SORBONNE UNIVERSITY

LABORATOIRE D'ARCHÉOLOGIE MOLÉCULAIRE ET STRUCTURALE, UMR 8220, Paris, 75005, France

PHOTODEGRADATION OF HYDROXYANTHRAQUINONES BY COMBINING COMPUTATIONAL AND EXPERIMENTAL APPROACHES

by TRAN THANH HUYEN LINH

To obtain the grade of Doctor of Philosophy in
Physics and Chemistry of Materials

Directed by: Maguy JABER, Professor, Sorbonne University
Supervisor: Romain BERRAUD-PACHE, Associate Professor, Sorbonne University

JURY

Reviewers:

Adèle LAURENT, Research Director, Nantes University

Aurélie PERRIER, Professor, Paris-Cité University

Examiners:

Isabelle NAVIZET, Professor, Gustave Eiffel University (President of the jury)

Nicolas FERRÉ, Professor, Aix-Marseille University

Fabrice MONTI, PhD. CEO Microfactory

I dedicate this thesis to my parents, my brother, my grandmother and my two uncles as well as all my loved ones who encouraged me on this path.

« On ne voit bien qu'avec le cœur, l'essentiel est invisible pour les yeux »

Antoine de Saint-Exupéry

AKNOWLEDGEMENTS

This thesis would not have been well completed without the support, advice, and guidance of many people, whom I would like to sincerely thank.

Above all, I would like to express my deep gratitude to my two thesis advisors, Doctor Romain BERRAUD-PACHE and Professor Maguy JABER. Working with you on this thesis has been a real pleasure!

Thank you so much, Romain, for your availability, your scientific rigor, and the trust you placed in me from the very beginning of this adventure.

A big thank you to you, Maguy, director of the lab, for creating the perfect environment that allowed me to advance in my experiments and research.

I would also like to thank the members of the jury, Professor Isabelle NAVIZET, Professor Aurélie PERRIER, Doctor Adèle, Professor Nicolas FERRÉ, and Doctor and CEO Fabrice MONTI, for agreeing to evaluate this work and for their constructive comments, which have enriched my thinking.

I am also grateful to the ED397 doctoral school of physics and chemistry of materials for funding this thesis.

A warm thank you to all my colleagues in the laboratory. Thank you to Suzie, Alain, Sophie, Laurence, Rémi, Philippe, and all the doctoral students and interns for your advice and discussions.

Thanks to Laurent and Salomé for their help with HPLC-UV-MS. Thanks to Baptiste for his technical assistance with NMR. Thanks to Jacques and Fabrice from Mulhouse for the discussions on RPE.

Thanks to Professor Nadine WITKOWSKI and Professor Olivier DURUPTHY, ex-director and director of the Doctoral School ED397 for encouraging me to create and preside over the Materials'Doc association. Thanks to the sponsors whose support made the successful organisation of the Café Alumni Conference possible: Professor Xavier CARRIER, Professor Souhir BOUJDAY, Professor Jérôme TIGNON, and Professor Olivier DURUPTHY. Thanks to that experience, I grew from being an introverted person into a more confident and proactive researcher, which ultimately led to the publication of five articles.

I would like to thank Dr. Nicolas SISOURAT for inviting me to give an extended lecture of 45min during the "*Mardis de la Chimie*" seminars. I had the honour of being the first student to give a lecture in this seminar series. The event was a great success, with such a large audience that there were not enough chairs. The discussions with the participants were enriching and made me realise the genuine interest in my research. This experience dispelled the myth that research ends only in the library, and I am grateful for such a lively and inspiring opportunity.

Thank you, Dr. Claire VILAIN, Editor in chief of CultureSciences-Chimie for inviting me to write a vulgarisation article for a broad public. This opportunity made me realize that my research is also meaningful and useful to the public.

I am grateful to Dr. Agnès BÉE for the opportunity to teach at the Sorbonne University, Paris, at both the bachelor's and master's levels. This allowed me to share my knowledge and passion with students. I would also like to thank my students for their thoughtful questions and stimulating discussions.

Finally, a special thank you to my loved ones, especially my family and my parents, my brother, my grandmother, my two uncles, and all my relatives for their patience, their pride and unconditional support. It is an immeasurable force that gave me the strength and perseverance to overcome every challenge I faced while completing my thesis. This thesis is also a little bit yours. Because without you, I would not have gotten this far. Thank you for giving me the solid foundation on which I was able to build this path. I dedicate this important milestone in my life to you with deep emotion.

Finally, to all those who, in one way or another, have contributed to this intellectual and human adventure: thank you.



SUMMARY

SUMMARY 6

RÉSUMÉ EN FRANÇAIS / SUMMARY IN FRENCH 10

CHAPTER 1 | GENERAL INTRODUCTION -IN SEARCH OF LOST COLOURS: INVESTIGATION OF THE PHOTODEGRADATION OF NATURAL MADDER DYES 21

CHAPTER 2 | STATE OF THE ART 27

2.1. A history dyed in madder, a natural red dye 27
2.1.1. The colour red: a hue with deep social and symbolic meaning 27
2.1.2. Madder: a natural source of red dye 27
2.1.3. Madder as a dye or pigment/lake 29

2.2. Madder species 31
2.2.1. Identification technique of madder molecules in artworks 32
2.2.2. Composition of madder dyes 33

2.3. Overview state of the art on photodegradation of dyes 35
2.3.1. Reciprocity principle and accelerated photodegradation 37
2.3.2. Internal factors influencing photodegradation 38
2.3.3. External factors influencing photodegradation 38

2.4. Photodegradation of madder 41
2.4.1. Photodegradation of madder in artworks 41
2.4.2. Photodegraded products and proposed mechanism 42

2.5. Difficulty in studying madder dyes 43

2.6. Computational spectroscopy: state of the art 44
2.6.1. Computational chemistry in the cultural heritage field and dyes 44
2.6.2. Role of functional and basis set for TD-DFT vertical energy 45
2.6.3. Absorption spectra and colour: vibronic coupling 46
2.6.4. Predicting the influence of the environment 48

CHAPTER 3 | EXPERIMENTAL METHODS 51

3.1. Invasive methods 51
3.1.1. Liquid NMR 51
3.1.2. HPLC-UV-MS 52
3.1.3. EPR 53

3.2. Non-Invasive methods 55
3.2.1. Absorption spectra 55

3.3. The science of colour 57
3.3.1. Traditional colour theory 57
3.3.2. Colour science 57
3.3.3. Colourimetry 58
3.3.4. Measure colour: CIELAB space 58
3.3.5. Measure colour difference with ΔE 59

3.3.6. Converting UV-VIS spectra to CIELAB coordinates 61

3.4. Photodegradation 62

CHAPTER 4 | COMPUTATIONAL METHODS 67

4.1. The foundations of quantum mechanics 68

4.2. Computational chemistry methods 71
4.2.1. Ab initio methods 71
4.2.2. DFT methods 72
4.2.3. Semiempirical methods 72
4.2.4. Computing electronic molecular energy 72
4.2.5. Hartree-Fock method: a brief overview 72
4.2.6. The Roothaan equation 73
4.2.7. SCF procedure 74
4.2.8. Basis sets 75
4.2.9. Electron correlation 77

4.3. Density Functional Theory (DFT) 78
4.3.1. The exchange-correlation challenge and functionals 79
4.3.2. Hybrid functional 81
4.3.3. Double-hybrid functionals 81
4.3.4. Energy derivatives and Hessian matrix 82
4.3.5. Molecular properties and gradient method 82

4.4. Time-Dependent DFT (TD-DFT) 82
4.4.2. Linear response theory 84

4.5. Absorbance spectra simulation 84
4.5.1. Oscillator Strength (or Oscillator force) 85
4.5.2. Spectra intensity with vibrational effects 85
4.5.3. Vibronic couplings 88

4.6. Modelling the environment 95
4.6.1. Vacuo 95
4.6.2. Implicit model 95
4.6.3. Microsolvation model 97
4.6.4. Explicit solvation 98

CHAPTER 5 | ARTICLES 101

Article 1: "Rapid and complete prediction of alizarin in solution by combining experimental data with computational methods." 102

Article 2: "In-silico colour prediction process for natural dyes in Madder." 131

Article 3: "Progress in modeling complex dye mixtures: a case study for cultural heritage." 169

Article 4: "When light betrays colour: the photodegradation of alizarin, a historical dye, in basic media." 195

Article 5: "A computational insight into the photodegradation mechanism of alizarin " 245

CHAPTER 6 | CONCLUSION AND PERSPECTIVES 254

REFERENCES 259

ABBREVIATIONS 278

ABSTRACT 281

RÉSUMÉ EN FRANÇAIS/ FRENCH SUMMARY

Nous sommes entourés d'une abondante diversité de couleurs lorsque nous regardons les objets autour de nous. Ce foisonnement de couleurs, que nous contemplons aujourd'hui et que nous prenons parfois pour acquis, n'a rien d'anodin. Ces centaines de teintes sont le fruit d'un long processus d'expérimentation, de découvertes progressives de plusieurs civilisations, et parfois même d'accidents, avant de devenir les couleurs que nous connaissons aujourd'hui. Des chercheurs ont identifié les premiers pigments, créés il y a 40 000 ans. À cette époque, on combinait des éléments qu'on trouvait dans son environnement proche, comme de la terre, de la graisse animale, du charbon de bois brûlé ou de la craie, pour créer les pigments désirés. Cela a permis de créer une palette de couleurs de base. Celle-ci n'était constituée que de cinq couleurs principales : le rouge, le noir, le brun, le jaune et le blanc. Depuis ces premières tentatives empiriques, l'histoire de la quête de la couleur et des pigments s'est enrichie de découvertes continues. D'époque en époque, d'une civilisation à la suivante, les savoir-faire acquis se sont transmis, perfectionnés et diversifiés, bien au-delà du simple mélange brut initial. Ainsi, les techniques de fabrication des colorants, des pigments et leurs utilisations sont le fruit de connaissances complexes. Elles font partie de l'héritage culturel humain. Chaque couleur porte en elle les traces d'une époque, d'une invention, d'un symbole ou même d'un territoire. Dans le domaine du patrimoine culturel, les couleurs sont des témoins silencieux d'une civilisation. Ainsi, savoir les identifier et les décoder permet de remonter à l'histoire de l'œuvre, aux messages de l'auteur, aux savoir-faire, et même aux routes commerciales. Cela permet parfois aussi de dater une œuvre, d'authentifier mais aussi surtout d'aider à sa restauration et à sa conservation.

Pour voir la couleur, nous avons besoin de lumière. Lorsque celle-ci atteint un objet, une partie est absorbée tandis qu'une autre est réfléchie. Notre œil n'est capable de percevoir que celles réfléchies.

L'un des premiers à avoir découvert un lien entre la lumière et la couleur est Isaac Newton. Il a publié un premier article en 1672 sur la découverte que la lumière blanche est composée de plusieurs couleurs et que les couleurs des objets sont déterminées par la proportion et la nature de la lumière réfléchie vers nos yeux. Cette découverte a amené au 19ème siècle à la formulation des théories des couleurs et aux cercles chromatiques tels que nous les connaissons aujourd'hui. Ces bases théoriques ont ensuite permis de représenter les couleurs de manière plus structurée, à travers par exemple des systèmes de coordonnées colorimétriques tels que le RGB, ou encore le CIELAB. Ce dernier est utilisé dans cette thèse pour sa prise en compte de la sensibilité de l'œil humain aux différentes longueurs d'onde. La couleur peut aussi être analysée via son spectre d'absorbance, qui nous donne une empreinte spectrale de chaque colorant ou pigment et permet donc de les identifier mais aussi de les quantifier. Ce spectre renseigne sur la nature des colorants, les transitions électroniques, les longueurs d'onde absorbées ou transmises, la concentration, les altérations éventuelles. L'étude du spectre d'absorbance fait donc partie de l'approche méthodologique de cette thèse.

Malheureusement, la lumière dégrade les molécules, en particulier organiques, et donc leur couleur avec le temps. Ce phénomène est appelé la photodégradation. Cela entraîne

le ternissement, le changement ou même la disparition des couleurs. Un des exemples célèbres de la photodégradation dans les peintures est le brunissement du jaune dans « *Sunflowers* » de Vincent van Gogh, ou encore le violet de la chambre « *Bedroom* » qui est devenu bleu. De même, le rouge dans « *Madame Léon Clapisson* » de Pierre-Auguste Renoir a complètement disparu.

C'est pourquoi ce sujet a mobilisé l'attention de chimistes, physiciens, restaurateurs et industriels. À partir de 1888, les travaux pionniers de Russell et Abney comptent parmi les premières études à reporter les effets de la photodégradation dans les musées anglais, suscitant un regain d'intérêt parmi les scientifiques. Pour étudier ces processus et leurs mécanismes de photodégradation, les scientifiques ont eu recours à plusieurs techniques. On peut les classer en deux catégories principales : les techniques invasives (avec prélèvement d'un échantillon) et non invasives (sans extraction d'échantillon). Ces dernières sont privilégiées dans le domaine du patrimoine pour leur caractère non destructif ou faiblement destructif, un critère important pour la préservation des biens irremplaçables, historiques et culturels. Parmi elles, la spectroscopie, qui regroupe l'ensemble des méthodes d'analyse basées sur l'interaction entre la matière et les rayonnements électromagnétiques, occupe une place centrale. En analysant la façon dont un objet absorbe, réfléchit ou transmet la lumière à différentes longueurs d'onde, il est possible d'en déduire sa composition chimique et les dégradations éventuelles. Grâce à toutes ces approches, il a été découvert que lorsque les colorants absorbent des photons, cela peut conduire à des réactions photophysiques et/ou photochimiques au sein de la molécule colorée, pouvant entraîner des changements ou ruptures de liaisons chimiques. Ce phénomène touche à la fois des molécules inorganiques et organiques et un large éventail de domaines tels que la cosmétique, la pharmaceutique, l'agroalimentaire ou encore les matériaux polymères. L'étude des mécanismes de photodégradation et la compréhension de l'influence de divers facteurs ont permis d'élaborer des stratégies d'exposition des objets historiques plus sûres et mieux contrôlées. De plus, cela permet d'informer tant les historiens que le grand public sur l'apparence originale des œuvres d'art, ce qui est important pour saisir l'intention de l'artiste.

La compréhension des mécanismes précis des processus de photodégradation reste encore incomplète jusqu'à aujourd'hui. Comment se dégradent les colorants ? Par quels mécanismes et comment les prévenir de façon durable ?

Ce problème revêt un intérêt particulier aujourd'hui, où les enjeux écologiques sont devenus primordiaux. Malgré des technologies de pointe disponibles, la compréhension fine de ces mécanismes demeure complexe et fastidieuse. Sur le plan théorique, le principal obstacle réside dans les approximations inhérentes aux calculs, qui deviennent d'autant plus problématiques que le système étudié est de grande taille.

C'est dans ce contexte que s'inscrit cette thèse, qui vise à étudier la photodégradation des colorants issus de la garance (*Rubia tinctorum* L.), une plante riche en histoire et connue pour ses colorants dans la gamme rouge. Cette thèse a été effectuée au LAMS (Laboratoire d'Archéologie Moléculaire et Structurale), basé à Sorbonne Université, Paris. Le LAMS est spécialisé dans l'analyse scientifique des matériaux du patrimoine, en particulier les analyses moléculaires et structurales. Les sujets sont au croisement de la chimie, de la

physique et de l'histoire de l'art pour étudier la composition, la dégradation et la préservation des œuvres d'art et archéologiques.



Figure 1. (À gauche) Photo de la garance des teinturiers (Rubia tinctorum L.*). (À droite) Structure moléculaire de l'anthraquinone et de l'alizarine, un des principaux composants de la garance.*

La garance (Figure 1) est une plante tinctoriale utilisée très tôt depuis l'Antiquité pour produire une large gamme de rouges allant du rose pâle au rouge profond. Les textes anciens témoignent de son utilisation dans les textiles, tapis, uniformes militaires, manuscrits et objets décoratifs. Les recettes écrites sont rares car parfois jalousement gardées secrètes pour leur valeur commerciale. Les principaux colorants extraits de la racine sont des anthraquinones naturelles, parmi lesquels l'alizarine et la purpurine. Ces molécules sont responsables des propriétés colorantes de la garance. Leur structure chimique complexe, riche en doubles liaisons conjuguées, favorise l'absorption dans le visible, mais les rend également vulnérables à la photodégradation, phénomène au cœur de cette étude. Présents dans de nombreuses œuvres historiques, les colorants à base de garance posent aujourd'hui des défis en matière de conservation, en raison de leur fugacité à la lumière.

Les racines de garance renferment un mélange complexe d'une trentaine à une soixantaine de colorants selon l'espèce et l'âge de la plante. Certaines de ses molécules colorantes sont difficiles à identifier, sont instables, photosensibles, d'autres sont difficiles à synthétiser ou non disponibles commercialement. Par conséquent, des références sont manquantes pour leur identification dans les œuvres d'art par spectroscopie. Cependant, pour comprendre comment ces colorants se dégradent sous l'effet de la lumière, il est nécessaire de faire une identification précise des molécules présentes et construire une base de données de leurs signatures spectroscopiques. Ceci est une des clefs pour mieux comprendre leur mécanisme de photodégradation. En effet, chaque composé interagit différemment avec la lumière, et la présence simultanée de plusieurs substances peut influencer leur comportement individuel.

Ainsi, l'objectif de cette thèse est double : d'une part, développer des outils prédictifs de couleur et de spectre, dont les résultats sont utiles pour la conservation du patrimoine, et compléter la base de données manquante pour aider à l'identification de molécules ; et d'autre part, comprendre les mécanismes de photodégradation des colorants issus de la garance.

Pour atteindre les objectifs de la thèse, en nous basant sur la chimie computationnelle, nous avons choisi dans cette thèse de combiner à la fois une approche théorique et expérimentale. Grâce aux ressources de calcul haute performance et au développement de la chimie quantique, il est désormais possible d'accéder aux propriétés chimiques de molécules difficiles à obtenir expérimentalement. Cette approche est très utilisée dans de nombreux domaines, mais très peu dans le domaine du patrimoine.

Nous avons commencé par chercher et optimiser un protocole computationnel adapté aux colorants de la garance. En nous basant sur la littérature et les benchmarks, notamment de Denis Jacquemin, Carlo Adamo et Ilaria Ciofini parmi de nombreux autres, nous avons trouvé que les méthodes basées sur la Théorie de la Fonctionnelle de la Densité (DFT en anglais) et son analogue dépendant du temps (TD-DFT) conviennent à notre système : les anthraquinones. Étant donné la complexité et l'hétérogénéité des mélanges de pigments présents dans les œuvres, nous avons choisi une modélisation en solution. Ce choix permet de simplifier le système, en étudiant la photodégradation du colorant seul dans un solvant, tout en assurant un meilleur contrôle des paramètres expérimentaux tels que le pH, la concentration, la température ou la présence d'oxygène. Il facilite également l'analyse par des techniques telles que la spectroscopie UV-VIS, la Résonance Paramagnétique Électronique (RPE) ou la Chromatographie Liquide Haute Performance couplée à l'UV et au spectre de masse (HPLC-UV-MS). Ce modèle en solution permet une étude de la photodégradation en milieu homogène.

Le protocole computationnel optimisé sert à prédire les spectres d'absorbance et les coordonnées colorimétriques de l'alizarine, l'un des principaux colorants de la garance, et de son analogue, l'alizarine red S plus soluble dans l'eau. Ce protocole a permis de prédire rapidement et efficacement la couleur de l'alizarine à différents pH et en prenant en compte l'effet du solvant. Il comprend l'analyse des conformères, et des formes déprotonées pour chaque pH via un calcul de l'énergie moléculaire. L'effet du solvant (en particulier l'eau) a été étudié à l'aide de différents modèles théoriques, afin d'évaluer son influence sur les propriétés spectrales des colorants. Plusieurs approches ont été considérées : un modèle implicite, dans lequel le solvant est représenté comme un continuum diélectrique, et plusieurs modèles explicites, qui prennent en compte plusieurs configurations de molécules d'eau interagissant directement avec la molécule d'intérêt. Les résultats ont montré que le modèle de solvant optimal dépend de la structure de la molécule (protonée de couleur jaune à $pH < 5.5$, mono-déprotonée rouge à $5.5 < pH < 10$ et di-déprotonée violet à $pH > 10$) et de son interaction avec le solvant. Si la molécule interagit moins avec l'eau, comme la forme protonée (pH acide) de l'alizarine, un modèle implicite suffit et permet un gain de temps de calcul. Pour les formes déprotonées qui forment des liaisons d'hydrogène avec l'eau, le modèle implicite couplé à 3 molécules d'eau explicites/groupe hydroxy est optimal tout en gardant un temps de calcul raisonnable. Le modèle le plus précis, applicable sur une large gamme de pH, est un

modèle hybride multi-échelle de type Quantum Mechanics Tight Binding QM/xTB avec 100 molécules d'eau explicites, couplé à une dynamique quantique, mais le temps de calcul est long. En effet, à la fin de la dynamique, une centaine de snapshots ont été pris pour calculer l'énergie verticale selon le schéma QM/xTB en prenant l'alizarine et 3 molécules d'eau par -OH pour la partie QM et le reste des solvants avec xTB. Ces calculs ont permis d'obtenir de bons accords avec les résultats expérimentaux. Un des paramètres importants dans la modélisation des spectres d'absorption concerne la prise en compte des effets vibrationnels, qui déterminent en partie la forme du spectre. Ces effets peuvent être modélisés selon plusieurs approches. L'approche Vertical Gradient (VG), plus rapide en termes de temps de calcul, repose uniquement sur les propriétés de l'état fondamental telles que l'énergie, les dipôles de transition et le hessien. En revanche, l'approche Adiabatic Hessian (AH) requiert le calcul de ces mêmes propriétés à la fois pour l'état fondamental et l'état excité. Bien qu'elle soit plus coûteuse en ressources computationnelles, l'approche AH fournit des résultats plus proches des données expérimentales. Une fois le spectre calculé, il peut ensuite être converti en coordonnées de couleur CIELAB, permettant une comparaison directe avec les couleurs expérimentales afin de valider le protocole. L'approche AH s'est révélée plus proche des spectres expérimentaux avec un écart de moins de 0,2 eV de la littérature. L'écart colorimétrique ΔE_{2000} est inférieur à 13 pour l'alizarine (moins soluble dans l'eau) et un $\Delta E_{2000} < 8$ pour l'alizarine red S, indiquant une bonne concordance avec la couleur réelle. Ces résultats ont permis de valider la robustesse du protocole computationnel, non seulement pour la prédiction des spectres d'absorbance et des couleurs, mais également pour la modélisation précise des déplacements chimiques en RMN ^{13}C et ^{1}H, pouvant contribuer ainsi à l'identification structurale des pigments. Ceci a fait l'objet d'un premier article publié intitulé "Rapid and complete prediction of alizarin in solution by combining experimental data with computational methods."

Par la suite, en nous appuyant sur ce protocole, nous avons pu constituer une base de données regroupant les caractéristiques colorimétriques et les signatures spectroscopiques de 31 colorants présents dans différentes espèces de garance, à différents pH. En effet, la garance se décline en plusieurs espèces, dont 6 principales : *Rubia peregrina* L., *Rubia tinctorum* L, *Rubia cordifolia* L. *Akane Nakai*, *Rubia sikkimensis Kurz*, *Rubia yunnanensis Diels.* Ces espèces contiennent des molécules similaires, mais à des proportions différentes. On en décompte une trentaine dans la littérature. Ils peuvent être sous forme glycosidique (avec un groupe glucose) ou sous leur forme aglycone (où le groupe de glucose est hydrolysé et remplacé par un groupe hydroxy). Le protocole établi précédemment a été validé sur une dizaine de colorants accessibles dans la littérature et un écart entre les valeurs calculées et expérimentales était toujours en dessous de 0,2 eV, la barre d'erreur théorique dans la littérature. Ensuite, en nous basant sur une liste de 31 colorants identifiés par Richard S. Blackburn et en utilisant le protocole validé du premier article, nous avons établi une base de données incluant les spectres d'absorbance, leurs λ_{max}, la couleur, et cela dans les deux conditions les plus souvent utilisées lors de l'extraction ou de l'utilisation de la garance : solution aqueuse acide et neutre. Ce protocole a permis d'accéder à des informations qui sont difficiles à obtenir expérimentalement pour certaines molécules comme la munjistine ou la pseudopurpurine en raison de leur instabilité chimique, de la complexité de leur synthèse

ou de leur indisponibilité commerciale. Cette base de données a permis de mettre en évidence les relations entre structure et couleur et spectres : l'effet du pH, l'effet de l'hydrolyse glycosidique. À pH acide, les colorants de la garance sont majoritairement jaunes ou transparents dans le visible. À pH neutre, certaines deviennent plus rougeâtres. L'hydrolyse glycosidique a un effet légèrement bathochrome. Cela peut expliquer l'étape de macération de la garance, qui permet aux enzymes présentes d'hydrolyser les molécules pour les rendre plus rouges. Ces résultats ont fait l'objet d'un deuxième article publié, intitulé « In-silico color prediction process for natural dyes in Madder ».

La garance étant en réalité un mélange complexe de plusieurs colorants, l'étape suivante de notre projet a consisté à modéliser les propriétés de ce mélange. Nous avons ainsi réussi à prédire les propriétés spectrales de combinaisons de colorants présents en proportions variables dans un extrait de garance de l'espèce *Rubia tinctorum* L., la plus répandue en France et la plus souvent utilisée en teinture et peinture. Ce travail a abouti à la publication d'une troisième communication intitulée « Progress in modeling complex dye mixtures: a case study for cultural heritage ». Pour cela, nous avons fait une analyse HPLC-UV-MS afin de déterminer les molécules présentes ainsi que leur proportion relative. Nous avons ensuite appliqué le protocole précédemment établi afin de prédire les spectres de chacune des molécules étudiées. Plusieurs approches ont été testées pour prendre en compte les effets vibrationnels : AH (sans sucre), VG et AHAS (Adiabatic Hessian After Step). Certaines molécules contenant un groupe glucosidique, comme la lucidine primeveroside, ne permettaient pas une convergence satisfaisante avec l'approche AH. Pour contourner ce problème, nous avons substitué le groupe glucoside par un groupement méthoxy, car il n'intervient pas dans des transitions électroniques, ce qui a permis d'assurer la convergence des calculs. Par ailleurs, les méthodes AHAS et VG ont également été testées en conservant les groupes glucosidiques. Au terme de ces comparaisons, les approches les plus performantes se sont révélées être AH (sans sucre) et AHAS, la première donnant les résultats les plus proches des données expérimentales avec un écart $\Delta E_{2000} = 3{,}7$, contre $\Delta E_{2000} = 4{,}5$ pour AHAS et $\Delta E_{2000} = 10{,}4$ pour VG. Cette méthode ouvre la voie à de nouvelles perspectives prometteuses pour la prédiction des couleurs et des spectres de mélanges moléculaires complexes. Sa capacité à prévoir avec précision les spectres UV-VIS et les couleurs des colorants naturels en fait un outil particulièrement précieux, notamment dans le domaine du patrimoine culturel, où les analyses non invasives sont essentielles sur des objets complexes et hétérogènes. En prédisant les couleurs de différentes molécules, cette approche computationnelle permet un meilleur contrôle des palettes chromatiques, une réduction du gaspillage de colorants et contribue à l'avancement de la recherche sur les colorants naturels. Elle pourrait également orienter de futures études sur une variété de systèmes chimiques, incluant les mélanges de colorants, les tautomères, les isomères, les produits de dégradation ainsi que d'autres mélanges complexes.

Une fois les propriétés spectroscopiques déterminées et la nature des transitions électroniques élucidées, cela nous a menés à l'étude de la photodégradation. Pour cela, nous avons choisi l'alizarine comme modèle, dont nous suivons la dégradation de manière accélérée avec des lampes LED ou au xénon à l'aide de mesures d'absorbance UV-VIS. Bien que l'alizarine soit couramment utilisée sur des supports solides, son emploi en milieu aqueux, notamment dans les aquarelles, justifie une étude en solution, où la

photodécoloration est plus marquée. Travailler en solution permet également un contrôle précis du pH, de l'oxygène et des effets de solvant, tout en facilitant la détection d'espèces transitoires grâce à des techniques sensibles comme l'EPR et la HPLC-UV-MS. Nous avons d'abord étudié l'influence de différents solvants sur la photodégradation. Nous avons observé que l'alizarine ne se dégrade quasiment pas ou faiblement dans le glycérol et le DMSO, alors que dans le DMF, au contraire, la cinétique de photodégradation est accélérée. Cela rend compte de la nature radicalaire du mécanisme de photodégradation en solution. Dans le domaine du patrimoine, l'eau est souvent utilisée, nous avons donc choisi de nous focaliser sur ce solvant ainsi que sur l'alizarine, un des principaux colorants de la garance, comme point de départ. Nous avons ensuite étudié la cinétique de photodégradation en solution aqueuse à différents pH et avec différentes lampes (chambre xénon dont le spectre est proche de celui du solaire, LED UVA et LED visible). L'alizarine n'étant pas totalement soluble dans l'eau, nous avons fait cette étude sur son analogue, l'alizarine red S, plus soluble. Cette étude a permis de montrer l'influence de la lampe et du pH sur la photodégradation. En particulier, à pH basique, la photodégradation est plus rapide. L'alizarine étant soluble à pH basique, nous avons aussi étudié sa photodégradation sous différentes lampes.

Un aspect clé de notre étude est l'analyse de la forme basique (di-déprotonée) de l'alizarine, encore peu explorée, la littérature n'étant principalement portée que sur la forme rouge et en laque solide. Grâce à l'analyse par résonance paramagnétique électronique (RPE), nous avons identifié des radicaux centrés sur l'oxygène. Par ailleurs, des analyses complémentaires par HPLC-UV-MS ont permis d'identifier plusieurs produits de photodégradation sous irradiation UVA, à différents temps d'exposition. Certains de ces produits n'avaient pas encore été rapportés dans la littérature, comme l'acide 3-hydroxy-1,4-dioxo-1,4-dihydronaphthalene-2-carboxylique. D'autres auteurs ont rapporté la présence d'acide phtalique parmi les photoproduits pour la forme neutre de l'alizarine. Plusieurs mécanismes ont été proposés en se basant sur les résultats RPE et les produits dégradés identifiés. Ces mécanismes passent par un intermédiaire endoperoxyde, contenant une liaison C-O-O-C formant un pont sur le cycle. Les liaisons O-O viennent d'un dioxygène. Ils pourraient se rompre en présence de lumière ou de chaleur pour ensuite conduire à des biradicaux, qui sont les précurseurs des produits de photodégradation. Ce travail a montré l'influence du pH, du solvant, de la lumière et de l'oxygène sur la photostabilité de l'alizarine et de son analogue, l'alizarine red S, des connaissances utiles pour des applications industrielles telles que la teinture textile ou l'extraction en bain alcalin. Il a fait l'objet d'un article soumis, intitulé « When light betrays color: the photodegradation of alizarin, a historical dye, in basic media ».

Une suite logique consisterait à approfondir la compréhension des mécanismes de photodégradation en modélisant les états excités et les voies réactionnelles à l'aide de méthodes de chimie quantique avancées. Des outils tels que l'analyse de la coordonnée de réaction intrinsèque (IRC), la méthode du Nudged Elastic Band (NEB), sont en cours pour explorer les paysages énergétiques, localiser les éventuelles intersections coniques. De telles études permettraient de mieux cerner la nature mécanistique de la dégradation induite par la lumière : quelles sont les barrières énergétiques ? Des intermédiaires réactifs sont-ils formés ? Existe-t-il un chemin réactionnel privilégié ? Il sera l'objet d'un cinquième article en cours de préparation.

Cette approche combinée entre expérience et théorie, permet de mieux comprendre l'origine et l'évolution des propriétés colorimétriques de la garance. Le protocole computationnel optimisé a permis de prédire les spectres d'absorbance et les couleurs d'une trentaine de colorants présents dans différentes espèces de garance, y compris celles inaccessibles expérimentalement. Ce protocole a conduit à l'établissement d'une base de données facilitant leur identification et la caractérisation de leur empreinte spectrale. Le protocole permet également d'évaluer l'influence de paramètres tels que le pH, le solvant, l'hydrolyse du groupe glucose ou encore la présence de mélanges sur les propriétés spectrales et la couleur perçue, contribuant à une meilleure compréhension des relations structures-propriétés optiques. Enfin, l'étude sur la photodégradation de l'alizarine a révélé un mécanisme impliquant des espèces radicalaires, notamment un oxygène réactif en milieu basique. Ce protocole présente un fort potentiel d'extension à l'étude d'autres colorants. En parallèle, cette étude contribue à la valorisation de l'utilisation des plantes tinctoriales.

Dans la continuité de ce travail, plusieurs pistes s'ouvrent. La photostabilité d'autres colorants dérivés de la garance, tels que la purpurine, ou de composés glycosylés comme l'acide rubérythrique et la lucidine primevéroside, pourrait être étudiée sous différentes conditions environnementales. En combinant la cinétique de photodégradation avec la RPE et la HPLC-UV-MS, il serait possible d'identifier les structures intermédiaires et photoproduits, élargissant ainsi la compréhension de la stabilité et de la vulnérabilité des colorants dans le contexte du patrimoine culturel. L'approche de modélisation développée dans ce travail pour les mélanges de colorants pourrait également être utilisée en sens inverse : en comparant les spectres expérimentaux aux spectres calculés théoriquement, il serait envisageable de déconvoluer des mélanges complexes de colorants et d'estimer les concentrations relatives de chaque espèce. Appliquer cette méthode à des mélanges contenant plusieurs colorants et leurs produits de photodégradation fournirait des informations précieuses sur l'évolution chimique des colorants, tout en apportant un éclairage sur les mécanismes de décoloration. Un axe de recherche majeur et essentiel pour le domaine du patrimoine culturel concerne l'étude de la structure et photodegradation de pigments à l'état solide, en particulier les pigments laqués à base de garance, qui sont des assemblages complexes formés par la coordination de colorants avec des cations métalliques (appelés mordants). D'un point de vue interdisciplinaire, ces protocoles computationnels et expérimentaux pourraient être appliqués à d'autres domaines utilisant des colorants d'origine biologique, tels que les cosmétiques, le textile, l'agroalimentaire ou la pharmacie. Il constitue un cadre propice au développement durable, à la réduction des tests expérimentaux, et à la formulation raisonnée de colorants stables et respectueux de l'environnement.

Enfin, cette recherche contribue à la redécouverte et à la valorisation des plantes tinctoriales. À une époque où les savoir-faire ancestraux suscitent un regain d'intérêt, de tels outils peuvent orienter le choix des espèces et des procédés de transformation en assurant à la fois qualité esthétique, stabilité chimique, et respect de leur héritage culturel.

CHAPTER 1

GENERAL INTRODUCTION

IN SEARCH OF LOST COLOURS:

INVESTIGATION OF THE PHOTODEGRADATION OF NATURAL MADDER DYES

CHAPTER 1 | GENERAL INTRODUCTION
IN SEARCH OF LOST COLOURS
PHOTODEGRADATION OF NATURAL MADDER DYES

When we look around us, colours are everywhere in our lives. Today, we can admire with the naked eye the beauty of a palette with hundreds of colours and infinite shades. This rich diversity of today's palette seems given, but in reality, these colours are the result of a long journey of experimentation, discoveries, ancestral know-how, and … sometimes accidents, before becoming part of the palette we now know. The primary palette of colours was composed of five main colours: red, yellow, brown, black, and white. They were used around 40,000 years ago, when humans randomly combined materials available in their daily environment, such as earth, burnt charcoal, bones, or animal fat. Since then, the quest for colour and pigments has been marked by continuous discoveries, driven by scientific progress and contributions from diverse civilisations, far beyond the basic initial mixtures. In fact, colours have even become a part of human history. They tell the story of a scientific and cultural adventure, where each shade carries a trace of a period, an invention, a symbol, or a country.

Colour is intrinsically linked to light. Without light, colour cannot be perceived by the human eye. In 1704, Sir Isaac Newton published a treatise, "Opticks", about the nature of light and colours. He discovered that white light is a combination of different wavelengths, each corresponding to a specific colour, and that the objects' colour was determined by the proportion and colour of light reflected. This finding helped us understand that the colours we perceive result from the interaction of light with matter and the way our visual system interprets these wavelengths. These findings laid the theoretical foundation for a more structured representation of colours through absorbance spectra or colour coordinate systems like the CIELAB, which is used in this thesis.

Yet, many research studies have proved that light, particularly sunlight, causes chemical changes in colourants, leading to their alteration or fading over time, a process known as photodegradation. This affects both natural and synthetic dyes. This problem has been widely discussed in artworks, particularly those by renowned artists such as Vincent van Gogh, and Pierre-Auguste Renoir, among others. They found, for example, that the painting "*Sunflowers*" by V. van Gogh has faded upon exposure to light. Researchers have attributed this to the reduction of chromium (Cr^{IV} to Cr^{III}), present in the yellow chrome pigment. The loss of red colour in "*Madame Leon Clapisson*" of Pierre-Auguste Renoir is due to the photodegradation of carminic acid. The exact mechanism is not well understood, but it is believed to proceed through a radical pathway in the presence of oxygen. The Bayeux Tapestry, conserved in Normandy, France, is 70m long and has been found to be dyed with madder, among others. It has been found that areas exposed to sunlight have shown signs of photodegradation. These alterations significantly impact both the aesthetic and economic value of artworks. Therefore, studying photodegradation is a scientific and cultural challenge that has long engaged chemists, physicists, and

conservators. In order to preserve cultural artifacts, we need to have a firm understanding of the materials, dyes, and pigments used, and this depends on their correct identification.

This thesis fits into that context, focusing on the photodegradation of madder (*Rubia tinctorum* L.), a renowned plant-based dye that has been used since antiquity for its vibrant red colours. Widely used in the 18th century, madder played a significant economic role before declining with the advent of synthetic dyes. Madder dyes belong to the anthraquinone family. Its roots contain a complex mixture of compounds, ranging from around thirty to sixty different dyes in varying proportions depending on the plant species and age, some of which are unstable and photosensitive, making their analysis difficult. This is the case, for example, with munjistin and pseudopurpurin. During the dyeing process or extraction for analytical purposes, these compounds can be easily decarboxylated, converting into xanthopurpurin and purpurin, respectively. Furthermore, some compounds are not commercially available or are too costly to synthesise, resulting in a lack of reference standards for their identification in artworks through spectroscopy. This is the case of ruberythric acid, the precursor of alizarin, one of the main compounds in madder. Therefore, to understand and study their photodegradation, it is essential to precisely identify the molecules present and their spectroscopic signatures. Indeed, each compound interacts differently with light, and the simultaneous presence of multiple substances can influence their behaviour. That is why, for decades, the study of photodegradation has drawn the attention of many research laboratories and conservators, both in the heritage field and in industry. Several analytical techniques can be used to identify and characterise dyes to explain this process. These can be grouped into two main categories: invasive techniques, which require sampling, and non-invasive techniques, which allow analysis without extraction. This latter method is preferred in cultural heritage science due to its non-destructive or minimally destructive nature. Among them, spectroscopy plays a key role. In this thesis, we will focus on the absorbance spectra, as they provide information on both colour and electronic transitions following light absorption. The precise mechanism of photodegradation remains incomplete, particularly for natural dyes, and their accurate identification still encounters challenges.

To contribute to this ongoing effort, we used a combined computational and experimental approach. The objective of this thesis is twofold: first, to develop predictive tools for colour and spectral properties to complete the missing database, and second, to understand the mechanisms of photodegradation.

Computational chemistry based on quantum mechanics theory, although applied in many other fields in material science, remains underused in the cultural heritage field. This could be due to a lack of a reference dataset to validate computational predictions in heritage-specific contexts and interdisciplinary barriers. Meanwhile, thanks to high-performance computing resources and advances in quantum chemistry, it is now possible to access the chemical properties of molecules that are difficult to obtain experimentally. Based on this, we developed a computational protocol specifically adapted to madder dyes.

This thesis is organised as follows:

-It begins with a state-of-the-art review, including a brief history of red colour, an overview presentation of madder, and before focusing specifically on the photodegradation of madder.

-The following sections detail the materials and methods employed, covering both experimental and theoretical approaches. An overview of the main principles of quantum chemistry and its application in chemistry, as well as the computational protocol, is presented.

-The results build on the previous section and are presented through five articles: three published, one under review, and one currently under preparation.

-Finally, it concludes with a summary of findings and perspectives for future research.

The articles are presented as follows: guided by the thesis goal presented above, we first established a computational protocol specifically adapted to madder dyes, guided by literature benchmarks (Article 1 published). Using this validated computational protocol, we will then compute colourimetric properties and build a complete database including spectroscopic signatures for 31 colourants identified in 6 different madder species, at different protonation states. It can help access spectra that are difficult to access experimentally and give theoretical reference standards to identify those compounds (Article 2 published). However, in reality, like many natural dyes, madder is a complex mixture of multiple colourants. Therefore, the next step focused on modelling the spectral properties of this mixture, using a madder extract of the species *Rubia tinctorum* L. It can be applied to predict the properties of other complex mixtures: mixtures of different colourants, tautomers, photodegraded products, impurities, etc. (Article 3 published). With the spectroscopic properties of the dyes elucidated, we finally turned to study photodegradation mechanisms using alizarin solution as a model, focusing on the basic conditions (Article 4 submitted). Those mechanisms could be confirmed by computational chemistry to study reactivity, reaction pathways, and energy barriers. (Article 5 under preparation).

CHAPTER 2

CHAPTER 2 | STATE OF THE ART 27

2.1. A history dyed in madder, a natural red dye 27
2.1.1. The colour red: a hue with deep social and symbolic meaning 27
2.1.2. Madder: a natural source of red dye 27
2.1.3. Madder as a dye or pigment/lake 29

2.2. Madder species 31
2.2.1. Identification technique of madder molecules in artworks 32
2.2.2. Composition of madder dyes 33

2.3. Overview state of the art on photodegradation of dyes 35
2.3.1. Reciprocity principle and accelerated photodegradation 37
2.3.2. Internal factors influencing photodegradation 38
2.3.3. External factors influencing photodegradation 38

2.4. Photodegradation of madder 41
2.4.1. Photodegradation of madder in artworks 41
2.4.2. Photodegraded products and proposed mechanism 42

2.5. Difficulty in studying madder dyes 43

2.6. Computational spectroscopy: state of the art 44
2.6.1. Computational chemistry in the cultural heritage field and dyes 44
2.6.2. Role of functional and basis set for TD-DFT vertical energy 45
2.6.3. Absorption spectra and colour: vibronic coupling 46
2.6.4. Predicting the influence of the environment 48

This chapter reviews the key literature and scientific background on madder dyes, focusing on their historical significance, chemical composition, and degradation processes. It begins with an overview of the cultural role of red dyes and madder's use as a natural pigment, followed by a discussion of madder species and identification techniques.

Photodegradation mechanisms, both general and madder-specific, are then explored, alongside the challenges of studying these complex materials. The chapter concludes with a brief survey of computational approaches used to model dye behaviour and predict spectroscopic properties in heritage science.

CHAPTER 2 | STATE OF THE ART

2.1. A history dyed in madder, a natural red dye

2.1.1. The colour red: a hue with deep social and symbolic meaning

Colours are important to humankind. They were vital for the survival of our species, as they helped humans interpret their environment, identifying danger and communicating. Among them, the colour red was one of the earliest colours mastered by humankind. Hence, it is one of the most interesting and studied colours. It is rich in many points of view: artistic, social, and symbolic.[1] In ancient times, humans produced the red colour from various natural sources available, such as plants, minerals, or insects: cochineal, kermes, brazilwood, ochre, lac dye, and madder.[2,3] Then later, humans learnt to synthesise colours.

Colours have always played a significant role in human history. In the field of cultural heritage in particular, each colour often carries social and symbolic meanings. These aspects have been extensively studied by historians and curators, among them Michel Pastoureau, a leading expert on the subject.[1,4] In antiquity, for the Greeks and Romans, red symbolised power, wealth, and majesty. During the Middle Ages, it acquired strong religious connotations, evoking both the blood of Christ and the flames of hell. The symbolism of red is sometimes contradictory: it could represent love, glory, and beauty, but also pride, violence, and lust. In the 16th century, Protestant reformers condemned the use of red, viewing it as indecent and immoral, and closely associated it with the theatrical rituals of the Catholic Church. As a result, the use of red declined across Europe in both everyday life and cultural artefacts, a trend that continued into the modern era. However, following the French Revolution, red took on new ideological and political meanings. It became a symbol of progressive and revolutionary movements and was later adopted by left-wing political groups, a symbolism that continues to resonate to this day.

The colour red is therefore rich in symbolic meaning that depends on culture and historical period. Its presence in many artworks makes their conservation essential for interpreting the intentions, messages of the artists, and know-how.

Among the natural red dye sources, madder (*Rubia tinctorum* L.), also known as rose madder, dyer's madder, or "*Garance des teinturiers*" in French, has been one of the most used red dyes since ancient times because of its low cost, wide availability, and fascinating shades.[5] This natural dye has gained increasing interest nowadays due to the rising ecological concerns.

2.1.2. Madder: a natural source of red dye

Madder is a herbaceous plant that grows in many parts of the world, including Europe, Asia, and Africa, and is harvested for its red dyes.[5] Those are extracted from madder ground-up roots and could produce pink to red and purple shades depending on the preparation process. Although the dyes extracted from madder have multiple

applications, their main use was in dyeing fibers for textiles, carpets, or making lakes and pigments for paintings (Figure 2).[2,6] One example of dyeing with madder is the Bayeux Tapestry, in Normandy, France, and the tapestry "*La dame à la licorne*" in the Cluny museum, Paris.[7,8] It was also used to stain objects like ivory game pieces.[9] The dyeing practice with madder was first documented around 2000 BC in Mesopotamian cuneiform texts, with traces also discovered on Egyptian mummy wrappings.[10,11] Moreover, historical records from the time of Herodotus (5th century BCE) mention its use in dyeing the cloaks of Libyan women.[12] In addition to its use in textiles, madder was also employed medicinally for the treatment of kidney and bladder stones. [5,13] Additionally, in the 19th century, it was accidentally discovered that animals consuming madder plants had their bones stained red, a phenomenon physiologists later exploited to study bone growth and cellular functions.[14] Between 1600 and 1900, the strong demand for madder drove substantial investment in its cultivation as a profitable cash crop throughout Eurasia. Madder cultivation expanded to Alsace, France, where it became a significant part of the economy.[5] It was used to dye French soldiers' trousers, though the bright red colour proved impractical on the battlefield. The French term for madder, "*garance*", derives from the French word "*garantie*" (warranty), as its production was once government-regulated.[15]

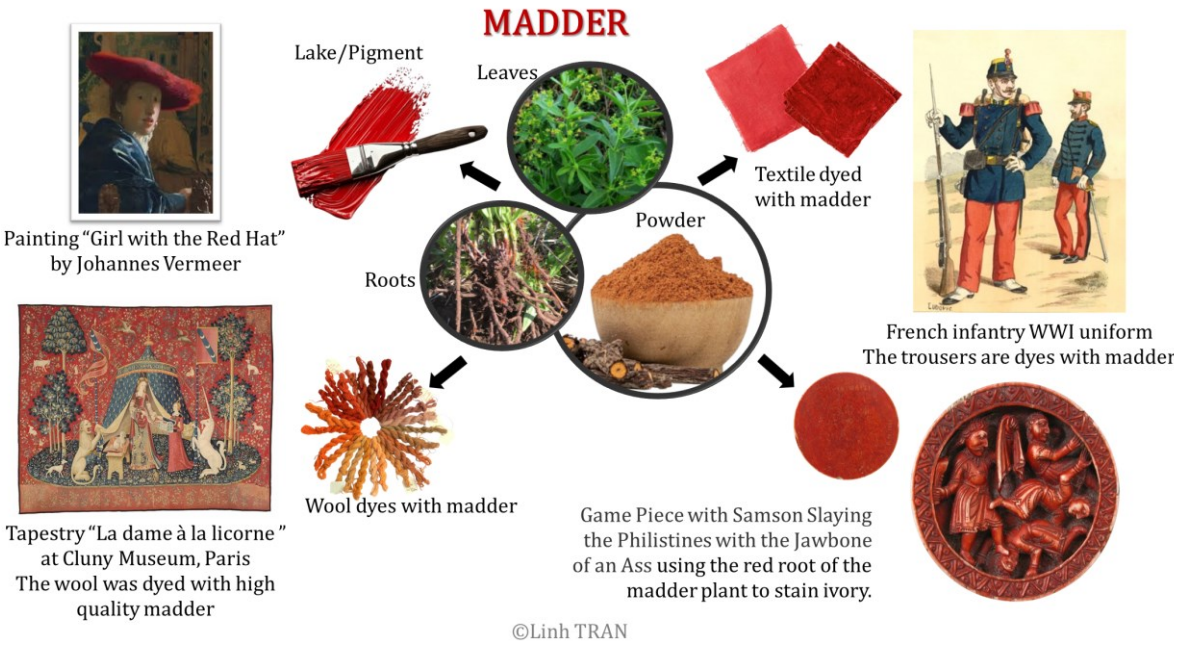


Figure 2. Illustration of main examples of madder dye applications: textiles, tapestry, paintings, and materials staining.

However, in 1868, Graebe and Liebermann discovered a method to synthesise alizarin, leading to a rapid decline in natural madder cultivation due to the lower cost of synthetic alternatives. [16] Despite its decline, growing concerns over the environmental impact of synthetic dyes have renewed interest in natural dyes. In recent years, madder has attracted increasing attention from researchers and industries seeking sustainable alternatives.[17–22] Other than its use in colour industry or for artistic purposes, madder has also applications in other fields such in other fields such as pharmaceuticals (inhibitory

effects towards cancer cells suggesting their application as a targeted drug for cancer), electrochemistry (madder dyes used as sensitizers in dye-sensitized solar cells), analytical chemistry (Alizarin, and its analogue alizarin red S are used as bone staining).[21,23–25]

2.1.3. Madder as a dye or pigment/lake

After the extraction of madder colourants from its roots, further processing can be done to produce madder pigment or lake used in paintings or to dye textiles. There are various types of colourants, distinguished primarily by their physical properties. A *dye* is a colourant that is soluble in its application medium, allowing it to penetrate the material being coloured. In contrast, a *pigment* is an insoluble substance, typically ground into a fine powder and combined with a binder to adhere to a surface (like Arabic gum, acrylic, water, oil, etc.).[26] A lake pigment is a type of pigment made by precipitating a soluble dye onto an inert substrate (called a mordant or binder, usually a metal salt).

a) Madder as a dye

As stated before, one of the most common applications of madder (*Rubia tinctorum* L.) is as a dye for textiles. Hence, the name *Rubia* means "red," while *tinctorum* is derived from the Latin word for "dyeing." Similarly, its French name is "*Garance des teinturiers*".



Figure 3. (left) Dyers soaking red cloth in a heated barrel, taken from the book Des Propriétés des Choses (Volume II) (1482).[27] (right) A madder-based dyeing recipe, taken from the book The Art of Dyeing Wool and Woolen Fabrics in Great and Small Dyeing by Jean Hellot, 1750 (p. 377).[28]

One of the first written French memoirs on dyeing, published in 1750: The Art of Dyeing Wool and Woolen Fabrics in Great and Small Dyeing by Jean Hellot (Figure 3).[28] In this book, before dyeing, the wool was treated with a mordant (alum and cream of tartar) to help fix the dye to the fibre. The mordant and dye were boiled in water for 2 hours. Madder root is finely ground and soaked in warm water to extract the dye. Then, the pre-mordanted wool is added to the warm madder bath. It is left to soak for 1h without boiling, and if necessary, the dye bath is boiled again at the end of the operation for 4- 5 min.

b) Madder as a pigment

When madder is used in a dye bath with alum as a mordant, any excess alum can react with the dye components, leading to the formation of a precipitate that settles as sediment at the bottom of the vessel. The dyers discovered that this insoluble red precipitate could be used for painting as a pigment. This led to the creation of madder lake. A Roman naturalist, Pliny the Elder, described this practice among the Tyrians (present-day Lebanon), who reportedly dried the precipitate and sold it at a high price to painters and artists.[29] Later, in the 19th century, despite the discovery of synthetic dye, natural madder lake was still manufactured at large scale by Winsor & Newton (W&N), a famous artist's colourman at that time. This company was founded in England by William Winsor, a chemist, and Henry Newton, a talented artist. Madder lake pigments are produced by precipitating madder dyes with salts such as alum ($AlK(SO_4)_2 \cdot 12H_2O$), resulting in aluminium (Al^{3+}) coordination complexes whose exact structures are still under debate.[6,30]

An example of a W&N recipe is shown below for Rose Madder, in a laboratory scale by T. Veiga *et al.,* and inspired by the book "*The artist's assistant: oil painting instruction manuals and handbooks in Britain 1800-1900*" by L. Carlyle:[30,31]

<u>Recipe for Madder Rose:</u>

1. Wash 5 g of madder powder with 100 mL of water (repeat 5 times);
2. Dissolve 8.57 g of alum in 30 mL of hot water (near boiling);
3. Pour the above solution into the madder roots;
4. Decant the supernatant liquor into 88 mL of distilled water;
5. Slowly add an aqueous solution of carbonate of ammonia (1.2 M) and stir occasionally;
6. Filter the precipitate.

<u>Recipe for madder lake:</u>

1. Let 5 g of madder powder rest in 100 mL of water for 1 h;
2. Decant the supernatant and let the madder rest in 100 mL of water for 30 min;
3. Decant the supernatant and wash the madder 3 times;
4. Dissolve 10.48 g of alum in 30 mL of hot water (80 °C);
5. Run through the above solution into the madder roots placed in a cotton filter bag and collect the red extract;
6. To the red extract, add immediately, one at a time, three solutions of sodium borate: 1.9 g in 10 mL, 0.24 g in 5 mL and 0.06 g in 2.5 mL;
7. Filter the precipitate.

<u>Recipe for Rose Madder:</u>

1. Prepare a solution of 0.5 mL of sulfuric acid in 4 mL of water;
2. Pour the solution into the madder roots after step 5 of the Madder Lake recipe.
3. After 30 days, add 16 mL of boiling water;
4. Add immediately, one at a time, three solutions of sodium borate: 0.5 g in 2.5 mL, 0.06 g in 1.25 mL, and 0.03 g in 1.25 mL;
5. Filter the precipitate.

Those recipes are summarised in the Figure 4 below.

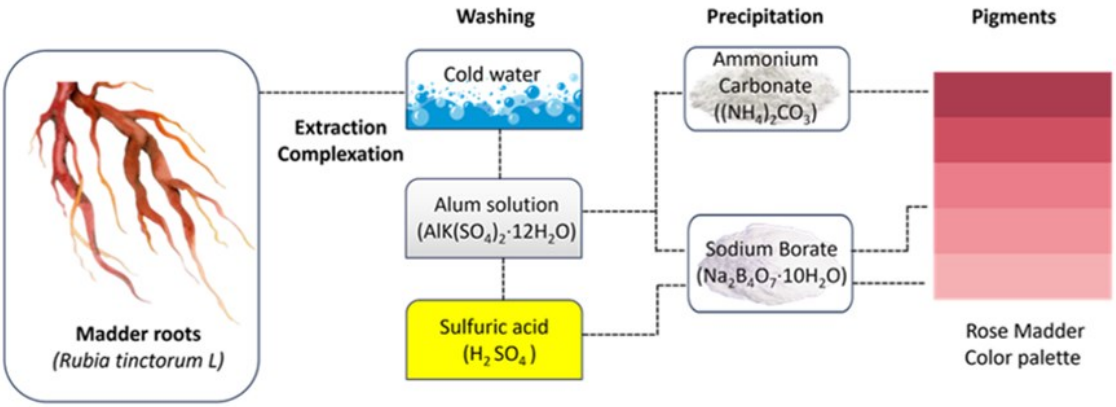


Figure 4. Scheme of the main steps of the 19th century manufacture of madder red lake pigments. [30,31]

2.2. Madder species

Madder belongs to the genus *Rubia* and the *Rubiaceae* family, which includes bedstraw and coffee plants (Figure 5). There are more than 70 species. [32] Madder can be found across many places from the Mediterranean, Africa, to Asia and America, with notable varieties including *Rubia tinctorum* L. *(Europe, Middle East), Rubia peregrina* L. *(Europe, North Africa), Rubia cordifolia* L. *(India), Rubia sikkimensis Kurz (India), Rubia akane Nakai (Asia), Rubia Yunnamensis* Diels *(Asia)* and *Rubia fructicosa* Aiton (Europe) (Figure 6).[5,10,15,25,33,34] Madder is an herbaceous plant with small leaves that intertwine as it grows. It is cultivated for its roots, which are long and red-coloured. The root's internal structure reveals varying distributions of red dye. The central region, or pith, is composed of large vascular bundle cells and contains only a small amount of dye.[20] In contrast, the intermediate layer holds the highest concentration, while the outer layer shows little to no presence of red dye. The stalks of the madder often sprawl along the ground, similar to bedstraw. Its full height is about 60–100 cm. The flowers are star-shaped yellow and bloom in the second or third year, typically in June.

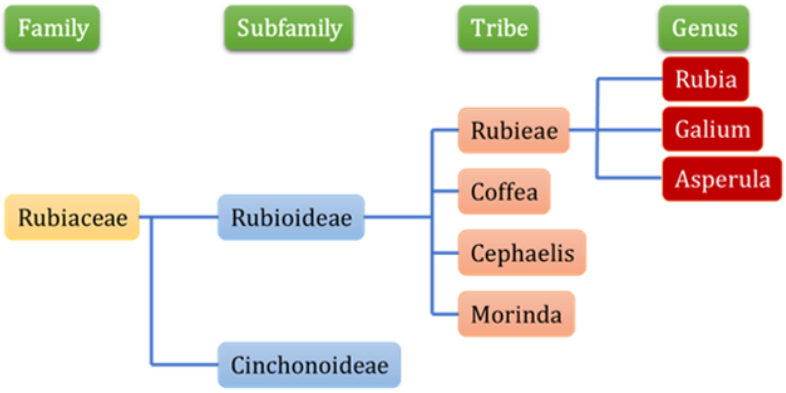


Figure 5. Botanical classification of the madder (Rubia) within the Rubiaceae family.[20]

The colourants extracted from the madder plant are grouped in the Colour Index as C.I. Natural red 8. For *Rubia tinctorum* L. and *Rubia peregrina* L. For dyes extracted from the specie *Rubia cordifolia* L. the colour index is C.I. Natural red 16.

Plant common name	Distribution	Plant species	Other synonyms	Color Index Name
Dyer's madder	Europe Middle East	*Rubia tinctorum* L.	R. acaliculata; R. iberica; R. sativa; R. sylvestris, Garance des teinturiers in French	C. I. Natural Red 8
Wild madder	Europe North Africa	*Rubia peregrina* L.	Galium peregrinum L. Garance voyageuse In French	C. I. Natural Red 8
Indian madder	India Asia	*Rubia cordifolia* L.	R. oncotricha; R. manjith; R. munjista; R. sylvatica, R.alata	C. I. Natural Red 16
Naga madder	India	*Rubia sikkimensis* Kurz	None	None
Japanese madder	Asia	*Rubia akane nakai*	R. argyi	None
Xiao hong can (no English name)	China Asia	*Rubia yunnanensis* Diels	R. ustulata	None
None	Spain, portugal	*Rubia fructicosa* Aiton	R. canariensis Poir.	None

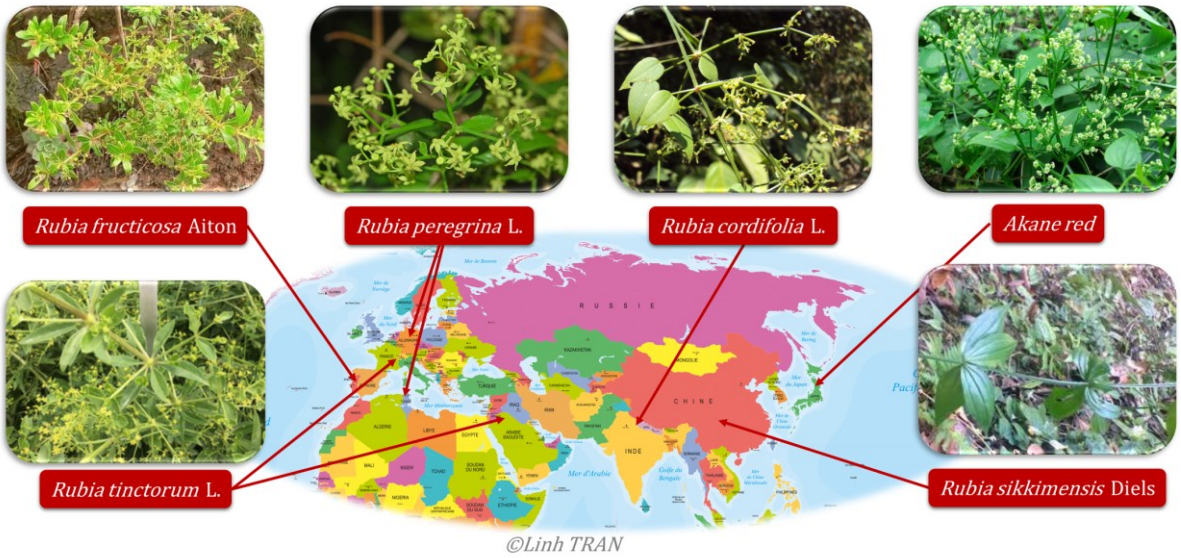


Figure 6. Different madder species around the world.[33]

In the 20th century, chemical analytical techniques have been developed such as chromatography and spectroscopic techniques that have helped identify the madder colouring compounds.

2.2.1. Identification technique of madder molecules in artworks

There are many methods to analyse dyes and identify them in historical artworks and materials such as textiles, paintings, etc. They can be classified into two categories: invasive and non-invasive.

a) Invasive techniques

Invasive methods require the extraction of a sample for further analysis, hence altering or destroying the artwork. For instance, we can cite High-Performance Liquid Chromatography with Mass Spectra (HPLC-MS), Gas Chromatography-Mass Spectrometry (GC-MS).[35–37] Those techniques are often coupled with UV-VIS detection, such as in HPLC-UV-MS or HPLC-DAD.[38] The identification of compounds often needs comparison with a standard reference, which is not always available.[34,35,39] Other techniques include Nuclear magnetic resonance (NMR), Time-of-Flight Secondary Ion

Mass Spectrometry (TOF-SIMS), Scanning Electron Microscopy with Energy Dispersive X-ray Spectroscopy (SEM-EDS), etc.[40–42]

Chromatographic techniques were widely used to identify molecules in the plants or in historic textiles, where a microfiber is extracted from the artwork for further analytical procedures. [33,43] They enable the separation and hence identification of components. However, the final composition depends on the extraction method. One of the most commonly used in literature is acidic extraction using 37% hydrochloric acid:methanol:water in a ration 2:1:1. [19] Other techniques have been developed to be done in more neutral or mild environment.[34,44]

b) Non-invasive

On the contrary, non-invasive methods preserve the sample's integrity without sample extraction. They include techniques like Raman or Surface-Enhanced Raman Spectroscopy (SERS), UV-Visible (UV-VIS) spectroscopy, Fourier-Transform InfraRed (FTIR), Fiber Optic Reflectance Spectroscopy (FORS), fluorescence, or hyperspectral imaging.[35,39,45–53] The latter combines reflectance spectroscopy and imaging, allowing the acquisition of multiple spectra at each point of the image. It can therefore spatially resolve chemical information on the dyes and pigments.

Non-invasive techniques are preferred in the field of cultural heritage, as they preserve the precious and irreplaceable historical artwork. However, the increasing development of more high-resolution spectroscopy techniques, such as hyperspectral imaging, results in hundreds of spectra that are difficult to analyse without a standard dye reference. Those references are sometimes difficult to obtain experimentally, especially in the case of madder, as discussed later in *subsection 2.5*, making the correct non-invasive identification of madder challenging.[39,54]

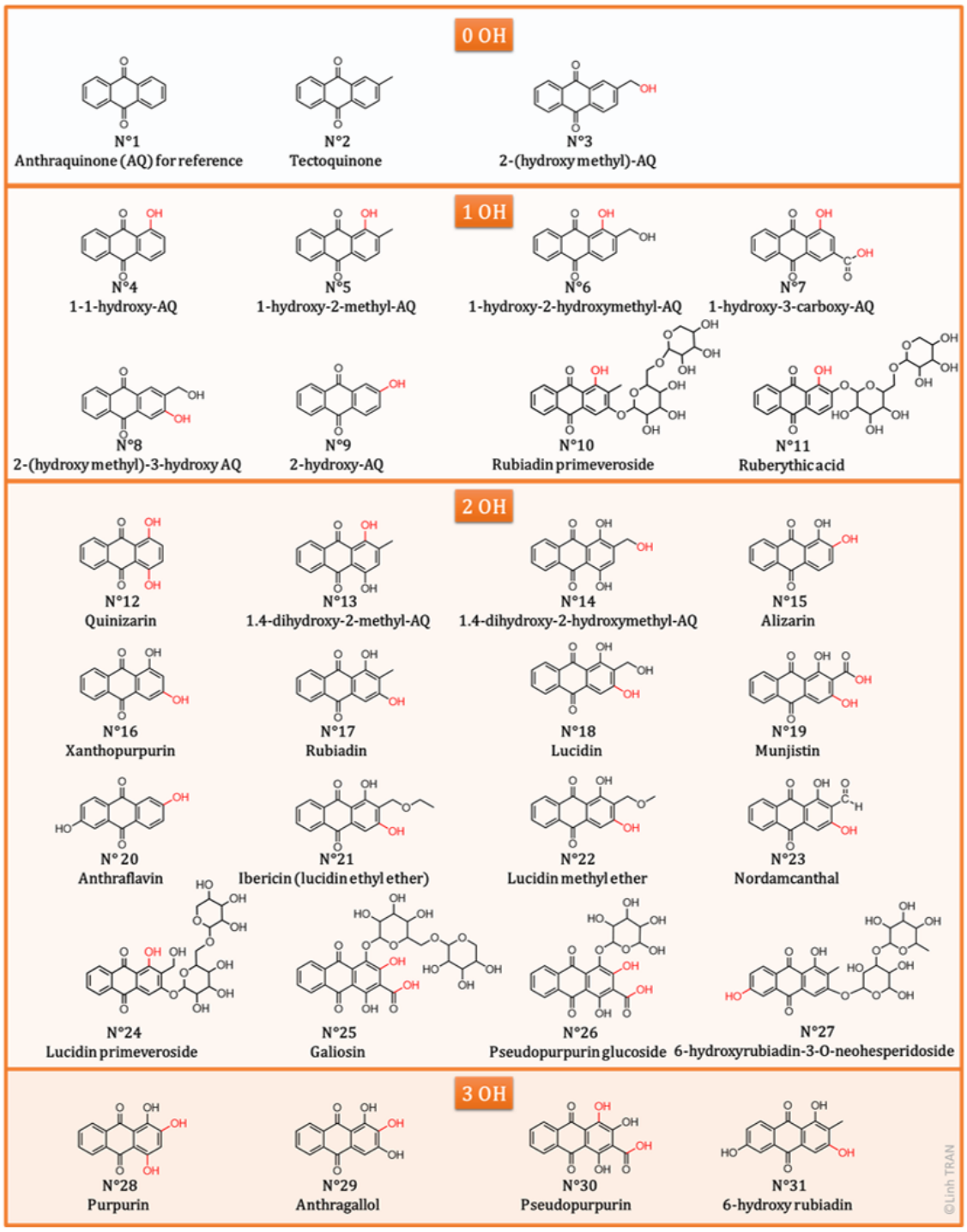
2.2.2. Composition of madder dyes

Unlike synthetic dyes, which are typically composed of single, well-defined molecules, natural dyes consist of a complex mixture of molecular compounds. Madder, as a natural dye, exemplifies this complexity, with its composition varying based on factors such as species, age, cultivation methods, soil conditions, temperature, pH and extraction method.[33,55,56] In the last century, with the development of previously presented analytical techniques, such as chromatography or spectroscopy, researchers have investigated the molecular composition and structure of madder. The madder colourants are from the anthraquinone family. They differ by the position and number of substituents on the anthraquinone moiety. Each species has a similar molecule but differs in its proportions.[19] While some studies have identified up to 68 different molecules, most findings report approximately 30.[33,34,55–58] Blackburn *et al.* have compiled and summarised the identified molecules for each madder species, as shown in Table 1. Their identification in artefacts is valuable, as it provides key insights into important information such as species, which then leads to ethnographic origins, the techniques used by painters or dyers, the geographical source of the dye, and historical trading routes. Madder dyes are present in two primary forms in the plant's roots: glycosidic compounds where a sugar

moiety is attached (Table 1, N°10, 11, 24 to 27), and aglycones, where the glycoside is replaced by a hydrogen atom (all other molecules in Table 1). Glycosidic compounds are more concentrated in younger madder plants and are more soluble in water. [33,44,52,59] As the plant matures or during the dye preparation process, glycosides are broken down by enzymes or through chemical extraction, resulting in the release of aglycons, which produce a redder hue and intensify the dye colour.[20]

Table 1 . Compounds detected in the roots of six main Rubia species studied in the literature. ++ indicates the main compound in the plants, + significant amount, tr. very low or trace amount. (Reference for each compound can be found in Richard S. Blackburn's article). [32,60,61] The table below shows the molecular structure of each compound.

N°	Species / Name	*R. tinctorum* L.	*R. cordifolia* L.	*R.peregrina* L.	*R. Sikkimensis* Diels	*R. Yunnanensis*	*R. akane Nakai*
1	Anthraquinone						
2	Tectoquinone	+	+				
3	2-(Hydroxymethyl)-AQ	+				+	
4	1-Hydroxy-AQ	+	+				
5	1-Hydroxy-2-methyl-AQ	+	+	+		+	+
6	1-Hydroxy-2-hydroxymethyl-AQ		+				
7	1-Hydroxy-3-carboxy-AQ		+				
8	2-(Hydroxymethyl)-3-hydroxy AQ					+	
9	2-Hydroxy-AQ	+					
10	Rubiadin primeveroside	+		+			
11	Ruberythric acid (alizarin primeveroside)	++	tr.				+/tr.
12	Quinizarin	+					
13	1,4-Dihydroxy-2-methyl-AQ		+				
14	1,4-Dihydroxy-2-hydroxymethyl-AQ		+			+	
15	Alizarin	+					
16	Xanthopurpurin	+/tr.	+	+	+	+	tr.
17	Rubiadin	+	tr.	+		+	
17	Rubianine	+					
18	Lucidin			+			
19	Munjistin	+	++	+	++	+	++
20	Anthraflavin	+					
21	Ibericin (lucidin ethyl ether)	tr.					
22	Lucidin methyl ether	+					
23	Nordamnacanthal	+/tr.		+			
24	Lucidin primeveroside	++	tr.	++	+	+	
25	Galiosin (pseudopurpurin primeveroside)	+		++			
26	Pseudopurpurin glucoside	+	+				+
27	6-Hydroxyrubiadin-3-O-neohesperidoside		+			++	++
28	Purpurin	+	+	++	+		+/tr.
29	Anthragallol	+					
30	Pseudopurpurin	+	++	++	++	+	++
30	Pseudopurpurin	+					
31	6-Hydroxyrubiadin		+			+	



2.3. Overview state of the art on photodegradation of dyes

Some dyes gradually lose or change their colour over time, particularly when exposed to light, a process known as photodegradation. Madder, like many natural dyes, is photosensitive and is prone to fading under light exposure. This presents challenges for the preservation of textiles, as well as manuscripts, statues, paintings, and other artworks containing it. This photodegradation is present for both organic and inorganic dyes.

