Food Research International 236 (2026) 119262

Contents lists available at ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

An *in silico* study of binding of new lutein esters to nuclear endocrine receptors

Marjan Vračko [a], Alen Albreht [b,*]

[a] *Laboratory for Cheminformatics, Theory Department, National Institute of Chemistry, Hajdrihova 19, SI-1000 Ljubljana, Slovenia*
[b] *Laboratory for Food Chemistry, Department of Analytical Chemistry, National Institute of Chemistry, Hajdrihova 19, SI-1000 Ljubljana, Slovenia*

ARTICLE INFO

Keywords:
Lutein ester
Endocrine-related activity
Neural network
Prediction
Antioxidant

ABSTRACT

Here we present the results of an *in silico* bioactivity study of 13 structurally diverse lutein diesters that were previously synthesized in the laboratory to improve the inherently poor chemical stability of lutein, an important food nutrient. Using predictive models, we aimed to identify potential receptor-related activity patterns relevant to endocrine signalling. Because orally administered lutein diesters may undergo partial or complete gastrointestinal hydrolysis before absorption, the corresponding carboxylic acids were also evaluated as potentially relevant exposure forms. Most lutein esters showed predicted receptor-activity classifications similar to lutein, whereas esters containing long-chain fatty acid moieties showed more distinct prediction patterns. Lutein di(pentafluoropropanoate) is of particular concern because it readily hydrolyses, releasing pentafluoropropionic acid, a corrosive hydrolysis product. The remaining carboxylic acids were not classified as notably concerning within the scope of the applied prediction models. Overall, the results identify lutein esters that merit further investigation as candidate stabilized lutein derivatives, while highlighting specific structures that should be prioritized for targeted experimental follow-up.

1. Introduction

Lutein, a naturally abundant xanthophyll carotenoid, plays an important role in human health due to its many positive bioactive properties. Humans cannot produce lutein themselves, but must ingest this nutrient through food. Its chemical structure, characterized by a central conjugated double bond system, gives it potent antioxidant activity, making it an important component of the body's defence against oxidative stress. The multiple benefits of lutein include not only its antioxidant abilities, but also its anti-cancer, anti-inflammatory and, most importantly, vision-supporting effects (Xu et al., 2023). The antioxidant effect of lutein is particularly noteworthy as it plays a crucial role in neutralizing reactive oxygen species (ROS) and reactive nitrogen species (RNS), thereby protecting cellular components from oxidative damage (Rao & Rao, 2007). Lutein can also alleviate chronic diseases such as cardiovascular disease, age-related macular degeneration (AMD) and certain types of cancer (Seddon et al., 1994; Zhang, Liang, et al., 2024). The ability of lutein to modulate inflammatory processes further increases its therapeutic relevance (Pap et al., 2021).

Lutein is present in many green leafy plants such as vegetables, but its bioavailability from these common food sources is generally poor due to low water solubility of lutein and its interaction with dietary fiber (Metličar et al., 2019; Yong et al., 2023). This highlights the need to also consider alternative sources such as lutein supplements or fortified foods to ensure adequate intake (Ma & Lin, 2010). However, the commercial use of lutein is hampered by its inherent chemical instability, especially when exposed to different environmental factors and processing conditions. The susceptibility of lutein to degradation limits its uptake from dietary supplements and functional foods, thereby compromising its efficacy and bioavailability (Manupa et al., 2023). The compound is susceptible to light, air (oxygen), elevated temperatures, metals, and acids, which accelerate the degradation of lutein into less bioactive compounds, reducing its health benefits (Becerra et al., 2020).

In order to solve this issue, various stabilization strategies have been proposed in recent decades. In this context, it is reported that encapsulation of lutein in emulsion-based delivery systems using milk proteins (the caseins) and phospholipids (soy lecithin) as emulsifiers significantly enhanced the chemical stability of lutein during storage (Mora-Gutierrez et al., 2018). However, each stabilization strategy, including encapsulation technology, has its own advantages and limitations, thus further research is needed to fine-tune these approaches for specific applications (Zhang, Li, et al., 2024).

* Corresponding author.
E-mail address: alen.albreht@ki.si (A. Albreht).

https://doi.org/10.1016/j.foodres.2026.119262
Received 1 October 2025; Received in revised form 9 April 2026; Accepted 15 April 2026
Available online 16 April 2026

Another stabilization possibility, which mimics a process occurring naturally in plants, is esterification of lutein, providing lutein mono- or diesters (Metličar et al., 2021; Metličar & Albreht, 2022). Bio-derived lutein esters (LE), especially those present in marigold (*Tagetes erecta* L.) and other plant materials, have been intensively studied, but in most cases their stability was assessed within complex plant or food matrices. Under such conditions, the influence of the ester moiety itself is difficult to separate from matrix effects, which complicates direct comparison across studies (Olmedilla-Alonso et al., 2024; Song et al., 2024). By contrast, the use of synthetic lutein esters as intentionally designed and potentially more stable lutein derivatives has only recently begun to be explored in a systematic way, for example with respect to their bioaccessibility (Krstić et al., 2026). Because this research area is still at an early stage, and because such compounds are not yet established as commercial nutritional ingredients to our knowledge, their application potential, industrial feasibility, and cost-effectiveness remain difficult to predict.

To address some of the safety-related endpoints and obtain a more realistic view of their practical potential, we present here the results of an *in silico* bioactivity assessment of 13 lutein diesters that were previously synthesized. These compounds were selected as a structurally diverse test set spanning short-chain, branched, unsaturated, aromatic, long-chain, and fluorinated acyl substituents. They are not intended to represent all possible lutein derivatives, but rather a chemically diverse panel suitable for first-step comparative screening. The human endocrine system plays a central role in regulating physiological processes and behaviour of an organism by controlling hormone production and release. Hence, we assessed the predicted receptor-related activity of these new derivatives using selected endocrine-relevant targets. Since oral administration is the principal intended route for lutein-based food and feed ingredients, the studied lutein esters are expected to pass through the gastrointestinal tract, where partial or complete hydrolysis may occur. Thus, exposure after ingestion may comprise both intact ester molecules and their hydrolytic products. For this reason, we evaluated the parent lutein esters and the corresponding carboxylic acids as interrelated chemical forms that may jointly determine the biological profile of these compounds after oral intake. The *in silico* prediction modelling presented here was intentionally designed as a logical, rapid, informative, non-invasive, and comparatively inexpensive first-step screening approach for identifying possible receptor-level bioactivity signals and prioritizing novel lutein compounds for further experimental evaluation. An additional advantage of this approach is that it enables parallel assessment of both intact lutein esters and their likely hydrolysis products within the same study design before progressing to more resource-intensive experimental studies. This is particularly relevant because experimental discrimination between parent esters and hydrolysis-derived species may require multiple complementary digestion and bioactivity assays, as hydrolysis can be partial and matrix-dependent.

