

Design and Synthesis of New Antioxidants Predicted by the Model Developed on a Set of Pulvinic Acid Derivatives

Antoine Le Roux,^{†,‡} Igor Kuzmanovski,^{§,||} Damien Habrant,[†] Stéphane Meunier,[†] Pierre Bischoff,[‡] Brice Nadal,[†] Sophie A.-L. Thetiot-Laurent,[†] Thierry Le Gall,[†] Alain Wagner,^{*,†} and Marjana Novič^{*,§}

[†]Laboratoire des Systèmes Chimiques Fonctionnels, UMR 7199, Faculté de Pharmacie, 74 route du Rhin, BP 24, 67401 Illkirch-Graffenstaden, France

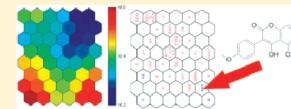
[‡]Laboratoire de Radiobiologie, EA 3430, Université de Strasbourg, Centre Régional de Lutte contre le Cancer Paul Strauss, 3 rue de la Porte de l'Hôpital, BP 42, 67065 Strasbourg, France

[§]Laboratory of Chemometrics, National Institute of Chemistry, Hajdrihova 19, POB 660, SI-1001, Ljubljana, Slovenia

^{||}Institut za hemija, PMF, Univerzitet "Sv. Kiril i Metodij", P.O. Box 162, 1001 Skopje, Macedonia

[†]CEA Saclay, iBiTecS, Service de Chimie Bioorganique et de Marquage, 91191 Gif-sur-Yvette, France

ABSTRACT: Antioxidative activity expressed as protection of thymidine has been investigated for a set of 30 pulvinic acid derivatives. A combination of in vitro testing and in silico modeling was used for synthesis of new potential antioxidants. Experimental data obtained from a primary screening test based on oxidation under Fenton conditions and by an UV exposure followed by back-titration of the amount of thymidine remaining intact have been used to develop a computer model for prediction of antioxidant activity. Structural descriptors of 30 compounds tested for their thymidine protection activity were calculated in order to define the structure–property relationship and to construct predictive models. Due to the potential nonlinearity, the counter-propagation artificial neural networks were assessed for modeling of the antioxidant activity of these compounds. The optimized model was challenged with 80 new molecules not present in the initial training set. The compounds with the highest predicted antioxidant activity were considered for synthesis. Among the predicted structures, some coumarine derivatives appeared to be especially interesting. One of them was synthesized and tested on in vitro assays and showed some antioxidant and radioprotective activities, which turned out as a promising lead toward more potent antioxidants.



1. INTRODUCTION

The ionizing radiation or UV light exposure has a harmful impact on living cells and may affect entire populations. Biological effects of radiation on living cells may result in various outcomes, among which the most harmful is irreparable DNA damage leading to apoptosis, nonlethal DNA mutations that may contribute to the formation of a cancer, or replication and transcriptional errors that may trigger viral interactions. The DNA damage of the radiation or UV light exposure can be tested *in vivo*, which is, however, expensive, time-consuming, and questionable due to ethical reasons (sacrifice of animals). Instead, one may substitute the *in vivo* test by the *in vitro* screening test based on oxidation under Fenton conditions or by UV exposure of thymidine. The titration of the amount of thymidine remaining intact after exposure provides the measure of DNA damage. It is of great interest to find effective radioprotective substances and to determine their protective activity. Many natural and synthetic compounds show antioxidant potency, which can be employed as radioprotective agents. But as testing them on *in vitro* assays represents a huge amount of work, the use of a predicting model, which could rapidly evaluate the potential antioxidant activity of each of them, could appear as an interesting alternative in the course of the research of new antioxidant and radioprotective agents.

In the present study 30 compounds, pulvinic acid derivatives (Table 1) were tested *in vitro* under Fenton conditions and by

UV exposure of thymidine. On the basis of the experimentally determined antioxidative potency of the compounds (physical chemical properties) and their molecular structure (structural properties), the quantitative structure–property relationship models were developed. The optimized models, based on counter-propagation artificial neural networks (CPANN)^{1,2} were tested for prediction of antioxidative potency of a series of new compounds, unknown from the viewpoint of the models developed on the training set of 30 compounds. The goal was to make a prioritization list of 80 chemicals for further consideration of carrying out chemical synthesis. Among four hits from the top with the highest predicted antioxidant activity, one compound, a coumarine derivative, has been synthesized, which turned out to have excellent protective potency under UV irradiation and was moderate under Fenton conditions. It also showed radioprotective activity on a cellular assay.

2. EXPERIMENTAL SECTION

2.1. Training Set Preparation. Compounds present in the training set are derivatives of pulvinic acids, which are antioxidants derived from γ -butenolide, which bear an aromatic group in position 3, a hydroxy group in position 4, and a methylene

Received: May 7, 2011

Published: November 12, 2011

Table 1. Structures of Compounds Used for the Training Set

compound reference	structure	R1	R2	X
t1		/	/	/
t2		/	/	/
type I	t3	Ph	-	OMe
	t4	4-F-Ph	-	OMe
	t5	4-Br-Ph	-	OMe
	t6	4-MeO-Ph	-	OMe
	t7	4-iPr-Ph	-	OMe
	t8	(E)-Ph-CH=CH-	-	OMe
	t9	4-MeO-Ph	-	OH
	t10	4-MeO-Ph	-	Me
	t11	4-MeO-Ph	-	
	t12	4-MeO-Ph	-	OH
	t13	4-MeO-Ph	-	
	t14	4-MeO-Ph	-	
	t15	4-MeO-Ph	-	NH-(CH ₂) ₃ -NH ₂
	t16	4-MeO-Ph	-	N(Me)OMe
	t17	4-MeO-Ph	-	NH-(CH ₂) ₃ -CH ₃
	t18	4-MeO-Ph	-	O <i>i</i> Pr
type II	t19	-	Ph	OMe
	t20	-	4-Br-Ph	OMe
	t21	-	4-MeO-Ph	OMe
	t22	-	4-MeO-Ph	O <i>i</i> Pr
	t23	-	4-MeO-Ph	NH-(CH ₂) ₃ -CH ₃
	t24	4-MeO-Ph	CO ₂ Et	OEt
	t25	4-Br-Ph	4-Br-Ph	OMe
	t26	4-MeO-Ph	4-MeO-Ph	OMe
	t27	Ph	Ph	OMe
	t28	4-F-Ph	4-F-Ph	OMe
	t29		-	-
	t30		-	-

group in position 5, which is functionalized with a carboxylic acid and a second aromatic ring (Figure 1). The hydroxyl function in position 4 is responsible of the radical-trapping activity, as the O—H bond can easily split to form a radical on the oxygen atom (Figure 2).³ This radical is then delocalized in position 2, where it is stabilized by the two adjacent carbonyl groups and the aromatic ring in position 3.

