Radical scavenging activity of natural antioxidants and drugs: Development of a combined machine learning and quantum chemistry protocol

Radical scavenging activity of natural antioxidants and drugs: development of a combined machine learning and quantum chemistry protocol

Cecilia Muraro,[1, a)] Mirko Polato,[2, a)] Marco Bortoli,[1] Fabio Aiolli,[2, b)] and Laura Orian[1, c)]
[1)]*Dipartimento di Scienze Chimiche Università degli Studi di Padova Via Marzolo 1 35131 Padova Italy*
[2)]*Dipartimento di Matematica "Tullio Levi-Civita" Università degli Studi di Padova Via Trieste 63 35121 Padova Italy*

(Dated: 20 July 2022)

Many natural substances and drugs are radical scavengers that prevent the oxidative damage to fundamental cell components. This process may occur via different mechanisms, among which, one of the most important, is hydrogen atom transfer (HAT). The feasibility of this process can be assessed *in silico* using quantum mechanics to compute the ΔG°_{HAT} . This approach is accurate, but time consuming. The use of machine learning allows to reduce tremendously the computational cost of the assessment of the scavenging properties of a potential antioxidant, almost without affecting the quality of the results. However, in many machine learning implementations the description of the relevant features of a molecule in a machine-friendly language is still the most challenging aspect. In this work, we present a newly developed machine-readable molecular representation aimed at the application of automatized machine learning algorithms. In particular, we show an application on the calculation of ΔG°_{HAT} .

I. INTRODUCTION

The term radical was used in chemistry since the 18 th century. However, it acquired its modern meaning in the early 20 th century first, with the discovery of a carbon trivalent species[1] and then, when molecules containing unpaired electrons were seen to be responsible of a very peculiar chemical reactivity.[2] It was however from the beginning of the 21 st century, when the efforts of chemistry and biology fused together to investigate the functioning of living systems, that the recognized importance of radicals in biological processes lead to the birth and development of *redox biology*.[3]

In this relatively young field, radicals, especially free radicals, are commonly associated to oxidative stress, another term that has a somewhat liquid definition,[3] and belong to a wider class of compounds known as reactive oxygen species (ROS) or reactive nitrogen species (RNS). There is a common misconception about free radicals that sees only their harmful effect. On the contrary, radicals play an essential role in the process of redox signaling[4,5] and are therefore the natural products of many biological processes.[6] However, due to their, in many cases, high reactivity, accumulation of these species can indeed result in fatal damage to cellular activity.[6,7] Therefore, the intricate and delicate balance present inside the cell must be maintained at all cost, lest the insurgence of cellular damage leading to severe pathological conditions.[7–11] Sometimes the endogenous mechanisms that maintain this balance can falter and an external source of antioxidants is required to prevent the insurgence of high levels of oxidative stress. Many natural compounds were known since ancient times to cure or alleviate conditions that were much later discovered being caused by oxidative stress, and many other were synthesized in modern times to be used as antioxidant drugs.[12–17]

Unfortunately, there is a vast variety of different molecules that possess known antioxidant properties but, in most cases, the exact mechanism of their action has not been completely elucidated. Among these, flavonoids, a widespread class of polyphenols found mainly in plants, have been recognized as beneficial compounds and employed in different pharmaceutical and cosmetical applications.[18,19] Although their efficiency *in vitro* as antioxidants has been confirmed,[20,21] their *in vivo* role is not completely clear as both prooxidant and antioxidant effects have been observed.[22,23] Different mechanisms for the activity of these molecules have been modeled and in some cases the same compound was found to react in multiple ways, depending on the free radical targeted in the scavenging process.[24–27] Hydrogen atom transfer (HAT) was recognized as one of the most frequent reaction pathway found in flavonoids radical scavenging.[28] It involves the transfer of an hydrogen atom from the scavenger (AH) to the free radical ($R^{\bullet}$):

$$AH + R^{\bullet} \longrightarrow A^{\bullet} + RH \tag{1}$$

A thorough evaluation of free reaction energies (thermodynamics) and activation energies (kinetics) would be, in principle, necessary to correctly assess the most favorable reaction pathway and hence the overall antioxidant capacity of a compound. However, such an investigation on a large number of compounds, each possessing multiple reaction sites is extremely time consuming. As suggested in previous works,[28,29] a simplification of the analysis is possible, provided some quite general assumptions hold in the cases under investigation. For example, this means assuming the validity of the Bell-Evans-Polanyi principle, *i.e.* that the most thermodynamically favored reactions are those that have the highest rates, thus considering only the free reaction energies in the evaluation of the antioxidant capacity of a molecule. Indeed, once the best compounds have been identified, a complete study comprising both thermodynamics and kinetics is of course needed to confirm the preliminary results.