2. Data and models

The following models have been applied to endocrine nuclear receptors to predict activity: androgen (AR)-agonist, AR-antagonist, oestrogen-α (ERα)-agonist, ERα-antagonist, oestrogen-β (ERβ)-agonist, ERβ-antagonist, glucocorticoid (GR)-agonist, and GR-antagonist. The models for androgen and oestrogen receptors are described in detail in references (Stanojević, Sollner Dolenc, & Vračko, 2023; Stanojević, Vračko, & Sollner Dolenc, 2023). The data were taken from the CompTox Chemicals Dashboard (US Environmental Protection Agency, 2016), a publicly available web-based application developed by the US Environmental Protection Agency to provide access to systematically compiled and consolidated chemical data (molecular structures and some basic chemical data), toxicity and exposure data for approximately 1 million chemicals. The number of all chemicals in the database and the number of chemicals used in the modelling is shown in Supplementary Table S1. The compounds are categorised as active or inactive considering the thresholds for activity from *in vitro* measurements according to the Tox21 protocols (US Environmental Protection Agency, Center for Computational Toxicology and Exposure, 2022). To include lutein and the thirteen LEs in the modelling, their structures were generated in the PubChem platform and represented with SMILES codes. The DRAGON package was used to compute over 3000 structural descriptors that quantitatively described molecular properties without needing experimental data (Talete Srl, 2014).

In the first step, the descriptors were analysed using principal component analysis (PCA) to reduce the number of variables while preserving the information content of the data (the variance). The new variables were ordered according to the variance they carried. As a rule, a small number of new variables carry a substantial part of the total variance (Johnson et al., 1988). After the transformation, the objects were displayed graphically in score plots for the eight studied nuclear receptor models. To predict the activities, we used counter-propagation artificial neural network (CPANN) models as described in references (Stanojević, Sollner Dolenc, & Vračko, 2023; Stanojević, Vračko, & Sollner Dolenc, 2023). CPANN models are often used to analyse the structure of multidimensional datasets. Since information on their architecture and learning strategy is readily available (Novič, 2023; Vracko, 2005), we give here only a brief description. A CPANN was composed of two layers of neurons arranged in two-dimensional rectangular matrices. The input or Kohonen layer received the input variables, *i.e.* the first twenty principal components representing the compounds. During the learning process, the target values (the membership of an active/inactive class) were passed to the output layer, which had the same topological arrangement of neurons as the Kohonen layer. Learning in a Kohonen layer of the CPANN took place in the same way as in Kohonen networks. This means that all neurons were presented with a vector of input variables. The algorithm selected the neuron (*i.e.* winning neuron) whose weights were closest to the input values. After correcting the weights in the Kohonen layer, the position of the winning neuron was transferred from the Kohonen layer to the output layer and the weights in the output layer were corrected according to the specified target value. Iteratively, the weights in the Kohonen and output layers were corrected to become more and more similar to the input variables. Once the weights were stabilized, the CPANN was considered trained. The compounds located on the same neuron are considered as a cluster. With regard to the class of compounds in the cluster, two basic situations can occur. First, all compounds that form the cluster belong to the same class. In such a case, one can define a group of compounds that are unanimously classified. Secondly, the compounds that form the cluster are classified inconsistently. In such a case, it is known that the model cannot classify the particular group of compounds. In the prediction phase, when a new structure with unknown effect is presented to the system, it is first placed in the Kohonen layer, the position of the winning neuron found is projected into the output layer, and the activity class is drawn from this position in the output layer. The CPANN models were optimized using the leave-one-out (LOO) test strategy. In fact, we followed the previously described models (Stanojević, Sollner Dolenc, & Vračko, 2023), which exclusively applied the LOO method to validate the models. The models described in the reference (Stanojević, Vračko, & Sollner Dolenc, 2023) were also tested using the leave-20%-out method where 20% of the objects are used as an external set. The procedure is repeated so that each object appears once in the external set.

The models were evaluated considering sensitivity (Se), specificity (Sp), accuracy (Acc) and Matthew's correlation coefficient (Mcc). The expressions for these parameters are given in eqs. 1–4. The calculations are based on predictions in the LOO test procedure, which can be true positive (TP), false positive (FP), true negative (TN) or false negative (FN).

$$Sp = \frac{TN}{(TN + FP)} \quad (1)$$

$$Se = \frac{TP}{(TP + FN)} \quad (2)$$

$$Acc = \frac{TP + TN}{TP + FN + TN + FP} \quad (3)$$

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \quad (4)$$

The models applied for the clustering of LEs are described in detail in references (Stanojević, Sollner Dolenc, & Vračko, 2023; Stanojević, Vračko, & Sollner Dolenc, 2023), which also contain supplementary information about the compound training sets. It should be emphasized that the models were constructed with different training sets and therefore the predictions for each model must be evaluated individually. The results of the LOO test are summarized in Table S2.

In addition, we evaluated the binding affinity of 13 carboxylic acids, which make the 13 studied LEs, to nuclear receptors using six models trained in the VEGA-HUB platform (Benfenati et al., 2013). Briefly, VEGA-HUB was developed to support the European chemicals legislation REACH. It is a battery of *in-silico* models for predicting the various physicochemical, biological, toxicological and ecotoxicological properties of molecules based on their molecular structures. The models were developed in different laboratories as part of various European research projects. The models were built using different methods, from traditional multiple linear regression to advanced machine learning methods, including artificial neural networks, decision trees, random forest, support vector machine, *etc.* Various databases were used to create the models in order to cover as much of the chemical space as possible. In addition to the predicted value, all models provide comprehensive information that supports the prediction, *e.g.* similar compounds from the training set, structural fragments and information about the application area of the model.

