



**HAL**  
open science

# Rapid and Complete Prediction of Alizarin in Solution by Combining Experimental Data with Computational Methods

Thanh Huyen Linh Tran, Baptiste Rigaud, Maguy Jaber, Romain Berraud-Pache

## ► To cite this version:

Thanh Huyen Linh Tran, Baptiste Rigaud, Maguy Jaber, Romain Berraud-Pache. Rapid and Complete Prediction of Alizarin in Solution by Combining Experimental Data with Computational Methods. *Dyes and Pigments*, In press, 10.1016/j.dyepig.2024.112242 . hal-04586447

**HAL Id: hal-04586447**

**<https://hal.science/hal-04586447v1>**

Submitted on 29 May 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# 1 Rapid and Complete Prediction of Alizarin in Solution by Combining 2 Experimental Data with Computational Methods

3 Thanh Huyen Linh TRAN<sup>a</sup>, Baptiste RIGAUD<sup>b</sup>, Maguy JABER<sup>a,c</sup>, Romain BERRAUD-PACHE<sup>a\*</sup>

4 <sup>a</sup>Laboratoire d'Archéologie Moléculaire et Structurale (LAMS), Sorbonne Université, CNRS, UMR 8220, Paris, 75005, France

5 <sup>b</sup>Fédération de Chimie des Matériaux de Paris-Centre FCMat

6 <sup>c</sup>IUF-Institut Universitaire de France, Paris, France

7 \*corresponding author: romain.berraud-pache@sorbonne-universite.fr

8

## 9 ARTICLE INFO

10 Keywords: Computational chemistry, organic dyes, cultural heritage, UV-Vis spectrum, NMR, color prediction, solvation model

## 11 ABSTRACT

12 In the field of cultural heritage, understanding the nature of pigments and dyes is of great importance for  
13 conservation issues. Among the many natural dyes used in paintings or textiles, madder was one of the most  
14 important sources of red color at the time and was present in a wide variety of objects. However, due to  
15 extraction challenges, impurities, and high costs, a lot of information are not easily available. In this  
16 publication we designed a computational protocol able to reproduce the spectroscopic properties of two  
17 dyes, alizarin and alizarin red S. This framework allows us to compute both Ultraviolet-Visible absorption  
18 spectra and Nuclear Magnetic Resonance (NMR) with good accuracy as well as reproduce with precision the  
19 CIELAB color. We have also explored different types of interactions that impact these properties, notably the  
20 solvation effect. We found that microsolvation is sufficient to reproduce the experimental measurements  
21 made in water. The high accuracy of the computation method makes this technique particularly promising  
22 for a non-destructive study of dyes on works of art and the preservation of cultural heritage.

23

## 24 INTRODUCTION

25 The analytical investigation of dyes is considered of utmost importance in the field of cultural heritage.  
26 Indeed, through the identification of the corresponding dye, we are often able to gain insight into several  
27 important aspects: the characterization of the dye and the artist's techniques, the degradation pathways,  
28 the preservation state of the artifact, the dating and the determination of the provenance of objects [1–4].  
29 Furthermore, it provides precious details to plan correct restoration, conservation, and display strategy in  
30 museums and art galleries.

31 Both traditional and cutting-edge techniques such as ultraviolet-visible (UV-Vis), infrared (IR), and Raman  
32 spectroscopies, as well as nuclear magnetic resonance (NMR), mass spectrometry (MS), liquid  
33 chromatography, Fiber Optics Reflectance Spectra (FORS), X-Ray diffraction (XRD) or hyperspectral  
34 spectroscopy have been reported for the analytical investigation of dyes [2,5–14]. While the accuracy and  
35 sensitivity of these techniques have allowed researchers to gain a huge amount of insight into these  
36 systems, most of the technologies are sampling-based and need complex dye extraction and chemical  
37 preparation methods. Moreover, sample collection, preparation, and chemical analysis are often challenging  
38 due to material complexity, limited sample sizes, dye degradation, and lack of appropriate standards needed  
39 for UV-Vis detection [8,15–19]. Therefore, the identification of historical dyes is still one of the most  
40 challenging tasks in the realm of heritage preservation. In this context, computational methods have  
41 emerged as an appealing tool to decipher complex experimental data [20]. These theoretical methods aid in  
42 the interpretation of individual contributions to the experimental spectrum. Unfortunately, their application

43 in the field of cultural heritage is not as widespread as other fields of research (such as catalysis,  
44 astrochemistry, electrochemistry, etc).

45 Representing one of the most famous natural dyes since antiquity, madder was selected as the focus of this  
46 study [21–23]. From the roots of this plant, around 36 red colorants derivating from the anthraquinone  
47 family can be extracted in various quantities. We will focus here on alizarin, one of the main constituents of  
48 madder [24,25].

49 In order to study the associated dyes, the combination of NMR with UV-Vis is often used in the heritage  
50 science field because of their complementary nature. While UV-Vis is good at characterizing the  
51 environmental effect of the dye or its degradation (color fading), NMR can differentiate analog molecules. It  
52 is also very precise in attesting chemical modifications, notably when lake or pigment structures are created  
53 by association of a colored dye with a metallic ion.

54 For the computation of spectroscopic properties, density functional theory (DFT) and Time Dependent-DFT  
55 are often the methods of choice due to their reasonable costs and good precision. In fact, there have been  
56 many reports regarding the prediction of Alizarin color and properties using these methods [26–37]. Still, the  
57 role of the environment is not clearly understood or deciphered. Furthermore, some key properties like  
58 reproducing the UV-Vis or NMR spectra are not always present.

59 When examining molecules using spectroscopic analysis, a significant portion of their observable  
60 characteristics displays varying degrees of sensitivity to the surrounding environment (solvent, pH, binding  
61 partners, mordant, substrate, temperature, light, etc) [38,39]. This influence of the solvent on spectroscopic  
62 properties arises from multiple factors such as the polarity of the solvent or the presence of hydrogen  
63 bonds. It is therefore crucial to find the optimal solvation model in order to correctly predict the UV-Vis  
64 spectrum or other properties of these organic molecules.

65 In this study, our attention will be directed toward the impact of solvents on a precious dye, alizarin. The aim  
66 is to introduce a holistic computational protocol designed to predict color and chemical properties,  
67 developed through a collaborative effort with experimental data. The protocol integrates multiple facets,  
68 including the exploration of conformers and deprotonated forms, effective selection of optimal solvation  
69 models based on the solute-solvent system of study (implicit, microsolvated, or multilevel), prediction of UV-  
70 Vis spectrum, color prediction with CIELAB coordinates and <sup>13</sup>C and <sup>1</sup>H NMR shift predictions. This protocol  
71 will be applied to two dyes: alizarin and alizarin red S.

72

## 73 **1. MATERIALS & METHODS**

74

### 75 **2.1. EXPERIMENTAL METHOD**

#### 76 *UV Spectrum*

77 Alizarin (97% purity) was purchased from Sigma Aldrich and used without further purification. Alizarin red S  
78 pure (sodium alizarinsulfonate) was purchased from Thermoficher.

79 10 mg of alizarin were dissolved in 40 mL of deionized water (MilliQ). The alizarin solution was then heated  
80 for 2 hours at 90°C, and filtered using a Buchner funnel. 5mg of alizarin red S was added to 80 mL of distilled  
81 water and stirred for 10min. The obtained solution was used directly without heating or filtering. Then, the  
82 pH was adjusted using either hydrochloric acid (HCl, 0.1M) or sodium hydroxide (NaOH, 1M). Three pH  
83 values were considered for alizarin (pKa<sub>1</sub>=5.5 and pKa<sub>2</sub>=10) : 3, 8, and 12. For alizarin red S (pKa<sub>1</sub>=5.8 and  
84 pKa<sub>2</sub>=10.8), the following pH were used: 2, 7 and 12 [40,41].

85 The UV-Vis spectra were measured directly after each preparation. It was done using an Ocean Optics Flame  
86 spectrometer (Model: FLMS00699) with a wavelength range from 300 to 850 nm. The integration time for  
87 each measurement ranged from 5 milliseconds to 10 milliseconds. The light source used was a Deuterium-  
88 Tungsten Halogen lamp covering the range of 300 nm to 850 nm. The blank was done using deionized water.  
89 This light source was coupled to the spectrometer using 400 $\mu$ m diameter optical fibers and positioned  
90 through a 1 cm quartz cuvette holder. The acquisition of spectra was facilitated by the Ocean View software.

### 91 *NMR Spectrum*

92 NMR spectra were acquired with a Bruker Advance Neo 500 spectrometer at the frequency of 500.07 MHz  
93 for  $^1\text{H}$  and 125.76 MHz for  $^{13}\text{C}$  equipped with a 5 mm double resonance broadband prodigy probe use  
94 nitrogen-cooled RF coils and preamplifiers. Chemical shifts were calibrated using the Tetramethylsilane (0  
95 ppm). For  $^1\text{H}$ , we used a one  $90^\circ$  pulse with 64 scans and with a recycling time of 1s. For  $^{13}\text{C}$ , we used a one  
96  $90^\circ$  pulse with a  $^1\text{H}$  decoupling waltz16 with 4096 scans and with a recycling time D1 of 2s. The  $^1\text{H}$  and  $^{13}\text{C}$   
97 spectra were recorded in exactly the same conditions at room temperature and processed with a zero-filling  
98 factor of 2 and without an exponential decay. 5mg of either alizarin or alizarin red S were dissolved in 0.5mL  
99 of acetone- $\text{D}_6$ . The  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra were also measured for alizarin red S in  $\text{D}_2\text{O}$  as it was more  
100 soluble than alizarin.

101

## 102 **2.2. COMPUTATIONNAL METHOD**

103 All calculations were carried out using the version 5.0.3 of the ORCA quantum chemistry program package  
104 and the version 6.5.1 of xTB [42,43].

### 105 *Optimization Procedure*

106 Ground and excited states properties were computed using density functional theory (DFT) and its time-  
107 dependent (TD-DFT) extension [44,45]. Optimized Geometries and Hessians were obtained using the PBE0  
108 functional [46] complemented by the D3(BJ) dispersion correction [47,48]. Tight thresholds were employed  
109 for the self-consistent field (SCF) and optimization (tightSCF and tightopt keywords in ORCA nomenclature)  
110 [47,49]. The triple zeta def2-TZVP basis set, coupled with matching auxiliary basis sets, was consistently  
111 employed throughout all the computations [47,49,50]. All the optimized geometries can be found in the  
112 supplementary data 14.