Variations on the substituents in position 3 and on the exocyclic double-bond (modification of R₂ and of the carbonyl function) lead to significant effect on the protection of thymidine under Fenton oxidation and UV irradiation. In particular, replacing the aromatic ring in position 3 by a methyl leads to an important loss of thymidine protecting activity.³ As a consequence of this effect a structure–property relationship study with several pulvinic acid derivatives have been performed, and the results have been used for the prediction model elaboration. Compounds with sufficient diversity have been synthesized, with different R₁, R₂ and X-R₃ groups (Figure 1). When R₁ is an aromatic group, compounds are designated as “type I” products (Table 1,

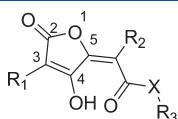


Figure 1. General structure of pulvinic acids.

compounds t3–t18). When R₂ is aromatic, the compounds are designated as “type II” products (compounds t19–t24). When both R₁ and R₂ are the same groups, we thus have products designated as “symmetric” (compounds t25–t28). To this set of compounds have also been added norbadione A (a dimer of pulvinic acids with known antioxidant properties, compound t1),⁴ trolox (vitamin E derivative, compound t2), tetronic acid (compound t29), and an alcohol which is a precursor in pulvinic acid synthesis (compound t30).⁵

2.2. In Vitro Tests. The thymidine protection assay consists of exposing thymidine to two different oxidative stresses in the presence of an antioxidant followed by the measurement of the proportion of intact thymidine remaining after oxidation.^{6,7} The two conditions of degradation are called Fenton condition and UV irradiation: In the first case, thymidine is degraded by oxidants species generated by a mixture of FeSO₄/EDTA/H₂O₂ (1/1/100), and in the second case, thymidine is degraded by UV irradiation in the presence of H₂O₂. The measurement of the concentration of intact thymidine by competitive immunoassay gives the percentage of protection of the tested antioxidant; the higher the percentage, the more protective the compound. Table 2 reports the experimental data obtained with each compound from the training set.

As previously reported,³ type I compounds have a better protecting effect, with 12 compounds over 15 which protect more than 60% of the thymidine under Fenton conditions. On the

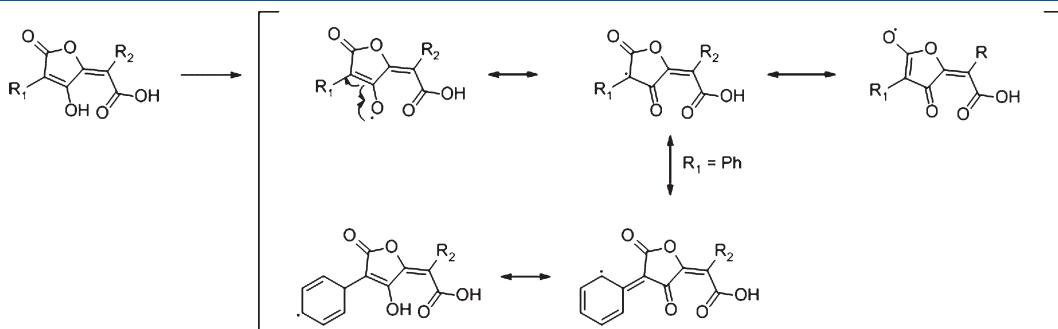


Figure 2. Radical formation and stabilization with pulvinic acids.

Table 2. Experimental Data of Training Set Products, Expressed As the Percentage of Protected Thymidine Under Fenton Conditions and UV Irradiation

compound reference	% protection Fenton	% protection UV	compound reference	% protection Fenton	% protection UV
t1	13	52	t16	84	47
t2	81	10	t17	100	47
t3	91	34	t18	36	27
t4	84	29	t19	19	23
t5	68	26	t20	11	21
t6	89	38	t21	7	19
t7	59	29	t22	11	19
t8	36	31	t23	57	25
t9	69	28	t24	68	18
t10	62	25	t25	100	50
t11	77	35	t26	92	45
t12	72	30	t27	93	42
t13	83	39	t28	90	40
t14	74	31	t29	45	18
t15	37	39	t30	53	23