The models assessing activity at androgen, oestrogen, glucocorticoid, RBA and thyroid receptors were: Androgen receptor-mediated action (IRFMN/COMPARA)1.0.0 (ANDROGEN_COMPARA) (Mansouri et al., 2020), Oestrogen receptor-mediated action (IRFMN/CERAPP) 1.0.0 (OESTROGEN_CERAPP) (Mansouri et al., 2016), P-glycoprotein activity model (NIC) 1.0.0 (PGP_NIC) (Mora Lagares et al., 2019), and oestrogen receptor relative binding affinity model (IRFMN) 1.0.1 (RBA_IRFMN), thyroid receptor alpha effect (NRMEA) 1.0.0 (TRALPHA_NRMEA), and thyroid receptor beta effect (NRMEA) 1.0.0 (TRBETA_NRMEA) (Benfenati et al., 2013).

The in-house CPANN models were applied to the intact lutein esters, whereas VEGA-HUB was applied to the corresponding hydrolysis products (carboxylic acids). Because these approaches are based on different modelling frameworks, training sets, and applicability domains, their outputs (activity of compounds) were interpreted as complementary rather than directly comparable. In this context, the terms *active* and *inactive* refer only to model-predicted receptor-related activity classifications within the scope of the applied models and should not be interpreted as direct evidence of toxicity or endocrine disruption *in vivo*.

3. Results and discussion

As there are no toxicity data available for most of the LE derivatives studied here, we used computational models which classify compounds as active or inactive with respect to eight endocrine-receptor-related endpoints in order to compare the predicted receptor-activity profiles of these compounds with that of free lutein. In this context, agonists are compounds predicted to activate a receptor, whereas antagonists are compounds predicted to inhibit receptor activation. In biological systems, such activities may have different physiological implications depending on exposure, potency, metabolism, and context; however, the present results should be interpreted only as model-predicted receptor-related activity classifications. The models used in this study are based on selected training sets whose active compounds have a documented affinity for a given endocrine receptor.

The PCA score plots (Fig. 1) for eight nuclear receptor models (AR-agonist, AR-antagonist, ERα-agonist, ERα-antagonist, ERβ-agonist, ERβ-antagonist, GR-agonist, and GR-antagonist) illustrate the embedding of lutein and LEs into the domains of the models. The original number of 3716 variables (molecular descriptors) was reduced to twenty PC variables, which were used for further modelling. In general, PCA showed some similarities for all models where the first PC axis carried 43% to 56% of the total variance, while the second and third PC axes carried about 8% and 6% of the total variance, respectively. The first twenty PC variables used in further modelling carried over 80% of the total variance. The score plots reveal similar patterns with the models where LEs occupy the regions that do not strongly overlap with regions of compounds from the models' training sets. It should be emphasized that although the models were created on different training sets, which consisted of chemicals of different molecular size, polarity, structure, and functionalities, LEs clearly represent a unique group of highly hydrophobic compounds. They have specific chemical structures that possess the same core chemical moiety, *i.e.* the conjugated double-bond chain of lutein, but they differ in the side-chains that form esters with the lutein's hydroxyl functional groups. Nonetheless, in all score plots shown in Fig. 1 LEs are located in regions, which are populated with predominantly active compounds. This indicates that these compounds are predicted to show receptor-related activity, either as agonists or antagonists, similarly to the free lutein.

In the trained CPANN models, the LEs were mostly grouped and they occupied three to six neurons, which is not particularly surprising considering their similar core structures (details in Table S3). In some cases, LEs were located on neurons that were occupied by connections from the training set of the model, resulting in direct determination of the prediction. In other cases, LEs were located on empty neurons and the prediction of their activity was influenced by the connections from the neighbouring neurons in the neural network. When a certain share of these connections was active and the other inactive, no decision could be reached for the LE, as in the case of the predicted agonist activity of lutein dipalmitate and lutein dioleate toward ERβ, which were located on the same neuron. Some additional specific situations are discussed below.

In all models, a large cluster with seven to eleven LEs was found, which reasonably indicates that the studied compounds could be considered quite similar structurally. Although lutein was localised on a separate neuron in AR-agonist, AR-antagonist, and ERβ-agonist models, it otherwise fell within this larger cluster of LEs in other models. In the ERα-antagonist model, the majority of LEs were clustered and predicted to be inactive, with the exception of lutein dipalmitate, lutein dioleate, and lutein di(pentafluoropropanoate) which were located on two neurons. Their activity was determined by five compounds, which are listed in Table S3. The GR-agonist model predicts all LEs as active. The largest cluster comprises seven compounds whose activity was determined based on retinal. The other seven compounds were distributed across four different neurons. In the GR-antagonist model, eleven compounds formed a single cluster that was predicted to be inactive. The predictions were influenced by two compounds: vitamin K and tocopherol. Exceptions were lutein dioleate and lutein di(pentafluoropropanoate) which were active in both GR models.

The predicted activity classifications for LEs are summarized in Table 1. For six of the studied models, the majority of LEs are predicted to be active. The exceptions are the models for ERα-antagonist and GR-antagonist, where most of the LEs are predicted to be inactive. Lutein diacetate, lutein dipropanoate, lutein di(2,2-dimethylpropanoate), lutein di(2-methylpropanoate), lutein di(3-methylbutanoate), lutein divalerate, lutein di(pent-4-enoate), lutein dibenzoate, and lutein

diphthalate showed predicted receptor-activity classifications similar to lutein across the evaluated models. Lutein itself has a well-established oral safety profile in nutritional use; however, the present results should not be interpreted as direct evidence of toxicological safety for the corresponding esters. On the other hand, the predicted activities of lutein didecanoate, lutein dipalmitate, lutein dioleate, and lutein di (pentafluoropropanoate) were found to be different from that of lutein. The former two were predicted to be inactive as AR-antagonists (lutein was active), while the latter two showed activity as GR-antagonists (lutein was inactive). Furthermore, lutein dipalmitate, lutein dioleate, and lutein di(pentafluoropropanoate) were found to be active ERα-antagonists, while all other LEs and lutein were found to be inactive. These four LEs therefore displayed a receptor-activity pattern different from lutein in one or more of the evaluated models.