Famous examples of photodegradation can be seen in Vincent van Gogh's paintings like "*Bedroom*", "*Sunflowers*", or "*Roses*". It can also be seen in Pierre-Auguste Renoir's painting named "*Madame Léon Clapisson*". As we can see in Figure 7, the colours have changed through time. The blue of the bedroom was initially violet. The brown of sunflowers was once yellow and the red in other paintings has completely disappeared.

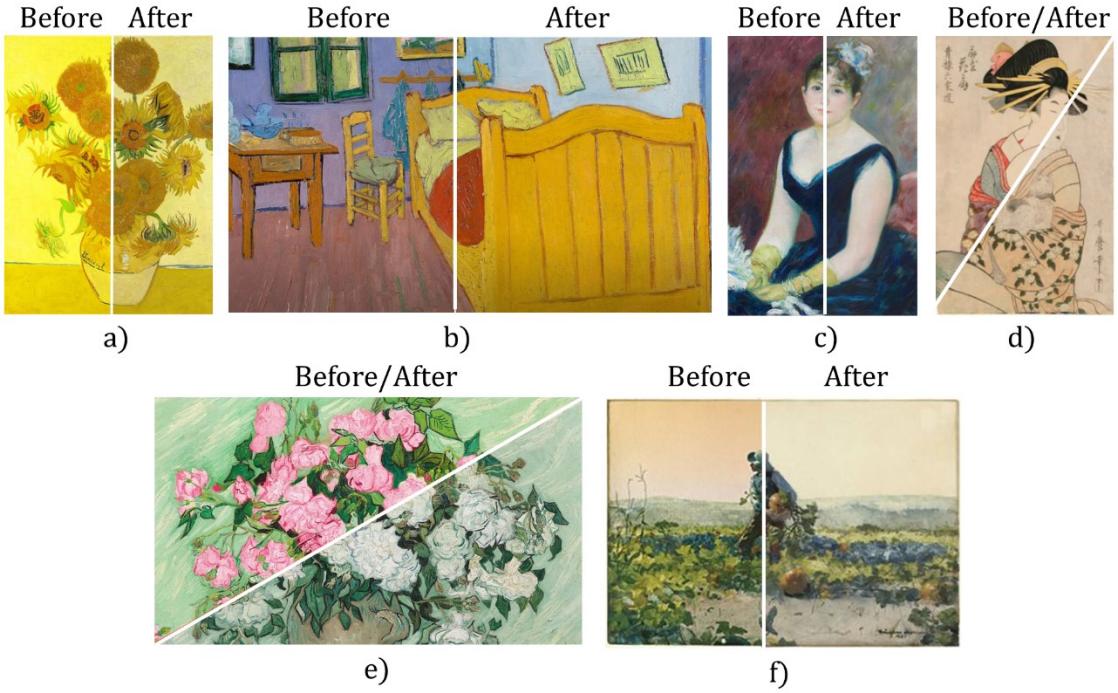


Figure 7. Illustration of some examples of paintings before and after photodegradation. The left side of each picture is a reconstruction. a) "Sunflowers" by V. van Gogh. b) "Bedroom" by V. van Gogh, c) "Madame Léon Clapisson" by P.A. Renoir, d) Japanesse woodblock by Utamaro, e) Roses by V. van Gogh, f) "To be a Farmers Boy" by Homer. The pictures are adapted from [62–65] (with permission from authors).[1]

Preserving the original colours is both a cultural and economic priority, engaging chemists and physicists in ongoing research efforts.[40] The studies in photodegradation help develop better conservation strategies and prevent unwanted photodegradation, but also provoke desired photodegradation of wastewater dyes from colouring industries, for example.[66] Studies about the photofading of madder is scarce. We will therefore present an overview of the state of the art about photodegradation studies for dyes and pigments in general before focusing on madder.

Dyes and pigments' photodegradation are influenced by diverse environmental factors.[67] I. Groeneveld and al. have reviewed the main parameters that affect the photodegradation of dyes and pigments in solution and substrate. [68] S. Khan *et al.* have studied the different photocatalytic dye degradation from textile wastewater and the parameters affecting it.[66]

[1] Hendriks, E., Brokerhof, A. W., & van den Meiracker, K. (2017). Valuing Van Gogh's Colours: From Past to Future. In J. Bridgland (Ed.), ICOM-CC : 18th triennial conference, Copenhagen, 4-8 September 2017

Thissection provide a brief overview of the current state of knowledge on this topic, before focusing on the photodegradation of madder.

Colourants can interact with light in different ways. When a molecule absorbs a photon, it transitions to an excited state and must release the excess energy. This can occur through different pathways to bring the molecule back to its ground state (Figure 8) through: [69]

-non-radiative relaxations without emission of light (collisions, vibrations, etc)

-a radiative pathway by emitting light (fluorescence or phosphorescence)

-photochemical reactions that cause molecular breakdown and the formation of new compounds.[70]

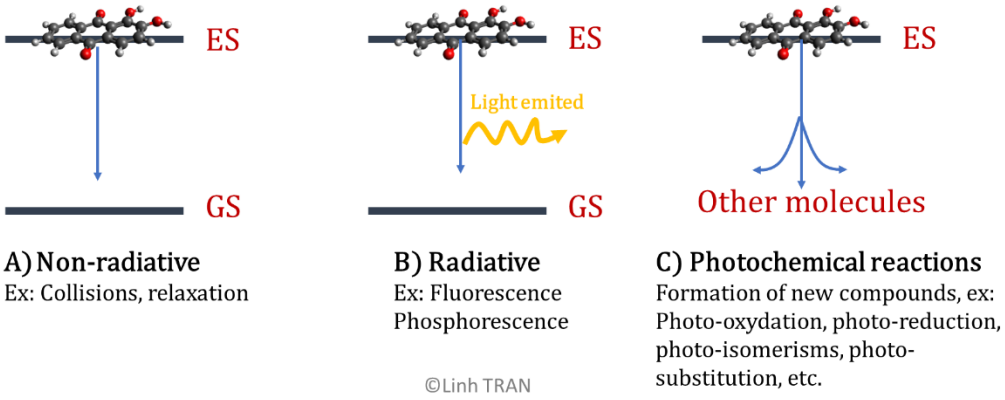


Figure 8. After a molecule absorbs light, it can be promoted to an excited state. It then releases energy through different processes: a) non-radiative without emission of light, such as collisions, relaxation, vibrations, b) radiative with emission of light, like phosphorescence or fluorescence or c) undergoes photochemical reactions to form new compounds.

Photochemical reactions can be classified into two main types: direct and indirect. In the direct pathway, a dye interacts with light, and is brought to an excited state. The dye molecules are not chemically stable in their photoexcited forms and therefore undergo intra- or intermolecular reactions to form photodegraded products. In the *direct* photodegradation pathway, the dye decomposes without interacting with other substances.[67] In contrast, the *indirect* pathway involves another component, such as a photocatalyst, photosensitiser, or solvent, that first becomes excited upon light exposure and subsequently transfers this energy to the nearby dye molecule, initiating its degradation.[40,70–73].

2.3.1. Reciprocity principle and accelerated photodegradation

Usual photodegradation of dyes is typically a slow process, occurring over several months to even years.[51] To study this phenomenon within a practical timeframe, researchers use accelerated ageing techniques under controlled laboratory conditions to ensure reproducibility. This often involves the use of commercial light sources, such as xenon arc lamps, in chambers where parameters like temperature, humidity, light intensity, and wavelength can be precisely regulated, often through the use of selective filters, LED

sources for photolysis with different maximum absorption wavelengths.[74] The reciprocity principle justifies this acceleration, assuming the material's degradation response is linear with respect to light dose (irradiance x time). The effect of photodegradation is the same, whether with high intensity in a short time or low intensity in a long time. [75]

Using this principle, several studies have found that the photodegradation mechanisms can be influenced by both internal and external factors, which are presented below.

2.3.2. Internal factors influencing photodegradation

The photodegradation process is not universal and depends on the type of molecule. Lachman *et al.* demonstrated that the extent of photofading of dyes in pharmaceutical pills varied among different water-soluble organic dyes and did not exhibit a consistent trend.[76] The acceleration in the fading rate ranged from 1.5 times faster for indigotin to 240 times faster for alizarin cyanine green F. The photodegradation kinetics also depends on the dye's concentration. Less concentrated dye degrades faster.

2.3.3. External factors influencing photodegradation

External factors have been studied for several dyes. They can be listed as pH, dye concentration, light wavelength, intensity, source, presence of oxygen, temperature, catalysis, type of substrate (solution or solid), solvents and mordants.[68] They are discussed below.

Influence of pH

During the degradation of dye through heterogeneous photocatalysis, pH plays a crucial role as an operational parameter. It influences several key factors: the charge of the dye molecules, the surface charge of the photocatalyst, the aggregation behaviour of the photocatalyst particles, and the positions of the valence and conduction bands. For photodegradation in solution, redox active groups can be protonated or deprotonated in acidic or alkaline pH, respectively, resulting in forms that may behave differently under light exposure. Seixas de Melo *et al.* studied indigotin in their keto (oxidised form) and leuco-forms (where the two carbonyls are reduced in the presence of alkali).[77] They found different relaxation pathways from the excited to the ground state, resulting in different photochemistry, which could be attributed to different excited-state geometries for each form. The effects of pH depend on the dye structure and its pKa. But in general, the more molecules are charged, the more it is prone to initiate photodegradation reactions. Therefore, the photodegradation of acidic dyes, such as carminic acid, alizarin and phycocyanin, has been shown to increase at higher pH values where these molecules are mostly negatively charged. L. Jepersen found that exposure of phycocyanin at pH 5 and 7 leads to 80% degradation, and its use in the food industry should take that into account for their storage and production.[78] K. Jørgensen *et al.* have studied carminic acid light sensitivity for different pH and found that for three different monochromatic lights, photobleaching increases with increasing pH.[79]

Influence of dye concentration

Photodegradation is significantly affected by dye concentration in both solution and on substrates. Higher dye concentrations often lead to greater aggregation, reducing the surface area exposed to light and environmental factors, and thus slowing photofading. This has been confirmed for various dyes on paper, textiles, and synthetic materials.[80–82] Increased concentrations may also weaken dye–substrate interactions and enhance surface absorption, further reducing degradation rates.

Influence of light

As stated before, researchers often use accelerated photodegradation with artificial lamps. However, the choice of light source (intensity and wavelength) can significantly influence the photodegradation pathways. These pathways depend not only on the spectral distribution and intensity of the irradiation but also on the absorption characteristics of the dye and the duration of exposure. Light intensity significantly influences the rate of photodegradation, as higher intensities result in a greater flux of photons available to interact with the system. However, not all absorbed photons lead to photochemical reactions. The absorbed energy can also be dissipated through non-reactive pathways, such as heat release or fluorescence, rather than triggering chemical transformations. [70] To assess and compare the light stability of various compounds, researchers commonly use the quantum yield of photodegradation, defined as the ratio of the number of absorbed photons that result in a photochemical reaction to the total number of photons absorbed. A high quantum yield indicates a greater probability that photon absorption will lead to photoreaction, whereas a low quantum yield implies that photoreactions are less probable, with energy more likely dissipated through alternative pathways.[79] For instance, Lachman *et al.* demonstrated that the extent of photofading acceleration under increased light intensity varies among different water-soluble organic dyes. [76]In their study on dye tablets, they found no consistent pattern across compounds: the increase in fading rate ranged from a modest 1.5-fold for indigotin to a dramatic 240-fold for alizarin cyanine green F, highlighting the dye-specific nature of photodegradation behaviour. [76] Exposure time also influences photodegradation, with longer durations leading to increased photofading. Accelerated aging tests often rely on the reciprocity principle, which assumes that the same light dose (intensity × time) produces a similar degree of photodegradation, whether delivered through high-intensity light over a short time or lower intensity over a longer period.

Influence of incident wavelength

Wavelength is another key factor in photodegradation. Some dyes degrade readily under visible light, while others require the higher energy of UV light. This behaviour depends on the dye's absorption spectrum: the irradiance spectrum must overlap the dye's absorbance spectrum to be effectively absorbed. [67] For example, Batchelor *et al.* demonstrated that both UV and visible light can induce photodegradation of azo dyes and phthalocyanines under solar conditions. [80] However, visible light was more damaging for azo dyes, while UV light was dominant for phthalocyanines, due to structural differences influencing their photochemistry

Influence of oxygen

Photodegradation is is also influenced by oxygen and the effects vary depending on the dye structure and substrate. Several studies have examined the role of oxygen in dye photodegradation. Arney *et al.* showed that some dyes (like carmin oil paint, felt pen green marker, geranium lake oil paint, alizarin watercolour) fade linearly in proportion to oxygen concentration, while others follow a square-root relationship, like gamboge dye on paper ISO Blue dye on wool, felt pen yellow marker, and mauve watercolour. In protein-based substrates like silk and wool, where reduction dominates, oxygen has less impact. Saito *et al.* confirmed this by observing stronger oxygen effects in cotton (where oxidation occurs) than in silk for six natural anthraquinones. [83] They also tested the fading behaviour of the dyes in ethanol solution and found the following decreasing order of fading rate: purpurin > quinalizarin > alizarin > chrysazin > quinizarin > anthrarufin. This trend correlated well with the oxidation index of the dyes, indicating that a higher oxidation index corresponds to a faster fading rate in solution. Interestingly, Schwen *et al.* found that some dyes, such as diazacyanine blue, degraded faster without oxygen, contrary to typical trends, also seen with Basic Blue 3 and certain azo pigments. [84] Arney *et al.* have found that vermillion and Prussian blue fade more rapidly in absence of oxygen. [85] Generally, the absence of oxygen slows down the fading of organic dyes, as reactive oxygen species (ROS) such as singlet oxygen ($^1O_2$) and superoxide ($O_2^{\bullet-}$), key drivers of degradation, are not formed. Under aerobic conditions, the excited dye molecule (^{3}Dye*) can transfer energy to ground-state oxygen ($^3O_2$), generating highly reactive $^1O_2$ that accelerates degradation. Alternatively, electron transfer between the excited dye and O_2 can produce superoxide radicals ($O_2^{\bullet-}$). In contrast, under anaerobic conditions, the dye may undergo reduction or oxidation depending on the surrounding redox agents, but degradation tends to proceed more slowly in the absence of ROS. These findings highlight that no universal rule applies, and photodegradation must be studied case by case, considering dye type, substrate, and environmental conditions.

Influence of temperature

Temperature has a relatively minor impact on the photofading of organic dyes and pigments. While higher temperatures slightly increase fading rates, the effect is less significant than other factors. Temperature changes do not alter the photodegradation pathways, likely because thermal energy is negligible compared to photon energy. [86,87] Giles *et al.* showed that in fibre substrates, temperature affects fading in two opposing ways: higher temperatures reduce moisture (decreasing photodegradation) but also break down dye aggregates (increasing fading). This aggregation effect is reversible and depends on the fibre type.[88] In solution, elevated temperatures reduce dissolved oxygen levels, but this has minimal influence on degradation pathways or efficiency. [86]

Influence of photocatalysts

Photocatalysts can significantly enhance photodegradation rates by generating reactive radical species upon light irradiation with energy exceeding the catalyst's bandgap. This excitation creates electron-hole pairs, which convert adsorbed oxygen and water into superoxide ions ($O_2^{\bullet-}$) and hydroxyl radicals ($HO^{\bullet-}$), highly reactive species that break down organic dyes. [66,73] While some dyes fully degrade under light, others have a constant

efficiency over time due to the blockage of active catalyst sites. The effectiveness of photodegradation in heterogeneous systems largely depends on catalyst properties such as light absorption range, redox potential, charge separation efficiency, and stability.

Influence of mordants

Mordants are substances that bind dyes to fabrics by forming coordination complexes, greatly influencing lightfastness. Studies show mordants often play a larger role than dye structure in determining colour stability, particularly for similar dye types. Common mordants include tannic acid and metal salts of Al, Cr, Cu, Fe, and Sn. Cr, Cu, and Fe generally enhance lightfastness, whilexas Sn reduce it. [17,89,90] Factors such as mordant concentration and method (pre- or post-mordanting) also affect stability. Different coordination modes, such as Al's bidentate versus Pb's unidentate/bidentate, can affect lightfastness and substrate degradation, as shown by Beltran *et al.* for geranium lake pigments.[91]

Substrate v.s. Solution

Dye photofading in solution does not always reflect behaviour on substrates, as substrate type (e.g., proteinaceous, cellulosic, synthetic) and solvent properties (e.g., proton exchange, polarity) significantly influence photodegradation. Nonetheless, solution or gel systems are sometimes valid models for studying photodegradation mechanisms and kinetics in solids. For example, studies by Confortin *et al.* showed that crystal violet have similar degradation products in solution and on paper, while Sousa *et al.* found that indigo degraded to isatin across both liquid and gel media, consistent with ancient Andean textiles. [92,93] However, photofading in solution is often faster due to greater dye mobility

2.4. Photodegradation of madder

2.4.1. Photodegradation of madder in artworks

One studied example madder photodegradation can be seen in the Bayeux Tapestry, which red parts are dyed in madder. This tapestry is about 70m long and the exposed parts appear to be faded under natural light exposure. Chavanne *et al.* replicated the photodegradation of madder on wool samples and found that munjistin degraded the fastest, followed by rubiadin, purpurin, and then alizarin.[94] These results suggest that anthraquinones with more substituent groups degrade more readily. Consequently, the rate of photodegradation depends on the specific dyes present and their relative concentrations. Additionally, other dye families, such as flavonoids, exhibit even faster degradation rates than anthraquinones.[7,94] Similarly, Casadio *et al.* identified madder colourants in a painting by Homed untitled "To be a Farmer's Boy".[54] The sky of the painting seems unfinished, but it was revealed with Raman SERS that madder was initially present and has faded under light exposure, leading to a resulting white sky.

2.4.2. Photodegraded products and proposed mechanism

The photodegradation of alizarin, a key component of madder dye, has been studied experimentally. The natural light is replaced by an artificial light for reproducibility, as natural light depends on the weather. Under xenon light exposure, Jiang *et al.* found that alizarin oxidises into colourless molecules, leading to fading. [74] They also found that the fading process is pH-dependent, for example, a higher pH leads to a decrease in photostability. The authors have attributed it to a decrease in the HOMO-LUMO energy gap. The oxygen also influences the photofading by increasing its photodegradation rate (the percentage of alizarin that degrades respective to its initial concentration). The photoproducts were found to be $C_9H_8O_4$ (molar mass=180 g/mol), $C_7H_6O_5$ (molar mass=170 g/mol), $C_8H_6O_4$ (molar mass=166 g/mol), $C_6H_6O_4$ (molar mass=142 g/mol), and $C_{10}H_8O_5$ (molar mass=208 g/mol) using HPLC-MS (Figure 9). They proposed the following mechanism: alizarin acts as a photo initiator, which could easily be excited from the valence band (VB) to the conduction band (CB) upon irradiation, and lead to the formation of holes (h^+) in the VB. Then $O_2^{\bullet -}$ formed by the combination of O_2 and e^- produced in the excitation process, may be further converted into highly active ·OH radicals. Then ·OH tend to oxidise and decompose alizarin.

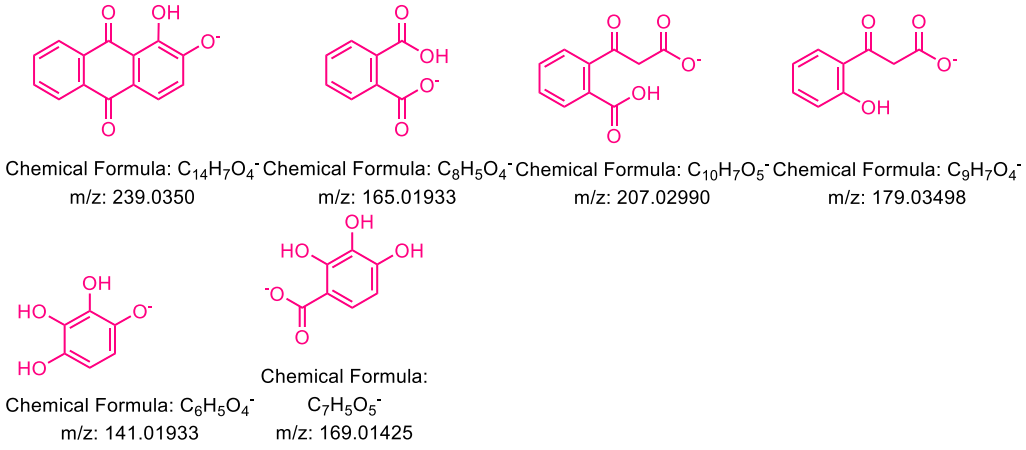


Figure 9. Photodegraded products found in the publication of Jiang et al. [95]

Furthermore, Manhita *et al.* found that the nature of the mordant ion (alum, copper, iron), their concentration, and mordanting technique (premordanting or simsultaneaous mordanting) influenced photodegradation.[89] Copper ions (Cu^{2+}) showed highest binding affinity for wool fiber. Pre-mordanting technique (wool in a mordant bath with metallic cation prior to adding madder dyes) typically yields stronger hues and better photostability due to higher metal uptake. Wool samples photodegradated under 480h with alum mordants shows faster colour fading. Ultimately, dye stability on fabrics depends on the strength of dye-fiber-mordant complexes.

Several other studies have tried to gather hints about the photodegradation mechanism. Guanzheng *et al.* have investigated the structure and degradation of alizarin lake pigments using a surface study approach with aluminium oxide thin films.[46] They demonstrated the

oxidation of catechol groups to carbonyl, resulting in dissociation of the conjugative system. Z. Machatova *et al.* have used EPR spectroscopy and detected alizarin radicals after exposition under UV LED at λ_{max}=365nm. They found that, under photoexcitation, several hydroxyanthraquinone forms reactive oxygen species (ROS).[96]

The findings have demonstrated that the lightfastness of dyes is strongly influenced by several factors such as pH and the presence of oxygen or mordant. Nevertheless, the precise mechanisms underlying their photodegradation remain the subject of ongoing research.

2.5. Difficulty in studying madder dyes

Studying madder and its historical dyeing processes is challenging for many reasons.

First, there are limited written records documenting its cultivation and use. Knowledge was primarily passed down orally from one generation to another. The Greek philosopher Democritus (c. 460 – c. 370 BC) is believed to be one of the first to have documented madder dyeing, but his works have not been conserved. [59] Later, Pedanius Dioscorides a botanist and pharmacologist, wrote "*De Materia Medica*" to document madder dyeing properties. [97,98] Despite these early mentions, little was known about the production and dyeing or lake manufacturing techniques in earlier times. It was not until the 16th century that more detailed madder recipes and technical knowledge were recorded. They became parts of national heritage, like the Turkey red (also known as Adrianople red), which flourished in the 18th to 19th century. [5] The madder dyeing techniques were closely guarded secrets, particularly under Turkish rule. The red textiles were obtained through complex and tedious recipes incorporating ingredients like urea, ash, soda, plus sheep's dung, blood, alum, and olive oil.[99] Their dyeing recipe and techniques were highly protected due to their economic value, reflecting the country's industry and production. At the time, colour was achieved solely through artisans' expertise, with little understanding of the chemical composition or structure of madder dyes.

Secondly, studying natural madder dye is challenging due to its inherently complex composition, which includes a mixture of multiple molecular compounds. As stated previously, their composition varies based on species, age, cultivation or extraction methods, soil conditions, and geographical location. For example, European madder (*Rubia tinctorum* L.) primarily contains alizarin (1,2-dihydroxyanthraquinone), whereas Indian madder (*Rubia cordifolia* L.) is richer in purpurin (1,2,4-trihydroxyanthraquinone), though purpurin is also found in European madder. [89]

Finally, another major challenge in studying natural madder dye lies in the fragility of certain compounds, such as glucosidic ones, which are highly fragile, prone to hydrolysis, decarboxylation, or degradation when exposed to heat, enzymes, chemicals, or light. This is the case, for example, with munjistin and pseudopurpurin, a precursor of purpurin. During the dyeing process or extraction for analytical purposes, they can be easily decarboxylated into xanthopurpurin and purpurin, respectively. Additionally, some key dye compounds, such as ruberythric acid (a precursor of alizarin, one of madder's principal colourants), are either not commercially available or difficult to synthesise.

The limited availability and fragility of those compounds lead to a lack of a reference database for their identification by spectroscopy techniques. [34,55,57,100,101] Moreover, their colour and absorbance spectra are highly dependent on pH and solvent conditions, further adding to the complexity of their characterisation.

2.6. Computational spectroscopy: state of the art

To help overcome the experimental difficulties outlined in the previous subsection, we chose to combine the experimental data with a computational investigation. In this thesis, we are interested in predicting the optical properties of dyes. These properties refer to how dye molecules interact with light, in particular how they absorb, emit or reflect light across various wavelengths. We will focus on the absorbance spectra and colour of the dyes.

The first step is to find a computational protocol that can predict these properties for madder molecules to complete the missing database. To do so, a study of the computational benchmark was done.

2.6.1. Computational chemistry in the cultural heritage field and dyes

Computational chemistry is widely used in materials science, but its application in the cultural heritage field is relatively new. [102] This can be explained by the inherent complexity of artistic materials presents significant challenges for computational modelling. Their composition can be unclear, and they often undergo unpredictable changes over time, making accurate simulations difficult.

Spectroscopic techniques allow non-invasive probing of molecular structure, properties, and dynamics across various conditions. Using multiple methods across the electromagnetic spectrum provides a more complete understanding of molecular systems. As these techniques become more advanced, interpreting results often requires computational support. Computational molecular spectroscopy, originally a branch of quantum chemistry, predicts spectroscopic features and aids interpretation. It has since evolved into a widely used tool, even among experimental researchers. [103]

In recent years, computational approaches, such as density functional theory (DFT), has started to gain attention, to complement experimental investigations in cultural heritage studies.[101] These methods help model the structural, electronic, and spectroscopic properties of complex systems, including organic pigments like alizarin-based compounds. The predicted spectra are: IR/Raman, UV-VIS, X-rays, EPR, NMR, Lifetimes (fluorescence/phosphorescence).[103] There are many software programs developed, and the power of supercomputers is always increasing, thus democratising the use of computational chemistry. The most used software is Gaussian and ORCA.[103] In this thesis, we choose ORCA, a free software developed by Franck Neese. [104]

For instance, Pagliai *et al.* have computed the IR and Raman spectra of alizarin, purpurin, kermesic acid, and carminic acid, using DFT at the B3LYP/6-31G(d) level.[49] Computed vibrational frequencies guided spectral assignment and interpretation of the dyes'

aluminum complexes. Similarly, Silva *et al.* have combined experimental data with computed NMR shifts to find a structure for madder lake pigments. [6] Tissier *et al.* have computed the UV-VIS absorbance spectra and NMR chemical shifts of alizarin, purpurin. [48] Z. Machatová *et al.* have used DFT to predict EPR g-tensor and hyperfine couplings for alizarin in DMSO, to access its photodegradation intermediates and interpret the experimental EPR result. [96] Other methods combining molecular dynamics and a newly developed approach is quantum mechanics/molecular dynamics (QM/MM). [105]

2.6.2. Role of functional and basis set for TD-DFT vertical energy

The simulation UV-VIS absorbance spectra of organic and inorganic molecules is often done using Time-Dependent Functional Density (TD-DFT).[106] This method is presented in the *section 4.4*. The choice of functional and basis sets to compute the vertical energy (VE), which can be approximated as the maximum absorption of the spectra, is important. This choice is tailored to the specific family of dyes being studied.[107] In the literature, there are benchmark reviews available for different dyes. Benchmarks have been conducted to assess the accuracy of various exchange-correlation functionals for predicting vertical excitation energies, oscillator strengths, and excited-state geometries.[108–110] In 2009, D. Jacquemin *et al.* published an extensive benchmark on more than 500 organic compounds and excited states.[108] He found that global hybrids with 22-25% exact exchange and a long-range-corrected hybrid (X3LYP, B98, PBE0, and mPW1PW91) provide the best match with reference data. Notably, A. Laurent and D. Jacquemin have published a review on the available functional benchmarks in 2013.[106] This paper reviews TD-DFT benchmarks from the last decade, focusing on identifying the most accurate exchange-correlation functionals by comparing TD-DFT results with CC2 calculations and experimental data for biochromophore models, and identifies B3LYP, PBE0, or M06 as providing the most satisfying transition energies in general with a deviation about 0.25eV. Adiabatic transition energies are accurately computed with functionals like BMK, BHLYP or M06-2X. For large organic dyes, Goerigk and Grimme have shown that double-hybrid functionals like B2GP-PLYP gave promising results, while conventional hybrids with ~40% Fock exchange and range-separated functionals like CAM-B3LYP also perform well.[111]

For anthraquinone compounds in particular, D. Jacquemin *et al.* benchmarks on industrial anthraquinone dataset show that the choice of functional significantly influences transition energies and vibronic shapes, while the basis set has a limited impact. [107] The functional ωB97XD is identified as a good compromise for achieving accurate positions and intensities of bands. Other studies have shown that TD-DFT performance varies depending on the functional used, with B3LYP and PBE0 generally performing well for singlet states but underestimating triplet energies. [109,112] Similarly E. Anouar et al have tested different functionals for synthetic anthraquinone dyes. He found out that PBE0 or its combination with B3LYP and B3P86 hybrid functionals are the most suitable functionals for this family of dye.

M. Dierksen and S. Grimme studied the influence of TD-DFT functionals for a large set of various closed-sell and open-shell molecules. [113] They found that the accuracy of vibronic absorption spectra simulated using TD-DFT critically depends on the fraction of exact

Hartree–Fock exchange (EEX) included in the hybrid functional. Specifically, functionals with ~30–40% Hartree–Fock exchange generally produce the most reliable agreement with experimental vibronic spectra across a wide variety of molecules (including dyes, biomolecules, and radical species).

PBE0 developed by Adamo and Barone, is a more physically motivated hybrid functional. [114] They have been tested by several benchmarks on anthraquinones and other dyes, and thus will be tested and used for this thesis.

2.6.3. Absorption spectra and colour: vibronic coupling

The computation of vertical energy results in a spectrum with vertical narrow lines, which is difficult to compare to experimental spectra. Furthermore, additional peaks can arise from the coupling between electronic and vibrational degrees of freedom. Therefore, we need to simulate the spectra broadening, which arises from several factors: electronic transitions accompanied by vibrational and sometimes rotational transitions (also known as vibronic coupling, or vibrational effects), solvent effects, temperature, or conformational flexibility.

In the literature, several methods are used to simulate the spectra broadening in UV-VIS spectra. We can find empirical methods, which depend on experimental data or fitted parameters, and non-empirical methods, which rely purely on theory, which are fitted for rigid molecules. For larger molecules, Molecular Dynamics (MD) and Monte Carlo (MC) can be used.

a) Empirical methods

Among empirical methods, Gaussian function broadening has been used for its computational efficiency. Other functions can be found, such as Lorentzian or pseudo-Voigt function. [115–117] For example, B. Tirri and E. Brémond applied these functions to predict the absorption spectra and perceived colour of industrial dyes, including anthraquinone derivatives.[116,118] The full width at half maximum (FWHM) of the convolution functions was fixed in the case of Gaussian functions to 0.2–0.3 eV. However, the accuracy of this empirical method depends on prior benchmarking, and its transferability across diverse dye systems is limited. Tirri *et al.* did a comparative study of empirical and non-empirical broadening techniques for rigid dye molecules like anthraquinone and indigo.[116] They demonstrated that empirical functions like Gaussian and pseudo-Voigt approximations yield satisfactory agreement with experimental data. The Euclidean colour difference (ΔE), which ranged from 16 to 75 for anthraquinones, depending on the functional and vibronic treatment used. PBE0 functional with pseudo-Voight model is the most accurate. The empirical approach with PBE0/AHFC yields a similar colour to the PBE0/Gaussian approach.

b) *Non-empirical methods*

The broadening of the spectra using an empirical approach, with a fixed width and fitted parameters, leading to spectra similar to the experimentally recorded one, lacks physical

information. [119] Additional peaks from vibronic couplings cannot be obtained with this method.

In contrast, empirical approaches that explicitly account for vibrational effects, such as the Vertical Gradient (VG) and Adiabatic Hessian (AH) Franck–Condon (FC) approximations, offer higher accuracy in capturing the fine structure and band shapes of absorption spectra. These methods, however, are computationally demanding, particularly for large or flexible molecular systems. [116] The details of these methods are found in *section* 4.5.3. A study by A. Laurent, C. Adamo, and D. Jacquemin provides an overview of applications of time-dependent density functional theory (TD-DFT) to determine excited-state properties of dyes, elucidate band shapes, focusing on optical absorption and emission spectra.[112] Results were obtained using both vertical and adiabatic (vibronic) approximations to include vibronic effects. TD-DFT can also be used to optimise solar cell dyes, predict fluorescence wavelengths, study optically active biomolecules, and improve photochromes. Tissier *et al.* have tested different vibrational effects, such as Vertical Gradient (VG) and Adiabatic Hessian (AH), for alizarin and purpurin, the main madder dye, in DMSO. He found that AH performs better. [48]

c) *Other methods for large and flexible molecules*

All those above-mentioned methods perform well for rigid or semi-rigid molecules.[116] However, for larger molecules and flexible ones, the computation of AH with Franck-Condon approximation requires lots of computational resources, and sometimes does not converge. In that case, other methods like Molecular Dynamics (MD) or Monte Carlo (MC) simulations have been developed and applied to the prediction of properties of these types of dyes. [120–122] Molecular dynamics are often used to compute vertical energy for a set of structural snapshots (hundreds to thousands) along the MD trajectory. The final absorption spectrum is constructed by averaging multiple spectra, each broadened using Gaussian or Lorentzian functions. This method is particularly effective when the absorption profile results from ground-state various conformers. However, it is computationally intensive and neglects quantum nuclear effects, leading to the absence of peaks arising purely from vibrational transitions.[119] E. Tapavicza *et al.* developed a real-time generating function method to avoid complex FC approaches to compute vibronic spectra.[115,119] They also used Born–Oppenheimer molecular dynamics simulations (BOMD) to simulate vibronic spectra. Contrary to MD, BOMD considers the quantum nature of nuclei.[115] They found that the quantum nuclear motion is necessary in the case of TCNQ anion computation, and that classical MD cannot accurately model some experimental bands.

While widely used, MD methods can suffer from poor sampling due to high energy barriers between molecular conformations (bond rotation, geometrical rearrangements in aggregates) that trap the system in local minima. Techniques such as high-temperature simulations, simulated annealing approaches, restarting the dynamics at different initial conditions, or enhanced sampling methods (e.g., metadynamics, umbrella sampling) can help, but are computationally expensive and may lead to artificially broadened spectra.[123–125] The details of these methods can be found in the original references.
Other methods include Monte-Carlo (MC) simulations for the prediction of vibronic and solvent effects in UV-VIS spectra. This method is used for large, flexible molecules that

possess many conformations and a strong solvatochromism.[122] MC is a statistical technique used to explore the configurational space of molecules by generating a large number of different molecular geometries and conformations by randomly sampling possible arrangements and accepting the new configuration according to the Metropolis sampling algorithm.[122] For instance, B. Tirri used a combination of MC sampling of the configurational space and time-dependent density functional theory (TD-DFT) for calculating vertical excitation energies and to simulate the absorption spectra of a series of merocyanine dyes in both their monomeric and dimeric forms.[122]

2.6.4. Predicting the influence of the environment

The spectra and colour of dyes can be impacted by the presence of solvents, such as water. Those solvents can form hydrogen bonds, participate in excited proton transfer, or change the molecule's vibration or molecular geometry due to charge or steric effects. Hence, they must be considered when computing absorbance spectra.

Carta *et al.* have studied the influence of environment, such as the presence of tautomers, pH, mordant, and solvents, for the prediction of free alizarin and its complexes with Mg(II) and Al(III).[101] By accounting for solvation and metal coordination effects, the research reveals how environmental and complexation conditions influence pigment colour. The computational results show good semi-quantitative agreement with experimental UV-VIS spectra, highlighting the power of in silico methods in interpreting and predicting the behaviour of cultural heritage materials. Le Person *et al.* study revealed that intramolecular proton transfer in alizarin occurs not only in the excited state but also partially in the ground state. [126] Using UV–Vis electronic spectroscopy in methanol and time-dependent DFT (TD-DFT) calculations, the authors identified a tautomeric form of alizarin (1,10-keto form) coexisting with the more stable normal form (9,10-keto form) at room temperature, comprising approximately 10% of the total population. The ground-state tautomeric equilibrium is temperature-dependent and independent of concentration, with a calculated Gibbs free energy difference of 1.2 kcal·mol^{-1}. These findings challenge previous assumptions that tautomerism in alizarin was confined solely to the excited state. Similarly, Douma *et al.* have tested two solvation models for alizarin: implicit PCM and explicit. The authors found that inclusion of explicit solvent changes the predicted UV-VIS spectra and takes into account proton transfer processes. [127]

CHAPTER 3

CHAPTER 3 | EXPERIMENTAL METHODS 51

3.1. Invasive methods 51
3.1.1. Liquid NMR 51
3.1.2. HPLC-UV-MS 52
3.1.3. EPR 53

3.2. Non-Invasive methods 55
3.2.1. Absorption spectra 55

3.3. The science of colour 57
3.3.1. Traditional colour theory 57
3.3.2. Colour science 57
3.3.3. Colourimetry 58
3.3.4. Measure colour: CIELAB space 58
3.3.5. Measure colour difference with ΔE 59
3.3.6. Converting UV-VIS spectra to CIELAB coordinates 61

3.4. Photodegradation 62

In this chapter, we describe the experimental methods used in this thesis. It is divided into invasive and non-invasive methods, colourimetry and photodegradation subsections.

CHAPTER 3 | EXPERIMENTAL METHODS

3.1. Invasive methods

3.1.1. Liquid NMR

a) Principle

Nuclear Magnetic Resonance (NMR) spectroscopy is an analytical technique used to determine the structure, purity, and dynamics of molecules based on the magnetic properties of atomic nuclei. [128,129] Not all nuclei are equally visible in NMR. The sensitivity depends on two factors: the nuclear spin and natural abundance. The abundance of ^{1}H is 99.98%, so the sensitivity is high. For ^{13}C, its natural abundance is 1.1%, so it is less sensitive. Nuclear spin quantum number I=0, such as ^{12}C, ^{16}O, ^{2}H, are NMR-inactive and do not produce a signal. In contrast, nuclei with spin of I=1/2, like ^{1}H, ^{13}C, ^{19}F, and so on, are NMR-active and produce a signal. In NMR, a sample is placed in a strong magnetic field, causing certain nuclei, most commonly hydrogen (^{1}H) or carbon (^{13}C), to align with or against the field. When radiofrequency (RF) pulses are applied, these nuclei absorb energy and transition between magnetic energy levels. As they relax back to their original state, they emit signals that are detected and translated into a spectrum.[129] The chemical shifts, coupling constants, and signal intensities provide detailed information about the chemical environment and connectivity of atoms within the molecule. In liquid NMR, the solvents also have ^{1}H, ^{13}C, and so on, and they may interfere with the signals. Therefore, deuterated solvents (with Deuterium ^{2}H) are used not to interfere with the sample's ^{1}H protons. The sample must be soluble in the solvent. In NMR, the choice of magnetic field B_0 strength is directly related to the properties of the nucleus under investigation, particularly its magnetic moment and gyromagnetic ratio (γ). When a nucleus with a non-zero spin ($I \neq 0$) is placed in an external magnetic field B_0, its energy levels split due to the Zeeman effect (Figure 10). The energy difference between these levels is given by $\Delta E=\gamma\hbar B_0$, which determines the frequency at which the nucleus will resonate when exposed to radiofrequency radiation. This resonance frequency must match the Larmor frequency, $\nu=\gamma B_0/2\pi$, specific to each type of nucleus. Therefore, the magnetic field must be carefully selected to match the gyromagnetic ratio of the nucleus to be studied, ensuring that the resonance occurs within the detectable radiofrequency range. Higher magnetic fields increase the energy separation ΔE, improving both sensitivity and resolution of the NMR signal, which is especially important for nuclei with low natural abundance or low gyromagnetic ratios, such as ^{13}C.

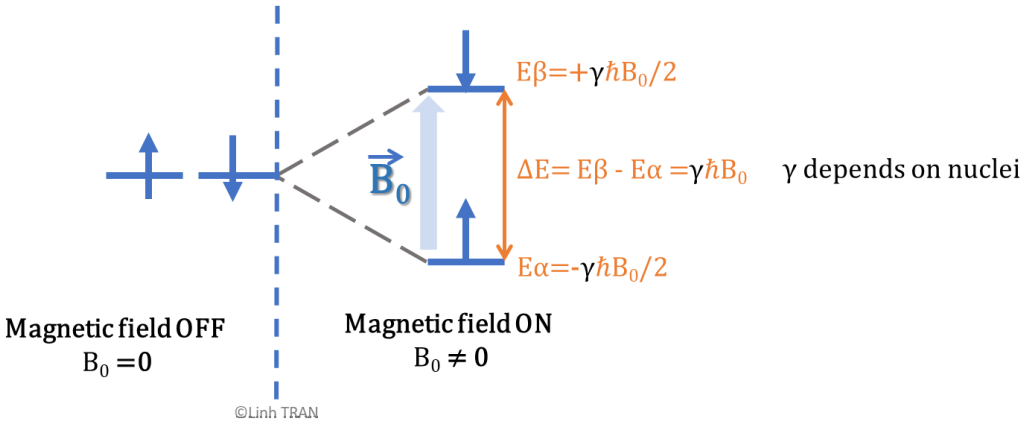


Figure 10. Schematic diagram of NMR resonance condition.

b) Protocol for alizarin and alizarin red S

In this thesis, liquid NMR spectra were acquired with a Bruker Advance Neo 500 spectrometer at the frequency of 500.07 MHz for ^{1}H and 125.76 MHz for ^{13}C equipped with a 5 mm double resonance broadband prodigy probe use nitrogen-cooled RF coils and preamplifiers. Chemical shifts were calibrated using the Tetramethysilane (0 ppm). For ^{1}H, we used a one 90◦ pulse with 64 scans and with a recycling time of 1s. For ^{13}C, we used a one 90◦ pulse with a ^{1}H decoupling waltz16 with 4096 scans and with a recycling time D1 of 2s. The ^{1}H and ^{13}C spectra were recorded in exactly the same conditions at room temperature and processed with a zero-filling factor of 2 and without an exponential decay. 5 mg of either alizarin or alizarin red S were dissolved in 0.5 mL of acetone-D6. The ^{1}H and ^{13}C NMR spectra were also measured for alizarin red S in D_2O as it was more soluble than alizarin.

3.1.2. HPLC-UV-MS

a) Principle

The HPLC-UV-MS/MS method combines High-Performance Liquid Chromatography (HPLC) with Ultraviolet (UV) detection and tandem Mass Spectrometry (MS/MS) to enable sensitive, selective, and accurate analysis of complex mixtures. [130] In this technique, HPLC first separates the components of a sample based on their interactions with the stationary and mobile phases. The separated compounds are then detected by a UV detector, which measures absorbance at specific wavelengths, providing initial qualitative and quantitative data. Following UV detection, the analytes are introduced into the mass spectrometer, where they are ionised, fragmented, and analysed based on their mass-to-charge (m/z) ratios. The MS/MS system allows for further fragmentation of selected ions, offering high specificity for structural identification and quantification. [58]

b) Protocol for madder

In $H_2O + NH_4OH$

Solutions of alizarin (Sigma Aldrich, 97%) were prepared with a concentration of 10^{-2} M, adjusted at pH=12 with NH_4OH. Liquid chromatography-UV-mass spectrometry (LC-UV-MS) experiments were performed on a 1290 Infinity II LC system coupled to a Q-TOF 6540 mass spectrometer equipped with an ESI source operated in the negative and positive ionization modes. The chromatographic separation was carried out on a Zorbax 300SB-C8 (150 × 2.1 mm, 5 µm) column heated at 35 °C at a flow rate of 300 µl/min with water at 0.1% of formic acid as solvent A and acetonitrile as solvent B. Five microliters of samples in H_2O + NH_4OH at pH12/H_2O were injected using the following gradient: 0 min (5% B), 20 min (40% B), 22 min (100% B), 25 min (100% B), 26 min (5% B)and 6 min of equilibration time. Ion source conditions were set for negative ion mode: capillary voltage 3.5 kV, Nozzle voltage 1250 V, Fragmentor voltage 125 V, Skimmer 65 V, drying gas at 10 l/min, 275 °C, nebuliser 40 psi, and sheath gas flow at 10 l/min, and 325 °C. UV detection was performed at 260 and 410 nm for quantitation at 260 nm. The TOF was operated in the 40−1700 m/z mass range. Data Acquisition 9.0 and Qualitative Analysis 8.0 software were used to process data.

In MeOH/H_2O 50:50 with and without O_2

Six aqueous samples containing alizarin (1mg/mL) were irradiated at different times were centrifuged, and the supernatant was diluted by a factor of 2 in MeOH before LC-UV-MS analysis. Liquid chromatography-UV-mass spectrometry (LC-UV-MS) experiments were performed on a 1290 Infinity II LC system coupled to a Q-TOF 6540 mass spectrometer equipped with an ESI source operated in the negative and positive ionisation modes. The chromatographic separation was carried out on a Zorbax 300SB-C8 (150 × 2.1 mm, 5 µm) column heated at 35 °C at a flow rate of 300 µl/min with water at 0.1% of formic acid as solvent A and acetonitrile as solvent B. Five microliters of samples in MeOH/H2O 50:50 were injected using the following gradient: 0 min (10% B), 25 min (55% B), 30 min (100% B), 35 min (100% B), 36 min (10% B)and 6 min of equilibration time. Ion source conditions were set for negative ion mode: capillary voltage 3.5 kV, Fragmentor voltage 125 V, Skimmer 65 V, Drying gas at 10 l/min, 275 °C, and sheath gas flow at 10 l/min, and 325 °C. UV detection was performed at 232 and 260 nm for quantitation at 260 nm. The TOF was operated in the 50−1700 m/z mass range. Data Acquisition 9.0 and Qualitative Analysis 8.0 software were used to process data.

3.1.3. EPR

a) Principle

Electron Paramagnetic Resonance (EPR), also known as Electron Spin Resonance (ESR), is a spectroscopic technique used to study materials or molecules that contain unpaired electrons, such as free radicals, transition metal complexes, or defects in solids. In EPR, the sample is exposed to a strong magnetic field and microwave radiation. Unpaired electrons have a magnetic moment due to their spin, and under the magnetic field, they can align in two energy states. When the energy of the microwave radiation matches the energy difference between these states, resonance occurs, and the electrons absorb energy. This absorption is measured and plotted as an EPR spectrum. The resulting signals provide information on the number, environment, and interactions of unpaired

electrons, as well as on nearby atomic nuclei through hyperfine splitting (Figure 11). Two key parameters in EPR are the g-factor (g constant) and the hyperfine coupling constant (a_n).

The resonance condition in EPR is: $h\nu = g\mu_B B + m_I a_n$, where

- h = Planck's constant
- ν = Microwave frequency (GHz range)
- g= g-factor (dimensionless; varies with the environment of the electron)
- μ_B = Bohr magneton
- B = Magnetic field strength
- m_I= Nuclear spin projection quantum number (e.g., for I = 1/2, m_I=±1/2)
- a_n = Hyperfine coupling constant (in MHz or mT)

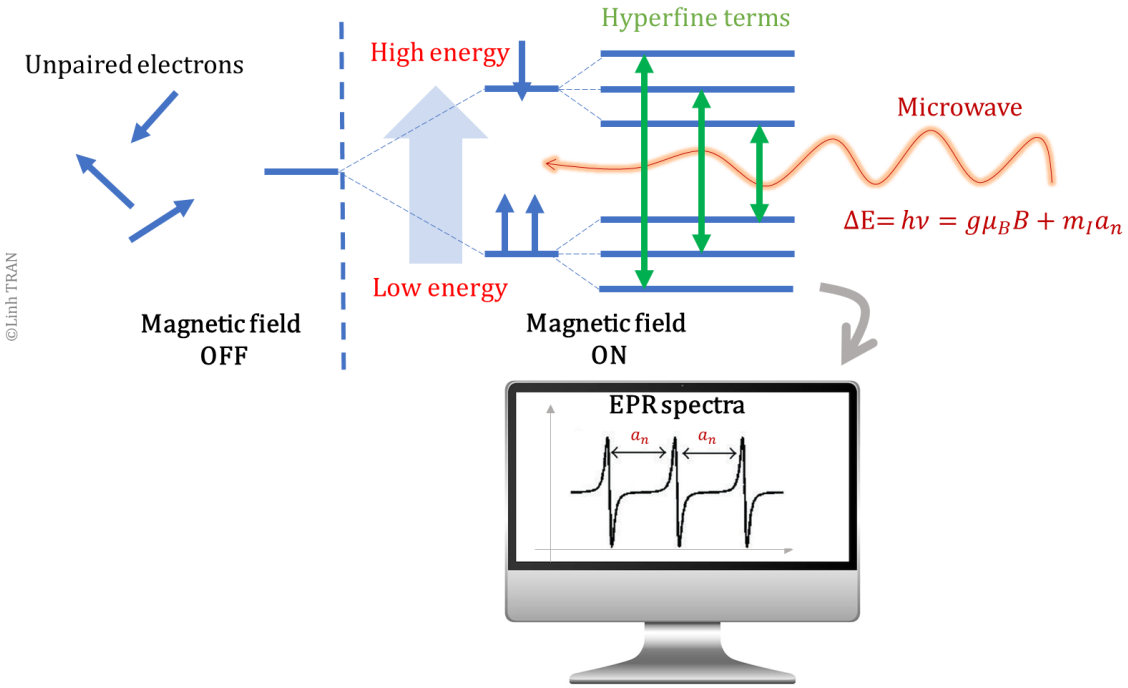


Figure 11. Schematic diagram of EPR resonance condition and hyperfine coupling.

b) Protocol for alizarin photodegradation product

Alizarin was dissolved in water at pH=12, adjusted using NH_4OH. The formation of paramagnetic intermediates upon irradiation of alizarin was monitored by cw-EPR spectroscopy using an EMX Plus X-band EPR spectrometer (Bruker, Germany). The samples were irradiated at 293 K directly in the EPR resonator (High Sensitivity Probe-head, Bruker), and the EPR spectra were recorded in situ using a capillary in an EPR tube. The irradiation source was a UV lamp from Hamamatsu with an interferential filter at 365 nm and an intensity of 200 mW/cm2. The acquisition parameters are: receiver gain is 1.0.105. The sweep width is 25G (central field: 3488 G). The modulation amplitude is 3 G. The time constant is 5.12 ms. The resolution is 1024 pts. The power is 6.325 mW (15dB). Scans: 150. Acquisition time is 1500s.

3.2. Non-Invasive methods

3.2.1. Absorption spectra

An absorption spectrum represents how much light is absorbed by a substance at different wavelengths. This absorption is governed by quantum mechanical principles and depends on the electronic, vibrational, and rotational energy levels of the absorbing species.[131] This spectroscopy will be detailed in the method section.

a) Jablonski Diagram

The Jablonski diagram is a useful tool to explain the processes involved in the absorption and emission of light by molecules, particularly in the context of absorption spectra. It illustrates the energy levels of a molecule and the transitions that occur between them. When a molecule absorbs a photon, it undergoes an electronic transition from the ground state (S_0) to an excited singlet state (e.g., S_1 or S_2). This absorption corresponds to a specific wavelength of light and appears as a peak in the absorption spectrum. The energy difference between the ground and excited states determines the position of the peak. After absorption, the molecule may undergo non-radiative relaxation (internal conversion, intersystem crossing or vibrational relaxation) to the lowest vibrational level of the excited state, followed by fluorescence (if emission occurs from the excited singlet state) or phosphorescence (from the triplet state, T_1). The Jablonski diagram helps visualize these transitions and energy losses, explaining why absorption and emission spectra often differ (Figure 12).

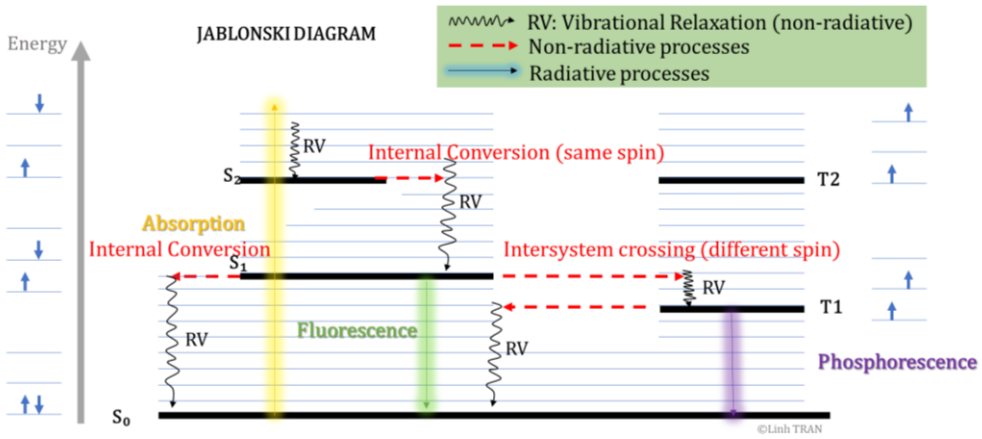


Figure 12. Jablonski diagram, the different electronic transitions: absorption, fluorescence, phosphorescence, internal conversion, intersystem conversion, vibrational relaxation.

b) Principle: Ultraviolet-Visible (UV-VIS) spectra

UV-VIS spectroscopy works by measuring the absorption of ultraviolet (UV) and visible light by a sample. When light passes through or is reflected from a sample, certain wavelengths of light are absorbed by the sample's molecules. The spectrophotometer thus measures the absorbance A of a solution at a wavelength λ by taking the logarithmic ratio of the intensity of the emitted monochromatic light to the intensity of the transmitted

light: $A_\lambda = log\left(\frac{I}{I_0}\right)_\lambda$. The Figure 13 presents a schematic diagram of a UV-VIS Absorption spectrophotometer.

The absorption occurs when electrons in the molecules are excited from a lower energy state to a higher energy state. The specific wavelengths of light that are absorbed depend on the electronic structure of the molecules. The resulting spectrum displays the intensity of light absorbed at various wavelengths, and peaks in the spectrum indicate the wavelengths where absorption is strongest. By analysing these peaks, we can gain information about the sample's electronic structure, such as energy levels and band gaps.

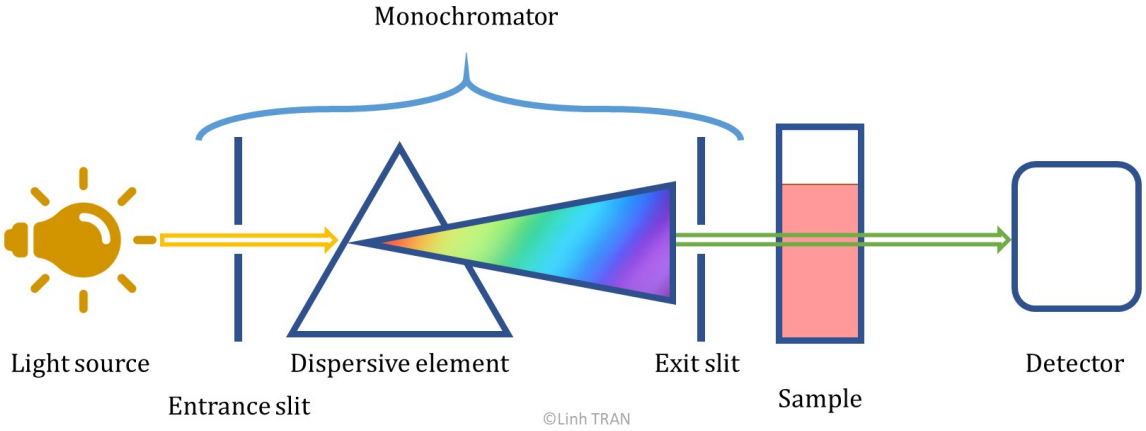


Figure 13. Schematic Diagram of a UV-Visible Absorption Spectrophotometer

UV-VIS spectroscopy, when combined with the Beer-Lambert Law, allows for the quantitative analysis of a sample's concentration based on its absorption. The Beer-Lambert Law states that the absorbance (A) of light by a sample is directly proportional to the concentration (C) of the absorbing species, the path length (l) of the light through the sample, and the molar absorptivity (ε), which is a constant specific to the substance and wavelength of light being used. The Beer-Lambert Law equation is: $A = \varepsilon . l . [C]$.

c) Use in the analysis of madder dyes

In this study, UV-VIS (Ultraviolet-Visible) spectroscopy was utilized as a primary analytical technique to investigate the optical properties of the madder samples. Samples were prepared by dissolving the substances in appropriate solvents, e.g., water or other detailed in the articles, at a concentration of 10^{-4} M. Their UV-VIS spectra were measured immediately using an Ocean Optics Flame spectrometer (Model: FLMS00699) with a wavelength range from 300 to 850 nm. The integration time for each measurement ranged from 5 ms to 10 ms. The light source used was a Deuterium-Tungsten Halogen lamp covering the range of 300 nm–850 nm. The blank was done using deionized water. This light source was coupled to the spectrometer using 400 µm diameter optical fibers and positioned through a 1 cm quartz cuvette holder. The acquisition of spectra was facilitated by the Ocean View software.

Calibration curves, when necessary, were constructed to quantify the concentration of specific species present in the samples.

3.3. The science of colour

While colour holds profound cultural and historical significance, it is also a physico-chemical phenomenon, and therefore a subject of scientific examination. Numerous efforts to understand colour from this perspective have contributed to the development of a dedicated scientific discipline: colour science.

3.3.1. Traditional colour theory

The science of colour initially took shape through traditional colour theory, which provided the foundational frameworks for understanding colour relationships. Traditional colour theory is a historical knowledge describing colour effects, the mixing of colour, their harmony, colour schemes, and symbolism. It tends to be more subjective than colour science and is used by artists, designer for artistic applications. The traditional colour theory dates back at Aristotle's 'On colours", which says that all colours are derived from combinations of black and white. [132] Since then, several other colour theories were proposed by artists (Figure 14). They differ by the set of "primary colours" from which other colours are created, the classification of colours, the "colour harmony", the relationship between colour or the combination guidelines for colour applications.

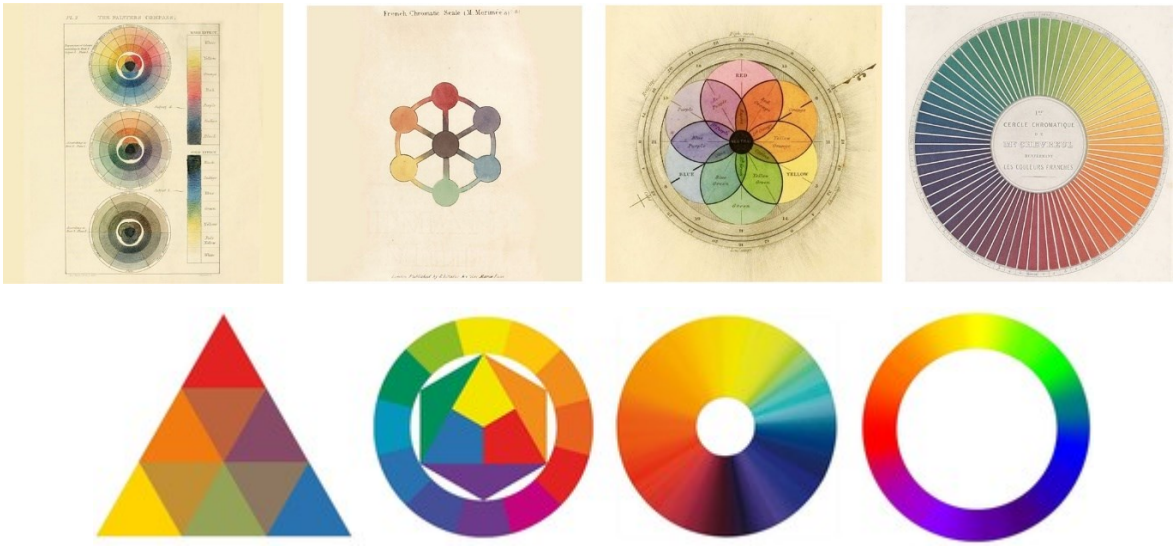


Figure 14. (Top From left to right) Hue circles of Hayter (1826), Mérimée (1830), Field (1835), and Chevreul (1855). (Bottom From left to right) Colour illustrations, Albers (1963), Itten (1961), Eliasson (2009) and O'Connor (2020). Illustration adapted from [132] (with author permission).

3.3.2. Colour science

Colour science is the scientific study of colour including optics, lights, measurement of colour, colour reproduction, the physiology and psychophysics of colour vision. It is the continuation of the traditional colour theory. One of the first formalisation of colour science was done by Isaac Newton's treatise on "Opticks: or, A Treatise of the Reflexions, Refractions, Inflexions and Colours of Light" in 1704. [133,134] Several optics and spectroscospic techniques have been developed to study light and its interaction with matter. Among them, absorbance spectra, reflectance spectra and more recently

hyperspectral imaging are the most used to study colour. In this thesis, we will mainly focus on absorbance spectra. [39,45,135]

3.3.3. Colourimetry

Colour observation differs from one person to another. Therefore, a system of colourimetry is necessary to describe and reproduce colour unambiguously. Colorimetry is the "branch of colour science concerned with specifying numerically the colour of a physically visual stimulus", as defined by Wyszecki and Stiles.[136] There are many international organisations that have built standard systems to measure colour or light references. A known example is the "Commission Internationale de l'Éclairage" (known as CIE). They have established the CIElab system colour coordinates, which are widely used in the cultural heritage field, as well as in other colour-related fields. It will be used in this thesis, and the coordinates as well as their formula are detailed in the *Methods* section.

Colour perception and visible region of electromagnetic light

Visible light is a form of electromagnetic radiation perceptible to the human eye that spans wavelengths from 380-700nm (Figure 15).[131] It includes all colours of the rainbow that the human eye can perceive. We see different colours depending on the wavelengths of visible light that reach us. When light from a source (like the Sun or a bulb) hits an object, some wavelengths are absorbed, while others are reflected or transmitted. The reflected visible light is converted to colour by human eye through our photoreceptor cells. Thus, the colour we see are the opposite of the absorbed colours. For instance, wavelengths of approximately 680-740 nm look red while a wavelength of 550 nm is yellow-green. A way to study colour and characterize it is to study absorption spectra.

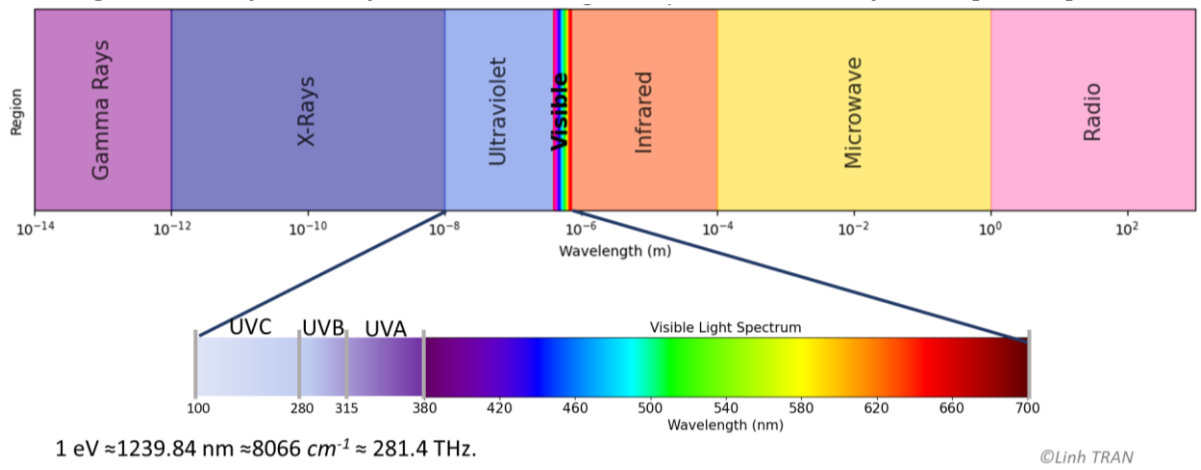


Figure 15. Illustration of electromagnetic spectrum, with a zoom on the visible part of the spectrum.

3.3.4. Measure colour: CIELAB space

To assist in quantifying colour differences, measurement systems have been developed. By assigning each colour a coordinate in colour space, we can measure in quantitative numbers the difference between two colours. There are numerous colour spaces available, but the most common for textile applications are CIE L*a*b* and CMC. The CIE L*a*b* (or CIELAB) colour space (Figure 16) is a colour representation model developed

by the International Commission on Illumination (CIE) in 1976.[137] It is designed to be perceptually uniform, meaning that equal changes in values correspond to roughly equal changes in human colour perception. It is widely used in industries like printing, textiles, photography, and colour science because it represents colours independently of devices (unlike RGB or CMYK, which depend on specific screens or printers).[136,138,139]

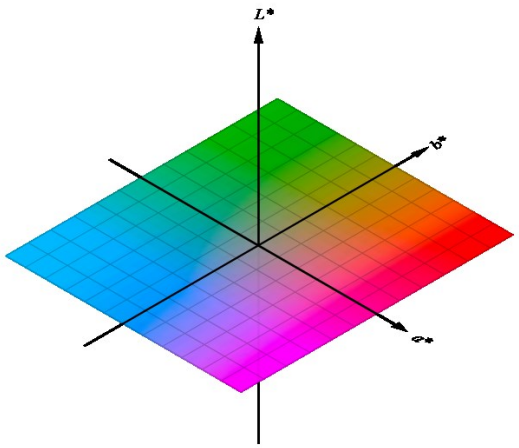


Figure 16. CIELAB colour space representation.

The CIELAB model is a three-dimensional colour space with the following components:

L* → Lightness (Brightness)

Ranges from 0 (black) to 100 (white)

Measures the perceived brightness of a colour, independent of its hue or saturation

a* → Green-Red Axis

Negative values (−) indicate green

Positive values (+) indicate red

b* → Blue-Yellow Axis

Negative values (−) indicate blue

Positive values (+) indicate yellow

Thus, each colour can be represented by a unique combination of L*, a*, and b* values.

3.3.5. Measure colour difference with ΔE

To quantify the difference between two colours, the **ΔE_{76}(Delta E)** formula is used:[138]

$$\Delta E_{76} = \sqrt{(L_2^* - L_1^*)^2 + (a_2^* - a_1^*)^2 + (b_2^* - b_1^*)^2}$$

This formula dates from 1976. This formula is used in ISO procedures like 12647-2 for process control in halftone separations, proofing, and printing, allowing better communication of colour differences under standard conditions.[140] The equation used the

Lab* colour space, but it had limitations. It overlooked the fact that the human eye is more sensitive to colour differences in some areas of the colour wheel than others. As a result, a ΔE of 1.0 could represent a small difference in dark blues but a large difference in light pastels. This led to the release of a more precise formula in the year 2000, which is noted $\boldsymbol{\Delta E}_{2000}$ (Figure 17). The new formula introduces correction terms to consider the human perceives differences in the chroma C* (color saturation), lightness L, and hue h* (type of color; red, yellow, green, and so on). The formula is detailed below , and the full explanation can be found in this paper by G. Sharma *et al.*[141] In this formula, $K_L = K_C = K_H = 1.0$ by default. ΔL', ΔC' and Δh' are differences in lightness, chroma and hues, respectively. The improvement of $\boldsymbol{\Delta E}_{2000}$ formula includes parametric weighting functions for L, C, and H, respectively, such as S_L, S_c, and S_H, a rotation term R_T accounts for interaction between chroma and hue differences, particularly important for colours near the blue region. It also includes hue adjustments. This colour difference method will be used in this thesis.

$$\Delta E_{00}^{*} = \sqrt{\left(\frac{\Delta L'}{k_L S_L}\right)^2 + \left(\frac{\Delta C'}{k_C S_C}\right)^2 + \left(\frac{\Delta H'}{k_H S_H}\right)^2 + R_T \frac{\Delta C'}{k_C S_C}\frac{\Delta H'}{k_H S_H}}$$

$$\Delta L' = L_2^* - L_1^*$$

$$\bar{L} = \frac{L_1^* + L_2^*}{2} \quad \bar{C} = \frac{C_1^* + C_2^*}{2}$$

$$a_1' = a_1^* + \frac{a_1^*}{2}\left(1 - \sqrt{\frac{\bar{C}^7}{\bar{C}^7 + 25^7}}\right) \quad a_2' = a_2^* + \frac{a_2^*}{2}\left(1 - \sqrt{\frac{\bar{C}^7}{\bar{C}^7 + 25^7}}\right)$$

$$\bar{C}' = \frac{C_1' + C_2'}{2} \text{ and } \Delta C' = C_2' - C_1' \quad \text{where } C_1' = \sqrt{a_1'^2 + b_1^{*2}} \quad C_2' = \sqrt{a_2'^2 + b_2^{*2}}$$

$$h_1' = \operatorname{atan2}(b_1^*, a_1') \mod 360^\circ, \quad h_2' = \operatorname{atan2}(b_2^*, a_2') \mod 360^\circ$$

$$\Delta h' = \begin{cases} h_2' - h_1' & |h_1' - h_2'| \le 180^\circ \\ h_2' - h_1' + 360^\circ & |h_1' - h_2'| > 180^\circ, h_2' \le h_1' \\ h_2' - h_1' - 360^\circ & |h_1' - h_2'| > 180^\circ, h_2' > h_1' \end{cases}$$

$$\Delta H' = 2\sqrt{C_1' C_2'}\sin(\Delta h'/2), \quad \bar{H}' = \begin{cases} (h_1' + h_2' + 360^\circ)/2 & |h_1' - h_2'| > 180^\circ \\ (h_1' + h_2')/2 & |h_1' - h_2'| \le 180^\circ \end{cases}$$

$$T = 1 - 0.17\cos(\bar{H}' - 30^\circ) + 0.24\cos(2\bar{H}') + 0.32\cos(3\bar{H}' + 6^\circ) - 0.20\cos(4\bar{H}' - 63^\circ)$$

$$S_L = 1 + \frac{0.015\left(\bar{L} - 50\right)^2}{\sqrt{20 + \left(\bar{L} - 50\right)^2}} \quad S_C = 1 + 0.045\bar{C}' \quad S_H = 1 + 0.015\bar{C}'T$$

$$R_T = -2\sqrt{\frac{\bar{C}'^7}{\bar{C}'^7 + 25^7}}\sin\left[60^\circ \cdot \exp\left(-\left[\frac{\bar{H}' - 275^\circ}{25^\circ}\right]^2\right)\right]$$

Figure 17. ΔE_{2000} formula to measure colour difference between two points. [141]

The interpretation of colour difference varies by field. In cultural heritage, the following ΔE ranking is often used:[137,142]

- ≤ 1.0: Not perceptible by the human eye
- 1–2: Perceptible through close observation

• 2–10: Perceptible at a glance

• 11–49: Colours are more similar than the opposite

• 100: Colours are exactly the opposite

3.3.6. Converting UV-VIS spectra to CIELAB coordinates

Human vision relies on two types of photoreceptors located in the retina: rods and cones. Rods are primarily responsible for vision under low-light or dark conditions, while cones enable vision in bright light. More specifically, there are three types of cones, S, M, and L, each sensitive to different wavelengths of light: short (S), medium (M), and long (L), respectively. These cones play a crucial role in colour perception; when stimulated by light, they send signals that are processed by the eye and interpreted by the brain, allowing us to perceive colour. The goal is to quantify colour perception numerically, which is the object of colourimetry, enabling the objective measurement and representation of colour. [137,141] The CIELAB colour coordinates integrate this human perception of colour and the human sensitivity difference for each colour.