With the enormous advancement in computing power

[a)]These authors equally contributed to this work.
[b)]Electronic mail: fabio.aiolli@unipd.it (corresponding author)
[c)]Electronic mail: laura.orian@unipd.it (corresponding author)

achieved since the beginning of the 21 st century, computer-aided molecular design, nowadays a routine practice in many research groups and laboratories, can be a powerful tool to assess the antioxidant efficiency of a compound and determine its mechanism of action.[26,29–31] As more details on the activity of efficient antioxidant molecules emerge, new rational routes on the improvement of both the natural and synthetic species can be developed to obtain even more efficient and more biocompatible structures. Due to the huge number of compounds that could display antioxidant activity and the presence of different mechanisms for the radical scavenging,[24–27] a challenging problem arises when evaluating the direction in which to go to obtain the desired enhancements. Even though quantum mechanics (QM) calculations are becoming more and more affordable and computing power is constantly increasing, the time needed to describe completely the energetics of the mechanism of action of a common antioxidant is still in the order of magnitude of hours, if not days. This is mostly due because, if no prior knowledge has been obtained, radical scavenging reactions could in principle take place involving different sites of an antioxidant, creating a substantial number of possible pathways. For example, in the specific case of HAT, different H• can be transferred in the scavenging of the free radicals, with each process having a different Gibbs free energy of reaction. A possible way to decrease the computational cost to study these reactions would be to perform a preemptive screening of the different sites to determine which are reactive (*i.e.* which are very likely to have a negative free energy of reaction). In this case, far lesser expensive techniques could be employed and accurate calculations could be performed only on a reduced subset of the original possible routes. To this end, machine learning (ML) presents a very appealing method that can act as an initial, computationally inexpensive approach to separate the favored from the unfavored reaction paths.

Machine learning in chemistry has lately attracted much interest due to the very good results it can achieve with very limited computational effort.[32–35] More specifically, the branch of techniques belonging to the supervised learning area[36] has found wide application in the estimation of enthalpies of formation,[37] melting points,[38] reorganization energies,[39] solubilities,[40] and other relevant chemical properties.[41–45] These methods are able to infer a function that can predict the (initially unknown) properties of a general set of molecules, provided a starting set of similar molecules with known properties is given.[36] A very challenging aspect of the application of these algorithms in chemistry is to provide a suitable machine-readable descriptor that contains all the required chemical information, *i.e.* develop a correct data representation. Many options on how to perform this classification have been explored (for an exhaustive overview, which is out of the scope of this work, the reader can refer to a recent review article[46]) and in some cases machine learning itself has been applied to recognize the most meaningful molecular descriptor for a given set of molecules.[47] However, in most cases data representations remain deeply linked to the particular task they are applied to and to some *a priori* knowledge. Therefore, human input is in many cases required to *tailor* the representation to the required application.

In this work, we have developed a new model to create a simple and machine-friendly representation that can capture the relevant features needed to describe a large set of chemically different hydrogen atoms. This representation was applied in a combined QM/ML protocol for the assessment of the reactivity of HAT reactions on a class of mainly natural antioxidant compounds, namely flavonoids. Additionally, some very well known molecules, developed originally to address other conditions, were recently shown to possess promising antioxidant properties[48–53] ascribed to HAT mechanism.[26,52] Therefore,a set of analogues of fluoxetine (N-methyl-3-phenyl-3-[4-(trifluoromethyl)phenoxy]propan-1-amine), were also included in this study. We chose a selection of free radicals that contains the most common species encountered in cellular biology to obtain a general picture of the overall reactivity of a compound in the different situations that can be found *in vivo*. We selected a machine learning algorithm based on a classification method to categorize the hydrogen atoms into two groups: reactive (*i.e.* those with $\Delta G^{\circ}_{HAT} < 0$) and non-reactive (*i.e.* those with $\Delta G^{\circ}_{HAT} > 0$). A database of Gibbs free energies of reaction for a large set of HAT processes was obtained through DFT calculations in a automatized manner. This allowed a successful training of the machine learning algorithm and the prediction of the antioxidant capacity of a large set of compounds was obtained with satisfactory accuracy and with a minimal computational effort.

II. THE QM/ML PROTOCOL

A. QM Computational protocol

All geometry optimizations were performed in gas phase without any constraint, using the M06-2X functional combined with the 6-31G(d) basis set, as implemented in Gaussian16.[54] This functional is a hybrid meta-generalized gradient-approximations (hybrid meta-GGA) functional, developed by Truhlar *et al.*[55,56] M06-2X has been chosen because it has already been used in literature for similar systems and has provided very satisfactory results.[31,52,57] Spin contamination was checked for the doublet ground state species to assess the reliability of the wave function. Frequency calculations were carried out at the same level of theory (M06-2X/6-31G(d)) in order to confirm the nature of the energy minima (all positive frequencies) and to obtain the thermodynamic corrections at 1 atm and 298 K. Single point energy calculations were performed at M06-2X/6-311+G(d,p) in gas phase in order to obtain more accurate energy values, and subsequently, at the same level of theory, in benzene and water using the SMD continuum model.[42] This level of theory is denoted in the text (SMD)-M06-2X/6-311+G(d,p)//M06-2X/6-31G(d,p). Benzene and water mimic an apolar and a polar environment, respectively,[57] which could in principle modify the reactivity of an antioxidant.[29]

B. Learning method

1. Data representation

In supervised machine learning, specifically in classification contexts, the data set is composed of a set of n pairs $\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)\}$ where $\mathbf{x}_i \in \mathscr{X}$ are called examples (or patterns), and $y_i \in \{-1, 1\}$ are the labels associated to the examples. Given an example $\mathbf{x}'$, the task of the ML method is to classify $\mathbf{x}'$ correctly, i.e., it has to associate to $\mathbf{x}'$ with the label y'. The set $\mathscr{X}$ is the so-called input space in which examples lie, and each dimension of an input example is called feature. To successfully apply a ML algorithm, the chosen representation for the examples is crucial. In our task at hand, i.e., reactive vs. non reactive hydrogen atoms in a molecule, we represent examples in term of hydrogen atoms themselves. So, our representation is a site-based representation rather than the usual molecule-based one. Thus, an example is characterized by features extracted for a particular site in a molecule. Importantly, the representation of the sites is needed only in the starting antioxidant molecule (AH in (1)) and no information is required on the topology of the hydrogen atoms found in the radicals formed after the hydrogen transfer. This is another clear advantage over the quantum mechanical method used to study chemical reactivity as it requires the precise knowledge of the structures (and hence the energies) of both the reactants and products of a reaction.