The interpretation of lutein dipalmitate and lutein dioleate requires particular caution, as these long-chain lutein esters occur naturally in

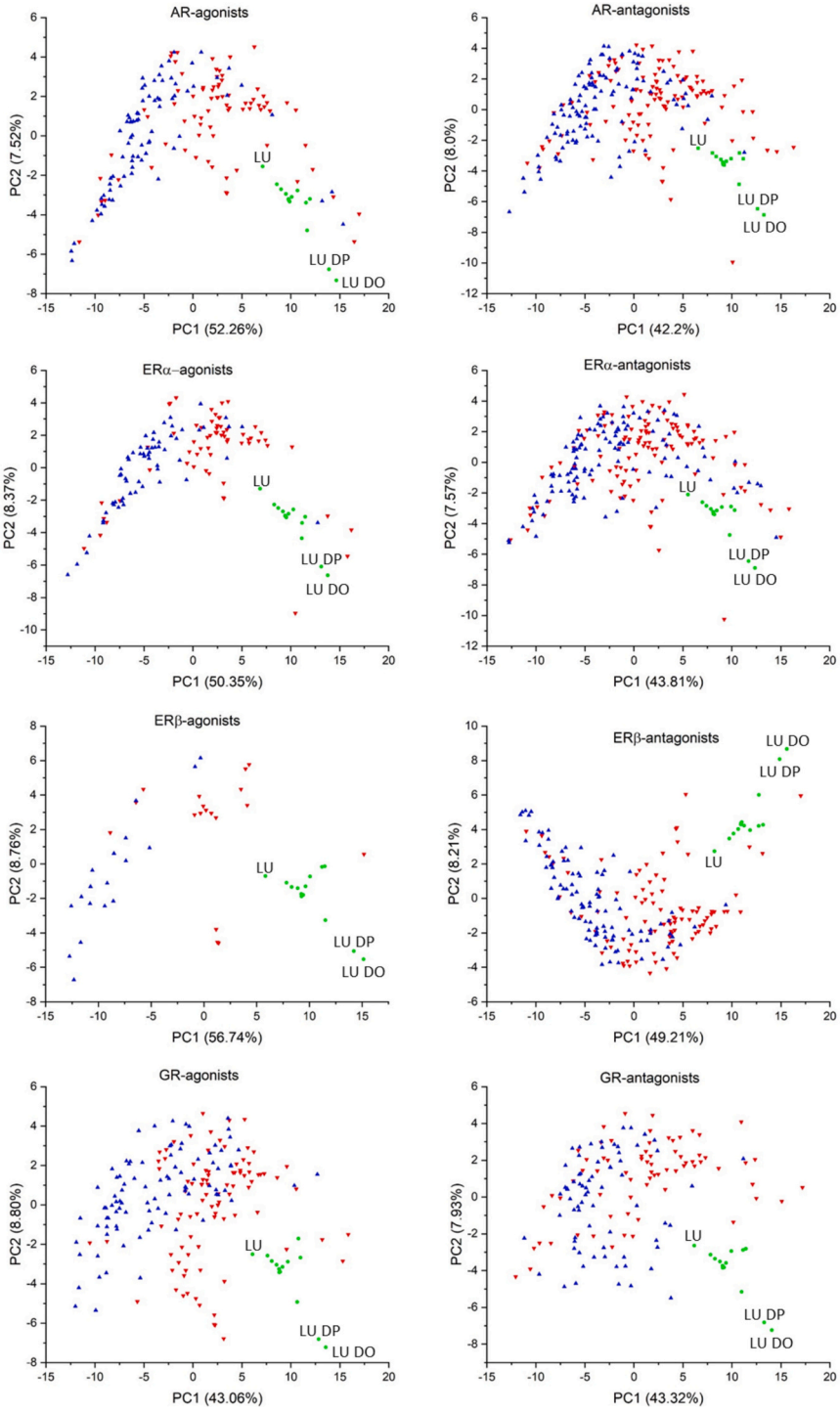

Fig. 1. Score plots for eight nuclear receptor models showing the first and the second principal components (PC1, PC2) that carry the largest portion of the dataset's overall variability: inactive compounds (blue), active compounds (red), and lutein and LEs (green). The bordering cases of the xanthophyll region, represented by lutein dipalmitate (LU DP), lutein dioleate (LU DO), and lutein (LU), are highlighted. The percentages of the total variance are given in the axes brackets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Predictions for 8 activities for 13 LEs using the CPANN models. The differences relative to lutein are marked by an asterisk (*).

Compound	AR-ag	AR-an	ERα-ag	ERα-an	ERβ-ag	ERβ-an	GR-ag	GR-an
Lutein	+	+	+	−	+	+	+	−
Lutein diacetate	+	+	+	−	+	+	+	−
Lutein dipropanoate	+	+	+	−	+	+	+	−
Lutein di(2,2-dimethylpropanoate)	+	+	+	−	+	+	+	−
Lutein di(2-methylpropanoate)	+	+	+	−	+	+	+	−
Lutein di(3-methylbutanoate)	+	+	+	−	+	+	+	−
Lutein divalerate	+	+	+	−	+	+	+	−
Lutein di(pent-4-enoate)	+	+	+	−	+	+	+	−
Lutein dibenzoate	+	+	+	−	+	+	+	−
Lutein didecanoate	+	−*	+	−	+	+	+	−
Lutein dipalmitate	+	−*	+	+*	ND	+	+	−
Lutein dioleate	+	+	+	+*	ND	+	+	+*
Lutein di(pentafluoropropanoate)	+	+	+	+*	+	+	+	+*
Lutein diphthalate	+	+	+	−	+	+	+	−