Table 3. Set of Compounds Used for External Evaluation of the Developed Models

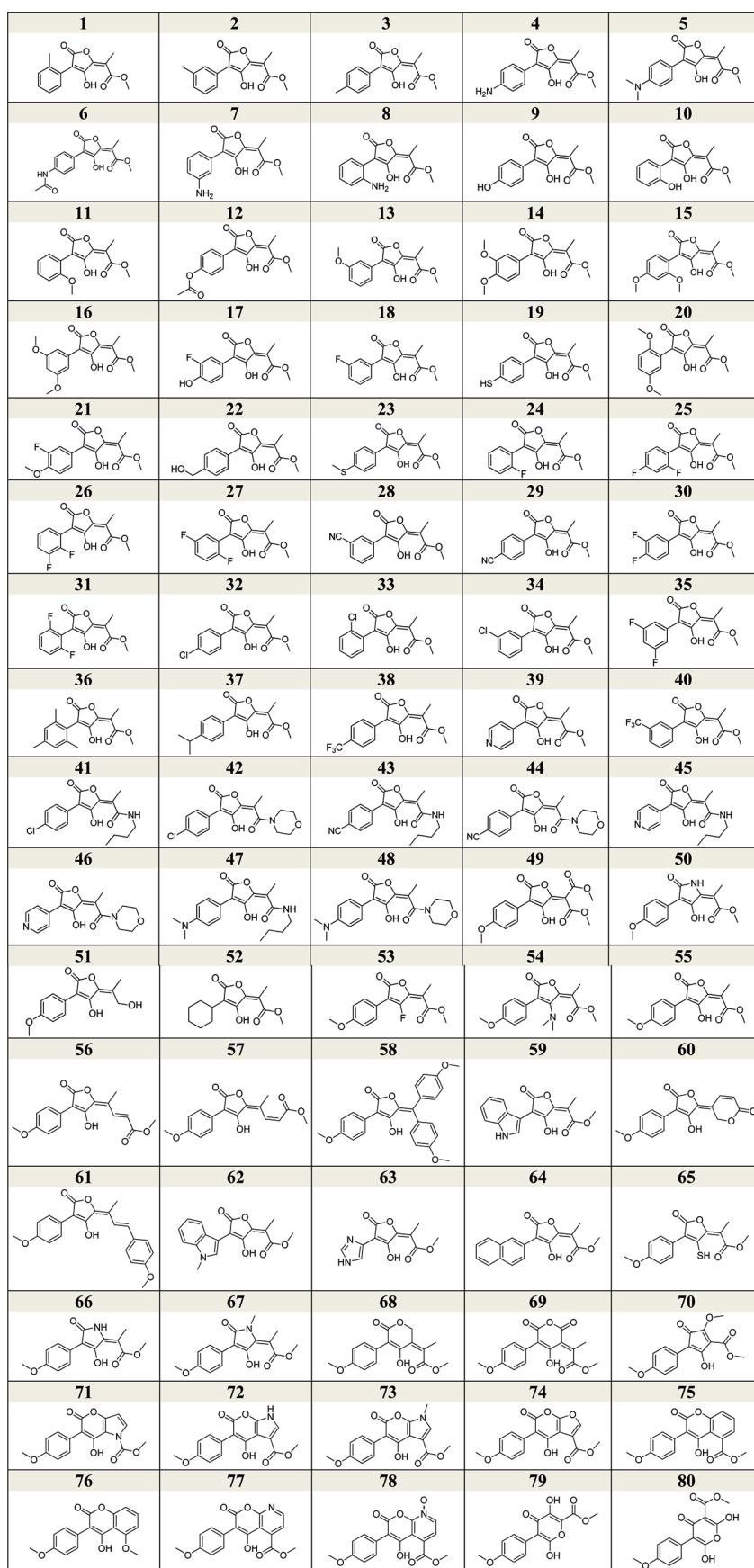
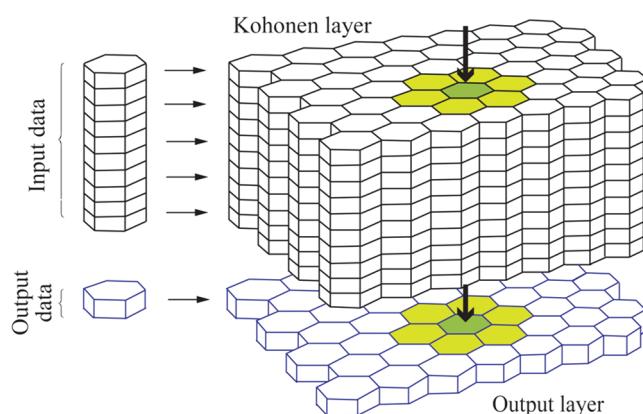


Table 4. The Selected Set of Descriptors Used for Modeling of Thymidine Protective Activity Obtained under Fenton Conditions

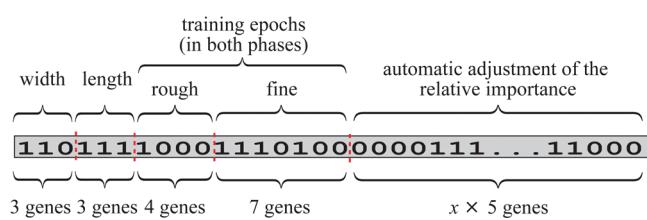
	abbreviation	description
1	JGI8	mean topological charge index of order 8
2	JGI10	mean topological charge index of order 10
3	BELp2	lowest eigenvalue n . Two of Burden matrix/weighted by atomic polarizabilities
4	BELv2	lowest eigenvalue n . Two of Burden matrix/weighted by atomic van der Waals volumes
5	GATS8p	Geary autocorrelation—lag 8/weighted by atomic polarizabilities
6	SIC4	structural information content (neighborhood symmetry of 4 order)
7	MLOGP	Moriguchi octanol—water partition coefficient
8	TPSA(NO)	topological polar surface area using N, O polar contributions
9	Hypertens-80	Ghose—Viswanadhan—Wendoloski antihypertensive-like index at 80%
10	IC3	information content index (neighborhood symmetry of 3 order)
11	BELp3	lowest eigenvalue n . Three of Burden matrix/weighted by atomic polarizabilities
12	GATS4v	Geary autocorrelation—lag 4/weighted by atomic van der Waals volumes

Table 5. Selected Set of Descriptors Used for Modeling of the Thymidine Protective Activity under UV Exposure Conditions

	abbreviation	description
1	Ui	unsaturation index
2	nCbH	number of unsubstituted benzene C(sp^2)
3	ICR	radial centric information index
4	AECC	average eccentricity
5	IDE	mean information content on the distance equality
6	BELm2	lowest eigenvalue n . Two of Burden matrix/weighted by atomic masses
7	TI2	second Mohar index
8	HVcpx	radial centric information index
9	BEHp2	highest eigenvalue n . Two of Burden matrix/weighted by atomic polarizabilities
10	BEHv2	highest eigenvalue n . Two of Burden matrix/weighted by atomic van der Waals volumes
11	BELv2	lowest eigenvalue n . Two of Burden matrix/weighted by atomic van der Waals volumes
12	BELp2	lowest eigenvalue n . Two of Burden matrix/weighted by atomic polarizabilities
13	Yindex	Balaban Y index
14	BELe2	lowest eigenvalue n . Two of Burden matrix/weighted by atomic Sanderson electronegativities

**Figure 3.** Graphical representation of the CPANN in the training phase. The training vector (consisting of input and output data) is introduced to the network. After the best-matching neuron is found, its weights along with the weights of the neighboring neurons are adjusted.

contrary, type II derivatives **t19–t22** do not protect more than 20% of the thymidine in the same conditions. Symmetric compounds led to excellent protecting effect, with more than 90% of the thymidine which remains intact after oxidation.