Since we wanted to build a generic representation that can be used with any classifier, we needed to define a unified sites representation. For this reason the number of features of the proposed representation is molecule independent. Specifically, we included two sets of features:

- **statistical**: this set of features contains statistics about the geometrical characteristics of the site w.r.t. the other atoms in the molecule. Table I summarizes these features. As evident in the table, some features are computed using the overall distances from the sites and types of atoms, while other are computed independently for each type of atoms.
- **domain-specific**: this set of features has been added after some trials, and has been directly suggested by the experts. As of many other applied ML contexts, the *a priori* knowledge is always beneficial to improve the model accuracy. In particular these features include the number of atoms that are within a well-defined range of distances from the site. Three intervals has been selected: up to 2.2 Å, from 2.2 to 2.7 Å and from 2.7 to 3.0 Å.

At the end of this pre-processing, we extracted, for each site, a total of 79 features, i.e., $\mathbf{x}_i \in \mathbb{R}^{79}$.

Finally, the label associated to an example, i.e. the site, is (ΔG°_{HAT}) for that particular site. We created a new example-label pair for all the free radical/environment combinations (15 values total, *vide infra*). This value was then used to generate the classification label according to the threshold 0, i.e., $\Delta G < 0$ the site has negative label, positive otherwise. There is an exception for the HO$^{\bullet}$ ROS because the threshold 0 produced an highly unbalanced data set. In fact, due to the high reactivity of HO$^{\bullet}$ (the most reactive of the radicals here considered) only few of the sites are above such threshold and hence we moved it to -30.

Overall	For each atom type
average distance	average distance
variance of distance	variance of distance
total number of atoms	number of atoms
	minimum distance
	maximum distance

TABLE I. Statistical features extracted from the sites/molecule. All the distances are w.r.t. the site.

		Prediction outcome		
		p	**n**	**total**
actual value	**p′**	**T**rue **P**ositive	**F**alse **N**egative	P′
	n′	**F**alse **P**ositive	**T**rue **N**egative	N′
	total	P	N	

TABLE II. Binary classification confusion matrix used to compute accuracy and balanced accuracy.

2. Classification

To classify between reactive and non-reactive sites, we employed a soft-margin SVM[58]. In our experiments, two kernel functions have been tested: RBF (Radial Basis Function), and a polynomial kernel[58]. The hyper-parameters have been validated using a nested 5-fold cross validation procedure with grid search. The hyper-parameter γ, that controls the width of the Gaussian function for the RBF kernel, has been validated in the range $[2^{-15}, \ldots, 2^{4}]$, while the degree of the polynomial kernel has been validated in the set $\{1, 2, \ldots, 6\}$. Finally, the trade-off parameter C of the soft-margin SVM has been validated in the range $[10^{-2}, \ldots, 10^{5}]$.

To evaluate the proposed method, we used both accuracy and balanced accuracy that are computed as follows

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{P} + \text{N}}$$

$$\text{balanced accuracy} = \frac{1}{2}\frac{\text{TP}}{\text{P}}\frac{\text{TN}}{\text{N}}$$

where TP, TN, P and N are defined by the standard confusion matrix for binary classification shown in Table II.

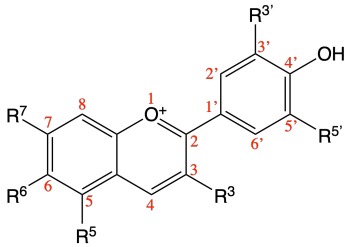

FIG. 1. General structure of an anthocyanidin

Balanced accuracy is preferred to the standard accuracy because in unbalanced databases the examples belonging to the less represented class are often misclassified but the effect on the standard accuracy would be negligible.

The results in Table III are the average (with standard deviation) over 10 repeated experiments with different partitioning of 80% training and 20% testing data. We only report the results achieved with the RBF kernel because it has always provided better performance than the polynomial kernel.

III. RESULTS AND DISCUSSION

We selected five species that represent the most common types of free radicals found in biological environments, namely $HO^{\bullet}$, $CH_3O^{\bullet}$, $HOO^{\bullet}$, $CH_3HOO^{\bullet}$ and $CH_2=CHOO^{\bullet}$. The reaction energies were obtained with DFT calculations employing the M06-2X meta hybrid functional.[56] Three different reaction environments were included, *i.e.* gas-phase, water and benzene. For reactions in condensed phase the solvation model based on solute electron density (SMD) was employed.[42] We therefore obtained the ΔG°_{HAT} for $3N$ possible reactions, where N is the number of sites from which a hydrogen atom can be removed.