Legend: + - active; − - inactive; ND - no decision.

plants (Metličar et al., 2019), especially in marigold, which is a major industrial source of lutein for food and feed applications. In addition to the experimental toxicology study of Harikumar et al. (2008), more recent regulatory evaluations provide broader safety context for marigold-derived lutein materials. EFSA concluded that lutein esters were not of concern with respect to genotoxicity in the assessed preparations (EFSA Panel on Food Additives and Nutrient Sources added to Food (ANS) et al., 2011; EFSA Panel on Additives and Products or Substances used in Animal Feed (FEEDAP) et al., 2019; EFSA Panel on Additives and Products or Substances used in Animal Feed (FEEDAP) et al., 2024), and a recent FDA animal GRAS notice response stated that the agency had no questions regarding the notifier's conclusion that marigold extract is generally recognized as safe under the intended conditions of use (US Food and Drug Administration, Center for Veterinary Medicine, 2025). These assessments support the general safety profile of marigold-derived lutein preparations under their evaluated conditions of use, but they do not directly address the receptor-level prediction endpoints considered in the present work. Therefore, the differing model outputs observed for lutein dipalmitate and lutein dioleate identify these ester forms as priorities for targeted experimental investigation. Their distinct prediction patterns may plausibly reflect the presence of long-chain fatty acid moieties, which substantially increase steric bulk and lipophilicity relative to free lutein and the shorter-chain esters. This interpretation is indirectly supported by the VEGA-HUB results (*vide infra*), in which palmitic acid and oleic acid were the only hydrolysis products classified as active, and by emerging evidence that long-chain fatty acids can modulate oestrogen-receptor-related biology (Frago et al., 2017; Ogata et al., 2025). However, the present outputs are qualitative and do not account for receptor potency, internal exposure, gastrointestinal hydrolysis kinetics, absorbed dose, or tissue concentrations after oral intake. The physiological relevance of these predicted activity patterns therefore remains to be established experimentally.

In contrast, lutein di(pentafluoropropanoate) does not occur naturally and is expected to rapidly hydrolyse (Metličar & Albreht, 2022) to the relatively strong pentafluoropropionic acid (pKa ~ 0.38), a compound predicted here to be active in several models; its activity profile is therefore of greater concern. This halogenated acid is also hazardous to human health according to the European Chemicals Agency (ECHA).

Looking at the influential chemicals from all models (Table S3), we find that retinal occurred in four models: AR-antagonists, ERα-agonists, ERβ-antagonists and GR-agonists. Retinal is produced by living organisms through oxidative cleavage of carotenoids by a dioxygenase. Other influential chemicals were not so closely related to carotenoids as retinal, but were sourced from high-quality public chemistry resources for supporting improved predictive toxicology such as US EPA Distributed Structure-Searchable Toxicity (DSSTox) Database.

Since LEs present in foods as well as in dietary and feed supplements are administered orally, they pass through the gastrointestinal tract before they are finally absorbed into the lymphatic system. During this process, they undergo enzymatic or acidic/alkaline hydrolysis to variable extents. The hydrolysis yields lutein and the carboxylic acid that constitutes the initial LE. According to EFSA and FDA, lutein has a well-established general oral safety profile in nutritional and regulatory contexts and evidence is steadily accumulating that it exerts numerous health-beneficial effects; however, a biological response to the cleaved carboxylic acids should be carefully considered. The degree of hydrolysis and uptake of a particular LE is governed by multiple variables, such as the molecular structure of the compound, food/supplement matrix, specific physiology and metabolism, the delivery method, microbiome composition, age, health conditions, and stress, therefore, it can vary considerably between individuals.

Although some of them are readily found in nature, the thirteen carboxylic acid fragments were evaluated using VEGA-HUB models to assess their predicted activity toward endocrine-relevant nuclear receptor pathways (Table 2). For 46% of these evaluations, VEGA-HUB provided existing experimental annotations from its underlying data sources, which offer additional support for interpretation of the model outputs. For the remaining cases, activity toward androgen, oestrogen, glucocorticoid, RBA, as well as thyroid receptor was predicted and in most cases the compounds were found to be inactive. There were, however, a few exceptions. Predictions for saturated short-chain carboxylic acids (C2-C5) and pentafluoropropionic acid did not exclude the possibility of these compounds to be active toward oestrogen receptors. Yet, nuclear hormone receptors typically bind larger ligands which have aromatic or steroid-like structures. Short-chain carboxylic acids are therefore not expected to interact directly with nuclear receptors in the same way as larger ligands, but increased concentrations of these acids may influence hormone regulation indirectly through metabolic and signalling pathways. It should be stressed that with the exception of pivalic acid (trimethylacetic acid) these small carboxylic acids represent endogenous metabolites which are produced mainly in the gut by fermentation of dietary fiber (Gasaly et al., 2021). Therefore, in healthy individuals these endogenous metabolites would not generally be expected to raise major concern at typical physiological concentrations. Acetic acid and propionic acid are two major gut acids with typical physiological gut concentrations between 10 and 80 mM, while isobutyric, isovaleric, and valeric acids are found at lower levels (< 5 mM) (den Besten et al., 2013; Niccolai et al., 2019). Pivalic acid is an exogenous compound and is mainly released in the liver as a metabolic byproduct of pivaloyl-containing prodrugs (Brass, 2002). The main reason for using pivaloyl structural moiety in the pharmaceutical industry is to improve the drug properties such as absorption, stability, and overall oral bioavailability (Brass, 2002; Bundgaard et al., 1988; Yoshimura et al., 1985). However, there is a possibility that long-term treatment with pivaloyl prodrugs might cause side effects in populations with pre-existing metabolic disorders or malnutrition; pivalic

Table 2
Predictions or experimental values of endocrine activities as provided by six models from VEGA-HUB.