To convert absorbance spectra into CIE Lab (L*a*b*) colour coordinates, the first step is to transform the absorbance data into transmittance or reflectance, depending on the context. This is typically done using the relation $T(\lambda)=10\text{-}A(\lambda)$. The resulting spectral data is then multiplied by a standard illuminant spectrum (e.g., D65 for daylight) and the CIE 1931 colour matching functions to obtain the tristimulus values X, Y, and Z. The colour appearances of a given object (in our case the molecules) under a light source (i.e., the illuminant), it is derived from the tristimulus values as follow:

$$\bar{X} = \int T(\lambda)\ S(\lambda)\,\bar{x}(\lambda)\,d\lambda$$

$$\bar{Y} = \int T(\lambda)\ S(\lambda)\,\bar{y}(\lambda)\,d\lambda$$

$$\bar{Z} = \int T(\lambda)\ S(\lambda)\,\bar{z}(\lambda)\,d\lambda$$

$$k = 100 \Big/ \int S(\lambda)\,\bar{y}(\lambda)\ \ d$$

Where $S(\lambda)$ is the spectral power distribution of the Illuminant, $T(\lambda)$is the transmission spectrum of the object and k is a normalising constant. These integrals approximate how the human eye perceives colour under a given light source. Once the XYZ values are calculated, they are converted into the CIE Lab coordinates using a standard nonlinear transformation that references a defined white point.

$$L* = 116(Y/Yn)1/3 - 16 if Y/Yn > 0.00856$$

$$L* = 903.3(Y/Yn) \quad if\ Y/Yn \leq 0.00856$$

$$a* = 500[f(X/Xn) - f(Y/Yn)]$$

$$a* = 200[f(Y/Yn) - f(Z/Zn)]$$

$$with\ f(x) = 1/3\ for\ x > 0.00856\ and\ f(x) = 7.787 + 16/116\ for\ x \leq 0.00856$$

The terms Xn, Yn, and Zn are the nominally white object colour stimulus (i.e., the reference white).

3.4. Photodegradation

To study the photodegradation of dyes, accelerated photodegradation experiments were conducted in the laboratory, as natural photodegradation is typically a slow process. This approach is based on the reciprocity theorem, which states that the extent of photodegradation is equivalent when using high light intensity for a short duration or low intensity for a longer period. [75] Several light sources were used in this study, including a UVA LED lamp (Hamamatsu, irradiance: 10 mW/cm^2), a visible LED lamp, and a xenon lamp housed in a controlled chamber simulating daylight. The spectral output of each lamp is shown in *section 3.4*. Dyes were solubilized in water at a concentration of 10^{-4} mol. L^{-1} and distributed into glass vials. These solutions were then exposed to UVA, visible, or xenon light under various durations and pH conditions described below. Absorbance spectra were recorded immediately after exposure. A non-exposed control solution was used for comparison and is labelled as "0 h" in the results section. In addition, the degradation products of alizarin were analysed by HPLC-UV-MS to identify potential photoproducts. The Xenon lamp is SUNTEST XLS+ with Daylight option, the unit was operated with constant cooling, and the intensity of the lamp is 50 mW/cm^2. Visible lamp is a polychromatic white LED TSPOT4-WHI-9-DS from TPL Vision (measured irradiance of about 10 mW/cm^2). Finally, the UVA lamp is a UV-LED L11921-410 from Hamamatsu Photonics with a maximal emission at 369 nm and a measured irradiance of about 10 mW/cm^2.

The exposition was followed by UV-VIS spectra every hour for 5 hours or every 24h until 72h. A 5-hour exposure at 50,000 lux is equivalent to approximately 0.43 years (approximately 5 months) of natural ageing under typical museum lighting conditions at 200 lux.

Alizarin photodegradation (Sigma Aldrich, 97%) was done in water and at pH=12 because it is the most soluble at this pH at 10^{-4} mol. L^{-1}. To study the influence of photodegradation kinetics on pH, we used an alizarin alternative, alizarin red S (sodium alizarinsulfonate, Thermoficher), for its higher water-solubility at different pH (2, 6.5, and 12) adjusted with NH_4OH or HCl and with a concentration of 10^{-4} mol. L^{-1}. Three replicates of each kinetic experiment were done, and the error deviation was calculated. The influence of solvent was also tested for alizarin in toluene, acetone, ethanol, DMSO, DMF, 2-propanol, and glycerol (solvent grade). The number of photons per cm^2 as a function of irradiance is shown in Table 1. It can be used to normalise the experiments with different light intensities. An example of the calculation of the number of photons for 10mW/cm^2 at λ=550 nm is presented below:

$$E_{\text{photon}} = \frac{hc}{\lambda} = \frac{6.626 \times 10^{-34}\, J{\cdot}s \times 3.0 \times 10^{8}\, m/s}{520 \times 10^{-9}\, m} = 3.825 \times 10^{-19}\, J$$

$$E_{\text{total}} = 10\,\text{mW/cm}^2 = 0.01\,\text{W/cm}^2$$

$$N_{\text{photons}} = \frac{E_{\text{total}}}{E_{\text{photon}}} = \frac{0.01\,\text{W}}{3.825 \times 10^{-19}\,\text{J}} \approx 2.61 \times 10^{16}\ \text{photons/s}$$

N_{photons} for 1h exposition: $\approx 2{,}61 \times 10^{16}$ photons/s *3600s= $9.40^{*}\ 10^{19}$photons/cm^2

Table 2. Photons absorbed as a function of irradiance and wavelength.

Wavelength (nm)	Irradiance	$N_{\text{photons}}/\text{cm}^2$ 1h exposition
520	50mW/cm^2 (xenon)	4.72×10^{20} photons/cm^2
520	10mW/cm^2 (LED)	9.40×10^{19} photons/cm^2
430	50mW/cm^2 (xenon)	7.80×10^{17}photons/cm^2
430	10mW/cm^2 (LED)	1.56×10^{17}photons/cm^2

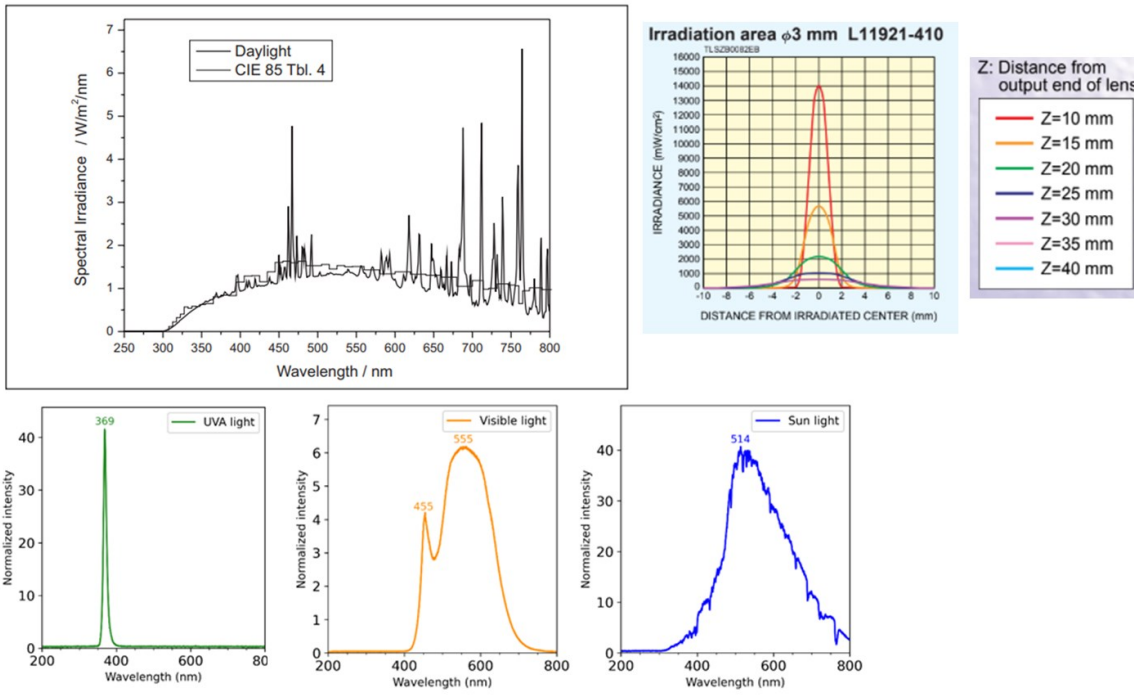


Figure 18. Spectral distribution of daylight Xenon Chamber SUNTEST XLS+, UVA lamp, and visible lamp used in this thesis. The sunlight is shown for reference. Spectra are taken from the furnisher or the S.I. of the article https://doi.org/10.1016/j.saa.2024.125371.

CHAPTER 4

CHAPTER 4 | COMPUTATIONAL METHODS 67

4.1. The foundations of quantum mechanics 68

4.2. Computational chemistry methods 71
4.2.1. Ab initio methods 71
4.2.2. DFT methods 72
4.2.3. Semiempirical methods 72
4.2.4. Computing electronic molecular energy 72
4.2.5. Hartree-Fock method: a brief overview 72
4.2.6. The Roothaan equation 73
4.2.7. SCF procedure 74
4.2.8. Basis sets 75
4.2.9. Electron correlation 77

4.3. Density Functional Theory (DFT) 78
4.3.1. The exchange-correlation challenge and functionals 79
4.3.2. Hybrid functional 81
4.3.3. Double-hybrid functionals 81
4.3.4. Molecular properties and gradient method 82
4.3.5. Energy derivatives and Hessian matrix 82

4.4. Time-Dependent DFT (TD-DFT) 82
4.4.2. Linear response theory 84

4.5. Absorbance spectra simulation 84
4.5.1. Oscillator Strength (or Oscillator force) 85
4.5.2. Spectra intensity with vibrational effects 85
4.5.3. Vibronic couplings 88

4.6. Modelling the environment 95
4.6.1. Vacuo 95
4.6.2. Implicit model 95
4.6.3. Microsolvation model 97
4.6.4. Explicit solvation 98

This chapter presents the computational methods used in this thesis, focusing on quantum chemistry approaches for modelling madder dyes. Key topics include DFT, TD-DFT, absorbance (vibronic) spectra simulation, and strategies for representing the molecular environment.

CHAPTER 4 | COMPUTATIONAL METHODS



Computational chemists employ the principles of quantum mechanics to theoretically investigate chemical reactivity and to interpret the full range of spectroscopic techniques.[103] They use computer simulations to solve chemical problems. It has been used in a wide range of fields such as drug design in medicine, solar energy harvesting, design catalytic process, but is not much used in the field of cultural heritage.[143]

Our ambition is to understand the colourant's molecular structure, their properties and how they react with the environment (such as light, pH, and solvent). To a larger extent, this means understanding "what the electrons are doing", because it is the electrons which are forming bonds, the electrons which are rearranged when a reaction occurs, and a change in molecular 'energy' is in part responsible for the absorption spectra. Those behaviours can be explained and predicted using quantum mechanics. This is a key to a deeper understanding of all phenomena in chemistry.

Indeed, quantum chemistry is built on a set of postulates and equations that describe the behaviour of atoms and molecules. By applying these principles and resulting equations, it becomes possible to predict and interpret key molecular properties, such as molecular geometry, electronic structure, energy levels, and therefore spectroscopic signatures, which is one of the focuses of this thesis, because it offers a non-invasive way to study colours in precious artworks. Moreover, today, advances in experimental spectroscopic techniques have led to increasingly sophisticated methods, and the combination of multiple modern techniques can provide a more complete picture of the system studied. However, the resulting experimental high-resolution spectra may consist of hundreds or thousands of peaks, making their interpretation and analysis highly challenging. With the help of quantum chemistry, it can predict spectral properties. Moreover, it can aid in the interpretation of complex spectra and filling gaps in data that are either incomplete or experimentally difficult to access. It can also be used to predict molecular properties before synthesis, reducing the cost and formulation time based on trials and errors.

This section introduces the fundamental principles used in quantum chemistry. It begins by defining the key terms used and the fundamental postulates of quantum mechanics. We then present the Schrödinger equation and the Self-Consistent Field (SCF) procedure. Then we define the concept of basis sets. Along the way, we explore the Density Functional Theory (DFT) and its Time-Dependent (TD-DFT) counterparts' methods, which are the basis of the theoretical methods used in this thesis. The remaining parts build on them and explain how they are applied to solve chemical problems: to extract molecular properties. We will particularly focus on the prediction of absorption spectra with vibrational effects and the Franck-Condon approximation.

These sections and equations are inspired by the books by Szabo *et al.*, Peter Atkins *et al.*, J.Keeler *et al.* as well as publications by the ORCA software authors and manual. [128,144–146]

4.1. The foundations of quantum mechanics

In the world of very small objects, particles like electrons behave in different ways than objects in our everyday world, which is governed by classical mechanics (also called Newton mechanics). On the other hand, the electrons are described by quantum mechanics, which follows some postulates.

One of the most important results from quantum mechanics is that the energy levels of particles like the electron is *quantised*. This means that energy cannot take any values and is restricted to certain values called *energy levels*. The second important idea is that we cannot tell or measure precisely the position of an electron. All we can do is give the probability of that electron being in a particular position. Both quantisation of energy and the idea of the probability of a position are closely related to the notion of the *wavefunction*.

a) Wavefunction

In quantum mechanics, the wavefunction contains everything there is to know about the particle. If we know the wavefunction, we can extract properties like energy and the probability of finding the electron. It is a mathematical function of relevant variables such as the electrons' position coordinates (such as vector $\mathbf{r_1}=(x_1,y_1,z_1)$ for electron 1, $\mathbf{r_2}=(x_2,y_2,z_2)$ for electron 2, and so on). It is usually denoted with the Greek letter 'psi', Ψ.

b) Probability density

The probability of finding an electron at position $\mathbf{r}=(x,y,z)$, in a small volume dV, is defined as the square of the module of the wavefunction: $dP= |\Psi(x,y,z)|^2 dV = |\Psi(\mathbf{r})|^2 dV$.

The probability of finding an electron in all space is certain, so the integral of dP over the volume is equal to 1. This is why the wavefunctions are normalised by a constant.

The probability density is the previous probability divided by the volume dV.

The wavefunction can be complex, so the probability density can be written as $\Psi*(\mathbf{r})\Psi(\mathbf{r})dV$.

c) Energy

Transition between different levels of energy gives rise to different spectra. Energy comes from many forms, but the two we are most concerned with are kinetic and potential energy. Kinetic energy is the energy related to motion. Potential energy is the 'energy stored' and is present in diverse forms: gravitational potential energy, elastic potential energy in a spring, for example, or charged particles, there is the electrostatic potential energy. This last one is considered for electrons.

d) Introduction to Schrödinger equation

Now that we have introduced the key concepts, such as energy and wavefunctions, the question is, how do we find or compute them?

Quantum mechanics tells us that they can be found by solving the Schrödinger equation. This equation depends on the system we are studying. Its time-dependent (*1*) and time-independent (*2*) forms can be written, as:

$$\hat{H}\Psi(r,t) = i\hbar\frac{\partial}{\partial t}\Psi(r,t) \qquad (1)$$

$$\hat{H}\Psi = E\Psi \qquad (2)$$

where $\Psi(\mathbf{r},t)$is the time-dependent wavefunction of the system, $\widehat{H}$ is the total Hamiltonian operator, $\mathbf{r} = \mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3}...$ the position coordinates all electrons, and $\hbar$ is the reduced Planck constant, E is the energy. The details of this equation and the notion of operator are explained in the following sections.

This equation can be solved exactly in simple cases, such as for example the hydrogen atom. Calculating the eigenvalues E of this equation gives access to energy levels, and the eigenvector gives the wavefunction Ψ. But it cannot be solved exactly for atoms or molecules that contain more than one electron. We have to resort to approximations and computer-based solutions as the mathematics becomes rapidly complicated. For big molecules like alizarin and other anthraquinones, a powerful computer is needed, such as a supercomputer. In this thesis, we use the MESU supercomputer from Sorbonne University and the Jean Zay French national supercomputer.

e) Operators

An observable is any variable that can be measured, such as position, momentum, energy, or spin. In classical mechanics, they are represented by functions, for example, the function position as a function of time. In quantum mechanics, they are represented by operators. In mathematics, an operator is a symbol for an instruction to carry out an action such as an operation $\frac{df}{dx}$. The operators are often represented by the symbol 'uppercase omega', $\widehat{\Omega}$, associated with the observable Ω. But other letters with a hat can be used, such as $\hat{H}, \hat{V}, \hat{T}$, $\hat{f}$, etc. The Hamiltonian operator $\hat{H}$, potential $\hat{V}$, or kinetic operators $\hat{T}$ are detailed in the following section.

Before delving into more complex equations for the approximation of the solution, we need to first put some postulates

f) Postulates of quantum mechanics

Like many theories, quantum mechanics is based on postulates. A postulate is a theoretical statement that cannot be proved but is consistent with experimental results. They may seem rather distant from the application to the molecules, which we are interested in, but it serves as a better comprehension about what those theories are based on.

The postulates in quantum mechanics can be formulated simply as below: [145]

1) The state of a quantum mechanical system is completely described by the wavefunction $\Psi(\mathbf{r}, t)$ that depends on the coordinates of the particle, $\mathbf{r}$ and the time, t.

2) For a system described by the wavefunction $\Psi(\mathbf{r})$, the probability of finding this particle in the volume dV is $|\Psi(\mathbf{r})|^2$ dV.

3) To every observable in classical mechanics there corresponds a linear, Hermitian operator in quantum mechanics.

4) In any measurement of the observable associated with operator $\hat{\Omega}$ the only values that will ever be observed are the eigenvalue, ω, that satisfy the eigenvalue equation: $\hat{\Omega}\Psi = \omega\Psi$

5) When the value of an observable Ω is measured for a system that is described by a linear combination of eigenfunctions of $\hat{\Omega}$, with coefficients c_k, each measurement gives one of the eigenvalues ω_k of $\hat{\Omega}$, with a probability proportional to $|c_k|^2$.

g) The Hamiltonian operator

For the remaining parts, atomic units (a.u.) will be used, in which fundamental constants such as $(\hbar)$, (m_e), and (e) are set to 1, simplifying the mathematical expressions.

The Hamiltonian $\hat{H}$ represents the total energy of the system and includes several terms: the kinetic energy of electrons $\widehat{T_e}$, the kinetic energy of nuclei $\widehat{T_N}$, the electron-nucleus interactions $\widehat{V_{Ne}}$, nucleus-nucleus repulsions $\widehat{V_{NN}}$, electron-electron repulsions $\widehat{V_{ee}}$, and, if applicable, an external potential $\widehat{V_{ext}}$.

$$\hat{H} = \widehat{T_e} + \widehat{T_N} + \widehat{V_{Ne}} + \widehat{V_{NN}} + \widehat{V_{ee}} + \widehat{V_{ext}} \qquad (3)$$

In Equation (3) above, $(\widehat{T_e})$ and $(\widehat{T_N})$ represent the kinetic energy operators for the electrons and nuclei, respectively. These are defined as (in a.u.):

$$\widehat{T_e} = -\sum_{i=1}^{N_e} \frac{1}{2}\nabla_i^2 \;\; and \;\; \widehat{T_N} = -\sum_{A=1}^{N_n} \frac{1}{2M_A}\nabla_A^2 \qquad (4)$$

Where ∇ is the nabla operator, (with partial derivatives). For this equation and throughout the remainder of this manuscript, the system under consideration will consist of (N_e) electrons, each with mass (m_e) and charge $(-e)$ where e denotes the elementary electric charge, and (N_n) nuclei, each with mass (M_A) and charge $(Z_A e)$. The spatial coordinates of the electrons are denoted by $(\boldsymbol{r}_i)$ and those of the nuclei by $(\boldsymbol{R}_A)$.

The potential energy of attraction operators $(\widehat{V_{Ne}})$, $(\widehat{V_{NN}})$, and $(\widehat{V_{ee}})$ correspond to the nucleus–electron, nucleus–nucleus, and electron–electron Coulomb interactions, respectively. They are expressed as (in a.u.):

$$\widehat{V_{Ne}} = -j_o \sum_{i=1}^{N_e}\sum_{A=1}^{N_n} \frac{Z_A}{|r_i - R_A|}, \; \widehat{V_{NN}} = \sum_{A=1}^{N_n}\sum_{B>A}^{N_n} \frac{Z_A Z_B}{|R_A - R_B|}, \; \widehat{V_{ee}} = j_o \sum_{i=1}^{N_e}\sum_{j>i}^{N_e} \frac{1}{|r_i - r_j|} \qquad (5)$$

Where ($|r_i - R_A|$) is the distance between electron (i) and nucleus (A), ($|R_A - R_B|$) is the distance between nuclei (A) and (B), and ($|r_i - r_j|$) is the distance between electrons (i) and (j) , $j_o = e^2/4\pi\varepsilon_o$ is a constant encountered in computational chemistry. The final term, $(\widehat{V_{ext}})$, represents the interaction potential between the system (electrons and nuclei) and any external field, such as an applied electric or magnetic field. If this term is equal to zero, the system is considered isolated, and the Hamiltonian becomes time-independent.

h) The Born-Oppenheimer approximation

As stated before, the exact solution of the Schrödinger equation (Eq. 1) of molecules cannot be solved, except for the Hydrogen atom. Therefore, all theories about molecular structures make some simplifications and approximations to solve this equation. Here we adopt the Born-Oppenheimer approximation, which supposes that the nuclei, being much heavier than the electrons, move relatively very slowly and can be treated as stationary, while the electron moves. This simplifies the Schrödinger equation by treating the nuclear kinetic energy term $(\widehat{T_N})$ and the nuclear-nuclear repulsion term $(\widehat{V_{NN}})$ as constants. The Hamiltonian operator is then reduced to electronic Hamiltonian: $\widehat{H_e} = \widehat{T_e} + \widehat{V_{Ne}} + \widehat{V_{ee}}$.

4.2. Computational chemistry methods

Computational methods use the power of computers to predict molecular structure and reactivity, as well as spectroscopic properties. This section introduces some of the computational techniques to solve the Schrödinger equation. We present the equations that are used and describe some of the approximations that make the computation feasible, as well as the different methods used.

The starting point is the Born-Oppenheimer approximation and the focus on finding the solution of the electronic Schrödinger equation: $\hat{\mathrm{H}}\,\Psi = \mathrm{E}\,\Psi$. (2)

There are two main approaches for the solution of the Schrödinger equation: *ab-initio* and semi-empirical method. The Density Functional Theory (DFT) has become a dominant *ab initio* procedure. [103,106,147]DFT methods are the foundation of the methods used in this thesis.

4.2.1. Ab initio methods

In this method, a model is chosen for the electronic wavefunction and equation (*2*) is solved using as input only the values of the fundamental constants and the atomic numbers of the nuclei (the term *ab initio* means 'from the beginning'). Its accuracy depends on the model chosen for the wavefunction. For large molecules, accurate *ab initio* is computationally expensive and semi-empirical methods have been developed to study those molecules.

4.2.2. DFT methods

Contrary to other methods which relies on wavefunctions, Density Functional Theory relies on the electron density ρ. This drastically reduces the computational complexity and still capturing essential quantum behaviour. It is now the most used computational procedure for molecular electronic structure calculation. This method is fully detailed in a following subsection.

4.2.3. Semiempirical methods

This method uses a simplified form of the Hamiltonian, approximations and adjustable parameters, derived from experimental data to improve accuracy and reduce computation time. They are mixed between *ab initio* methods (which rely only on theory) and empirical methods (based entirely on data). One example of semiempirical method is extended Tight Binding (xTB), developed by Grimme and co-workers. Commonly used versions are GFN-xTB, GFN2-xTB. xTB method is used in this thesis for molecular quantum dynamics and conformers generation.

4.2.4. Computing electronic molecular energy

We will first begin by computing the electronic energy. From that, many other properties can be determined like the potential energy surface (PES), and from that we can find the equilibrium structure of a molecule. Vibrational frequencies can be determined from gradients or PES, and from that we can extract spectroscopic properties.

To compute molecular electronic energy and geometry optimisation necessary before all computations, the self-consistent field (SCF) method is the starting point of many of the ab initio molecular methods used in this thesis. Therefore, we introduce it here briefly. We begin by first presenting the Hartree-Fock approach and basis set before the SCF iteration procedure. Then we will move on to the DFT and TD-DFT methods along with vibronic spectra methods.

4.2.5. Hartree-Fock method: a brief overview

The electronic wavefunction of many-electron molecules $\Psi(\mathbf{r_1}, \mathbf{r_2}, \dots)$ is a function of the positions of all the electrons. In the orbital approximation, it is supposed that each electron occupies an orbital, allowing the wavefunction to be rewritten as $\phi_1(\mathbf{r_1})\phi_2(\mathbf{r_2})\dots$ or more simply $\phi_1(\mathbf{1})\phi_2(\mathbf{2})\dots$ Here we suppose that electron 1 occupies orbital ϕ_1 with a spin α, electron 2 occupies the same orbital with a spin β, and so on. Hence, the overall wavefunction Ψ can be written as a product: $\Psi = \phi_1^{\alpha}(\mathbf{1})\phi_2^{\beta}(\mathbf{2})\dots$ where each spinorbital such as $\phi_1^{\alpha}(\mathbf{1})$ is a product of the spatial wavefunction and the spin state, for example: $\phi_1^{\alpha}(\mathbf{1}) = \phi_1(\mathbf{1})\alpha(1)$. To ensure that the wavefunction follows the Pauli principle, it can be written as a sum of all possible permutations. It can be formulated as a Slater determinant[148]:

$$\Psi(1,2,\dots,N_e) = \frac{1}{\sqrt{N_e!}}\begin{vmatrix} \phi_1^{\sigma_1}(r_1) & \phi_2^{\sigma_2}(r_1) & \dots & \phi_N^{\sigma_N}(r_1) \\ \phi_1^{\sigma_1}(r_2) & \phi_2^{\sigma_2}(r_2) & \dots & \phi_N^{\sigma_N}(r_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1^{\sigma_1}(r_{N_e}) & \phi_2^{\sigma_2}(r_{N_e}) & \dots & \phi_N^{\sigma_N}(r_{N_e}) \end{vmatrix}, with\ \sigma_i \in \{\alpha,\beta\} \quad (6)$$

The $\frac{1}{\sqrt{N_e!}}$ ensures the wavefunction is normalised.

Following the variational principle, the form of Ψ is the one with the lowest energy, as the spin-orbitals ϕ are varied. D.R. Hartree and V. Fock showed that the minimisation of the energies of each ϕ leads to:

$$\hat{\mathrm{f}}(i)|\phi_{\mathrm{k}}\rangle = \epsilon_{\mathrm{k}}|\phi_{\mathrm{k}}\rangle \quad (7)$$

Where $\hat{\mathrm{f}}(i)$ is the Fock operator, which is defined in terms of Coulomb operator $\widehat{J_l}$, and exchange operator, $\widehat{K_l}$:

$$\hat{\mathrm{f}}(i) = -\frac{\nabla_i^2}{2} - \sum_A^M \frac{Z_A}{|r_i - R_A|} + \sum_l^N [\widehat{J_l}(i) - \widehat{K_l}(i)] \quad (8)$$

$$\widehat{J_l}(i)\ \phi_k(i) = \frac{\int \phi_{\mathrm{k}}(i)\phi_l^*(i')\phi_l(i')dV'}{|r_i - r_{i'}|} \quad (9)$$

$$\widehat{K_l}(i)\ \phi_{\mathrm{k}}(\mathrm{i}) = \frac{\int \phi_l(i)\phi_l^*(i')\phi_k(i')\mathrm{dV}'}{|r_i - r_{i'}|} \quad (10)$$

The Coulomb operator considers the Coulombic repulsion between the electrons, and the exchange operator represents the modification of this energy that can be attributed to the effects of spin correlation. The sum in (8) represents the average potential energy of electron 1 due to the presence of other N_e -1 electrons. Each molecular orbital must be obtained by solving equation (7) by adopting an iterative type of solution, it starts by guessing the initial form of the wavefunction ϕ_l and use them to defince $\widehat{J_l}$ and $\widehat{K_l}$. Then the HF equation (7) are solved. This process is iterated using newly found wavefunctions until the cycle of calculation leaves the energies ϵ_{k} and wavefunctions ϕ_{k} unchanged within a chosen criterion. This is the origin of the term Self-Consistent Field (SCF) in general and of HF-SCF for the approach based on the orbital approximation.

4.2.6. The Roothaan equation

The difficulty of HF-SCF procedure is that its numerical solution is a complex task, even for powerful computer. Therefore, a modification must be used to be applied for large molecules. In 1951, Roothaan developed LCAO-MO (Linear Combination of Atomic Orbital for expressing Molecular Orbital). [149] He found a way to convert the HF equations for the molecular orbitals into equations that appear in LCAO to simulate molecular orbital. He wrote:

$$\phi_k = \sum_{o=1}^{N_b} c_{om}\chi_o \tag{11}$$

Where c_{om} are unknown coefficents and χ_o are atomic orbitals. N_b is the number of basis sets, which corresponds to the number of atomic orbitals (such as 1s, 2s, 2px, 2py, 2pz, etc.) This formalism leads to a set of simultaneaous equations for the coefficients called the Roothaan equations.

This led to equation:

$$\hat{\mathrm{f}}(i)\sum_{o=1}^{N_b} c_{om}\chi_o = \epsilon_{\mathrm{k}}\sum_{o=1}^{N_b} c_{om}\chi_o \tag{12}$$

These equations can be reformulated as the following form:

$$\sum_{\nu=1}^{K} F_{\mu\nu}C_{\nu k} = \varepsilon_k \sum_{\nu=1}^{K} S_{\mu\nu}C_{\nu k} \tag{13}$$

With the overlap integral $S_{\mu\nu}$ and Fock integral $F_{\mu\nu}$ defined as:

$$S_{\mu\nu} = \langle \chi_\mu \mid \chi_\nu \rangle \tag{14}$$

$$F_{\mu\nu} = \langle \chi_\mu | \hat{\mathrm{f}} | \chi_\nu \rangle \tag{15}$$

4.2.7. SCF procedure

To solve equation (13), which could be written as its matrix form FC=εSC, a iterative procedure is used because the Fock operator depends on the coefficients c_{om}. The solutions are non-trivial only if the following equation is satisfied:

$$det|F - \varepsilon S| = 0 \tag{16}$$

This is the SCF procedure and is summarised in the illustration below.

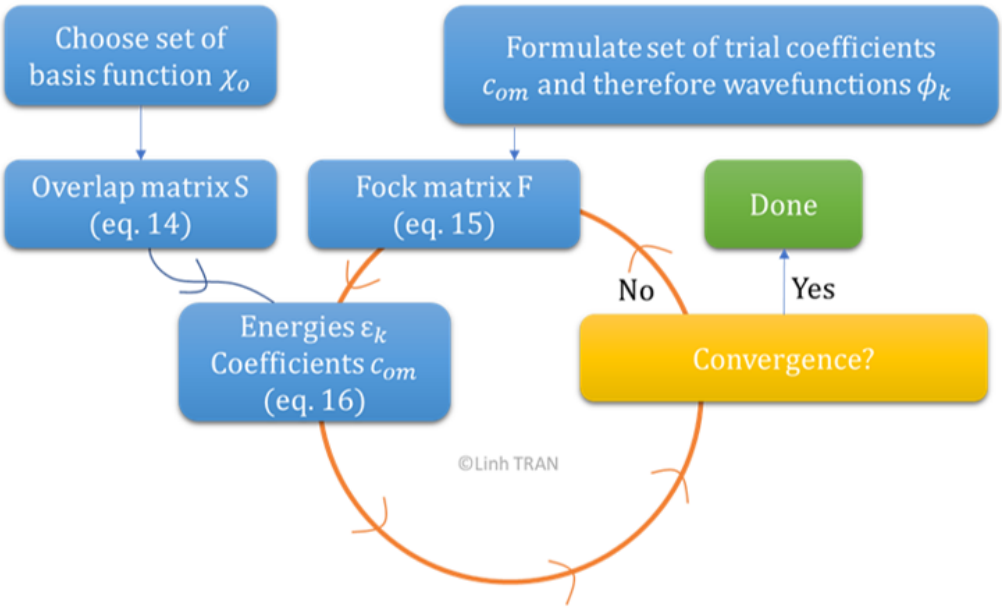


Figure 19. Illustration of the iteration procedure for Hartree-Fock SCF calculation.

4.2.8. Basis sets

In theory, a complete basis function (an infinite number) must be used to represent the molecular orbital exactly. However, this is not computationally feasible and a finite number of basis sets is always used. The error due to the incompleteness of the basis is the basis-set truncation error. The goal is to choose the basis set effectively to have a low number of basis sets but also a small truncation error. There are many types of orbitals. One of the first choices of orbitals is the Slater-type Orbitals (STO), and then later came the Gaussian-type orbitals (GTO), which are used in this thesis. The Slater orbital centred on each of the atomic nuclei has the following form:

$$\psi_{n,l,m}(r,\theta,\phi) = N \cdot r^{n-1} e^{-\zeta r} Y_l^m(\theta,\phi) \qquad (17)$$

With N the normalisation constant, r, θ, ϕ, the spherical coordinates, (n, l, m) the quantum numbers ζ ("zeta") is the Slater constant, an $Y_l^m(\theta,\phi)$ is the spherical harmonics. The use of STO in HF-SCF calculations on molecules is computationally heavy.

The introduction of the GTO by Samuel Francis Boys in 1950 has played a major role in making ab initio computations feasible.[150] Cartesian Gaussians are functions of the form:

$$g_{ijk} = N x^i y^j z^k e^{-\alpha r^c} \qquad (18)$$

Where the origin of the coordinate is the nucleus of atoms, i, j and k are non-negative integers, N is a normalisation constant, and α is a positive exponent. A s-type GTO has i=k=0; a p-type GTO has i+j+k=1, a d-ype GTO has i+j+k=2, etc. (see Figure 20).

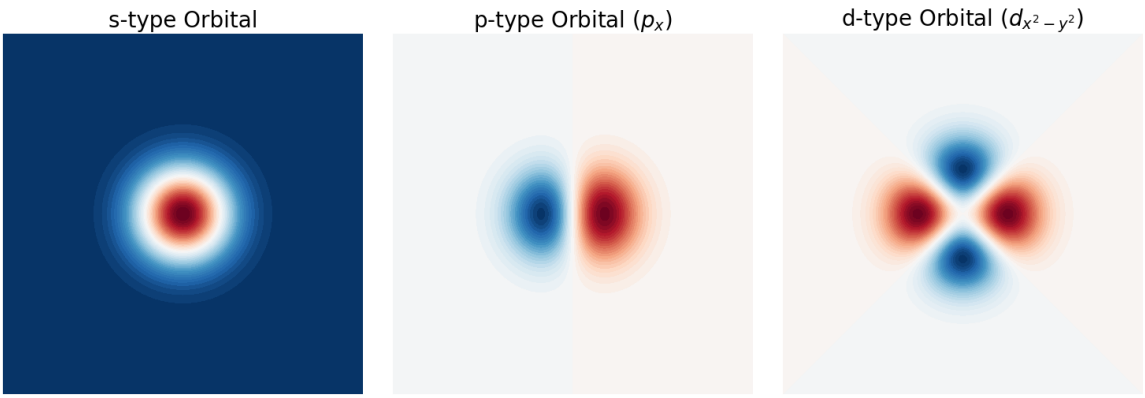


Figure 20. Illustration of contour plot of Gaussian-Type Orbitals (GTO): s-type, p-type and d-type. Red means positive, and blue is negative.

The central advantage of Gaussians is that the product of two GTOs at different centres is equivalent to a single Gaussian function centred at a point between the two centres. Therefore, two-electron integrals on three and four different atomic centres can be reduced to integrals over two different centres, which is easier to compute.

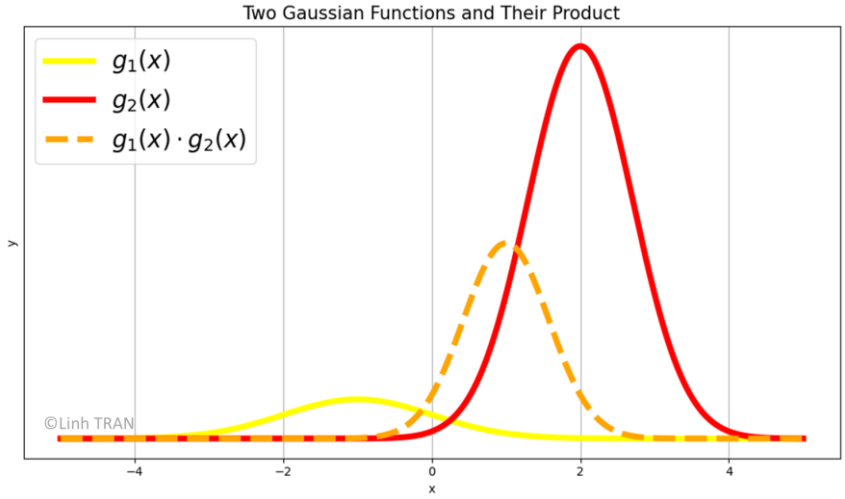


Figure 21. Illustration for the product of two gaussians, situated at the centre of the two contributing gaussians. The scale of the product has been increased for better visualisation.

However, GTOs give a poorer representation of the orbitals at the atomic nuclei; a larger basis set must be used to achieve higher accuracy. This is done by grouping together several GTOs, which gives contracted Gaussian functions, centred on the same atomic nucleus. The molecular orbitals are then expressed as a combination of the contracted Gaussians. The use of this type of contracted functions rather than the primitive ones decreases the number of unknown coefficients and saves computer time without a loss of accuracy if they are well chosen.

Improvements in the agreement between electronic structure calculations and experiment can often be achieved by increasing the size of the basis set.

a) N-zeta basis sets

In a double-zeta (DZ) basis set, each basis function in the minimal basis is replaced by two functions; in a triple-zeta (TZ) basis set, by three functions, and so on. For example, a double-zeta basis function for H_2O consists of 14 functions: 2 basis sets for each 1s of Hydrogen orbitals (total is 4), 2 basis sets for each orbital of the Oxygen, 1s, 2s, 2px, 2py, and 2pz (with a total of 10).

b) Split-Valence (SV) basis set

In a Split-Valence (SV) basis set, each inner-shell (core) atomic orbital is represented by one basis set function and each valence atomic orbital by two basis set functions. Logically, the valence electrons are more responsible for the molecules' reactivity and properties, so they are more precisely described. It means that for C atoms, for example, there is one contracted function representing the 1s orbital and two representing the 2s orbital, and two for each 2p orbital.

c) Polarisation functions in basis

Further improvements to the accuracy of the electronic structure calculations can often be achieved by including polarisation functions in the basis. For instance, the polarisation function in a calculation of CH_4 includes basis functions representing d orbitals on carbon or p orbitals on hydrogen. Polarisation functions often lead to improved results because atomic orbitals are distorted (or polarised) by adjacent atoms when bonds form in molecules. The addition of polarisation functions to a DZ basis set results in a double-zeta-plus polarisation basis (DZP basis). For example, in a DZP for CH_4, a set of three 2p-functions is added to each H atom, and a set of six 3d-functions is added to the C atom.

d) def2-TZVP basis set

In this thesis, we use a def2-TZVP basis set, which corresponds to a triple-zeta-valence-polarisation basis. [151–154] The def2 term corresponds to the 2nd generation of the default basis set of the Ahrichs group. It is a Gaussian-type orbital with three functions per valence orbital, optimised for valence electrons and with diffuse polarised functions (p-orbital on H atoms and d-orbital on C atoms). In this thesis, auxiliary basis sets are also used, such as def2-tzvp/C, def2/J. The /C stands for correlation fitting, and /J stands for Coulomb fitting. They are used to speed up expensive parts of the calculation. They are part of the Resolution of Identity (RI) approximation. Further details can be found in the original articles. [155,156]

4.2.9. Electron correlation

The ground-state HF wavefunction is not the 'exact' wavefunction because it does not consider instantaneous Coulombic interaction between electrons, but rather treats them as an average field of electrons. This approximation ignores the electron correlation. The difference between the exact energy E_0 and the HF limit is the correlation energy:

$$E_{\text{corr}} = E_0 - E_{\text{HF}} \qquad (19)$$

Many theories have been developed to take electron correlation into account such as: configuration interaction (CI), multiconfigurational methods, Møller-Plesset (MP) perturbation theory, coupled-cluster (CC), Complete Active Space Self Consistent Field (CASSCF), etc. They are not detailed here.

4.3. Density Functional Theory (DFT)

DFT is one of the primary computational methods used in recent years. Its advantage resides in lower computational time, and in some cases, like d-metal complexes. The central focus of DFT is not the wavefunction but the electron probability density ρ(r) and has its roots in the Thomas-Fermi model. This model assumes that the electron density determines the properties of the molecules and that energy is correctly given by a variation principle. A functional is a function of a function. Hence, the term 'functional' refers to the fact that the energy of the molecule E[ρ(r)] is a function of electron density, which itself is a function of the position of electrons ρ(r).

DFT was put on a solid theoretical footing by Pierre Hohenberg and Walter Kohn in 1964, starting with the following theorems[157,158]:

1. **Hohenberg and Kohn's existence theorem**: The ground-state energy and all other ground-state electronic properties are uniquely determined by the electron density
2. **Hohenberg and Kohn's variational theorem**: For a trial density function ρ'(r), the energy functional $E_o[\rho'(r)]$ cannot be less than the true ground-state energy of the molecule.

The Hamiltonian of many-electron is repeated here:

$$\widehat{H_e} = -\sum_{i=1}^{N_e} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N_e}\sum_{A=1}^{N_n} \frac{Z_A}{|r_i - R_A|} + \sum_{i=1}^{N_e}\sum_{j>i}^{N_e} \frac{1}{|r_i - r_j|}$$

The interaction between the electrons and nuclei (second term) will be noted $v(r)$ for each electron:

$$\widehat{H_e} = -\sum_{i=1}^{N_e} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N_e} \mathrm{v(r)} + \sum_{i=1}^{N_e}\sum_{j>i}^{N_e} \frac{1}{|r_i - r_j|}$$

The first theorem implies that there exists a one-to-one correspondence between the external potential $v(r)$(and thus the Hamiltonian and wavefunction) and ρ(r). It is possible to compute $v(r)$ from ρ(r), then the Hamiltonian $\widehat{H}$, and finally the energy E. It follows that the ground-state electron density determines (in principle) all properties of the molecules.

The second theorem follows the variational principle that implies that the minimisation of the energy functional E[ρ] gives the ground-state energy E_0:

$$E_0 = min_\rho E[\rho] \quad (20)$$

We have seen previously that the total energy can be written as the sum of the kinetic energy (T) of the electrons and the potential energy V. We can express T and V as functionals of the electronic density ρ. In 1964, Hohenberg and Kohn were able to prove that that the exact ground state of energy of N_e electron molecule is uniquely determined by the electron probability density ρ [158] :

$$\mathrm{E}[\rho] = \mathrm{E}[\rho]^{\mathrm{classical}} + E_{xc}[\rho] \quad (21)$$

with $\mathrm{E}[\rho]^{\mathrm{classical}}$ defined as:

$$\mathrm{E}[\rho]^{\mathrm{classical}} = T_s[\rho] + V_{ne}[\rho] + V_{ee}^{\mathrm{classical}}[\rho] \quad (22)$$

Where $V_{ne}[\rho]$ is the electron-nucleus potential energy, and $V_{ee}[\rho]$ the electron-electron potential energy, $T_s[\rho]$ being the kinetic energy of non-interacting electrons, and:

$$E_{xc}[\rho] = V_{ee}^{\text{non-classical}}[\rho] + \Delta T \quad (23)$$

$E_{xc}[\rho]$ is the exchange-correlation energy. It considers all the non-classical electron-electron effects due to spin and applies a small correction to the kinetic energy due to electron-electron interactions. The $\Delta T = \mathrm{T} - \mathrm{T_s}$, it accounts for the difference between the true kinetic energy and that of non-interacting electrons. The Hohenberg-Kohn theorem proves the existence of $E_{xc}[\rho]$ but does not show how to compute them.

The first step is to compute electron density ρ. In 1965, Kohn and Lu Jeu Sham[158], showed that ρ can be expressed as a sum of each electron contribution as:

$$\rho = \sum_{k=1}^{N} \phi_k^{*} . \phi_k \quad (24)$$

with φ_k the Kohn-Sham orbitals, which are solutions of the Kohn-Sham equations. These are Schrödinger-like equations:

$$\hat{\mathrm{f}}_{\mathrm{KS}}\phi_k = [T_s + V_{\mathrm{ext}}]\phi_k = \varepsilon_k\phi_k \quad (25)$$

We find ourselves in the same situation as with the Hartree-Fock method. Therefore, we can use a self-consistent field procedure to solve these equations.

4.3.1. The exchange-correlation challenge and functionals

The challenge is to find the functional $E_{xc}[\rho]$. Numerous approaches have been developed to obtain approximate forms of the functional exchange-correlation energy. The main source of DFT errors comes from the approximate nature of $E_{xc}[\rho]$.

Currently, there exist around hundreds of different exchange-correlation functionals, and the search for more accurate functionals is still under research: Local Density Approximation (LDA), Generalised Gradient Approximation (GGA), meta-GGA (mGGA), hybrid functionals, double-hybrid functional, etc.[143,147,159]
It is common to classify functional in a ladder called the Jacob's ladder of approximations to the exact exchange-correlation functional. [160,161] This ladder progresses from the Hartree world, with no exchange-correlation, toward the "heaven" of chemical accuracy (within 1 kcal/mol for energetics).

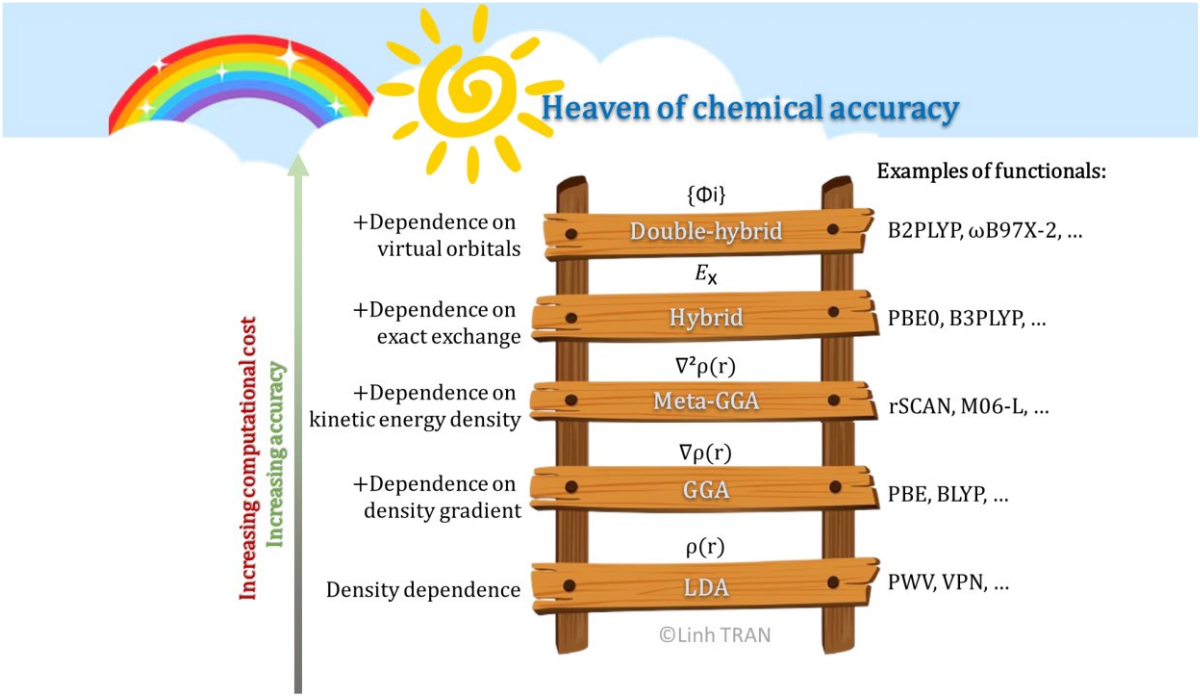


Figure 22. Illustration of Jacob's ladder of DFT functionals.

LDA was one of the first functional, derived from Monte Carlo simulations for the uniform gas. It assumes the $E_{xc}[\rho]$ depends only on the local electron density. Then GGA functionals were developed to improve upon LDA. It includes gradient of the density and captures spatial variations, which better describes molecules. The meta-GGA brought about a massive improvement by an order of magnitude of binding energies.[160] It incorporates kinetic energy density or second derivatives of the density. Hybrid functionals is the last big advance in exchange-correlation functionals. It includes some of HF-exchange. The energy can be expressed as:

$$E_{\mathrm{xc}}[\rho] = \int_0^1 \langle \Psi_\lambda | \widehat{V_{\mathrm{xc}}} | \Psi_\lambda \rangle \, d\lambda \qquad (26)$$

A fixed percentage of exact exchange is then chosen to be included in the functional being constructed. The two most widely used hybrid functionals are B3LYP and PBE0. Then later development came to double-hybrid (or hyper-GGA) functionals which involves unoccupied virtual orbitals. Hybrid (like PBE0) and double hybrid functionals (like B2PLYP) are used in this thesis and is detailed in the following sections.

4.3.2. Hybrid functional

a) PBE0 hybrid functional

PBE0 [162,163] was established by Adamo and Barone, and at the same time Ernzerhof and Scuseria.[164] It is created using GGA, Perdew-Burke-Ernzerhof (PBE) functionals and ¼ Hartree-Fock exchange. It is a purely theoretical functional, and does not rely on any fitted parameters. It is used in this thesis for optimization of molecular structure geometries and TD-DFT (which is presented in a later section).

$$E_{\mathrm{xc}}^{\mathrm{PBE0}}[\rho] = E_{xc}^{\mathrm{PBE}}[\rho] + 1/4(E_x^{HF}[\rho] - E_x^{PBE}[\rho]) \qquad (27)$$

With $E_{xc}^{\mathrm{PBE}}[\rho]$ the total exchange-correlation energy in PBE functional, $E_x^{HF}[\rho]$ the exact exchange energy calculated using HF theory and $E_x^{PBE}[\rho]$ the exchange energy from the PBE GGA functional.

b) B3LYP hybrid functional

Another hybrid functional B3LYP[165], combines Becke's three-parameter exchange functional (B) [166] with 1/5 Hartree-Fock exchange and the Lee-Yang-Parr correlation functional (LYP)[167,168]. The general expression for the B3LYP functional is then given by:

$$E_{\mathrm{xc}}^{\mathrm{B3LYP}} = aE_x^{\mathrm{HF}} + (1-a)E_x^{\mathrm{LSDA}} + b\,\Delta E_x^{\mathrm{B88}} + c\,E_c^{\mathrm{LYP}} + (1-c)E_c^{\mathrm{LSDA}} \qquad (28)$$

with coefficients a = 0.2, b = 0.72, and c = 0.81. E_x^{LSDA} *is the* Local spin-density approximation exchange.

4.3.3. Double-hybrid functionals

a) B2PLYP / RI-B2PLYP double hybrid functional

In DFT, a double hybrid functional includes a certain amount of HF exchange and PT2 correlation (thus the term *double hybrid*). This idea was first proposed by Donald Truhlar in 2004 and more well-known, by Stefan Grimme in 2006.[169] It is based on a mixing of standard generalized gradient approximations (GGAs) for exchange by Becke (B) and for correlation by Lee, Yang, and Parr (LYP) with Hartree-Fock (HF) exchange and a perturbative second-order correlation part (MP2) that is obtained from the Kohn-Sham (GGA) orbitals and eigenvalues. This virtual orbital-dependent functional contains only two global parameters that describe the mixture of HF and GGA exchange a_x and of the PT2 and GGA correlation a_c, respectively.

$$E_{xc}^{\mathrm{B2PLYP}} = a_x E_x^{\mathrm{HF}} + (1-a_x)E_x^{\mathrm{B88}} + (1-a_c)E_c^{\mathrm{LYP}} + a_c E_c^{\mathrm{MP2}} \qquad (29)$$

Where E_c^{MP2} is the Møller-Plesset (MP) perturbative correlation energy, $a_x = 0.53$ *and* $a_c = 0.27$.

In this thesis, we use RI-B2PLYP, which has the RI(Resolution of Identity) approximation that reduces computational cost of MP2 calculation.[155,156]

4.3.4. Energy derivatives and Hessian matrix

Since 1969, when P.Pulay developed the first computer program for analytically computing derivatives of self-consistent field (SCF) energies, gradient-based methods have become fundamental in quantum chemical calculations. Determining a molecule's equilibrium structure involves locating the minimum on the potential energy surface (PES), which corresponds to a point where the first derivatives of the energy with respect to nuclear coordinates, meaning the derivatives equal zero. However, a zero gradient merely indicates a stationary point and does not distinguish between a minimum, maximum, or saddle point. To make this distinction, the second derivatives must be evaluated. These second derivatives, taken with respect to nuclear coordinates, are collectively represented in the Hessian matrix. When there is one negative frequency, the molecule is at its minimum (equilibrium) or maximum (transition state TS). When there are two negative frequencies, the system is at a saddle point.

4.3.5. Molecular properties and gradient method

Once the electronic energy is obtained by solving the Schrödinger equation, various molecular properties can be computed. One of them is the equilibrium structure, which is the first step to most of the computations in this thesis. To do so, we must first compute the derivatives of the PES with respect to nuclear coordinates. It can be done numerically, but it is onerous in computation time. On the other hand, the gradient method, which computes derivatives analytically, is more accurate and faster than numerical differentiation. Energy derivatives can also be used to compute the normal mode frequencies within the harmonic approximation.

4.4. Time-Dependent DFT (TD-DFT)

DFT methods are used to solve the time-independent Schrödinger equation. However, predicting spectroscopic properties requires accounting for time-dependent phenomena, such as light, which is studied in this thesis. Indeed, when a molecule is exposed to electromagnetic radiation which oscillates continuously during exposure, it undergoes a time-dependent perturbation. Therefore, a time-dependent theoretical framework, such as Time-Dependent DFT (TD-DFT), becomes necessary to compute spectral properties and excited states.

The Schrödinger time-dependent equation (1) is reminded here:

$$\hat{H}\Psi(r,t) = i\hbar\frac{\partial}{\partial t}\Psi(r,t)$$

The time-dependent Hamiltonian is:

$$\hat{H}(t) = \underbrace{\hat{T}_e + \hat{V_{NN}} + \hat{V_{Ne}} + \hat{V_{ee}}}_{\hat{H_e}} + \hat{V_{pert}}(t) \qquad (30)$$

where the first part is defined in equations (3)-(4), and $\widehat{V_{pert}}(t)$ is the external perturbation operator. Only the last term is time-dependent and can represent a perturbation like light exposition. It is expressed as:

$$\widehat{V_{pert}}(t) = \sum_{i=1}^{N} v_{pert}(\boldsymbol{r}_i, t) \quad (31)$$

where $v_{pert}(\boldsymbol{r}_i, t)$ is the potential describing the perturbation experienced by an electron i:

$$v_{pert}(\boldsymbol{r}_i, t) = E_{el}(t).\alpha.\boldsymbol{r}_i \quad (32)$$

with $E_{el}(t)$ the periodic field time-dependent and α the associated polarisation.

Each Kohn-Sham monoelectronic orbital can be expressed as:

$$\psi_k(\boldsymbol{r}_i, t) = \varphi_k(\boldsymbol{r}_i)\phi_k(t) \quad (33)$$

By analogy, the electronic density time-dependent can be expressed as:

$$\rho(\boldsymbol{r}_i, t) = \sum_{k=1}^{N} \psi_k(\boldsymbol{r}_i, t)^*.\psi_k(\boldsymbol{r}_i, t) \quad (34)$$

a) Runge and Gross theorem

The Runge–Gross theorem is the foundational theorem for Time-Dependent DFT(TD-DFT), much like the Hohenberg–Kohn theorem is for ground-state DFT. In 1984, Runge and Gross has proven the that the electronic density time-dependent can be used to solve the Schrödinger time-dependent equation (1). [170] This has led to the formal foundation of of TD-DFT theory.
This theorem states that there is a one-to-one mapping between the potential v_{pert} and the electronic density time-dependent $\rho(\boldsymbol{r}_i, t)$. If two external potential v_{pert} and v_{pert}' differ by a constant c(t), then those potentials generate two electronic different densities.

$$v_{pert} \neq v'_{pert} \Rightarrow \rho(\boldsymbol{r}_i, t) \neq \rho'(\boldsymbol{r}_i, t)$$

In TD-DFT, the time-dependent Kohn-Sham equations can thus be written as:

$$E[\rho, t] = T_s[\rho, t] + V_{ext}[\rho, t] \quad (35)$$

$$\hat{f}_{KS}\phi_k(t) = [T_s + V_{\text{ext}}(t)]\phi_k(t) = \varepsilon_k\phi_k(t) \quad (36)$$

It is possible to solve this equation using the linear response approximation.

This approximation assumes that the system is subjected to a small-amplitude perturbation V_{ext}. Under this assumption, one can analyze the linear response of the electron density to the perturbation, as well as the response function of the Kohn-Sham orbitals.

By performing a first-order expansion of these equations, one can determine the excitation energies, and thus obtain the energies of the system's excited states.

4.4.2. Linear response theory

Linear response theory is central to TD-DFT as it describes how quantum system responds to a weak, time-dependent perturbation. For instance, in absorption spectroscopy, we study how a molecule in its ground state responds to a weak external electromagnetic field. When the perturbation is small, the system's response can be approximated using Linear Response Theory (LRT). The time-dependent electron density $\rho(r,t)$ is expanded in a Taylor series around the ground-state density ρ_0 and for a weak perturbation, only the first-order term is retained $\rho_{(1)}$ as it is greater than higher orders:

$$\rho(r,t) \approx \rho_0(r) + \rho_{(1)}(\boldsymbol{r},t) \qquad (37)$$

According to the Runge–Gross theorem, there is a one-to-one correspondence between the time-dependent external potential and the time-dependent density. Thus, the first-order density response is written as:

$$\rho_{(1)}(r,t) = \int\int \chi_{KS}(r,r',t-t')\delta v_{\text{eff}}(r',t')\ dr'\,dt \qquad (38)$$

where χ_{KS} is the density–density response function in the Kohn-Sham scheme. In TD-DFT, this response is computed using the Kohn–Sham scheme, where the interacting response function is built from the non-interacting response χ_0 and the exchange-correlation kernel f_{xc}:

$$\chi = \chi_0 + \chi_0 \cdot f_{\text{xc}}\,\chi \qquad (39)$$

Transforming to the frequency domain gives access to the system's excitation energies. These are determined by the poles of the response function and correspond to transitions between occupied and virtual orbitals. In the Kohn–Sham framework, the non-interacting response is given by:

$$\chi_0(\omega) = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \frac{f_j - f_k}{\omega - (\varepsilon_k - \varepsilon_j) + i\eta}\,\phi_j^*(r)\phi_k(r)\phi_k^*(r')\phi_j(r') \qquad (40)$$

where j and k run over occupied and virtual orbitals, respectively, ε the orbital energies, η a postive infinitesimal, and f_j , f_k are occupation numbers. $(\varepsilon_k - \varepsilon_j)$ represents the energy gap between the occupied and unoccupied orbitals.

4.5. Absorbance spectra simulation

Modelling electronic absorption or emission spectra involves computing two key quantities: the excitation energies and the corresponding oscillator strengths for transitions from the ground to excited states. The excitation energies give the electronic transitions and therefore the position of the peaks, and the oscillator strengths of a peak

are proportional to the intensity of that peak. Solving the Schrödinger time-dependent equation (1) seen previously in the TD-DFT section yields the electronic energies between the ground state and excited states, known as the vertical energies (VE). These values are represented in the form of discrete spectral lines. However, real absorption and emission bands are not infinitely sharp; instead, they span a finite wavelength range. This line broadening arises primarily from three factors: natural (lifetime) broadening, environmental effects such as solvent interactions, and vibronic (vibrational) coupling. The following subsections discuss these quantities, and the theoretical equations are inspired by the thesis of Bernardino Tirri and the Orca software manual. [104,146,171]

4.5.1. Oscillator Strength (or Oscillator force)

Once the vertical energies are computed, we then compute the oscillator force (*fosc*) which tells how strongly a molecule absorbs light. A high oscillator force for a certain vertical energy means a high probability of absorption at that energy. It is therefore proportional to the intensity of each absorption peak. The *fosc* is proportional to the absorption coefficient ε in the Beer-Lambert Law:

$$\mathrm{A} = \log_{10}\left(\frac{I_0}{I}\right) = \varepsilon \cdot l \cdot C \qquad (41)$$

where l is the cuve length, C, the molar concentration. I_o and I the incident and transmitted light intensity respectively.

fosc is related to the transition probability between the initial state (i) to a final state (j), as:

$$\mathrm{P_{ij}} = \left(\frac{8\pi^3 e^2}{3h^2 c}\right) G_j D_{ij} \qquad (42)$$

where e is the electron charge, h the Planck constant, G_j the statistic weight of the final state j and D_{ij} the transition dipole intensity, defined as:

$$D_{ij} = \left(\int \langle \psi_i | \mu | \psi_j \rangle \, d\tau\right)^2 \qquad (43)$$

If the molecule or chromophore possesses a dipole moment μ that can oscillate at the same frequency as the incident electromagnetic radiation, absorption or emission can occur.

The relation between computed oscillation force f and experimental absorption molar coefficient is:

$$\mathrm{f} = \mathrm{I} - I_o = 4.315.10^{-9} \int \varepsilon(\nu)\, d\nu \qquad (44)$$

4.5.2. Spectra intensity with vibrational effects

When we predict the full spectra, we need to add the vibrational energies, with their corresponding intensity. Those intensities are computed in Orca software using the

Excited States Dynamics (ESD) module. The vibrational effects are presented in the later subsection. Here, we will present how the rates are computed. The idea behind *fosc* started with the equation from the quantisation of the electromagnetic field for the transition rates between an initial state $\Psi_i$and a final stateΨ_f. It can be calculated using Fermi's Golden rule:[172,173]

$$k(\omega)_{if} = \frac{4\omega^3 n^2}{3\hbar c^3} \left|\langle \Psi_i | \hat{\mu} | \Psi_f \rangle\right|^2 \delta\left(E_i - E_f \pm \hbar\omega\right) \quad (45)$$

with $\hbar\omega$ being the energy of the photon, $\hat{\mu}$ the dipole operator and n the refractive index of the solvent, as suggested by Strickler and Berg. [174]

One approach is to compute it in the frequency domain by evaluating the Franck-Condon factors between all initial and final states that satisfy the Dirac delta condition in equation (45) while appropriately weighting the thermally accessible initial states.

$$k^{obs} = \int k(\omega) d\omega\,, \qquad k(\omega) = \sum_{if} P_i(T) k_{if}(\omega) \quad (46)$$

where $P_i(T) = e^{-\frac{\epsilon_i}{k_B T}}/Z$ is the Boltzmann population of a given initial state at temperature T , ϵ_i is the total vibrational energy of state i and Z is the vibrational partition function.

This approach can require the inclusion of a very large number of states, especially in the presence of low-frequency modes. Orca Software adopt an alternative strategy by transitioning to the time domain, using the Fourier transform representation of the Dirac delta function:

$$\delta(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\omega t} dt \quad (47)$$

The equation to solve in atomic units can be written as:

$$k(\omega) = \frac{2\omega^3}{3\pi c^3 Z} \sum_{i,f} e^{-\frac{\epsilon_i}{k_B T}} \langle \Theta_i | \overrightarrow{\mu^e} | \overline{\Theta_f} \rangle \langle \Theta_f | \overrightarrow{\mu^e} | \overline{\Theta_i} \rangle \int e^{i(E_i - E_f - \omega)t} dt \quad (48)$$

with $\overrightarrow{\mu^e}$ the electronic transition dipole and $|\overline{\Theta}\rangle$ the vibrational wavefunction of the initial or final state.

After additional steps, including change of time variable and the insertion of the resolution of identity, the equation (48) can be simplified to a discrete Fourier transform of a correlation function $\chi(t)$ with a timestep Δt, multiplied by a prefactor α:

$$
\begin{aligned}
k(\omega) &= \alpha \int Tr\left(\overrightarrow{\mu^e} e^{-i\hat{H}\tau} \overrightarrow{\mu^e} e^{-i\widehat{\widehat{H}\tau}}\right) e^{i\Delta E t}\, e^{-i\omega t} dt \\
&= \alpha \int \chi(t) e^{-i\omega t} dt \\
&= 2\alpha \Re e \int_0^{\infty} \chi(t) e^{-i\omega t} dt \\
&\simeq 2\alpha \Delta t \;\; \Re e\, DFT\{\chi(t)\}
\end{aligned}
\qquad (49)
$$

and this correlation function is then calculated using path integrals analytically at each time point t.[172,175]

If the electronic component of the transition dipole moment is allowed to vary with nuclear displacements, then it can be expressed as:

$$
\overrightarrow{\mu^e}(\boldsymbol{Q}) = \overrightarrow{\mu_0^e} + \sum_i \left.\frac{\partial \overrightarrow{\mu^e}}{\partial Q_i}\right|_{\boldsymbol{Q}=0} Q_i + \cdots, \qquad (50)
$$

This approach also allows for the inclusion of vibronic coupling, specifically the so-called Herzberg–Teller (HT) effect. The corresponding correlation function for HT can be efficiently derived recursively from the Franck–Condon (FC) case, enabling an effective computational treatment.

a) Franck–Condon (FC) Principle

The Franck–Condon (FC) principle states that electronic transitions occur much faster than nuclear motion, meaning that during an electronic transition (e.g., absorption or emission), the nuclei in the molecule are essentially "frozen" in place. As a result, the transition probability is highest between vibrational states whose nuclear configurations (geometries) overlap the most.

b) Herzberg–Teller (HT) Effect

While the Franck–Condon principle assumes that the electronic transition dipole moment is independent of nuclear coordinates (the Condon approximation), this is not always the case. The Herzberg–Teller (HT) effect accounts for the vibrational dependence of the electronic transition dipole moment. In situations where the transition is symmetry-forbidden or weak under the Condon approximation, the Herzberg–Teller effect can introduce intensity to otherwise forbidden transitions. This occurs when the transition dipole moment varies linearly (or higher order) with nuclear displacements, effectively allowing vibronic coupling to "borrow" intensity from nearby allowed transitions.

Together, the FC and HT effects provide a more complete description of the intensity distribution in electronic spectra, particularly for systems with weak or forbidden transitions.

c) Duschinsky rotation

Another point in this theory is that, to solve the path integrals[172,175], we need to choose a single set of coordinates either for the initial state Q or the final state $\bar{\boldsymbol{Q}}$ (where the bar marks the final coordinates). Because we're dealing with a transition between two states, we have to convert one coordinate system into the other. In 1930s, Duschinsky proposed that the two sets of coordinates are related by:

$$\boldsymbol{Q} = \boldsymbol{J}\bar{\boldsymbol{Q}} + \boldsymbol{K} \qquad (51)$$

where rotation matrix $\boldsymbol{J}$ and the displacement vector $\boldsymbol{K}$ defined as:

$$\boldsymbol{J} = \mathbf{L}_{\mathbf{x}}^{\mathbf{T}}\bar{\boldsymbol{L}}_{\boldsymbol{x}},\ \boldsymbol{K} = \mathbf{L}_{\mathbf{x}}^{\mathbf{T}}(\overline{\boldsymbol{q_0}} - \boldsymbol{q_0}) \qquad (52)$$

with $\mathbf{L_x}$ being the matrix containing the normal modes, here described in Cartesian coordinates (x), and $\boldsymbol{q_0}$ the mass weighted coordinates ($q_i = \sqrt{m_i}x_i$).

The program starts by reading the initial and final geometries and their Hessians. It then calculates the Duschinsky rotation matrix and the displacement vector. Next, it computes the transition dipole derivatives and uses them to calculate the correlation function. After that, a DFT calculation is performed, and the transition rates are computed and printed if needed.

Predicting transition rates requires at least the geometries and Hessians of both the ground state (GS) and an excited state (ES). In ORCA, there are seven available methods to approximate the excited-state potential energy surface (PES): **AHAS, VH, VG, HHBS, HHAS, UFBS,** and **UFAS**.

4.5.3. Vibronic couplings

To model absorption or emission spectra, vertical excitation energies and oscillator strengths are required. However, to accurately capture spectral shape and broadening, vibrational (or vibronic) couplings must be included. Vibronic couplings are responsible for the shape of the absorbance spectra, and therefore to the colour of the compound.

As discussed before, vertical electronic transitions give rise to absorption lines. These transitions, however, are often accompanied by vibrational transitions. In absorption, a vibrational transition involves a vertical excitation from a vibrational state in the ground electronic state (S_0) to one in an excited electronic state (S_1), as illustrated in Figure 23. Conversely, during emission, the vibrational transition corresponds to a vertical transition from an excited vibrational state back to a vibrational level in the ground electronic state.

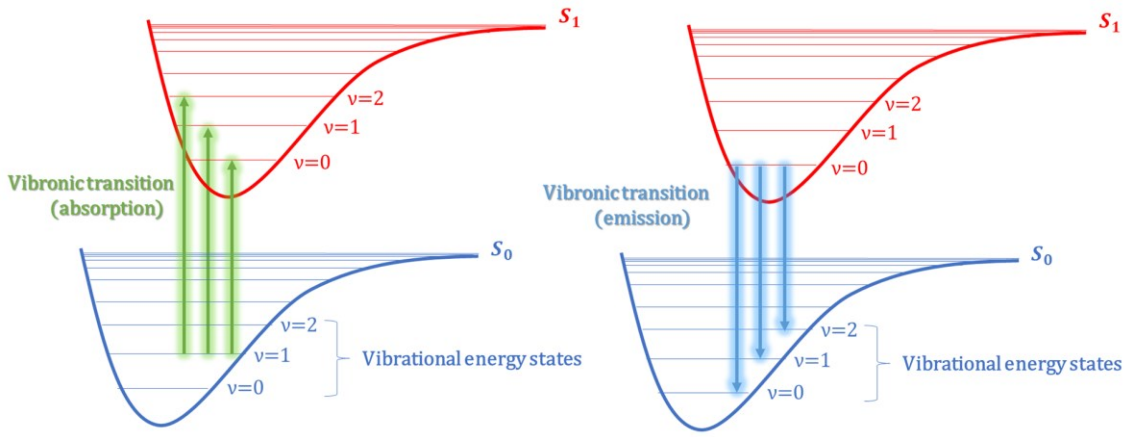


Figure 23. Scheme of the vibrionic transitions between two singlet states. Absorption (solid green line) and emission (solid blue line).