**Figure 4.** Graphical representation of the encoding of the chromosomes (x represents the number of descriptors).

Trolox also possesses a good effect, with 81% of protection. Norbadione A, **t24**, **t29**, and **t30** have few to moderate effect on the protection of thymidine. When degradation is conducted under UV irradiation, protection values are less contrasted. Type I compounds have better protection values than type II derivatives (25–47% for type I compounds, 19–25%), and symmetric compounds still have excellent protection values with 40–50% protection of thymidine. Trolox and norbadione A have reverse effects, compared to the Fenton oxidation, with, respectively, 10% and 52% of protection. Based on those data, good protection is obtained when more than 80% of the thymidine is protected under Fenton conditions and when more than 35% of protection is obtained under UV irradiation.

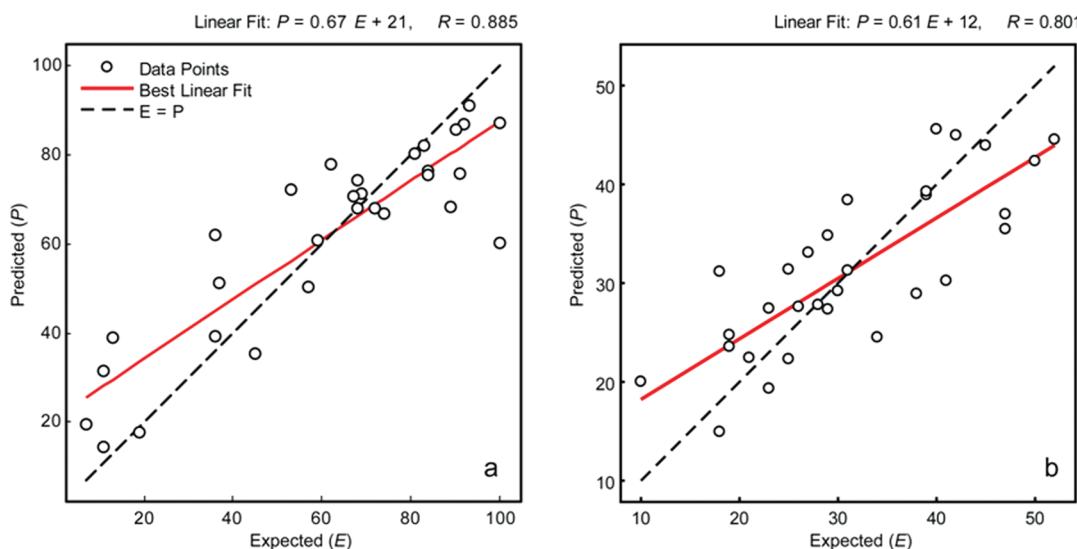


Figure 5. Expected vs found values for the thymidine protective activity: (a) Fenton conditions and (b) UV exposure conditions using the best models obtained using cross-validation leave-two-out.

Table 6. CPANN Parameters (Size of the Network and Number of the Training Epochs) for the Best Models with the RMSEC and for the RMSECV Selected on the Basis of the Lowest Cumulative RMSE from a Large Pool of Tested Models

modeled property	size of the CPANN		training epochs in		RMSEC	RMSECV
	width	length	rough training phase	fine-tuning phase		
Fenton	6	8	13	70	9.08%	13.93%
UV	8	6	22	73	6.40%	11.63%

2.3. Test Set Preparation. The test set is made of 80 compounds. Among them, 67 are similar to pulvinic acids as they all bear a butenolide ring. They diverge from pulvinic acids by the nature of the aromatic R₁ group (differently substituted phenyl, naphthyl, or heterocyclic groups, see Table 3), the replacement of oxygen atoms on the butenolide moiety by other heteroatoms (nitrogen, sulfur, halogen), or the modification or removal of the carbonyl function on the exocyclic double bond (amides, esters, ketones, alcohol, and alkenes). As they are less interesting regarding to their weaker activity, type II compounds were not taken into account in this test set. The 13 other compounds of this set are analogues of pulvinic acids. They have been designed by extending the butenolide ring to six atoms and inserting of the newly created bond in another ring or functionalization of the sixth atom. Those structures can be known in the literature (such as the coumarin core of 75 and 76) or unknown (such as compound 80).

3. COMPUTATIONAL METHODS

In order to develop models suitable for selection of new molecules with high antioxidant activity, which later will be considered for synthesis, as previously stated, predictive modeling was performed using structural descriptors in combination with CPANN.^{1,2} The automated search for the most appropriate network size and the training epochs was performed using genetic algorithms.^{8–10}

3.1. Calculation and Selection of Descriptors. The descriptors used for modeling of the antioxidant activity were calculated

using Dragon 5.5.¹¹ Among all descriptors which were calculated, those with the largest absolute value for the correlation coefficients with the thymidine protective activity obtained under Fenton conditions or UV radiation were selected for further use. Additionally, in order to further reduce the number of descriptors, pair wise correlation coefficients among the descriptors were calculated. Further, among each pair of descriptors with absolute value of the correlation coefficient larger than 0.8, the one descriptor was removed. Among the pairs which fulfilled this condition, whenever possible, the descriptor which was easier for interpretation was kept for further use. Using this approach, the data sets used for modeling of the thymidine protective activity obtained under Fenton conditions were reduced down to 12 descriptors (Table 4), and the data set used for modeling the thymidine protective activity under UV exposure conditions was reduced down to 14 descriptors (Table 5).