LE	LE fragment	Androgen_COMPARA		Oestrogen_CERAPP		RBA_IRFMN		PGP_NIC		TRα_NRMEA		TRβ_NRMEA	
		Pred.	Exp.	Pred.	Exp.	Pred.	Exp.	Pred.	Exp.	Pred.	Exp.	Pred.	Exp.
2	Acetic acid	0	/	1	/	0	/	0	/	/	0	/	0
3	Propionic acid	0	/	1	/	0	/	0	/	/	0	/	0
4	Pivalic acid	0	/	1	/	0	/	0	/	/	0	/	0
5	Isobutyric acid	0	/	1	/	0	/	0	/	/	0	/	0
6	Isovaleric acid	0	/	1	/	0	/	0	/	/	0	/	0
7	Valeric acid	0	/	1	/	0	/	0	/	/	0	/	0
8	4-Pentenoic acid	0	/	0	/	0	/	0	/	/	0	/	0
9	Benzoic acid	/	0	/	0	0	/	0	/	/	0	/	0
10	Decanoic acid	/	0	/	0	0	/	0	/	/	0	/	0
11	Palmitic acid	/	0	0	/	2	/	0	/	/	0	0	/
12	Oleic acid	/	0	/	0	2	/	0	/	/	0	/	0
13	Pentafluoropropionic acid	0	/	1	/	0	/	0	/	0	/	0	/
14	Phthalic acid	/	0	/	0	/	0	0	/	/	0	/	0

Legend: 0 – non-active; 1 – probably non-active; 2 – active.

acid readily conjugates with carnitine which is efficiently excreted in urine leading to a potential depletion of carnitine and impaired energy metabolism (Brass, 2002). On the other hand, lutein di(pentafluoropropionate) is readily hydrolysed, yielding a perfluoroalkyl acid of greater toxicological concern than the other hydrolysis products considered here (Liu et al., 2019). Although VEGA-HUB did not indicate a strong likelihood of oestrogen-receptor-related activity for this compound, its highly acidic nature and chemical stability still make it concerning from a chemical safety perspective, particularly with respect to potential gastrointestinal injury. There is not much toxicity data available on the compound and there are no established regulatory consumption limits for it, but a much more studied PFAS analogue, perfluorooctanoic acid, is associated with numerous health concerns and is classified as a possible carcinogen to humans by the International Agency for Research on Cancer (IARC). Its maximum permitted concentration in drinking water is set at 4 ng/L by the US EPA CCTE (US Environmental Protection Agency, Center for Computational Toxicology and Exposure, 2022).

Palmitic and oleic acids were the only two compounds classified as active by the VEGA-HUB models, with the IRFMN model predicting notable relative binding affinity for oestrogen receptors.

Considered together with the CPANN results (Table 1), these findings identify lutein dipalmitate and lutein dioleate as the lutein esters most likely to display predicted receptor-activity profiles distinct from lutein. Rather than being contradictory, the two model sets provide complementary information on different exposure forms after oral intake: CPANN describes the intact esters relative to lutein, whereas VEGA-HUB addresses the corresponding hydrolysis products. In this way, the selected receptor-centred panel offers an efficient and biologically meaningful basis for first-step screening and prioritization within the relatively well studied endocrine-relevant pathway space recognized by the Organisation for Economic Co-operation and Development (OECD, 2025). Additional predictive models addressing other relevant targets are currently under development in our laboratory and will support broader future assessment of lutein esters and their hydrolysis products.

4. Conclusion

This *in silico* screening showed that most of the studied lutein esters displayed predicted receptor-activity classifications similar to lutein across the evaluated CPANN models and therefore did not exhibit a markedly divergent receptor-activity profile within the scope of this study. In contrast, lutein didecanoate, lutein dipalmitate, lutein dioleate, and lutein di(pentafluoropropanoate) showed more distinct prediction patterns either as intact esters or through their corresponding gastrointestinal hydrolysis products. Among these, lutein di(pentafluoropropanoate) is of particular concern, because it is rapidly hydrolysed and releases pentafluoropropionic acid, a corrosive compound that can cause severe tissue injury, and should therefore be avoided. Lutein didecanoate, lutein dipalmitate, and lutein dioleate warrant targeted experimental follow-up. For the latter two, interpretation requires particular caution, because these are naturally occurring long-chain lutein esters found in marigold-derived lutein preparations that have an established general safety context. Accordingly, the present predictions should not be interpreted as evidence that currently used marigold-derived lutein products are unsafe, but rather as an indication that these specific ester forms merit closer mechanistic and safety-related evaluation. The remaining hydrolysis products (carboxylic acids) were not classified as notably concerning within the scope of the applied VEGA-HUB models.

Overall, the present work provides a prioritization framework for the further evaluation of lutein esters as potential stabilized lutein ingredients. In particular, it identifies which ester structures should be prioritized for future validation in receptor-specific *in vitro* assays, digestion/bioaccessibility studies, and mammalian cell models, thereby supporting a more evidence-based assessment of their practical application potential.

CRediT authorship contribution statement

Marjan Vračko: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Conceptualization. **Alen Albreht:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank the Slovenian Research and Innovation Agency (ARIS) for financial support under contracts J4-3092, P1-0017, and P1-0005.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodres.2026.119262.

Data availability

Data will be made available on request.

References

Becerra, M. O., Contreras, L. M., Lo, M. H., Díaz, J. M., & Herrera, G. C. (2020). Lutein as a functional food ingredient: Stability and bioavailability. *Journal of Functional Foods., 66*, Article 103771. https://doi.org/10.1016/j.jff.2019.103771

Benfenati, E., Manganaro, A., & Gini, G. (2013). VEGA-QSAR: AI inside a platform for predictive toxicology. In proceedings of the workshop popularize artificial intelligence 2013, Turin, Italy, December 5, 2013. *CEUR Workshop Proceedings, Vol.*, 1107. Available at: https://www.vegahub.eu/.

den Besten, G., van Eunen, K., Groen, A. K., Venema, K., Reijngoud, D. J., & Bakker, B. M. (2013). The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *Journal of Lipid Research., 54*, 2325–2340. https://doi.org/10.1194/jlr.R036012

Brass, E. P. (2002). Pivalate-generating prodrugs and carnitine homeostasis in man. *Pharmacological Reviews, 54*, 589–598. https://doi.org/10.1124/pr.54.4.589