For many chromophores commonly used in the dye industry, such as anthraquinones, vibrational coupling significantly influences the spectral band structure. In this context, we examine the key factors required to derive the vibrational fine structure based on the molecule's vibrational modes.

It can be modelled via several approaches. A simple method involves placing a Gaussian function centred on vertical energies, but it neglects the computed vibrational contributions. Another option is molecular dynamics, where hundreds of snapshots are computed to average vertical energies and consider the surrounding explicit water molecules. While this method is more accurate, it is very computationally expensive.

In this work, we adopt a stationary method to compute vibrational energies, offering a good balance between accuracy and efficiency. A comparison with the dynamic method is presented in Article 1.

All vibrational effect calculations were carried out using the Excited State Dynamics (ESD) module in ORCA. This thesis focuses on three vibrational coupling approaches: Vertical Gradient (VG), Adiabatic Hessian After a Step (AHAS), and Adiabatic Hessian (AH). These methods differ in how they treat the potential energy surfaces (PES) of the ground and excited electronic states and how vibrational normal modes are computed. The details of the methods can be found in reference papers and are described below. [117,172,175–178]

a) Harmonic approximation and vibronic models

The theoretical framework of vibrational coupling relies on an accurate description of the potential energy surfaces (PES) associated with the electronic states involved in a transition. However, for large molecular systems (hundreds to thousands of nuclear coordinates), a full PES mapping through high-level electronic structure calculations is computationally out of reach. As a result, approximate models based on local Taylor expansions around key geometries are used to represent these surfaces. Quadratic expansions are the simplest and physically acceptable description of PES and is the basis of harmonic approximation. These approximations remain the most practical option for handling large systems.

Two main harmonic models are used: the adiabatic (A) and vertical (V) approaches. Both share the same initial-state description, obtained through equilibrium geometry optimisation and normal mode analysis, but differ in how the final-state surface is treated. The Adiabatic Hessian (AH) model involves a full harmonic expansion around the excited-state minimum, while the Vertical Hessian (VH) model builds the final surface using data (energy, gradient, Hessian) evaluated at the ground-state equilibrium geometry. The "adiabatic" refers to the adiabatic Born–Oppenheimer approximation BO), thus neglecting direct nonadiabatic couplings. These models account for normal modes rotation upon electronic transition, the Duschinsky effect. To reduce computational demands, simplified alternatives such as the Vertical Gradient (VG), also known as LCM (Linear Coupling Model) and Adiabatic Shift (AS), also called Adiabatic Hessian After a Step (AHAS), models have been proposed (Figure 24). These assume shared normal modes between states and require only minimal data, such as energy gradients or equilibrium shifts. The idea behind these approximations is to do a geometry update step ($\boldsymbol{\Delta S = -gH^{-1}}$ for Quasi-Newton and $\Delta S = -g(H+S)^{-1}$ for Augmented Hessian, where $\boldsymbol{g}$ is the gradient and $\boldsymbol{H}$ the matrix Hessian.

Method	Step on the ES* geometry	ES Hessian in Orca
Vertical Gradient, VG	1 Augmented Hessian* step on the ES PES	GS hessian is used
Adiabatic Hessian After Step, AHAS	1 Augmented Hessian* step on the ES PES	ES hessian is computed on the new ES geometry
Adiabatic Hessian, AH	Relaxed geometry on the ES PES	ES hessian at the relaxed geometry

*ES means Excited Stated, GS means Grounds State, the Augmented Hessian step corresponds to a very small displacement on the ES PES

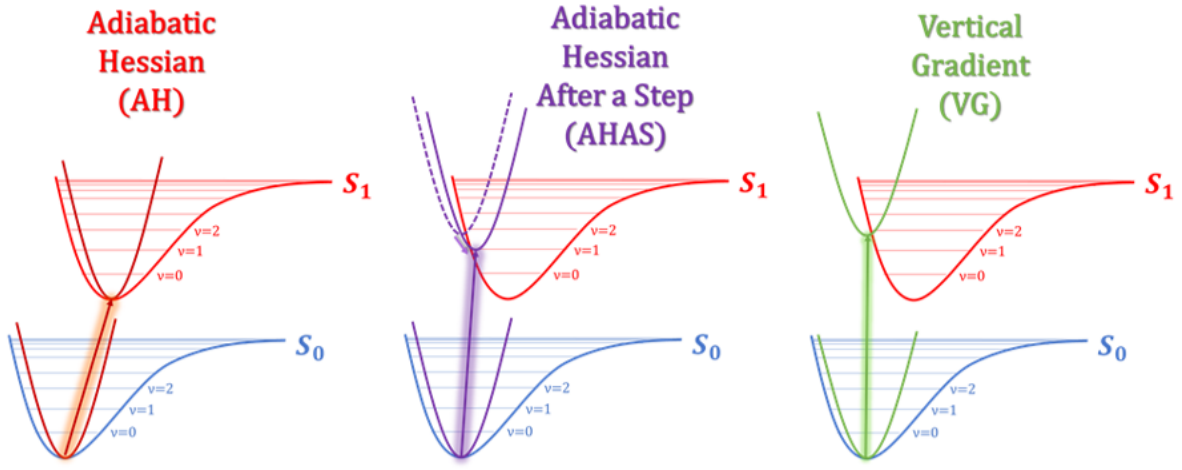


Figure 24. Illustration of the different methods for vibronic couplings: Vertical Gradient (green), Adiabatic Hessian After a Step (purple), and Adiabatic Hessian (red).

Intermediate schemes like VG and AS allow frequency changes while maintaining mode alignment.

We will now present the theoretical details behind those computations. They can be found in the original article.[177]

Let us consider a molecular system undergoing an electronic transition from an initial electronic state, denoted $|\boldsymbol{e_0}\rangle$ to a final state $|\boldsymbol{e_i}\rangle$. From this point onward, a prime (′) will be used to label quantities associated with the initial state. The nuclear configuration of the system is described by a 3N-dimensional column vector x (cartesian coordinates of all nuclei). We denote by $\boldsymbol{x_0}'$ the equilibrium geometry of the initial electronic state. Within the harmonic approximation, the PES of initial state V' can be expressed in matric as: [177]

$$V'(Q') = \frac{1}{2} Q'^{\mathrm{T}} \Omega'^2 Q' \qquad (53)$$

where Ω is the diagonal matrix of the normal frequencies of vibration and Q' is the vector of the N associated normal coordinates (N = $3N_a$-6 except for linear molecules where N = $3N_a$-5, with N_a is the number of atoms) defined as:

$$L'Q' = M^{1/2}(\boldsymbol{x} - \boldsymbol{x}_0') \qquad (54)$$

M being the diagonal matrix of the atomic masses.

b) Adiabatic Hessian (AH)

The Adiabatic Hessian (AH) method involves optimising the geometries of both the ground and excited electronic states and computing the Hessian (second derivative matrix) at each equilibrium geometry. The PES of final state V is expanded around its own equilibrium geometry $\boldsymbol{x}_0$. By definition, the gradient of V at $\mathbf{x_0}$ is zero, so only the Hessian of V is computed to construct its harmonic approximation, analogous to the procedure used for the initial state:

$$V(\boldsymbol{Q}) = E_{\mathrm{ad}} + \frac{1}{2} \boldsymbol{Q}^{\mathrm{T}} \Omega^2 \boldsymbol{Q} \qquad (55)$$

$$\boldsymbol{LQ} = \boldsymbol{M}^{1/2}(\boldsymbol{x} - \boldsymbol{x}_0) \qquad (56)$$

Where the adiabatic energy E_{ad} represents the minimum energy of the final state PES with respect to the energy of the initial state PES in its own minimum taken as a reference. It is possible to express Q' as a function of Q (or vice versa), by extracting x from (56) and inserting it into (54). We obtain:

$$\boldsymbol{Q'} = \boldsymbol{JQ} + \boldsymbol{K} \quad (57)$$

$$\mathbf{J} = \mathbf{L'^{-1}}\boldsymbol{L} \quad (58)$$

$$\boldsymbol{K} = \boldsymbol{L'^{-1}}\boldsymbol{M^{1/2}}(\boldsymbol{x} - \boldsymbol{x'_0}) \quad (59)$$

where J is the rotation matrix known as the Duschinsky matrix, and K is the equilibrium position displacement vector. It is worth noting that while the 3 translational coordinates can be rigorously eliminated in (58). In general, this is not possible in an exact way for the rotational coordinates.

The final state PES can be written as a function of the normal coordinates of the initial state, Q':

$$\mathrm{V}(\boldsymbol{Q'}) = E_{\mathrm{ad}} + \frac{1}{2}\boldsymbol{K}^{\mathrm{T}}\mathbf{J}\boldsymbol{\Omega^2}\boldsymbol{J}^{\mathrm{T}}\mathbf{K} - \boldsymbol{K}^{\mathrm{T}}\mathbf{J}\boldsymbol{\Omega^2}\boldsymbol{J}^{\mathrm{T}}\boldsymbol{Q'} + \frac{1}{2}\boldsymbol{Q'}^{\mathrm{T}}\mathbf{J}\boldsymbol{\Omega^2}\boldsymbol{J}^{\mathrm{T}}\boldsymbol{Q'} \quad (60)$$

This latter is a quadratic function but not in canonical form, exhibiting linear terms and a force constant matrix $F_A = J\Omega^2 J^T$ that is, in general, non-diagonal.

$$\mathrm{E}_{(\mathrm{r})}^{\mathrm{AH}} = \frac{1}{2}\mathbf{K}^{\mathrm{T}}\mathrm{J}\Omega^2\mathrm{J}^{\mathrm{T}}\mathbf{K} \quad (61)$$

is the reorganisation energy on the final state PES at the ground state geometry, and therefore

$$E_{(V)}^{\mathrm{AH}} = E_{\mathrm{ad}} + E_{(r)}^{\mathrm{AH}} \quad (62)$$

r is the estimate of the vertical excitation energy at the initial state geometry according to the AH model. Being based on a harmonic PES, $E_{(V)}^{\mathrm{AH}}$is in general an approximation of the exact value $\mathrm{E_V}$ (by ''exact'' we mean the value computed directly at $\boldsymbol{x'_0}$ with the same level of electronic theory adopted at $\boldsymbol{x}_0$).

Although highly accurate, the AH method is computationally demanding due to the need for excited-state geometry optimization and Hessian evaluation. If the molecule is large and flexible, this can lead to difficulty in convergence.

c) Adiabatic Hessian After a Step (AHAS)

The Adiabatic Hessian After Step (AHAS) method simplifies the computation by avoiding excited-state optimizations. Instead, it uses the ground-state normal modes and computes the gradient of the excited-state PES at the ground-state equilibrium geometry. So instead of evaluating the excited-state properties strictly at the ground-state geometry, the system is first displaced along the excited-state potential energy surface (PES) through a step of geometry optimization. At this new excited-state minimum, the excited-state

Hessian is computed. By definition, $J^{(AS)} = 1$ and $\Omega^{(AS)} = \Omega'$, K retains the same expression given in (59) while the reorganization energy becomes

$$E_{(r)}^{AH} = \frac{1}{2}\boldsymbol{K^{T} J \Omega'^{2} K} \quad (63)$$

and

$$E_{(V)}^{AS} = E_{ad} + E_{(r)}^{AS} \quad (64)$$

The AS expression for V PES is simplified, and its difference with respect to V' is just a linear function of Q':

$$V(Q') = V'(\boldsymbol{Q'}) + E_{(V)}^{AS} + \boldsymbol{K'^{T} \Omega'^{2} Q'} \quad (65)$$

d) Vertical approach

Vertical Hessian

Vertical gradient is based on the Vertical Hessian approach. This approach assumes that the shape of the potential energy surface does not change significantly upon excitation, neglecting Duschinsky rotation. As a result, the Vertical method is less computationally intensive but generally less accurate, especially for systems with significant geometric changes upon excitation.

An quadratic model for the final-state potential energy surface (PES), V, can be constructed using exclusively vertical data, that is, quantities evaluated at the equilibrium geometry of the initial state, $\boldsymbol{x}_0'$. Specifically, this includes the vertical excitation energy E_V, the gradient ($\mathbf{g_x}$) and Hessian (F_x) of V, where x means cartesian coordinates. Using these quantities, a harmonic approximation of the final-state PES can be formulated as follows:

$$V(x) = E_V + g_x^T x + \frac{1}{2} x^T F_x x \quad (66)$$

and, exploiting, it can be expressed in terms of the ground state normal coordinates Q' as:

$$V(Q') = E_V + \mathbf{g_x^T} Q' + \frac{1}{2} Q'^{T} \mathbf{F_x} Q' \quad (67)$$

$$g = M^{1/2} L' g_x \quad (68)$$

$$F = L'^{T} M^{-1/2} F_x L' M^{1/2} \quad (69)$$

This model is named VH and the expression in (67) can be easily put in canonical form by determining the orthogonal transformation $Q'' = (J^{(VH)})^T Q'$ that diagonalizes the force

constants matrix F, $J^{(VH)}FJ^{(VH)T} = (\Omega^{(VH)})^2$ and then simply locating the minimum by setting to zero the first derivatives and solving for Q_j variables,

$$Q''_{j0} = \frac{\left(-\boldsymbol{g}^{\boldsymbol{T}} J^{(VH)}\right)_j}{\left(\Omega_j^{(VH)}\right)^2} \quad (70)$$

Final state normal coordinates are then obtained by a translation of the origin $Q_j^{(VH)}=Q_j''-Q_{j0}''$, where 'VH' indicates that $Q^{(VH)}$ coordinates of the final state are in principle different from those obtained with AH model for the same state Q. They coincide where the harmonic approximation of V is exact.

A relationship between $Q^{(VH)}$ and Q' can be written as:

$$Q' = J^{(VH)}Q^{(VH)} + K^{(VH)} \quad (71)$$

$$K^{(VH)} = -J^{(VH)}Q^{(VH)} - 2\left(J^{(VH)}\right)^T \boldsymbol{g} \quad (72)$$

Vertical Gradient (VG)

In analogy with the adiabatic models, a simplified alternative to the Vertical Hessian (VH) model can be defined by assuming that the normal modes of the initial and final states are identical, differing only by a displacement. This approach leads to a model that requires only the gradient of the final-state potential energy surface V, evaluated at the equilibrium geometry of the initial state, x0'.

We refer to this approximation as the Vertical Gradient (VG) model, also commonly known in the literature as the Linear Coupling Model (LCM). In this model, the final-state normal coordinates and frequencies are taken to be the same as those of the initial state, i.e., $Q(VH) \approx Q'$ and $\Omega(VG) \tilde{} \Omega'$, then $J(VG)=1$ and $K(VG) = -(\Omega')-2g$.

Overall, these methods yield a new set of vibrational normal modes and frequencies, characteristic of the excited state. These excited-state vibrations directly influence the vibronic progression in the simulated spectrum, often improving the accuracy of spectral features. In particular, AHAS can recover fine spectral details, such as shoulders or relative intensities, that simpler models like VG might miss. Each method represents a trade-off between computational cost and accuracy, and the choice depends on the molecular system under study and the level of detail required in the spectral simulation. This is discussed in the results section.

4.6. Modelling the environment

In computational chemistry, solvent models are essential for accurately simulating chemical systems in environments that resemble real-world conditions. In the presence of solvents, molecular behaviour can be influenced by various interactions, including hydrogen bonding, electrostatic forces, steric repulsion, and proton transfer processes, all of which can affect molecular geometry and conformational stability. We have to take these effects into account to accurately model reality. There are several ways to model the solvent. These models vary in complexity and computational cost, ranging from simplified to highly detailed representations. The ones used in this thesis are vacuo, implicit, microsolvatation (1 or 3 explicit solvent per OH) or explicit solvation with a box of 100 water molecules. (Figure 25). They are detailed in the following subsections.

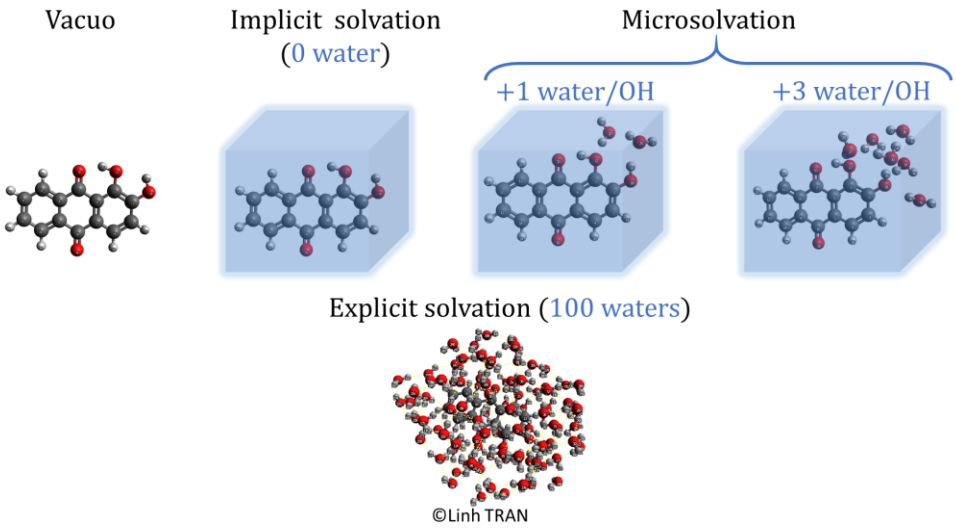


Figure 25. Different solvation models used in this thesis: vacuo, implicit solvation, microsolvation, and explicit solvation.

4.6.1. Vacuo

In vacuo calculations entirely neglect solvent effects, treating molecules in isolation within a vacuum, which can be computationally efficient but often lacks realism for reactions occurring in solution. It can be used to model experiments done in a vacuum. In this thesis, it is used to compare the results with other solvent models to understand the environmental effects.

4.6.2. Implicit model

In computational chemistry, implicit solvation models simulate the effect of a solvent on a solute without representing individual solvent molecules. Instead, the solvent is treated as a continuous, homogeneous dielectric medium characterised by its dielectric constant (ε). This ε constant varies for each solvent. The Polarizable Continuum Model (PCM) is among the most widely used implicit models and serves as a conceptual and mathematical framework for many others, including COSMO and SMD.[179,180] Models using a

homogeneous medium are grouped in the self-cohering reaction fields (SCRF) category. This approach captures bulk solvent effects while maintaining low computational cost.

In this thesis, the CPCM (solvent) keyword in Orca is used for implicit solvation and is detailed here. In CPCM, the solute is placed in a cavity (typically defined by overlapping van der Waals spheres) within a dielectric continuum representing the solvent. The solvent modifies the electrostatic potential experienced by the solute by polarising in response to the solute's charge distribution. This polarisation induces a reaction field, which in turn interacts with the solute, altering its energy and properties.

In this model, the free energy of solvation, which can be assimilated to all the average interactions between the solute and the solvent, can be written as :

$$\Delta G = \Delta G_{\text{el}} + \Delta G_{\text{cav}} + \Delta G_{\text{rep}} + \Delta G_{\text{dis}} \qquad (73)$$

In this equation the electrostatic energy (ΔG_{el}) represents the interaction between the solute's charge distribution and the polarization charges induced in the surrounding dielectric medium at the solute-cavity boundary. The cavitation energy (ΔG_{cav}) accounts for the work required to form cavity in the solvent continuum to accommodate the solute. The repulsion (ΔG_{rep}) and dispersion energies (ΔG_{dis}) describe non-electrostatic interactions at the solute-solvent interface, specifically short-range Pauli repulsion between overlapping electron clouds and long-range London dispersion forces, respectively.

The use of this model requires defining the shape of the cavity that encloses the solute (Figure 26). Ideally, this cavity should closely follow the solute's molecular surface while remaining simple enough to model efficiently. A common approach is to represent each atom in the solute as a sphere cantered on its nucleus, with a radius typically corresponding to its van der Waals radius. In this framework, the solvent is modelled as a probe sphere that "rolls" over the surface formed by the solute atoms. Two key surfaces are then defined: the Solvent Excluded Surface (SES) and the Solvent Accessible Surface (SAS). The SES corresponds to the surface formed by the contact points between the solvent probe and the solute atoms, effectively marking the boundary that the solvent cannot penetrate. The SAS, on the other hand, traces the path of the centre of the solvent probe as it moves around the solute surface. Additionally, the van der Waals Surface (vdWS) can be used as an alternative approach, providing a smooth and continuous representation of the molecular surface.

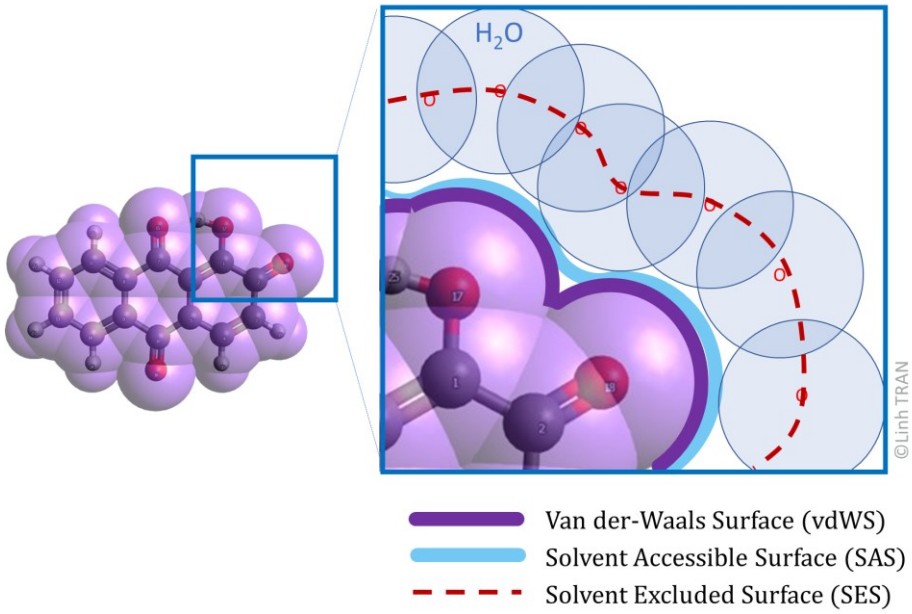


Figure 26. Definition for different surfaces: Van der-Waals Surface (vdWS) in bold violet, Solvent Accessible Surface (SAS) in bold blue, and Solvent Excluded Surface (SES) in dashed red.

4.6.3. Microsolvation model

Alternatively, the solvent can be represented using explicit solvents, in which individual solvent molecules are included in the simulation. This approach enables the accurate description of specific solute–solvent interactions, such as hydrogen bonding, ion pairing, π–π interactions, and solvent structuring. The primary advantage of explicit solvation lies in its ability to capture these detailed, localised effects that are inaccessible to continuum models. However, this increased realism comes at a considerable computational cost, as it requires handling a much larger number of particles and performing extensive conformational sampling to obtain statistically meaningful results.

In this thesis, a hybrid explicit–implicit microsolvation approach was employed, combining both solvation models to balance computational efficiency with chemical accuracy. In this framework, the regions of the solute involved in strong solvent interactions, particularly functional groups such as hydroxyl and carboxyl, were treated explicitly by placing water molecules in their neighbourhood to capture specific interactions like hydrogen bonding. The other parts of the solute and surrounding environment was modelled using an implicit continuum solvent CPCM, allowing for the inclusion of bulk solvent effects without significantly increasing computational cost. Two configurations were tested: one with one water molecule per hydroxyl group, and another with three water molecules per hydroxyl group. The latter configuration, involving three water molecules per OH group, was found to provide greater accuracy and is discussed in detail in the Results section. This hybrid strategy provides a more realistic representation of solvation phenomena while maintaining manageable computational demands.

4.6.4. Explicit solvation

An alternative strategy was also investigated, which involved incorporating a greater number of explicit water molecules (e.g., 100) into the system and performing ab initio molecular dynamics (AIMD) simulations. By coupling these simulations with quantum dynamics techniques such as QM/xTB (extended tight-binding), various solvent conformations could be effectively sampled. In this approach, the environment is represented by a molecular mechanics (MM) force field the solute and a portion of the solvent are treated quantum mechanically. In this thesis the QM region is the dye and 3 neighbouring waters making hydrogen bonds with the hydroxy groups (Figure 27). Subsequently, multiple snapshots were extracted from the trajectory to compute the desired properties, and an average over these configurations was used to obtain the final result. This solvation model is more realistic but is more expensive in computation time.

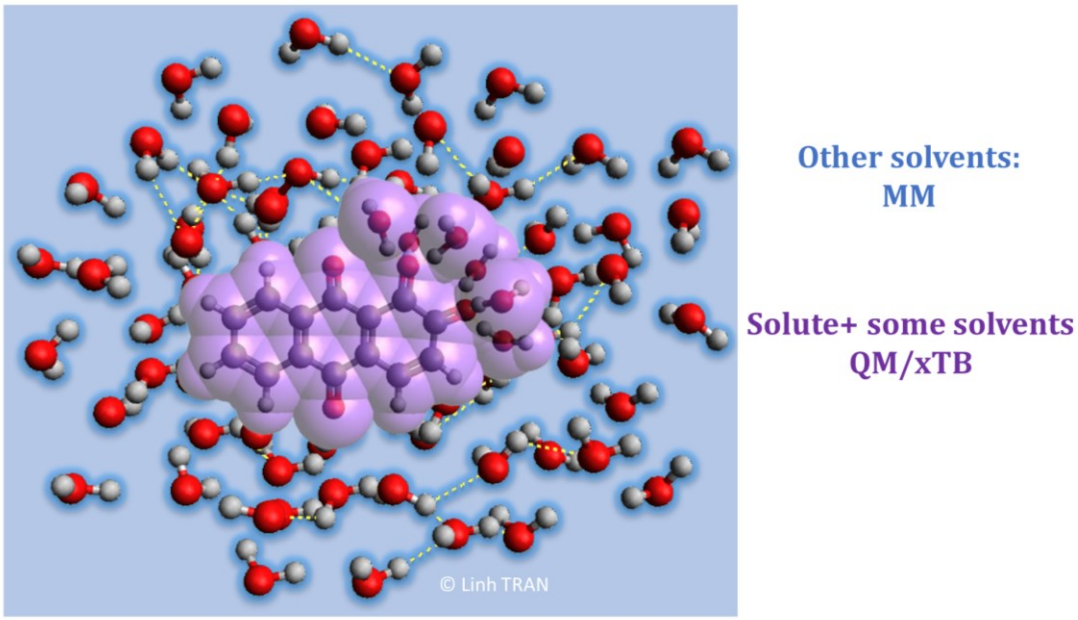


Figure 27. Hybrid QMxTB/MM solvation model. The QM/xTB is alizarin with 3 water/hydroxy groups, and the other remaining water are treated with MM.

CHAPTER 5

CHAPTER 5 | ARTICLES 102

Article 1: "Rapid and complete prediction of alizarin in solution by combining experimental data with computational methods." 102

Article 2: "In-silico colour prediction process for natural dyes in Madder." 131

Article 3: "Progress in modeling complex dye mixtures: a case study for cultural heritage." 169

Article 4: "When light betrays colour: the photodegradation of alizarin, a historical dye, in basic media." ..195

Article 5: "A computational insight into the photodegradation mechanism of alizarin " 245

This chapter presents the research findings in the form of five articles. Articles 1 to 3 have been published, Article 4 is currently under review, and Article 5 is in preparation.

CHAPTER 5 | ARTICLES

This thesis is structured around five scientific articles, three already published, one currently under review, and the last one under preparation. They all collectively articulate the research trajectory and key findings around the prediction of spectral and colourimetric properties of complex mixtures of colourants in madder, as well as their photodegradation behaviour.

This thesis contributes to the study of madder (*Rubia tinctorum* L.), a natural dye historically important since antiquity for its red colourants. Madder roots contain 30–68 colourants that vary by species and age. Many are unstable, photosensitive, not available commercially, or difficult to synthesise. Therefore, there is a lack of reference standards, complicating spectroscopic analysis. Meanwhile, understanding how these dyes degrade under light exposure requires precise identification of the molecules present and their spectroscopic signatures. Each compound interacts differently with light, and the presence of multiple substances can affect individual behaviour. In this thesis, we chose to combine experimental methods with computational chemistry, which is underused in heritage science.

This study has two main objectives: first, to develop a predictive computational method for colour and spectral properties to aid in dye identification and characterisation. This allows access to the missing experimental database; second, to deepen the understanding of photodegradation mechanisms of madder dyes, which are not yet elucidated. This approach supports preservation strategies in cultural heritage and is also applicable to other fields that use colourants.

With this goal in mind, we first established a computational protocol specifically adapted to madder dyes, guided by literature benchmarks (Article 1 published). Using this validated computational protocol, we then compute optical properties and build a complete database including spectroscopic signatures for 31 colourants identified in 6 different madder species, at different pH levels (Article 2 published). However, in reality, like many natural dyes, madder is a complex mixture of multiple colourants. Therefore, the next step focused on modelling the spectral properties of this mixture, using a madder extract of the species *Rubia Tinctorum* L. To our knowledge, this represents a novel contribution in the field of theoretical chemistry. It can be applied to predict properties of other complex mixtures: mixtures of different colourants, tautomers, photodegraded products, impurities, etc. (Article 3 published). With the spectroscopic properties of the dyes elucidated, we finally turned to study photodegradation mechanisms using alizarin solution as a model, focusing on the basic conditions (Article 4 submitted). Those mechanisms could be confirmed by computational chemistry to study reactivity, reaction pathways and energy barriers. (Article 5 under writing). The summaries of each article are presented below, before each article.

SUMMARY

Article 1: "Rapid and complete prediction of alizarin in solution by combining experimental data with computational methods."

In the context of this thesis, we first developed a computational protocol tailored to madder dyes, based on literature benchmarks, particularly from Adèle Laurent, Denis Jacquemin, Carlo Adamo. [106,112] Density Functional Theory (DFT) and its time-dependent analogue (TD-DFT) were identified as well-suited for anthraquinone systems, with PBE0 and B2PLYP as best functionals. This computational protocol can predict the absorption spectra, colourimetric coordinates, and ^{1}H and ^{13}C NMR of alizarin, one of madder's principal colourants. The protocol enabled accurate prediction of alizarin's colour across different pH levels, incorporating solvent effects. It includes conformer analysis and deprotonated forms based on molecular single-point energy (SPE) calculations. Solvent effects can change the molecules' geometry, participate in proton transfer or form hydrogen bonds, particularly with water solvent. The effects of water solvents were evaluated using different theoretical models: implicit, explicit, and hybrid QM/MM(xTB), the latter coupled with quantum dynamics simulations to model multiple water molecule configurations.

The results showed that the optimal solvent model depends on the molecular form and its interaction with the solvent. For example, the protonated form of alizarin (acidic pH), which interacts less with water, can be modelled efficiently using an implicit approach. For deprotonated forms, which form hydrogen bonds with water, a combined model using implicit solvent and three explicit water molecules per hydroxyl group offered the best compromise between accuracy and computational cost. The most accurate model across all pH conditions was the QM/MM(xTB) approach with 100 explicit water molecules, though it was computationally intensive. Following ab initio molecular dynamics simulations, about 100 snapshots were extracted to compute vertical excitation energies using QM/MM(xTB), treating alizarin and three water molecules per hydroxyl group quantum-mechanically, and the remaining solvent molecules with xTB.

These calculations yielded excellent agreement with experimental results. Vibrational effects were included using both the Vertical Gradient (VG) and Adiabatic Hessian (AH) approaches to model the visible absorption spectra, later converted into CIELAB colour coordinates. The AH method was found to be more consistent with experimental spectra. This led to highly realistic colour predictions, with a colour deviation $\Delta E < 13$ for alizarin (less water-soluble) and $\Delta E < 8$ for alizarin red S. This validated the computational protocol for accurate prediction of both absorbance spectra and colours, as well as ^{13}C and ^{1}H NMR chemical shifts, critical for pigment identification and structural elucidation. These findings were presented in a first published article.

ARTICLE 1

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

Thanh Huyen Linh TRAN[a], Baptiste RIGAUD[b], Maguy JABER[a,c], Romain BERRAUD-PACHE[a*]

[a] Laboratoire d'Archéologie Moléculaire et Structurale (LAMS), Sorbonne Université, CNRS, UMR 8220, Paris, 75005, France
[b] Fédération de Chimie des Matériaux de Paris-Centre FCMat
[c] IUF-Institut Universitaire de France, Paris, France
[*]corresponding author: romain.berraud-pache@sorbonne-universite.fr

Dyes and Pigments 2024;228:112242. https://doi.org/10.1016/j.dyepig.2024.112242.

ARTICLE INFO

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

ABSTRACT

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

1. INTRODUCTION

The analytical investigation of dyes is considered of utmost importance in the field of cultural heritage. Indeed, through the identification of the corresponding dye, we are often able to gain insight into several important aspects: the characterization of the dye and the artist's techniques, the degradation pathways, the preservation state of the artifact, the dating and the determination of the provenance of objects [54,181–183]. Furthermore, it provides precious details to plan correct restoration, conservation, and display strategy in museums and art galleries.

Both traditional and cutting-edge techniques such as ultraviolet-visible (UV-VIS), infrared (IR), and Raman spectroscopies, as well as nuclear magnetic resonance (NMR), mass spectrometry (MS), liquid chromatography, Fiber Optics Reflectance Spectra (FORS), X-

Ray diffraction (XRD) or hyperspectral spectroscopy have been reported for the analytical investigation of dyes [6,19,37,45,54,184–189]. While the accuracy and sensitivity of these techniques have allowed researchers to gain a huge amount of insight into these systems, most of the technologies are sampling-based and need complex dye extraction and chemical preparation methods. Moreover, sample collection, preparation, and chemical analysis are often challenging due to material complexity, limited sample sizes, dye degradation, and lack of appropriate standards needed for UV-VIS detection [19,22,34,42,44,190]. Therefore, the identification of historical dyes is still one of the most challenging tasks in the realm of heritage preservation. In this context, computational methods have emerged as an appealing tool to decipher complex experimental data [103]. These theoretical methods aid in the interpretation of individual contributions to the experimental spectrum. Unfortunately, their application in the field of cultural heritage is not as widespread as other fields of research (such as catalysis, astrochemistry, electrochemistry, etc).

Representing one of the most famous natural dyes since antiquity, madder was selected as the focus of this study [5,95,191]. From the roots of this plant, around 36 red colorants derivating from the anthraquinone family can be extracted in various quantities. We will focus here on alizarin, one of the main constituents of madder [18,59].
In order to study the associated dyes, the combination of NMR with UV-VIS is often used in the heritage science field because of their complementary nature. While UV-VIS is good at characterizing the environmental effect of the dye or its degradation (color fading), NMR can differentiate analog molecules. It is also very precise in attesting chemical modifications, notably when lake or pigment structures are created by association of a colored dye with a metallic ion.
For the computation of spectroscopic properties, density functional theory (DFT) and Time Dependent-DFT are often the methods of choice due to their reasonable costs and good precision. In fact, there have been many reports regarding the prediction of Alizarin color and properties using these methods [107,117,126,192–200]. Still, the role of the environment is not clearly understood or deciphered. Furthermore, some key properties like reproducing the UV-VIS or NMR spectra are not always present.
When examining molecules using spectroscopic analysis, a significant portion of their observable characteristics displays varying degrees of sensitivity to the surrounding environment (solvent, pH, binding partners, mordant, substrate, temperature, light, etc) [101,201]. This influence of the solvent on spectroscopic properties arises from multiple factors such as the polarity of the solvent or the presence of hydrogen bonds. It is therefore crucial to find the optimal solvation model in order to correctly predict the UV-VIS spectrum or other properties of these organic molecules.

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

2. MATERIALS & METHODS

2.1. EXPERIMENTAL METHOD

UV Spectrum

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

10 mg of alizarin were dissolved in 40 mL of deionized water (MilliQ). The alizarin solution was then heated for 2 hours at 90°C, and filtered using a Buchner funnel. 5mg of alizarin red S was added to 80 mL of distilled water and stirred for 10min. The obtained solution was used directly without heating or filtering. Then, the pH was adjusted using either hydrochloric acid (HCl, 0.1M) or sodium hydroxide (NaOH, 1M). Three pH values were considered for alizarin (pKa_1=5.5 and pKa_2=10) : 3, 8, and 12. For alizarin red S (pKa_1=5.8 and pKa_2=10.8), the following pH were used: 2, 7 and 12 [202,203].

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

NMR Spectrum

NMR spectra were acquired with a Bruker Advance Neo 500 spectrometer at the frequency of 500.07 MHz for ^{1}H and 125.76 MHz for ^{13}C equipped with a 5 mm double resonance broadband prodigy probe use nitrogen-cooled RF coils and preamplifiers. Chemical shifts were calibrated using the Tetramethysilane (0 ppm). For ^{1}H, we used a one 90° pulse with 64 scans and with a recycling time of 1s. For ^{13}C, we used a one 90° pulse with a ^{1}H decoupling waltz16 with 4096 scans and with a recycling time D1 of 2s. The ^{1}H and ^{13}C spectra were recorded in exactly the same conditions at room temperature and processed with a zero-filling factor of 2 and without an exponential decay. 5mg of either alizarin or alizarin red S were dissolved in 0.5mL of acetone-D_6. The ^{1}H and ^{13}C NMR spectra were also measured for alizarin red S in D_2O as it was more soluble than alizarin.

2.2. COMPUTATIONAL METHOD

All calculations were carried out using the version 5.0.3 of the ORCA quantum chemistry program package and the version 6.5.1 of xTB [204,205].

Optimization Procedure

Ground and excited states properties were computed using density functional theory (DFT) and its time-dependent (TD-DFT) extension [167,206]. Optimized Geometries and Hessians were obtained using the PBE0 functional [114] complemented by the D3(BJ) dispersion correction [207,208]. Tight thresholds were employed for the self-consistent field (SCF) and optimization (tightSCF and tightopt keywords in ORCA nomenclature) [207,209]. The triple zeta def2-TZVP basis set, coupled with matching auxiliary basis sets, was

consistently employed throughout all the computations [151,207,209]. All the optimized geometries can be found in the supplementary data S.I. 14.

Conformer and deprotonated Forms Search

Using Grimme's open-source Conformer-Rotamer Ensemble Sampling Tool (CREST), which relies on the xTB method, the exploration for deprotonated forms and conformer for the fully protonated form was conducted [205,210]. The geometries of the generated conformer or deprotonated forms were then subjected to optimization using the PBE0 functional. Finally, their single point energy (SPE) were obtained by using DLPNO-CCSD(T) computations with def2-TZVP basis sets with matching auxiliary basis sets and CPCM (water) [211]. Their SPE energies are then compared for each deprotonated or conformer forms to identify the most energetically stable forms (the one with the lowest SPE energy). By comparing the SPE energies, we can calculate the Boltzmann population distribution of each form. The identified stable forms correspond to the highest Boltzmann distribution. Those forms are dominant and are therefore selected for the following computations.

Solvation Models

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

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

-Implicit solvation (water or acetone) using the linear response Conductor-like Polarizable Continuum Model (C-PCM) [212]. No solvents appear but it treats the solvent as a continuous medium surrounding the solute. This model accounts for the electrostatic and non-polar interactions between the solute and the solvent without explicitly representing individual solvent molecules. Their bulk properties are replaced by some macroscopic quantities (dielectric constant or solvent radius).

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

-Furthermore, a hybrid approach called Quantum Mechanics/Molecular Mechanics-2 (QM/QM2) [105,213,214] calculations were conducted by engaging between 70-100 water molecules It combines quantum mechanics (QM) for a region of interest (here the dye molecule) and the semi-empirical GFN2-xTB method [205,210] for the rest of the system (here the surrounding water molecules). The optimization was accomplished using the multiscale module in ORCA [104,204] together with the QM/QM2 ONIOM procedure. This allows detailed treatment for the electronic structure of the solute while considering the surrounding solvent at a less computationally expensive level. To consider the solvent effect, we also performed a Quantum Molecular Dynamics simulation of 25,000 steps with a 0.5 fs time step (other parameters selected as default such as a NHC (Nosé–Hoover chains) thermostat of 300K) [215]. An ellipsoid repulsive wall was added to keep the water molecules close to the substrate during the MS. From this simulation, 50 snapshots were selected, each 200 steps starting at 800. They were then optimized and the mean Vertical Energy (VE) of these snapshots calculated.

UV-VIS Spectrum computation

For Time-Dependent Density Functional Theory (TD-DFT) calculations concerning the UV-VIS spectrum, the B2PLYP functional was used with the resolution of identity method (RI) [107,108,166]. This choice is justified by previous theoretical benchmarks and preliminary calculations (see supplementary data S.I.2) [106,110]. The UV-VIS VE calculations were done with the first 15 roots. The experimental spectra of the dyes studied (alizarin and alizarin red S) has only one band on the visible region. In all calculations, only the first excited state was considered for the VE calculation, as it can be compared to the experimental maximum absorption of the Visible region of the UV-VIS spectrum. The topology and form of the experimental spectra are due to vibrational and solvent effects which are not present when calculating the VE but will be accounted using another calculation. The color of the compound depends on the form of the spectra, it is therefore necessary to take it into account.

The vibrational contributions, including Herzberg–Teller effects [176,178,216], was accomplished using the Excited State Dynamics (ESD) module within ORCA [175]. Within this framework, the Adiabatic Hessian (AH) and Vertical Gradient (VG) approaches were used. As AH is a more complete and accurate method, computed spectra within the main text will be performed using this scheme [217]. VG spectra can be found in supplementary data S.I.5. The AH calculations require the geometry of both Ground State (*GS)* and the first singlet excited state (S_1*)* geometries and frequencies, the VG ones only required the *GS* parameters, all of them computed at the DFT PBE0 level + CPCM. The VE energy and transition dipole moments were computed using the most accurate values *i.e.* RI-B2PLYP together with the appropriate solvation scheme (3 water molecules per OH). The linewidth was created using a Gaussian distribution and was fine-tuned to align with the Full Width at Half Maximum (FWHM) parameter of the experimental spectra, which depends on the concentration and other experimental parameters. To reduce the impact of experimental parameters like concentration, the computed and experimental spectra were normalized. This allows a direct comparison between the two spectra.

The CIELAB color coordinates was calculated for the UV-VIS spectra with an in-house script with illuminant D65 [134,218,219]. To quantify the quality of the color prediction, the color difference was quantified with ΔE_{2000} which is the latest ISO standard by CIE [138,140].

NMR Spectrum computation

Geometry optimization of tetramethylsilane (TMS), alizarin, alizarin red S molecules was conducted within vacuo, an implicit acetone or water solvent environment using the CPCM method or with explicit acetone or water molecules as described before. Subsequent computation of ^{1}H-NMR and ^{13}C-NMR shielding employed the revTPSS functional with D4 correction [152,159,220]. Basis set pcSseg-3 was selected for these calculations. Next, the isotropic shifts were used to compute the shifts that are comparable to the experimental by using this formula:

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

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

3. RESULTS AND DISCUSSION

Fig.1 presents the structure of the studied molecules: alizarin (later referred as *H_2A*) and alizarin red S (later referred as *H_2ARS*). Both alizarin and alizarin red S have two acidic protons which can be subject to deprotonation depending on the solutions' pH. Alizarin (alizarin Red S) has pKa_1=5.5 and pKa_2=10 (pKa_1=5.8 and pKa_2=10.8) respectively. Therefore, in this study, we have chosen the following pH for each form: pH=3, pH=8, and pH=12 for alizarin and pH=2, pH=7 and pH=12 for alizarin red S [202,203].

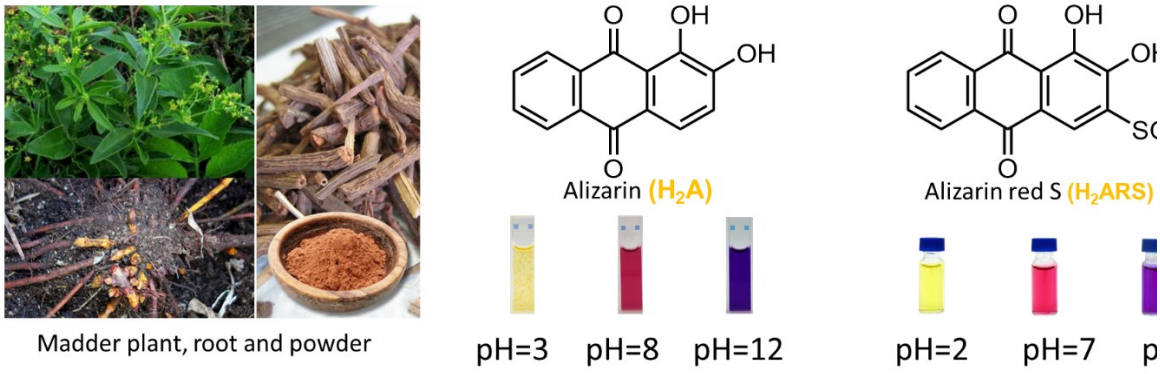


Fig. 1. (Left) Madder plant, its roots and powder. (Middle) Chemical structure of alizarin, abbreviated as H_2A and its alternative more soluble in water, alizarin red S (right), abbreviated as H_2ARS. The color of each molecule for different pH level are depicted below each structure.

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

3.1. ALIZARIN

3.1.1. UV-VIS spectra

3.1.1.1. *Conformer and deprotonated forms search*

The initial step involves exploring the conformer and deprotonated forms of alizarin for each pH range in order to identify the most stable forms.
In the protonated form (neutral) at pH=3, various studies have addressed the presence of different conformers [126,221]. We first generated the conformer of the neutral protonated form of alizarin and computed their Single Point Energy (SPE). The one with the lowest SPE energy after optimization is the most stable and is taken as a reference. The other conformer forms are listed by energy (in kcal/mol) with respect to the most stable form in Fig 2. We have found that the dominant form is the one with the hydrogen of both the α and β-hydroxyl groups making hydrogen bonds with the neighboring oxygen. It will be later referred to as *H_2A*. Furthermore, for the ground state, the forms with the proton transfer on the ketone as reported in previous publications are much higher in energy (14.07 kcal/mol respective to *H_2A)*, meaning that in the ground state, there is likely no

proton transfer for the neutral form in water. The remaining conformer forms do not exist in solution as they are higher in energy than the most stable one.

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

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

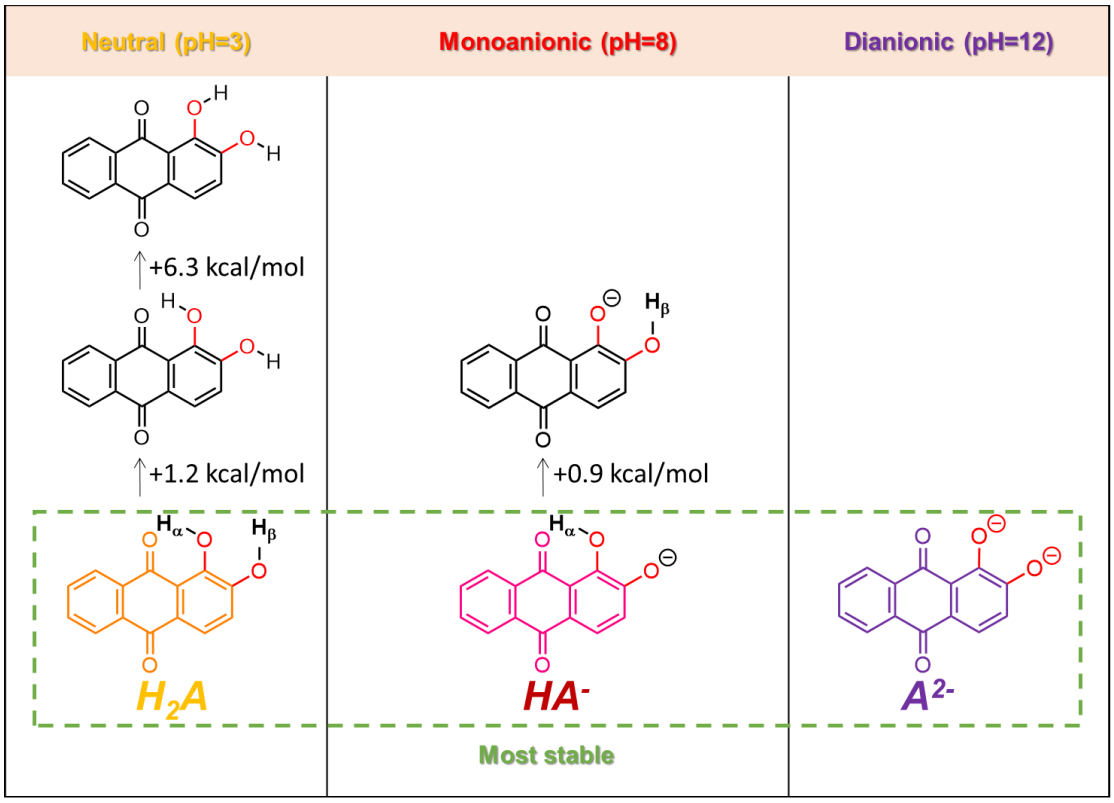


Fig. 2. Alizarin rotamers and deprotonated forms for each pH range and their relative energies. For the fully protonated form at pH=3. The geometries are sorted from the most stable form to the less stable form. The most stable forms are denoted H_2A (neutral form), HA^- (monoanionic form), and A^{2-}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

3.1.1.2. Vertical energy and choice of solvation model

In computational chemistry, the choice of solvation model is critical as it determines the accuracy of the predicted UV-VIS spectra [222]. In fact, the geometry of the molecules can be

influenced by the surrounding solvent via specific interactions like hydrogen bonds, ion pairs, π-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 data S.I. 3)*.

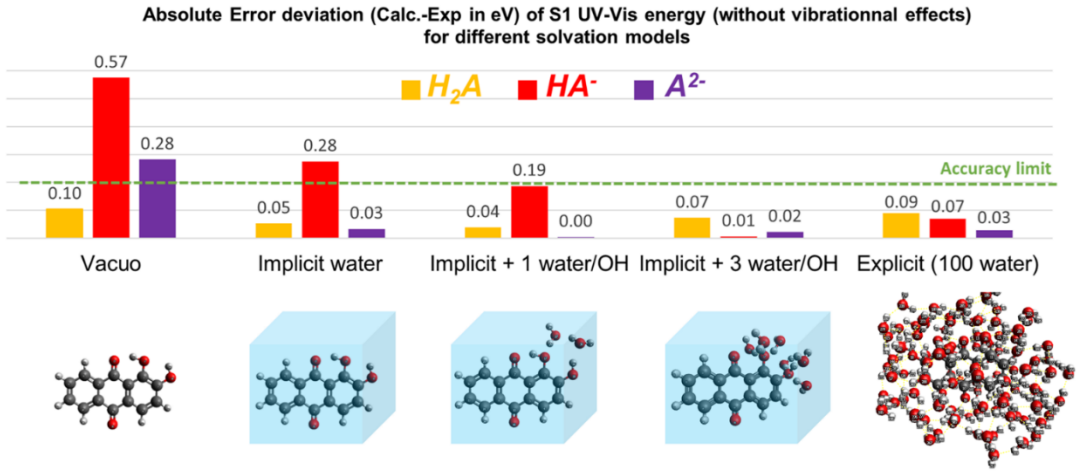


Fig. 3. (Top) Absolute energy difference between the calculated VE and the experimental absorption energy for the different solvation models of alizarin. The yellow stands for the H_2A form, the red for HA^- one and the purple for A^{2-}. The accuracy threshold is fixed at 0.2 eV according to the literature. (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).

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

A first approximation of the solvent effect can be done by using an implicit model. For alizarin, implicit solvation gives better agreement with experimental data by reducing the error by at least half (it falls to 0.05eV for H_2A, 0.28eV for HA^- and 0.03eV for A^{2-}). It can be used to test different solvent environments while being computationally efficient. This approach is useful for the case where there is no apparent interaction between the solvent and the dye, like the neutral form H_2A in water or for apolar solvents. However, the HA^- form is still above the accuracy limit (with 0.28 eV error deviation) in implicit calculation. Therefore, we need to account for specific solvent effects. For those forms, the model can be refined by adding some explicit solvent molecules near the hydroxyl sites. Explicit

water molecules (either 1 or 3) per hydroxy group were tested. This considerably reduces the calculated and experimental energy difference for the deprotonated forms. Among them, the best model is the one with 3 explicit water molecules near the OH sites (error deviation of 0.01 eV for HA^- and 0.02 eV for A^{2-}). This shows that for the deprotonated forms, solvent interactions like hydrogen bonds are mandatory to accurately model the absorption energy. Therefore, this combined model can be a good compromise to accurately predict the UV-VIS energies for a low computational cost. It can be used when there is a specific solute-solvent interaction.

Finally, the last model tested was QM/QM2 with 100 explicit water which represents a theoretical best estimate for our protocol. For each form, a quantum molecular dynamic was performed followed by computing the VE of 50 snapshots. Here, it was observed that across all the forms, the QM/QM2 results 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-}). This implies that QM/QM2 is a reliable method for predicting the properties of new molecules for which the solute-solvent interactions are predominant but requires a much higher computational cost, at least 50 times higher. More information about the MD and the selected snapshots can be found in supplementary data S.I.4.

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

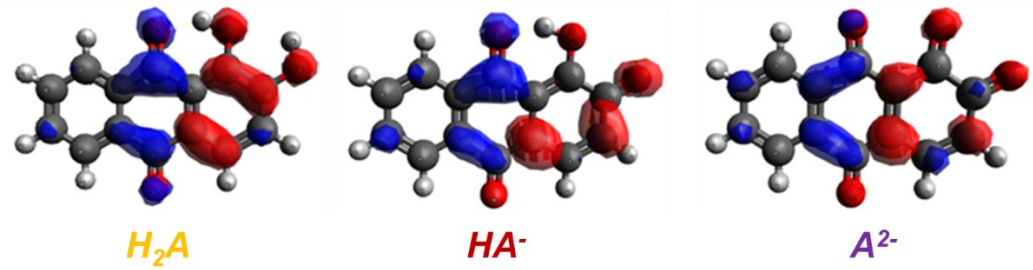


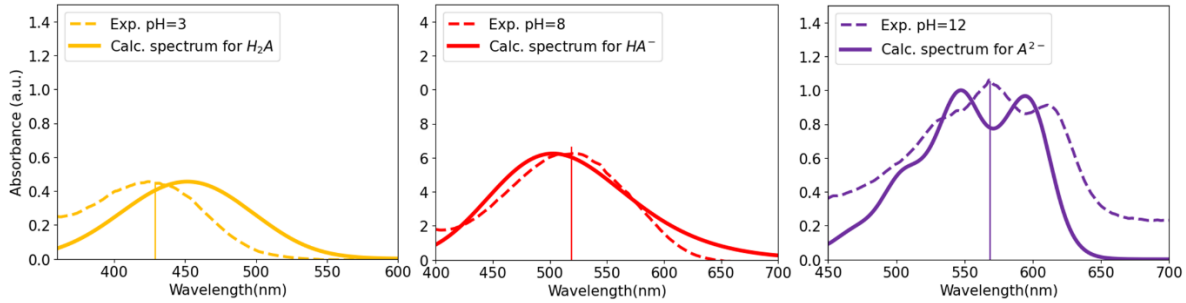
Fig. 4. Difference in electronic densities for the GS-> S_1 transition for alizarin for different protonation states. The blue orbital represents a decrease in electronic density. The red one represents an increase in electronic density. The isovalue is 0.02 a.u.

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

For the neutral form H_2A, the transition involves a combined n->π (owing to the presence of oxygen) and π-> π* transition, with a more pronounced π-> π*character. As we progress to the deprotonated form, HA^- and A^{2-}, the transition nature shifts more towards

π-> π* transition rather than n->π* as the deprotonated hydroxy groups become part of the conjugated system.

3.1.1.3. Computed absorption spectra and CIELAB color



	Experimental (normalized)				Calculated (normalized)				ΔE_{2000}
Form	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 ΔE_{2000}.*

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

Overall, the computed spectra with AH are in good agreement with experimental data for both the maximum and topology, in particular for HA^- and A^{2-}. For H_2A, the computed spectrum is shifted by about (27 nm) but it can be explained by the low solubility of the neutral form in water, preventing us from having a reliable UV-VIS spectrum.
The A^{2-} replicates the experimental topology with three distinct shoulders. The VE corresponds to one electronic transition from $GS \rightarrow S_1$. The middle peak is therefore attributed to one electronic transition, while the shoulders around 540 nm and 610 nm result from vibronic effects.

While the absorption spectra are really useful in spectroscopic studies, in the cultural heritage or dyeing field, the color of the molecule is the key property. In order to gain access to this information we can convert the UV-VIS spectra to color CIELAB coordinates. CIELAB color space is a color model introduced by the Commission Internationale de l'Eclairage (International Commission on Illumination or CIE) to provide a uniform and consistent representation of colors [218]. It is designed to be independent of devices and illuminants, making it useful for comparing colors across different conditions. Here, we will use it to compare the computed and experimental color. The colors are expressed in L*a*b* coordinates. L stands for Lightness and corresponds to the perceived brightness of the color. The values range from 0 (black) to 100 (white). The a* stands for the position on the red(positive) to green (negative). b* stands for blue (negative) to yellow (positive). The CIELAB color space is represented in supplementary data S.I.6. The bottom part of Fig.5 illustrates the predicted and experimental along with their respective CIELAB coordinates. ΔE_{2000} is the latest standard created by the CIE and quantifies the color difference. E stands for "Empfindung" which is the German word for sensation and 2000 stands for the version of standard (older versions exist like ΔE_{76} or ΔE_{94}) [134,138]. The standard perception ranges as follows:

- <= 1.0: Not perceptible by the human eye
- 1-2: Perceptible through close observation
- 2-10: Perceptible at a glance
- 11-49: Colors are more similar than the opposite
- 100: Colors are exactly the opposite

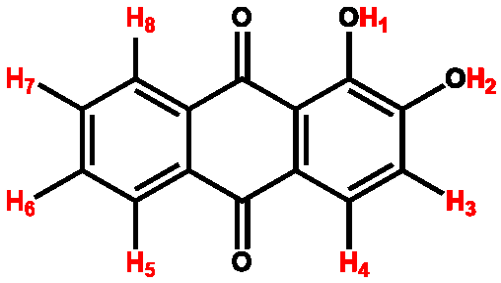
The computed color closely resembles the experimental ones. A higher ΔE_{2000}=13.3 is observed for the H_2A form and HA^- (ΔE_{2000}=7.57), possibly attributed to the lower solubility of alizarin in water, where a non-negligible precipitate can be observed. In contrast, for A^{2-} which is more soluble in water, its computed color is similar to the experimental (ΔE_{2000}=4.1). That means that the color is similar but the difference can be perceptible to the human eye at a glance. While the CIELAB colors are predominantly used for non-liquid systems, our approach shows that computations can be modeled with a reasonably good accuracy.

3.1.2. NMR shift prediction

Another key property of chemical systems is their NMR spectra. It is often used to decipher the structure of molecules in a complex environment, like pigments used in paintings, which are composed of a dye complexed with a metallic cation. Our protocol will aim at reproducing with good accuracy the experimental peaks.
Here, we present a study case with alizarin. The computed ^{1}H shift values were compared to the experimental spectra of alizarin in acetone-D_6 (chosen because alizarin is not soluble in water) in Table 1. We tested the influence of the solvation model with implicit and 3 explicit acetones following the UV spectrum protocol. The calculated ^{1}H NMR shift differs by less than 0.16 ppm on average with respect to the experimental one for aromatic hydrogen in implicit acetone and 0.22 ppm on average for explicit and implicit acetone.

This suggests a good level of accuracy for the prediction of proton chemical shifts for both solvation models. The prediction of the ^1H shifts for hydrogen H3, H4, H6, and H7 are better with implicit than explicit. However, it is noted that the accuracy is worse when dealing with the labile hydrogen in the hydroxy (-OH) group (H1 and H2) with implicit solvation. Labile hydrogen is prone to exchange with the surrounding solvents, it is therefore necessary to add explicit acetones near the hydroxy groups. For labile hydrogen, the explicit solvation model yields better accuracy. Furthermore, H5-H8 are better with explicit than implicit models. Explicit solvation is therefore more suited for labile hydrogen. The experimental spectra are depicted in supplementary data S.I.7, along with its attribution, which was done by confronting the computed shifts.

Table 1. Computed and experimental ^1H chemical shift of alizarin in acetone-D_6 with the respective error deviation in ppm or % for implicit or implicit and explicit solvation model.



	Exp. (ppm)	Calc. (ppm)		Calc.-error (ppm)		Calc.-error (%)	
Attribution	In acetone-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

3.1.3. Experimental uncertainty due to solubility

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

3.2. ALIZARIN RED S

As we have seen, the limited solubility of alizarin in water can pose challenges when measuring their experimental UV-VIS and NMR spectra due to the potential formation of

precipitates. To address this concern, a more soluble analogue, alizarin red S, has been studied.

3.2.1. Conformer and deprotonated forms search

For the acidic form (pH=2) of alizarin red S, different conformer forms can exist. We computed the energies of the possible conformer and sorted them in Fig. 6. Like alizarin, we do not observe any proton transfer for the ground state. The most stable form shows a hydrogen on the α position pointing towards the ketone, making a hydrogen bond, while the β-hydrogen is pointing on the opposite side towards the negatively charged sulfonic group. It will be noted as *H_2ARS*. The other conformer forms do not exist as they are much 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 β-hydroxyl deprotonated is the most stable configuration, the α-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-}*.

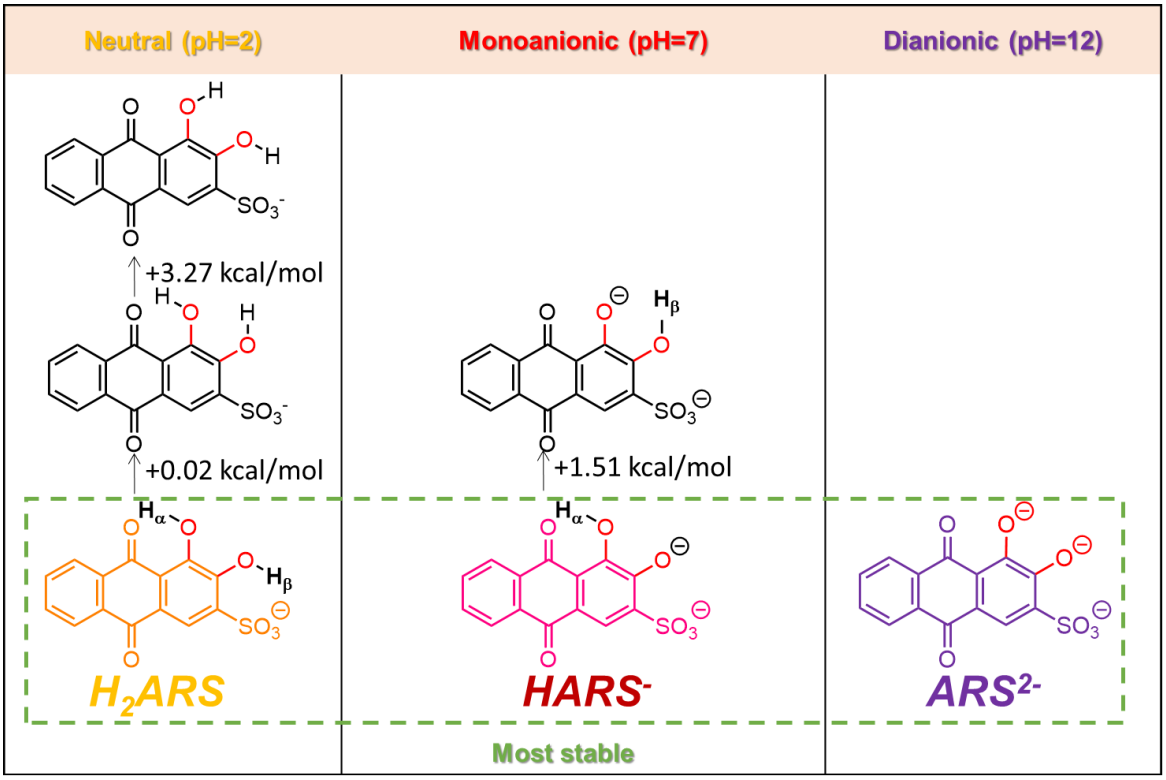
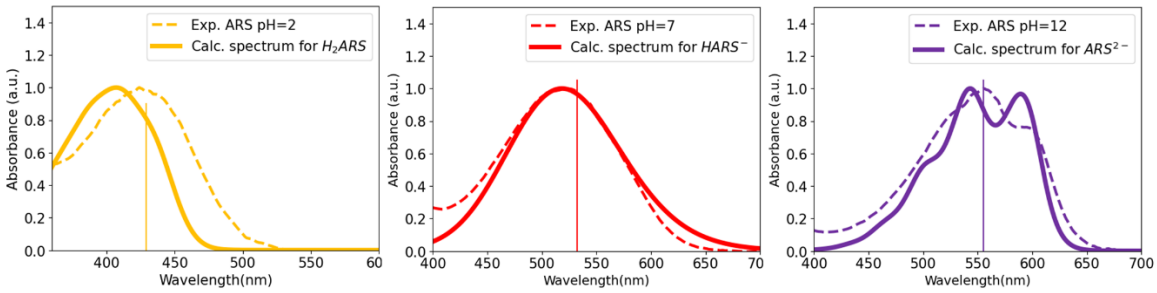


Fig. 6. Alizarin red S rotamer and deprotonated for each pH range, along with their relative energies. For the acidic form, the other forms are much higher in energy (>17 kcal/mol) and are not considered. The most stable forms are denoted H_2ARS (acidic form), $HARS^-$ (mono-deprotonated form), and ARS^{2-}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

3.2.2. UV-VIS spectra prediction

The experimental spectra have maximum absorption energy at 425nm for *H_2ARS*, 521nm for *$HARS^-$* and 560 nm for *ARS^{2-}* with two shoulders at 600nm and 530nm. The computed spectra were done considering the method and protocol used for alizarin. The results are

depicted in Fig.7. There is a better agreement between the experimental and computed spectra for the *H_2ARS and $HARS^-$* form than alizarin, which can be explained by a higher solubility of alizarin red S. The computed VE matches perfectly the experimental one. Similarly, the AH spectra of the *$HARS^-$* and *ARS^{2-}* are in adequation with the experimental ones, in both the topology and position of the spectra.



	Experimental (normalized)				Calculated (normalized)				ΔE_{2000}
Form	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 ΔE_{2000}.*

To assess the quality of the color prediction, the calculated and experimental color CIELAB are represented in Fig.7. The calculated color matches the observed experimental color, with higher accuracy for the *H_2ARS* form (3.31 for alizarin red S and 13.32 for alizarin, see Fig.5). This can be explained again by its higher solubility compared to alizarin.

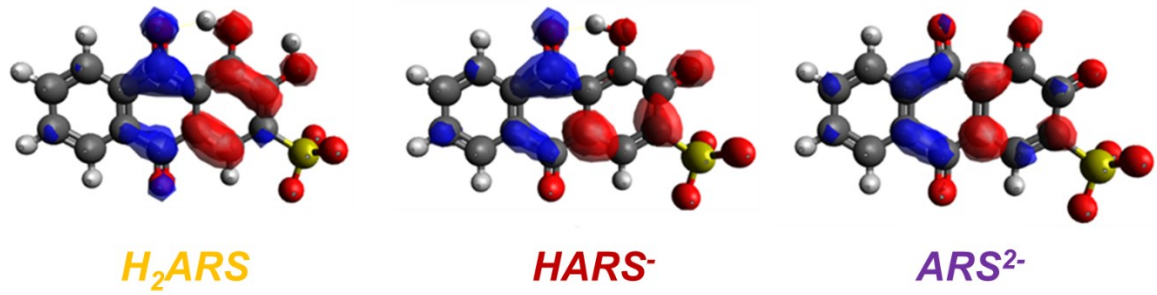


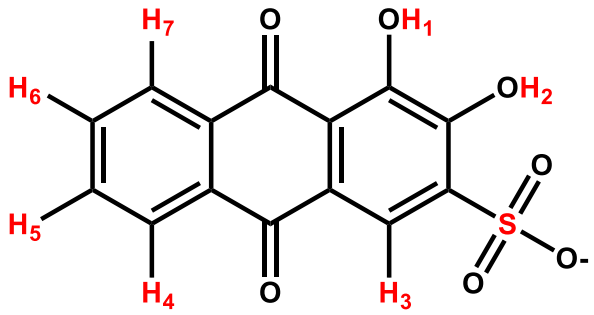
Fig. 8. Difference electronic densities for the GS-> S_1 transition for alizarin red S for different protonation states. The blue orbital represents a decrease in electronic density. The red one represents an increase in electronic density. The isovalue is 0.02 a.u.

The additional sulfonic group does not change the color of the compound compared to alizarin. This can be explained by the difference densities plotted in Fig.8. The sulfonic group does not participate in the electron delocalization for the *$GS \rightarrow S_1$* transition.

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

Thanks to the higher solubility of alizarin red S in water, the measure of NMR spectra in D_2O is possible. We have obtained the ^{1}H NMR spectra of *H_2ARS* in acetone-D_6 only as the one in water gives very broadened signals, probably due to exchange processes at room temperature which is consistent with other publications [192]. Following the best computational protocol for alizarin, the calculations were done using an explicit solvation scheme. All the geometries are accessible in the supplementary data S.I.14. The comparison between the computed ^{1}H chemical shifts is represented in Table 2 (in acetone-D_6) and ^{13}C in Table 3 (for both acetone-D_6 and D_2O). The error deviation of aromatic hydrogen (H3 to H7) is less than 0.2 ppm in average (which represents 2.1% in mean error deviation). This error is less than for alizarin (which has 2.8% mean error deviation). There is a better agreement with experimental data for alizarin red S, again explained by its higher solubility compared to alizarin. The labile hydrogen are less accurate with 17% of error deviation for H2 and 11% for H2. The experimental and computed shifts can be found for different alizarin red S forms in data S.I.8 to S.I.13.

Table 2. Computed and experimental chemical shift of alizarin red S in acetone-D_6 with the respective error deviation in ppm or %.



Attribution	Exp. (ppm) in acetone-D_6	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

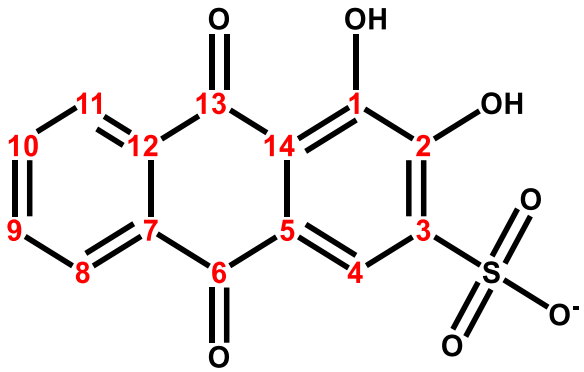
The comparison between the computed and experimental ^{13}C shifts in D_2O for alizarin red S is represented in Table 3. It shows that the error deviation of Carbons C4-C14 are less than 6 ppm on average (which represents 4.4 % in mean error deviation). There is a

higher error deviation for carbons that are next to a hydroxy or sulfonic group (C1-C3) with up to 5.5% error deviation.

