0.08	-7.6	
PPAR-gamma	Low	0.32	-7.8	
Phosphodiesterase 4D	Low	0.00	-6.1	
Phosphodiesterase 10A	Uncertain	0.66	-8.2	

Figure 2: PanScreen results for Tartrazine.
"Predicted Affinity" is measured in kcal/mol. Values in the "Flags" column correspond to the following messages: "1": Compound not binding,
"2": Compound could not be docked by all programs,
"3": Compound has very poor docking scores.

Bioavailability scores from SwissADME: for Tartrazine, it is **0.17**; for Sulfanilic acid, it is **0.56**. Tartrazine's bioavailability score indicates that this compound is unlikely to be absorbed efficiently when administered orally, which is also supported by its violation of 2 of the 5 Lipinski's rules (more than 10 hydrogen bond acceptors; molecular weight is more than 500 g/mol) and its GI absorption, as shown in Table 2. Considering the fact that Sulfanilic acid is most likely formed in the gut (see Metabolism), this metric will not be considered for further analysis of this compound.

PAINS and Brenk alerts were raised for Tartrazine, shown in Figure 8, which can indicate that the compound contains fragments known to cause adverse effects or interfere with biological assays leading to false positives (for more information see Table 5). For Sulfanilic acid, Brenk alerts remain in effect. This might also indicate poor drug-likeness for both compounds.

At the same time, Tartrazine is a P-gp substrate (see Table 2), which means that it can be more rapidly eliminated from the body, reducing its systemic exposure and potentially reducing toxicity [36]. Also, Figure 7 shows high values of saturation degree (INSATU) for both compounds, which can be linked to a lower likelihood of being toxic [37]. And a high TPSA value suggests poor cell membrane permeability for Tartrazine, see POLAR in Figure 7 and Table 4.

or a zwitterion." We consider Tartrazine in the ionized form that exists in solution and, most probably, inside the body, and expect that some SwissADME predictions may not be accurate.

Target name	Probability of interaction	Probability active	Predicted affinity	Flags
Phosphodiesterase 10A	Low	0.03	-6.6	
Beta-2 adrenergic receptor	Low	0.00	-6.5	
Androgen receptor	Low	0.00	-5.7	
Dopamine receptor D2	Low	0.10	-7.6	
Tyrosine-protein kinase JAK2	Low	0.00	-6.6	
Estrogen receptor alpha	High	1.00	-11.0	4
Estrogen receptor beta	Low	0.01	-5.1	
Glucocorticoid receptor	Low	0.00	-7.0	
Tyrosine-protein kinase JAK1	Uncertain	0.95	-8.6	4
Substance-P receptor	High	1.00	-15.0	4
5HT receptor 1B	Low	0.06	-7.6	
PPAR-gamma	Low	0.00	-7.1	
Melatonin receptor 1B	Low	0.00	-6.9	
Phosphodiesterase 4D	Low	0.00	-5.5	

Figure 3: PanScreen results for Sulfanilic acid.
"Predicted Affinity" is measured in kcal/mol. Values in the "Flags" column correspond to the following messages: '4': Calculated applicability score is very low.