113

### 114 *Conformer and deprotonated Forms Search*

115 Using Grimme's open-source Conformer-Rotamer Ensemble Sampling Tool (CREST), which relies on the xTB  
116 method, the exploration for deprotonated forms and conformer for the fully protonated form was  
117 conducted [43,51]. The geometries of the generated conformer or deprotonated forms were then subjected  
118 to optimization using the PBE0 functional. Finally, their single point energy (SPE) were obtained by using  
119 DLPNO-CCSD(T) computations with def2-TZVP basis sets with matching auxiliary basis sets and CPCM (water)  
120 [52]. Their SPE energies are then compared for each deprotonated or conformer forms to identify the most  
121 energetically stable forms (the one with the lowest SPE energy). By comparing the SPE energies, we can  
122 calculate the Boltzmann population distribution of each form. The identified stable forms correspond to the  
123 highest Boltzmann distribution. Those forms are dominant and are therefore selected for the following  
124 computations.

### 125 *Solvation Models*

126 Different solvation models were evaluated for UV-Vis and NMR spectra prediction. These models included  
127 the following:

128 -In vacuo: the molecule is studied in the absence of any solvent.

129 -Implicit solvation (water or acetone) using the linear response Conductor-like Polarizable Continuum Model  
130 (C-PCM) [53]. No solvents appear but it treats the solvent as a continuous medium surrounding the solute.  
131 This model accounts for the electrostatic and non-polar interactions between the solute and the solvent  
132 without explicitly representing individual solvent molecules. Their bulk properties are replaced by some  
133 macroscopic quantities (dielectric constant or solvent radius).

134 -A combined approach involving both implicit and explicit water molecules (either 1 water molecule per  
135 hydroxyl (OH) group or 3 water molecules per OH group). The water molecules were positioned to make  
136 hydrogen bonds with each hydroxyl group.

137 -Furthermore, a hybrid approach called Quantum Mechanics/Quantum Mechanics 2 (QM/QM2) [54–56]  
138 calculations were conducted by engaging between 70-100 water molecules. It combines quantum mechanics  
139 (QM) for a region of interest (here the dye molecule) and the semi-empirical GFN2-xTB method [43,51] for  
140 the rest of the system (here the surrounding water molecules). The optimization was accomplished using the  
141 multiscale module in ORCA [42,57] together with the QM/QM2 ONIOM procedure. This allows detailed  
142 treatment for the electronic structure of the solute while considering the surrounding solvent at a less  
143 computationally expensive level. To consider the solvent effect, we also performed a Quantum Molecular  
144 Dynamics simulation of 25,000 steps with a 0.5 fs time step (other parameters selected as default such as a  
145 NHC (Nosé–Hoover chains) thermostat of 300K) [58]. An ellipsoid repulsive wall was added to keep the  
146 water molecules close to the substrate during the MS. From this simulation, 50 snapshots were selected,  
147 each 200 steps starting at 800. They were then optimized and the mean Vertical Energy (VE) of these  
148 snapshots calculated.

#### 149 *UV-Vis Spectrum computation*

150 For Time-Dependent Density Functional Theory (TD-DFT) calculations concerning the UV-Vis spectrum, the  
151 B2PLYP functional was used with the resolution of identity method (RI) [30,59,60]. This choice is justified by  
152 previous theoretical benchmarks and preliminary calculations (see supplementary data 2) [61,62]. The UV-  
153 Vis VE calculations were done with the first 15 roots. The experimental spectra of the dyes studied (alizarin  
154 and alizarin red S) has only one band on the visible region. In all calculations, only the first excited state was  
155 considered for the VE calculation, as it can be compared to the experimental maximum absorption of the  
156 Visible region of the UV-Vis spectrum. The topology and form of the experimental spectra are due to  
157 vibrational and solvent effects which are not present when calculating the VE but will be accounted using  
158 another calculation. The color of the compound depends on the form of the spectra, it is therefore necessary  
159 to take it into account.

160 The vibrational contributions, including Herzberg–Teller effects [63–65], was accomplished using the Excited  
161 State Dynamics (ESD) module within ORCA [66]. Within this framework, the Adiabatic Hessian (AH) and  
162 Vertical Gradient (VG) approaches were used. As AH is a more complete and accurate method, computed  
163 spectra within the main text will be performed using this scheme [67]. VG spectra can be found in  
164 supplementary data 5. The AH calculations require the geometry of both Ground State (GS) and the first  
165 singlet excited state ( $S_1$ ) geometries and frequencies, the VG ones only required the GS parameters, all of  
166 them computed at the DFT PBE0 level + CPCM. The VE energy and transition dipole moments were  
167 computed using the most accurate values *i.e.* RI-B2PLYP together with the appropriate solvation scheme (3  
168 water molecules per OH). The linewidth was created using a Gaussian distribution and was fine-tuned to  
169 align with the Full Width at Half Maximum (FWHM) parameter of the experimental spectra, which depends  
170 on the concentration and other experimental parameters. To reduce the impact of experimental parameters  
171 like concentration, the computed and experimental spectra were normalized. This allows a direct  
172 comparison between the two spectra.

173 The CIELAB color coordinates was calculated for the UV-Vis spectra with an in-house script with illuminant  
174 D65 [68–70]. To quantify the quality of the color prediction, the color difference was quantified with  $\Delta E_{2000}$   
175 which is the latest ISO standard by CIE [71,72].

### 176 *NMR Spectrum computation*

177 Geometry optimization of tetramethylsilane (TMS), alizarin, alizarin red S molecules was conducted within  
178 vacuo, an implicit acetone or water solvent environment using the CPCM method or with explicit acetone or  
179 water molecules as described before. Subsequent computation of  $^1\text{H}$ -NMR and  $^{13}\text{C}$ -NMR shielding employed  
180 the revTPSS functional with D4 correction [73–75]. Basis set pcSseg-3 was selected for these calculations.  
181 Next, the isotropic shifts were used to compute the shifts that are comparable to the experimental by using  
182 this formula:

$$\sigma_{\text{calculated}} = \sigma_{\text{TMS}} - \sigma_{\text{targetmolecule}}$$

183 A summary of the computational protocol is summarized in a flow chart in the supplementary data 1.

184

## 185 2. RESULTS AND DISCUSSION

186

187 Fig.1 presents the structure of the studied molecules: alizarin (later referred as  $H_2A$ ) and alizarin red S (later  
188 referred as  $H_2ARS$ ). Both alizarin and alizarin red S have two acidic protons which can be subject to  
189 deprotonation depending on the solutions' pH. Alizarin (alizarin Red S) has  $\text{pK}_{a1}=5.5$  and  $\text{pK}_{a2}=10$  ( $\text{pK}_{a1}=5.8$   
190 and  $\text{pK}_{a2}=10.8$ ) respectively. Therefore, in this study, we have chosen the following pH for each form: pH=3,  
191 pH=8, and pH=12 for alizarin and pH=2, pH=7 and pH=12 for alizarin red S [40,41].

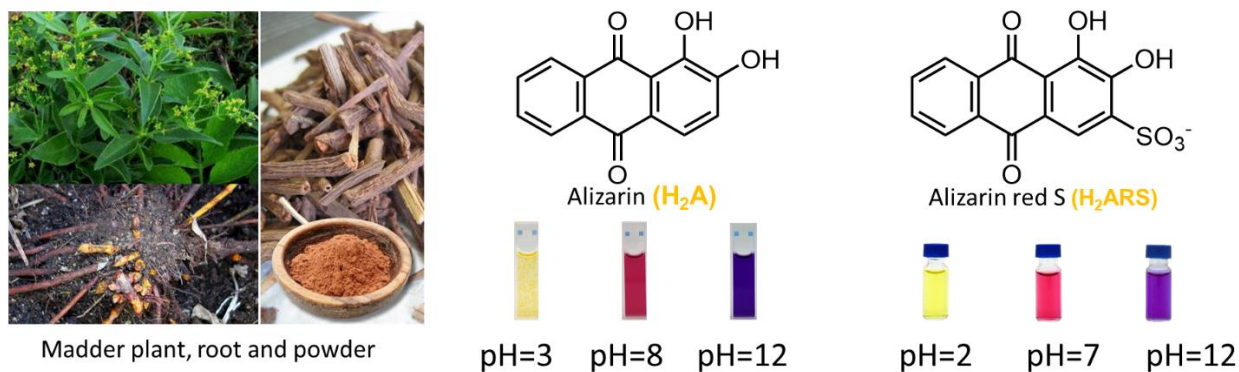


Fig. 1. (Left) Madder plant, its roots and powder. (Middle) Chemical structure of alizarin, abbreviated as  $H_2A$  and its alternative more soluble in water, alizarin red S (right), abbreviated as  $H_2ARS$ . The color of each molecule for different pH level are depicted below each structure.

192 This study aims to establish a reliable computational protocol in order to reproduce their spectroscopic  
193 properties (UV-Vis spectra, color and NMR spectra) and provide us with relevant chemical information.

194

195

### 196 3.1. ALIZARIN

197

#### 198 3.1.1. UV-Vis spectra

199

##### 200 3.1.1.1. Conformer and deprotonated forms search

201 The initial step involves exploring the conformer and deprotonated forms of alizarin for each pH range in  
202 order to identify the most stable forms.

203 In the protonated form (neutral) at pH=3, various studies have addressed the presence of different  
204 conformers [36,76]. We first generated the conformer of the neutral protonated form of alizarin and  
205 computed their Single Point Energy (SPE). The one with the lowest SPE energy after optimization is the most  
206 stable and is taken as a reference. The other conformer forms are listed by energy (in kcal/mol) with respect  
207 to the most stable form in Fig 2. We have found that the dominant form is the one with the hydrogen of  
208 both the  $\alpha$  and  $\beta$ -hydroxyl groups making hydrogen bonds with the neighboring oxygen. It will be later  
209 referred to as  $H_2A$ . Furthermore, for the ground state, the forms with the proton transfer on the ketone as  
210 reported in previous publications are much higher in energy (14.07 kcal/mol respective to  $H_2A$ ), meaning  
211 that in the ground state, there is likely no proton transfer for the neutral form in water. The remaining  
212 conformer forms do not exist in solution as they are higher in energy than the most stable one.

213 For the mono-deprotonated form (monoanionic) at pH=8, two potential forms are possible due to the  
214 presence of two deprotonation sites (see Fig. 2). However, by comparing the computed SPE, the most stable  
215 configuration is the one with the  $\beta$ -hydroxyl being deprotonated. The other form with the  $\alpha$  hydroxyl  
216 deprotonated is higher in energy by 0.9 kcal/mol. This corresponds to the Boltzmann distribution of 82% in  
217 favor of the most stable form. This can be explained by the formation of intramolecular hydrogen bonds  
218 between the ketone and the  $\alpha$  hydroxyl group, making it more stable and therefore less favorable to  
219 deprotonation than the  $\beta$ -hydroxyl group. Furthermore, the hydrogen in the  $\beta$  position can be more  
220 accessible to the surrounding solvent, making it more susceptible to deprotonation. The most stable form  
221 will be used to do further calculations as it will be the most dominant. We will denote it as  $HA^-$ .