Bundgaard, H., Buur, A., Chang, S.-C., & Lee, V. H. L. (1988). Timolol prodrugs: Synthesis, stability and lipophilicity of various alkyl, cycloalkyl and aromatic esters of timolol. *International Journal of Pharmaceutics, 46*, 77–88. https://doi.org/10.1016/0378-5173(88)90012-9

EFSA Panel on Additives and Products or Substances used in Animal Feed (FEEDAP), Bampidis, V., Azimonti, G., de Lourdes Bastos, M., Christensen, H., Dusemund, B., … Aquilina, G. (2019). Safety and efficacy of lutein and lutein/zeaxanthin extracts from tagetes erecta for poultry for fattening and laying (except turkeys). *EFSA Journal., 17*, Article e05698. https://doi.org/10.2903/j.efsa.2019.5698

EFSA Panel on Additives and Products or Substances used in Animal Feed (FEEDAP), Villa, R. E., Azimonti, G., Bonos, E., Christensen, H., Durjava, M., … Holczknecht, O. (2024). Safety and efficacy of a feed additive consisting of lutein-rich extract of tagetes erecta L. for turkeys for fattening (EW nutrition). *EFSA Journal., 22*, Article e9027. https://doi.org/10.2903/j.efsa.2024.9027

EFSA Panel on Food Additives and Nutrient Sources added to Food (ANS), Aguilar, F., Dusemund, B., Galtier, P., Gilbert, J., Gott, D. M., … Woutersen, R. A. (2011). Scientific opinion on the re-evaluation of lutein (E 161b) as a food additive. *EFSA Journal., 9*, Article 2144. https://doi.org/10.2903/j.efsa.2011.2144

Frago, L. M., Canelles, S., Freire-Regatillo, A., Argente-Arizón, P., Barrios, V., Argente, J., … Chowen, J. A. (2017). Estradiol uses different mechanisms in astrocytes from the hippocampus of male and female rats to protect against damage induced by palmitic acid. *Frontiers in Molecular Neuroscience., 10*, 330. https://doi.org/10.3389/fnmol.2017.00330

Gasaly, N., de Vos, P., & Hermoso, M. A. (2021). Impact of bacterial metabolites on gut barrier function and host immunity: A focus on bacterial metabolism and its relevance for intestinal inflammation. *Frontiers in Immunology., 12*, Article 658354. https://doi.org/10.3389/fimmu.2021.658354

Harikumar, K. B., Nimita, C. V., Preethi, K. C., Kuttan, R., Shankaranarayana, M. L., & Deshpande, J. (2008). Toxicity profile of lutein and lutein ester isolated from marigold flowers (tagetes erecta). *International Journal of Toxicology., 27*, 1–9. https://doi.org/10.1080/10915810701876265

Johnson, M., Basak, S. C., & Maggiora, G. (1988). A characterization of molecular similarity methods for property prediction. *Mathematical and Computer Modelling, 11*, 630–634. https://doi.org/10.1016/0895-7177(88)90569-9

Krstić, Đ., Bolko Seljak, K., Grilc, B., German Ilić, I., & Albreht, A. (2026). In vitro bioaccessibility of lutein diesters in relation to their chemical structure. *Food Chemistry., 509*, Article 148572. https://doi.org/10.1016/j.foodchem.2026.148572

Liu, W., Wu, J., He, W., & Xu, F. (2019). A review on perfluoroalkyl acids studies: Environmental behaviors, toxic effects, and ecological and health risks. *Ecosystem Health and Sustainability., 5*, 1–19. https://doi.org/10.1080/20964129.2018.1558031

Ma, L., & Lin, X. M. (2010). Effects of lutein and zeaxanthin on aspects of eye health. *Journal of the Science of Food and Agriculture., 90*, 2–12. https://doi.org/10.1002/jsfa.3785

Mansouri, K., Abdelaziz, A., Rybacka, A., Roncaglioni, A., Tropsha, A., Varnek, A., … Judson, R. S. (2016). CERAPP: Collaborative estrogen receptor activity prediction project. *Environmental Health Perspectives., 124*, 1023–1033. https://doi.org/10.1289/ehp.1510267

Mansouri, K., Kleinstreuer, N., Abdelaziz, A. M., Alberga, D., Alves, V. M., Andersson, P. L., … Judson, R. S. (2020). CoMPARA: Collaborative modeling project for androgen receptor activity. *Environmental Health Perspectives., 128*, 27002. https://doi.org/10.1289/EHP5580

Manupa, W., Wongthanyakram, J., Jeencham, R., & Sutheerawattananonda, M. (2023). Storage stability and antioxidant activities of lutein extracted from yellow silk cocoons (bombyx mori) in Thailand. *Heliyon, 9*, Article e16805. https://doi.org/10.1016/j.heliyon.2023.e16805

Metličar, V., & Albreht, A. (2022). Esterification of lutein from japanese knotweed waste gives a range of lutein diester products with unique chemical stability. *ACS Sustainable Chemistry & Engineering., 10*, 6072–6081. https://doi.org/10.1021/acssuschemeng.2c01241

Metličar, V., Kranjc, K., & Albreht, A. (2021). Utilization of plant-based wastes for a sustainable preparation of xanthophyll esters via acid anhydrides using β-pinene as a bio-derived solvent. *ACS Sustainable Chemistry & Engineering., 9*, 10651–10661. https://doi.org/10.1021/acssuschemeng.1c04032

Metličar, V., Vovk, I., & Albreht, A. (2019). Japanese and bohemian knotweeds as sustainable sources of carotenoids. *Plants, 8*, 384. https://doi.org/10.3390/plants8100384

Mora Lagares, L., Minovski, N., & Novič, M. (2019). Multiclass classifier for p-glycoprotein substrates, inhibitors, and non-active compounds. *Molecules, 24*, 2006. https://doi.org/10.3390/molecules24102006

Mora-Gutierrez, A., Attaie, R., & Núñez de González, M. (2018). Lutein-enriched emulsion-based delivery system: Impact of casein-phospholipid emulsifiers on chemical stability. In L. Q. Zepka, E. Jacob-Lopes, & V. V. De Rosso (Eds.), *Progress in carotenoid research*. IntechOpen. https://doi.org/10.5772/intechopen.78601.