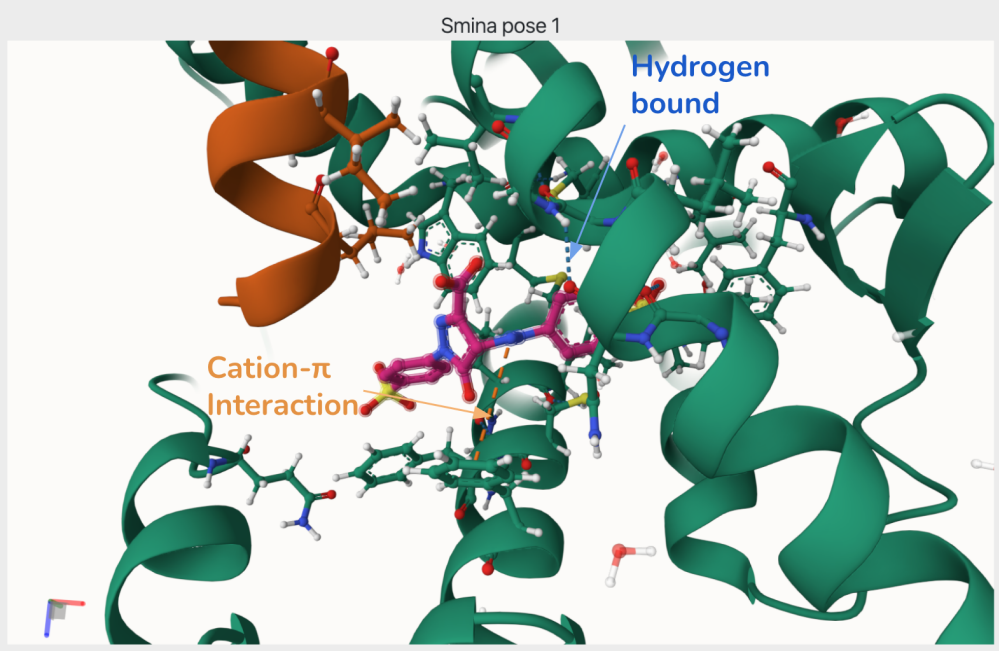
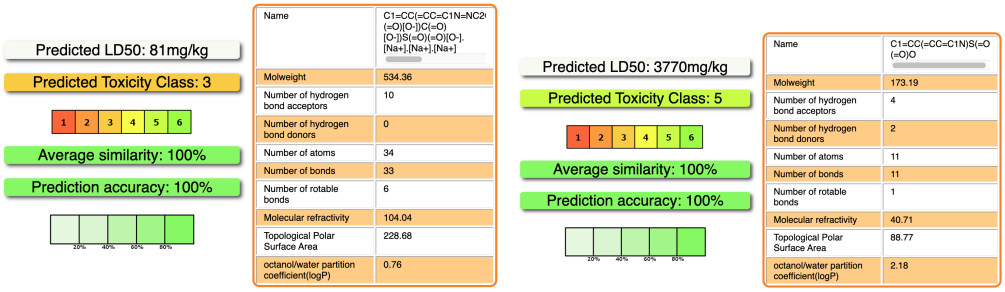


Figure 4: Ligand-protein configuration (Tartrazine-Glucocorticoid receptor) obtained from PanScreen.

	Tartrazine	Sulfanilic acid
GPCR Ligand	-0.13	-0.61
Ion Channel Modulator	-0.23	-0.04
Kinase Inhibitor	-0.49	-1.04
Nuclear Receptor Ligand	-0.72	-1.82
Protease Inhibitor	-0.04	-0.47
Enzyme Inhibitor	0.01	-0.00

Figure 5: Molinspiration results for bioactivity scores.
Red color indicates bioactivity (score is greater than 0),
and green indicates moderate bioactivity (score is between -5 and 0).

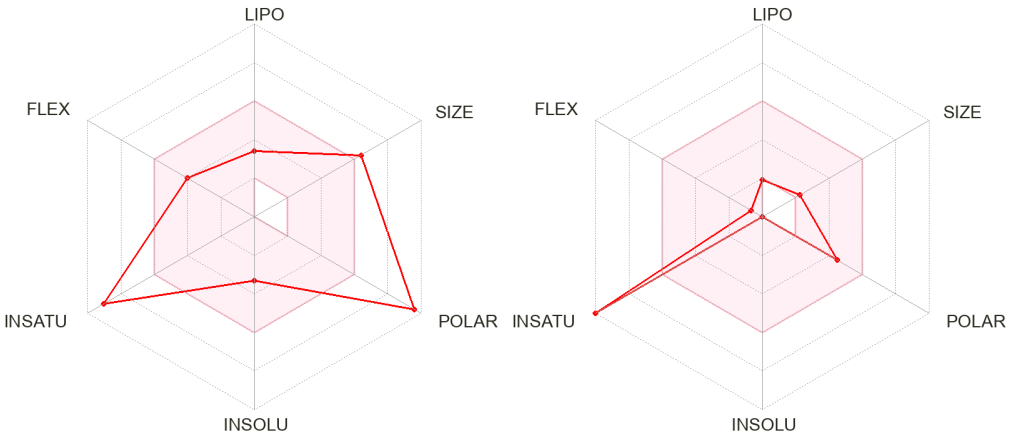


(a) Tartrazine **(b)** Sulfanilic acid

Figure 6: ProTox-3.0 oral toxicity prediction results.