First, an initial explorative attempt was made to apply this novel model to a limited group of molecules, namely anthocyanidins, a group of natural dyes belonging to the family of flavonoids (Figure 1).[59] They were chosen because they have high antioxidant activity against ROS, verified *in vivo*[60,61], with potential beneficial effects such as anti-inflammatory and chemoprotective and cancer prevention properties[62]. We chose a set of 12 anthocyanidins for a total of 162 reactive sites. We ascertained the reactivity of these 162 sites towards our selection of five ROS and in the three chosen environments (gas-phase, water and benzene). Each anthocyanidin was fully optimized in gas-phase and all the radicals obtained after hydrogen transfer were optimized too, for a total of 162 calculations. Energies in the condensed phase were obtained with single point computations on the gas-phase geometry. From this data, ΔG°_{HAT} were calculated considering all five free radicals. On average,calculation on each site took about 30 minutes on a 16 core machine. This has to be multiplied for the total number of sites (162) affording about 81 hours (approximately 1300 core-hours). Results confirmed the predicted reactivity of molecular sites towards HAT, as OH moieties were seen to result in those with the largest negative ΔG°_{HAT} (Figure 2 and Supporting Information). On the other hand, aliphatic and aromatic groups were calculated to have a positive ΔG°_{HAT} with all the radicals except $HO^{\bullet}$ which was seen to react with negative ΔG°_{HAT} in all the cases considered. (Figure 2).

The obtained ΔG°_{HAT} values were divided in 15 different data sets (based on environment and free radical involved), using which the ML algorithm was trained. For reference time needed for the evaluation of ΔG°_{HAT} with the ML method was less than a minute per data set on a 4 core machine.

After obtaining very good results for this first group of compounds (see Supplementary Information), the data set has been extended with new molecules from two main families of compounds, chain breaker natural antioxidants, mainly flavonoids, and psychoactive molecules analogous to fluoxetine, due to the fact that both classes of compounds have been found to react mainly via the HAT mechanism. Thus the full database contains 148 molecules with 2118 sites.

We evaluated the reactivity in HAT reaction involving five different ROS and in three separate environments (gas-phase, water and benzene). All the scavengers were represented with a site-based description that contained data divided into three groups.

Table III shows the performance of the ML algorithm. Results are very encouraging since for all 15 combinations of ROS/environment an accuracy higher than 90% has been achieved, with an overall performance of 93.3%. This means that nine time out of ten our method is able to correctly predict if a HAT process is thermodynamically favored, based only on the starting structure of the involved scavenger. Balanced accuracy is, on average, lower than the accuracy (with the only exception of $CH_3O^{\bullet}$) but its overall score is higher than 80%. This is a rather promising result since some of the classification tasks were highly unbalanced (e.g., $\sim$ 90%-10%). We want to stress the fact that these numbers indicate the predictive power of the procedure, *i.e.* the ability to discriminate between reactive ($\Delta G^{\circ}_{HAT} < 0$) and nonreactive ($\Delta G^{\circ}_{HAT} > 0$) sites. In fact, our method is not aimed at predicting the values of ΔG°_{HAT} directly using ML but at recognizing the reactive sites of a potential antioxidant.

Accuracy-wise the best results were obtained with the most reactive free radicals (*i.e.* $CH_3O^{\bullet}$, $HO^{\bullet}$ and $CH_3OO^{\bullet}$) whereas the lowest performance was found for $CH_2=CHOO^{\bullet}$. This behavior can be explained looking at how many reaction sites have a ΔG that falls within the $\pm 2\,kcal\,mol^{-1}$. These "borderline" reactions are the hardest to classify as they are very close to the boundary chosen to discriminate between reactive and non reactive sites. For example, the least performing set (*i.e.* the reactions with $CH_2=CHOO^{\bullet}$ in gas-phase) presents a total of 162 sites in the $+2\ -2\,kcal\,mol^{-1}$ interval (effectively making up for 7.9 % of the overall sites of the set) whereas in a better performing series of reactions($CH_3O^{\bullet}$ in gas-phase) only 2.6 % of the processes fall within these limits (*i.e.* 53 reactions). From a balanced accuracy perspective, performance is around 80% (4 out of 5 times the classifier is right) with peaks of 96% for $CH_3O^{\bullet}$. The worst performing experiments are the ones concerning $CH_3OO^{\bullet}$ (on average 68%) that is the most unbal-