3.2. Counter-Propagation Artificial Neural Networks. The learning strategy of for CPANN is very similar to the one used for training of Kohonen artificial neural networks.^{2,12,13} However, unlike Kohonen artificial neural networks, CPANN are composed of two layers.^{2,12,14} The first layer (also called Kohonen layer, Figure 3) serves for mapping of the multidimensional input data (in our case structural descriptors) into two-dimensional map of neurons, while the second layer (also called Grossberg layer)^{15,16} serves as a response surface accessed from the Kohonen layer through a pointing contrivance. During the training phase, after the winning neuron for a training vector is selected, using only the weight levels that correspond to Kohonen layer, the weight levels in both layers are corrected simultaneously.

Once the CPANN is trained the new input vectors, using only Kohonen layer, are mapped, and the predicted values for the new antioxidants are picked up from the output layer.

3.3. Genetic Algorithms. Genetic algorithms (GA) have been proven as an effective algorithm for optimization purpose,^{8–10} allowing relatively fast convergence without the need of evaluation of every possible permutation of the variables. In chemometrics this algorithm is used mainly as a tool for variable selection.^{17–20} In order to avoid the use of trial and error approach in the optimization of the CPANN models, in this study we used GA for finding the most suitable size of the CPANN and for finding the most appropriate number of training epochs for optimization of the CPANN. But also, in order to simplify the models, we used GA for automatic adjustment for the relative importance of the selected descriptors.²¹

In this work we used populations composed of 100 chromosomes. The chromosomes were encoded as presented in Figure 4. The different part of the chromosomes were used for finding the most appropriate size (width and length) of the CPANN and for finding the most appropriate number of epochs in the rough training phase (where a larger neighborhood was affected by the change and larger learning rates were used) and in the fine-tuning phase (where a smaller neighborhood was affected by the change using, at the same time, smaller learning rates). The remaining part of the chromosome was used for adjustment of the relative importance of the descriptors.

Since the models developed in this work are developed using small training data set (composed of 30 structures) and since these models are used for selection of potential substances with good antioxidant activity, the performances of the models were validated using cross-validation leaving two of the training objects out (as an internal validation set). Without using cross-validation, we can easily develop an overtrained model with poor generalization performances. However, using only cross-validation, it is still possible to overtrain the objects which are part of the internal validation set. In order to avoid this, the performance function (Perf) used for evaluation of the fitness of each chromosome was defined as presented here:

$$\text{Perf} = 0.5 \cdot \text{RMSEC} + 0.5 \cdot \text{RMSECV}$$

In the above equation, RMSEC is the root-mean-square error for the prediction of the samples in the training set, while RMSECV represents root-mean-square error for the results obtained using cross-validation. In addition to this, since cross-validation leave-one-out is also prone to overfitting the objects in the internal validation set, in this study we performed cross-validation leaving two samples in the internal validation set.

4. RESULTS AND DISCUSSION

The optimization of the models using GA lasted for 600 generations. Due to the stochastic character of the GA, the entire optimization procedure for modeling of each of the properties representing the thymidine protective activity was repeated several times. The expected vs found values for the modeled properties predicted with cross-validation (using the best models) are presented in Figure 5. The size of the CPANN, the training epochs as well as the RMSEC for the results obtained using cross-validation are presented in Table 6.

The weight level for the modeled variable of the CPANN that corresponds to the model used for prediction of thymidine protective activity obtained under Fenton conditions is presented

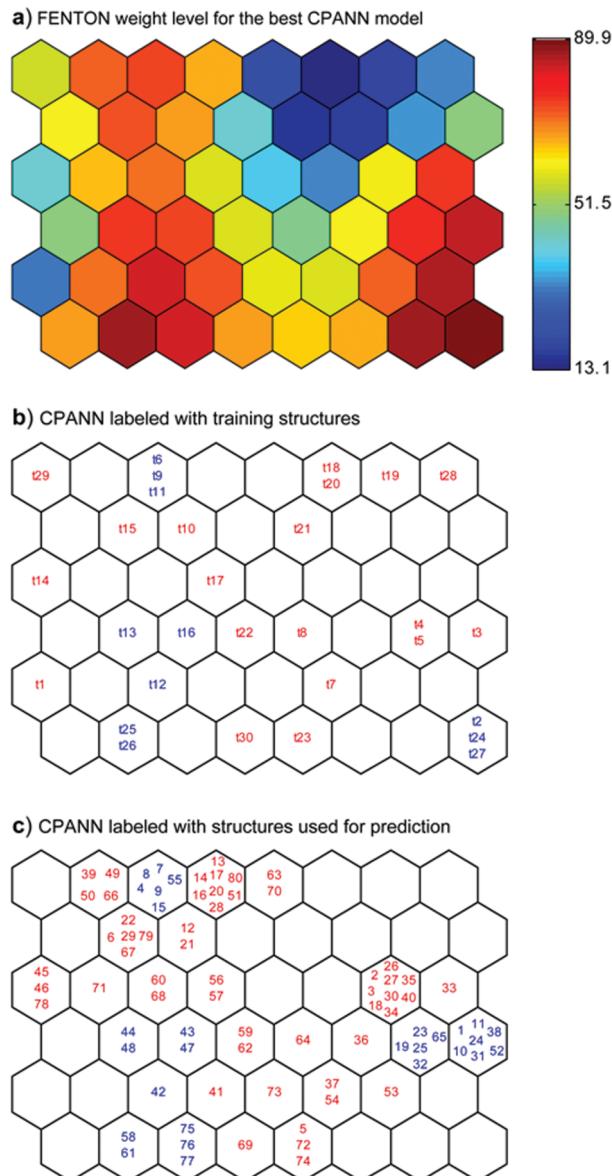


Figure 6. CPANN used for prediction of thymidine protective activity obtained under Fenton conditions: (a) The regions of this weight level labeled with red and dark red color correspond to high protective abilities regions of the CPANN. (b) Top map for the same CPANN with the training compounds mapped on it (marked with the labels taken from Table 1). Blue labels correspond to the structures with the highest protective activity. Red labels correspond to structures with low protective activity. (c) Top map for the same CPANN with the test compounds mapped on it (labels correspond to Table 3). Blue labels correspond to the compounds mapped in the regions of the CPANN that correspond to the regions with higher predicted protective activity. Red labels for the test compounds represent structures mapped in the regions of the CPANN that correspond to small or medium protective activity.

in Figure 6. The weight level for the modeled variable for prediction of thymidine protective activity obtained under UV exposure is presented in Figure 7.