222 For the di-deprotonated (dianionic) form at pH=12, there is only one stable form without other conformers.  
223 It will be denoted as  $A^{2-}$ .



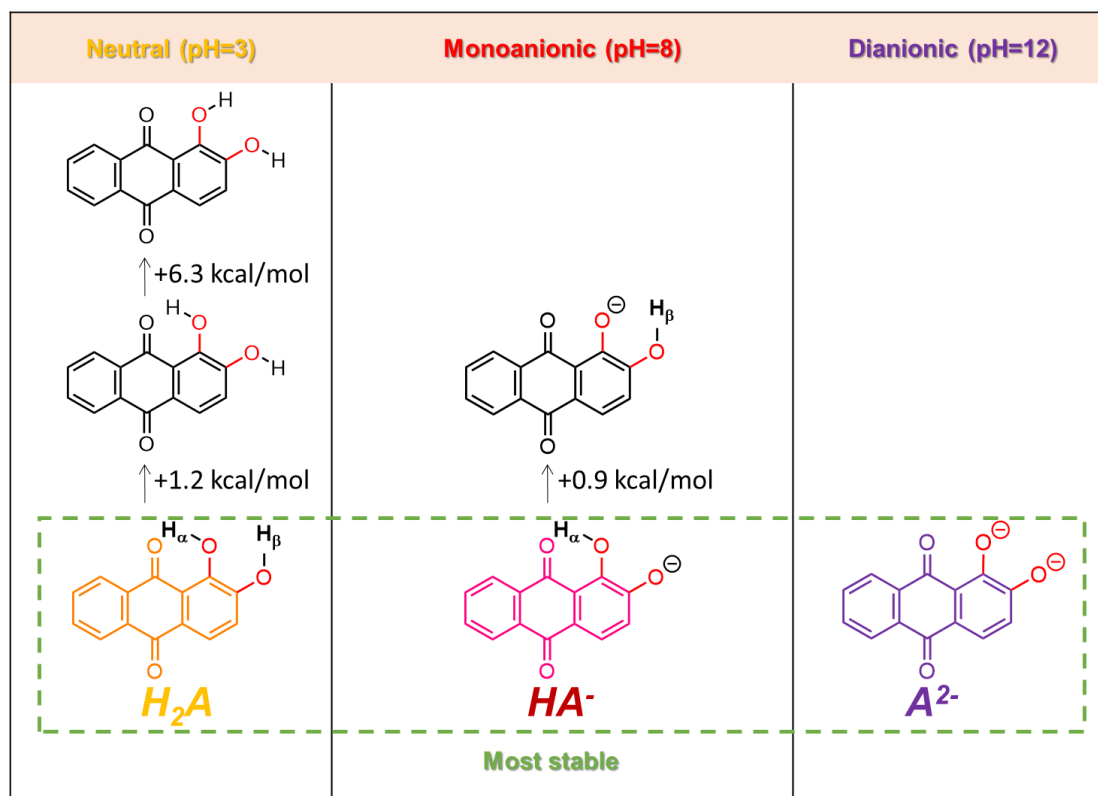


Fig. 2. Alizarin rotamers and deprotonated forms for each pH range and their relative energies. For the fully protonated form at pH=3. The geometries are sorted from the most stable form to the less stable form. The most stable forms are denoted  $H_2A$  (neutral form),  $HA^-$  (monoanionic form), and  $A^{2-}$  (dianionic form).

224

### 3.1.1.2. Vertical energy and choice of solvation model

225

In computational chemistry, the choice of solvation model is critical as it determines the accuracy of the predicted UV-Vis spectra [77]. In fact, the geometry of the molecules can be influenced by the surrounding solvent via specific interactions like hydrogen bonds, ion pairs,  $\pi$ -interactions, etc. Here, we have tested different solvation models to compute the VE between the electronic GS and  $S_1$  and compare it to the maximum absorption of the experimental visible spectra. The calculations were first done in vacuo, then with implicit water, a combined implicit and explicit, and finally with a multi-level approach such as QM/QM2 in combination with Molecular Dynamics. All computational details of those solvation models are outlined in the *Methods* section. Fig.3 presents a condensed view of the different models used in this study and the associated results for the alizarin forms  $H_2A$ ,  $HA^-$  and  $A^{2-}$  (Additional information can be found in supplementary data 3).

234

235



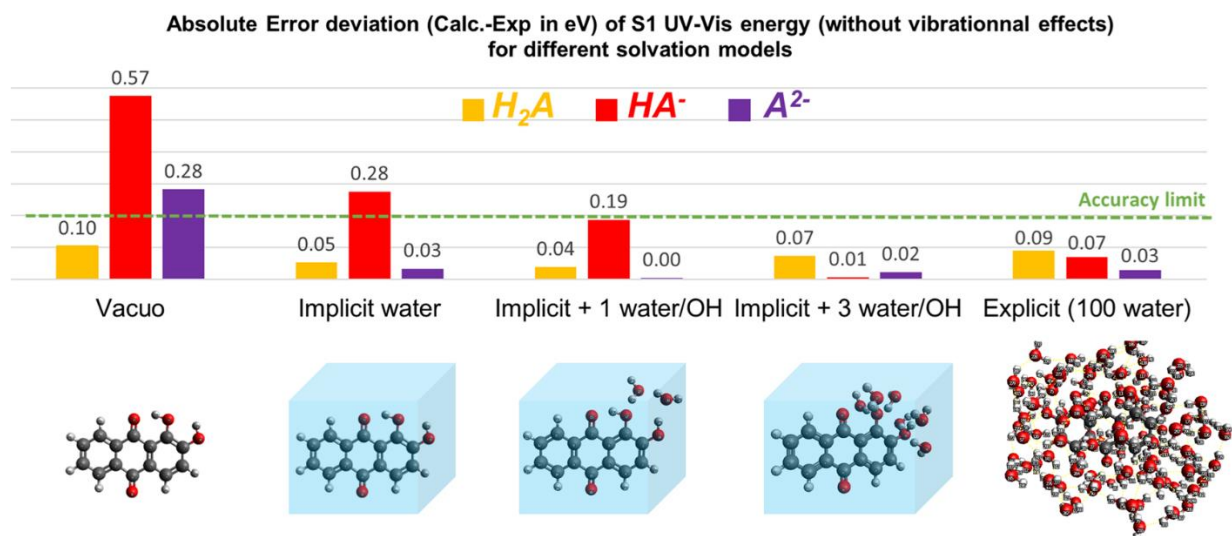


Fig. 3. (Top) Absolute energy difference between the calculated VE and the experimental absorption energy for the different solvation models of alizarin. The yellow stands for the  $H_2A$  form, the red for  $HA^-$  one and the purple for  $A^{2-}$ . The accuracy threshold is fixed at 0.2 eV according to the literature. (Bottom) Different solvation models were used. From left to right: alizarin in vacuo, implicit water, implicit with 1 water per hydroxy group, implicit with 3 water per hydroxy group, QM/QM2 with 100 explicit water (over 50 snapshots).

236 We first did the calculations in vacuo, without any solvent as a study case. However, it does not provide  
 237 insights into the effects of a surrounding medium, which is often crucial for realistic scenarios. Here, it was  
 238 observed that for the neutral form, the vacuo model is accurate enough (the error deviation falls below the  
 239 0.2eV accuracy limit which is the standard in the literature). However, for the deprotonated forms  
 240 (monoanionic  $HA^-$  or dianionic  $A^{2-}$ ), the vacuo model is not accurate enough (error deviation of 0.57eV for  
 241  $HA^-$  and 0.28eV for  $A^{2-}$ ). It is therefore necessary to include solvent effects for these forms.

242 A first approximation of the solvent effect can be done by using an implicit model. For alizarin, implicit  
 243 solvation gives better agreement with experimental data by reducing the error by at least half (it falls to  
 244 0.05eV for  $H_2A$ , 0.28eV for  $HA^-$  and 0.03eV for  $A^{2-}$ ). It can be used to test different solvent environments  
 245 while being computationally efficient. This approach is useful for the case where there is no apparent  
 246 interaction between the solvent and the dye, like the neutral form  $H_2A$  in water or for apolar solvents.  
 247 However, the  $HA^-$  form is still above the accuracy limit (with 0.28 eV error deviation) in implicit calculation.  
 248 Therefore, we need to account for specific solvent effects. For those forms, the model can be refined by  
 249 adding some explicit solvent molecules near the hydroxyl sites. Explicit water molecules (either 1 or 3) per  
 250 hydroxy group were tested. This considerably reduces the calculated and experimental energy difference for  
 251 the deprotonated forms. Among them, the best model is the one with 3 explicit water molecules near the  
 252 OH sites (error deviation of 0.01 eV for  $HA^-$  and 0.02 eV for  $A^{2-}$ ). This shows that for the deprotonated forms,  
 253 solvent interactions like hydrogen bonds are mandatory to accurately model the absorption energy.  
 254 Therefore, this combined model can be a good compromise to accurately predict the UV-Vis energies for a  
 255 low computational cost. It can be used when there is a specific solute-solvent interaction.

256  
 257 Finally, the last model tested was QM/QM2 with 100 explicit water which represents a theoretical best  
 258 estimate for our protocol. For each form, a quantum molecular dynamic was performed followed by  
 259 computing the VE of 50 snapshots. Here, it was observed that across all the forms, the QM/QM2 results  
 260 consistently fall below the 0.2 eV accuracy threshold (0.09 eV for  $H_2A$ , 0.07 eV for  $HA^-$  and 0.03 eV for  $A^{2-}$ ).  
 261 This implies that QM/QM2 is a reliable method for predicting the properties of new molecules for which the  
 262 solute-solvent interactions are predominant but requires a much higher computational cost, at least 50

263 times higher. More information about the MD and the selected snapshots can be found in supplementary  
 264 data 4.

265  
 266 From the previous paragraphs, it is clear that the accuracy of the VE is tightly bound to how the  
 267 environmental effects are modeled. As water is a highly polar solvent and involved in multiple inter-  
 268 molecular interactions, one should expect that explicit molecules have to be added. Interestingly, for some  
 269 forms, implicit solvation is accurate enough, notably for  $A^{2-}$ , where two charged oxygens are present.  
 270 However, only the microsolvation with 3 added water molecules per OH is sufficient to get close to the  
 271 experimental measurements for all forms. This model also gives similar results in comparison to a full explicit  
 272 model, allowing us to ensure that it is reliable and transferable to similar dyes.