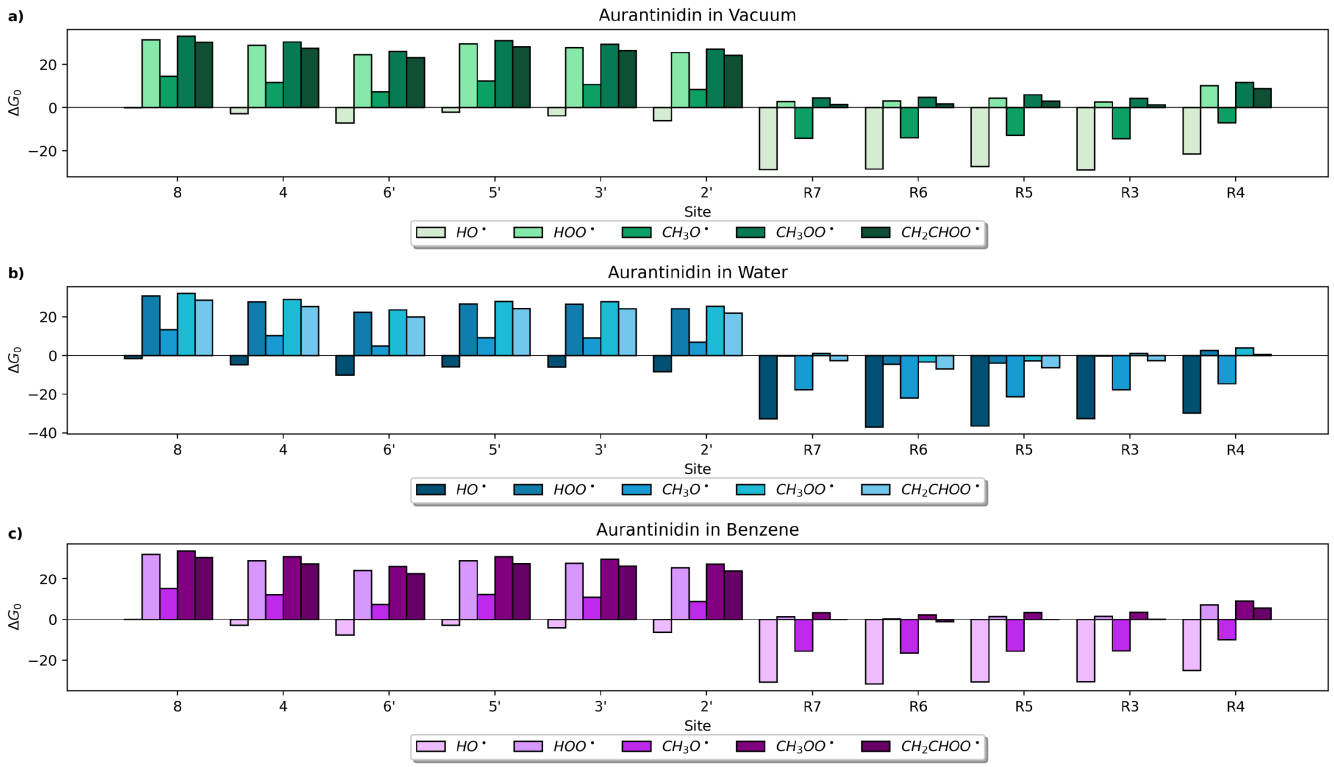

FIG. 2. ΔG°_{HAT} for aurantinidin (R3', R5'= H; R5, R6, R7 = OH) in gas-phase (a), in water (b) and in benzene (c). Site numbering is shown in Figure 1. Level of theory (SMD)-M06-2X/6-311+G(d,p)//M06-2X/6-31G(d,p)

ROS	solvent	accuracy [%]	balanced acc. [%]
$CH_2{=}COO^{\bullet}$	Benzene	92.3 ±0.98	81.4 ±3.69
	Gas-phase	91.2 ±1.65	73.1 ±6.86
	Water	91.5 ±1.01	83.3 ±2.51
$CH_3O^{\bullet}$	Benzene	96.6 ±0.63	96.4 ±0.67
	Gas-phase	96.7 ±0.66	96.4 ±0.73
	Water	96.4 ±0.51	96.2 ±0.53
$CH_3OO^{\bullet}$	Benzene	93.8 ±0.86	63.7 ±5.01
	Gas-phase	94.2 ±0.84	63.4 ±5.06
	Water	93.2 ±0.76	76.2 ±4.14
$HO^{\bullet}$ (*)	Benzene	92.8 ±0.83	83.1 ±2.96
	Gas-phase	92.3 ±1.05	76.8 ±3.08
	Water	91.3 ±1.07	84.1 ±1.82
$HOO^{\bullet}$	Benzene	92.3 ±0.89	74.9 ±3.64
	Gas-phase	92.2 ±1.01	70.0 ±4.38
	Water	92.5 ±1.05	82.5 ±3.11
Overall		93.3	80.1

TABLE III. Accuracy and balanced accuracy (along with standard deviation) of the proposed ML method for the 15 combinations of solvents and ROS. (*) for $HO^{\bullet}$, we used ΔG°_{HAT} > -30 to divide positive and negative class for unbalancing reasons.

anced ROS in all the solvent.

IV. CONCLUSIONS

The use of machine learning in chemistry is getting more and more popular, thus opening new horizons for computational chemistry. A new and challenging area of research has come to light: the translation of the molecular features of a molecule from human-friendly to machine-friendly representations. In this work, we have presented a novel machine-readable representation based on simple chemically relevant parameters to describe the molecular topology. This newly derived definition of the structure of a molecule was employed in the application of a ML algorithm on a series of biologically relevant reactions, namely a series of hydrogen atom transfers involving a wide range of antioxidant molecules selected among flavonoids, a class of natural antioxidants, and analogues of fluoxetine, a well known drug that showed promising radical scavenging properties. Five ROS were considered and the processes were evaluated in three different environments to have a broad spectrum of reactions that cover most of the possible *in vivo* scenarios.