As previously explained (in Section 3.2), the trained CPANN is composed of the response surface accessible in only one point for each input object, thus serving as a pointing device to the

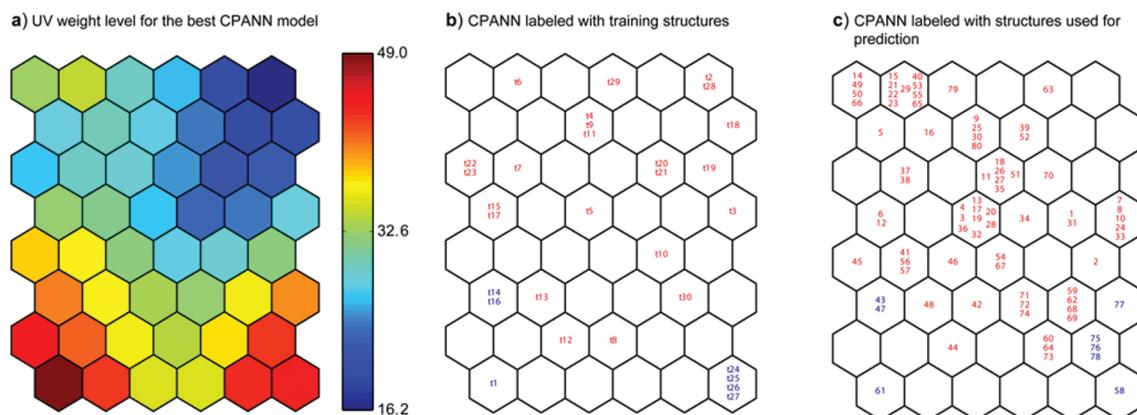


Figure 7. CPANN used for prediction of the thymidine protective activity obtained under UV exposure conditions: (a) The regions of this weight level labeled with red and dark-red color correspond to high-protective abilities regions of the CPANN. (b) Top map for the same CPANN with the training compounds mapped on it. Blue labels correspond to the structures with the highest protective activity. Red labels correspond to structures with low protective activity. (c) Top map for the same CPANN with the test compounds mapped on it (labels correspond to Table 3). Blue labels correspond to the compounds mapped in the regions of the CPANN that correspond to the regions with higher predicted protective activity. Red labels for the test compounds represent structures mapped in the regions of the CPANN that correspond to small or medium protective activity.

Table 7. Structures of the Compounds with the Highest Predicted Activity in UV Irradiation

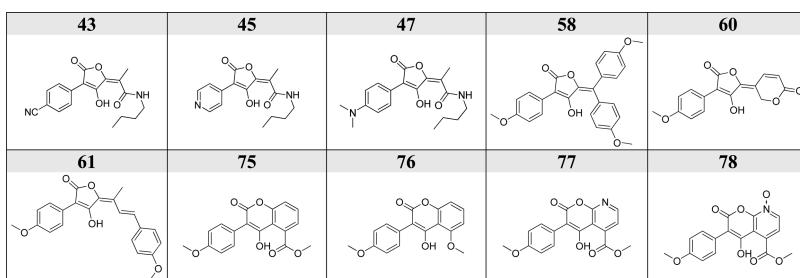
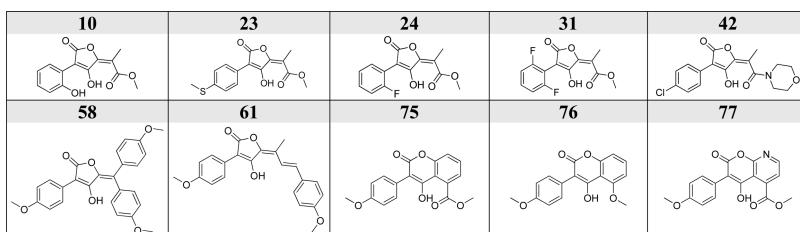


Table 8. Structures of the Compounds with the Highest Predicted Activity in Fenton Condition



trained response area (Figures 6a and 7a). So, the training samples (Figures 6b and 7b) were positioned on the CPANN as a result of the training of the network. However, the 80 structures which were used for the selection of the potential anti-oxidants suitable for synthesis (Table 3) were mapped on the trained CPANN (Figure 6c and 7c), and the predicted antioxidant activity for these structures was picked up from the output layer of the network (Figure 6a and 7a).

The most promising compounds considered for synthesis were not selected using only one model. In order to get more reliable results, the structures with highest potential antioxidant activities were selected on the base of several models. In the case of the selection of new structures with the highest antioxidant

potency obtained under UV exposure conditions, we used nine best models. These models predicted that structures with the highest predicted values for this property were the structures labeled (see Table 7) with numbers: 43, 45, 47, 58, 60, 61, 75, 76, 77, and 78.

Using the 6 best models for prediction of the thymidine protective activity obtained under Fenton conditions, we selected 10 lead compounds. These structures (see Table 8) are labeled as: 10, 23, 24, 31, 42, 58, 61, 75, 76, and 77.

According to those results, we can classify the most promising structures into three categories: the first one is made with pulvinic acid derivatives (structures 10, 23, 24, 31, 42, 43, 45, and 47). These compounds have all in common to bear the

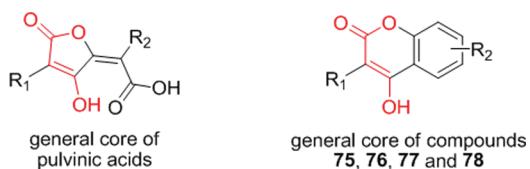


Figure 8. Similarities between pulvinic acids core and 4-hydroxycoumarines core.