273  
 274  
 275

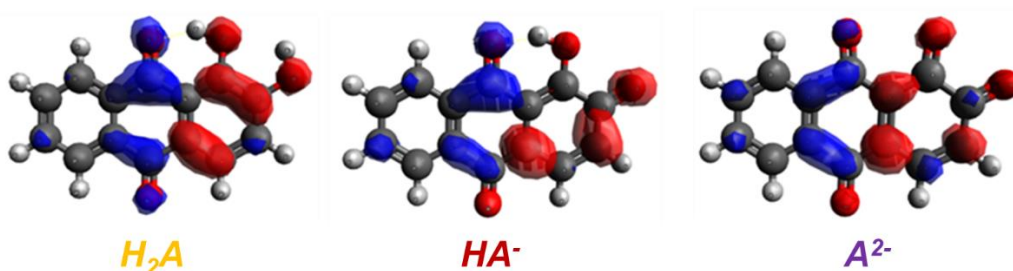
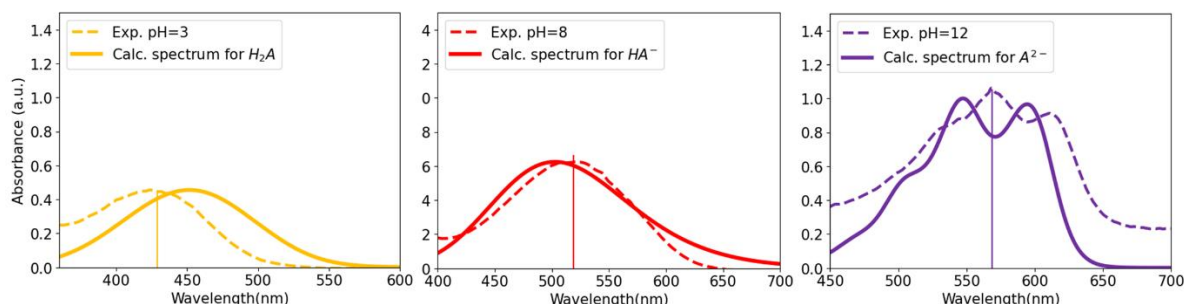


Fig. 4. Difference in electronic densities for the  $GS \rightarrow S_1$  transition for alizarin for different protonation states. The blue orbital represents a decrease in electronic density. The red one represents an increase in electronic density. The isovalue is 0.02 a.u.

276 To further delve into the nature of electronic transitions, the difference in electronic densities has been  
 277 calculated for the transition  $GS \rightarrow S_1$  (which corresponds to the absorption in the visible part of the UV-Vis  
 278 spectrum). The result for each alizarin form is depicted in Fig.4.

279 For the neutral form  $H_2A$ , the transition involves a combined  $n \rightarrow \pi^*$  (owing to the presence of oxygen) and  $\pi \rightarrow$   
 280  $\pi^*$  transition, with a more pronounced  $\pi \rightarrow \pi^*$  character. As we progress to the deprotonated form,  $HA^-$  and  
 281  $A^{2-}$ , the transition nature shifts more towards  $\pi \rightarrow \pi^*$  transition rather than  $n \rightarrow \pi^*$  as the deprotonated  
 282 hydroxy groups become part of the conjugated system.

### 283 3.1.1.3. Computed absorption spectra and CIELAB color



| Form     | Experimental (normalized) |       |       |       | Calculated (normalized) |      |       |       | $\Delta E_{2000}$ |
|----------|---------------------------|-------|-------|-------|-------------------------|------|-------|-------|-------------------|
|          | L*                        | a*    | b*    | color | L*                      | a*   | b*    | color |                   |
| $H_2A$   | 99.0                      | -13.4 | 68.4  |       | 90.7                    | 6.9  | 80.4  |       | 13.3              |
| $HA^-$   | 67.5                      | 53.3  | 0.9   |       | 59.1                    | 45.5 | 3.2   |       | 7.5               |
| $A^{2-}$ | 54.1                      | 51.9  | -50.0 |       | 57.0                    | 61.7 | -53.6 |       | 4.1               |

Fig. 5. (Top) The calculated and experimental UV-Vis spectra for alizarin at different pH. The dotted line is the experimental one. The

full lines are the calculated spectrum with the Adiabatic Hessian (AH) method. The vertical lines correspond to the computed VE. (Bottom) Predicted and computed colors of different alizarin forms with CIELAB coordinates ( $L^*$ ,  $a^*$ ,  $b^*$ ) and the calculated color variation between computed and experimental color coordinates with  $\Delta E_{2000}$ .

284 The VE alone is not enough to predict the full spectrum because it does not consider vibrational effects  
285 which can be responsible for the UV-Vis spectra topology (shape). In Fig.5. we have computed the  
286 absorption spectrum of alizarin at different pH and compared them with their experimental UV-Vis spectra.  
287 First, if we look at the different pH forms of alizarin we can see that their UV-Vis spectra shape and color  
288 differ. When the pH increases, we see a red shift of the spectra together with a hyperchromic shift (increase  
289 in absorption intensity). For pH=3, the spectrum has a maximum absorption peak at 430nm, resulting in a  
290 yellow solution. For pH=8, the solution turns red, with the maximum absorption at 530 nm. For pH=12, the  
291 solution takes a violet hue, with a maximum peak at 570 nm and two shoulders at 530 nm and 610 nm.  
292 These additional features could be attributed to several electronic transitions or a combination of electronic  
293 and vibrational effects. To discern the nature of each transition, we used the Adiabatic Hessian method (AH).  
294 The details of the computation can be found in the *Methods* section.

295 Overall, the computed spectra with AH are in good agreement with experimental data for both the  
296 maximum and topology, in particular for  $HA^-$  and  $A^{2-}$ . For  $H_2A$ , the computed spectrum is shifted by about (27  
297 nm) but it can be explained by the low solubility of the neutral form in water, preventing us from having a  
298 reliable UV-Vis spectrum.

299 The  $A^{2-}$  replicates the experimental topology with three distinct shoulders. The VE corresponds to one  
300 electronic transition from  $GS \rightarrow S_1$ . The middle peak is therefore attributed to one electronic transition, while  
301 the shoulders around 540 nm and 610 nm result from vibronic effects.

302

303 While the absorption spectra are really useful in spectroscopic studies, in the cultural heritage or dyeing  
304 field, the color of the molecule is the key property. In order to gain access to this information we can convert  
305 the UV-Vis spectra to color CIELAB coordinates. CIELAB color space is a color model introduced by the  
306 Commission Internationale de l'Eclairage (International Commission on Illumination or CIE) to provide a  
307 uniform and consistent representation of colors [69]. It is designed to be independent of devices and  
308 illuminants, making it useful for comparing colors across different conditions. Here, we will use it to compare  
309 the computed and experimental color. The colors are expressed in  $L^*a^*b^*$  coordinates. L stands for  
310 Lightness and corresponds to the perceived brightness of the color. The values range from 0 (black) to 100  
311 (white). The  $a^*$  stands for the position on the red(positive) to green (negative).  $b^*$  stands for blue (negative)  
312 to yellow (positive). The CIELAB color space is represented in supplementary data 6. The bottom part of Fig.5  
313 illustrates the predicted and experimental along with their respective CIELAB coordinates.  $\Delta E_{2000}$  is the latest  
314 standard created by the CIE and quantifies the color difference. E stands for "Empfindung" which is the  
315 German word for sensation and 2000 stands for the version of standard (older versions exist like  $\Delta E_{76}$  or  $\Delta E_{94}$ )  
316 [68,71]. The standard perception ranges as follows:

- 317
- 318 •  $\leq 1.0$ : Not perceptible by the human eye
  - 319 • 1-2: Perceptible through close observation
  - 320 • 2-10: Perceptible at a glance
  - 321 • 11-49: Colors are more similar than the opposite
  - 322 • 100: Colors are exactly the opposite

323 The computed color closely resembles the experimental ones. A higher  $\Delta E_{2000}=13.3$  is observed for the  $H_2A$   
324 form and  $HA^-$  ( $\Delta E_{2000}=7.57$ ), possibly attributed to the lower solubility of alizarin in water, where a non-  
325 negligible precipitate can be observed. In contrast, for  $A^{2-}$  which is more soluble in water, its computed color  
326 is similar to the experimental ( $\Delta E_{2000}=4.1$ ). That means that the color is similar but the difference can be

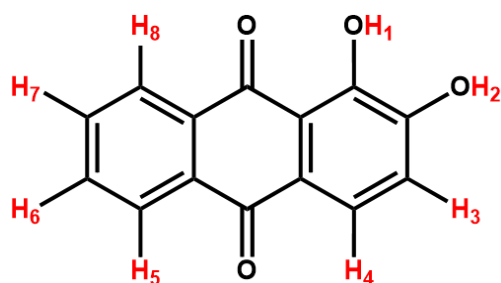
327 perceptible to the human eye at a glance. While the CIELAB colors are predominantly used for non-liquid  
328 systems, our approach shows that computations can be modeled with a reasonably good accuracy.

### 329 3.1.2. NMR shift prediction

330 Another key property of chemical systems is their NMR spectra. It is often used to decipher the structure of  
331 molecules in a complex environment, like pigments used in paintings, which are composed of a dye  
332 complexed with a metallic cation. Our protocol will aim at reproducing with good accuracy the experimental  
333 peaks.

334 Here, we present a study case with alizarin. The computed  $^1\text{H}$  shift values were compared to the  
335 experimental spectra of alizarin in acetone- $\text{D}_6$  (chosen because alizarin is not soluble in water) in Table 1. We  
336 tested the influence of the solvation model with implicit and 3 explicit acetones following the UV spectrum  
337 protocol. The calculated  $^1\text{H}$  NMR shift differs by less than 0.16 ppm on average with respect to the  
338 experimental one for aromatic hydrogens in implicit acetone and 0.22 ppm on average for explicit and  
339 implicit acetone. This suggests a good level of accuracy for the prediction of proton chemical shifts for both  
340 solvation models. The prediction of the  $^1\text{H}$  shifts for hydrogens H3, H4, H6, and H7 are better with implicit  
341 than explicit. However, it is noted that the accuracy is worse when dealing with the labile hydrogens in the  
342 hydroxy (-OH) group (H1 and H2) with implicit solvation. Labile hydrogen is prone to exchange with the  
343 surrounding solvents, it is therefore necessary to add explicit acetones near the hydroxy groups. For labile  
344 hydrogens, the explicit solvation model yields better accuracy. Furthermore, H5-H8 are better with explicit  
345 than implicit models. Explicit solvation is therefore more suited for labile hydrogens. The experimental  
346 spectra are depicted in supplementary data 7, along with its attribution, which was done by confronting the  
347 computed shifts.