Property	Tartrazine	Sulfanilic acid	Description
GI Absorption	Low	High	Absorption through the gastrointestinal (GI) tract.
BBB Permeant	No	No	The compound effectively crosses the blood-brain barrier (BBB).
Pgp Substrate	Yes	No	Indicates that the compound is a substrate for P-glycoprotein (Pgp), a membrane transporter protein that pumps foreign substances out of cells.
CYPs Inhibitor	No	No	This indicates that the compound does not inhibit the cytochrome the corresponding enzyme (from the list CYP1A2, CYP2C19, CYP2C9, CYP2D6, CYP3A4).

Table 2: Pharmacokinetic and pharmacodynamic properties from SwissADME



(a) Tartrazine **(b)** Sulfanilic acid

Figure 7: SwissADME Bioavailability Radar plots.

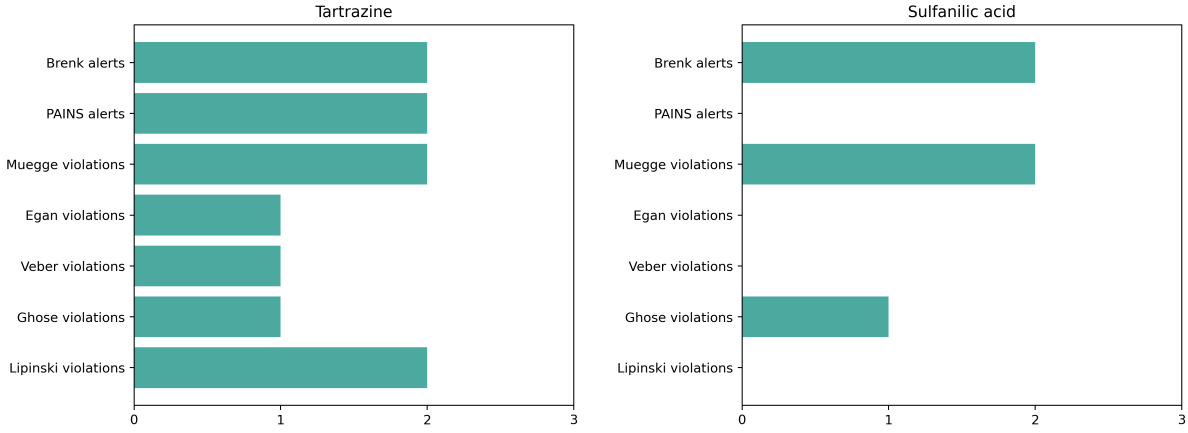


Figure 8: Violations and alerts from SwissADME.

4 Discussion

Most software tools concur that the studied compounds are not highly toxic. ProTox-3.0 classifies Tartrazine as toxicity class 3 (moderately hazardous) and Sulfanilic acid as non-toxic. VirtualToxLab also indicates a low toxic potential for both compounds.

The Molinspiration results for Tartrazine and its metabolite Sulfanilic acid show moderate bioactivity across multiple protein targets. Specifically, Molinspiration predicts that Tartrazine functions as an enzyme inhibitor. Similarly, Sulfanilic acid exhibits moderate enzyme inhibitory activity. This suggests that these compounds might interfere with and inhibit various enzymes in the body, potentially leading to adverse effects. However, it is important to note that bioactivity scores are probabilistic predictions based on structural similarities and require validation through additional methods.