Table 9. Comparison between Predicted Activity and Experimental Data for Compound 76

	prediction ^a	experimental ^a
Fenton	83.00%	48.50% \pm 1.38
UV	43.00%	46.56% \pm 0.33

^a Results are given as the percentage of intact thymidine after oxidation.

butenolide ring substituted by a —OH function in position 4 and also to bear a carboxylic acid-derived function (ester or amide) on the double bond. They are directly related to the structures present in the training set. The second category is made of compounds close to pulvinic acids but without the carboxyl acid-derived function on the exocyclic double bond (structures 58, 60, and 61). As explained (Section 2.1), this function plays an important role in the formation of the radical by lowering the energy of the O—H bond, making it more able to undergo homolytic ruptures (see Figure 2). The third category is made with compounds having new structures, different from pulvinic acids (structures 75–78). Those four compounds are derived from 4-hydroxycoumarins. Among the 10 best compounds in Fenton conditions and in UV irradiation, 5 of them have both high predicted activities in UV irradiation and in Fenton conditions: 58, 61, 75, 76, and 77. Two of them are pulvinic acid derivatives, and three are coumarin derivatives. Although coumarines are different from pulvinic acids, those compounds bear in their structure the same bonds which are involved in the radical stabilization (Figure 8). This is interesting, as they potentially have the same mode of action as pulvinic acids. This finding also concerns an important issue of prediction models, the applicability domain, which is defined as the chemical descriptor space spanned by a particular training set of chemicals in which reliable predictions are expected. Our results indicate that the applicability domain of our models was extended beyond simple derivatives of pulvinic acid by a smart selection of descriptors that represent the chemical structure of the compounds.

4-Hydroxycoumarines with R₁ = aryl are not known for having an antioxidant or radioprotective activity. We thus decided to focus on those structures and started with the synthesis of 76. It was tested on the thymidine protection assay, under both Fenton conditions and UV irradiation. Results are presented in Table 9.

Comparison of the experimental assay and predicted values for the antioxidant activity gives us predictions with different quality for the results obtained under Fenton conditions and UV irradiation. Although the expected protection rate in Fenton condition is 83.00%, experimentally only 48.50% of the thymidine is protected, which classified this coumarine derivative as a moderate antioxidant. However, under UV irradiation, a considerably better agreement with the experimental value is observed: with 43.00% expected and 46.56% observed protection rate.

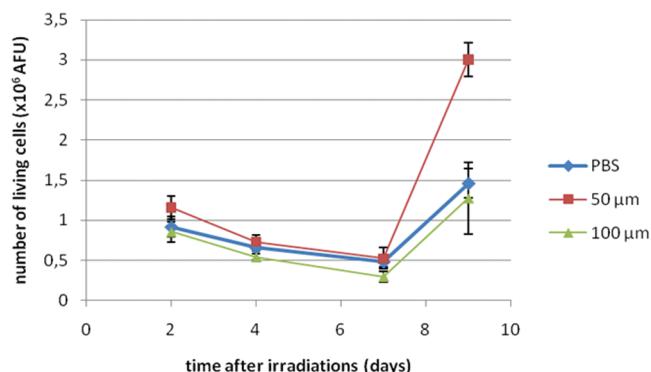


Figure 9. Evolution of the number of living cells irradiated by X-rays (8 Gy, produced by a 6 MV linear medical accelerator), in absence of compounds (PBS) and in presence of 76 at 50 or 100 μ M. Counting is made by fluorescence, and results are expressed as arbitrary fluorescence units (AFU). The number of living cells is proportional to the absorbance.

This result is encouraging, as good prediction has been performed by the model in UV irradiation. With the aim to find and study new radioprotective agents, we tested compound 76 on a radioprotection assay developed in our laboratory.²² It consists of irradiating radiosensitive cell line TK6 with X-rays at a dose of 8 Gy and then to follow the number of living cells in the days after the irradiation. Compound 76 has been tested at two concentrations, 50 and 100 μ M. Results are presented in Figure 9.

During the first seven days, the number of living cells is decreasing, due to a massive loss of viability at this dose of radiation. At the ninth day, both treated and nontreated cells had resumed proliferation. Nontreated cells and cells treated with 76 at 100 μ M have similar proliferative activity, but cells treated by 76 at 50 μ M have an important increasing of the proliferation, with two times more living cells than in the nontreated group. This result indicates a radioprotective effect of compound 76. However the loss of activity at 100 μ M may indicate a possible pro-oxidant activity at high concentration.

5. CONCLUSIONS

We have developed a data-driven prediction model of the antioxidant potency of compounds based on the experimental results obtained on 30 derivatives of pulvinic acids tested on a thymidine protection assay. The entire set of compounds was used for training the prediction model in order to encompass completely the structural domain of the data available. For this reason, only the internal cross-validation was applied to assess the constructed models for their predictive ability. The constructed models enabled us to test a larger set of 80 compounds for their antioxidative potency and to identify the hits from top of the resulting rank list as potential new leads. The rank list served as a prioritization scheme for further synthesis of new antioxidants. An appropriate selection of influential molecular descriptors extended the applicability domain of the model beyond the pulvinic acids derivatives. This model allowed us to discover the antioxidant activity of a new family of compounds, the 4-hydroxycoumarines, because the selected chemical descriptors reflected satisfactorily the antioxidative properties of the molecular subunits common to both families of compounds.

A hypothesis of a mode of action in the radical stabilization was open by analyzing structural features common to the two

families of compounds: pulvinic acids and coumarins. One of the 4-hydroxycoumarins was synthesized and tested on the same assay and showed good antioxidant properties and more interestingly a very good correlation with predicted activity under UV irradiation. Moreover this compound appeared to have radioprotective activity on a cellular assay. Although little correlation was found in Fenton conditions, more refinement should improve its predictive ability. Based on this study, we created an *in silico* model from one family of compounds, and it allowed the discovery the activity of a new family.