Overall, there is a good agreement between the computed and experimental chemical shifts for both ^{1}H and ^{13}C NMR spectra. This protocol proves to be efficient in predicting NMR spectra.

Table 3. Computed and experimental chemical ^{13}C shift of alizarin red S in acetone-D_6 or D_2O with the respective error deviation in ppm or %.

	Acetone-D_6		D_2O	
Attribution	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



4. CONCLUSION

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

5. AUTHOR INFORMATION

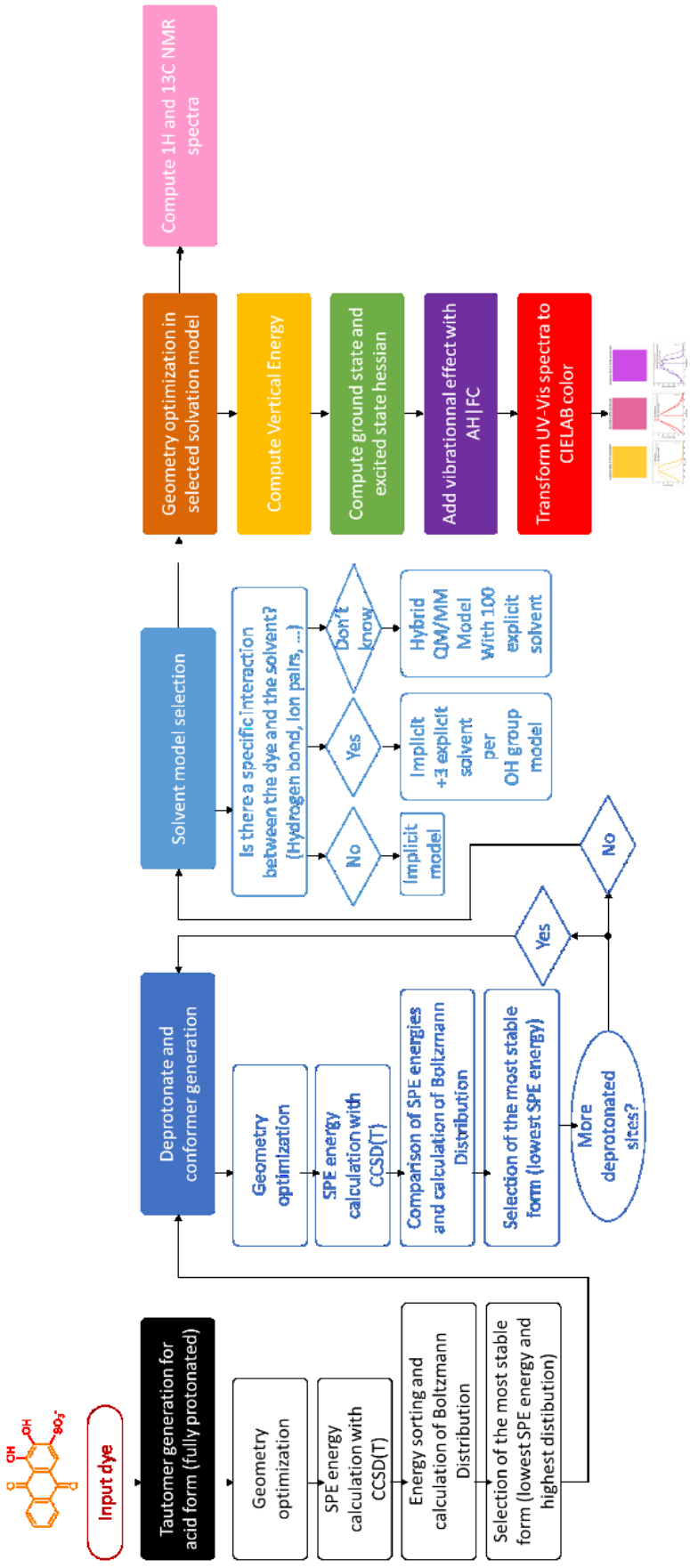
Linh Thanh Huyen TRAN: linh.tran@sorbonne-universite.fr

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

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

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

SUPPLEMENTARY DATA



S.I.1. *Detailed flow chart of the optimized protocol for the automatic computation of UV-VIS spectra, dye's color and NMR spectra.*

	B3LYP		B2PLYP		Exp.
ALIZARIN RED S	Calc. eV(nm)	Absolute Error deviation (eV) calc.-exp.	Calc. eV(nm)	Absolute Error deviation (eV) calc.-exp.	eV(nm)
H2ARS implicit	2.81 (441.3)	0.11	2.93 (425.4)	-0.02	2.91 (425.4)
HARS- implicit	2.12 (584.8)	0.26	2.17 (521.7)	0.21	2.38 (521.7)
ARS2- implicit	2.19 (565)	0.02	2.17 (559.7)	0.05	2.22 (559.7)
HARS- explicit	2.28 (543.5)	0.10	2.33 (533)	0.05	2.38 (521.7)
ARS2- explicit	2.29 (542.5)	0.07	2.27 (545.6)	-0.06	2.22 (559.7)

	B3LYP		B2PLYP		Exp.
ALIZARIN	Calc. eV(nm)	Absolute Error deviation (eV) calc.-exp.	Calc. eV(nm)	Absolute Error deviation (eV) calc.-exp.	eV(nm)
H2A implicit	2.82 (439.7)	0.06	2.94 (422.2)	0.06	2,88 (430)
HA- implicit	2.08 (595.7)	0.30	2.11 (587.9)	0.27	2,38 (520)
A2- implicit	2.17 (571)	0.01	2.14 (579.1)	0.04	2,18 (570)
HA- explicit	2.38 (584.8)	0.00	2.39 (518.5)	-0,01	2,38 (520)
A2- explicit	2.25 (565)	-0.07	2.2 (564)	0,02	2,18 (570)

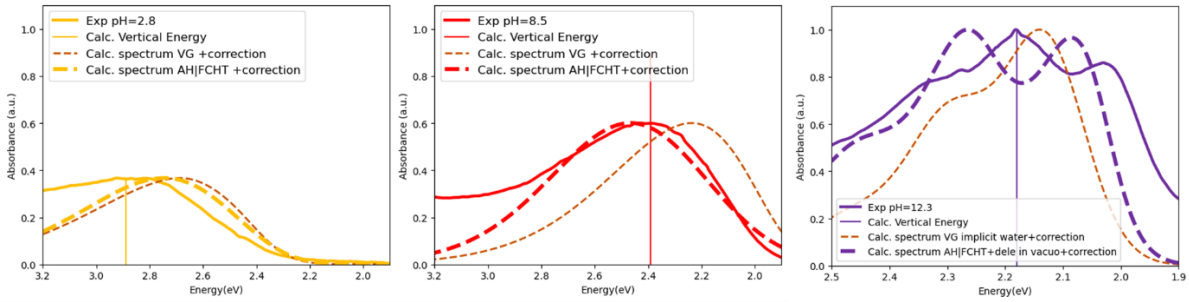
S.I.2. *Comparison between computed VE with B3LYP and B2PLYP functional for alizarin red S and alizarin.*

Solvation model	H_2A	HA^-	A^{2-}
Vacuo	0.10	-0.57	-0.28
Implicit water	0.05	-0.28	-0.03
Implicit + 1 water/OH	-0.04	-0.19	0.00
Implicit + 3 water/OH	0.07	0.01	0.02
Explicit (100 water)	0.09	-0.07	0.03

S.I.3. *Error (Calc.-Exp) of S1 UV-VIS energy of alizarin (without vibrational effects) for different solvation models*

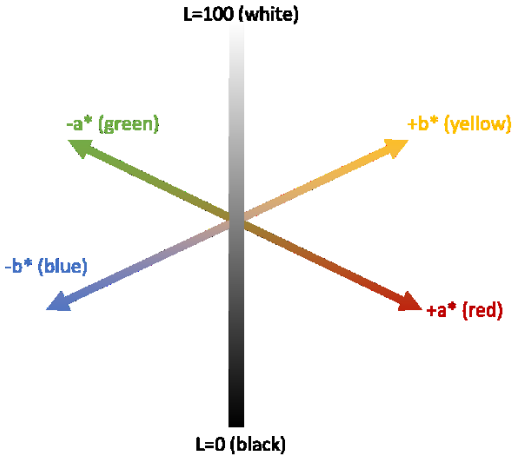
H_2A		HA^-		A^{2-}	
Calc. nm (eV)	Calc.-exp (eV)	Calc. nm (eV)	Calc.-exp (eV)	Calc. nm (eV)	Calc.-exp (eV)
416.4 (2.98)	0.10	527.4 (2.35)	-0.03	569.5 (2.18)	0.00
416.3 (2.98)	0.10	542.9 (2.28)	-0.10	566 (2.19)	0.01
405.3 (3.06)	0.18	548.7 (2.26)	-0.12	577 (2.15)	-0.03
409.7 (3.03)	0.15	584.2 (2.12)	-0.26	568.2 (2.18)	0.00
411.9 (3.01)	0.13	579.3 (2.14)	-0.24	569.2 (2.18)	0.00
405.8 (3.06)	0.18	541.4 (2.29)	-0.09	575.1 (2.16)	-0.02
431.6 (2.87)	-0.01	535.6 (2.32)	-0.06	573.9 (2.16)	-0.02
428.9 (2.89)	0.01	526.8 (2.35)	-0.03	575.5 (2.15)	-0.03
428.9 (2.89)	0.01	566.2 (2.19)	-0.19	569.5 (2.18)	0.00
425.6 (2.91)	0.03	546.4 (2.27)	-0.11	563.2 (2.2)	0.02
403.4 (3.07)	0.19	585.8 (2.12)	-0.26	560.3 (2.21)	0.03
410.8 (3.02)	0.14	564 (2.2)	-0.18	568.2 (2.18)	0.00
423 (2.93)	0.05	557.3 (2.23)	-0.15	565.7 (2.19)	0.01
412.8 (3)	0.12	540.9 (2.29)	-0.09	559.4 (2.22)	0.04
417.8 (2.97)	0.09	594.4 (2.09)	-0.29	583.6 (2.12)	-0.06
412.4 (3.01)	0.13	562.3 (2.21)	-0.17	554.9 (2.23)	0.05
413.9 (3)	0.12	535 (2.32)	-0.06	554.9 (2.23)	0.05
418.3 (2.96)	0.08	511.5 (2.42)	0.04	559.6 (2.22)	0.04
418.3 (2.96)	0.08	533.7 (2.32)	-0.06	553.7 (2.24)	0.06
406.8 (3.05)	0.17	521.3 (2.38)	0.00	555.2 (2.23)	0.05
427.3 (2.9)	0.02	541.3 (2.29)	-0.09	550.6 (2.25)	0.07
427.3 (2.9)	0.02	522.9 (2.37)	-0.01	554.9 (2.23)	0.05
435.5 (2.85)	-0.03	546.5 (2.27)	-0.11	557.9 (2.22)	0.04
437 (2.84)	-0.04	541.9 (2.29)	-0.09	539.8 (2.3)	0.12
438.3 (2.83)	-0.05	539.2 (2.3)	-0.08	542.4 (2.29)	0.11
431.8 (2.87)	-0.01	519.3 (2.39)	0.01	544.8 (2.28)	0.10
431.8 (2.87)	-0.01	504.2 (2.46)	0.08	540.3 (2.3)	0.12
417.6 (2.97)	0.09	522.7 (2.37)	-0.01	563.4 (2.2)	0.02
434.1 (2.86)	-0.02	524.5 (2.36)	-0.02	570.7 (2.17)	-0.01
416.5 (2.98)	0.10	512.5 (2.42)	0.04	549.2 (2.26)	0.08
424.9 (2.92)	0.04	525.9 (2.36)	-0.02	575.2 (2.16)	-0.02
403.3 (3.07)	0.19	545.1 (2.27)	-0.11	557.2 (2.23)	0.05
415.6 (2.98)	0.10	535.6 (2.32)	-0.06	550.3 (2.25)	0.07
422.7 (2.93)	0.05	521.9 (2.38)	0.00	562.7 (2.2)	0.02
430.7 (2.88)	0.00	529.1 (2.34)	-0.04	559.7 (2.22)	0.04
412.7 (3)	0.12	533.4 (2.32)	-0.06	569.9 (2.18)	0.00
415.6 (2.98)	0.10	503.1 (2.46)	0.08	545.7 (2.27)	0.09
383.3 (3.24)	0.36	529.1 (2.34)	-0.04	559.8 (2.22)	0.04
383.3 (3.24)	0.36	487.8 (2.54)	0.16	565.2 (2.19)	0.01
406.4 (3.05)	0.17	552.3 (2.25)	-0.13	565.1 (2.19)	0.01
399.8 (3.1)	0.22	548.6 (2.26)	-0.12	553.2 (2.24)	0.06
411.7 (3.01)	0.13	518.1 (2.39)	0.01	544.9 (2.28)	0.10
415.9 (2.98)	0.10	479.5 (2.59)	0.21	555.2 (2.23)	0.05
402.9 (3.08)	0.20	555.4 (2.23)	-0.15	552.4 (2.24)	0.06
402.9 (3.08)	0.20	514.6 (2.41)	0.03	551.4 (2.25)	0.07
433.1 (2.86)	-0.02	541.6 (2.29)	-0.09	577.2 (2.15)	-0.03
421.4 (2.94)	0.06	550.5 (2.25)	-0.13	568 (2.18)	0.00
432.5 (2.87)	-0.01	539 (2.3)	-0.08	552 (2.25)	0.07
430.4 (2.88)	0.00	522.6 (2.37)	-0.01	558.7 (2.22)	0.04
426.2 (2.91)	0.03	579.2 (2.14)	-0.24	569.6 (2.18)	0.00
Mean	0,09	**Mean**	-0,07	**Mean**	0,03

S.I.2*. Computed VE and its error deviation with 50 QM/MM snapshots for different alizarin forms.*

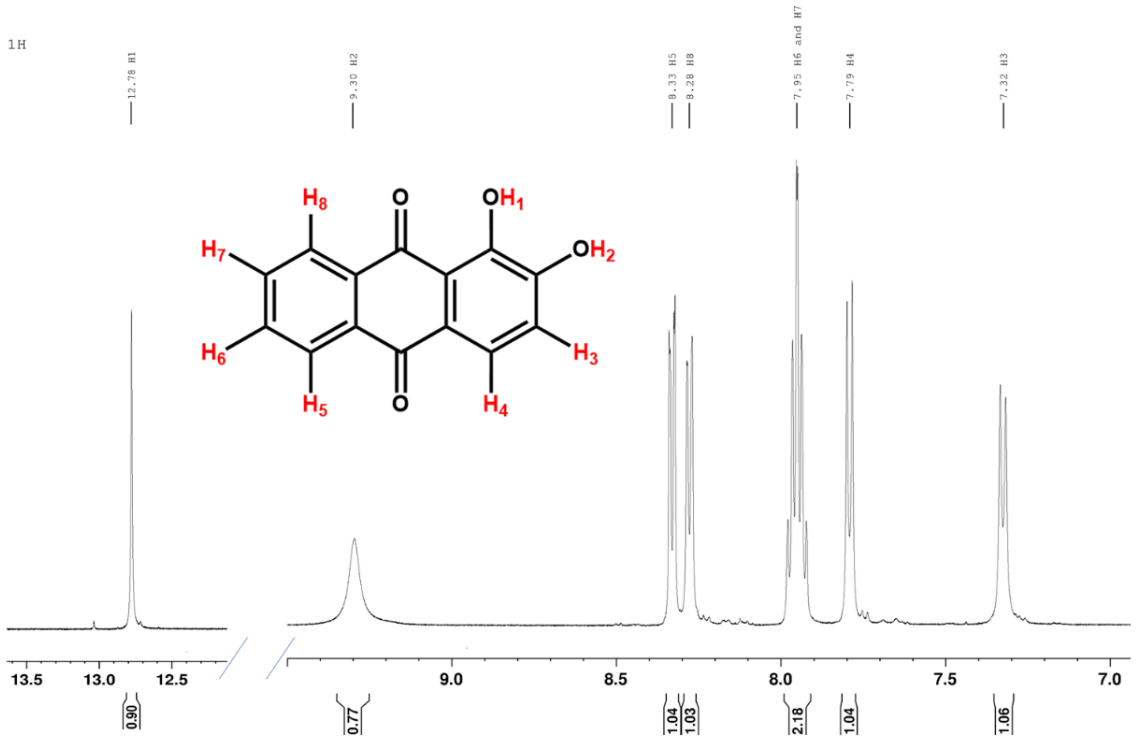


	Experimental (normalized)				Calculated VG (normalized)				ΔE_{2000}
Form	L*	a*	b*	color	L*	a*	b*	color	
H_2A	99.0	-13.4	68.4		90.8	8.1	80.3		**13.9**
HA^-	67.5	53.3	0.9		48.1	49.4	-46.4		**26.8**
A^{2-}	54.1	51.9	-50.0		64.2	48.6	-54.8		**9.5**

S.I.3. *(Top) Computed UV-VIS spectra of alizarin with AH or VG methods (dotted lines) and experimental data (full line). The vertical line is the VE. (Bottom) Predicted and computed colors with VG of different alizarin forms with CIELAB coordinates (L*, a*, b*) and the calculated color variation between computed and experimental color coordinates with* ΔE_{2000}.

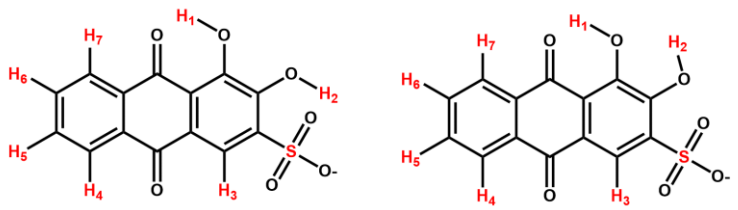


S.I.4. *CIELAB color space*



S.I.5. *Experimental ^{1}H NMR spectra of alizarin in acetone-D6 with its attribution*

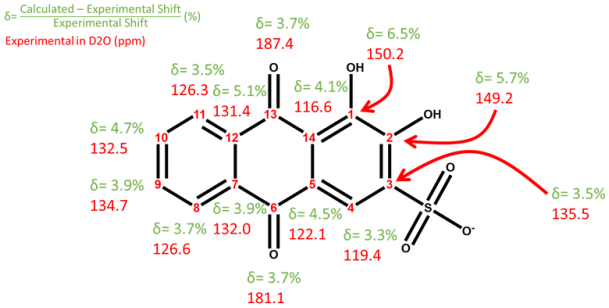
Attribution	Exp. (ppm) in acetone-D_6	Calc. (ppm) Tautomer 1	Error deviation (%) Tautomer 1	Calc. (ppm) Tautomer 2	Error deviation (%) Tautomer 2
H1	12.8	15.0	17.1	15.2	18.6
H2	11.9	13.2	11.1	8.1	32.0
H3	8.2	8.4	2.7	8.5	4.0
H4	8.3	8.5	2.7	8.6	3.2
H6-H5	7.9	8.0	1.3	8.2	3.3
H6-H5	7.9	8.0	1.3	8.1	2.8
H7	8.3	8.5	2.5	8.6	3.5

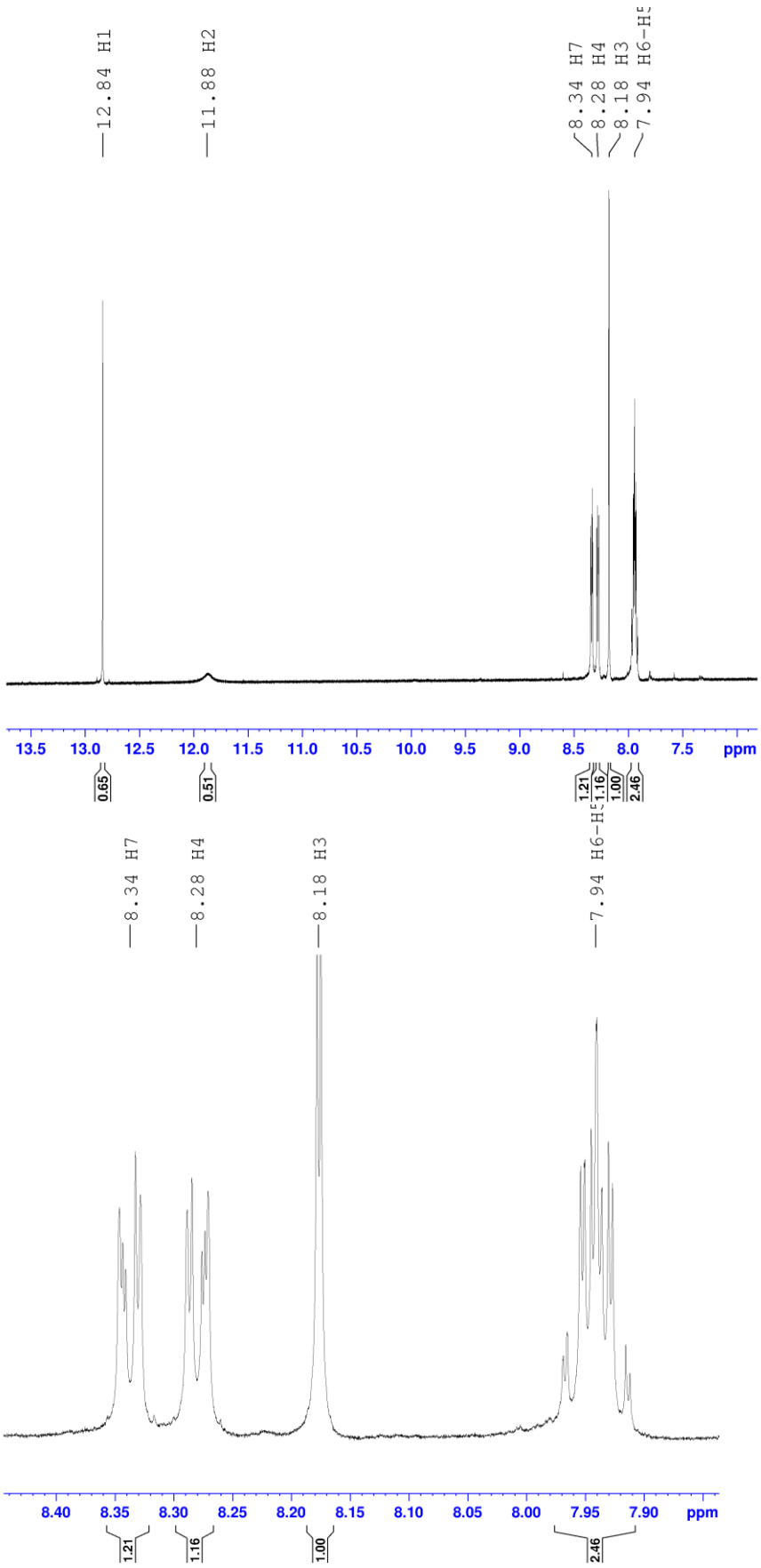


S.I.6. *Computed and experimental chemical shift of alizarin red S in acetone-D_6 with the respective error deviation in ppm or % for the two most stable tautomers.*

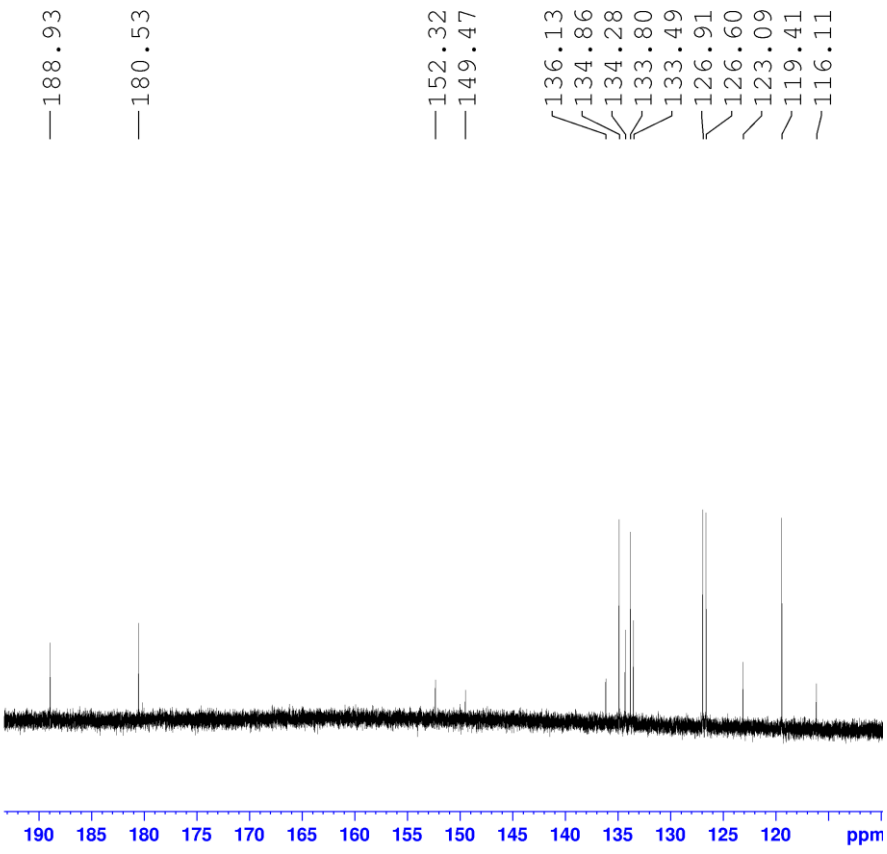
S.I.7. Table of computed ^{13}C shifts for alizarin red S for different forms.

			CALCULATED SHIFT (PPM)				ABS ERROR DEVIATION (%)			
C ATTRIBUTION	C Output number	Exp. shift in acetone D6	H_3ARS	H_2ARS	$HARS^-$	ARS^{2-}	H_3ARS	H_2ARS	$HARS^-$	ARS^{2-}
1	1	149.5	157.1	157.7	175.0	180.2	5.1	5.5	17.1	20.6
2	2	152.3	157.2	160.2	163.9	178.3	3.2	5.2	7.6	17.0
3	3	133.5	132.3	140.2	134.5	143.3	0.9	5.0	0.8	7.3
4	4	119.4	124.5	124.4	128.6	127.0	4.3	4.2	7.7	6.4
5	5	123.1	127.8	127.1	115.3	120.0	3.8	3.3	6.3	2.5
6	10	180.5	187.8	187.4	180.5	181.3	4.0	3.8	0.0	0.4
7	9	134.3	138.0	138.5	138.3	140.8	2.8	3.1	3.0	4.9
8	14	126.9	132.0	131.3	129.7	129.3	4.0	3.4	2.2	1.9
9	13	136.1	141.0	139.6	144.7	131.8	3.6	2.5	6.3	3.2
10	12	134.9	139.3	138.3	141.5	131.5	3.3	2.5	4.9	2.5
11	11	126.6	131.1	130.5	129.2	127.9	3.6	3.1	2.0	1.1
12	8	133.8	137.5	138.2	137.9	136.7	2.7	3.3	3.0	2.1
13	7	188.9	195.5	194.2	195.3	183.4	3.5	2.8	3.3	2.9
14	6	116.1	122.1	120.4	118.7	120.3	5.2	3.7	2.2	3.6

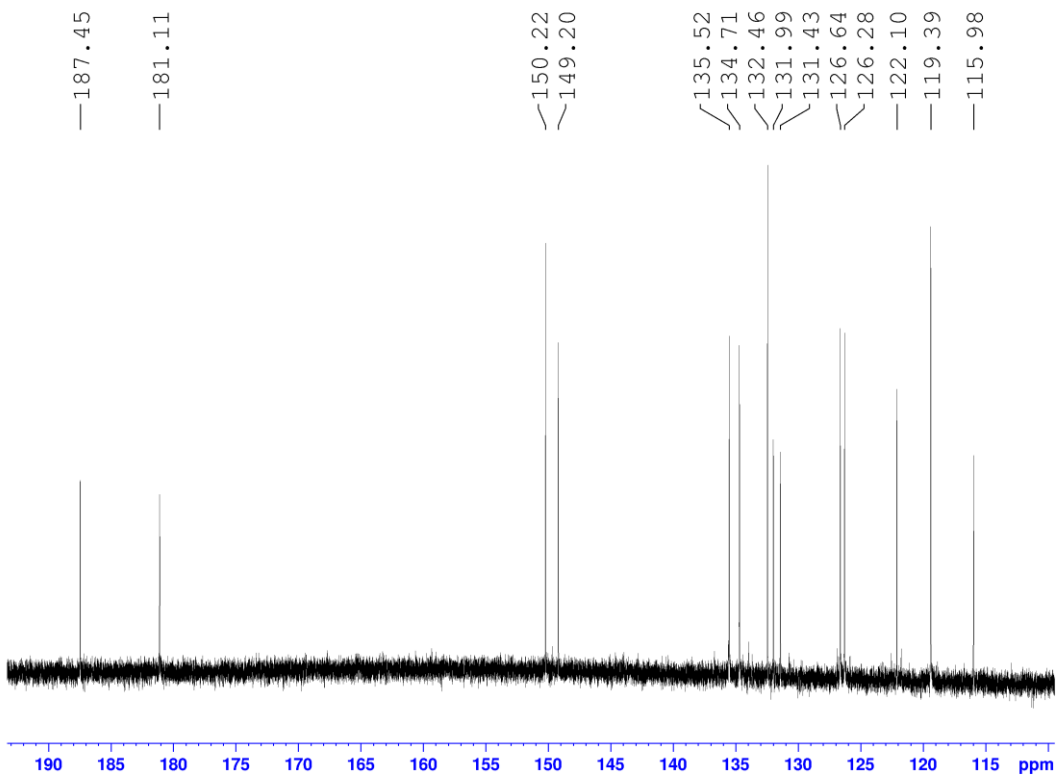

δ= Calculated – Experimental Shift / Experimental Shift (%)
Experimental in D2O (ppm)
δ= 3.7%
187.4
δ= 6.5%
150.2
δ= 3.5%
126.3
δ= 5.1%
131.4
δ= 4.1%
116.6
δ= 5.7%
149.2
δ= 4.7%
132.5
δ= 3.9%
134.7
δ= 3.5%
135.5
δ= 3.7%
126.6
δ= 3.9%
132.0
δ= 4.5%
122.1
δ= 3.3%
119.4
δ= 3.7%
181.1
OH
OH
O
O
O
O
S
O-
1
2
3
4
5
6
7
8
9
10
11
12
13
14



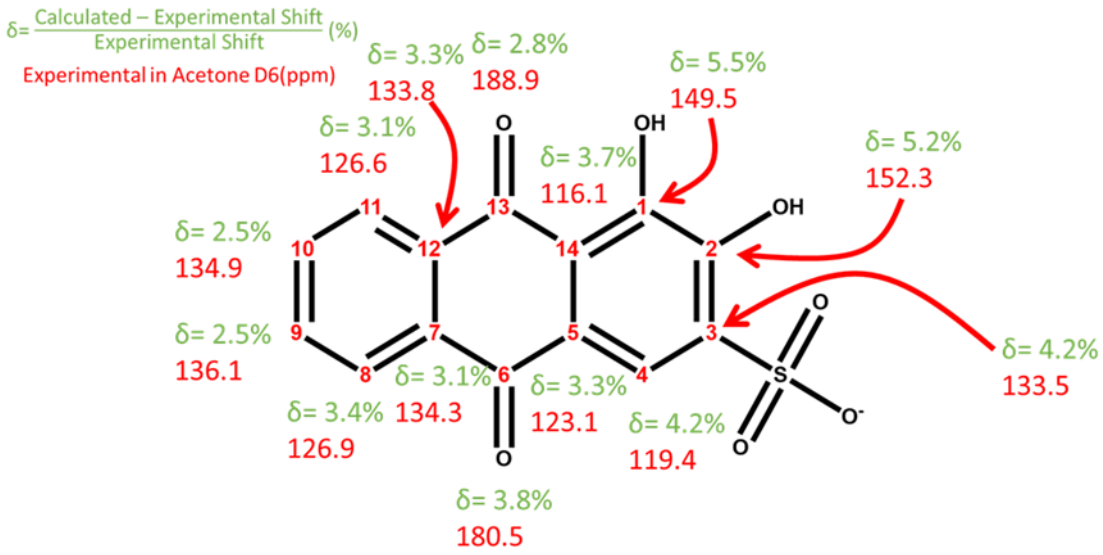
S.I.8. *Experimental ^{1}H NMR spectra of alizarin red S in acetone-D_6 (top) and zoomed (bottom).*



S.I.9*. ^{13}C experimental NMR spectra of alizarin red S in acetone-D6.*



S.I.10. *^{13}C experimental NMR spectra of alizarin red S in D_2O.*



S.I.11. *Experimental chemical shifts (red values) and the error deviation (green in %) with respect to the computed ^{13}C NMR shifts in acetone-D6 (top) or D_2O (bottom).*

SUMMARY

Article 2: "In-silico colour prediction process for natural dyes in Madder."

In this article, we used the previously optimised protocol to build a database of colourimetric characteristics and spectroscopic signatures for 31 colourants found in various madder species at different pH levels. This enables access to a theoretical standard that is missing due to experimental difficulties. There are six main species of madder: *Rubia peregrina* L., *Rubia tinctorum* L., *Rubia cordifolia* L., *Rubia Akane*, *Rubia sikkimensis* Kurz, and *Rubia yunnanensis* Diels. While these species share similar molecules, their compositions differ in proportions. Roughly 30-68 distinct colourants have been identified in the literature. Our previous protocol was validated on about ten colourants for which experimental data were available, consistently showing a calculation-experiment deviation of less than 0.2 eV, within the accepted theoretical error margin.

Using the list of 31 colourants identified by Richard S. Blackburn and the validated protocol, we established a comprehensive database including absorbance spectra, λ_{max} values, and colour predictions under the two most common conditions used in madder extraction or application: acidic and neutral aqueous solutions. This enabled us to obtain spectral and colour information for molecules that are otherwise difficult to characterise experimentally due to instability, synthetic challenges, or commercial unavailability, such as munjistin and pseudopurpurin.

This work highlighted structure–property relationships such as the influence of pH and glycosidic hydrolysis on spectral shifts. At acidic pH, madder colourants tend to be yellowish or colourless, while at neutral pH, some exhibit a redder hue due to bathochromic (red) shifts in their absorption spectra. Glycosidic hydrolysis also induces slight red shifts, potentially explaining the traditional maceration step in madder preparation, where natural enzymes hydrolyse glycosides, enhancing the red colouration. These findings were published in the second article.

ARTICLE 2

In-silico color prediction process for natural dyes in Madder

Thanh Huyen Linh TRAN[a], Maguy JABER[a,b], Romain BERRAUD-PACHE[a*]

[a] Laboratoire d'Archéologie Moléculaire et Structurale (LAMS), Sorbonne Université, CNRS, UMR 8220, Paris, 75005, France
[c] IUF-Institut Universitaire de France, Paris, France
[*]corresponding author: romain.berraud-pache@sorbonne-universite.fr

Dyes and Pigments 2025;237:112701. https://doi.org/10.1016/j.dyepig.2025.112701.

ABSTRACT

In the field of cultural heritage, accurately identifying the colors produced by madder root dyes without alteration is critical. In fact, textile dyeing using extracts from their roots has been performed for centuries. Many examples of madder dye remain today in museum and conservation collections. Madder is an interesting natural source of dye because it contains up to 68 different colorants. However, many of these dyes are extremely sensitive to hydrolysis, decarboxylation, and degradation from enzymes, extraction chemicals, and processing temperatures. The valuable chemical information embedded in the dye structure may be lost if extraction and analysis are too harsh, resulting in inaccurate color prediction with large color differences. Thus, they are often overlooked as colorants in historical textiles. Over the past decades, traditional extraction techniques have been used to analyze textiles but they are invasive and destructive. As a result, this method is expensive, time-consuming, and in some cases is not suited for the analysis of artwork. But in the last decades, computer techniques have been presented as a powerful tool that enables the prediction of UV-VIS signature spectra and color coordinates with high accuracy using Time-Dependent Density Functional Theory. The aim of this study is to elaborate an efficient color prediction process of madder's root without any chemical degradation of dyes. It will enable the creation of a robust analytical database to identify madder colors in heritage objects. Here, we provide a rapid method to obtain a complete database of predicted UV-VIS spectra and color of anthraquinone derivatives found in Madder. In the context of historical textiles, the detailed color of natural colorants present in the artwork is of paramount importance for conservation and restoration purposes.

1. INTRODUCTION

For centuries, colorants extracted from the roots of various *Rubia* species have been used as dyestuffs, collectively known as "madder", which are frequently encountered in archaeological analyses of prehistoric and historic textiles as well as classical texts and alchemical writings. For example, we have listed in Table 1 of supplementary information (S.I.) the identified madder species used for dyeing textiles and pigments fabrication.

Although madder has been a well-known dye plant since ancient times, its cultivation, processing, trade, molecular properties, and recipes were not well documented until recent history. The earliest occurrence of publication about madder dye was done by the Greek philosopher Democritus (c. 460 – c. 370 BC) is believed to have written four texts on dyeing methods, though these texts are now lost. Another Greek, Pedanius Dioscorides (around 50-70 C.E), documented the madder's dyeing properties. [97,98] Overall, little is known about how lake pigments were made in the ancient world. In the 16th century, madder recipes were more comprehensively documented and became part of national heritage. This historical color is created through a laborious dyeing process that has been shrouded in mystery and secrecy for centuries, due to their commercial value. It was not until the 19th century that these secrets were gradually decoded.[5] The best recorded madder dye data at that time was handwritten in a manuscript by Arthur Henry Newton, between 1846 and 1858, focusing on pigment manufacture. In this book, he named Madder under various names, like Madder Rose, Madder Lake, and Madder Carmine, but their molecular structures were not recorded. [30] Documenting about the identification of colors derived from madder roots is a challenging task and remains an active area of research to this day. Some articles from the state of the art highlight the lack and need for information on minor components or glucosidic molecules.[35,38,39] Indeed, identifying madder components in textiles or lake pigments is difficult due to their complex composition, low dye concentrations, and degradation.[18,20,33,42,44,57,59,181,182,185,187,190,223] Thanks to efforts to improve detection, many modern techniques have emerged, which can be classified into two groups: non-invasive techniques, such as Raman Microscopy, vibrational spectroscopies, Fiber Optic Reflectance Spectroscopy (FORS), and reflectance with Hyperspectral Imaging [45,49,53,182,189,190] or invasive techniques to determine composition like HPLC-DAD (Diode Array Detector) coupled with mass spectra and Surface-enhanced Raman spectroscopy (SERS). [36,38,46,53,187] Thanks to the development of those techniques, researchers have been able to study the composition of madder roots more precisely. In particular, Ram Singh *et al.* reported up to 68 anthraquinonoid compounds extractable from madder roots, although they noted that many may be artefacts of the extraction process. [224] E. S. B. Ferreira *et al.* summarized compounds in six main *Rubia* species, excluding those formed during extraction (S.I. Table 2). [225] Some compounds are present in low concentrations or result from harvesting and processing, with levels varying by age. [33] Another study reported 36 anthraquinones in *R. tinctorum* L .[33] But most of the studies are published with up to 15 molecules because they are the main ones, other are present in small amounts and are degraded easily. [33,57,59,100]

The analysis and characterization of these molecules can significantly enhance the identification of plant origins and unveil the hidden secret of historical artworks, aiding curators in uncovering information about the origin of artwork, the trade routes of dyes, the techniques used, and developing a suitable conservation strategy and restoration by understanding the original material. [181,182]

Unfortunately, in these studies, a complete database of UV-VIS spectroscopic signatures and the corresponding color of all madder molecules is still lacking. The complexity of experimental extraction techniques, which often degrade components, especially glucosides, contributes further to this gap. [35,38]

It is important to note that experimentally identifying the molecules responsible for color is not always an easy task because madder molecules contain glucosidic counterparts which degrade and hydrolyze easily by either enzyme present in the plant or by the employed extraction techniques. Furthermore, some madder dyes cannot be purchased or synthesized which makes difficult its characterization. With the rising power of computers and the development of quantum chemistry, computational chemistry has appeared as an appealing tool to circumvent experimental difficulties [103] and many molecules have been easily studied in vacuo on a computer. [20,22,33,35,44,56,147]

Several studies have focused on the prediction of UV-VIS spectra properties of rigid dyes. D. Jacquemin et al., Fehér et al. and Anouar et al. are well known for their extensive benchmarking on functionals for predicting UV-VIS spectra of organic chromophores including anthraquinones. [108,117,226–229] These benchmarks have confirmed that hybrid functionals (such as B3LYP or PBE0) often perform well for anthraquinones using Time-Dependent Density Functional Theory (TD-DFT). C. Ilaria, C. Adamo et al. have extended this analysis of rigid dyes by focusing on the perceived color and the vibrational effects. [118,199,200] But few studies have reported the prediction of UV-VIS spectra of molecules with glucosides. Giordano Mancini et al. have reported the prediction of emission and absorption spectra of flexible dye in solution by using General Liquid Optimized Boundary (GLOB) model and Molecular Dynamics followed by averaging of TD-DFT after multiple snapshots. [120] But none of these studies have reported the prediction of UV-VIS spectra properties or color of glycosidic anthraquinones, which are one of the main molecules in madder plant.

In this study, we aim to close the gap caused by the above-mentioned issues, by using computational methods to generate the spectra, providing a more complete and accessible resource. We used the Density Functional Theory (DFT) and its Time-Dependent variant (TD-DFT) for about 31 molecules present in Figure 1.

Thanks to this protocol, we have obtained a comprehensive database of UV-VIS spectroscopic signatures and the corresponding color codes of the main madder molecules giving a more complete theoretical reference standard for rapid and efficient madder species identification.

2. MATERIALS AND METHODS

2.1. COMPUTATIONAL METHOD

2.1.a. Modeling the UV–VIS spectra

The Figure 1 presents all the molecules computed in this study.

All calculations were carried out using the version 5.0.3 of the ORCA quantum chemistry program package. The version 6.5.1 of xTB was used for deprotonated forms exploration. [204,205]

For the neutral forms, implicit water solvation using the linear response Conductor-like Polarizable Continuum Model (C-PCM) was chosen. [212] For deprotonated forms, a

combined approach involving both implicit and explicit water molecules (3 water molecules per hydroxyl group and/or carboxy group) was used. The water molecules were positioned to establish hydrogen bonds with each hydroxyl group and carboxyl group. For purpurin, the di-deprotonated form was chosen to be closer to experimental data. For quinizarin, the molecule is symmetric and has the same pKa, so the two groups were both deprotonated.

For TD-DFT calculations concerning the UV-VIS spectrum, the PBE0 functional was used. [108,114,166] This choice is based on previous benchmark studies and is more computationally efficient than RI-B2PLYP or B3LYP without much deviation (see S.I. Table 3). [107,168,208,226,230,231] The triple zeta def2-TZVP basis set, coupled with matching auxiliary basis sets, was consistently employed throughout all the computations.[151]

The computed λ_{max} corresponds to the Vertical Energy (VE) of the first bright transition with the highest oscillation strength. It corresponds to an $S_0 \rightarrow S_1$ electronic transition mainly between the HOMO and LUMO and corresponding to a mixture of a $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$transition (depending on the deprotonation state).

The vibrational contributions, including Herzberg–Teller (HT) effects [176,178,216], were computed using the Excited State Dynamics (ESD) module within ORCA. [175] The impact of the HT effects on the spectra as well as their computation timing is provided in S.I.4. The Adiabatic Hessian (AH), a more complete and accurate method, where both GS and ES geometries and frequencies are considered [217], was used (unless stated otherwise). For some molecules containing sugar moieties, the AH spectra were computed by substituting the sugar with a methoxy group. This approach is explained in the discussion section. The linewidth was created using a Gaussian distribution and was set to 700 cm^{-1} for the whole database in Table 1. The maximum intensity of the computed spectra was scaled to match the maximum intensity of the experimental data, allowing for a more effective comparison of the spectral shapes.

2.1.b. Modeling the color

The CIELAB color coordinates were calculated for the UV-VIS spectra with an in-house script with illuminant D65. [134,218,219] The computed colors were normalized with respect to the molecule that has the most intense absorbance. The ΔE_{2000} formula (see S.I.5) was used to compute the color difference between the computed and experimental results. A ΔE_{2000} value below 2 indicates no visible difference, 2-10 suggests a noticeable but subtle difference, and values above 10 imply a more pronounced difference.

2.1.c. Modeling the mixture

To validate this computational protocol, we have computed the spectra of a mixed ratio of alizarin and purpurin and compared them to experimental spectra. The linewidth was created using a Gaussian distribution and was set to 2000 cm^{-1}. Three mass ratios of alizarin: purpurin tested were: 30:70, 50:50, and 70:30. The experimental pH and pKa=5.5 of alizarin (AH_2/ AH^-) and purpurin (pKa_1=4.6 for H_3P/ H_2P^- and pKa_2=5.5 for H_2P^-/HP^{2-}) were taken into account in the percentage following this equation:

$$pH = pKa + \log\left(\frac{[Base]}{[Acid]}\right)$$

For the ratio 50:50, the percentage contributions in the computed spectra are 25% AH_2, 25% AH^- and 50% H_2P^-. For the ratio 70:30, there is 16% AH_2, 54% AH^-, 7% H_2P^- and 23% HP^{2-}. For the ratio 30:70, there is 8.5% AH_2, 21.5% AH^-, 20% H_2P^-, and 50% HP^{2-}. The computed spectra were added by multiplying the intensity with these ratios, then added together and finally scaled to the maximum intensity of experimental data.

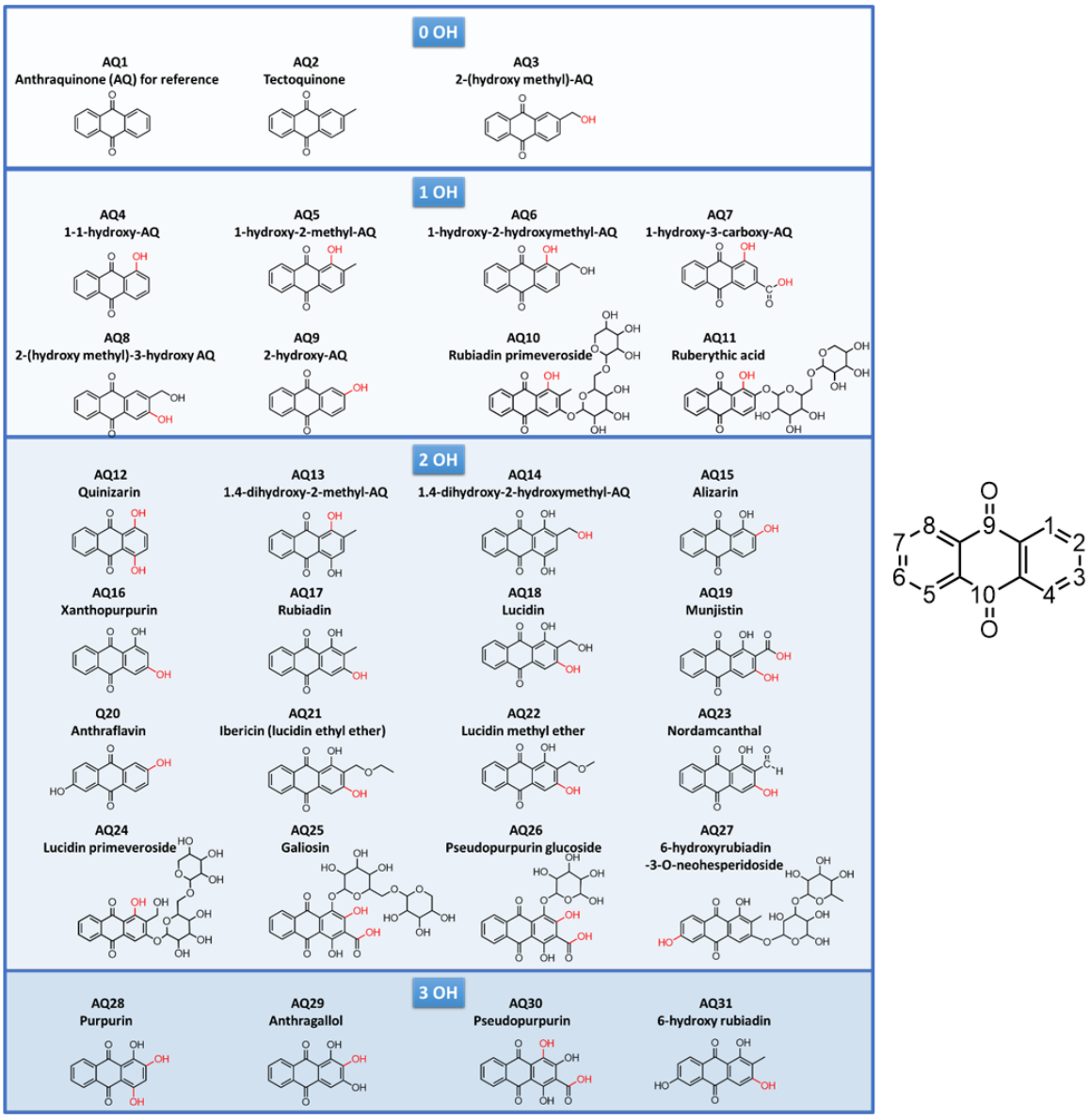


Figure 1. *Madder's anthraquinones (AQ) structures and numbering used in this article. The red hydroxyl groups are the ones that are deprotonated (carboxyl groups are always deprotonated as we are in water).*

2.2. EXPERIMENTAL METHOD

2.2.a. Reagents

Alizarin (97 % purity), purpurin (90%) and lucidin primeveroside (Phyproof) were purchased from Sigma Aldrich and used without further purification.

2.2.b. Procedure

30 mg of either alizarin (1,2-dihydroxyanthraquinone) or purpurin (1,2,4-trihydroxyanthraquinone) was dissolved in 75 mL of deionized MilliQ water and then boiled at 90°C for 2 hours. Then, the solutions were centrifuged for 15 minutes at 6000 rpm and room temperature. The supernatant from each preparation was collected, and then mixed in different mass ratios: 30:70, 50:50, and 70:30. The solutions were stirred for 5 min. Lucidin primeveroside was diluted in water at 0.5mg/mL.

2.2.c. UV-VIS spectra

Their UV-VIS spectra were measured immediately using an Ocean Optics Flame spectrometer (Model: FLMS00699) with a wavelength range from 300 to 850 nm. The integration time for each measurement ranged from 5 ms to 10 ms. The light source used was a Deuterium-Tungsten Halogen lamp covering the range of 300 nm–850 nm. The blank was done using deionized water. This light source was coupled to the spectrometer using 400 µm diameter optical fibers and positioned through a 1 cm quartz cuvette holder. The acquisition of spectra was facilitated by the Ocean View software.

3. RESULTS AND DISCUSSION

3.1. Madder coloring molecules belong to the anthraquinone (AQ) family.

Molecules from the Anthraquinone family exist in two primary forms in planta: glycosidic (see Figure 1: AQ10,11,24 to 27), where a sugar moiety is attached, and aglycones (all other molecules in Figure 1), where the glycoside is replaced by a hydrogen atom. Glycosidic compounds are more concentrated in younger madder plants and are more soluble in water. [33,44,52,59] As the plant ages or during dye preparation, glycosides are hydrolyzed by enzymes or during chemical extraction to enhance the dye color, as aglycons produce a redder hue.[20]

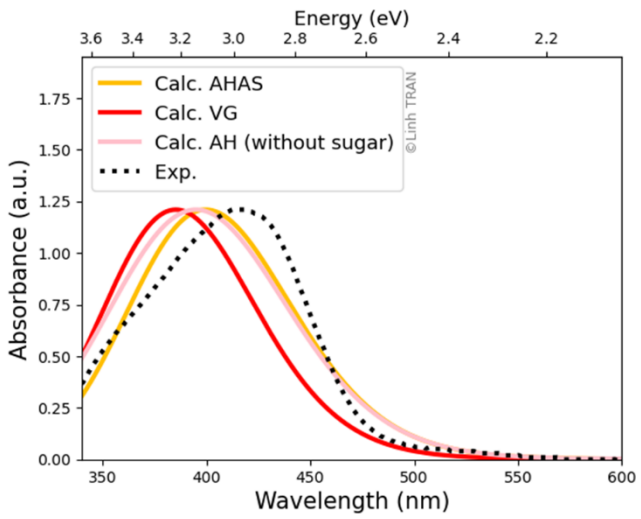
Some extraction techniques for madder dyes imply using an acidic solvent or water. [19,55,202,223,232] Therefore, we have chosen to compute the colors and λ_{max} at two pHs: in acidic conditions pH <5 and close to the one of distilled water where the carboxylic group is deprotonated as well as 1 hydroxyl group.

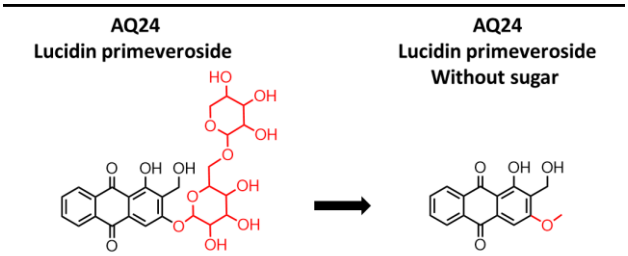
First, as a case study, we predict the spectrum of lucidin primeveroside, one of the main glycosidic molecules present in madder, and discuss the different methods used. Then, using the most accurate computational procedure available, we will provide a complete database (Table 1) on the color properties and UV-VIS spectroscopic signatures of the 31 most encountered madder molecules (Figure 1). Our protocol will be validated by comparing our results with available experimental data for 10 molecules (see S.I.6), and by mixing various ratios of alizarin and purpurin to simulate the different proportions found in madder species (Figure 4).

3.2. Computation method for glycosidic, molecules with sugar moieties (lucidin primeveroside as a study case)

Madder contains several glycosidic anthraquinones together with their corresponding aglycones form (without sugar moieties). Among these, lucidin primeveroside, AQ24, is particularly interesting, as it is prone to degradation when extracted.[56,233] In addition, due to the challenges in its synthesis, it is both very expensive and available in limited quantities, making its manipulation and experimental study complicated. This section explores various computational protocols for predicting its UV-VIS spectra and color.

The vibronic spectrum was modeled using the Vertical Gradient (VG), Adiabatic Hessian After Step (AHAS) and Adiabatic Hessian (AH) model.[234] However, the large flexibility of the sugar moieties results in a large geometry difference for the GS and ES minima thus breaking the harmonic approximation of the AH model and the spectrum itself. The VG and AHAS model, computed by only performing a small displacement on the Excited State Potential Energy Surface (ES PES), were however successful. In order to reach the best accuracy, the sugar moieties were replaced with methoxy groups, a reasonable approximation for the glycosidic groups, allowing successful AH spectrum calculation. The results, including deviations from the experimental spectrum, are shown in Figure 2. The experimental spectrum of lucidin primeveroside in water exhibits a peak at 3.02 eV (411 nm). The VG method predicts a peak at 3.22 eV (385 nm), while the AHAS and AH methods predict maxima at 3.10 eV (399 nm) and 3.14 eV (395 nm), respectively. The VG method shows the largest deviation from the experimental value (0.20 eV), whereas the AH (without sugar) and AHAS methods yield smaller deviations (0.12 eV and 0.08 eV, respectively). The AH method, which approximates sugar moieties with a methoxy group, is therefore validated by its close agreement with both the experimental data and the AHAS method.





Lucidin primeveroside	Calc. λ_{max}	Error deviation (calc-exp.)
VG	3.22 eV (385 nm)	0.20 eV (-26 nm)
AHAS	3.10 eV (399 nm)	0.08 eV (-12 nm)
AH (without sugar)	3.14 eV (395 nm)	0.12 eV (-16 nm)
Exp.	3.02 eV (411 nm)	-

©Linh TRAN

Figure 2. *(Left) Calculated VG (red full line), AHAS (yellow full line), AH without sugar (pink full line), and experimental (black dotted line) UV-VIS spectra for lucidin primeveroside AQ24. The AH spectrum was done by replacing the sugar moieties with a methoxy group. (Right) The scheme shows the molecule structures with or without sugar. The table below shows the error deviation between the computed and experimental spectra.*

The computed color coordinates are shown in Figure 3. The ΔE_{2000} is 6.3 for VG method, and is significantly lower for AHAS (0.8) and AH without sugar (0.9) methods. The computed color done with AHAS and the approximated AH (without sugar) is therefore closer to the one perceived experimentally.

CIELAB coordinate	Calculated				Experimental				Color difference
Method	L*	a*	b*	color	L*	a*	b*	color	ΔE_{2000}
VG	98.8	-11.9	38.8		97.9	-13.1	58.4		6.3
AHAS	97.4	-12.0	57.7		97.9	-13.1	58.4		0.8
AH (without sugar)	97.4	-11.7	56.1		97.9	-13.1	58.4		0.9

©Linh TRAN

Figure 3. *Calculated and experimental CIELAB color coordinates for lucidin primeveroside AQ24. The color deviation was computed with ΔE_{2000}.*

3.3. Color and λ_{max} of madder compounds

In this section, we aim to further validate our computational results by comparing them with the experimental data. The computed Vertical Energy (VE) values have been compared to the experimental data available in the literature[59], showing an error deviation of less than 0.10 eV. Details are provided in the S.I.6. This proves the accuracy of this method, even for more complex and bigger molecules containing glucosidic parts.

Understanding and predicting the color of these molecules involves considering various factors such as solvent effects, substituents, conformers, and pH conditions. The calculated λ_{max} and color for the acidic and first deprotonated form of 31 madder dyes, are shown in Table 1. Anthraquinone (AQ) is provided for reference. The numbers correspond to structures in Figure 1.

Table 1*. Computed color, λ_{max} of madder molecules in their acidic and basic form (first deprotonation). The numbers correspond to the structures in Figure 1. The corresponding color and their L*,a*,b* color coordinates are presented. Anthraquinone (AQ) is provided for reference. * stands for AH computation with sugar moieties replaced by a methoxy group.*

N°	ACIDIC FORM					DEPROTONATED FORM				
	λ_{max} nm (eV)	Color	Color coordinates L*	a*	b*	λ_{max} nm (eV)	Color	Color coordinates L*	a*	b*
AQ1	325 (3.82)		100	0	0	325 (3.82)		100	0	0
AQ2	405 (3.06)		100	0	0	-	-	-	-	-
AQ3	402 (3.08)		100	-1	1	409 (3.03)		100	-7	16
AQ4	370 (3.35)		100	0	1	480 (2.58)		94	5	55
AQ5	361 (3.43)		100	0	0	467 (2.66)		96	-9	91
AQ6	356 (3.48)		100	0	0	469 (2.64)		95	-7	88
AQ7	363 (3.42)		100	0	0	454 (2.73)		97	-15	88
AQ8	360 (3.44)		100	0	0	455 (2.73)		98	-12	57
AQ9	371 (3.34)		100	0	0	458 (2.71)		98	-13	56
AQ10	388 (3.20)*		100	-1	2	486 (2.55)		88	20	64
AQ11	411 (3.02)*		99	-19	62	484 (2.56)*		90	15	83
AQ12	453 (2.74)		95	-11	127	529 (2.34)		64	84	-18
AQ13	456 (2.72)		95	-8	129	523 (2.37)		68	77	-2
AQ14	455 (2.73)		95	-11	126	448 (2.77)		97	-15	107
AQ15	402 (3.08)		99	-16	43	520 (2.38)		76	60	10
AQ16	396 (3.13)		100	-9	19	499 (2.48)		90	22	24
AQ17	401 (3.09)		100	-12	28	477 (2.60)		93	9	44
AQ18	403 (3.08)		100	-10	24	478 (2.59)		92	11	44
AQ19	403 (3.08)		100	-11	26	422 (2.94)		99	-11	33
AQ20	332 (3.73)		100	0	0	438 (2.83)		100	-9	24
AQ21	401 (3.09)		100	-11	25	458 (2.71)		97	-9	70
AQ22	399 (3.11)		100	-12	27	457 (2.71)		97	-12	69
AQ23	413 (3.00)		100	-14	34	443 (2.80)		99	-17	69
AQ24	407 (3.05)*		100	-13	32	467 (2.66)		93	5	65
AQ25	419 (2.96)*		99	-19	59	493 (2.52)		86	32	28
AQ26	362 (3.43)*		100	-4	10	400 (3.10)*		99	-10	32
AQ27	396 (3.13)		100	-12	30	409 (3.03)		99	-10	34
AQ28	449 (2.76)		95	-10	125	532 (2.33)		55	93	-21
AQ29	379 (3.27)		100	-3	6	538 (2.30)		78	45	-17
AQ30	438 (2.83)		98	-21	101	490 (2.53)		81	39	68
AQ31	415 (2.99)		99	-17	47	469 (2.64)		96	0	43

©Linh TRAN

Impact of substituents' nature and position

For acidic forms, the results indicate that the color hue is predominantly yellow, with varying intensities (from transparent to deep yellow). The most intense yellow hues are observed for AQ12 to AQ14,28,30. These compounds share a common feature: the presence of hydroxyl groups at the 1 and 4 positions, adjacent to the ketone groups. Anthraquinones bearing hydroxyl groups at the 1 and 4 positions exhibit a pronounced redshift in their absorption spectra compared to analogs lacking these groups, which accounts for their more pronounced yellow hues. This effect arises from the formation of intramolecular OH···O=C hydrogen bonds at the peri positions of the anthraquinone. These hydrogen bonds enhance the electron-accepting character of the quinone unit and

promote a hydrogen bond–assisted charge transfer interaction, thereby inducing a significant bathochromic shift in the UV–Vis spectra. This explanation is supported by studies showing that the intramolecular hydrogen-bonded derivatives have altered frontier orbital energies leading to a redshift.[235]

The S.I.7 shows the difference densities of these molecules during the absorption transition. On the other hand, compounds lacking hydroxyl groups at the 1 or 4 positions exhibit the least intense hue, appearing almost transparent. This is the case for molecules AQ1 to AQ10. This underscores the importance of these specific positions on the color of AQ compounds in acidic environments.

A different behavior is observed for deprotonated forms at higher pH. When the dyes are deprotonated, the most intense hues are not always observed for the dyes with a hydroxyl group in positions 1 and 4 since the loss of hydrogen breaks the pseudo-cycle made by the hydrogen bond with the ketone. For example, the deprotonation occurs at the position 1 for AQ30.

Other substituents like methoxy or carboxy groups influence the λ_{max} values. For instance, for acid forms, AQ5,6 are red-shifted compared to AQ8,20. The presence of carboxy groups as seen in AQ7 also leads to redshifts, likely due to the increased electron-withdrawing effects enhancing the conjugation. For basic forms, molecules with hydroxyl and carboxyl groups at specific positions, such as AQ12,28,29, show significant redshifts upon deprotonation. These groups facilitate extensive π-conjugation and stabilization of the excited state. AQ6 shows one of the most significant redshifts, from 356 to 469 nm, highlighting its strong electron-donating properties when deprotonated. Similarly, AQ29 has the most intense red-violet color with a λ_{max} of 538 nm and show a shift of 159 nm upon deprotonation. Its most stable form is when the hydroxyl at position 2 is deprotonated.

Impact of hydrolysis of glycoside

As stated before, the glycosidic molecules hydrolyze to aglycone and tend to give a redder hue. The table 2 represents the shift in maximum absorbance and color before and after hydrolysis and in acidic or basic form.

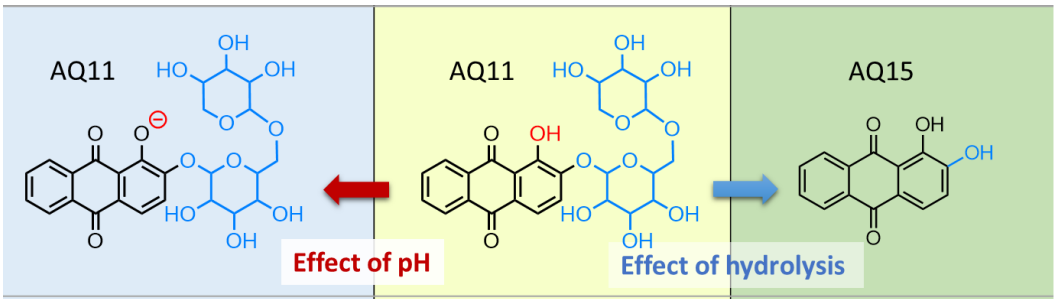
At acid pH (glycosidic form), the red hue is not observed, and the π-conjugation system remains unchanged after hydrolysis, resulting in similar colors. However, a slight redshift is noted in the computed λ_{max} for some of the aglycones. For instance, AQ10 with two glucoside groups has a λ_{max} of 388 nm, while the corresponding AQ17 without sugar has a λ_{max} of 401 nm. AQ24 with a sugar moiety has a λ_{max} of 407 nm, the corresponding molecule without sugar AQ18 has a λ_{max} of 403 nm. AQ25 has a λ_{max} of 419 nm, and AQ30 has a λ_{max} of 438 nm. AQ27 with a disaccharide has a λ_{max} of 396nm and shifts to 415nm after hydrolysis (AQ31). AQ24 is reported to be less prone to hydrolysis compared to AQ10 or AQ25. [56,223] Overall, aglycones exhibit a slight redshift compared to their glycosidic counterparts.

Impact of deprotonation (at higher pH)

For basic forms, both glycosidic and aglycone forms exhibit redshifts upon deprotonation. For instance, AQ11 shifts from 411nm to 484nm. AQ10 shifts to 486 nm from 388 nm, and AQ24 shifts to 467 nm from 407 nm. AQ25 has a λ_{max} of 419nm and 493nm upon deprotonation. AQ27 goes from 396 nm to 409 nm at higher pH. This can be explained by the extension of the conjugation system after deprotonation. This can lower the HOMO-LUMO energy gap, shifting the absorption towards a longer wavelength.

Overall, deprotonation generally causes a redshift (longer wavelengths) in λ_{max} values and color.

Table 2. *Effect of hydrolysis and pH on λ_{max} (nm) and color.*



Deprotonated form	Glycosidic form	Aglycon form
AQ10: 486	AQ10: 388	AQ17: 401
AQ24: 467	AQ24: 407	AQ18: 403
AQ25: 493	AQ25: 419	AQ30: 438
AQ11: 484	AQ11: 411	AQ15: 402
AQ27: 409	AQ27: 396	AQ31: 415

3.4. Modeling mixtures of different compound to simulate Garancine (50% Alizarin: 50% Purpurin).

To go further, we extended the analysis by comparing the measured spectra of an alizarin and purpurin mixture with the computed one using the same procedure (AH) as before. This mixture simulates the proportions found in different madder species. As within the same plant species, the amount of anthraquinone can vary depending on the plant's age, soil, and pH, making it difficult to predict the absolute color of one specific plant. We account for the presence of different deprotonated forms, considering the pKa values and experimental pH of the mixture (Figure 4).

As detailed in S.I.8, Garancine is a pigment obtained by treating the roots of *Rubia tinctorum* L. with sulfuric acid, which removes sugar moieties and impurities, leaving concentrated anthraquinone compounds, primarily alizarin, and purpurin. The experimental solution containing a 50:50 ratio of alizarin and purpurin has a pH of 5.5, aligning with the first pKa of alizarin. At this pH, there exists an equal concentration of

alizarin in its protonated form (AH_2) and its first deprotonated monoanionic form (AH^-). Consequently, the computed spectra were derived by combining contributions from 25% alizarin in its neutral state (AH_2) and 25% alizarin in its deprotonated monoanionic state (AH^-). For the 50% purpurin component, only its monoanionic form (H_2P^-) was considered, given that the pH is nearly one unit above purpurin's first pKa of 4.6. The computed mixed spectrum (indicated by the red bold line) and the experimental spectrum (represented by dark, dotted lines) are depicted in Figure 4. The details of the individual computed spectra contribution are in S.I.9. This analysis confirms the absence of contribution from the purpurin di-anionic form HP^{2-} at pH 5.5, as its spectra is much more redshifted compared to the experimental one. The computed spectra demonstrated significant agreement with the experimental one.

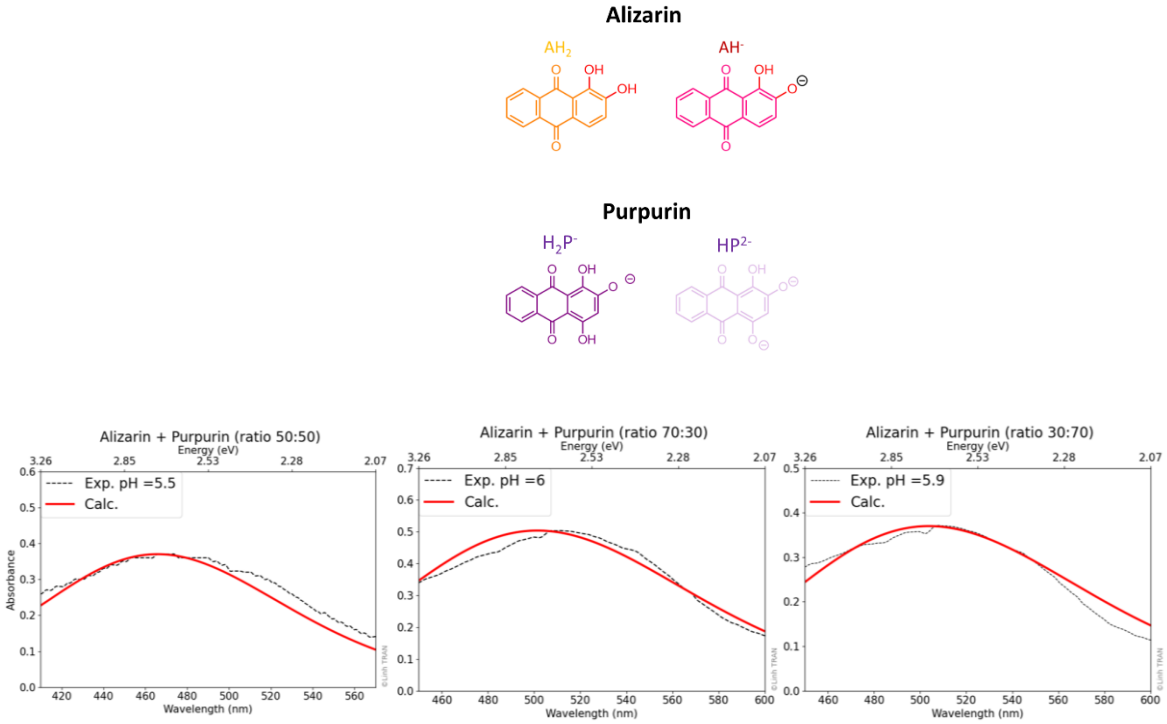


Figure 4. *Computed and experimental absorbance spectra of mixed ratios of alizarin and purpurin in different ratios (50:50, 30:70 or 70:30). The molecular structure of alizarin and purpurin forms are also represented.*

Similar calculations were conducted for different ratios of alizarin to purpurin, specifically 70:30 and 30:70, with the outcomes presented in Figure 4. The proportions of the molecules were adjusted in accordance with the pH. The computed spectra for these ratios also exhibited good agreement with the experimental ones. The Figure 5. presents the CIELAB color coordinates of the computed and experimental mixtures for different ratios and the color deviation ΔE_{2000}. Overall, the color difference ΔE_{2000} is less than 5. This deviation observed can be attributed to several factors. The increased deviation in ΔE_{2000} values, particularly for the 30:70 ratio, is mainly observed when the proportion of purpurin is higher. This is likely due to purpurin's lower purity (typically 90% when purchased, compared to 97% for alizarin). Additionally, the L*, a*, b* color representation does not fully reflect reality, as it depends on the choice of an illuminant (D65 in this case),

which may not perfectly match actual lighting conditions. The difference between ΔE_{2000} values of 4.6 and 3.1 is not necessarily indicative of a meaningful perceptual difference. Furthermore, experimental uncertainties contribute to variations in reproduced colors, adding to the observed deviations. Lastly, other transitions, notably in the UV range can also impact the shape of the absorption spectra but were not considered in our calculations as the color in Anthraquinone mainly comes from the unique electronic transition in the visible range. This validation serves to affirm the robustness of our protocol, which can now be utilized to explore other compositional variations with confidence.