Our focus was directed to a site based representation that was convenient to assess the reactivity of each hydrogen atom that could be transferred. The site-based description of each hydrogen atom taking part in a HAT consisted of a very simple collection of geometrical parameters (pairwise atomic distances from that particular site with the respective atomic numbers of the atoms considered) coupled with the ΔG°_{HAT} of that particular hydrogen atom transfer reaction with each

of the selected radicals. From this data, two sets of relevant features were identified and extracted. The first contained the distance of the closest and furthest atom to the site, the average distance of the atoms from the site, the total number of atoms in the molecule and the standard deviation of the atomic distances. The second included a broader set of features, namely the presence of atoms within certain distance intervals from the reaction site. These two models were tested in the application of a ML algorithm trained to predict the feasibility of a HAT on a particular site. The whole process of QM calculations, features extraction and reactivity prediction was completely automatized. While even the first and most basic set of relevant features gave good results, it was the second that, albeit being still very simple, improved noticeably the performance of the applied machine learning algorithm. In particular, the inclusion of the number of atoms present at specific distance intervals from the represented site was the key addition. This choice stemmed from the idea that the chemical properties of a particular hydrogen atom can be defined by the presence (or absence) of a certain number of atoms within well defined distance ranges. The improved results of the application of this new model confirmed this initial hypothesis.

Importantly, this study paves the way to relevant future activities, among which (a) extension of the database to include more molecular topologies to expand the applicability of the screening algorithm; (b) extension of the protocol to other reaction mechanisms which are important when assessing the radical scavenging potential of a molecule; (c) refinement and improvement of the molecular representation, an effort which requires strong synergy between information scientists and chemists. As first step, indeed this work represents an example of machine learning meeting chemistry!

ACKNOWLEDGEMENTS

L.O. acknowledges Università degli Studi di Padova for financial support P-DiSC (BIRD2018-UNIPD) project MAD^{3}S (Modeling Antioxidant Drugs: Design and Development of computer-aided molecular Systems). M. B. is grateful to Università degli Studi di Padova for a grant in the framework of the MC^4 (Molecole al Computer: Calcoli per la Chimica in Classe) and to the teachers of six secondary schools of regione Veneto who, in 2019, were trained with selected students in the fundamentals of the chemistry of antioxidant radical scavengers and could approach to supercomputing in chemistry. All the calculations were carried out on Galileo (CINECA: Casalecchio di Reno, Italy) thanks to the ISCRA Grant MC^4 (Molecole al Computer: Calcoli per la Chimica in Classe), P.I.: L.O.

DATA AVAILABILITY

The data that supports the findings of this study are available within the article and its supplementary material or from the corresponding author upon reasonable request.

REFERENCES

[1]M. Gomberg, "An instance of trivalent carbon: Triphenylmethyl," J. Am. Chem. Soc. **22**, 757–771 (1900).
[2]M. S. Kharasch and F. R. Mayo, "The Peroxide Effect in the Addition of Reagents to Unsaturated Compounds. I. The Addition of Hydrogen Bromide to Allyl Bromide," J. Am. Chem. Soc. **55**, 2468–2496 (1933).
[3]H. Sies, "Oxidative stress: A concept in redox biology and medicine," Redox Biol. **4**, 180–183 (2015).
[4]H. Sies, "Hydrogen peroxide as a central redox signaling molecule in physiological oxidative stress: Oxidative eustress," Redox Biol. **11**, 613–619 (2017).
[5]H. Sies, "On the history of oxidative stress: Concept and some aspects of current development," Current Opinion in Toxicology **7**, 122–126 (2018).
[6]J. Kehrer, J. Robertson, and C. Smith, "1.14 - free radicals and reactive oxygen species," in *Comprehensive Toxicology (Second Edition)*, edited by C. A. McQueen (Elsevier, Oxford, 2010) second edition ed., pp. 277 – 307.
[7]J. P. Silva and O. P. Coutinho, "Free radicals in the regulation of damage and cell death - basic mechanisms and prevention," Drug Discoveries Ther. **4**, 144–167 (2010).
[8]D. Dreher and A. F. Junod, "Role of Oxygen Free Radicals in Cancer Development," European Journal of Cancer **32**, 30–38 (1996).
[9]B. Halliwell, "Biochemistry of oxidative stress: Figure 1," Biochem. Soc. Trans. **35**, 1147–1150 (2007).
[10]S. Reuter, S. C. Gupta, M. M. Chaturvedi, and B. B. Aggarwal, "Oxidative stress, inflammation, and cancer: How are they linked?" Free Radical Biol. Med. **49**, 1603–1616 (2010).
[11]A. Phaniendra, D. B. Jestadi, and L. Periyasamy, "Free Radicals: Properties, Sources, Targets, and Their Implication in Various Diseases," Indian J. Clin. Biochem. **30**, 11–26 (2015).
[12]B. Uttara, A. Singh, P. Zamboni, and R. Mahajan, "Oxidative Stress and Neurodegenerative Diseases: A Review of Upstream and Downstream Antioxidant Therapeutic Options," Curr. Neuropharmacol. **7**, 65–74 (2009).
[13]L. Orian and S. Toppo, "Organochalcogen peroxidase mimetics as potential drugs: A long story of a promise still unfulfilled," Free Radical Biol. Med. **66**, 65–74 (2014).
[14]L. P. Wolters and L. Orian, "Peroxidase activity of organic selenides: Mechanistic insights from quantum chemistry," Curr. Org. Chem. **20**, 189–197 (2016).
[15]M. D. Tiezza, G. Ribaudo, and L. Orian, "Organodiselenides: Organic Catalysis and Drug Design Learning from Glutathione Peroxidase," Curr. Org. Chem. **23**, 1381–1402 (2019).
[16]K. Neha, M. R. Haider, A. Pathak, and M. S. Yar, "Medicinal prospects of antioxidants: A review," European Journal of Medicinal Chemistry **178**, 687–704 (2019).
[17]D. Komsiiska, "Oxidative stress and stroke: a review of upstream and downstream antioxidant therapeutic options," Comp. Clin. Pathol. **28**, 915–926 (2019).
[18]P. G. Pietta, "Flavonoids as antioxidants," J. Nat. Prod. **63**, 1035–1042 (2000).
[19]A. N. Panche, A. D. Diwan, and S. R. Chandra, "Flavonoids: An overview," J. Nutr. Sci. **5**, e47 (2016).
[20]V. Krishnamachari, L. H. Levine, and P. W. Pare, "Flavonoid oxidation by the radical generator aibn: a unified mechanism for quercetin radical scavenging," J. Agric. Food Chem. **50**, 4357–4363 (2002).
[21]D. Procházková, I. Boušová, and N. Wilhelmová, "Antioxidant and prooxidant properties of flavonoids," Fitoterapia **82**, 513–523 (2011).
[22]G. Cao, E. Sofic, and R. L. Prior, "Antioxidant and prooxidant behavior of flavonoids: Structure-activity relationships," Free Radical Biol. Med. **22**, 749–760 (1997).
[23]S. Eghbaliferiz and M. Iranshahi, "Prooxidant Activity of Polyphenols, Flavonoids, Anthocyanins and Carotenoids: Updated Review of Mechanisms and Catalyzing Metals," Phytother. Res. **30**, 1379–1391 (2016).
[24]F. Di Meo, V. Lemaur, J. Cornil, R. Lazzaroni, J. L. Duroux, Y. Olivier, and P. Trouillas, "Free radical scavenging by natural polyphenols: Atom versus electron transfer," J. Phys. Chem. A **117**, 2082–2092 (2013).
[25]S. B. Nimse and D. Pal, "Free radicals, natural antioxidants, and their reaction mechanisms," RSC Adv. **5**, 27986–28006 (2015).