Niccolai, E., Baldi, S., Ricci, F., Russo, E., Nannini, G., Menicatti, M., … Amedei, A. (2019). Evaluation and comparison of short chain fatty acids composition in gut diseases. *World Journal of Gastroenterology., 25*, 5543–5558. https://doi.org/10.3748/wjg.v25.i36.5543

Novič, M. (2023). Quantitative structure activity/toxicity relationship through neural networks for drug discovery or regulatory use. *Current Topics in Medicinal Chemistry., 23*, 2792–2804. https://doi.org/10.2174/0115680266251327231017053718

OECD. (2025). *Report on the state of the science to address endocrine disrupters under the globally harmonised system of classification and labelling. OECD series on testing and assessment. No. 427*. Paris: OECD publishing. https://doi.org/10.1787/9aa7cd81-en

Ogata, R., Luo, Y., Fujiwara-Tani, R., Sasaki, R., Ikemoto, A., Maehana, K., Sasaki, A., Sasaki, T., Fujii, K., Ohmori, H., & Kuniyasu, H. (2025). Long-chain fatty acids alter estrogen receptor expression in breast cancer cells. *International Journal of Molecular Sciences., 26*, 6722. https://doi.org/10.3390/ijms26146722

Olmedilla-Alonso, B., Granado-Lorencio, F., Castro-Feito, J., Herrero-Barbudo, C., Blanco-Navarro, I., & Estévez-Santiago, R. (2024). Bioavailability of lutein from marigold flowers (free vs. ester forms): A randomised cross-over study to assess serum response and visual contrast threshold in adults. *Nutrients, 16*, 1415. https://doi.org/10.3390/nu16101415

Pap, R., Pandur, E., Jarosa, G., Sipos, K., Agocs, A., & Deli, J. (2021). Lutein exerts antioxidant and anti-inflammatory effects and influences iron utilization of BV-2 microglia. *Antioxidants, 10*, 363. https://doi.org/10.3390/antiox10030363

Rao, A. V., & Rao, L. G. (2007). Carotenoids and human health. *Pharmacological Research, 55*, 207–216. https://doi.org/10.1016/j.phrs.2007.01.012

Seddon, J. M., Ajani, U. A., Sperduto, R. D., Hiller, R., Blair, N., Burton, T. C., … Willett, W. (1994). Dietary carotenoids, vitamins A, C, and E, and advanced age-related macular degeneration. *JAMA : The Journal of the American Medical Association, 272*, 1413–1420. https://doi.org/10.1001/jama.1994.03520180037032

Song, J., Zhang, Y., Wang, H., Wu, C., & Li, Y. (2024). The influence of food matrix and processing methods on the bioaccessibility of lutein: A review. *Journal of Food Bioactives., 26*, 7–23. https://doi.org/10.31665/JFB.2024.18376

Stanojević, M., Sollner Dolenc, M., & Vračko, M. (2023). Predictive models for compound binding to androgen and estrogen receptors based on counter-propagation artificial neural networks. *Toxics, 11*, 486. https://doi.org/10.3390/toxics11060486

Stanojević, M., Vračko, M., & Sollner Dolenc, M. (2023). Development of in silico classification models for binding affinity to the glucocorticoid receptor. *Chemosphere, 336*, Article 139147. https://doi.org/10.1016/j.chemosphere.2023.139147

Talete Srl. (2014). Dragon Software for Molecular Descriptor Calculation. http://www.talete.mi.it/products/dragon_description.htm.

US Environmental Protection Agency. (2016). CompTox Chemicals Dashboard. https://comptox.epa.gov/dashboard/.

US Environmental Protection Agency, Center for Computational Toxicology and Exposure. (2022). Toxicity Forecaster (ToxCast) in vitro assays: Assay documentation for non-guideline in vitro test methods. https://clowder.edap-cluster.com/files/6215520fe4b039b22c7a7836.

US Food and Drug Administration, Center for Veterinary Medicine. (2025). *GRAS response letter for AGRN 75: Tagetes (Aztec marigold) extract. College Park, MD, USA, September 15, 2025. Available at: https://www.fda.gov/media/189932/download*.

Vracko, M. (2005). Kohonen artificial neural network and counter propagation neural network in molecular structure-toxicity studies. *Current Computer-Aided Drug Design., 2*, 563–569. https://doi.org/10.2174/1573409052952224

Xu, L., Wang, W., Bai, K., Wu, Y., Ling, Y., Kong, X., Agusti, R., Qi, Q., Zheng, Z., Yuan, M., Chen, L., Liu, L., Weng, P., & Zhang, Y. (2023). Efficacy mechanisms research progress of the active components in the characteristic woody edible oils. *Food Frontiers., 4*, 1578–1605. https://doi.org/10.1002/fft2.314

Yong, K. T., Yong, P. H., & Ng, Z. X. (2023). Tomato and human health: A perspective from post-harvest processing, nutrient bio-accessibility, and pharmacological interaction. *Food Frontiers., 4*, 1702–1719. https://doi.org/10.1002/fft2.299

Yoshimura, Y., Hamaguchi, N., & Yashiki, T. (1985). Synthesis and relationship between physicochemical properties and oral absorption of pivaloyloxymethyl esters of parenteral cephalosporins. *International Journal of Pharmaceutics, 23*, 117–129. https://doi.org/10.1016/0378-5173(85)90003-1

Zhang, X., Liang, J., Lin, X., Chen, J., & Luo, X. (2024). A comprehensive review on the composition, processing methods, and sustainable utilization of tropical fruit seeds in food industry. *Food Frontiers., 6*, 644–669. https://doi.org/10.1002/fft2.493

Zhang, Y., Li, S., Feng, J., Binkley, L., Tan, L., & Kong, L. (2024). Enhancing lutein stability and bioaccessibility with high internal phase emulsions stabilized by octenylsuccinylated starch. *Food Frontiers., 6*, 316–328. https://doi.org/10.1002/fft2.475