Table 1. Computed and experimental  $^1\text{H}$  chemical shift of alizarin in acetone- $\text{D}_6$  with the respective error deviation in ppm or % for implicit or implicit and explicit solvation model.



| Attribution | Exp. (ppm)               | Calc. (ppm) |          | Calc.-error (ppm) |          | Calc.-error (%) |          |
|-------------|--------------------------|-------------|----------|-------------------|----------|-----------------|----------|
|             | In acetone- $\text{D}_6$ | Implicit    | Explicit | Implicit          | Explicit | Implicit        | Explicit |
| 1           | 12.8                     | 15.1        | 11.5     | 2.3               | -1.3     | 18.2            | 10.2     |
| 2           | 9.3                      | 7.2         | 10.0     | -2.1              | 0.6      | 22.5            | 7.0      |
| 3           | 7.3                      | 7.4         | 7.6      | 0.1               | 0.2      | 1.3             | 3.4      |
| 4           | 7.8                      | 8.0         | 8.2      | 0.2               | 0.4      | 3.0             | 5.4      |
| 5           | 8.3                      | 8.5         | 8.5      | 0.2               | 0.1      | 2.2             | 1.6      |
| 6           | 8.0                      | 8.1         | 8.1      | 0.1               | 0.1      | 1.6             | 1.9      |
| 7           | 8.0                      | 8.0         | 8.2      | 0.1               | 0.2      | 1.0             | 2.7      |
| 8           | 8.3                      | 8.5         | 8.4      | 0.3               | 0.1      | 3.0             | 1.5      |

348

### 349 3.1.3. Experimental uncertainty due to solubility

350 As raised multiple times in this part, alizarin has a poor solubility in water. The color of alizarin in solution is  
351 highly sensitive to experimental conditions such as concentration, pH and temperature. The maximum  
352 absorption remains constant but the shape can potentially alter. Consequently, the measured color may vary  
353 and a higher  $\Delta E$  can be calculated with respect to the computed color. To address this experimental  
354 obstacle, we opted for its alternative, alizarin red S which exhibits greater solubility.

355

## 356 **3.2. ALIZARIN RED S**

357

358 As we have seen, the limited solubility of alizarin in water can pose challenges when measuring their  
359 experimental UV-Vis and NMR spectra due to the potential formation of precipitates. To address this  
360 concern, a more soluble analogue, alizarin red S, has been studied.

361

### 362 *3.2.1. Conformer and deprotonated forms search*

363 For the acidic form (pH=2) of alizarin red S, different conformer forms can exist. We computed the energies  
364 of the possible conformer and sorted them in Fig. 6. Like alizarin, we do not observe any proton transfer for  
365 the ground state. The most stable form shows a hydrogen on the  $\alpha$  position pointing towards the ketone,  
366 making a hydrogen bond, while the  $\beta$ -hydrogen is pointing on the opposite side towards the negatively  
367 charged sulfonic group. It will be noted as  $H_2ARS$ . The other conformer forms do not exist as they are much  
368 higher in energy (hence low Boltzmann distribution).

For the mono-deprotonated form (pH=7), two possible forms can exist (see Fig.6). The comparison between energies point out that the  $\beta$ -hydroxyl deprotonated is the most stable configuration, the  $\alpha$ -hydroxyl making a hydrogen bond with the ketone and is less susceptible to deprotonation. This form will be noted as  $HARS^-$ . For the di-deprotonated form (pH=12), there is only one form that will be denoted as  $ARS^{2-}$ .

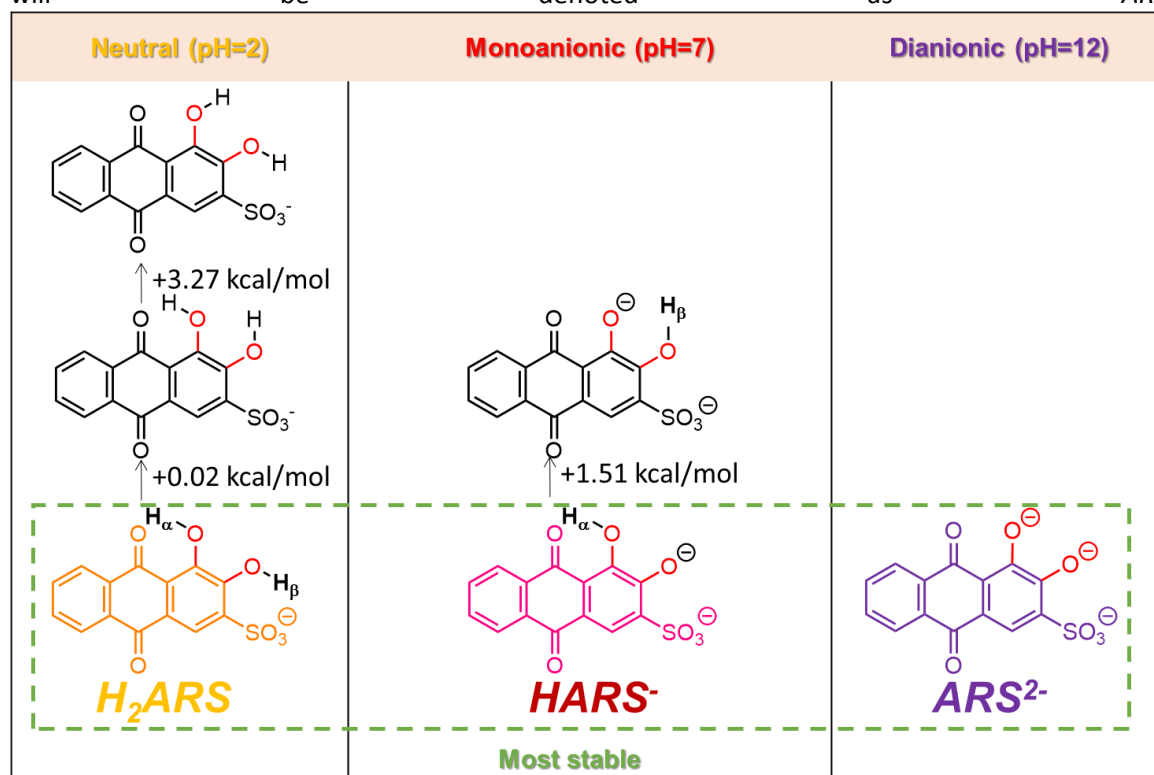


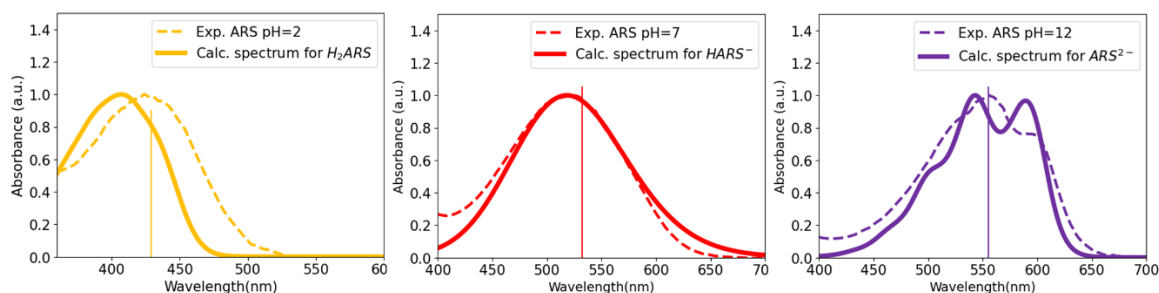
Fig. 6. Alizarin red S rotamer and deprotonated for each pH range, along with their relative energies. For the acidic form, the other forms are much higher in energy ( $>17$  kcal/mol) and are not considered. The most stable forms are denoted  $H_2ARS$  (acidic form),  $HARS^-$  (mono-deprotonated form), and  $ARS^{2-}$  (di-deprotonated form).

369

### 3.2.2. UV-Vis spectra prediction

370 The experimental spectra have maximum absorption energy at 425nm for  $H_2ARS$ , 521nm for  $HARS^-$  and 560  
 371 nm for  $ARS^{2-}$  with two shoulders at 600nm and 530nm. The computed spectra were done considering the  
 372 method and protocol used for alizarin. The results are depicted in Fig.7. There is a better agreement  
 373 between the experimental and computed spectra for the  $H_2ARS$  and  $HARS^-$  form than alizarin, which can be  
 374 explained by a higher solubility of alizarin red S. The computed VE matches perfectly the experimental one.  
 375 Similarly, the AH spectra of the  $HARS^-$  and  $ARS^{2-}$  are in adequation with the experimental ones, in both the  
 376 topology and position of the spectra.









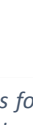
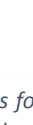
| Form       | Experimental (normalized) |       |       |   | Calculated (normalized) |       |       |   | $\Delta E_{2000}$ |
|------------|---------------------------|-------|-------|---|-------------------------|-------|-------|---|-------------------|
|            | L*                        | a*    | b*    | color   | L*                      | a*    | b*    | color   |                   |
| $H_2ARS$   | 99.6                      | -10.8 | 77.3  |    | 99.2                    | -14.0 | 68.0  |  | 3.3               |
| $HARS^-$   | 65.4                      | 54.8  | -3.2  |   | 59.1                    | 57.7  | -17.8 |  | 8.3               |
| $ARS^{2-}$ | 54.4                      | 51.0  | -50.9 |  | 57.0                    | 61.7  | -53.6 |  | 4.0               |

Fig. 7. (Top) The calculated and experimental UV-Vis spectra for alizarin red S at different pH. The dotted line is the experimental one. The full lines are the calculated spectrum with Adiabatic Hessian (AH) method. The vertical lines correspond to the computed VE. (Bottom) Predicted and computed colors of different alizarin red S forms with CIELAB coordinates ( $L^*$ ,  $a^*$ ,  $b^*$ ) and the calculated color variation between computed and experimental color coordinates with  $\Delta E_{2000}$ .

377 To assess the quality of the color prediction, the calculated and experimental color CIELAB are represented  
 378 in Fig.7. The calculated color matches the observed experimental color, with higher accuracy for the  $H_2ARS$   
 379 form (3.31 for alizarin red S and 13.32 for alizarin, see Fig.5). This can be explained again by its higher  
 380 solubility compared to alizarin.  
 381 The additional sulfonic group does not change the color of the compound compared to alizarin. This can be  
 382 explained by the difference densities plotted in Fig.8. The sulfonic group does not participate in the electron  
 383 delocalization for the  $GS \rightarrow S_1$  transition.