CIELAB coordinate	Experimental				Calculated				Color difference
Mixture	L*	a*	b*	color	L*	a*	b*	color	ΔE_{2000}
alizarin + purpurin 50%-50%	85.5	11.7	20.5		88.4	12.0	25.1		3.1
alizarin + purpurin 70%-30%	75.2	26.8	4.3		75.5	28.0	4.2		0.6
alizarin + purpurin 30%-70%	81.7	20.6	7.2		80.6	22.5	0.8		4.6

©Linh TRAN

Figure 5*. Experimental and computed (normalized) CIELAB color coordinated of corresponding mixed solutions, along with the color difference ΔE_{2000}.*

4. CONCLUSION

This study shows that our designed computational procedure is able to compute the absorption spectroscopic signatures and colors of all dyes extracted from madder, including their glycoside forms, with excellent accuracy. The ability to identify dyes inside a pigment without the need to take samples is important in all cultural heritage research. This work highlights that computational chemistry can provide precise identification of AQs and offer deeper insights into the lake substrate. The accuracy of this computational protocol was checked by comparing the computed λ_{max} with the experimental values for several madder compounds, and an error deviation of less than 0.10 eV was measured. Furthermore, the nature and position of substituents significantly influence the color hues and shifts of AQ compounds, with hydroxyl groups at positions 1 and 4 enhancing more redshifted in acidic conditions through stabilizing intramolecular hydrogen bonds and extending the π-conjugation. Additionally, the hydrolysis of glycosidic forms to aglycones generally results in a slight redshift, with deprotonation intensifying the redshift. These findings highlight the complex interplay of molecular structure, protonation state, and substituent effects on the optical properties of these compounds. Moreover, we experimentally measured the spectra of a mixture of alizarin and purpurin, finding good agreement with the computed one which validates our computational protocol. We believe that the proposed resource will aid researchers in historical and cultural preservation by providing insights into historical dyeing practices and facilitating the preservation and restoration of cultural artifacts, textiles, and artworks. The generated database would also benefit professionals in various coloring industries, such as quality control in dyeing processes in sectors like textiles, cosmetics, and food. The color database

can also help designers produce new materials, catering to their artistic aesthetics and color preferences.

5. ASSOCIATED CONTENT

SUPPORTING INFORMATION

List of compounds detected in different madder species, comparison of computed VE of some madder compounds with experimental data. Difference densities for some madder molecules. Relative percentage of compounds in madder Tinctorum, Peregrina species and in Garancine. Individual spectra of 31 madder molecules in acidic and deprotonated forms. Optimised geometries.

AUTHOR INFORMATION

Corresponding Author

* Dr. BERRAUD-PACHE Romain

Laboratoire d'Archéologie Moléculaire et Structurale (LAMS)

Sorbonne Université, CNRS, UMR 8220, Paris, 75005, France

E-mail: romain.berraud-pache@sorbonne-universite.fr

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENT

This work was granted access to the HPC resources of IDRIS under the allocation 2023-AD0814570 made by GENCI.

SUPPLEMENTARY DATA

S.I. Table 1. *Various Rubia species for textile dyeing. Taxonomy verified through Kew Gardens Medicinal Plant Names Services database.[236] The Color Index name represents the mixture of colorant compounds extracted from the plant roots. (Taken from Richard S. Blackburn).[60]*

Plant common name	Plant species	Scientific synonyms	Color Index Name
Dyer's madder	Rubia *tinctorum* L.	R. acaliculata; R. iberica; R. sativa; R. sylvestris	C. I. Natural Red 8
Wild madder	Rubia *peregrina* L.		C. I. Natural Red 8
Indian madder	Rubia *cordifolia* L.	R. oncotricha; R. manjith; R. munjista; R. sylvatica	C. I. Natural Red 16
Naga madder	Rubia *sikkimensis* Kurz		None
Xiao hong can (no English name)	Rubia *yunnanensis* Diels	R. ustulata	None
Japanese madder	Rubia *akane* Nakai	R. argyi	None


Rubia Tinctorum


Rubia Peregrina


Rubia Cordifolia


Rubia Akane Nakai

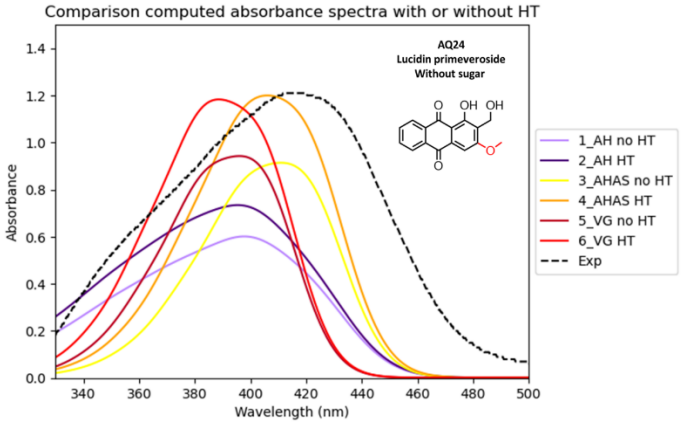
S.I. Table 2. *Compounds detected in the roots of six main Rubia species studied in literature. ++ indicates major compound in planta, + significant amount, tr. very low or trace amount. (Taken from Richard S. Blackburn's article).* [60]

		R.Tinctorum	R.cordifolia	R.peregrina	R.Sikkimensis	R. Yunnanensis	R. akane
1	Ruberythric acid (alizarin primeveroside)	++	tr.				
2	Alizarin	+	tr.	tr.			
3	Lucidin primeveroside	++	tr.	++	+	+	
4	Lucidin	tr.		+			
5	Nordamnacanthal	+/tr.		+			
6	Munjistin	+	++	+	++	+	++
7	Xanthopurpurin	+/tr.	+	+	+	+	tr.

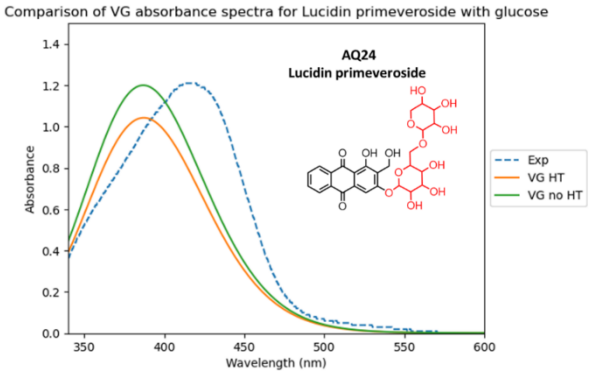
8	Galiosin (pseudopurpurin primeveroside)	+		++			tr.
9	Pseudopurpurin glucoside	+	+				+
10	Pseudopurpurin	+	++	++	++	+	++
11	Purpurin	+	+	++	+		tr.
12	Rubiadin primeveroside	+		+/tr.			
13	Rubiadin	+	tr.	+		+	
14	1-Hydroxy-AQ	+/tr.	+				
15	1-Hydroxy-2-methyl-AQ	+	+	+		+	+
16	1-Hydroxy-2-hydroxymethyl-AQ		+				
17	1-Hydroxy-3-carboxy-AQ		+				
18	1,4-Dihydroxy-2-hydroxymethyl-AQ		+			+	
19	1,4-Dihydroxy-2-methyl-AQ		+				
20	Anthragallol	+					
21	2-Hydroxy-AQ	+					
22	2-(Hydroxymethyl)-AQ	+				+	
23	2-(Hydroxymethyl)-3-hydroxy AQ					+	
24	6-Hydroxyrubiadin-3-O-neohesperidoside diac		+				+
25	6-Hydroxyrubiadin-3-O-neohesperidoside monoac		+			++	++
26	6-Hydroxyrubiadin-3-O-neohesperidoside		+			++	++
27	6-Hydroxyrubiadin-3-O-glucoside monoac		+			+	
28	6-Hydroxyrubiadin-3-O-glucoside		+				
29	6-Hydroxyrubiadin		+			+	
30	Rubianine	+					
31	Munjistin-3-O-glucoside		+				++

S.I. Table 3. *Benchmark: computed VE with different functionals for alizarin and their absolute error deviation.*

	B3LYP		B2PLYP		PBE0		Exp.
Alizarin	Calc. eV(nm)	Absolute Error deviation (eV) calc.-exp.	Calc. eV(nm)	Absolute Error deviation (eV) calc.-exp.	Calc. eV(nm)	Absolute Error deviation (eV) calc.-exp.	eV(nm)
Protonated	2.82 (439.7)	0.06	2.94 (422.2)	0.06	2.93 (423.2)	0.05	2,88 (430)
Mono-deprotonated	2.08 (595.7)	0.30	2.11 (587.9)	0.27	2.17 (572.7)	0.21	2,38 (520)



Method	**Computation time with HT effect**	**Computation time without HT effect**
AH	13min (opt)+14h (ES hessian)+ 3min(VE) + 6h (AH) Total≈20h	16min (opt+VE)+3min(AH) Total≈0.3h
AHAS	16min (opt+VE)+7h(AHAS) Total≈7.5h	16min (opt+VE)+3min(AHAS) Total≈0.3h
VG	16min (opt+VE)+9.5h(VG) Total≈10h	16min (opt+VE)+6min(VG) Total≈0.4h



	Computation time with HT effect	**Computation time without HT effect**
VG	8h (opt)+6 days 11h(spectra) Total≈163h	8h (opt)+1h(spectra) Total≈9h

S.I. 4. *Effect of Herzberg–Teller Correction and Computation Time*
(Top) Computed absorbance spectra of lucidin primeveroside (without sugar) using AH, AHAS, and VG methods, both with and without Herzberg–Teller (HT) effects. The experimental spectrum of lucidin primeveroside is shown as a black dashed line. A table below presents the computation time for each method. (Bottom) Computed absorbance spectra of lucidin primeveroside (with sugar) using the VG method, with and without HT effects. The experimental spectrum is shown as a blue dashed line. The table below summarizes the

computation time for each method. The cluster used is based on AMD EPYC 7443 (2.85 GHz) with 16 processors.

$$\Delta E_{2000} = \sqrt{(\frac{\Delta L'}{K_L S_L})^2 + (\frac{\Delta C'}{K_C S_C})^2 + (\frac{\Delta H'}{K_H S_H})^2}$$

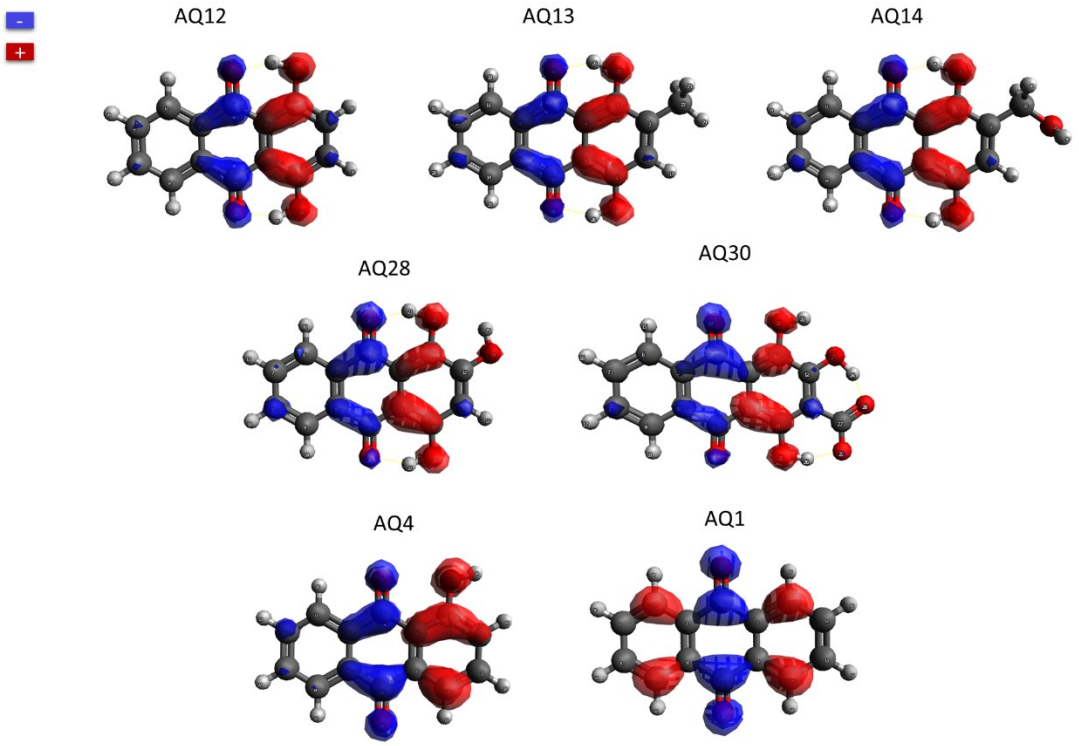
where:

$$\overline{L'} = (L_1 + L_2)/2,$$
$$\Delta L' = L_2 - L_1,$$
$$C_1 = \sqrt{a_1^2 + b_1^2},$$
$$C_2 = \sqrt{a_2^2 + b_2^2},$$
$$\overline{C} = (C_1 + C_2)/2,$$
$$G = (1 - \sqrt{\frac{\overline{C^7}}{\overline{C^7} + 25^7}})/2,$$
$$a_1' = a_1(1 + G),$$
$$a_2' = a_2(1 + G),$$
$$C_1' = \sqrt{a_1'^2 + b_1^2},$$
$$C_2' = \sqrt{a_2'^2 + b_2^2},$$
$$\overline{C'} = (C_1' + C_2')/2,$$
$$\Delta C' = C_2' - C_1',$$
$$\Delta H' = 2\sqrt{C_1' C_2'}\sin(\Delta h'/2),$$
$$S_L = 1 + \frac{0.015(\overline{L'} - 50)^2}{\sqrt{20 + (\overline{L'} - 50)^2}},$$
$$S_C = 1 + 0.045\overline{C'},$$
$$S_H = 1 + 0.015\overline{C'}T,$$
$$\Delta\theta = 30\exp\{-(\frac{\overline{H'} - 275^o}{25})\},$$
$$R_C = \sqrt{\frac{\overline{C^7}}{\overline{C^7} + 25^7}}),$$
$$R_T = -2R_C\sin(2\Delta\theta),$$
$$K_L = 1 - default,$$
$$K_C = 1 - default,$$
$$K_H = 1 - default.$$

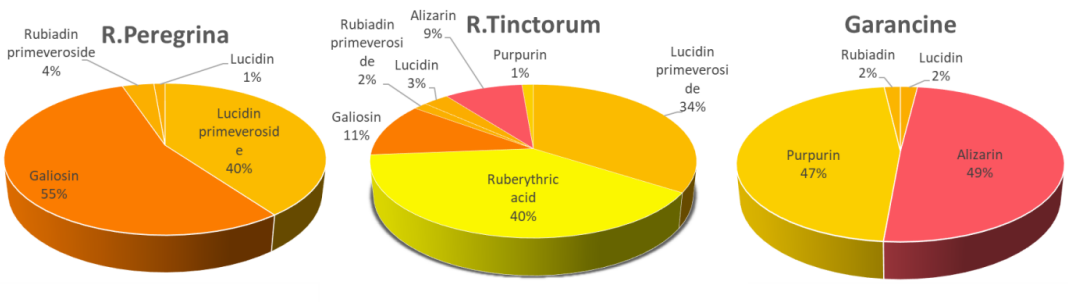
S.I. 5. *ΔE_{2000} formula for color difference.*[138,140]

N°	Molecule	Calc. VE (eV)	Exp. λmax (eV)	Calc.-Exp (eV)
24	Lucidin primeveroside	3.03	3.05	-0.02
11	Ruberythric acid	2.91	2.97	-0.06
25	Galiosin	2.82	2.86	-0.04
10	Rubiadin primeveroside	3.08	3.01	0.07
29	Anthragallol	3.10	3.00	0.10
18	Lucidin	3.02	3.00	0.02
15	Alizarin	2.93	2.90	0.03
28	Purpurin	2.59	2.58	0.01
16	Xanthopurpurin	3.05	2.98	0.07
17	Rubiadin	3.05	3.02	0.03

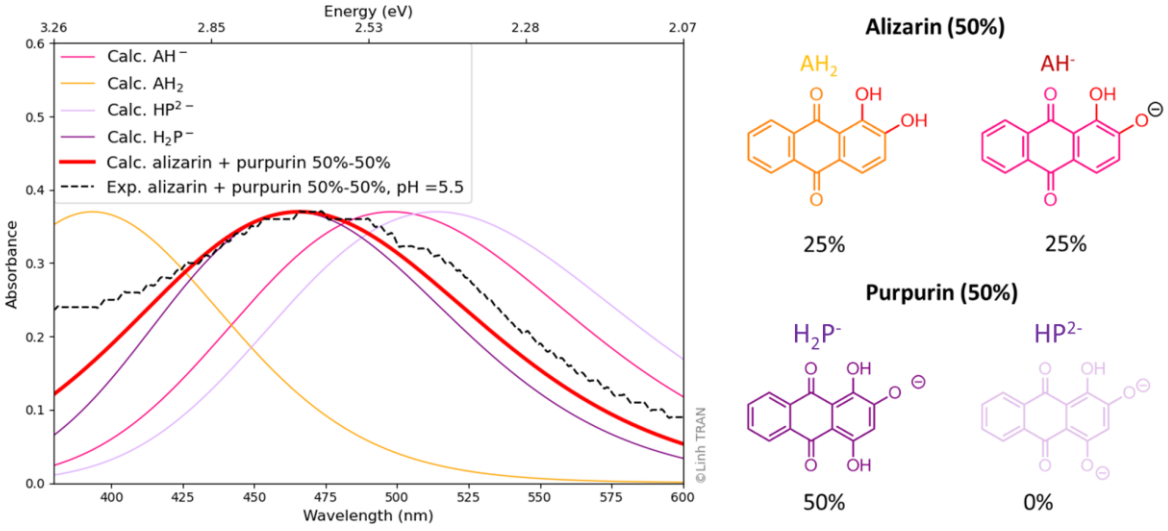
S.I. 6. *Experimental and computed maximum absorption energies of some madder molecules and their respective error deviation (in eV) which are less than 0.10 eV.*



S.I. 7. *The difference in electronic densities for the transition from the ground state (S_0) to the first excited state (S_1). This difference is obtained by subtracting the electronic density of the ground state (S_0) from that of the excited state (S_1), providing insight into the changes in electron distribution during the excitation. The calculations were performed using ORCA and the resulting electronic density difference is represented as a contour map to highlight the areas of increased and decreased electron density between the two states with isovalue =0.002 a.u. The blue (-) represents a decrease in electronic density and the red (+) an increase.*



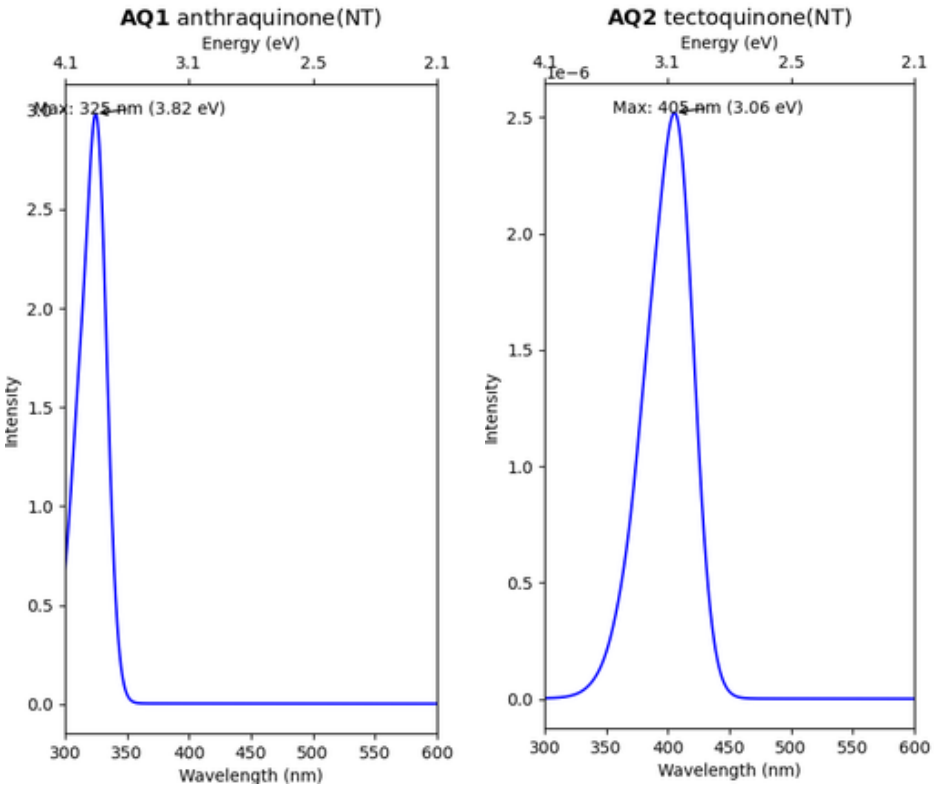
S.I. 8. *Pie chart showing the color and the percentage of molecules in different madder species (R. Tinctorum L. and R. Peregrina L.) and in Garancine (red pigment derived from the madder R. Tinctorum L.).* [59]

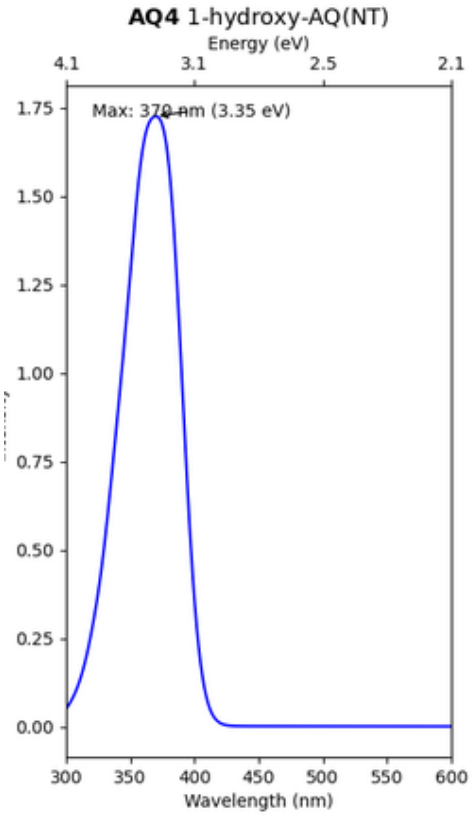


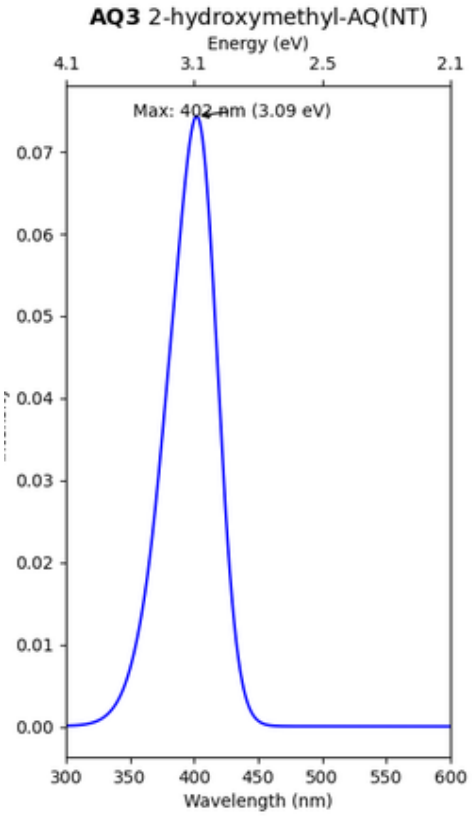
S.I. 9. *Computed (red) and experimental (dotted lines) absorbance spectra of mixed ratios of alizarin and purpurin (50%:50%). The individual spectra of each molecule are also represented, along with the molecule structure and individual percentage.*

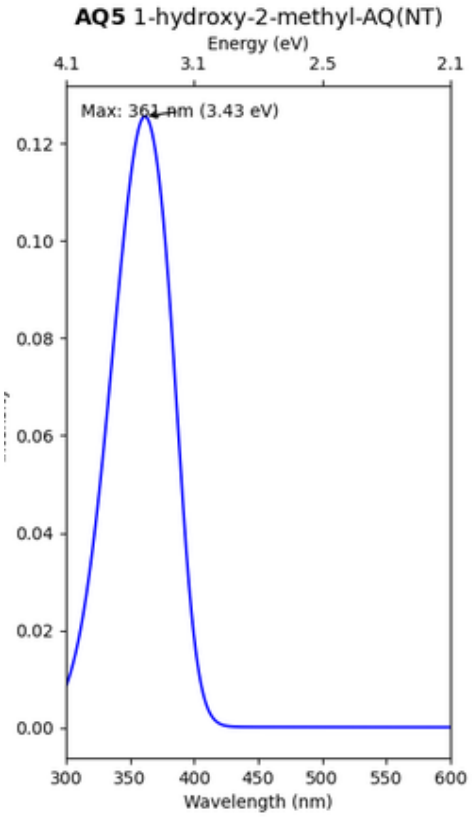
S.I. 10. *Computed UV-VIS spectra and geometries of studied molecules (.xyz files) for neutral (NT) and deprotonated (DP) forms.*

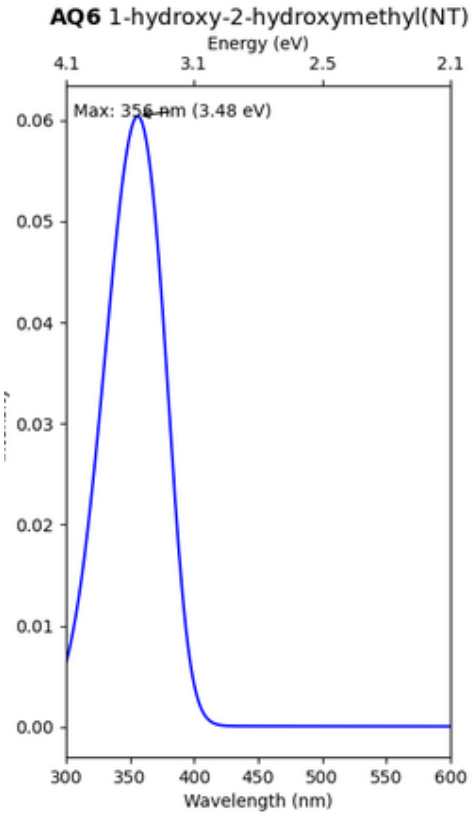
Neutral (NT) form UV-VIS computed spectra of 31 madder molecules

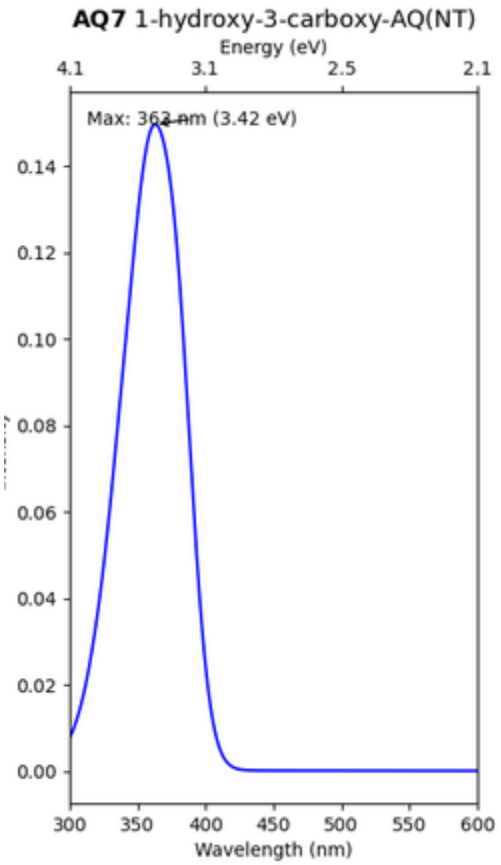


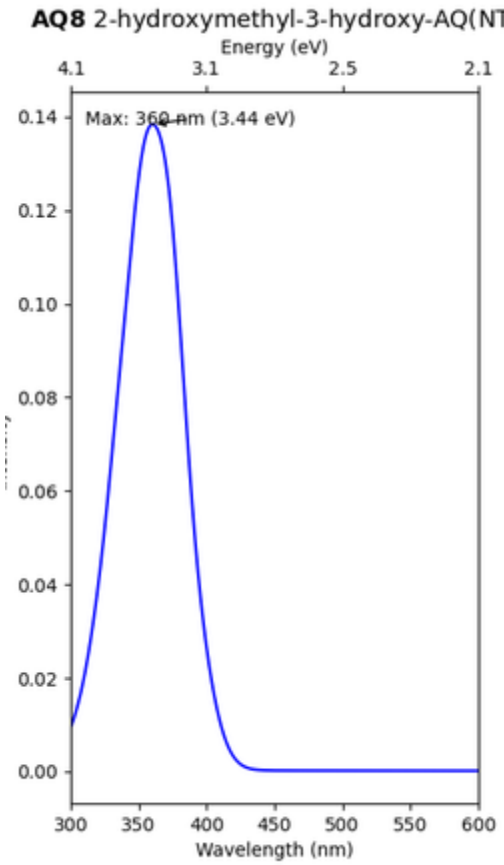

AQ4 1-hydroxy-AQ(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 370 nm (3.35 eV)
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
300
350
400
450
500
550
600
Wavelength (nm)

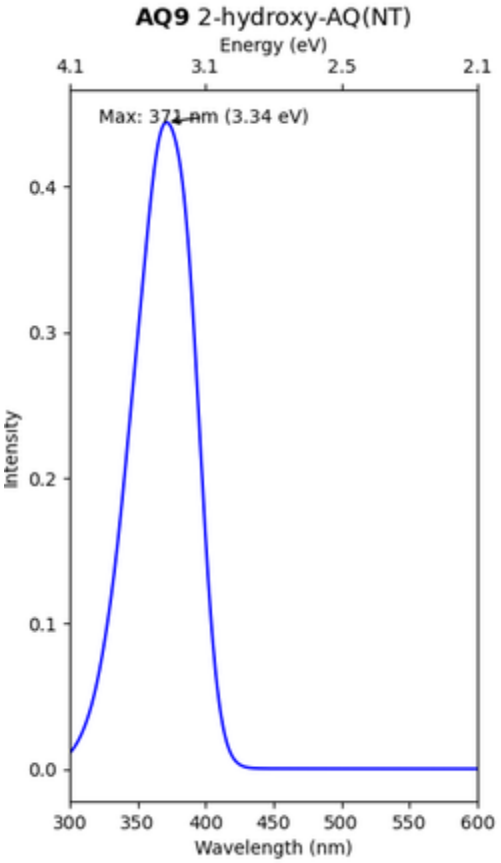

AQ3 2-hydroxymethyl-AQ(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 402 nm (3.09 eV)
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
300
350
400
450
500
550
600
Wavelength (nm)

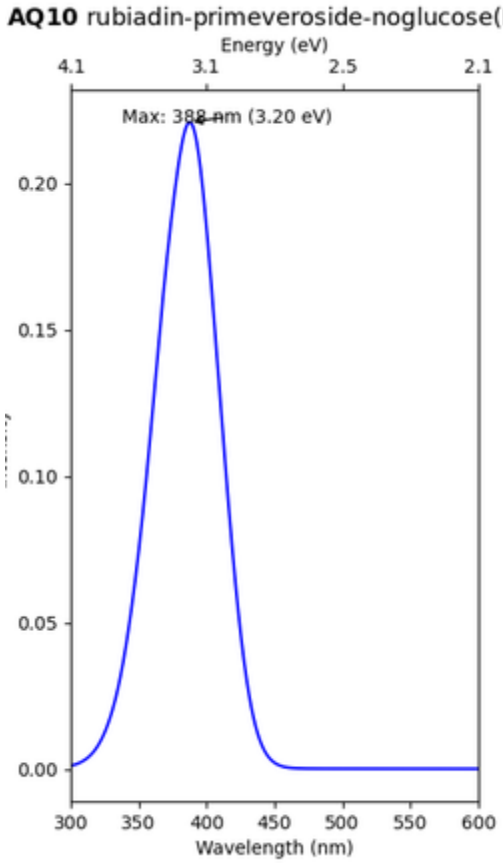

AQ5 1-hydroxy-2-methyl-AQ(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 361 nm (3.43 eV)
0.12
0.10
0.08
0.06
0.04
0.02
0.00
300
350
400
450
500
550
600
Wavelength (nm)

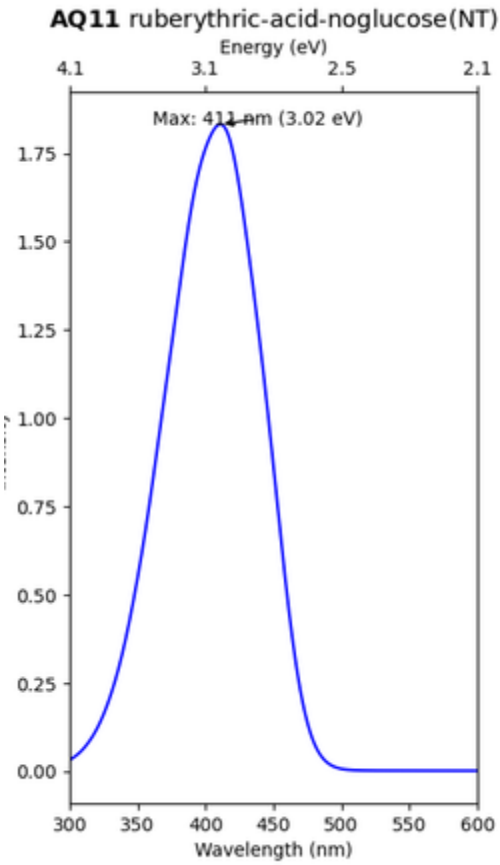

AQ6 1-hydroxy-2-hydroxymethyl(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 356 nm (3.48 eV)
0.06
0.05
0.04
0.03
0.02
0.01
0.00
300
350
400
450
500
550
600
Wavelength (nm)

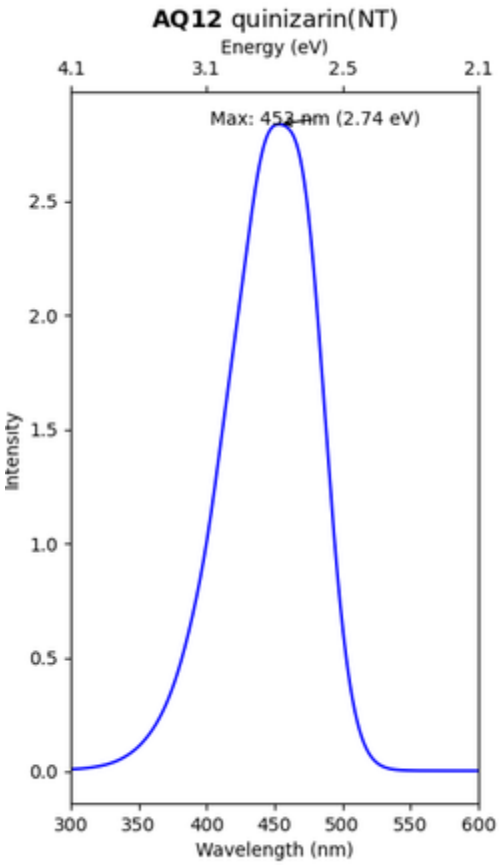

AQ7 1-hydroxy-3-carboxy-AQ(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 363 nm (3.42 eV)
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
300
350
400
450
500
550
600
Wavelength (nm)

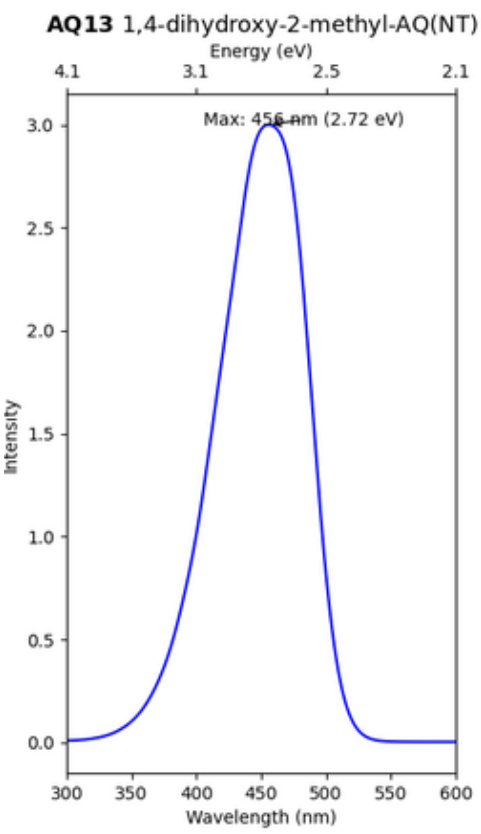

AQ8 2-hydroxymethyl-3-hydroxy-AQ(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 360 nm (3.44 eV)
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
300
350
400
450
500
550
600
Wavelength (nm)

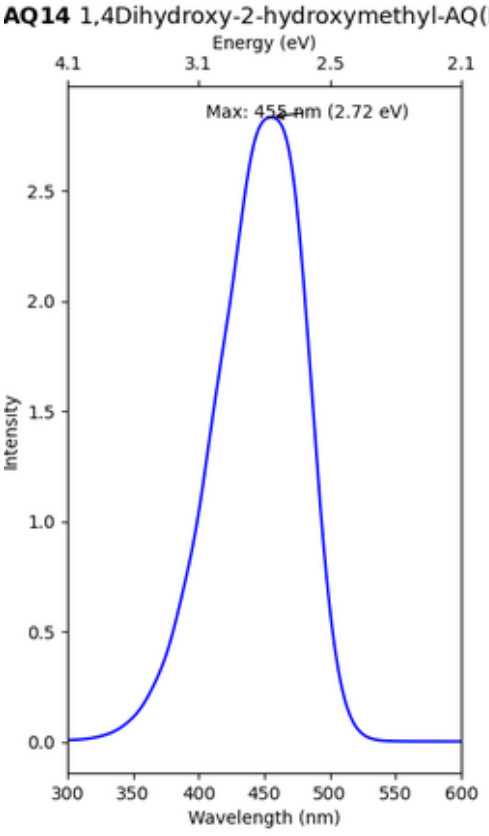

AQ9 2-hydroxy-AQ(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 371 nm (3.34 eV)
Intensity
0.4
0.3
0.2
0.1
0.0
300
350
400
450
500
550
600
Wavelength (nm)

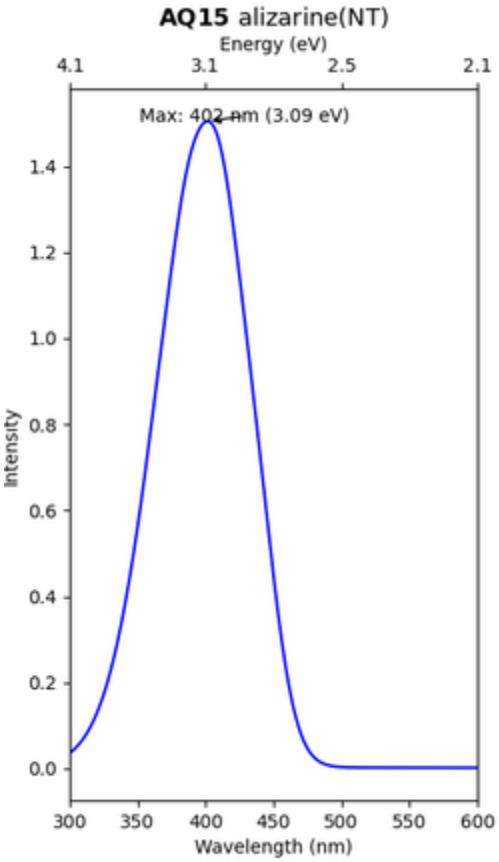

AQ10 rubiadin-primeveroside-noglucose(NT
Energy (eV)
4.1
3.1
2.5
2.1
Max: 388 nm (3.20 eV)
0.20
0.15
0.10
0.05
0.00
300
350
400
450
500
550
600
Wavelength (nm)

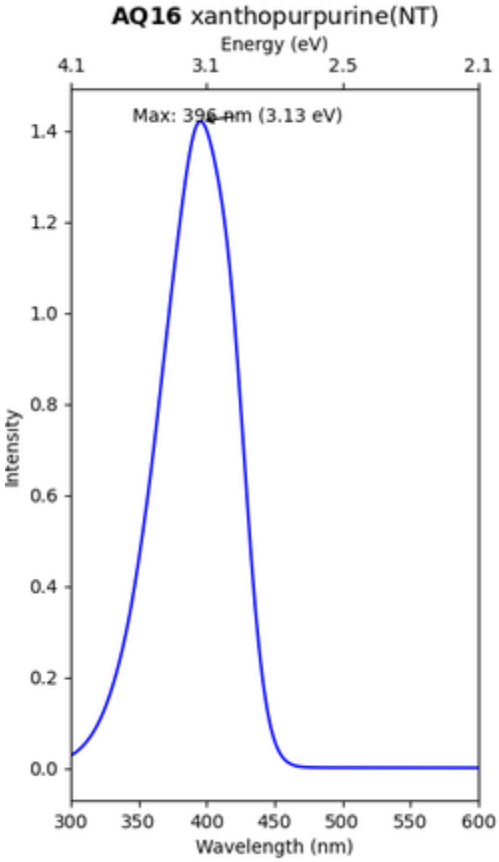

AQ11 ruberythric-acid-noglucose(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 411 nm (3.02 eV)
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
300
350
400
450
500
550
600
Wavelength (nm)

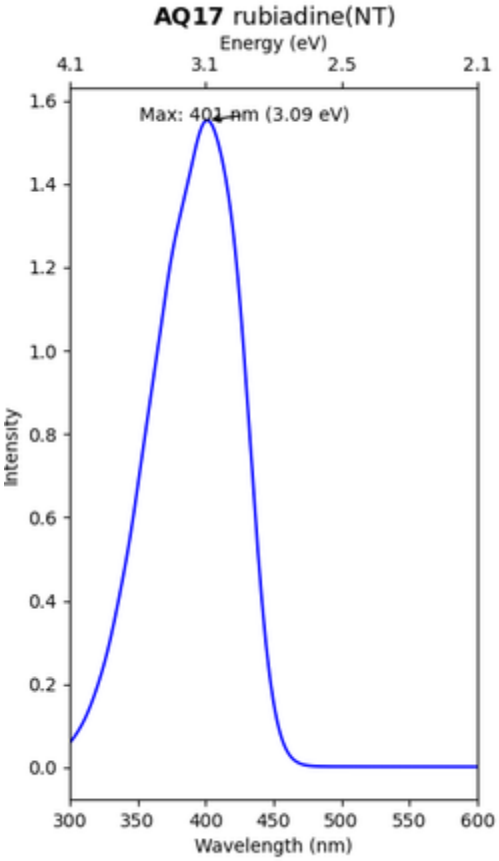

AQ12 quinizarin(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 453 nm (2.74 eV)
Intensity
2.5
2.0
1.5
1.0
0.5
0.0
300
350
400
450
500
550
600
Wavelength (nm)

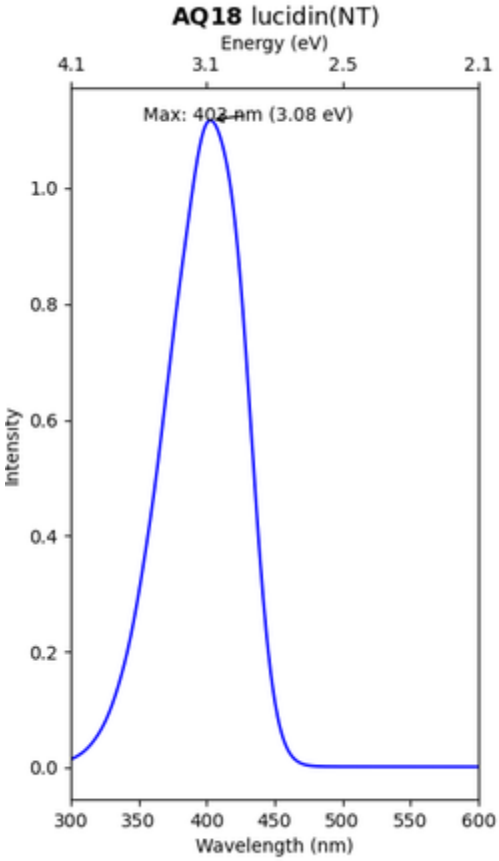

AQ13 1,4-dihydroxy-2-methyl-AQ(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 456 nm (2.72 eV)
Intensity
3.0
2.5
2.0
1.5
1.0
0.5
0.0
300
350
400
450
500
550
600
Wavelength (nm)

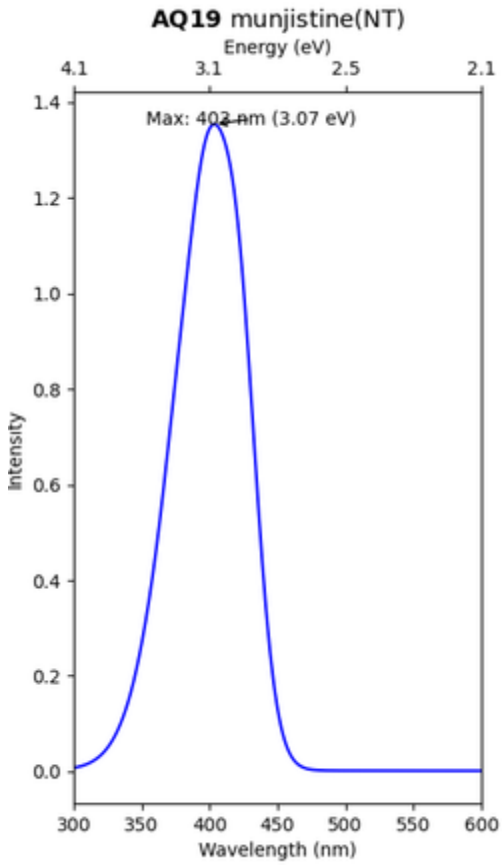

AQ14 1,4Dihydroxy-2-hydroxymethyl-AQ(NT
Energy (eV)
4.1
3.1
2.5
2.1
Max: 455 nm (2.72 eV)
Intensity
2.5
2.0
1.5
1.0
0.5
0.0
300
350
400
450
500
550
600
Wavelength (nm)

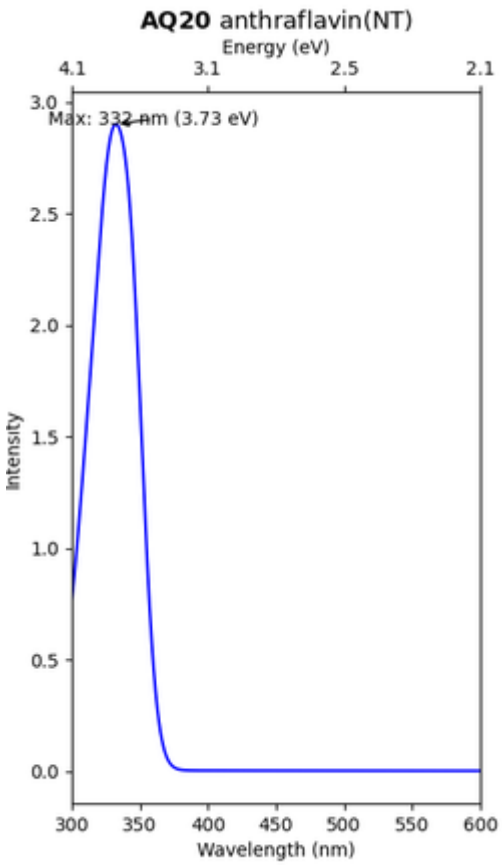

AQ15 alizarine(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 402 nm (3.09 eV)
Intensity
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
300
350
400
450
500
550
600
Wavelength (nm)


AQ16 xanthopurpurine(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 396 nm (3.13 eV)
Intensity
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
300
350
400
450
500
550
600
Wavelength (nm)

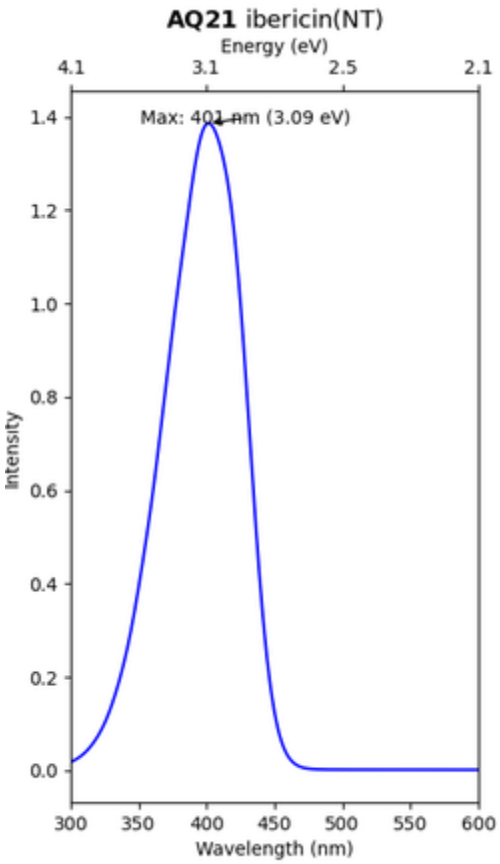

AQ17 rubiadine(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 401 nm (3.09 eV)
Intensity
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
300
350
400
450
500
550
600
Wavelength (nm)

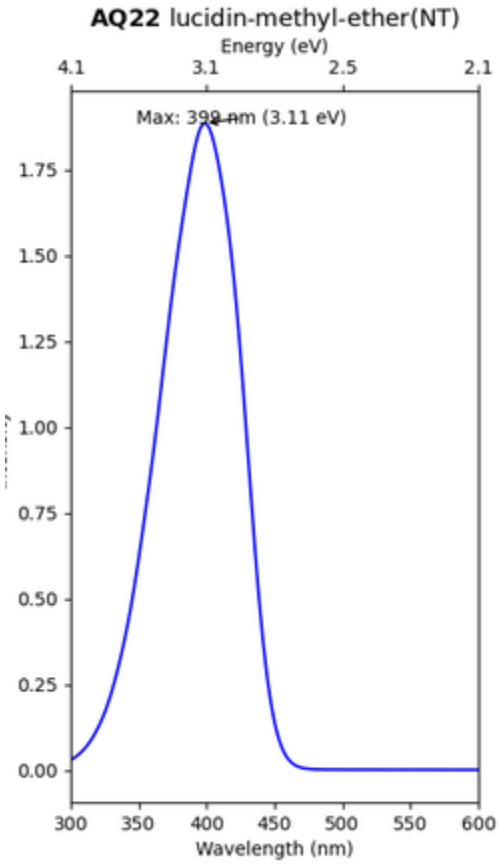

AQ18 lucidin(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 402 nm (3.08 eV)
Intensity
0.0
0.2
0.4
0.6
0.8
1.0
300
350
400
450
500
550
600
Wavelength (nm)

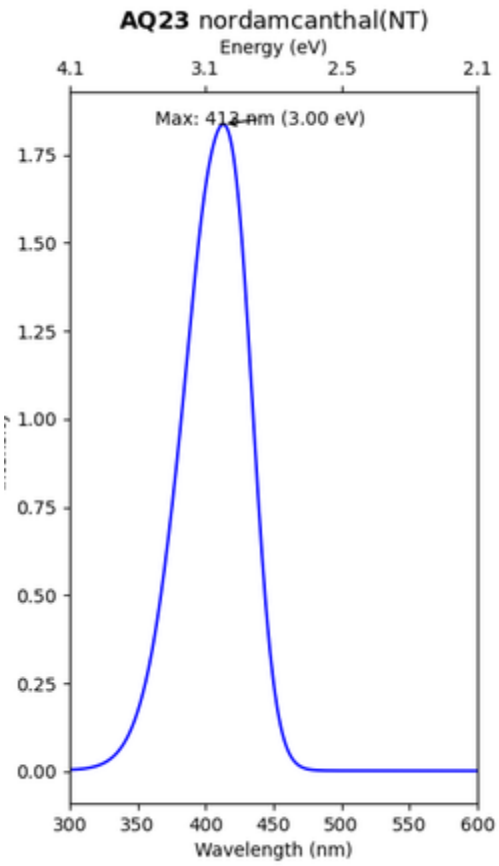

AQ19 munjistine(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 403 nm (3.07 eV)
Intensity
Wavelength (nm)

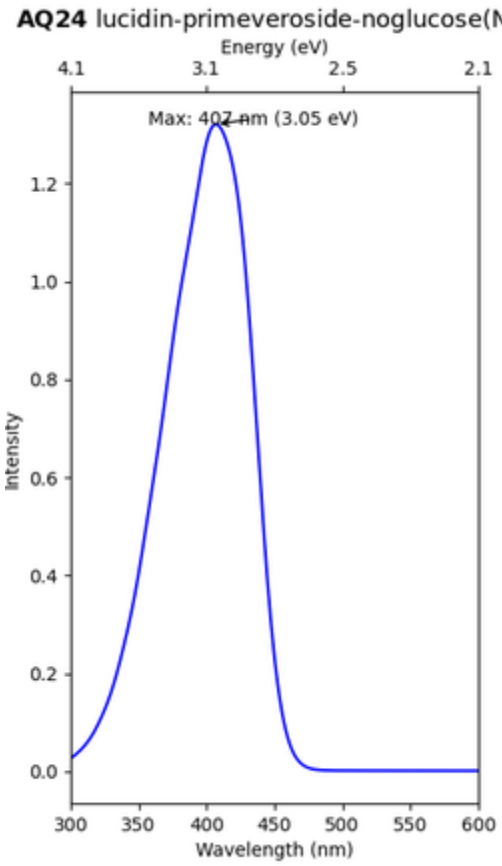

AQ20 anthraflavin(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 332 nm (3.73 eV)
Intensity
Wavelength (nm)

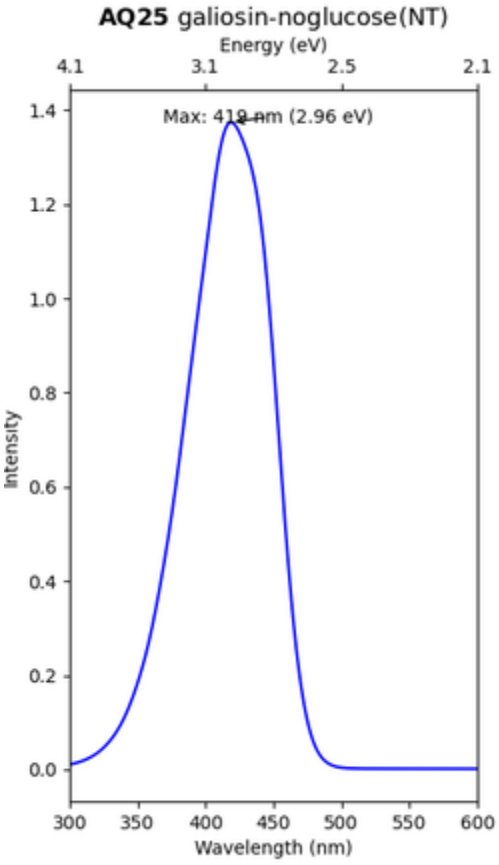
z

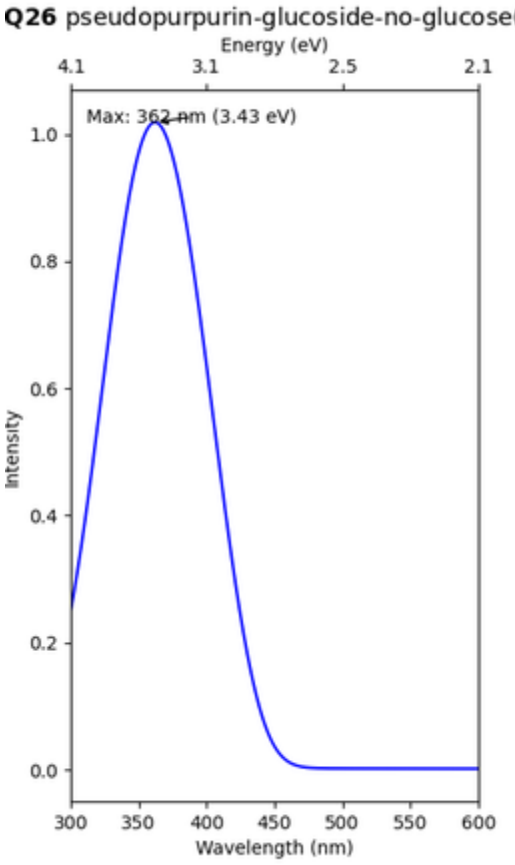

AQ21 ibericin(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 401 nm (3.09 eV)
Intensity
Wavelength (nm)

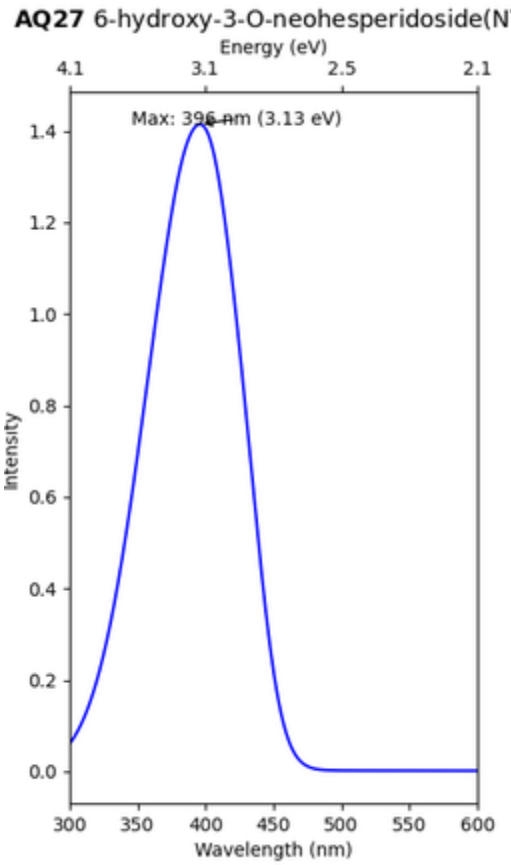

AQ22 lucidin-methyl-ether(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 398 nm (3.11 eV)
Wavelength (nm)

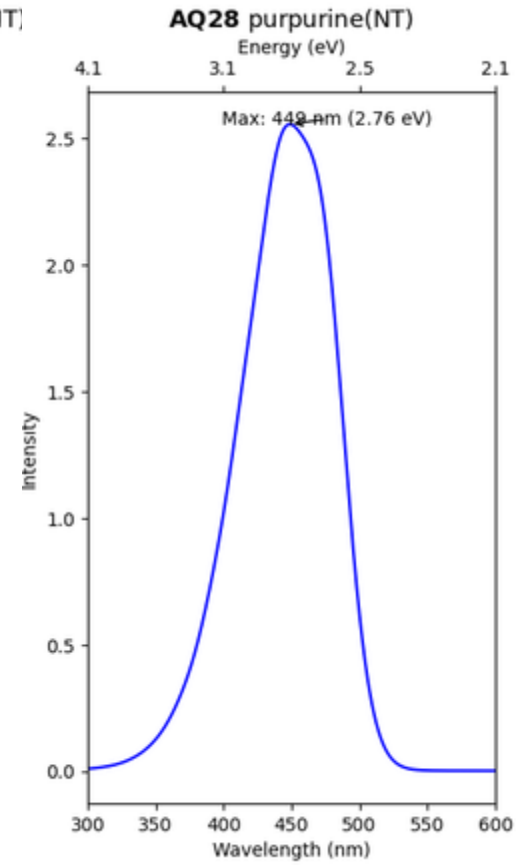

AQ23 nordamcanthal(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 413 nm (3.00 eV)
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
300
350
400
450
500
550
600
Wavelength (nm)

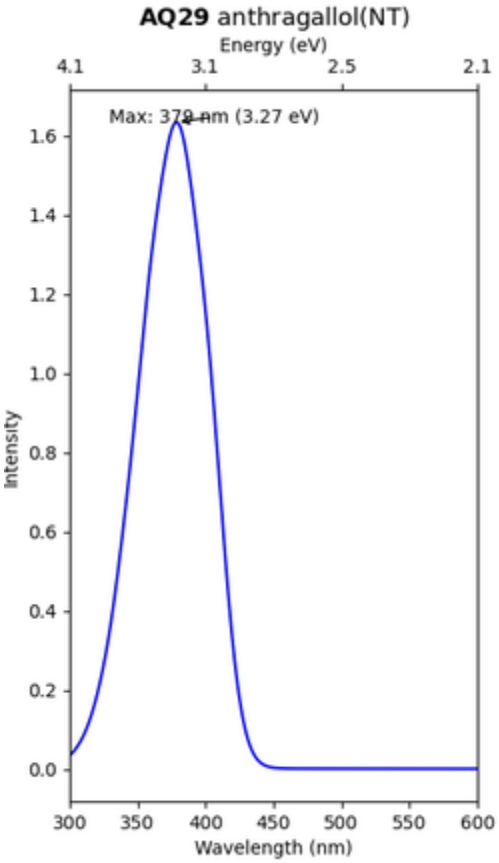

AQ24 lucidin-primeveroside-noglucose(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 407 nm (3.05 eV)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Intensity
300
350
400
450
500
550
600
Wavelength (nm)

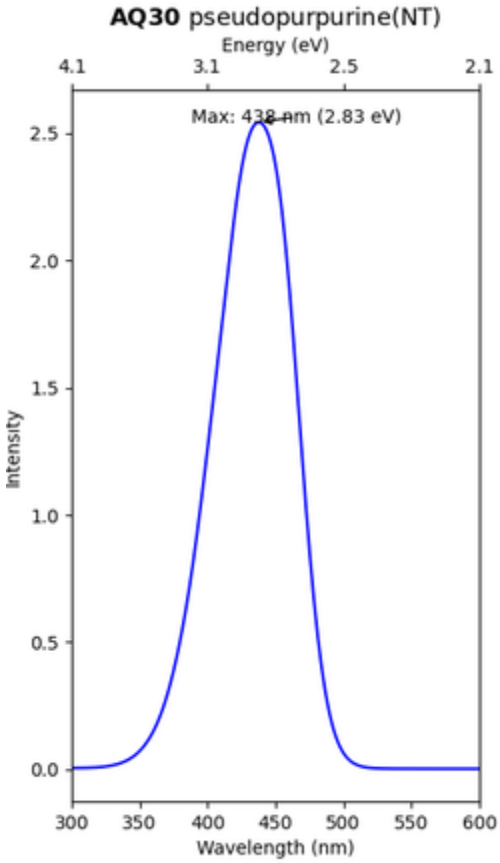

AQ25 galiosin-noglucose(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 418 nm (2.96 eV)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Intensity
300
350
400
450
500
550
600
Wavelength (nm)

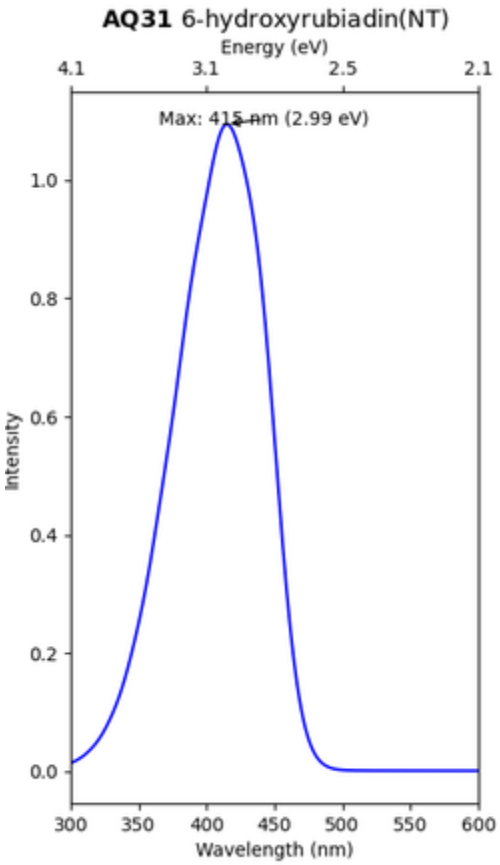

Q26 pseudopurpurin-glucoside-no-glucose(N
Energy (eV)
4.1
3.1
2.5
2.1
Max: 362 nm (3.43 eV)
1.0
0.8
0.6
0.4
0.2
0.0
Intensity
300
350
400
450
500
550
600
Wavelength (nm)


AQ27 6-hydroxy-3-O-neohesperidoside(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 396 nm (3.13 eV)
Intensity
Wavelength (nm)

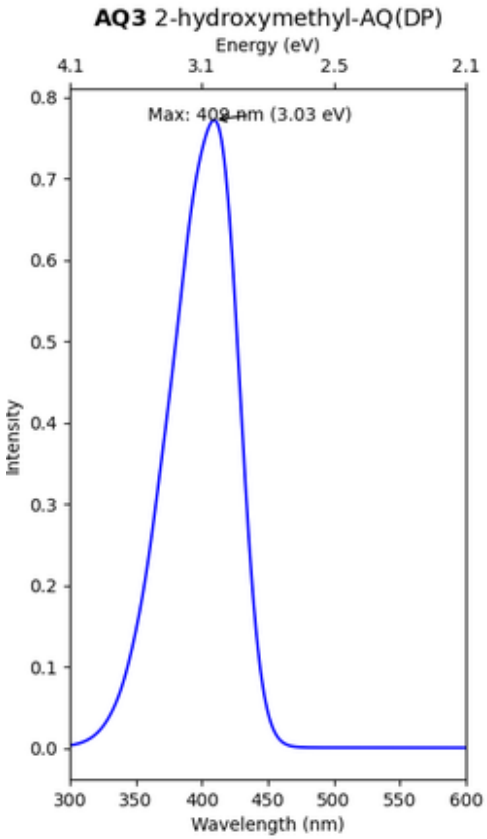

AQ28 purpurine(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 449 nm (2.76 eV)
Intensity
Wavelength (nm)

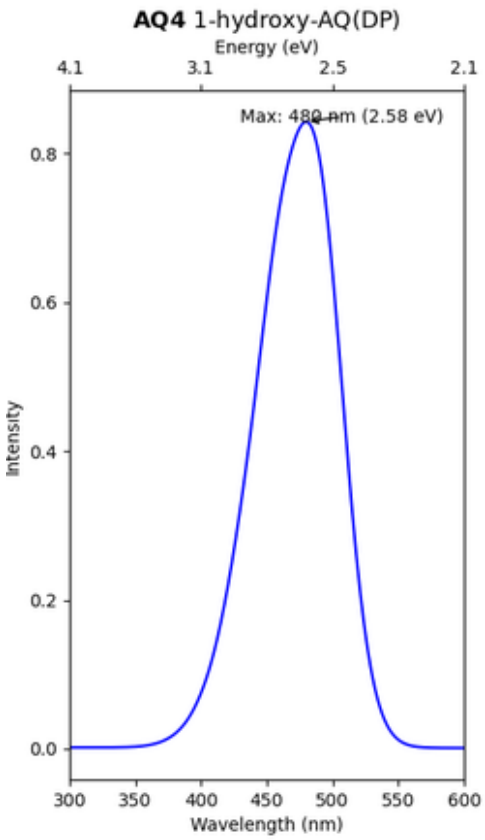

AQ29 anthragallol(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 379 nm (3.27 eV)
Intensity
Wavelength (nm)

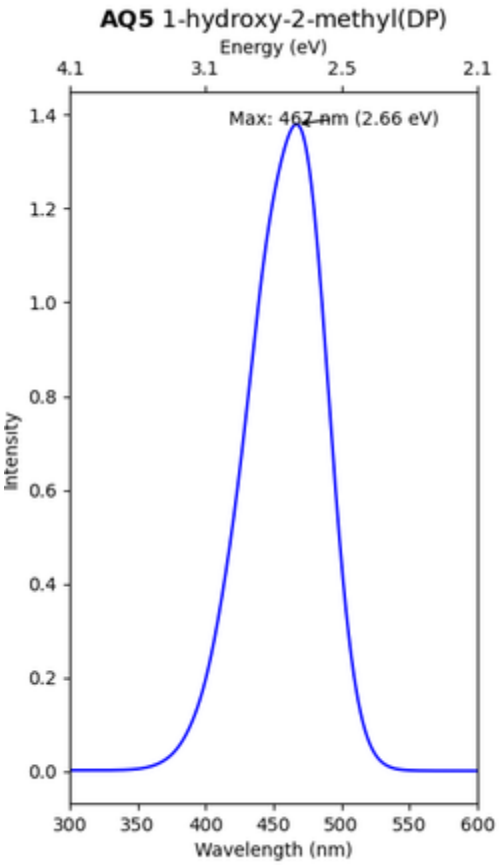

AQ30 pseudopurpurine(NT)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 438 nm (2.83 eV)
Intensity
Wavelength (nm)