[26] A. Galano, D. X. Tan, and R. J. Reiter, "Melatonin as a natural ally against oxidative stress: a physicochemical examination," J. Pineal Res. **51**, 1–16 (2011).

[27] A. Galano and R. J. Reiter, "Melatonin and its metabolites vs oxidative stress: From individual actions to collective protection," J. Pineal Res. **65**, e12514 (2018).

[28] A. Galano and J. Raúl Alvarez-Idaboy, "Computational strategies for predicting free radical scavengers' protection against oxidative stress: Where are we and what might follow?" Int. J. Quantum Chem. **119**, e25665 (2019).

[29] A. Galano and J. R. Alvarez-Idaboy, "A computational methodology for accurate predictions of rate constants in solution: Application to the assessment of primary antioxidant activity," J. Comput. Chem. **34**, 2430–2445 (2013).

[30] A. Galano, "Mechanism and kinetics of the hydroxyl and hydroperoxyl radical scavenging activity of N-acetylcysteine amide," Theor. Chem. Acc. **130**, 51–60 (2011).

[31] A. Galano and A. Martnez, "Capsaicin, a tasty free radical scavenger: mechanism of action and kinetics," J. Phys. Chem. B **116**, 1200–1208 (2012).

[32] S. Curtarolo, G. L. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, "The high-throughput highway to computational materials design," Nat. Mater. **12**, 191–201 (2013).

[33] M. Rupp, R. Ramakrishnan, and O. A. Von Lilienfeld, "Machine Learning for Quantum Mechanical Properties of Atoms in Molecules," J. Phys. Chem. Lett. **6**, 3309–3313 (2015), 1505.00350.

[34] D. Xue, P. V. Balachandran, J. Hogden, J. Theiler, D. Xue, and T. Lookman, "Accelerated search for materials with targeted properties by adaptive design," Nat. Commun. **7**, 11241 (2016).

[35] F. M. Paruzzo, A. Hofstetter, F. Musil, S. De, M. Ceriotti, and L. Emsley, "Chemical shifts in molecular solids by machine learning," Nat. Commun. **9**, 4501 (2018), arXiv:1805.11541.

[36] M. Rupp, "Machine learning for quantum mechanics in a nutshell," Int. J. Quantum Chem. **115**, 1058–1073 (2015).

[37] L. Hu, X. Wang, L. Wong, and G. Chen, "Combined first-principles calculation and neural-network correction approach for heat of formation," J. Chem. Phys. **119**, 11501–11507 (2003).

[38] M. Karthikeyan, R. C. Glen, and A. Bender, "General melting point prediction based on a diverse compound data set and artificial neural networks," J. Chem. Inf. Model. **45**, 581–590 (2005).

[39] M. Misra, D. Andrienko, B. Baumeier, J.-L. Faulon, and O. A. von Lilienfeld, "Toward quantitative structure–property relationships for charge transfer rates of polycyclic aromatic hydrocarbons," J. Chem. Theory Comput. **7**, 2549–2555 (2011).