The prediction of the antioxidant activity for the newly synthesized substance under Fenton condition was inferior compared to UV irradiation. In order to refine the models and consequently obtain better predictions, more derivatives of 4-hydroxycoumarins will be synthesized and tested on the thymidine assay under Fenton conditions and UV irradiation, and new data will be reinjected in the model. With such models we expect to be able to predict the antioxidant activity of a very large set of new structures, allowing rapidly focusing on the most interesting compounds and thus discovering new potential antioxidants and radioprotective agents.

AUTHOR INFORMATION

Corresponding Author

*E-mail: wagner@bioorga.u-strasbg.fr; marjana.novic@ki.si.

ACKNOWLEDGMENT

The financial support of the European Union in the framework of the Marie-Curie Research Training Network IBAAC project (MCRTN-CT-2003-505020) of the National Science Foundation of France and Direction Général des Armements as well as the Slovenian Ministry of High Education, Science, and Technology (grant P1-017) is gratefully acknowledged.

REFERENCES

- (1) Hecht-Nielsen, R. Counterpropagation networks. *Appl. Opt.* **1987**, *26*, 4979–4984.
- (2) Zupan, J.; Gasteiger, J. *Neural Networks in Chemistry and Drug Design*; WILEY-VCH Verlag GmbH: Weinheim, Germany, 1999.
- (3) Habrant, D.; Poigny, S.; Séguir-Derai, M.; Brunei, Y.; Heurtaux, B.; Le Gall, T.; Strehle, A.; Saladin, R.; Meunier, S.; Mioskowski, C.; Wagner, A. Evaluation of Antioxidant Properties of Monoaromatic Derivatives of Pulvinic Acids. *J. Med. Chem.* **2009**, *52*, 2454–2464.
- (4) Nadal, B.; Thetiot-Laurent, S.A.-L.; Pin, S.; Renault, J.-P.; Cressier, D.; Rima, G.; Le Roux, A.; Meunier, S.; Wagner, A.; Lion, C.; Le Gall, T. Synthesis and antioxidant properties of pulvinic acids analogues. *Bioorg. Med. Chem.* **2010**, *18*, 7931–7939.
- (5) Mallinger, A.; Gall, T. L.; Mioskowski, C. 3-Aryltetronic acids: efficient preparation and use as precursors for vulpinic acids. *J. Org. Chem.* **2009**, *74*, 1124–1129.
- (6) Meunier, S.; Desage-El Murr, M.; Nowaczyk, S.; Le Gall, T.; Pin, S.; Renault, J.-P.; Boquet, D.; Crémignon, C.; Saint-Aman, E.; Valleix, A.; Taran, F.; Mioskowski, C. A powerful antiradiation compound revealed by a new high-throughput screening method. *ChemBioChem* **2004**, *5*, 832–840.
- (7) Meunier, S.; Hanédanian, M.; Desage-El Murr, M.; Nowaczyk, S.; Le Gall, T.; Pin, S.; Renault, J.-P.; Boquet, D.; Crémignon, C.; Mioskowski, C.; Taran, F. High-throughput evaluation of antioxidant and pro-oxidant activities of polyphenols with thymidine protection assays. *ChemBioChem* **2005**, *6*, 1234–1241.
- (8) Holland, J. H. Outline for a logical theory of adaptive systems. *J. Assoc. Comput. Mach.* **1962**, *9*, 297–314.
- (9) Kermani, B.; Schiffman, S.; Nagle, H. T. Using neural networks and genetic algorithms to enhance performance in an electronic nose. *IEEE Trans. Biomed. Eng.* **1999**, *46*, 429–439.
- (10) Henderson, C. E.; Potter, W. D.; McClendon, R. W.; Hoogenboom, G. Predicting aflatoxin contamination in peanuts: A genetic algorithm/neural network approach. *Appl. Intell.* **2000**, *12*, 183–192.
- (11) Dragon for Windows (Software for Molecular Descriptor Calculations), version 5.5 – 2007; Talete SRL: Milano, Italy; <http://www.talete.mi.it/> (accessed October 12, 2011).
- (12) Zupan, J.; Novič, M.; Ruisánchez, I. Kohonen and Counter-propagation Artificial Neural Networks in Analytical Chemistry. *Chemometr. Intell. Lab. Syst.* **1997**, *38*, 1–23.
- (13) Kohonen, T. *Self-organizing maps*, 3rd ed.; Springer: Berlin, Germany, 2001.
- (14) Kuzmanovski, I.; Novič, M. Counter-propagation neural networks in Matlab. *Chemometr. Intell. Lab. Syst.* **2008**, *90*, 84–91.
- (15) Carpenter, G. A.; Grossberg, S. A massively parallel architecture for a self-organizing neural pattern recognition machine. *Comput. Vis. Graph. Image Process.* **1987**, *37*, 54–115.
- (16) Grossberg, S. Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks* **1988**, *1*, 17–61.
- (17) Leardi, R.; Lupiáñez González, A. Genetic algorithms applied to feature selection in PLS regression: how and when to use them. *Chemometr. Intell. Lab. Syst.* **1998**, *41*, 195–207.
- (18) Hasegawa, K.; Miyashita, Y.; Funatsu, K. GA Strategy for Variable Selection in QSAR Studies: GA-Based PLS Analysis of Calcium Channel Antagonists. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 306–310.
- (19) Smith, B. M.; Gemperline, P. J. Wavelength selection and optimization of pattern recognition methods using the genetic algorithm. *Anal. Chim. Acta* **2000**, *423*, 167–177.
- (20) Yoshida, H.; Leardi, R.; Funatsu, K.; Varmuza, K. Feature selection by genetic algorithms for mass spectral classifiers. *Anal. Chim. Acta* **2001**, *446*, 485–494.
- (21) Kuzmanovski, I.; Novič, M.; Trpkovska, M. Automatic adjustment of the relative importance of different input variables for optimization of counter-propagation artificial neural networks. *Anal. Chim. Acta* **2009**, *642*, 142–147.
- (22) Le Roux, A.; Meunier, S.; Le Gall, T.; Denis, J.-M.; Bischoff, P.; Wagner, A. Synthesis and Radioprotective Properties of Pulvinic Acid Derivatives. *ChemMedChem* **2011**, *6*, 561–569.