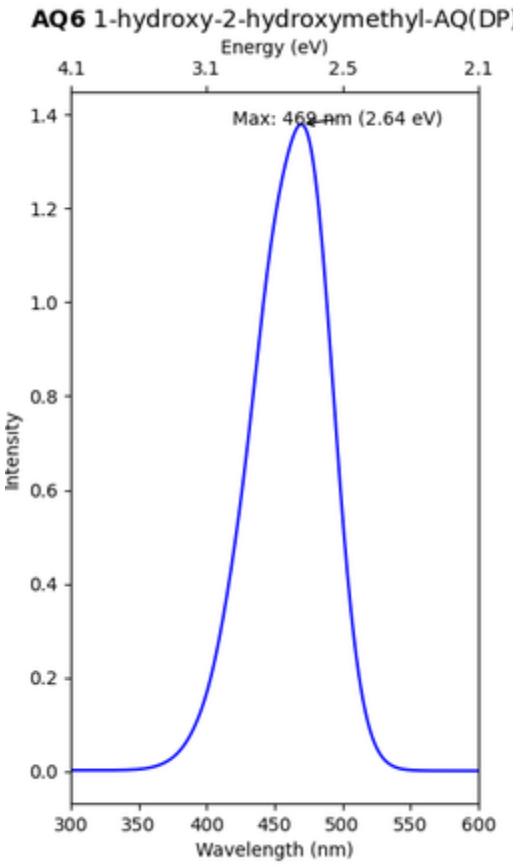


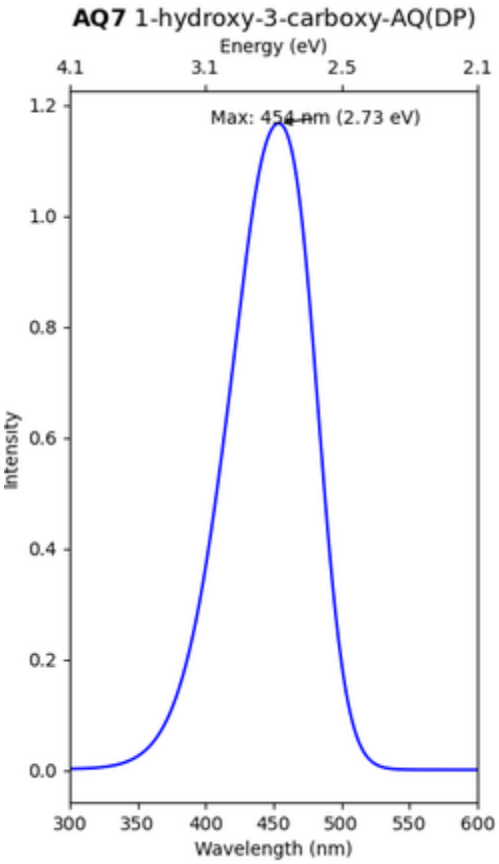
Deprotonated (DP) forms UV-VIS computed spectra of madder molecules

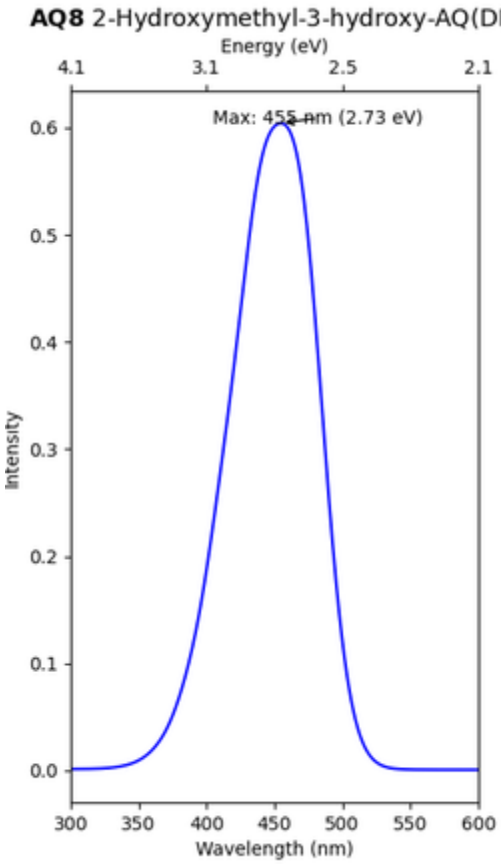


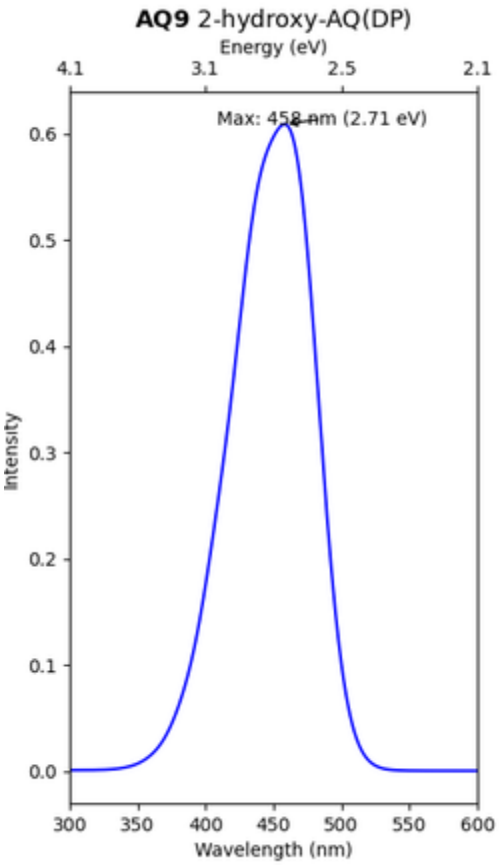


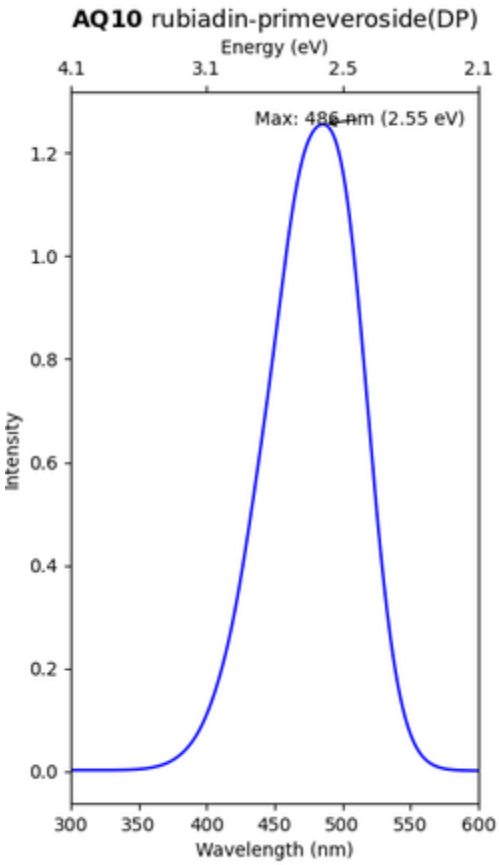

AQ5 1-hydroxy-2-methyl(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 467 nm (2.66 eV)
Intensity
Wavelength (nm)
300
350
400
450
500
550
600
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4

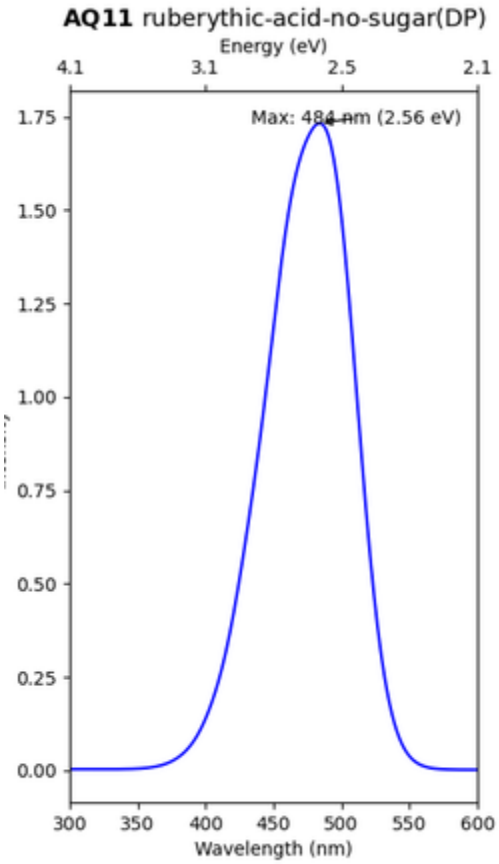

AQ6 1-hydroxy-2-hydroxymethyl-AQ(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 469 nm (2.64 eV)
Intensity
Wavelength (nm)
300
350
400
450
500
550
600
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4

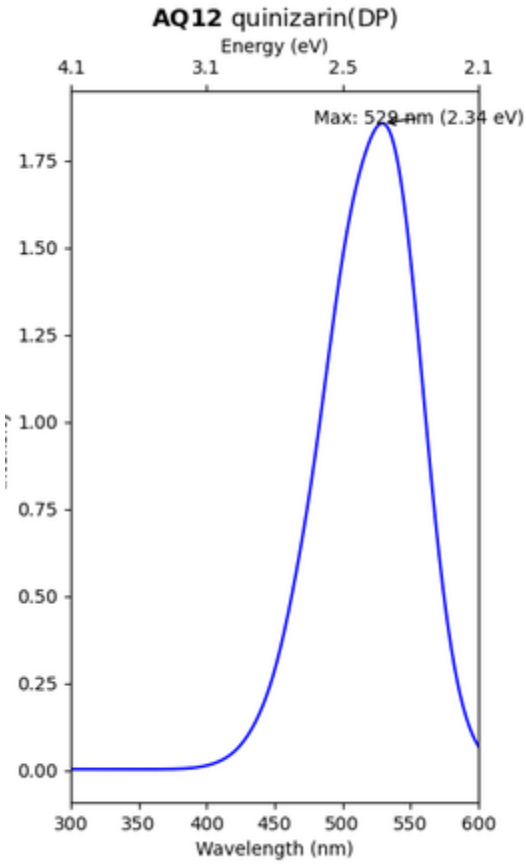

AQ7 1-hydroxy-3-carboxy-AQ(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 454 nm (2.73 eV)
Intensity
Wavelength (nm)
300
350
400
450
500
550
600
0.0
0.2
0.4
0.6
0.8
1.0
1.2

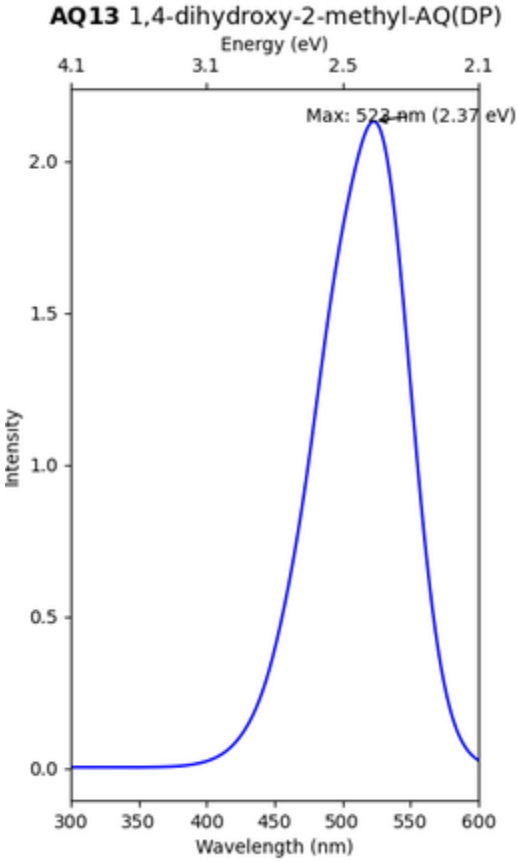

AQ8 2-Hydroxymethyl-3-hydroxy-AQ(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 455 nm (2.73 eV)
Intensity
Wavelength (nm)
300
350
400
450
500
550
600
0.0
0.1
0.2
0.3
0.4
0.5
0.6

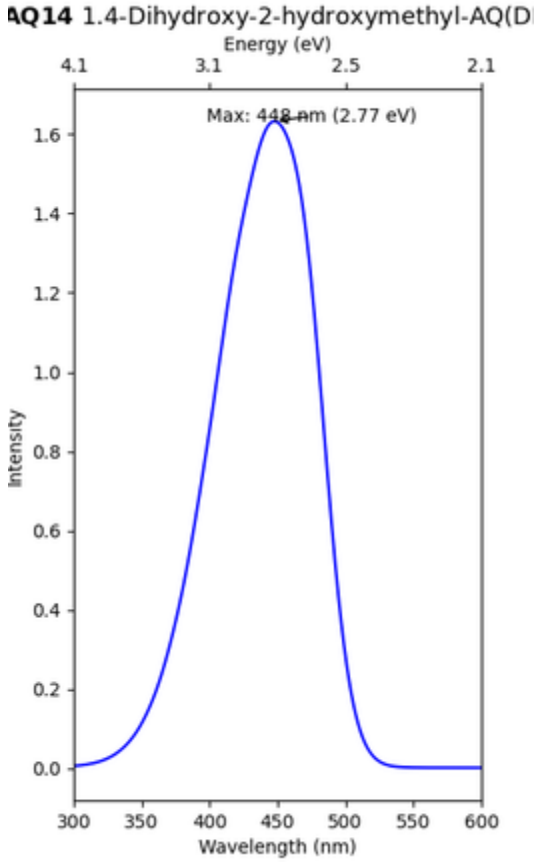

AQ9 2-hydroxy-AQ(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 458 nm (2.71 eV)
Intensity
0.0
0.1
0.2
0.3
0.4
0.5
0.6
300
350
400
450
500
550
600
Wavelength (nm)

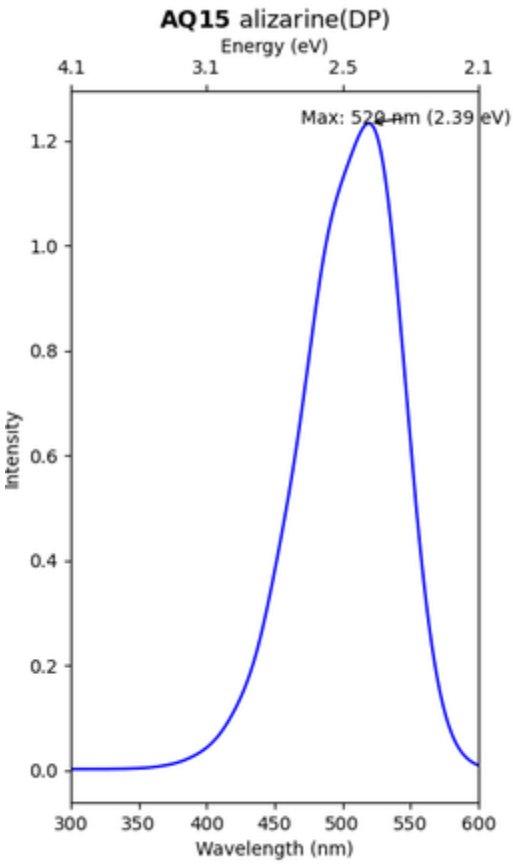

AQ10 rubiadin-primeveroside(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 486 nm (2.55 eV)
Intensity
0.0
0.2
0.4
0.6
0.8
1.0
1.2
300
350
400
450
500
550
600
Wavelength (nm)

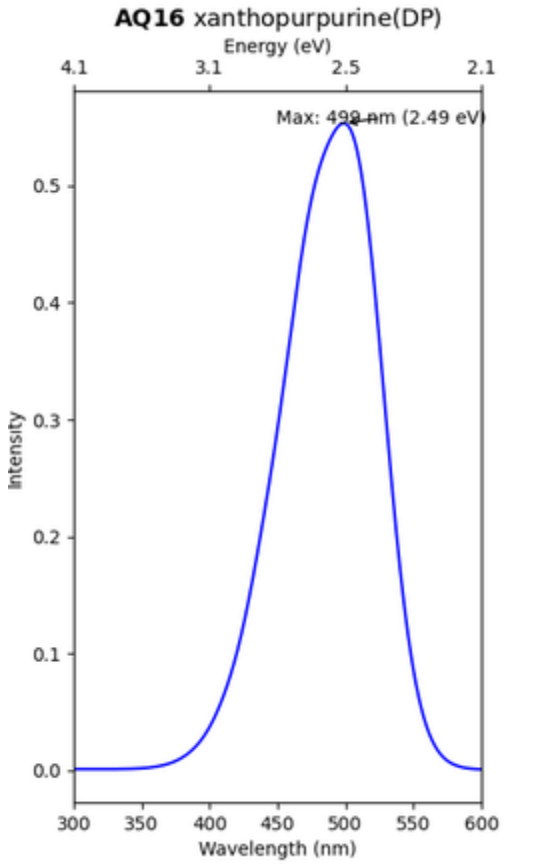

AQ11 ruberythic-acid-no-sugar(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 484 nm (2.56 eV)
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
300
350
400
450
500
550
600
Wavelength (nm)

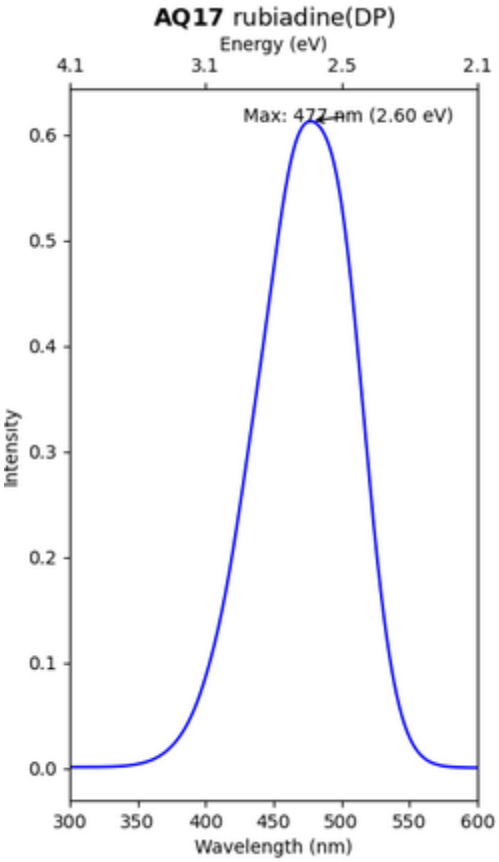

AQ12 quinizarin(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 528 nm (2.34 eV)
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
300
350
400
450
500
550
600
Wavelength (nm)

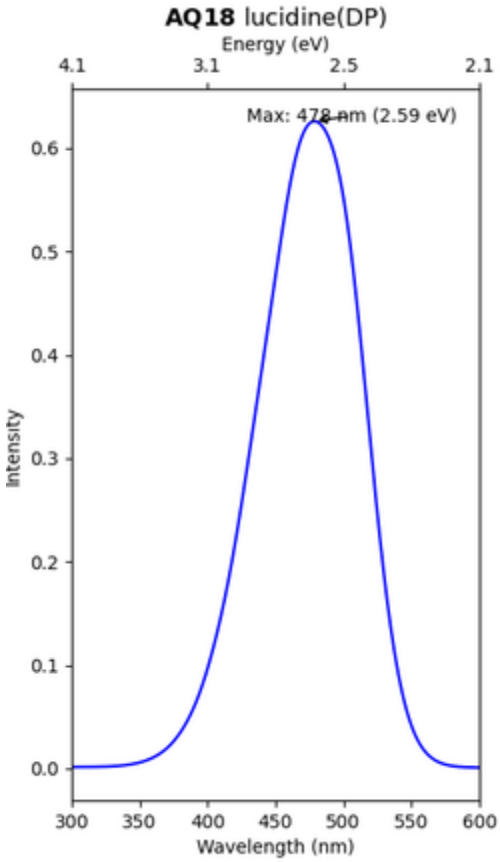

AQ13 1,4-dihydroxy-2-methyl-AQ(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 523 nm (2.37 eV)
Intensity
Wavelength (nm)
0.0
0.5
1.0
1.5
2.0
300
350
400
450
500
550
600

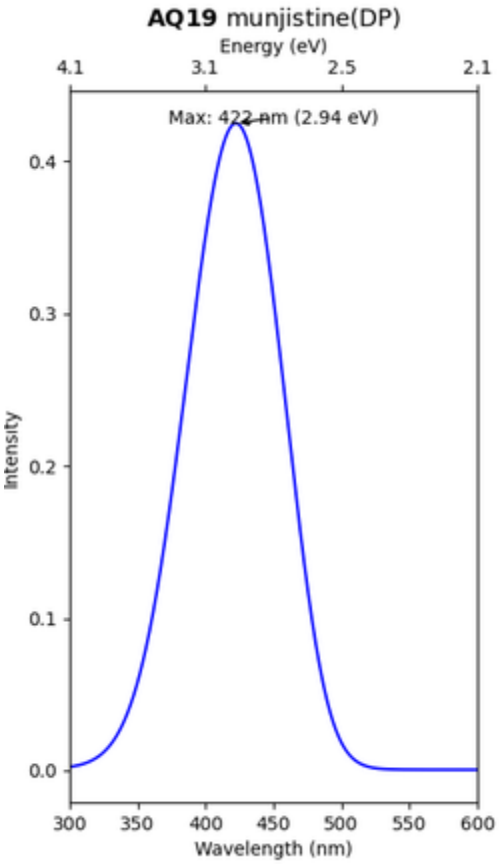

AQ14 1.4-Dihydroxy-2-hydroxymethyl-AQ(D
Energy (eV)
4.1
3.1
2.5
2.1
Max: 448 nm (2.77 eV)
Intensity
Wavelength (nm)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
300
350
400
450
500
550
600

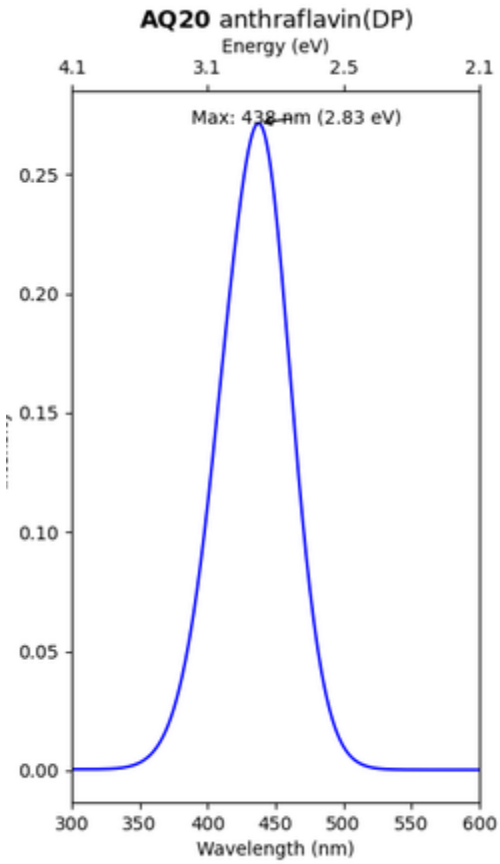

AQ15 alizarine(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 520 nm (2.39 eV)
Intensity
Wavelength (nm)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
300
350
400
450
500
550
600

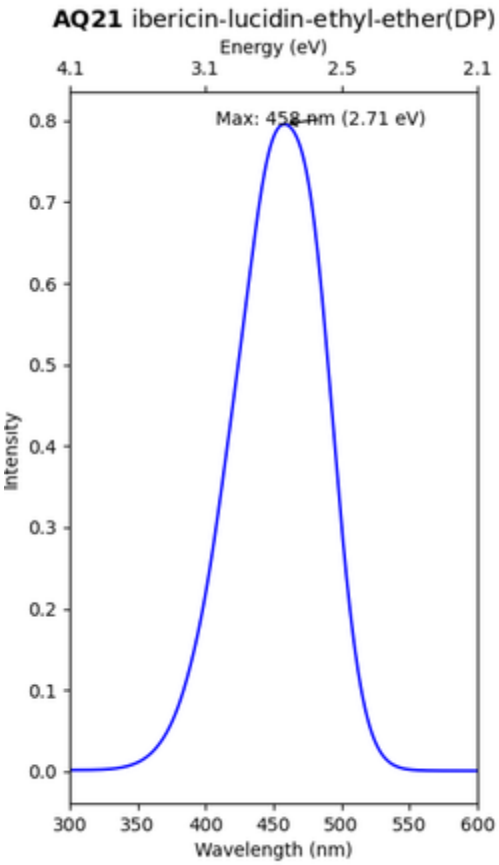

AQ16 xanthopurpurine(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 498 nm (2.49 eV)
Intensity
Wavelength (nm)
0.0
0.1
0.2
0.3
0.4
0.5
300
350
400
450
500
550
600

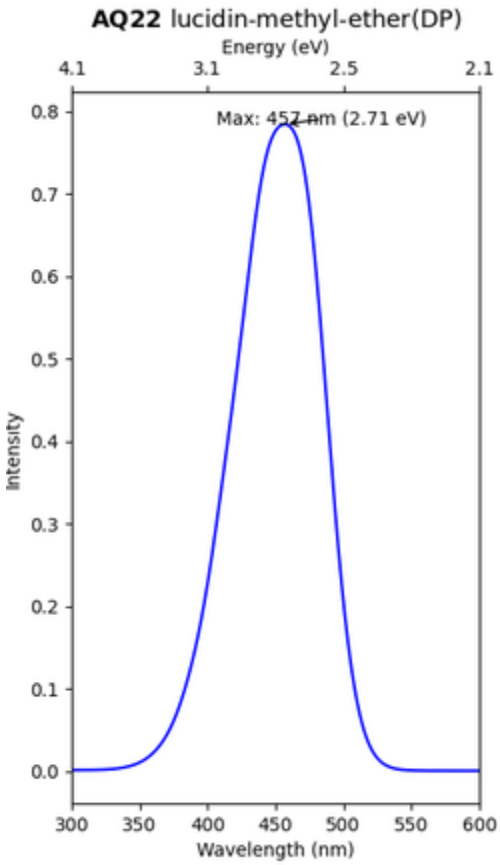

AQ17 rubiadine(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 477 nm (2.60 eV)
Intensity
0.0
0.1
0.2
0.3
0.4
0.5
0.6
300
350
400
450
500
550
600
Wavelength (nm)

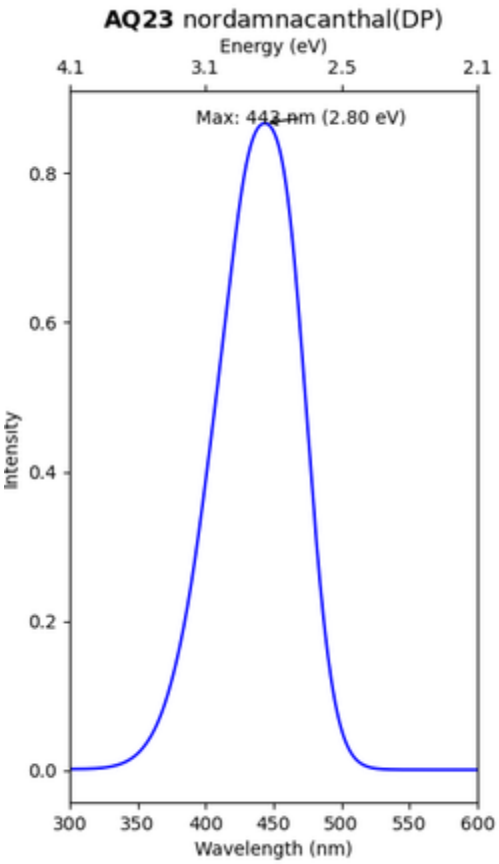

AQ18 lucidine(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 478 nm (2.59 eV)
Intensity
0.0
0.1
0.2
0.3
0.4
0.5
0.6
300
350
400
450
500
550
600
Wavelength (nm)

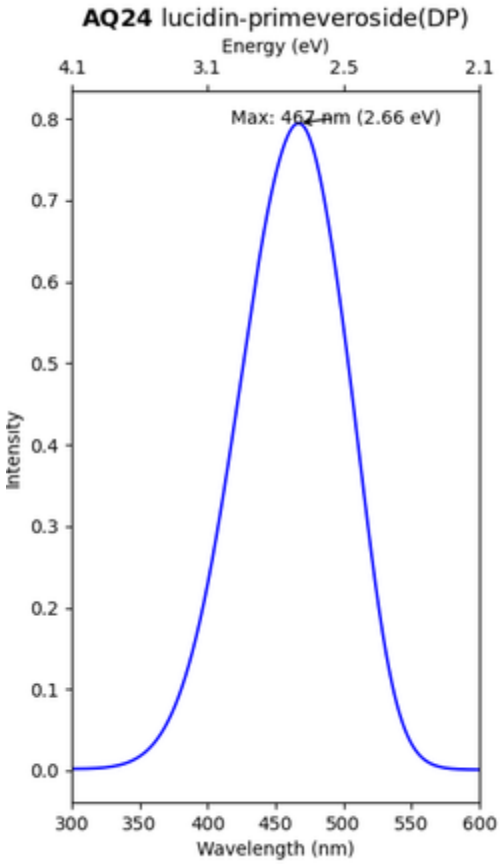

AQ19 munjistine(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 422 nm (2.94 eV)
Intensity
0.0
0.1
0.2
0.3
0.4
300
350
400
450
500
550
600
Wavelength (nm)

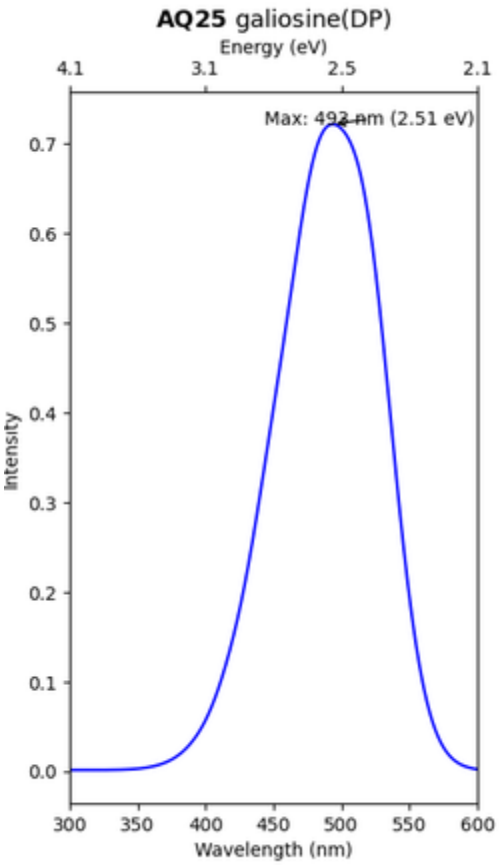

AQ20 anthraflavin(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 438 nm (2.83 eV)
0.00
0.05
0.10
0.15
0.20
0.25
300
350
400
450
500
550
600
Wavelength (nm)

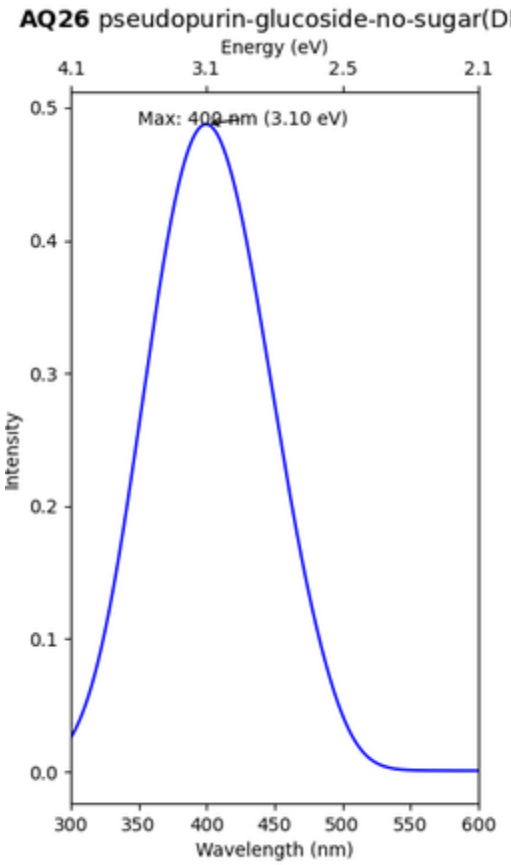

AQ21 ibericin-lucidin-ethyl-ether(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 458 nm (2.71 eV)
Intensity
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
300
350
400
450
500
550
600
Wavelength (nm)

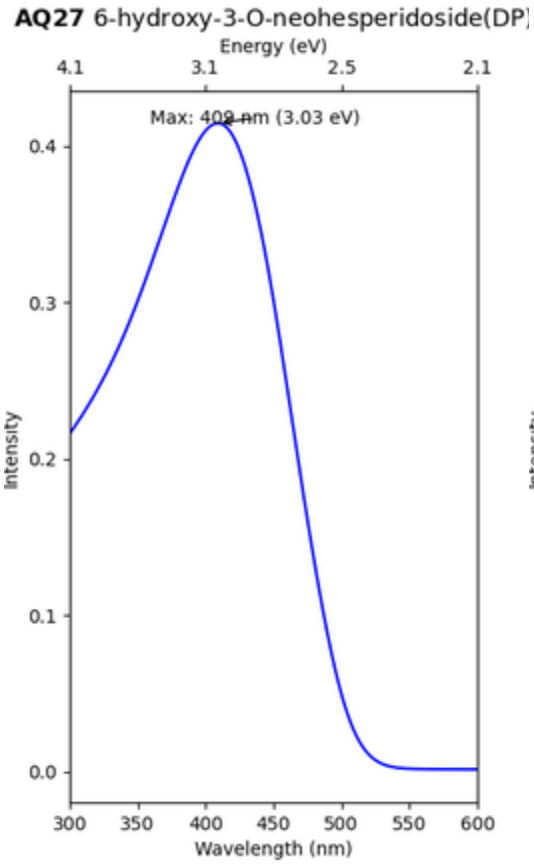

AQ22 lucidin-methyl-ether(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 457 nm (2.71 eV)
Intensity
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
300
350
400
450
500
550
600
Wavelength (nm)

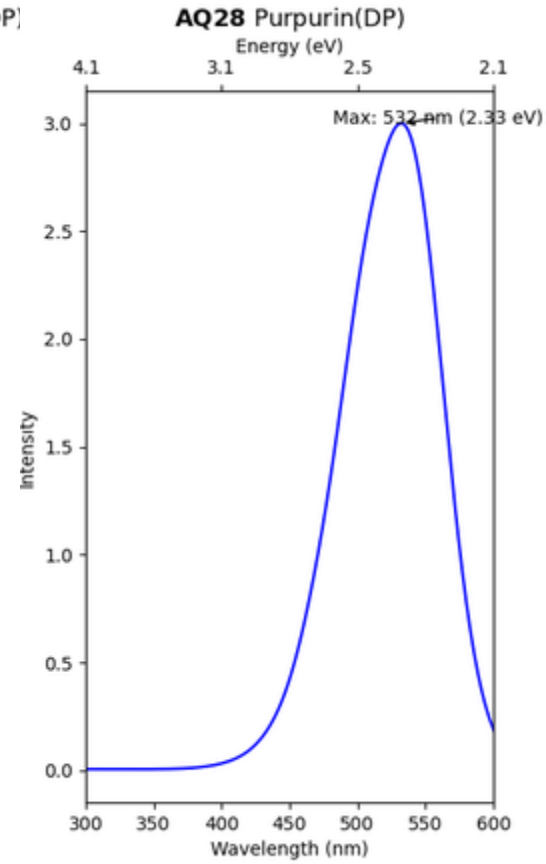

AQ23 nordamnacanthal(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 443 nm (2.80 eV)
Intensity
0.0
0.2
0.4
0.6
0.8
300
350
400
450
500
550
600
Wavelength (nm)

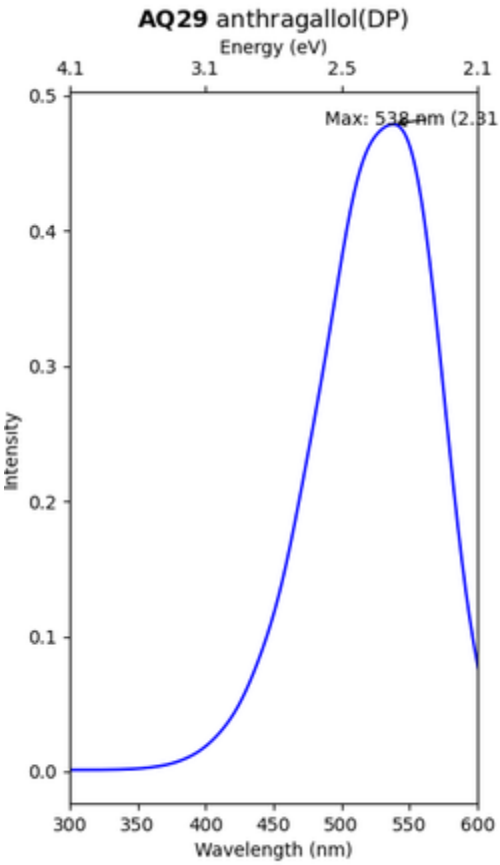

AQ24 lucidin-primeveroside(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 467 nm (2.66 eV)
Intensity
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
300
350
400
450
500
550
600
Wavelength (nm)

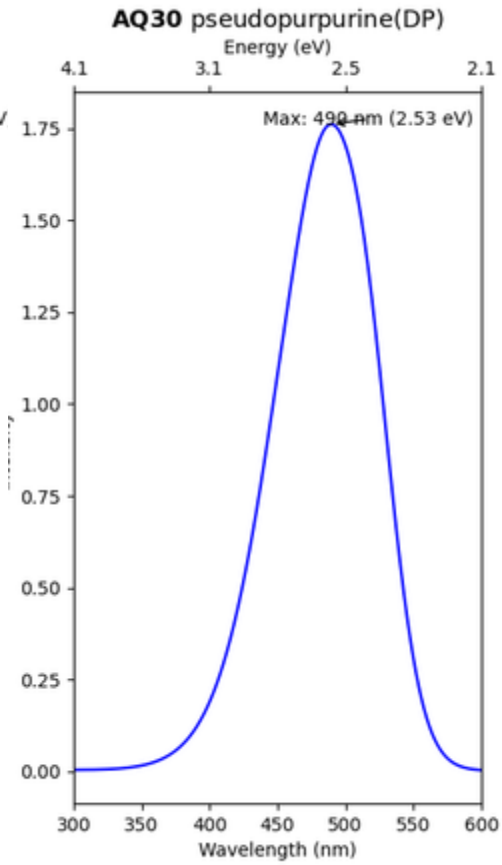

AQ25 galiosine(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 493 nm (2.51 eV)
Intensity
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
300
350
400
450
500
550
600
Wavelength (nm)

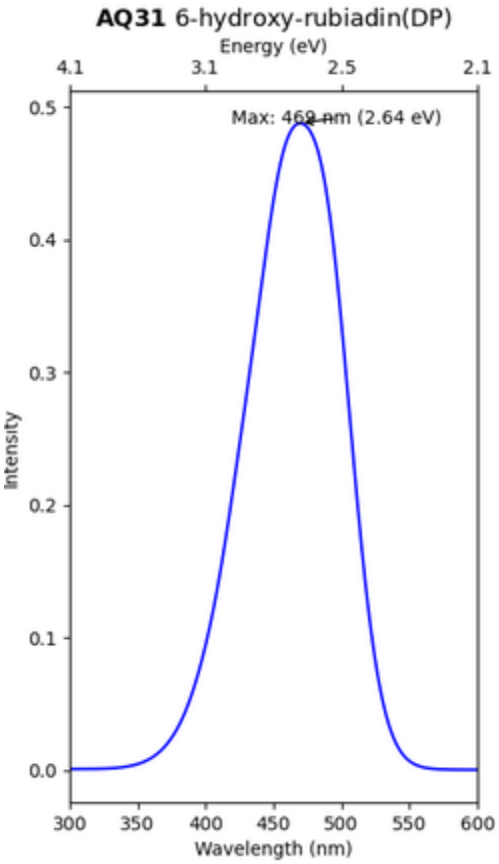

AQ26 pseudopurin-glucoside-no-sugar(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 400 nm (3.10 eV)
Intensity
0.0
0.1
0.2
0.3
0.4
0.5
300
350
400
450
500
550
600
Wavelength (nm)


AQ27 6-hydroxy-3-O-neohesperidoside(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 408 nm (3.03 eV)
Intensity
0.0
0.1
0.2
0.3
0.4
300
350
400
450
500
550
600
Wavelength (nm)


AQ28 Purpurin(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 532 nm (2.33 eV)
Intensity
0.0
0.5
1.0
1.5
2.0
2.5
3.0
300
350
400
450
500
550
600
Wavelength (nm)


AQ29 anthragallol(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 538 nm (2.31 eV
Intensity
0.0
0.1
0.2
0.3
0.4
0.5
300
350
400
450
500
550
600
Wavelength (nm)


AQ30 pseudopurpurine(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 490 nm (2.53 eV)
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
300
350
400
450
500
550
600
Wavelength (nm)


AQ31 6-hydroxy-rubiadin(DP)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 469 nm (2.64 eV)
Intensity
0.0
0.1
0.2
0.3
0.4
0.5
300
350
400
450
500
550
600
Wavelength (nm)

SUMMARY

Article 3: "Progress in modeling complex dye mixtures: a case study for cultural heritage."

In reality, like many natural dyes, madder is a complex mixture of multiple colourants. The next step involved modelling the spectral properties of dye mixtures. We successfully predicted the spectral properties of all dyes extracted from *Rubia tinctorum* L., the most commonly used madder species in France. The results were presented in a third paper. High-Performance Liquid Chromatography with UV and mass spectrometry (HPLC-UV-MS) was used to identify the constituent molecules and their relative proportions in the roots of available madder. We then applied the previously validated protocol to predict the spectra of each compound. Different approaches were tested to incorporate vibrational effects: AH (excluding sugars), VG, and AHAS. Some molecules containing glycosidic groups, like lucidin primeveroside, failed to converge with AH. To overcome this, the glycosidic group was replaced with a methoxy group (non-involved in electronic transitions), allowing successful convergence. AHAS and VG approaches were also tested with the sugar groups retained. Among these, AH (without sugars) and AHAS gave the most accurate results, with deviations of $\Delta E = 3.7$ and $\Delta E = 4.5$, respectively, while VG showed a larger deviation ($\Delta E = 10.4$). This method thus offers promising new avenues for modelling the spectra and colours of complex molecular mixtures. Its ability to accurately predict UV-VIS spectra and colour makes it particularly valuable for non-invasive analysis of heterogeneous cultural heritage objects. Beyond heritage, this approach enables better control of colour palettes, reduces dye waste, and advances research on natural dyes. It could also be extended to the study of tautomers, isomers, degradation products, and other complex dye systems.

ARTICLE 3

Progress in modeling complex dye mixtures: a case study for cultural heritage

Linh T.T. TRAN [a], Maguy JABER [a], Romain BERRAUD-PACHE* [a]

[a] Laboratoire d'Archéologie Moléculaire et Structurale (LAMS), Sorbonne Université, CNRS, UMR 8220, Paris, 75005, France

KEYWORDS: Color prediction • mixture • natural dye • madder • UV-VIS spectra • Computational Chemistry

New J Chem 2025;49:6114–9. https://doi.org/10.1039/D4NJ05450D.

ABSTRACT

Unlike synthetic dyes that produce a single hue, natural dyes offer a rich spectrum of colors. In old times, artisans controlled these colors through their expertise, making the results dependent on each dyer's skill and subjectivity. Nowadays, modern computational methods allow precise predictions of visible absorption spectra, enabling a more standardized color control. However, most studies focus on individual molecules, while natural dyes are made of complex mixtures. We address this gap by proposing improved methods for predicting the spectra of a mixture of natural dyes, focusing on the madder plant, a renowned source of complex colorants. Molecular compositions were determined using liquid chromatography, and the visible spectrum of the mixture was computed, converted to standard color coordinates, and compared with experimental data. Our approach achieved high accuracy, with deviations within accepted theoretical error margins, and successfully handled large molecules with sugar groups. Our workflow can be used for other various chemical systems, notably dye mixtures.

1. INTRODUCTION

Historically, the focus of dyers and painters was on achieving vibrant colors rather than on compound purity, resulting in a wide range of side products. They often used a blend of supplementary agents, sometimes combined with mordants, further adding to the system's complexity. [89] Moreover, even a single natural dye is a complex mixture of multiple chromophores. Over time, these dyes and their by-products degrade due to exposure to air, light, and humidity, while also interacting with each other and the materials they are applied to. This dynamic and ever-changing nature makes their accurate characterization challenging, particularly in the field of cultural heritage. [56,40,233,237,238]

Hence, sophisticated analytical methods are required for their identification. Many experimental methods for dye analysis exist and can be divided into two main groups: invasive and non-invasive techniques.[37,38,45,50,54,182,193,239] Non-invasive methods are

generally favored in the field of cultural heritage, as they preserve the integrity of the sample such as portable UV/IR or Fiber Optic Reflectance Spectroscopy (FORS).[45,48,182,240] On the contrary, invasive methods, like High-Performance Liquid Chromatography (HPLC), and Mass Spectrometry (MS), are used to determine the composition of historical samples but result in damaging them. For the last decades, HPLC combined with sensitive and selective detection techniques such as diode array detector (DAD) and more recently with MS has been widely applied to identify the colorants. [19,35,37,38,100] However, HPLC-DAD often requires reference standards that are often unavailable commercially, expensive, difficult to synthesize or degraded quickly. [19,39,42,55–57,100,118,233,237]

To overcome this gap, computational chemistry offers an alternative method for predicting properties like UV-VIS spectra of inaccessible dyes, deconvolute overlapping chromatographic peaks, and compensate for missing reference data.[103,147,241,242] Therefore, computational chemistry is widely applied in many different fields except in cultural heritage, where standard chemical techniques are favored. The commonly used computational approach is conducting classical molecular dynamics, and then refining the results with dozens of single-point calculations to accurately capture the excitation energy. These values are finally convoluted in a spectrum. While this approach is very accurate, especially for polar solvents, it is computationally expensive, particularly for a mixture of molecules. Another interesting approach is to compute the rates and spectra by solving the Fermi golden rule based on a path integral approach. [175] Despite this, computational studies predominantly focus on single-molecule properties, limiting their applicability in predicting the spectroscopic behavior of historic dye mixtures due to their complex composition, containing multiple molecules, impurities from historical dyeing practices, and degradation by-products.

The aim of this study is to bridge that gap and go beyond the single molecule approximation by introducing a novel approach for predicting the properties and colors of historic natural dye mixtures. To achieve this, we combine experimental analysis with computational modeling to accurately simulate both the color and absorbance spectra of the dyes in the madder plant. Madder roots are known to be composed of a complicated mixture.

2. MATERIALS & METHODS

HPLC-UV-MS and UV-VIS spectra

Alizarin (97%) and purpurin (90%) were purchased from Sigma Aldrich, while madder extract powder (Rubia tinctorum L.) was obtained from Ôkhra; more details are in the S.I. Samples were solubilized in MeOH/H_2O (50:50), sonicated for 10 minutes, and centrifuged before LC-UV-MS/MS and UV-VIS analysis. Chromatographic separation was performed on a Zorbax 300SB-C8 column with a 1290 Infinity II LC system coupled to a Q-TOF 6540 mass spectrometer. UV-VIS spectra of madder extract were recorded using an Ocean Optics Flame spectrometer with a Deuterium-Tungsten Halogen lamp.

Computational studies and modeling

Calculations have been performed using the ORCA quantum chemistry package. [146] A brief overview of the computational workflow is presented here (all the detailed computational

methods are provided in S.I.[176,178,216] First, molecular structure geometries were optimized using density functional theory (DFT) together with the PBE0 functional and dispersion correction and triple zeta def2-TZVP basis set.[107,151,153,166,196,208,243] Then, the excited state geometries and vertical energies were performed using the same procedure as before but with the time-dependent DFT algorithm. [107,117,176,196,244] Only one bright state is present in the visible region of the spectra and was selected accordingly. This transition can be described as a combined n->π* and π->π* located on the conjugated rings, with the difference densities plotted in SI.8. [48]

As the pH of the water solution is about 6.5, the considered madder dyes are either in neutral or deprotonated forms, with the one used in our calculations listed in Figure 1. For neutral forms, an implicit water model was used as it interacts less with the solvent.[234] For deprotonated molecules, 3 explicit water molecules were added to take into account the solvent effect, notably hydrogen bonds, as stated in a previous study. [234] For madder's molecules, the color derives from only one bright state The vibronic effects for each dye were added following different methods (VG, AHAS, or AH) and using the Excited State Dynamic module in order to have access to the computed absorption spectra.[48,107,117,178,245] AH requires geometry and frequencies at the ES minimum, while VG and AHAS approximate the excited-state structure via a short displacement of the ES potential energy surface. VG approach assumes that the GS is the same as the ES one, while AHAS computes the frequencies at the new position. To model the absorption spectra of Madder, each of the computed absorption spectra was weighted by the molecular proportions from Table 1 (including oscillation force), taking into account the ratios of deprotonated forms in function of pH, and then finally scaled to experimental data for effective comparison. Finally, the absorbance spectra were converted to color using CIELAB color coordinates. The color differences between the computed and experimental madder colors were quantified using ΔE_{2000} (details can be found in S.I.). A ΔE_{2000} value below 2 indicates no visible difference, 2-10 suggests a noticeable but subtle difference, and values above 10 imply a more pronounced difference.

3. RESULTS AND DISCUSSION

We applied this protocol to madder (*Rubia tinctorum* L.), a natural red dye with a long history of use. The anthraquinone colorants are extracted from its roots, which are present either in glycosidic form or as their aglycones counterparts. The glycosidic forms are known to be easily degraded by enzymatic hydrolysis or during chemical extraction processes, which complicates their experimental characterization.[56,100]

To determine the percentage of each molecule in our madder sample (*Rubia tinctorum* L.), we performed sonication extraction in a 50:50 water-ethanol mixture and measured the relative percentages using LC-UV-MS/MS. Table 1 compares our sample's percentages with literature values for the same madder species. Lucidin Primeveroside (31.1%) and Ruberythric Acid (31.8%) are the most prevalent in both, though Galiosin is lower in our sample (2% vs. 10.4%) while purpurin is higher (11.4% vs. 1.2%). The variation may be due to Galiosin hydrolyzing to Pseudopurpurin, which then decarboxylates to purpurin. The slight decrease in Ruberythric Acid (31.8% vs. 36.2%) could result from its hydrolysis to Alizarin.[29,39,94,246]

Table 1. *Relative percentage (%) of main madder molecules in a madder extract (Rubia tinctorum L.) measured by LC-UV-MS/MS compared to literature values.[239] The sum of relative percentage of identified madder is 95.6%. The remaining 4.4% was attributed to several minor components that could not be identified and was not taken into account. Below are their molecular structures.*

	Relative percentage (%)	
Detected molecules	Madder extract	Literature[239]
Lucidin primeveroside	31.1	31.1
Ruberythic acid	31.8	36.2
Galiosin	2.0	10.4
Rubiadin primeveroside	1.7	1.4
Alizarin	6.0	8.3
Purpurin	11.4	1.2
Lucidin	-	2.7
Munjistin	8.6	-
Pseudopurpurin	0.4	-
Nordamcanthal	2.6	-
Remaining Unidentified molecules	4.4	8.7

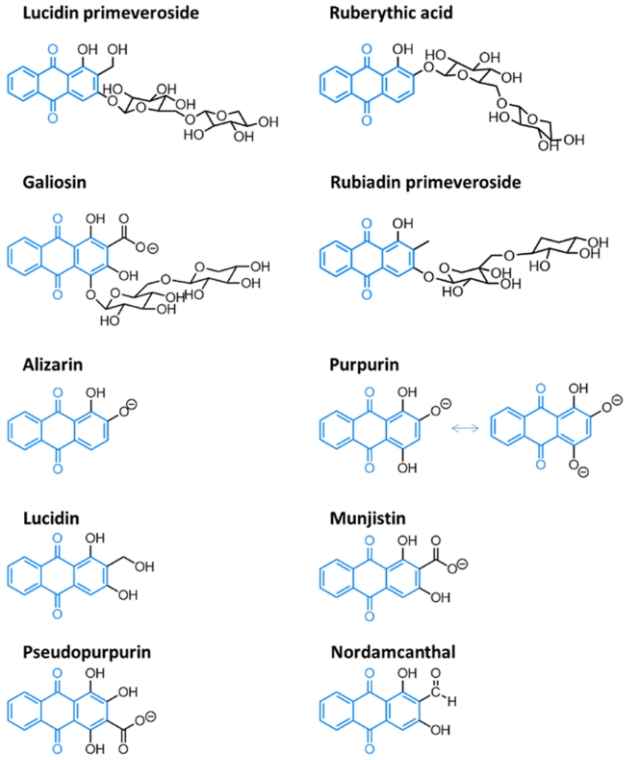


Figure 1. *Molecular structures used. The blue skeleton corresponds to the anthraquinone moiety.*

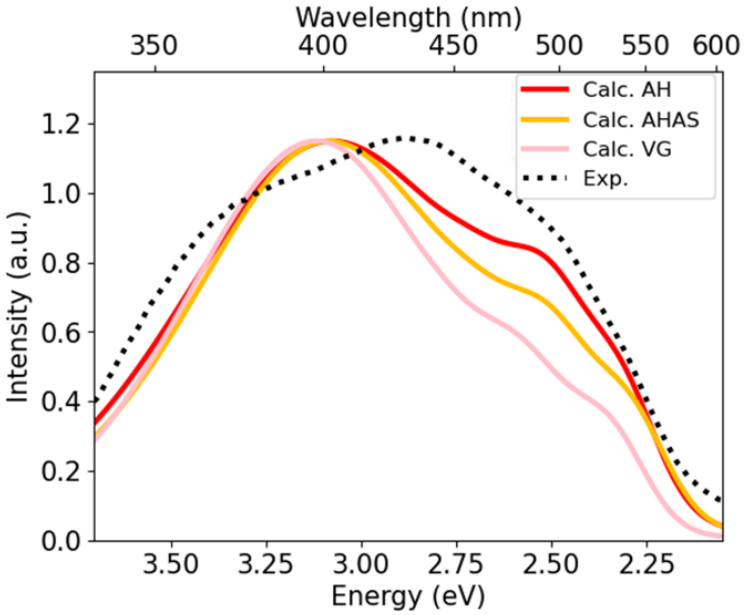


Figure 2. *Calculated VG (pink full line), calculated AHAS (yellow full line), calculated AH (red full line) and experimental (black dotted line) UV-VIS spectra for madder specie Rubia tinctorum L. in water.*

We have computed the absorbance spectra for madder molecules using the computational protocol described, except for dyes that contain a glucose moiety (lucidin primeveroside, ruberythric acid, galiosin, and rubiadin primeveroside). For these four, the sugar moiety was first replaced with a methoxy group to ensure convergence for AH before proceeding as above. The spectra were then summed based on the percentages from UV-LC-MS/MS (Table 1). Figure 2 compares the VG, AHAS, and AH methods with the experimental spectrum (black dotted line). The experimental peak is at 2.89 eV (429 nm) with shoulders at 2.48 eV (500 nm) and 3.35 eV (370 nm). Both AH (red) and AHAS (orange) closely match the experimental shape, while VG deviates more. VG peaks at 3.1 eV (400 nm), AHAS at 3.08 eV (402 nm), and AH at 3.07 eV (403 nm), with deviations of 0.21 eV (29 nm), 0.19 eV (27 nm), and 0.18 eV (26 nm), respectively. The AH spectra align more closely with the experimental data, particularly in the 480-550 nm shoulder region for both position and intensity. We can notice that VG and AHAS fail to reproduce the intensity of this shoulder. Figure 3 shows the individual contributions in the summed AH spectra. The 400 nm peak is primarily due to ruberythric acid and lucidin primeveroside, while the 480-550 nm shoulder is attributed to purpurin (monoanionic and dianionic forms) and alizarin (monoanionic form). Molecules present below 2% have minimal impact on the spectra

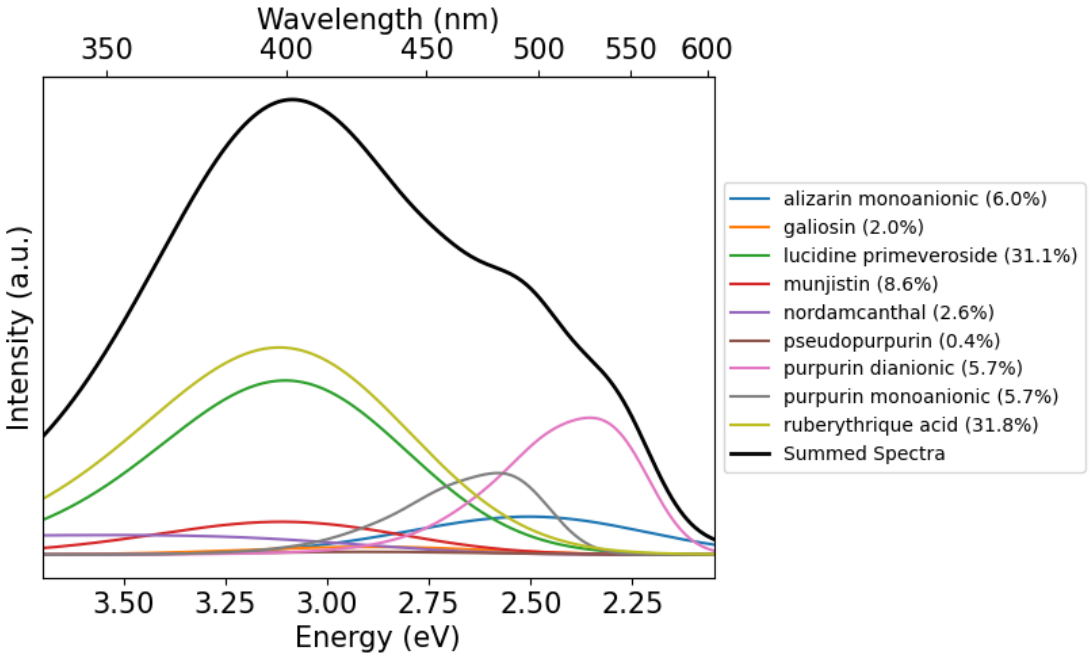


Figure 3. *Calculated UV-VIS spectra (not normalized) for madder with AHAS and the individual contribution of each component.*

The similarity of the UV-VIS spectra across different computational methods (VG, AHAS, or AH) suggests that all three approaches provide a reliable prediction of the absorption spectra and the resulting color. The choice of computational method can be tailored depending on the specific needs of the study. For instance, if computational efficiency is prioritized, the VG approximate method might be preferable due to its significantly reduced computational cost, while maintaining reasonable accuracy and ensuring convergence. This can be used for large molecules like glucosidic molecules or mixtures with a large number of different molecules. Conversely, if a more refined molecular description is required, the AH or AHAS methods might be more appropriate, as they better account for changes according to the excited states. The spectra of glycosidic molecules made with the AH protocol fail at first before substituting the sugar moieties with a methoxy group. It has however no impact on the VE due to the sugar not influencing the transition orbitals. AHAS, offers the best balance between accuracy and computational efficiency, particularly for mixtures containing flexible molecules, as it ensures convergence without significant loss of precision.

The shift in computed UV-VIS spectra may not be perceptible to the human eye due to varying color sensitivity. To evaluate this, both spectra were converted to CIELAB coordinates (L*, a*, b*) using an in-house script, with color difference measured by ΔE_{2000} (Figure 4). ΔE_{2000} values were 10.4 for VG, 4.5 for AHAS, and 3.7 for AH, indicating similar results, with AH being the most accurate for color prediction. A ΔE_{2000} between 2 and 5 is generally acceptable in fields like material science and art restoration

Madder *Rubia Tinctorum* L.	Calculated				Experimental				
CIELAB coordinate	L*	a*	b*	color	L*	a*	b*	color	ΔE_{2000}
VG	87.0	16.1	61.1		74.6	27.6	67.3		10.4
AHAS	80.2	26.7	59.9		74.6	27.6	67.3		4.5
AH	79.3	29.8	65.3		74.6	27.6	67.3		3.7

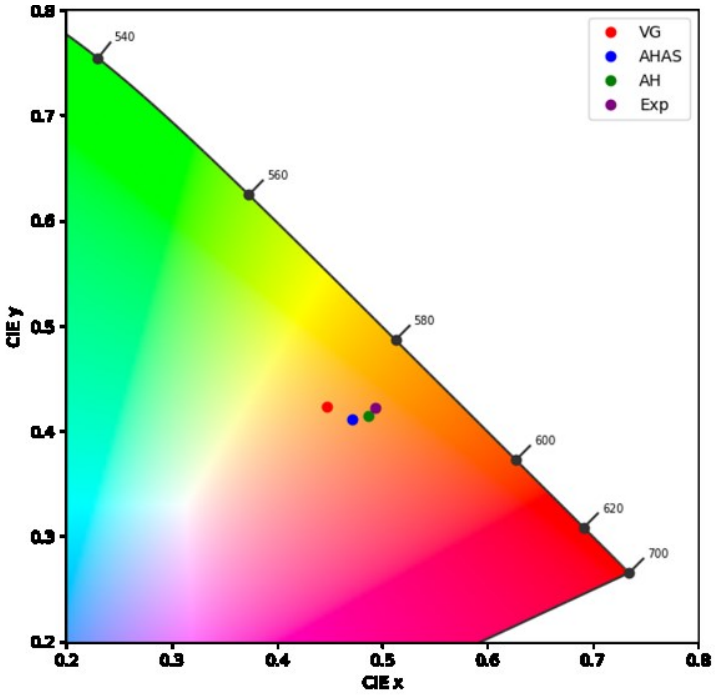


Figure 4. Calculated and experimental color and CIELAB coordinates for Rubia tinctorum L. are compared, with color difference assessed using ΔE_{2000}. The CIE chromaticity diagram follows below.

4. CONCLUSION

In conclusion, this breakthrough allows the prediction of a complete madder spectrum, including the contribution of several anthraquinones, which has never been achieved before. This method paves the way for exciting possibilities in predicting the colors and spectra of other molecular mixtures. Its precision in forecasting UV-VIS spectra and colors for natural dyes makes it highly valuable, particularly for historical dyes in cultural heritage, because it is non-invasive.

By predicting dye mixtures' spectra, we gain better control over color palettes, reduce dye waste, and advance natural dye research. Additionally, this approach could guide future studies on various chemical systems, including dye mixtures, tautomers, isomers, degraded products, and other chemical mixtures

5. ASSOCIATED CONTENT

Supporting Information available: Experimental and computational protocol and workflow graph, details of LC-UV-MS spectra, ion chromatogram, functional benchmark, difference densities, computation time, ΔE2000 formula, computed spectra and optimized geometries coordinates.

AUTHOR INFORMATION

Corresponding Author

* Dr. BERRAUD-PACHE Romain

Laboratoire d'Archéologie Moléculaire et Structurale (LAMS)

Sorbonne Université, CNRS, UMR 8220, Paris, 75005, France

E-mail: romain.berraud-pache@sorbonne-universite.fr

Author Contributions

Thanh Huyen Linh Tran: Writing – review & editing, Writing-original draft, Visualization, Software, Investigation, Formal analysis, Data curation, Methodology. Maguy Jaber: Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Conceptualization. Romain Berraud-Pache: Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Formal analysis, Data curation, Conceptualization

Author information

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

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

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

ACKNOWLEDGMENT

This work was granted access to the HPC resources of IDRIS under the allocation 2023-AD0814570 made by GENCI.

We thank Dr. Salomé POYER from the Institut de Chimie des Substances Naturelles (ICSN) and Dr. Alain BRUNELLE for their help in the analysis of LC-UV-MS/MS spectra

SUPPLEMENTARY DATA

1. *Computational method*

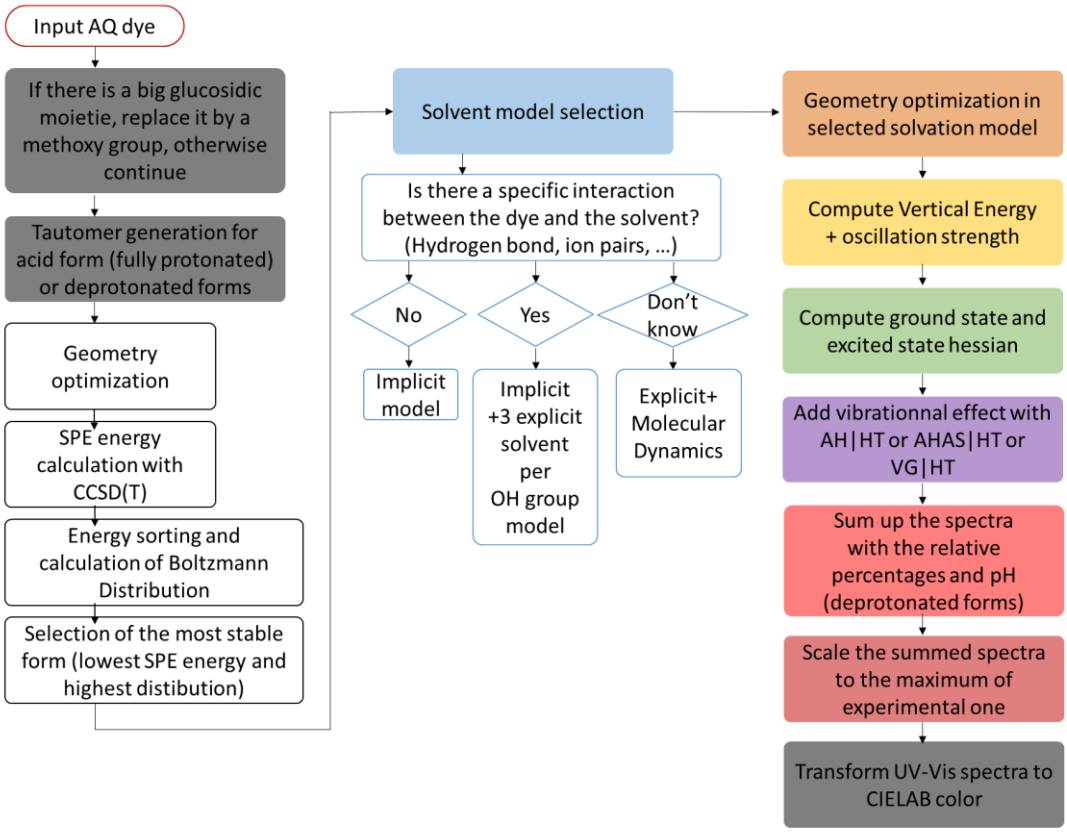
This computational protocol has been studied by us and compared with available experimental data. We have reported the accuracy of this protocol in a previous article. [234] The computational details can be found in the original article.

UV-VIS Spectrum computation

The relative percentage of madder extract was measured by LC-UV-MS/MS method. Using the measured percentages (Table 1), we computed the absorption spectra and color for all detected molecules. All calculations were carried out using a development version of the ORCA quantum chemistry program package.[146] We first optimized all molecules with the PBE0 functional with the D3(BJ) dispersion correction, along with def2-TZVP basis and matching auxiliary basis. [151,153,166,208] The same functional was used for time-dependent DFT (TDDFT) calculations. There is no imaginary frequencies above 10 cm^{-1}. For the neutral form, implicit water solvation was used with the linear response Conductor-like Polarizable Continuum Model (C-PCM). [234] For deprotonated forms, the latest Grimme's open-source Conformer-Rotamer Ensemble Sampling Tool (CREST), which relies on the xTB method was used to generate different deprotonated forms.[205,210,212] Then, the geometry was reoptimized with PBE0 and the single point energy (SPE) was computed using DLPNO-CCSD(T). [211] The deprotonated form which exhibit lowest SPE was taken into account for further computation. Then, a combined approach involving both implicit and explicit water molecules (3 water molecules per OH group) was used. The water molecules were positioned as to make hydrogen bonds with each hydroxyl group or carboxyl groups. All optimized geometries and other details can be found in the supplementary data. The same protocol was used to compute the vertical absorption energies and excited state (ES) geometries. For madder's molecules, only one bright state (S0→S1) contributes to the visible spectra and was selected accordingly in our calculations. The transition can be described as a combined n->π* and π->π* located on the conjugated rings with the difference densities plotted in SI.7. The SI.6 shows the computed vertical energy compared to experimental data found in the literature.[8] Finally, we have to take into account the pH of the solution of about 6.5 in water. For galiosin, pseudorpurpurin, and munjistin, this leads to the carboxy groups being deprotonated, and for alizarin and purpurin one of their hydroxy groups. The vibrational contributions, including Herzberg–Teller effects, were done using the Excited State Dynamics (ESD) module within ORCA. [176,178,216]

Three methods were tested: Vertical Gradient (VG), Adiabatic Hessian After-Step (AHAS), and Adiabatic Hessian (AH). For the computed spectra, the linewidth was created using a Gaussian distribution and was set with keyword inlinew 2000 cm^{-1}. AH requires geometry and frequencies at the ES minimum, while VG and AHAS approximate the excited-state structure via a short displacement of the ES potential energy surface. VG approach assumes that the GS is the same as the ES one, while AHAS computes the frequencies at

the new position. VG is efficient and reliable for mixture spectra prediction, but AH, though theoretically more accurate, had convergence issues with sugar-containing molecules due to large ES geometry displacements. To address this, sugar moieties were replaced with methoxy groups. AHAS offers a compromise, resolving convergence issues. For full spectra prediction, each computed spectrum (where the oscillator strength "fosc" is proportional to the extinction coefficient) was weighted according to the percentages in Table 1, summed without modifying the oscillator force, and then scaled against experimental data for effective comparison. CIELAB color coordinates were then calculated for the computed and experimental UV-VIS spectra[134,199,247] The CIELAB represents colors with L* (lightness), a* (green-red), and b* (blue-yellow). ΔE measures color differences, with ΔE_{2000} preferred for accuracy. This work will use ΔE_{2000} to compare computed and experimental madder mixture colors. [46]The experimental pH and pKa of alizarin (pKa_1 =5.2 and pKa_2=11.5) and purpurin (pKa_1= 5.2 and pKa_2=6.8) [248] were taken into account in the percentage following the Henderson-Hasselbach's equation.



S.I. 1. *Workflow graph for the computational protocol to compute dye mixture.*

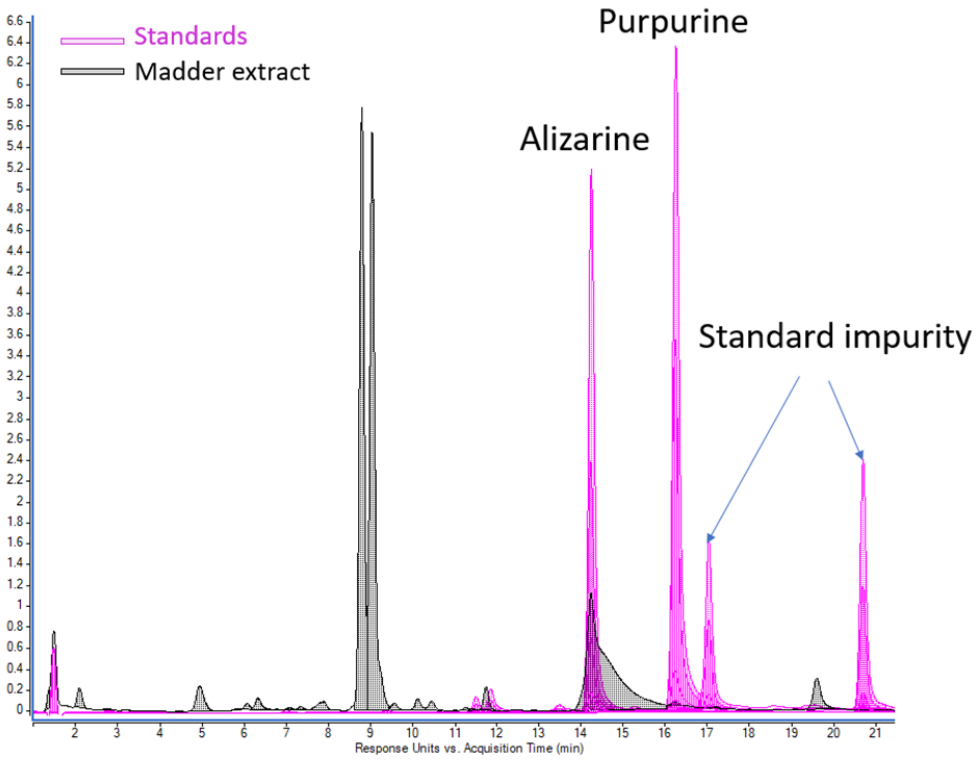
2. *Experimental method*

Alizarin (97 % purity), purpurin (90%) were purchased from Sigma Aldrich and used without further purification. Madder extract powder (*Rubia tinctorum* L.) is purchased from Ôkhra. Three samples were solubilized in MeOH/H2O 50:50 and sonicated 10 minutes. The Madder extract was diluted at 5 mg/mL, and the standards of Alizarine and Purpurin at 1 and 2 mg/mL respectively. Due to incomplete solubilization, samples were

centrifugated and 100 µL of the supernatant was used for LC-UV-MS/MS analysis, 1.5 mL is left for UV-VIS spectra measurement. Relative quantification was performed.