ADME analysis suggests low bioavailability for both compounds. Tartrazine, in particular, is identified as a substrate for P-glycoprotein, which facilitates its rapid elimination from the body. Brenk and PAINS alerts also can indicate poor drug-likeness, which is not an issue for food coloring, as it and its metabolite should not be effectively absorbed by the body. However, there is disagreement between ProTox-3.0 and SwissADME regarding the compounds' ability to cross the blood-brain barrier (BBB). ProTox-3.0 warns of potential BBB toxicity for both substances, whereas SwissADME does not. Given that SwissADME recommends submitting the neutral form of the molecule (which is not the case for Tartrazine), BBB toxicity remains a significant concern. BBB disruption can contribute to neurotoxicity by allowing toxins to enter the brain, corroborating findings of Tartrazine-induced neurotoxicity in rats [10].

PanScreen results reveal that Tartrazine can bind to the Glucocorticoid receptor, which is important for regulating gene expression [34]. It is known that the binding of synthetic glucocorticoids to this receptor can lead to neuropsychiatric [38] and metabolic [39] side effects, particularly with long-term exposure. Considering that Tartrazine is widely used as a food coloring and likely consumed regularly in low doses, its potential for chronic toxicity cannot be excluded.

5 Conclusions

In summary, Tartrazine and its metabolite can be considered to have low toxicity and bioavailability. At the same time, their enzyme inhibitory activity, potential for blood-brain barrier disruption, and Tartrazine's binding to the Glucocorticoid receptor raise concerns. Although Tartrazine is a permitted food colorant widely used in many food products, evidence from this in silico study suggests that it has a potential for inducing chronic toxic effects and highlights the need for further investigation into endocrine disruption and neurotoxicity with its long-term exposure.

References

[1] National Center for Biotechnology Information. Pubchem compound summary for cid 164825. https://pubchem.ncbi.nlm.nih.gov/compound/164825, 2024. Accessed: 2024-10-05.

[2] Carlos A Monteiro, Geoffrey Cannon, Renata B Levy, Jean-Claude Moubarac, Maria LC Louzada, Fernanda Rauber, Neha Khandpur, Gustavo Cediel, Daniela Neri, Euridice Martinez-Steele, et al. Ultra-processed foods: what they are and how to identify them. *Public health nutrition*, 22(5):936–941, 2019.

[3] Laure Schnabel, Emmanuelle Kesse-Guyot, Benjamin Allès, Mathilde Touvier, Bernard Srour, Serge Hercberg, Camille Buscail, and Chantal Julia. Association between ultraprocessed food consumption and risk of mortality among middle-aged adults in france. *JAMA internal medicine*, 179(4):490–498, 2019.

[4] Kevin D Hall, Alexis Ayuketah, Robert Brychta, Hongyi Cai, Thomas Cassimatis, Kong Y Chen, Stephanie T Chung, Elise Costa, Amber Courville, Valerie Darcey, et al. Ultra-processed diets cause excess calorie intake and weight gain: an inpatient randomized controlled trial of ad libitum food intake. *Cell metabolism*, 30(1):67–77, 2019.

[5] Fernanda Rauber, Maria Laura da Costa Louzada, Euridice Martinez Steele, Leandro FM de Rezende, Christopher Millett, Carlos A Monteiro, and Renata B Levy. Ultra-processed foods and excessive free sugar intake in the uk: a nationally representative cross-sectional study. *BMJ open*, 9(10):e027546, 2019.

[6] Bernard Srour, Léopold K Fezeu, Emmanuelle Kesse-Guyot, Benjamin Allès, Caroline Méjean, Roland M Andrianasolo, Eloi Chazelas, Mélanie Deschasaux, Serge Hercberg, Pilar Galan, et al. Ultra-processed food intake and risk of cardiovascular disease: prospective cohort study (nutrinet-santé). *bmj*, 365, 2019.