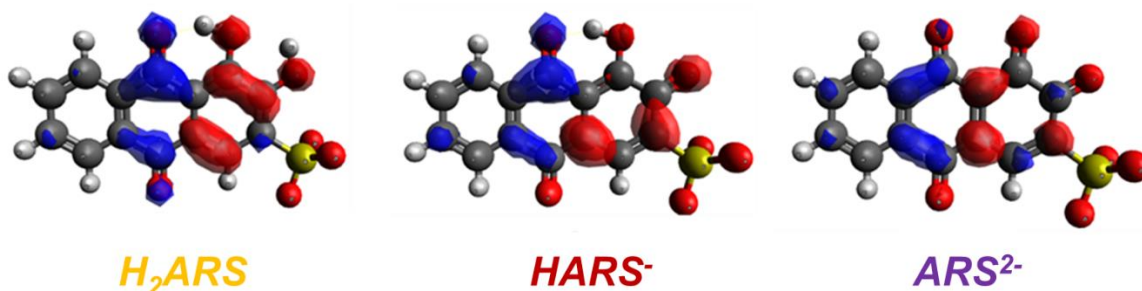


Fig. 8. Difference electronic densities for the  $GS \rightarrow S_1$  transition for alizarin red S for different protonation states. The blue orbital represents a decrease in electronic density. The red one represents an increase in electronic density. The isovalue is 0.02 a.u.

384

### 385 3.2.3. $^1H$ and $^{13}C$ NMR prediction

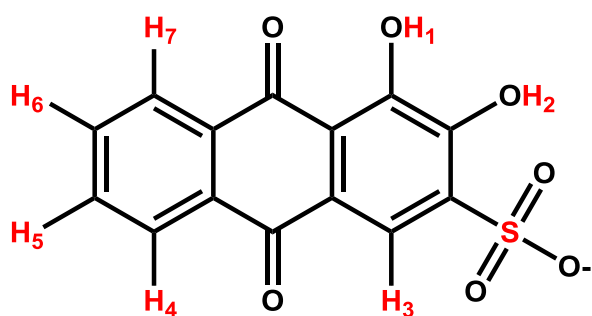
386

387 Thanks to the higher solubility of alizarin red S in water, the measure of NMR spectra in  $D_2O$  is possible. We  
 388 have obtained the  $^1H$  NMR spectra of  $H_2ARS$  in acetone- $D_6$  only as the one in water gives very broadened  
 389 signals, probably due to exchange processes at room temperature which is consistent with other  
 390 publications [26]. Following the best computational protocol for alizarin, the calculations were done using an  
 391 explicit solvation scheme. All the geometries are accessible in the supplementary data 14. The comparison  
 392 between the computed  $^1H$  chemical shifts is represented in Table 2 (in acetone- $D_6$ ) and  $^{13}C$  in Table 3 (for



393 both acetone-D<sub>6</sub> and D<sub>2</sub>O). The error deviation of aromatic hydrogens (H3 to H7) is less than 0.2 ppm in  
 394 average (which represents 2.1% in mean error deviation). This error is less than for alizarin (which has 2.8%  
 395 mean error deviation). There is a better agreement with experimental data for alizarin red S, again explained  
 396 by its higher solubility compared to alizarin. The labile hydrogens are less accurate with 17% of error  
 397 deviation for H2 and 11% for H2. The experimental and computed shifts can be found for different alizarin  
 398 red S forms in supplementary data 8 to 13.

Table 2. Computed and experimental chemical shift of alizarin red S in acetone-D<sub>6</sub> with the respective error deviation in ppm or %.

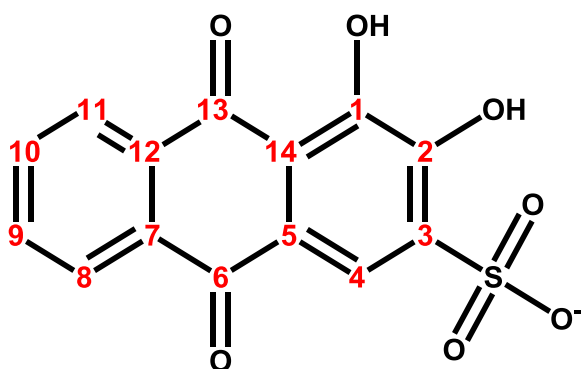


| Attribution | Exp. (ppm) in acetone-D <sub>6</sub> | Calc. (ppm) | Calc.-error (ppm) | Error deviation (%) |
|-------------|--------------------------------------|-------------|-------------------|---------------------|
| H1          | 12.8                                 | 15.0        | 2.2               | 17.1                |
| H2          | 11.9                                 | 13.2        | 1.3               | 11.1                |
| H3          | 8.2                                  | 8.4         | 0.2               | 2.7                 |
| H4          | 8.3                                  | 8.5         | 0.2               | 2.7                 |
| H6-H5       | 7.9                                  | 8.0         | 0.1               | 1.3                 |
| H6-H5       | 7.9                                  | 8.0         | 0.1               | 1.3                 |
| H7          | 8.3                                  | 8.5         | 0.2               | 2.5                 |

399  
 400 The comparison between the computed and experimental <sup>13</sup>C shifts in D<sub>2</sub>O for alizarin red S is represented in  
 401 Table 3. It shows that the error deviation of Carbons C4-C14 are less than 6 ppm on average (which  
 402 represents 4.4 % in mean error deviation). There is a higher error deviation for carbons that are next to a  
 403 hydroxy or sulfonic group (C1-C3) with up to 5.5% error deviation.  
 404 Overall, there is a good agreement between the computed and experimental chemical shifts for both <sup>1</sup>H and  
 405 <sup>13</sup>C NMR spectra. This protocol proves to be efficient in predicting NMR spectra.  
 406

Table 3. Computed and experimental chemical  $^{13}\text{C}$  shift of alizarin red S in acetone- $\text{D}_6$  or  $\text{D}_2\text{O}$  with the respective error deviation in ppm or %.

| Attribution | Acetone- $\text{D}_6$ |                     | $\text{D}_2\text{O}$ |                     |
|-------------|-----------------------|---------------------|----------------------|---------------------|
|             | Calc.-error (ppm)     | Error deviation (%) | Calc.-error (ppm)    | Error deviation (%) |
| 1           | 8.2                   | 5.5                 | 9.8                  | 6.5                 |
| 2           | 7.9                   | 5.2                 | 8.6                  | 5.7                 |
| 3           | 6.7                   | 5.0                 | 4.7                  | 3.5                 |
| 4           | 5.0                   | 4.2                 | 3.9                  | 3.3                 |
| 5           | 4.0                   | 3.3                 | 5.5                  | 4.5                 |
| 6           | 6.9                   | 3.8                 | 6.7                  | 3.7                 |
| 7           | 4.2                   | 3.1                 | 6.5                  | 4.9                 |
| 8           | 4.4                   | 3.4                 | 4.7                  | 3.7                 |
| 9           | 3.5                   | 2.5                 | 5.2                  | 3.9                 |
| 10          | 3.4                   | 2.5                 | 6.2                  | 4.7                 |
| 11          | 3.9                   | 3.1                 | 4.4                  | 3.5                 |
| 12          | 4.4                   | 3.3                 | 6.7                  | 5.1                 |
| 13          | 5.3                   | 2.8                 | 7.0                  | 3.7                 |
| 14          | 4.3                   | 3.7                 | 4.8                  | 4.1                 |



407

408

#### 4. CONCLUSION

409

410

411

412

413

414

415

416

417

Deciphering experimental spectra of natural dyes is a persistent challenge due to the presence of impurities, extraction difficulties, solubility issues, fast degradation, and so on. In this context, computational spectroscopy appears as a solution for the elucidation of complex structures of pigments. Here, we focus on Madder, a natural dye with historical importance. By combining computational insights with experimental data, our study has demonstrated the great efficiency of the optimized computational protocol for accurately predicting UV-Vis spectra, CIELAB color coordinates, and NMR shift of alizarin, the main coloring molecule of Madder, and its analog alizarin red S. In particular, we have found that the incorporation of conformer, deprotonated forms for each pH, solvation models, and careful consideration of solvent interactions resulted in accurate description when compared with experimental measurements. The

418 computed spectra closely match the experimental data, providing a reliable foundation for the prediction of  
419 other hydroxyanthraquinones and numerous colored dyes. The computation of CIELAB coordinates of these  
420 2 molecules also yields good results despite some issues due to solubility. This approach can be of interest to  
421 extracting the contribution of each dye in a mixture or for predicting new hues of colored molecules. By  
422 overcoming experimental complexities, this protocol accelerates the development of compounds tailored to  
423 specific needs by predicting with accuracy several properties. This optimized computational protocol  
424 outlined here offers a versatile tool for designing novel molecules across a wide variety of coloring  
425 industries, such as textiles, cosmetics, food, pharmaceuticals, and art. Furthermore, it can set the stage for  
426 deeper exploration into more complex structures of pigments, which are often challenging to elucidate by  
427 experimental methods.