[40] A. Lusci, G. Pollastri, and P. Baldi, "Deep architectures and deep learning in chemoinformatics: the prediction of aqueous solubility for drug-like molecules," J. Chem. Inf. Model. **53**, 1563–1575 (2013).

[41] G. Pilania, C. Wang, X. Jiang, S. Rajasekaran, and R. Ramprasad, "Accelerating materials property predictions using machine learning," Sci. Rep. **3**, 2810 (2013).

[42] A. V. Marenich, C. J. Cramer, and D. G. Truhlar, "Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions," J. Phys. Chem. B **113**, 6378–6396 (2009).

[43] E. O. Pyzer-Knapp, K. Li, and A. Aspuru-Guzik, "Learning from the harvard clean energy project: The use of neural networks to accelerate materials discovery," Adv. Funct. Mater. **25**, 6495–6502 (2015).

[44] X. Zheng, L. Hu, X. Wang, and G. Chen, "A generalized exchange-correlation functional: the neural-networks approach," Chem. Phys. Lett. **390**, 186–192 (2004).

[45] R. M. Balabin and E. I. Lomakina, "Support vector machine regression (ls-svm)—an alternative to artificial neural networks (anns) for the analysis of quantum chemistry data?" Phys. Chem. Chem. Phys. **13**, 11710–11718 (2011).

[46] T. F. G. G. Cova and A. A. C. C. Pais, "Deep Learning for Deep Chemistry: Optimizing the Prediction of Chemical Patterns," Front. Chem. **7**, 809 (2019).

[47] R. Winter, F. Montanari, F. Noé, and D. A. Clevert, "Learning continuous and data-driven molecular descriptors by translating equivalent chemical representations," Chem. Sci. **10**, 1692–1701 (2019).

[48] G. A. Behr, J. C. Moreira, and B. N. Frey, "Preclinical and clinical evidence of antioxidant effects of antidepressant agents: Implications for the pathophysiology of major depressive disorder," (2012).

[49] H. Erman, I. Guner, M. O. Yaman, D. D. Uzun, R. Gelisgen, U. Aksu, N. Yelmen, G. Sahin, and H. Uzun, "The effects of fluoxetine on circulating oxidative damage parameters in rats exposed to aortic ischemia-reperfusion," Eur. J. Pharmacol. **749**, 56–61 (2015).

[50] M. Herbet, M. Gawrońska-Grzywacz, M. Izdebska, and I. Piątkowska-Chmiel, "Effect of the interaction between atorvastatin and selective serotonin reuptake inhibitors on the blood redox equilibrium," Exp. Ther. Med. **12**, 3440–3444 (2016).

[51] V. Caiaffo, B. D. Oliveira, F. B. de Sá, and J. Evêncio Neto, "Anti-inflammatory, antiapoptotic, and antioxidant activity of fluoxetine," Pharmacol. Res. Perspect. **4**, 1–9 (2016).

[52] M. Bortoli, M. Dalla Tiezza, C. Muraro, C. Pavan, G. Ribaudo, A. Rodighiero, C. Tubaro, G. Zagotto, and L. Orian, "Psychiatric disorders and oxidative injury: Antioxidant effects of zolpidem therapy disclosed in silico," Comput. Struct. Biotechnol. J. **17**, 311–318 (2019).

[53] C. Muraro, M. Dalla Tiezza, C. Pavan, G. Ribaudo, G. Zagotto, and L. Orian, "Major Depressive Disorder and Oxidative Stress: In Silico Investigation of Fluoxetine Activity against ROS," Appl. Sci. **9**, 3631 (2019).

[54] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, (Jr.), J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, "Gaussian 16 revision b.01," (2016), gaussian Inc. Wallingford CT.

[55] Y. Zhao and D. G. Truhlar, "A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions," J. Chem. Phys. **125**, 194101 (2006).

[56] Y. Zhao and D. G. Truhlar, "The m06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four m06-class functionals and 12 other functionals," Theor. Chem. Acc. **120**, 215–241 (2008).

[57] A. Galano, "On the direct scavenging activity of melatonin towards hydroxyl and a series of peroxyl radicals," Phys. Chem. Chem. Phys. **13**, 7178–7188 (2011).

[58] B. Scholkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond* (MIT Press, Cambridge, MA, USA, 2001).

[59] O. Dangles and J.-A. Fenger, "The chemical reactivity of anthocyanins and its consequences in food science and nutrition," Molecules **23** (2018), 10.3390/molecules23081970.

[60] Y. Noda, T. Kaneyuki, A. Mori, and L. Packer, "Antioxidant activities of pomegranate fruit extract and its anthocyanidins: delphinidin, cyanidin, and pelargonidin," J. Agric. Food Chem. **50**, 166–171 (2002).

[61] J.-M. Kim, J.-S. Kim, H. Yoo, M.-G. Choung, and M.-K. Sung, "Effects of black soybean [glycine max (l.) merr.] seed coats and its anthocyanidins on colonic inflammation and cell proliferation in vitro and in vivo," J. Agric. Food Chem. **56**, 8427–8433 (2008).

[62] S. Meiers, M. Kemény, U. Weyand, R. Gastpar, E. von Angerer, and D. Marko, "The anthocyanidins cyanidin and delphinidin are potent inhibitors of the epidermal growth-factor receptor," J. Agric. Food Chem. **49**, 958–962 (2001).