[7] Thibault Fiolet, Bernard Srour, Laury Sellem, Emmanuelle Kesse-Guyot, Benjamin Allès, Caroline Méjean, Mélanie Deschasaux, Philippine Fassier, Paule Latino-Martel, Marie Beslay, et al. Consumption of ultra-processed foods and cancer risk: results from nutrinet-santé prospective cohort. *bmj*, 360, 2018.

[8] International Association of Color Manufacturers. Tartrazine - fd&c yellow no. 5. https://iacmcolor.org/color-profile/tartrazine-fdc-yellow-no-5/, 2024. Accessed: 2024-10-05.

[9] Verywell Health. The potential risks of tartrazine. https://www.verywellhealth.com/tartrazine-free-diet-83227, 2022. Accessed: 2024-10-05.

[10] Amany Abdel-Rahman Mohamed, Azza AA Galal, and Yaser HA Elewa. Comparative protective effects of royal jelly and cod liver oil against neurotoxic impact of tartrazine on male rat pups brain. *Acta Histochemica*, 117(7):649–658, 2015.

[11] Bruno Moreira Soares, Taíssa Maíra Thomaz Araújo, Jorge Amando Batista Ramos, Laine Celestino Pinto, Bruna Meireles Khayat, Marcelo De Oliveira Bahia, Raquel Carvalho Montenegro, Rommel Mario Rodríguez Burbano, and Andre Salim Khayat. Effects on dna repair in human lymphocytes exposed to the food dye tartrazine yellow. *Anticancer research*, 35(3):1465–1474, 2015.

[12] Jailson Rodrigues Dos Santos, Larissa de Sousa Soares, Bruno Moreira Soares, Marlene de Gomes Farias, Victor Alves de Oliveira, Natan Antônio Batista de Sousa, Helber Alves Negreiros, Felipe Cavalcanti Carneiro da Silva, Ana Paula Peron, Ana Carolina Landim Pacheco, et al. Cytotoxic and mutagenic effects of the food additive tartrazine on eukaryotic cells. *BMC Pharmacology and Toxicology*, 23(1):95, 2022.

[13] Gloryx predicting phase i and phase ii metabolites. https://nerdd.univie.ac.at/gloryx/about/. Accessed: 2024-06-03.

[14] Glory predicting cytochrome p450 metabolites. https://nerdd.univie.ac.at/glory/about/. Accessed: 2024-06-03.

[15] National Center for Biotechnology Information. Pubchem compound summary for sulfanilic acid. https://pubchem.ncbi.nlm.nih.gov/compound/8479, 2024. Accessed: 2024-10-05.

[16] Cheminfo. 2d stuctures from smiles. https://www.cheminfo.org/flavor/malaria/Utilities/SMILES_generator___checker/index.html, 2024. Accessed: 2024-10-05.

[17] Merriam-Webster. Toxicity, 2024. Accessed: 2024-05-24.

[18] Nigel Greene and William Pennie. Computational toxicology, friend or foe? *Toxicology Research*, 4(5):1159–1172, 2015.

[19] Arwa B Raies and Vladimir B Bajic. In silico toxicology: computational methods for the prediction of chemical toxicity. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 6(2):147–172, 2016.

[20] Angelo Vedani, Max Dobler, Zhenquan Hu, and Martin Smieško. Openvirtualtoxlab—a platform for generating and exchanging in silico toxicity data. *Toxicology letters*, 232(2):519–532, 2015.

[21] Panagiotis L Kastritis and Alexandre MJJ Bonvin. On the binding affinity of macromolecular interactions: daring to ask why proteins interact. *Journal of The Royal Society Interface*, 10(79):20120835, 2013.

[22] Duxin Sun, Wei Gao, Hongxiang Hu, and Simon Zhou. Why 90% of clinical drug development fails and how to improve it? *Acta Pharmaceutica Sinica B*, 12(7):3049–3062, 2022.

[23] Science Snail. The difference between ki, kd, ic50, and ec50 values. https://www.sciencesnail.com/science/the-difference-between-ki-kd-ic50-and-ec50-values, 2024. Accessed: 2024-05-24.