428

#### 429 ■AUTHOR INFORMATION

430 Linh Thanh Huyen TRAN: linh.tran@sorbonne-universite.fr

431 Baptiste RIGAUD: baptiste.rigaud@sorbonne-universite.fr

432 Maguy JABER: maguy.jaber@sorbonne-universite.fr

433 Romain BERRAUD-PACHE: romain.berraud-pache@sorbonne-universite.fr

#### 434 ■REFERENCES

- 435 [1] Baroni A, Comite V, Guglielmi V, Casanova M, Redegalli P, Fermo P. Identification of natural dyes in  
436 historical tapestries, a LC-MS/MS approach. *J Phys: Conf Ser* 2022;2204:012065.  
437 <https://doi.org/10.1088/1742-6596/2204/1/012065>.
- 438 [2] Casadio F, Leona M, Lombardi JR, Van Duyne R. Identification of Organic Colorants in Fibers, Paints, and  
439 Glazes by Surface Enhanced Raman Spectroscopy. *Acc Chem Res* 2010;43:782–91.  
440 <https://doi.org/10.1021/ar100019q>.
- 441 [3] Fonseca B, Schmidt Patterson C, Ganio M, MacLennan D, Trentelman K. Seeing red: towards an  
442 improved protocol for the identification of madder- and cochineal-based pigments by fiber optics  
443 reflectance spectroscopy (FORS). *Herit Sci* 2019;7:92. <https://doi.org/10.1186/s40494-019-0335-1>.
- 444 [4] Brommelle N. Material for a History of Conservation. The 1850 and 1853 Reports on the National  
445 Gallery. *Studies in Conservation* 1956;2:176. <https://doi.org/10.2307/1504963>.
- 446 [5] De Viguerie L, Michelin A, Radepont M, Pottier F, Alfeld M, Walter P, et al. Imagerie hyperspectrale  
447 pour l'analyse d'œuvres peintes. In: Benech C, Cantin N, Languille M-A, Mazuy A, Robinet L, Zazzo A,  
448 editors. *Instrumentation portable*, Editions des archives contemporaines; 2019, p. 43–65.  
449 <https://doi.org/10.17184/eac.2457>.
- 450 [6] Pronti L, Mazzitelli J-B, Bracciale MP, Massini Rosati L, Vieillescazes C, Santarelli ML, et al. Multi-  
451 technique characterisation of commercial alizarin-based lakes. *Spectrochimica Acta Part A: Molecular  
452 and Biomolecular Spectroscopy* 2018;200:10–9. <https://doi.org/10.1016/j.saa.2018.04.008>.
- 453 [7] Silva LC, Otero V, Melo MJ, Cabrita EJ, Mafra L. What's the Madder? Characterization of Old Fashioned  
454 Alizarin/Aluminum Red Pigments Using Liquid and Solid-State NMR. *Colorants* 2023;2:601–17.  
455 <https://doi.org/10.3390/colorants2040031>.
- 456 [8] Blackburn RS. Natural dyes in madder (*Rubia* spp.) and their extraction and analysis in historical  
457 textiles. *Coloration Technol* 2017;133:449–62. <https://doi.org/10.1111/cote.12308>.
- 458 [9] Janssens K, Van Der Snickt G, Vanmeert F, Legrand S, Nuyts G, Alfeld M, et al. Non-Invasive and Non-  
459 Destructive Examination of Artistic Pigments, Paints, and Paintings by Means of X-Ray Methods. *Top  
460 Curr Chem (Z)* 2016;374:81. <https://doi.org/10.1007/s41061-016-0079-2>.
- 461 [10] Creagh DC. The characterization of artefacts of cultural heritage significance using physical techniques.  
462 *Radiation Physics and Chemistry* 2005;74:426–42.  
463 <https://doi.org/10.1016/j.radphyschem.2005.08.011>.

- 464 [11] Krizsán K, Szókán Gy, Toth ZA, Hollósy F, László M, Khlafula A. HPLC Analysis of Anthraquinone  
465 Derivatives in Madder Root (*Rubia Tinctorum*) and Its Cell Cultures. *Journal of Liquid Chromatography*  
466 & Related Technologies 1996;19:2295–314. <https://doi.org/10.1080/10826079608017158>.
- 467 [12] Degano I, Tognotti P, Kunzelman D, Modugno F. HPLC-DAD and HPLC-ESI-Q-ToF characterisation of  
468 early 20th century lake and organic pigments from Lefranc archives. *Herit Sci* 2017;5:7.  
469 <https://doi.org/10.1186/s40494-017-0120-y>.
- 470 [13] Rafaelly L, Heron S, Nowik W, Tchaplá A. Optimisation of ESI-MS detection for the HPLC of  
471 anthraquinone dyes. *Dyes and Pigments* 2008;77:191–203.  
472 <https://doi.org/10.1016/j.dyepig.2007.05.007>.
- 473 [14] Cosentino A. FORS Spectral Database of Historical Pigments in Different Binders. *E-Cons* 2014;54–65.  
474 <https://doi.org/10.18236/econs2.201410>.
- 475 [15] Cruz Mojica K, León Santiago M, Varela García E, Reyes Lezama M, Tapia Mendoza E. New progress in  
476 the extraction and identification of Mexican cochineal (*Dactylopius coccus*) in fresh dyed wool and  
477 dyed wool with artificially accelerated aging and color description. *Color Research & Application*  
478 2023;48:151–64. <https://doi.org/10.1002/col.22828>.
- 479 [16] Baroni A, Comite V, Guglielmi V, Casanova M, Redegalli P, Fermo P. Identification of natural dyes in  
480 historical tapestries, a LC-MS/MS approach. *J Phys: Conf Ser* 2022;2204:012065.  
481 <https://doi.org/10.1088/1742-6596/2204/1/012065>.
- 482 [17] Whitney AV, Casadio F, Van Duyne RP. Identification and Characterization of Artists' Red Dyes and  
483 Their Mixtures by Surface-Enhanced Raman Spectroscopy. *Appl Spectrosc* 2007;61:994–1000.  
484 <https://doi.org/10.1366/000370207781745838>.
- 485 [18] Ford L, Rayner CM, Blackburn RS. Degradation of lucidin: New insights into the fate of this natural  
486 pigment present in Dyer's madder (*Rubia tinctorum* L.) during the extraction of textile artefacts. *Dyes*  
487 *and Pigments* 2018;154:290–5. <https://doi.org/10.1016/j.dyepig.2018.03.023>.
- 488 [19] Ford L, Henderson RL, Rayner CM, Blackburn RS. Mild extraction methods using aqueous glucose  
489 solution for the analysis of natural dyes in textile artefacts dyed with Dyer's madder (*Rubia tinctorum*  
490 L.). *Journal of Chromatography A* 2017;1487:36–46. <https://doi.org/10.1016/j.chroma.2017.01.053>.
- 491 [20] Barone V, Alessandrini S, Biczysko M, Cheeseman JR, Clary DC, McCoy AB, et al. Computational  
492 molecular spectroscopy. *Nat Rev Methods Primers* 2021;1:38. <https://doi.org/10.1038/s43586-021-00034-1>.
- 493
- 494 [21] Chenciner R. *Madder Red: A History of Luxury and Trade*. Routledge; 2003.
- 495 [22] Tian G, Cui R, Liu C, Hu X. Natural Dyeing with Madder: Exploring Traditional Techniques and Color  
496 Characteristics. *Fibres & Textiles in Eastern Europe* 2023;31:10–21. <https://doi.org/10.2478/ftce-2023-0031>.
- 497
- 498 [23] Hui-Yu J. Studies on the photofading of alizarin, the main component of madder | Elsevier Enhanced  
499 Reader 2020. <https://doi.org/10.1016/j.dyepig.2020.108940>.
- 500 [24] Ackacha MA, Połec-Pawlak K, Jarosz M. Identification of anthraquinone coloring matters in natural red  
501 dyestuffs by high performance liquid chromatography with ultraviolet and electrospray mass  
502 spectrometric detection. *Journal of Separation Science* 2003;26:1028–34.  
503 <https://doi.org/10.1002/jssc.200301484>.
- 504 [25] Cuoco G, Mathe C, Archier P, Vieillescazes C. Characterization of madder and garancine in historic  
505 French red materials by liquid chromatography-photodiode array detection. *Journal of Cultural*  
506 *Heritage* 2011;12:98–104. <https://doi.org/10.1016/j.culher.2010.05.005>.
- 507 [26] Justino LLG, Braz S, Ramos ML. Spectroscopic and DFT Study of Alizarin Red S Complexes of Ga(III) in  
508 Semi-Aqueous Solution. *Photochem* 2023;3:61–81. <https://doi.org/10.3390/photochem3010005>.
- 509 [27] Amat A, Miliari C, Romani A, Fantacci S. DFT/TDDFT investigation on the UV-vis absorption and  
510 fluorescence properties of alizarin dye. *Phys Chem Chem Phys* 2015;17:6374–82.  
511 <https://doi.org/10.1039/C4CP04728A>.
- 512 [28] Mech J, Grela MA, Szaćiowski K. Ground and excited state properties of alizarin and its isomers. *Dyes*  
513 *and Pigments* 2014;103:202–13. <https://doi.org/10.1016/j.dyepig.2013.12.009>.

- 514 [29] Cysewski P, Jeliński T, Przybyłek M, Shyichuk A. Color prediction from first principle quantum chemistry  
515 computations: a case of alizarin dissolved in methanol. *New J Chem* 2012;36:1836.  
516 <https://doi.org/10.1039/c2nj40327g>.
- 517 [30] Jacquemin D, Brémond E, Planchat A, Ciofini I, Adamo C. TD-DFT Vibronic Couplings in Anthraquinones:  
518 From Basis Set and Functional Benchmarks to Applications for Industrial Dyes. *J Chem Theory Comput*  
519 2011;7:1882–92. <https://doi.org/10.1021/ct200259k>.
- 520 [31] Anouar EH, Osman CP, Weber J-FF, Ismail NH. UV/Visible spectra of a series of natural and synthesised  
521 anthraquinones: experimental and quantum chemical approaches. *SpringerPlus* 2014;3:233.  
522 <https://doi.org/10.1186/2193-1801-3-233>.
- 523 [32] Berenbeim JA, Boldissar S, Owens S, Haggmark MR, Gate G, Siouri FM, et al. Excited state  
524 intramolecular proton transfer in hydroxyanthraquinones: Toward predicting fading of organic red  
525 colorants in art. *Sci Adv* 2019;5:eaaw5227. <https://doi.org/10.1126/sciadv.aaw5227>.
- 526 [33] Tissier R-C, Rigaud B, Thureau P, Jaber M, Ferré N. Stressing the differences in alizarin and purpurin  
527 dyes through UV- visible light absorption and <sup>1</sup>H-NMR spectroscopies n.d.:10.
- 528 [34] Jacquemin D, Brémond E, Ciofini I, Adamo C. Impact of Vibronic Couplings on Perceived Colors: Two  
529 Anthraquinones as a Working Example. *J Phys Chem Lett* 2012;3:468–71.  
530 <https://doi.org/10.1021/jz201552x>.
- 531 [35] Di Tommaso S, Bousquet D, Moulin D, Baltenneck F, Riva P, David H, et al. Theoretical approaches for  
532 predicting the color of rigid dyes in solution. *J Comput Chem* 2017;38:998–1004.  
533 <https://doi.org/10.1002/jcc.24774>.
- 534 [36] Le Person A, Cornard J-P, Say-Liang-Fat S. Studies of the tautomeric forms of alizarin in the ground  
535 state by electronic spectroscopy combined with quantum chemical calculations. *Chemical Physics*  
536 *Letters* 2011;517:41–5. <https://doi.org/10.1016/j.cplett.2011.10.015>.
- 537 [37] Tirri B, Turelli M, Boissonnat G, Ciofini I, Adamo C. Protocols for the in-silico screening of the perceived  
538 color of industrial dyes: Anthraquinones and indigos as study cases. *Dyes and Pigments*  
539 2023;208:110826. <https://doi.org/10.1016/j.dyepig.2022.110826>.
- 540 [38] Carta L, Biczysko M, Bloino J, Licari D, Barone V. Environmental and complexation effects on the  
541 structures and spectroscopic signatures of organic pigments relevant to cultural heritage: the case of  
542 alizarin and alizarin–Mg(ii)/Al(iii) complexes. *Phys Chem Chem Phys* 2014;16:2897.  
543 <https://doi.org/10.1039/c3cp50499a>.
- 544 [39] Pedone A, Biczysko M, Barone V. Environmental Effects in Computational Spectroscopy: Accuracy and  
545 Interpretation. *ChemPhysChem* 2010;11:1812–32. <https://doi.org/10.1002/cphc.200900976>.
- 546 [40] Miliani C, Romani A, Favaro G. Acidichromic effects in 1,2-di- and 1,2,4-tri- hydroxyanthraquinones. A  
547 spectrophotometric and fluorimetric study. *Journal of Physical Organic Chemistry* 2000;13:141–50.  
548 [https://doi.org/10.1002/\(SICI\)1099-1395\(200003\)13:3<141::AID-POC220>3.0.CO;2-J](https://doi.org/10.1002/(SICI)1099-1395(200003)13:3<141::AID-POC220>3.0.CO;2-J).
- 549 [41] Shalaby AA, Mohamed AA. Determination of acid dissociation constants of Alizarin Red S, Methyl  
550 Orange, Bromothymol Blue and Bromophenol Blue using a digital camera. *RSC Adv* 2020;10:11311–6.  
551 <https://doi.org/10.1039/C9RA10568A>.
- 552 [42] Software update: The ORCA program system—Version 5.0 - Neese - 2022 - WIREs Computational  
553 Molecular Science - Wiley Online Library n.d.  
554 <https://wires.onlinelibrary.wiley.com/doi/full/10.1002/wcms.1606> (accessed August 24, 2023).
- 555 [43] Bannwarth C, Ehlert S, Grimme S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent  
556 Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent  
557 Dispersion Contributions. *J Chem Theory Comput* 2019;15:1652–71.  
558 <https://doi.org/10.1021/acs.jctc.8b01176>.
- 559 [44] Parr RG, Weitao Y. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press; 1994.
- 560 [45] Marques MAL, Maitra NT, Nogueira FMS, Gross EKV, Rubio A. *Fundamentals of time-dependent*  
561 *density functional theory. Fundamentals of Time-Dependent Density Functional Theory*, 2012, p. 1–  
562 559.
- 563 [46] Adamo C, Barone V. Toward reliable density functional methods without adjustable parameters: The  
564 PBE0 model. *The Journal of Chemical Physics* 1999;110:6158–70. <https://doi.org/10.1063/1.478522>.