[24] Rajarshi Guha. On exploring structure–activity relationships. *In silico models for drug discovery*, pages 81–94, 2013.

[25] Xuan-Yu Meng, Hong-Xing Zhang, Mihaly Mezei, and Meng Cui. Molecular docking: a powerful approach for structure-based drug discovery. *Current computer-aided drug design*, 7(2):146–157, 2011.

[26] Scott A Hollingsworth and Ron O Dror. Molecular dynamics simulation for all. *Neuron*, 99(6):1129–1143, 2018.

[27] André Fischer, Martin Smiesko, Manuel Sellner, and Markus A Lill. Decision making in structure-based drug discovery: visual inspection of docking results. *Journal of Medicinal Chemistry*, 64(5):2489–2500, 2021.

[28] Angelo Vedani, Max Dobler, and Martin Smieško. Virtualtoxlab—a platform for estimating the toxic potential of drugs, chemicals and natural products. *Toxicology and applied pharmacology*, 261(2):142–153, 2012.

[29] Manuel S Sellner, Markus A Lill, and Martin Smiesko. Panscreen: A comprehensive approach to off-target liability assessment. *bioRxiv*, pages 2023–11, 2023.

[30] Tahmeena Khan, Shalini Dixit, Rumana Ahmad, Saman Raza, Iqbal Azad, Seema Joshi, and Abdul Rahman Khan. Molecular docking, pass analysis, bioactivity score prediction, synthesis, characterization and biological activity evaluation of a functionalized 2-butanone thiosemicarbazone ligand and its complexes. *Journal of chemical biology*, 10:91–104, 2017.

[31] Molinspiration. Virtual screening. https://www.molinspiration.com/docu/miscreen/virtualscreening.html, 2024. Accessed: 2024-06-03.

[32] Priyanka Banerjee, Emanuel Kemmler, Mathias Dunkel, and Robert Preissner. Protox 3.0: a webserver for the prediction of toxicity of chemicals. *Nucleic Acids Research*, page gkae303, 2024.

[33] Antoine Daina, Olivier Michielin, and Vincent Zoete. Swissadme: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific reports*, 7(1):42717, 2017.

[34] Robert H Oakley and John A Cidlowski. The biology of the glucocorticoid receptor: new signaling mechanisms in health and disease. *Journal of Allergy and Clinical Immunology*, 132(5):1033–1044, 2013.

[35] Toxicological endpoints: Bbb-barrier (protox-3.0). https://tox.charite.de/protox3/index.php?site=models#Toxicological_endpoints/model_bbb.

[36] Iman Imtiyaz Ahmed Juvale, Azzmer Azzar Abdul Hamid, Khairul Bariyyah Abd Halim, and Ahmad Tarmizi Che Has. P-glycoprotein: New insights into structure, physiological function, regulation and alterations in disease. *Heliyon*, 8(6), 2022.

[37] Wenxiu Wei, Srinivasulu Cherukupalli, Lanlan Jing, Xinyong Liu, and Peng Zhan. Fsp3: A new parameter for drug-likeness. *Drug Discovery Today*, 25(10):1839–1845, 2020.

[38] Anne-Sophie CAM Koning, Merel van der Meulen, Daphne Schaap, Djaina D Satoer, Christiaan H Vinkers, Elisabeth FC van Rossum, Wouter R van Furth, Alberto M Pereira, Onno C Meijer, and Olaf M Dekkers. Neuropsychiatric adverse effects of synthetic glucocorticoids: a systematic review and meta-analysis. *The Journal of Clinical Endocrinology & Metabolism*, 109(6):e1442–e1451, 2024.

[39] Michael J Garabedian, Charles A Harris, and Freddy Jeanneteau. Glucocorticoid receptor action in metabolic and neuronal function. *F1000Research*, 6, 2017.

A Additional data

Metric	Description
LIPO (Lipophilicity)	Measured by XLOGP3, a computational method for predicting lipophilicity. Acceptable range: -0.7 to +5.0. Affects ability to cross cell membranes.
SIZE	Refers to molecular weight (MW). Acceptable range: 150 - 500 g/mol. Influences absorption in the gastrointestinal tract.
POLAR (Polarity)	Measured by Topological Polar Surface Area (TPSA). Acceptable range: 20 - 130 Å^2. Affects solubility and permeability.
INSOLU (Insolubility)	Measured by Log S (ESOL), predicting solubility in water. Acceptable range: -6 to 0. Crucial for dissolution in gastrointestinal fluids.
INSATU (Insaturation)	Measured by Fraction Csp3, which is defined as the ratio of sp3 hybridized carbon atoms to the total number of carbon atoms in the molecule. Acceptable range: 0.25 - 1. Indicates degree of saturation, affecting 3D structure and flexibility.
FLEX (Flexibility)	Measured by number of rotatable bonds. Acceptable range: 0 - 9. Influences ability to fit into biological targets and overall stability.

Table 3: Description of SwissADME Radar metrics.

Property	Tartrazine	Sulfanilic acid	Description
MR	104.04	40.71	Molar refractivity (MR) is a measure of the total polarizability of a molecule and is related to its molecular volume. It is used to understand the dispersion forces in a molecule and can be calculated using atomic refractivity values.
TPSA	228.68	88.77	Topological Polar Surface Area (TPSA) is the surface sum over all polar atoms (usually oxygen and nitrogen) and their attached hydrogen atoms. It is a key descriptor for predicting a molecule's ability to permeate cell membranes. A high TPSA (>130) value suggests poor cell membrane permeability.
Consensus Log P	-7.72	-0.09	Consensus Log P is the arithmetic mean of multiple log P predictions (including iLOGP, XLOGP3, WLOGP, MLOGP, and Silicos-IT Log P). The values can span from highly negative (e.g. around -10), indicating very hydrophilic and water-soluble compounds, to highly positive values (e.g. above +5), indicating very lipophilic and lipid-soluble compounds
ESOL Log S	-3.3	0.11	ESOL Log S is a prediction of aqueous solubility based on molecular properties. The values can span from around -6 (indicating insoluble compounds) to 0 (indicating highly soluble compounds).

Table 4: Molecular Properties from SwissADME.

Metric	Description
Lipinski violations (Ro5)	Lipinski's rule of five provides a set of guidelines to evaluate drug-likeness based on molecular properties: 1) No more than 5 hydrogen bond donors; 2) No more than 10 hydrogen bond acceptors; 3) Molecular weight less than 500 g/mol. 4) LogP (octanol-water partition coefficient) is not greater than 5.
PAINS alerts	Pan-assay interference compounds are chemical compounds that often give false positive results in assays.
Abbott Bioavailability Score (ABS)	A semi-quantitative rule-based score used to predict the probability that a compound will have at least 10% oral bioavailability (in rats).
Brenk alerts	A number of structural alerts associated with toxicity and poor drug-likeness.
Ghose violations	Ghose's criteria for drug-likeness include: 1) Molecular weight between 160 and 480 g/mol; 2) Log P between -0.4 and 5.6; 3) Molar refractivity between 40 and 130; 4) Total number of atoms between 20 and 70.
Veber violations	Veber's criteria for oral bioavailability: 1) 10 or fewer rotatable bonds; 2) A topological polar surface area (TPSA) of 140 Å^2 or less.
Egan violations	Predicts the oral absorption of drug candidates, the Egan rules are: 1) Log P between -1 and 5; 2) TPSA of 132 Å^2 or less
Muegge violations	Wider set of drug-likeness rules: 1) Molecular weight between 200 and 600 g/mol; 2) Log P between -2 and 5; 3) TPSA of 150 Å^2 or less; 4) Number of rings between 3 and 7; 5) Number of carbon atoms between 4 and 35; 6) Number of heteroatoms between 1 and 15; 6) Number of rotatable bonds between 0 and 15 8) Number of hydrogen bond donors between 0 and 5; 9) Number of hydrogen bond acceptors between 2 and 20.

Table 5: Description of toxicity metrics in SwissADME.