- 565 [47] Effect of the damping function in dispersion corrected density functional theory - Grimme - 2011 -  
566 Journal of Computational Chemistry - Wiley Online Library n.d.  
567 <https://onlinelibrary.wiley.com/doi/10.1002/jcc.21759> (accessed December 3, 2023).
- 568 [48] Becke AD, Johnson ER. A density-functional model of the dispersion interaction. *J Chem Phys*  
569 2005;123:154101. <https://doi.org/10.1063/1.2065267>.
- 570 [49] A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D)  
571 for the 94 elements H-Pu | *The Journal of Chemical Physics* | AIP Publishing n.d.  
572 [https://pubs.aip.org/aip/jcp/article-abstract/132/15/154104/926936/A-consistent-and-accurate-ab-  
574 initio?redirectedFrom=fulltext](https://pubs.aip.org/aip/jcp/article-abstract/132/15/154104/926936/A-consistent-and-accurate-ab-<br/>573 initio?redirectedFrom=fulltext) (accessed December 3, 2023).
- 574 [50] Weigend F, Ahlrichs R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta  
575 valence quality for H to Rn: Design and assessment of accuracy. *Phys Chem Chem Phys* 2005;7:3297–  
576 305. <https://doi.org/10.1039/B508541A>.
- 577 [51] Pracht P, Bohle F, Grimme S. Automated exploration of the low-energy chemical space with fast  
578 quantum chemical methods. *Phys Chem Chem Phys* 2020;22:7169–92.  
579 <https://doi.org/10.1039/C9CP06869D>.
- 580 [52] Liakos DG, Guo Y, Neese F. Comprehensive Benchmark Results for the Domain Based Local Pair Natural  
581 Orbital Coupled Cluster Method (DLPNO-CCSD(T)) for Closed- and Open-Shell Systems. *J Phys Chem A*  
582 2020;124:90–100. <https://doi.org/10.1021/acs.jpca.9b05734>.
- 583 [53] Garcia-Ratés M, Neese F. Efficient implementation of the analytical second derivatives of hartree–fock  
584 and hybrid DFT energies within the framework of the conductor-like polarizable continuum model.  
585 *Journal of Computational Chemistry* 2019;40:1816–28. <https://doi.org/10.1002/jcc.25833>.
- 586 [54] Warshel A, Levitt M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric  
587 stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology*  
588 1976;103:227–49. [https://doi.org/10.1016/0022-2836\(76\)90311-9](https://doi.org/10.1016/0022-2836(76)90311-9).
- 589 [55] Groenhof G. Introduction to QM/MM Simulations. In: Monticelli L, Salonen E, editors. *Biomolecular*  
590 *Simulations: Methods and Protocols*, Totowa, NJ: Humana Press; 2013, p. 43–66.  
591 [https://doi.org/10.1007/978-1-62703-017-5\\_3](https://doi.org/10.1007/978-1-62703-017-5_3).
- 592 [56] Senn HM, Thiel W. QM/MM Methods for Biomolecular Systems. *Angewandte Chemie International*  
593 *Edition* 2009;48:1198–229. <https://doi.org/10.1002/anie.200802019>.
- 594 [57] Neese F, Wennmohs F, Becker U, Riplinger C. The ORCA quantum chemistry program package. *The*  
595 *Journal of Chemical Physics* 2020;152:224108. <https://doi.org/10.1063/5.0004608>.
- 596 [58] Martyna GJ, Tuckerman ME, Tobias DJ, Klein ML. Explicit reversible integrators for extended systems  
597 dynamics. *Molecular Physics* 1996;87:1117–57. <https://doi.org/10.1080/00268979600100761>.
- 598 [59] Jacquemin D, Wathelet V, Perpète EA, Adamo C. Extensive TD-DFT Benchmark: Singlet-Excited States of  
599 Organic Molecules. *J Chem Theory Comput* 2009;5:2420–35. <https://doi.org/10.1021/ct900298e>.
- 600 [60] Becke AD. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical*  
601 *Physics* 1993;98:5648–52. <https://doi.org/10.1063/1.464913>.
- 602 [61] Laurent AD, Jacquemin D. TD-DFT benchmarks: A review. *International Journal of Quantum Chemistry*  
603 2013;113:2019–39. <https://doi.org/10.1002/qua.24438>.
- 604 [62] Jacquemin D, Mennucci B, Adamo C. Excited-state calculations with TD-DFT: from benchmarks to  
605 simulations in complex environments. *Phys Chem Chem Phys* 2011;13:16987.  
606 <https://doi.org/10.1039/c1cp22144b>.
- 607 [63] Orlandi G, Siebrand W. Theory of vibronic intensity borrowing. Comparison of Herzberg-Teller and  
608 Born-Oppenheimer coupling. *The Journal of Chemical Physics* 2003;58:4513–23.  
609 <https://doi.org/10.1063/1.1679014>.
- 610 [64] Small GJ. Herzberg–Teller Vibronic Coupling and the Duschinsky Effect. *The Journal of Chemical Physics*  
611 2003;54:3300–6. <https://doi.org/10.1063/1.1675343>.
- 612 [65] Sussmann R, Neuhauser R, Neusser HJ. High-resolution UV spectroscopy of vibronic bands in p-  
613 difluorobenzene and p-difluorobenzene–Ar: The role of Herzberg–Teller coupling. *Can J Phys*  
614 1994;72:1179–86. <https://doi.org/10.1139/p94-151>.

615 [66] de Souza B, Neese F, Izsák R. On the theoretical prediction of fluorescence rates from first principles  
616 using the path integral approach. *The Journal of Chemical Physics* 2018;148:034104.  
617 <https://doi.org/10.1063/1.5010895>.

618 [67] Santoro F, Jacquemin D. Going beyond the vertical approximation with time-dependent density  
619 functional theory. *WIREs Computational Molecular Science* 2016;6:460–86.  
620 <https://doi.org/10.1002/wcms.1260>.

621 [68] Johnston-Feller R. Color science in the examination of museum objects: Nondestructive procedures.  
622 *Color Research & Application* 2002;27:456–7. <https://doi.org/10.1002/col.10107>.

623 [69] Colorimetry — Part 4: CIE 1976 L\*a\*b\* colour space | CIE n.d.  
624 <https://cie.co.at/publications/colorimetry-part-4-cie-1976-lab-colour-space-1> (accessed January 29,  
625 2024).

626 [70] Karma IGM. Determination and Measurement of Color Dissimilarity. *International Journal of*  
627 *Engineering and Emerging Technology* 2020;5:67. <https://doi.org/10.24843/IJEET.2020.v05.i01.p13>.

628 [71] Mokrzycki W, Tatol M. Color difference Delta E - A survey. *Machine Graphics and Vision* 2011;20:383–  
629 411.

630 [72] Habekost M. Which color differencing equation should be used? 2013.

631 [73] Santra G, Martin JML. Pure and Hybrid SCAN, rSCAN, and r2SCAN: Which One Is Preferred in KS- and  
632 HF-DFT Calculations, and How Does D4 Dispersion Correction Affect This Ranking? *Molecules*  
633 2022;27:141. <https://doi.org/10.3390/molecules27010141>.

634 [74] de Oliveira MT, Alves JMA, Braga AAC, Wilson DJD, Barboza CA. Do Double-Hybrid Exchange–  
635 Correlation Functionals Provide Accurate Chemical Shifts? A Benchmark Assessment for Proton NMR. *J*  
636 *Chem Theory Comput* 2021;17:6876–85. <https://doi.org/10.1021/acs.jctc.1c00604>.

637 [75] Jensen F. Segmented Contracted Basis Sets Optimized for Nuclear Magnetic Shielding. *J Chem Theory*  
638 *Comput* 2015;11:132–8. <https://doi.org/10.1021/ct5009526>.

639 [76] Preat J, Laurent AD, Michaux C, Perpète EA, Jacquemin D. Impact of tautomers on the absorption  
640 spectra of neutral and anionic alizarin and quinizarin dyes. *Journal of Molecular Structure: THEOCHEM*  
641 2009;901:24–30. <https://doi.org/10.1016/j.theochem.2008.12.032>.

642 [77] Giovannini T, Egidi F, Cappelli C. Molecular spectroscopy of aqueous solutions: a theoretical  
643 perspective. *Chem Soc Rev* 2020;49:5664–77. <https://doi.org/10.1039/C9CS00464E>.

644

645