Liquid chromatography-UV-tandem mass spectrometry (LC-UV-MS/MS) experiments were performed on a 1290 Infinity II LC system coupled to a Q-TOF 6540 mass spectrometer equipped with an ESI source operated in the negative and positive ionization mode. The chromatographic separation was carried out on a Zorbax 300SB-C8 (150 × 2.1 mm, 5 µm) column heated at 35 °C at a flow rate of 300 µl/min with water at 0.1% of formic acid as solvent A and acetonitrile as solvent B. Five microliters of samples in MeOH/H2O 50:50 were injected using the following gradient: 0 min (10% B), 25 min (55% B), 30 min (100% B), 35 min (100% B), 36 min (10% B)and 6 min of equilibration time. Ion source conditions were set for negative ion mode: capillary voltage 3.5 kV, Fragmentor voltage 125 V, Skimmer 65 V, Drying gas at 10 l/min, 275 °C, and sheath gas flow at 10 l/min, and 325 °C. UV detection was performed at 232 and 260 nm for quantitation at 260 nm. The TOF was operated in the 50−1700 m/z mass range using a medium isolation window and 20eV of collision energy for MS/MS. Data Acquisition 9.0 and Qualitative Analysis 8.0 software were used to process data.

The UV-VIS spectra of madder extract was measured using an Ocean Optics Flame spectrometer (Model: FLMS00699) with a wavelength range from 300 to 850 nm. The integration time for each measurement ranged from 5 ms to 10 ms. The light source used was a Deuterium-Tungsten Halogen lamp covering the range of 300 nm–850 nm. The blank was done using deionized water. This light source was coupled to the spectrometer using 400 µm diameter optical fibers and positioned through a 1 cm quartz cuvette holder. The acquisition of spectra was facilitated by the Ocean View software.



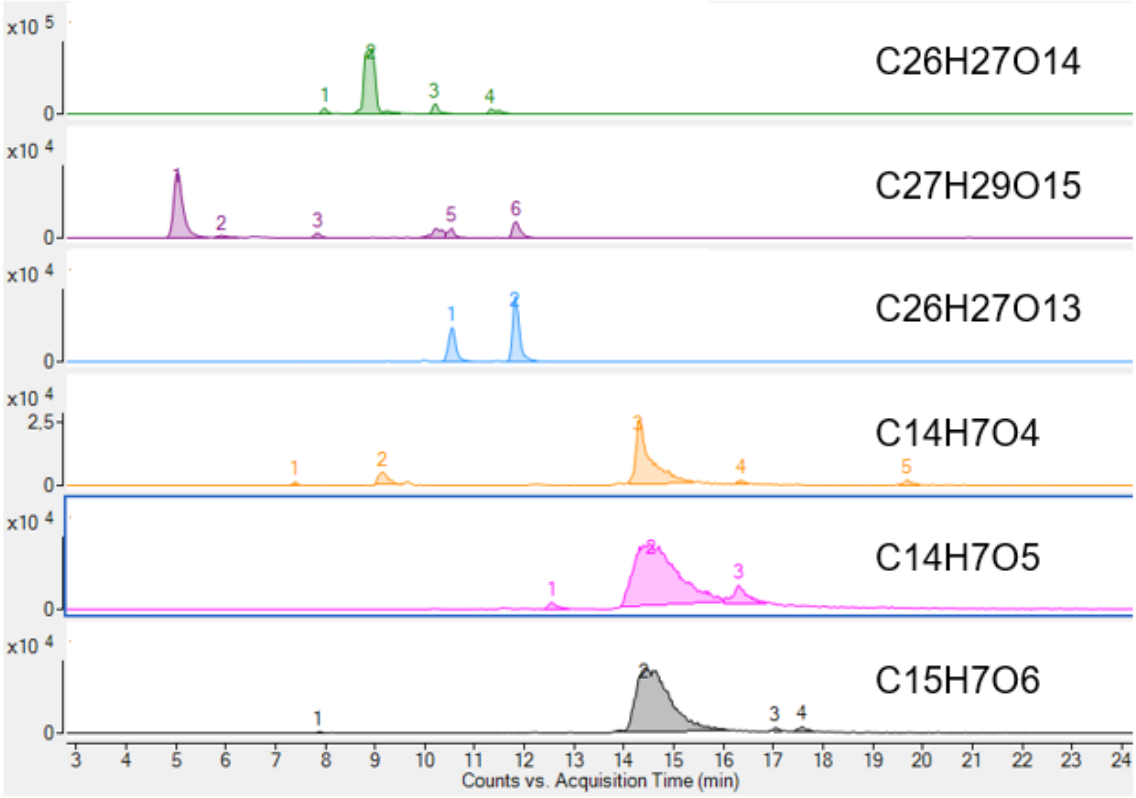
S.I. 2. *Overlay of LC-UV experiments at 260 nm of Madder extract (in black) and Alizarine and Purpurine standard solution at concentrations ranging from 420 µg/mL to 2 ng/mL (in pink).*

S.I. 3. *Relative percentage, retention time, mass/z, chemical formula, and attributed molecule for the madder extract (Rubia Tinctorum L.) using LC-UV-MS/MS.*

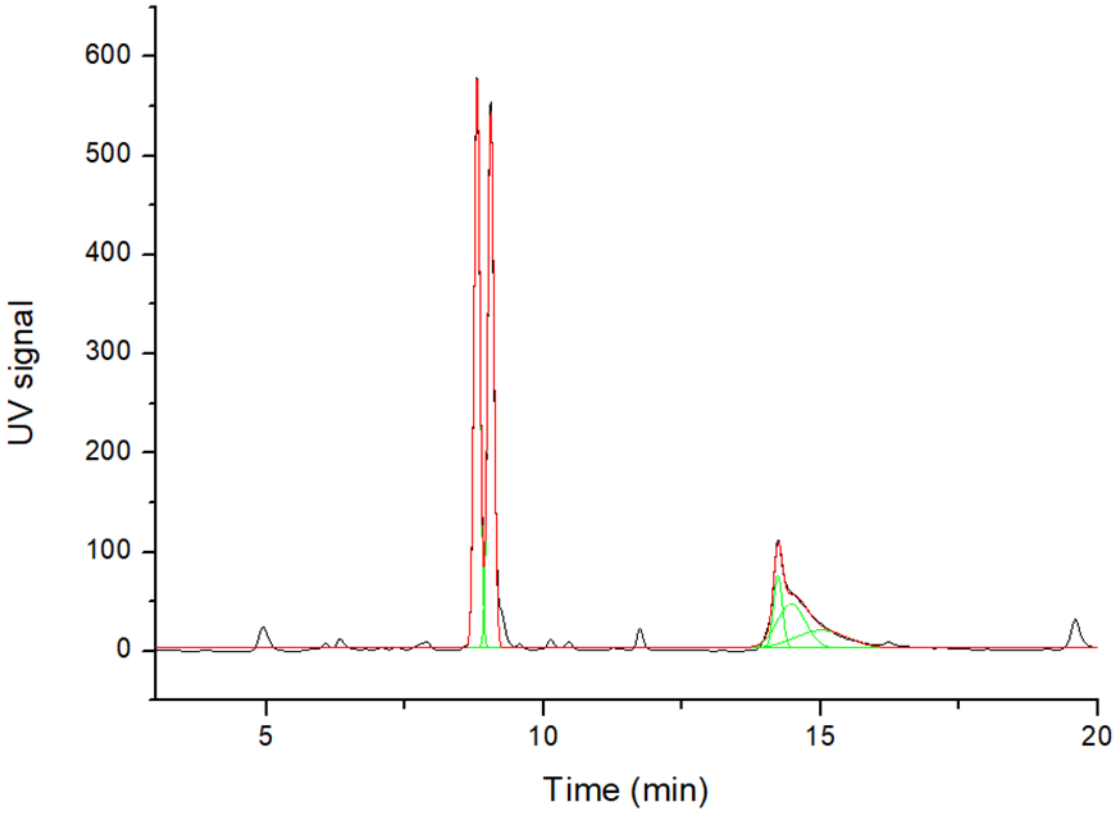
PEAK	RT	AREA SUM %	M/Z $[M-H]^-$	CHEMICAL FORMULAE	MOLECULE ATTRIBUTED
1	1.4	0.4	387.1145	ND	+46
2	1.7	0.05	377.0860	ND	ND
3	2.1	0.97	431.1196	ND	251.0557 + Glc
4	2.7	0.07	459.1142	ND	+46
5	3.2	0.04	387.1140	ND	ND
6	3.9	0.04	729.2606	ND	521.2022+GlcGlc+46
7	4.4	0.03	583.2026	ND	ND
8	4.9	2	593.1151	C26H25O16	Galiosin
9	5.8	0.04	461.0719	C21H17O12	255.0294+Glc+CO_2
10	6.1	0.33	497.1299	ND	ND
11	6.3	0.74	365.0875	ND	+163
12	6.8	0.02	461.0719	C21H17O12	255.0291+Glc+CO_2
13	7.1	0.11	655.2238	ND	+163
14	7.3	0.14	445.0772	C21H17O11	239.0397+Glc+CO_2
15	7.9	0.77	563.1405-593.1506-623.1615	C26H27O14 C27H29O15 C28H31O16	239.0347+Glc-Glc (6-hydroxyrubiadin-3-O-neohesperidoside) 253.0505+Glc-Glc (6-Hydroxyrubiadin-3-O-neohesperidoside) 283.0606+Glc-Glc
16	8.8	30.22	563.1411	C26H27O14	Lucidin primeveroside
17	9	31.75	533.1307	C25H25O13	Ruberythic acid
18	9.6	0.4	401.0873	C20H17O9	239.0339+Glc
19	10.1	0.69	563.1402-593.1505	C26H27O14 C27H29O15	Lucidin primeveroside 269.0460+Glc-Glc

20	10.5	0.55	547.1454-593.1505	C26H27O13 C27H29O15	Rubiadin primeveroside 251.0338+Glc-Glc
21	11	0.01	431.0978	ND	ND
22	11.3	0.13	563.1401	C26H27O14	Lucidin primeveroside
23	11.4	0.09	563.1398	C26H27O14	Lucidin primeveroside
24	11.7	1.19	547.1455-593.1507	C26H27O13 C27H29O15	Rubiadin primeveroside 253.0509+Xyl-Glc+46
25	12.4	0.02	447.1657	ND	+163
26	12.9	0.03	343.0819	ND	ND
27	14.2	6.00	239.0346-	C14H7O4	Alizarin
27	14.2	11.4	255.0300-	C14H7O5	Purpurin
27	14.2	8.6	283.0245-	C15H7O6	Munjistin
27	14.2	0.06	299.0201	C15H7O7	Pseudopurpurin
28	16.2	0.24	255.0298-299.0196	C14H7O5 C15H7O7	Purpurin Pseudopurpurin
29	17.2	0.08	293.1758	ND	
30	17.5	0.02	283.0243	C15H7O6	Munjistin
31	18	0.02	255.0297	C14H7O5	Purpurin
32	18.3	0.02	299.0195	C15H7O7	Pseudopurpurin
33	19.1	0.04	253.0504	C15H9O4	Rubiadin
34	19.6	2.55	267.0296	C15H7O5	Nordamcanthal
35	20.9	0.01	299.0194	C15H7O7	Pseudopurpurin
36	25.8	0.15	299.0192	C15H7O7	Pseudopurpurin

S.I. 4. *Extracted ion chromatogram of various isomers in the Madder extract.*



S.I. 5. *Deconvolution LC-UV experiments at 260 nm of Madder extract and fit of the peak at retention time RT=14.2min.*



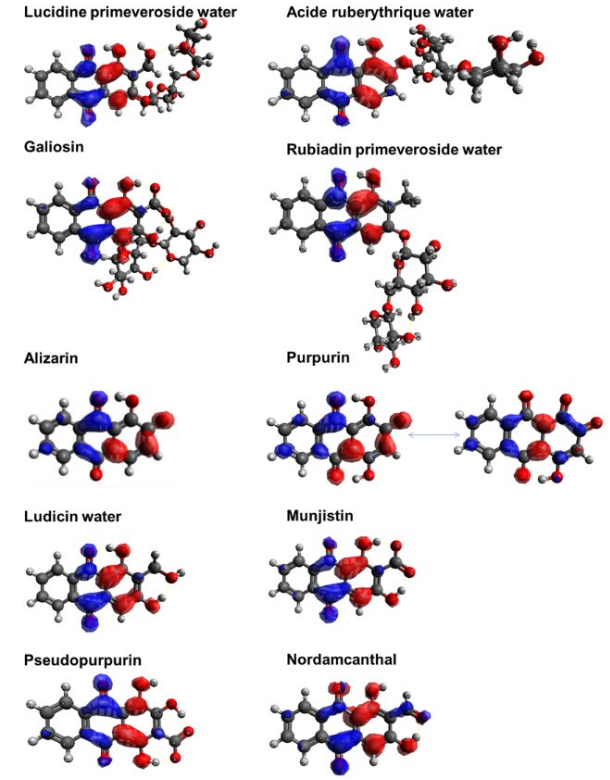
S.I. 6. *Benchmark of functional (B3LYP, B2PLYP and PBE0) for TD-DFT computation of Vertical energy (VE). The table represents the absolute error deviation (calc.-exp.) for computed VE of different deprotonated forms of alizarin.*

	Absolute error deviation (eV)		
	B3LYP	B2PLYP	PBE0
Alizarin neutral	0.06	0.06	0.05
Alizarin monoanionic	0.30	0.27	0.21
Alizarin dianionic	0.01	0.04	0.05

S.I. 7. *Experimental and computed maximum absorption energies of some madder molecules and their respective error deviation (in eV) which are less than 0.10 eV. The experimental data is extracted from the literature.* [59]

Molecules	Calc. VE (eV)	Exp. (eV)	Abs. Calc.-Exp (eV)
Lucidin primeveroside	3.03	3.05	0.02
Ruberythric acid	2.91	2.97	0.06
Galiosin	2.82	2.86	0.04
Rubiadin primeveroside	3.08	3.01	0.07
Anthragallol	3.10	3.00	0.10
Lucidin	3.02	3.00	0.02
Alizarin	2.93	2.90	0.03
Purpurin	2.59	2.58	0.01
Xanthopurpurin	3.05	2.98	0.07
Rubiadin	3.05	3.02	0.03

S.I. 8. *Difference in electronic densities for the first transition $S_0 \rightarrow S_1$ of madder molecules. The red represents an increase in electronic density and the blue a decrease. The isovalue is 0.002 a.u.*



S.I. 9. *Total computation time for each method (VG, AHAS and AH) for alizarin and lucidin primeveroside.*

Calculation time	Alizarin	Lucidin primeveroside
VG	23 h	164 h
AHAS	21 h	164 h
AH	29 h	26 h

©Linh TRAN

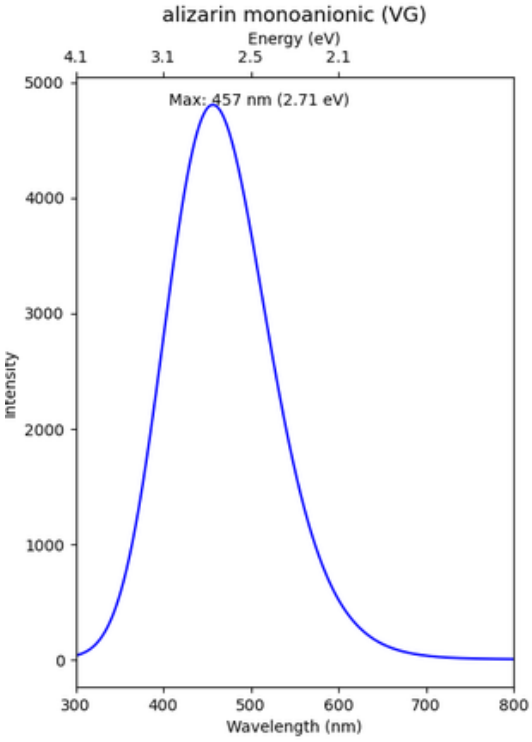
S.I. 10. *ΔE_{2000} formula for color difference between two L*, a*, b* coordinates of two samples.* [138,140]

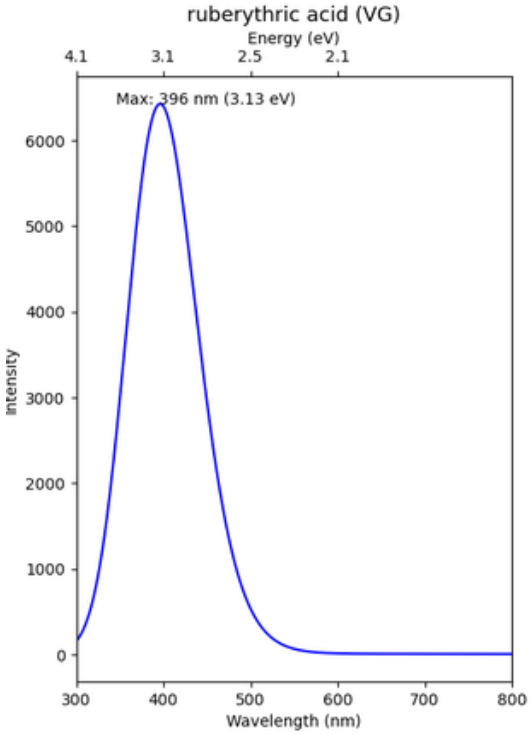
$$\Delta E_{2000} = \sqrt{(\frac{\Delta L'}{K_L S_L})^2 + (\frac{\Delta C'}{K_C S_C})^2 + (\frac{\Delta H'}{K_H S_H})^2}$$

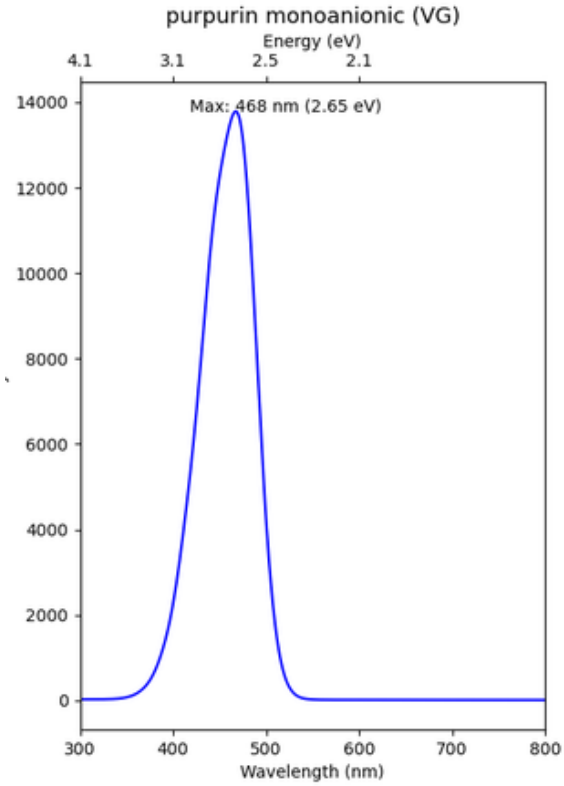
where:

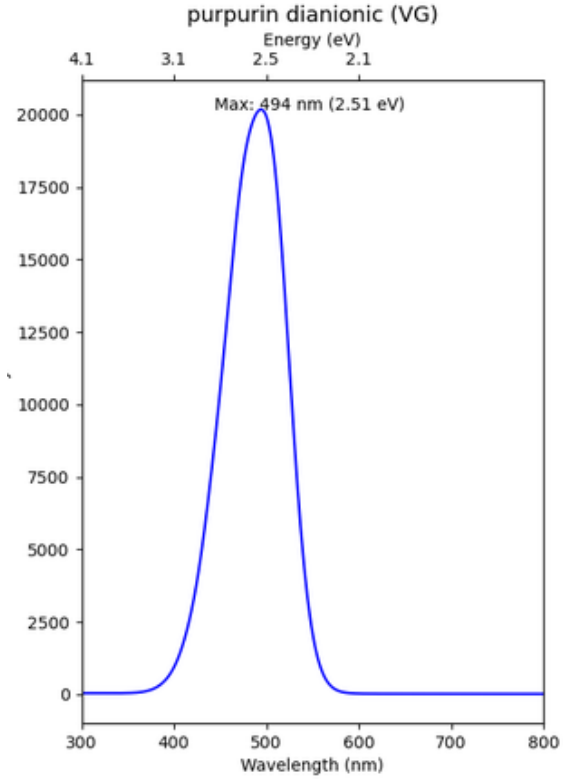
$$\overline{L'} = (L_1 + L_2)/2,$$
$$\Delta L' = L_2 - L_1,$$
$$C_1 = \sqrt{a_1^2 + b_1^2},$$
$$C_2 = \sqrt{a_2^2 + b_2^2},$$
$$\overline{C} = (C_1 + C_2)/2,$$
$$G = (1 - \sqrt{\frac{\overline{C^7}}{\overline{C^7} + 25^7}})/2,$$
$$a_1' = a_1(1 + G),$$
$$a_2' = a_2(1 + G),$$
$$C_1' = \sqrt{a_1'^2 + b_1^2},$$
$$C_2' = \sqrt{a_2'^2 + b_2^2},$$
$$\overline{C'} = (C_1' + C_2')/2,$$
$$\Delta C' = C_2' - C_1',$$
$$\Delta H' = 2\sqrt{C_1' C_2'}\sin(\Delta h'/2),$$
$$S_L = 1 + \frac{0.015(\overline{L}' - 50)^2}{\sqrt{20 + (\overline{L}' - 50)^2}},$$
$$S_C = 1 + 0.045\overline{C}',$$
$$S_H = 1 + 0.015\overline{C}'T,$$
$$\Delta\theta = 30\exp\{-(\frac{\overline{H'} - 275^o}{25})\},$$
$$R_C = \sqrt{\frac{\overline{C^7}}{\overline{C^7} + 25^7}}),$$
$$R_T = -2R_C\sin(2\Delta\theta),$$
$$K_L = 1 - default,$$
$$K_C = 1 - default,$$
$$K_H = 1 - default.$$

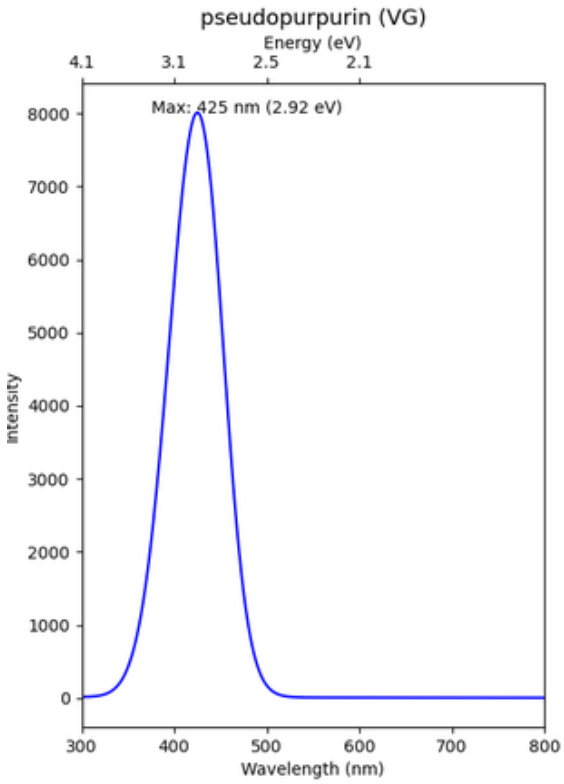
S.I. 11. *Computed individual spectra using Vertical Gradient (VG) method.*

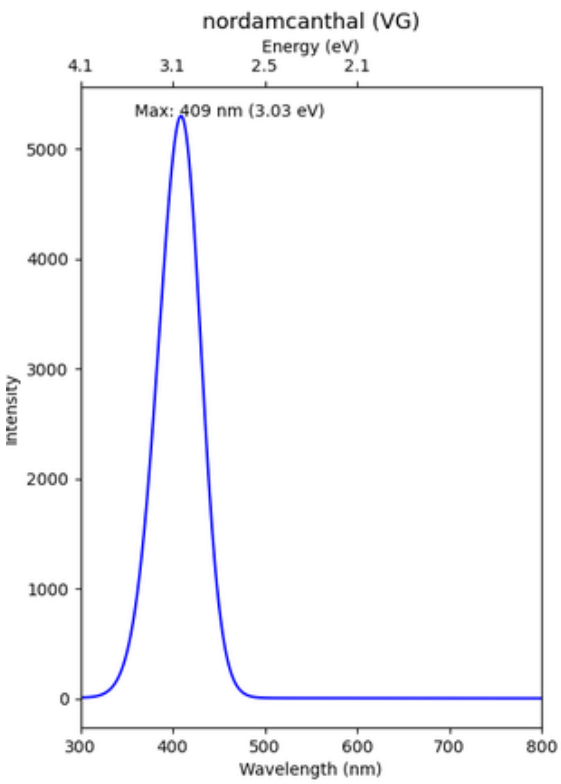


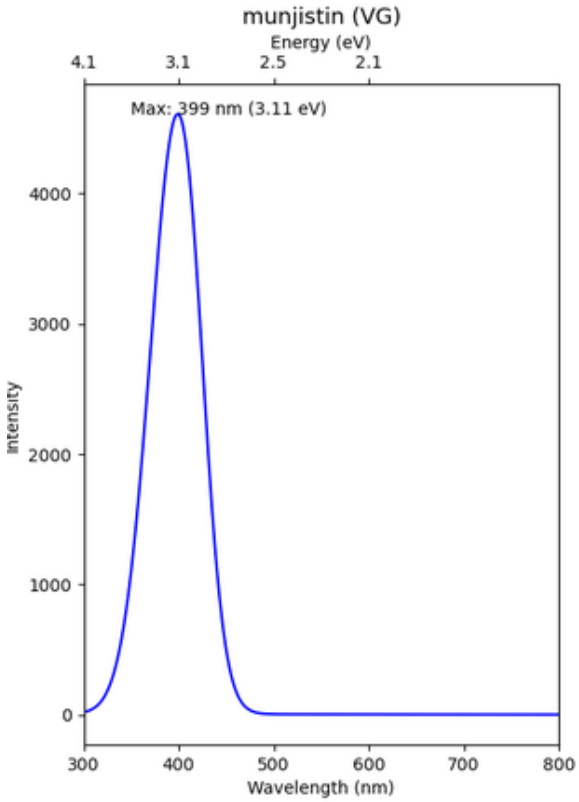


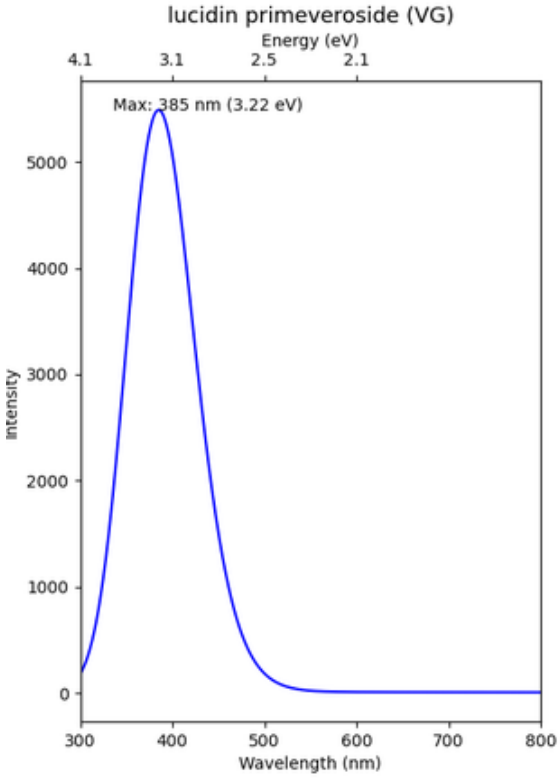


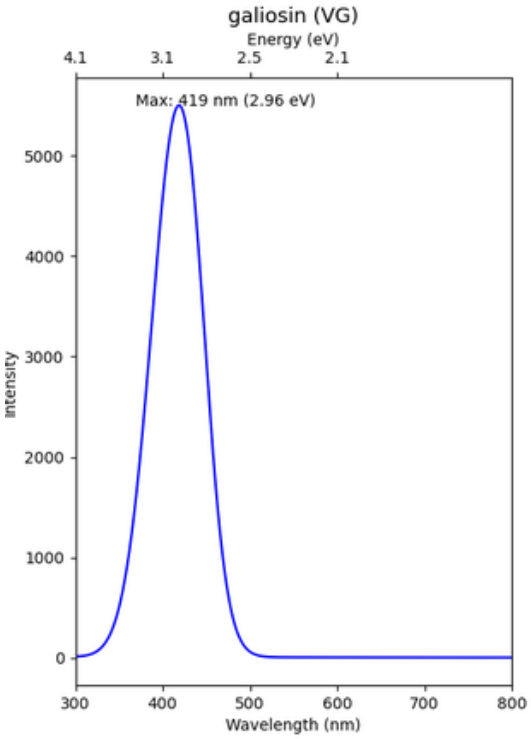



pseudopurpurin (VG)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 425 nm (2.92 eV)
8000
7000
6000
5000
4000
3000
2000
1000
0
Intensity
300
400
500
600
700
800
Wavelength (nm)

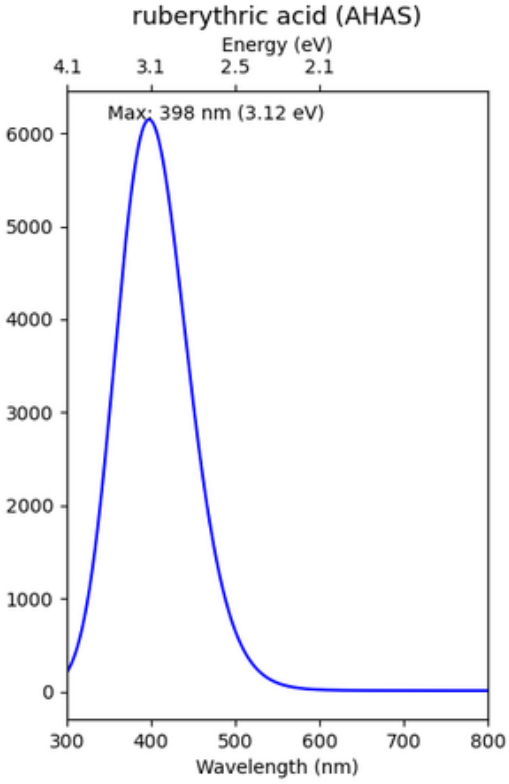

nordamcanthal (VG)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 409 nm (3.03 eV)
5000
4000
3000
2000
1000
0
Intensity
300
400
500
600
700
800
Wavelength (nm)

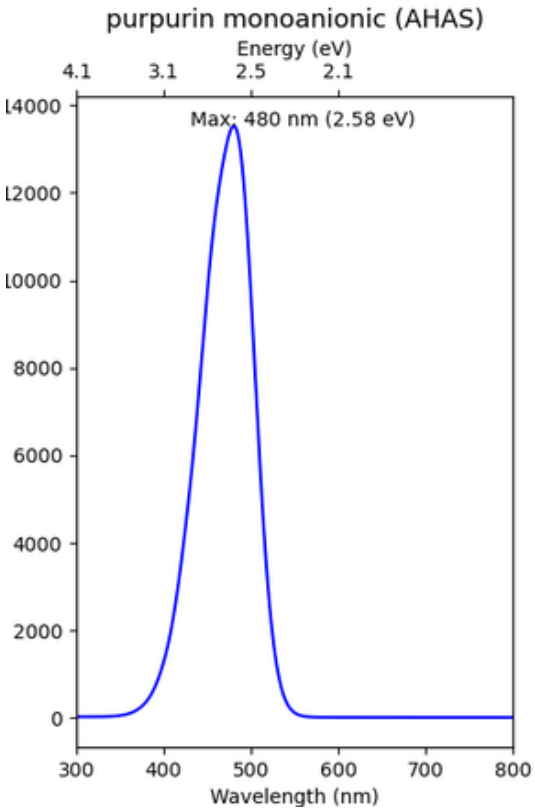

munjistin (VG)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 399 nm (3.11 eV)
4000
3000
2000
1000
0
Intensity
300
400
500
600
700
800
Wavelength (nm)

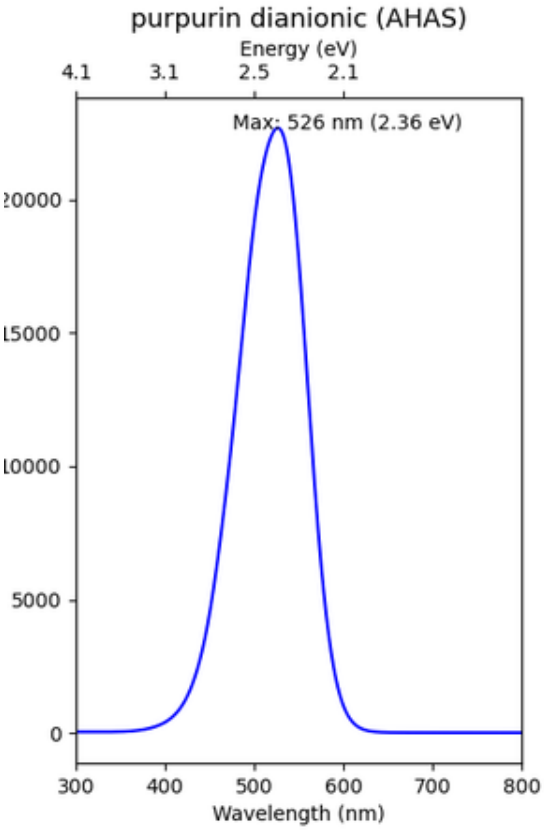

lucidin primeveroside (VG)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 385 nm (3.22 eV)
5000
4000
3000
2000
1000
0
Intensity
300
400
500
600
700
800
Wavelength (nm)

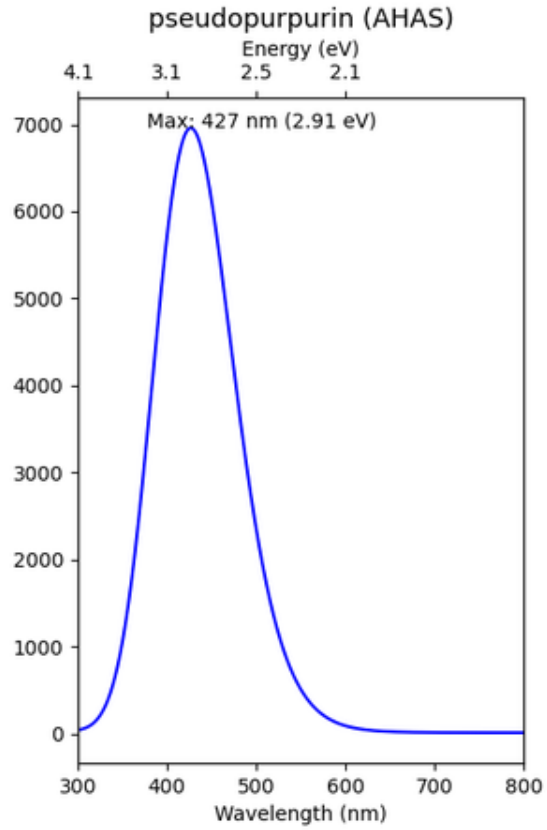


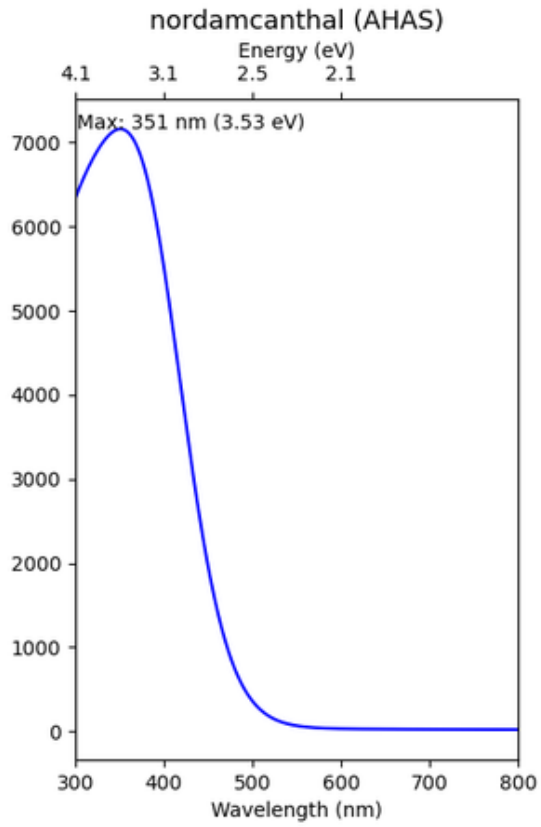
S.I. 12. *Computed individual spectra using Adiabatic Hessian After Step (AHAS)method.*

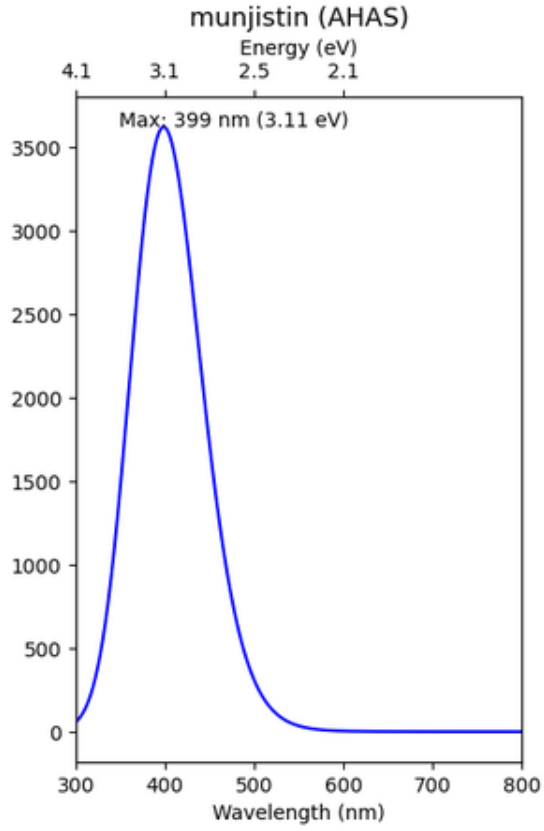


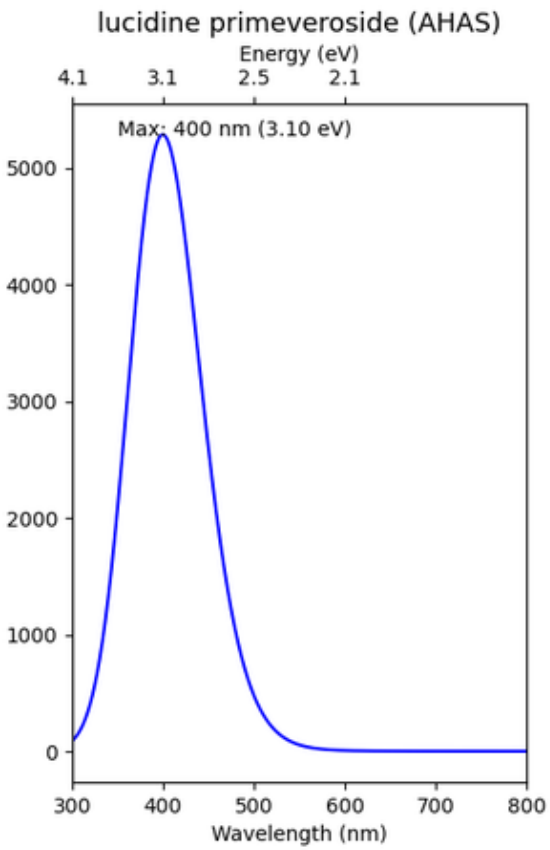


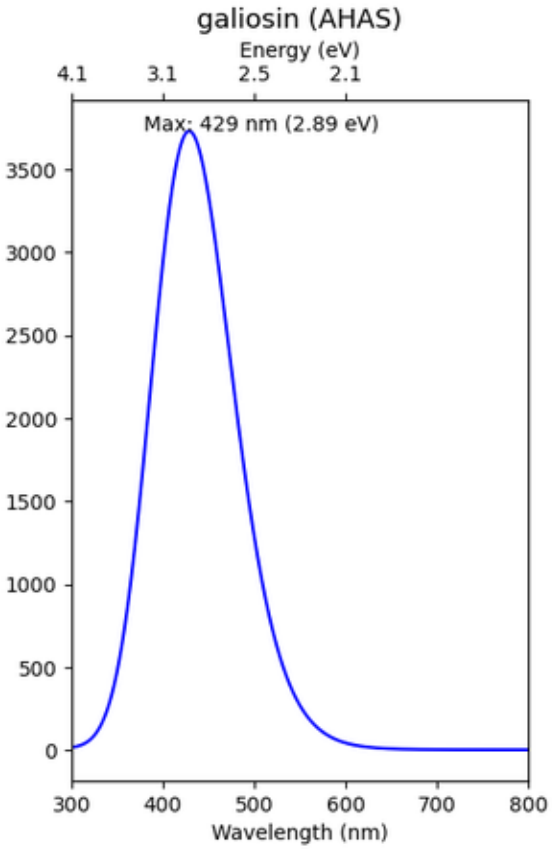

purpurin dianionic (AHAS)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 526 nm (2.36 eV)
20000
15000
10000
5000
0
300
400
500
600
700
800
Wavelength (nm)

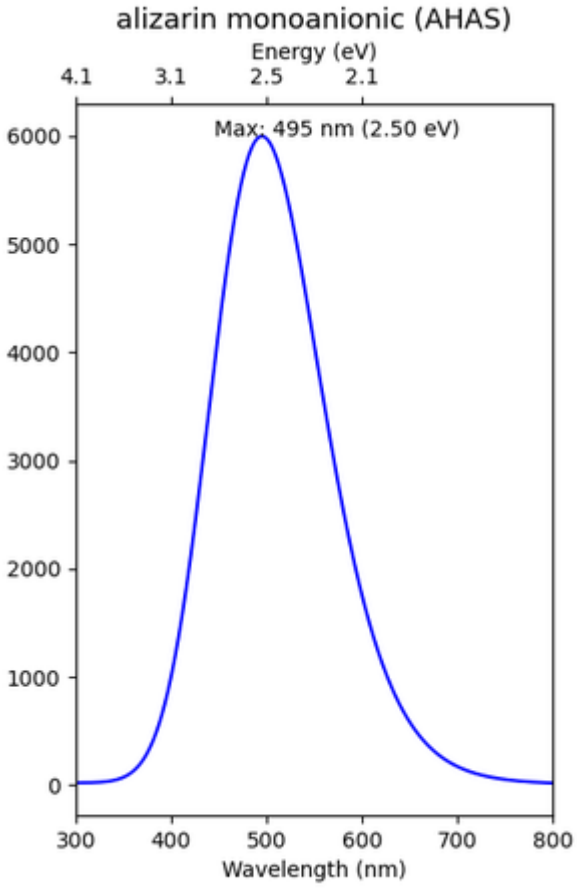

pseudopurpurin (AHAS)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 427 nm (2.91 eV)
7000
6000
5000
4000
3000
2000
1000
0
300
400
500
600
700
800
Wavelength (nm)


nordamcanthal (AHAS)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 351 nm (3.53 eV)
7000
6000
5000
4000
3000
2000
1000
0
300
400
500
600
700
800
Wavelength (nm)

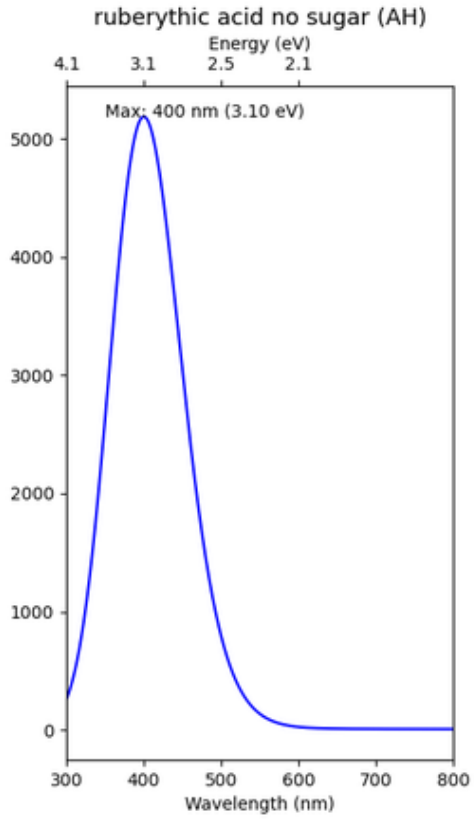

munjistin (AHAS)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 399 nm (3.11 eV)
3500
3000
2500
2000
1500
1000
500
0
300
400
500
600
700
800
Wavelength (nm)

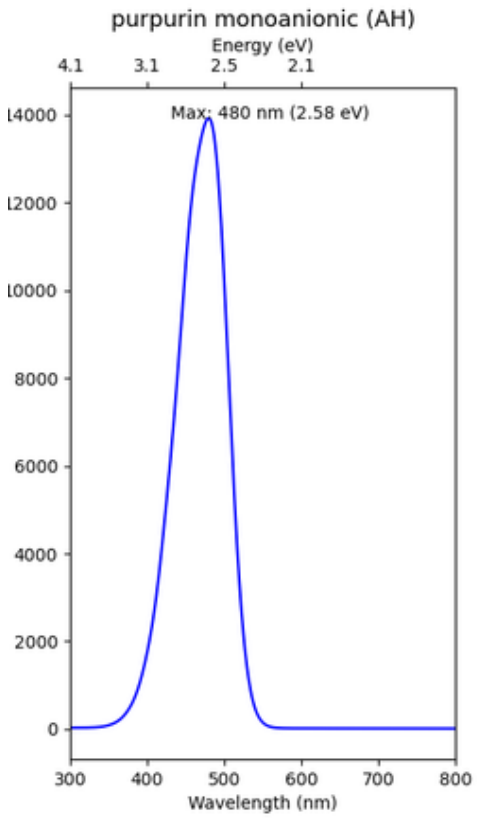

lucidine primeveroside (AHAS)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 400 nm (3.10 eV)
5000
4000
3000
2000
1000
0
300
400
500
600
700
800
Wavelength (nm)

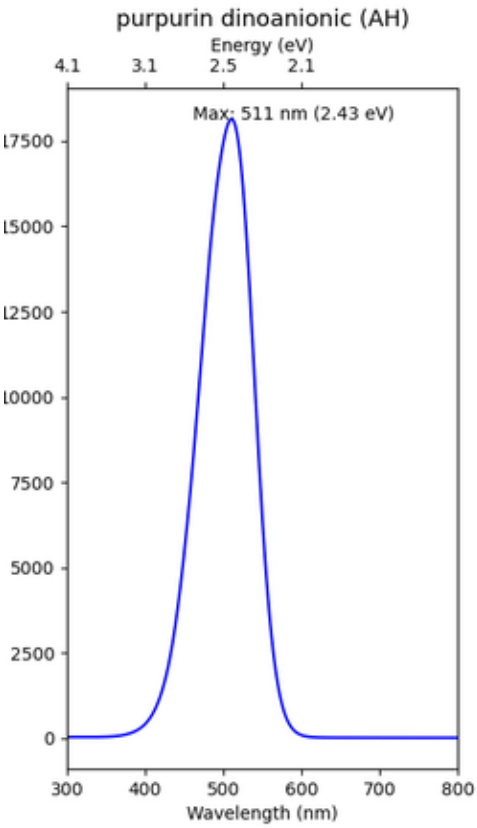

galiosin (AHAS)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 429 nm (2.89 eV)
3500
3000
2500
2000
1500
1000
500
0
300
400
500
600
700
800
Wavelength (nm)

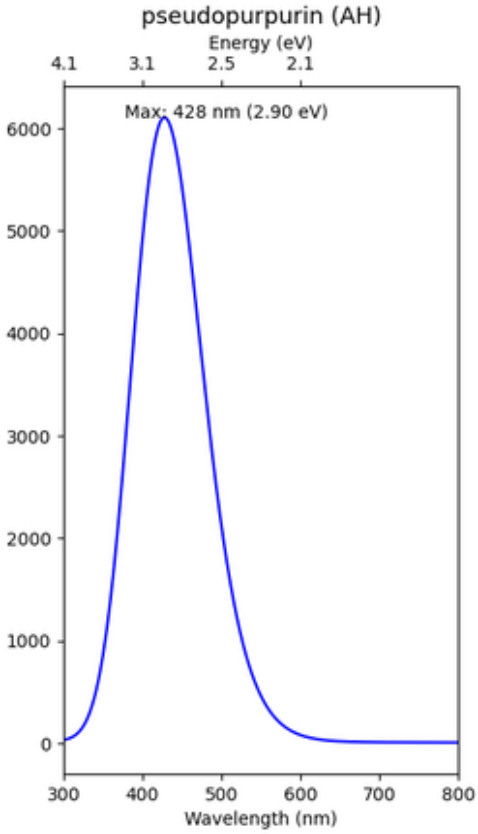

alizarin monoanionic (AHAS)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 495 nm (2.50 eV)
6000
5000
4000
3000
2000
1000
0
300
400
500
600
700
800
Wavelength (nm)

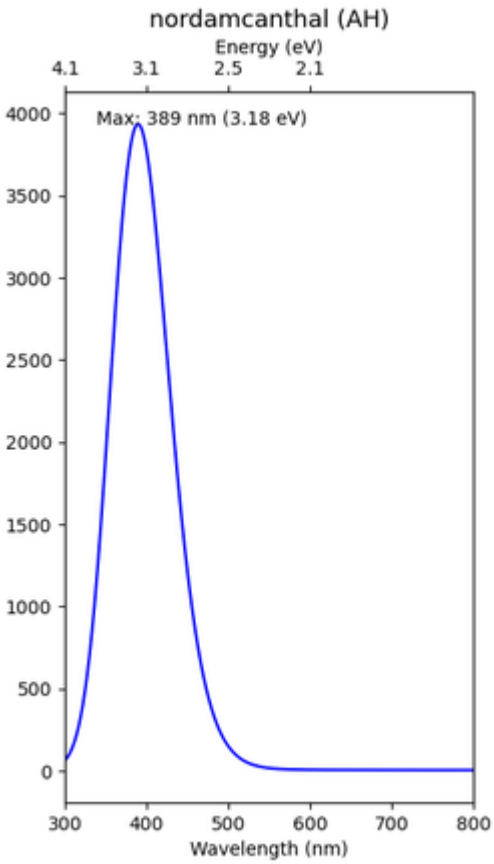
S.I. 13. *Computed individual spectra using Adiabatic Hessian (AH) method.*

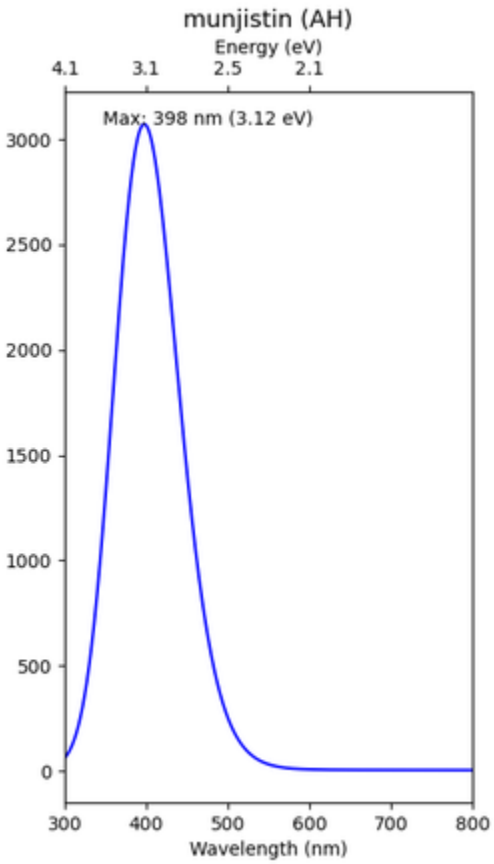


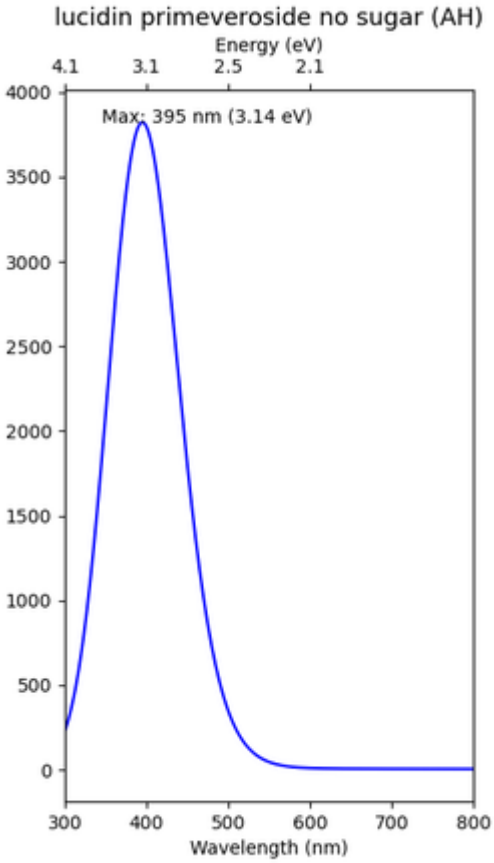


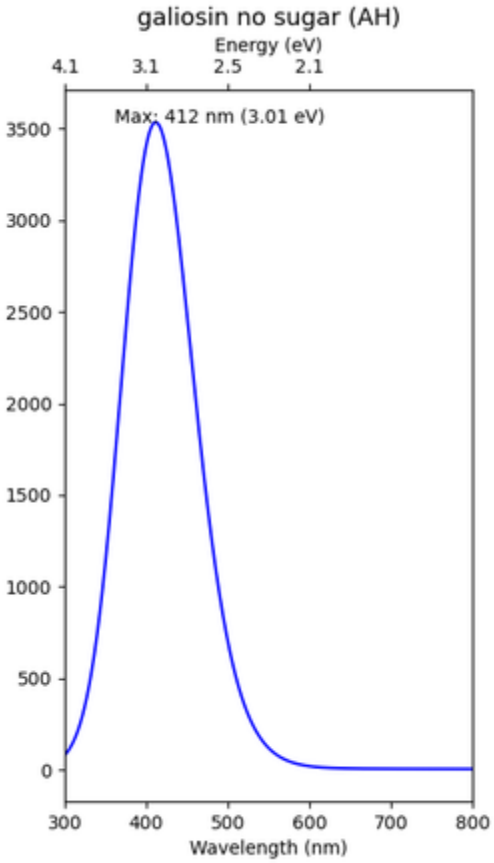


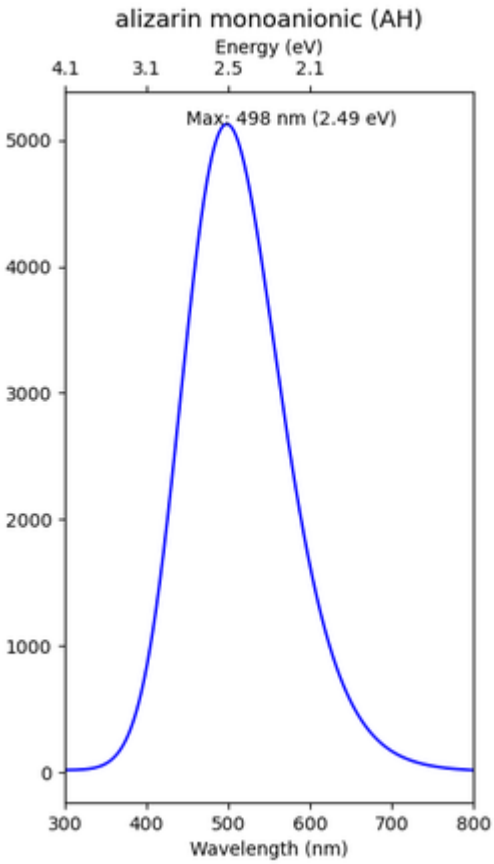



nordamcanthal (AH)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 389 nm (3.18 eV)
4000
3500
3000
2500
2000
1500
1000
500
0
300
400
500
600
700
800
Wavelength (nm)


munjistin (AH)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 398 nm (3.12 eV)
3000
2500
2000
1500
1000
500
0
300
400
500
600
700
800
Wavelength (nm)


lucidin primeveroside no sugar (AH)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 395 nm (3.14 eV)
4000
3500
3000
2500
2000
1500
1000
500
0
300
400
500
600
700
800
Wavelength (nm)


galiosin no sugar (AH)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 412 nm (3.01 eV)
3500
3000
2500
2000
1500
1000
500
0
300
400
500
600
700
800
Wavelength (nm)


alizarin monoanionic (AH)
Energy (eV)
4.1
3.1
2.5
2.1
Max: 498 nm (2.49 eV)
5000
4000
3000
2000
1000
0
300
400
500
600
700
800
Wavelength (nm)

SUMMARY

Article 4: "When light betrays colour: the photodegradation of alizarin, a historical dye, in basic media."

(Article 4 is currently under review by *Dyes and Pigments* at the time of thesis submission.)

With the spectroscopic properties and electronic transitions of the dyes elucidated, we proceeded to study photodegradation mechanisms using alizarin as a model. Photodegradation was accelerated under controlled conditions using UV-Visible absorbance monitoring. Though alizarin is commonly used on solid substrates, its use in aqueous media, where photofading is more pronounced. Studying in solution also allows control over pH, oxygen levels, and solvent effects, and facilitates the detection of transient species using techniques such as EPR and HPLC-UV-MS.

We first examined the influence of various solvents on alizarin photodegradation. Alizarin was found to be stable in glycerol and DMSO, but degraded rapidly in DMF, suggesting a radical-based photodegradation mechanism. As water is the solvent most frequently used in heritage contexts, we focused our study on aqueous systems using alizarin red S, a more water-soluble analogue, for kinetic analysis under different pH levels and light sources (xenon, UVA, and visible). We observed that degradation was significantly faster at basic pH. As alizarin is soluble at basic pH, we also studied its degradation under the same light sources.

A key focus was the underexplored basic form of alizarin (di-deprotonated), as literature mainly focuses on the red neutral form or its use in lake pigments. Electron Paramagnetic Resonance (EPR) analysis revealed oxygen-centered radicals at positions 1 and 2, supported by computed g-factors and hyperfine coupling constants. HPLC-UV-MS analysis identified photodegradation products after UVA exposure over time. Some products were newly identified (e.g., 3-hydroxy-1,4-dioxo-1,4-dihydronaphthalene-2-carboxylic acid), while others, like phthalic acid, previously observed in methanol, were found in aqueous basic conditions at different relative concentrations, suggesting a distinct degradation pathway at high pH. Several mechanisms were proposed based on EPR data and identified products. This work offers new insights into the influence of pH, solvents, light, and oxygen on dye stability, knowledge that is also relevant to industrial applications such as textile dyeing and alkaline bath extraction. These findings have been submitted in the fourth article.

ARTICLE 4

When light betrays colour: the photodegradation of alizarin, a historical dye, in basic media

Thanh Huyen Linh TRAN [a], Fabrice MORLET-SAVARY[c], Jacques LALEVEE[c], Romain BERRAUD-PACHE [*a]and Maguy JABER [a,d]

Laboratoire d'Archéologie Moléculaire et Structurale (LAMS), Sorbonne Université, CNRS, UMR 8220, Paris, 75005, France
Institut de Chimie des Substances Naturelles (ICSN)
Institut de Science des Matériaux de Mulhouse (IS2M)
Institut Universitaire de France (IUF), Paris, France

(Currently under review in *Dyes and Pigments*)

ABSTRACT

Archaeological and historical evidence indicates that natural dyes have been employed by human societies for millennia to achieve colouration in various materials. And light reveals this world of colours by painting everything it touches. However, there is much evidence that natural dyes photodegrade when exposed to light, fading cultural heritage work. Upon photon absorption, a molecule reaches an excited state, which promotes photochemical reactions leading to the formation of new compounds and thus the photodegradation. Understanding the underlying mechanisms of these processes helps elaborate effective strategies to mitigate them. These problematics have long been relevant across various fields. There have been many studies about the photodegradation of dyes, and this continues until today.

Our study focuses on alizarin, a historical natural dye that can be extracted from madder roots. It has been used in many fields, textiles, paintings, since antiquity. A distinctive aspect of our study is the investigation of alizarin's less-studied basic (di-deprotonated) form, using a solution model under controlled conditions. By using techniques such as EPR and HPLC-UV-MS, we identified radical oxygen intermediates and degradation products while evaluating the influence of pH, solvent and light wavelength. These findings have important implications for preserving alizarin-based artworks, improving industrial dyeing practices, and informing environmental efforts like wastewater treatment and shedding light on computational modelling as a crucial research field in these topics. Overall, the study deepens our understanding of anthraquinone dye degradation and promotes the sustainable use of natural dyes.

1. INTRODUCTION

Natural dyes are colourants obtained from renewable sources such as plants, invertebrates, and minerals. Before the invention of synthetic dye, across many cultures, natural dyes have been utilised in colouring textiles, food, and other items spans several centuries. In recent years, there has been a renewed interest in natural dyes due to their eco-friendliness, ready availability, non-toxicity, and sustainability. [249–251] In this study, we

focus on madder, and more specifically, alizarin, one of its major colourant compounds. This molecule from the hydroxyanthraquinone family shows distinct colours determined by the water pH, yellow (pH<5), red (5<pH<10) or purple (pH>10). The roots of madder plants such as *Rubia tinctorum* L., *Rubia cordifolia* L. and other species of madder have been widely applied in paintings, textiles, manuscripts, tapestries, and other culturally significant artworks, mostly using the red hue. Like other natural dyes, madder exemplifies natural and historical dyes that are sensitive to light-induced fading. [74] Famous examples of madder fading include the "*Bayeux Tapestry*", Winslow Homer's watercolour "*For to be a Farmer's Boy*", Japanese woodcuts, Van Gogh's "*Head of an Old Woman*", and many others. [6,7,64,94,252–254]. Its degradation poses serious challenges to the conservation of artworks, as the fading alters their visual appearance, historical interpretation, and economic value, while also complicating pigment identification. [40] This problem has engaged chemists, physicists, and biologists in ongoing research efforts. As early as 1888, pioneering work by Russell and Abney first reported the effects of photodegradation on paintings displayed in English museums, sparking growing scientific interest in this phenomenon. [255] Since then, numerous studies have explored the various factors influencing alizarin photodegradation, including the nature of the substrate, the presence of oxygen, mordants, and pH conditions. [68,74,101,256–259]

Most previous studies have focused on the red form of alizarin, often as a lake pigment due to its poor solubility in water.[46] As a result, the properties and photodegradation of alizarin in solution remain largely unexplored. A previous study by Jiang *et al.* studied the photodegradation of alizarin in methanol solution (yellow form) under xenon light. They proposed a mechanism where alizarin, upon photoexcitation, generates charge carriers (e^-, h^+), which react with O_2 to produce hydroxyl radicals (•OH) that drive oxidative degradation and have identified several products such as M/z=180 ($C_9O_4H_8$), M/z=170 ($C_7O_5H_6$), M/z=166 ($C_8O_4H_6$), M/z=142 ($C_6O_4H_6$), and M/z=208 ($C_{10}O_5H_8$) using mass spectroscopy. [74] As such, lightfastness refers to the ability of a dye to resist fading when exposed to light, and it is now evaluated using standardised methods. [260,261] Another study by Kristin L *et al.* have studied the photobleaching of alizarin and purpurin. [258] They found that alizarin and purpurin exhibit distinct blinking and photobleaching behaviours linked to their excited-state dynamics. Alizarin shows about four times more emissive events than purpurin due to excited-state intramolecular proton transfer (ESIPT), which promotes a fluorescent tautomer state and reduces electron transfer (ET) leading to photobleaching.

However, several critical aspects remain unclear or poorly understood, particularly the reaction mechanisms occurring in water and the nature of the degradation products formed. Although the violet form is unstable, it is less studied. [29,74,101] These conditions are not uncommon in industrial and historical contexts, for instance, during pigment or madder lake preparation, dyeing processes, or the extraction of dyes in alkaline baths. [46,55,262–264] For example, several studies have proposed mild extraction methods for madder dyes in basic media, often employing ammonia to preserve most of the compounds. [55,100]. In madder lake preparation, one key step involves raising the pH of the dye solution with sodium hydroxide or ammonia to deprotonate the dyes, thereby enhancing their ability to complex with metal cations such as aluminium from alum. [46]

This study aims to address these gaps by investigating the photodegradation of alizarin in different media, identifying the underlying reaction mechanisms and resulting products, and understanding how pH and solvent influence the rate and pathway of degradation. We will first describe how different solvents impact the photodegradation mechanism of alizarin. Then, using an analogue of alizarin which is more soluble in water, we will measure how various wavelengths influence the photodegradation. Finally, we will focus on the purple form in basic media and try to characterise the products of photodegradation and propose a mechanism.

2. MATERIALS AND METHODS

Accelerated photodegradation

The natural dyes can take a long time to fade under light exposure (from days to years). Therefore, we used accelerated photodegradation by using 3 different artificial lamps with different spectra: xenon lamp, UVA, or visible light (spectral distributions provided in the Supporting Information S.I. 1 & S.I. 2). The Xenon lamp is SUNTEST XLS+ with Daylight option (without filter), the unit was operated with constant cooling, and the intensity of the lamp is 50 mW/cm^2, the wavelengths spans from 300nm to 800 nm. Visible lamp is a polychromatic white LED TSPOT4-WHI-9-DS from TPL Vision (measured irradiance of about 10 mW/cm^2). Finally, the UVA lamp is a UV-LED L11921-410 from Hamamatsu Photonics with a maximal emission at 369 nm and a measured irradiance of about 10 mW/cm^2. The exposition was followed by UV-VIS spectra every hour for 5 hours. A 5-hour exposure at 50,000 lux is equivalent to approximately 0.43 years (about 5 months) of natural ageing under typical museum lighting at 200 lux (details in S.I.).

Alizarin photodegradation (Sigma Aldrich, 97%) was done in water and at pH=12 because it is the most soluble at this pH at 10^{-4} M. To study the influence of photodegradation kinetics on pH, we used an alizarin alternative, alizarin red S (sodium alizarinsulfonate, Thermoficher), for its higher water-solubility at different pH (2, 6.5, and 12) adjusted with NH_4OH or HCl and with a concentration of 10^{-4}M. A triplicate of each kinetic experiment was done and the error deviation was calculated. The influence of solvent was also tested for alizarin in toluene, acetone, ethanol, DMSO, DMF, 2-propanol and glycerol (solvent grade). The number of photons per cm^2 in function of irradiance is shown on Table 1. The computation details are in S.I.

Table 1. Photons absorbed in function of irradiance and wavelength.

Wavelength (nm)	Irradiance	$N_{photons}/cm^2$ 1h exposition
520	50mW/cm^2 (xenon)	4.72×10^{20} photons/cm^2
520	10mW/cm^2 (LED)	9.40×10^{19} photons/cm^2
430	50mW/cm^2 (xenon)	7.80×10^{17} photons/cm^2
430	10mW/cm^2 (LED)	1.56×10^{17} photons/cm^2

HPLC-UV-MS

Solutions of alizarin (Sigma Aldrich, 97%) were prepared with a concentration of 10^{-2} M, adjusted at pH=12 with NH_4OH. The solutions were irradiated using UVA light from a UV-LED source previously described. This specific lamp was selected to ensure consistency with parallel EPR experiments and due to the availability of compatible equipment. Liquid chromatography-UV-mass spectrometry (LC-UV-MS) experiments were performed on a 1290 Infinity II LC system coupled to a Q-TOF 6540 mass spectrometer equipped with an ESI source operated in the negative and positive ionization modes. The chromatographic separation was carried out on a Zorbax 300SB-C8 (150 × 2.1 mm, 5 µm) column heated at 35 °C at a flow rate of 300 µl/min with water at 0.1% of formic acid as solvent A and acetonitrile as solvent B. Five microliters of samples in H_2O + NH_4OH at pH12/H_2O were injected using the following gradient: 0 min (5% B), 20 min (40% B), 22 min (100% B), 25 min (100% B), 26 min (5% B)and 6 min of equilibration time. Ion source conditions were set for negative ion mode: capillary voltage 3.5 kV, Nozzle voltage 1250 V, Fragmentor voltage 125 V, Skimmer 65 V, Drying gas at 10 l/min, 275 °C, Nebulizer 40 psi, and sheath gas flow at 10 l/min, and 325 °C. UV detection was performed at 260 and 410 nm and the quantitation at 260 nm. The TOF was operated in the 40−1700 *m/z* mass range, Data Acquisition 9.0 and Qualitative Analysis 8.0 software were used to process data.

UV-VIS spectra

Their UV-VIS spectra were measured immediately using an Ocean Optics Flame spectrometer (Model: FLMS00699) with a wavelength range from 300 to 850 nm. The integration time for each measurement ranged from 5 ms to 10 ms. The light source used was a Deuterium-Tungsten Halogen lamp covering the range of 300 nm–850 nm. The blank was done using deionized water. This light source was coupled to the spectrometer using 400 µm diameter optical fibers and positioned through a 1 cm quartz cuvette holder. The acquisition of spectra was facilitated by the Ocean View software.

EPR

Alizarin was dissolved in water at pH=12 adjusted using NH_4OH. The formation of paramagnetic intermediates upon irradiation of alizarin was monitored by cw-EPR spectroscopy using an EMX Plus X-band EPR spectrometer (Bruker, Germany). The samples were irradiated at 293 K directly in the EPR resonator (High Sensitivity Probe-head, Bruker), and the EPR spectra were recorded in situ using a capillary in an EPR tube. The irradiation source was a UV lamp from Hamamatsu with an interferential filter at 365 nm and an intensity of 200 mW/cm^2. The acquisition parameters are: receiver gain set to $1.0.10^5$, the sweep width is 25G (central field: 3488 G), the modulation amplitude is 3 G, the time constant is 5.12 ms, the resolution is 1024 pts, the power is 6.325 mW (15dB). 150 scans were performed with an acquisition time of 1500s.

3. DISCUSSION AND RESULTS

The following sections present a stepwise investigation starting with the impact of the solvent on the photodegradation and its mechanism. We move then to the photofading of

an analogue of alizarin, which is more soluble in water, under various irradiation lamps before focusing on the purple form of the native alizarin, exploring its major degradation products via HPLC-UV-MS and probing the involvement of reactive intermediates using EPR spectroscopy.

3.1. Solvent Effect on Alizarin Photodegradation under Xenon Irradiation

In the following sections, we investigate the impact of the solvent on the photodegradation of alizarin.

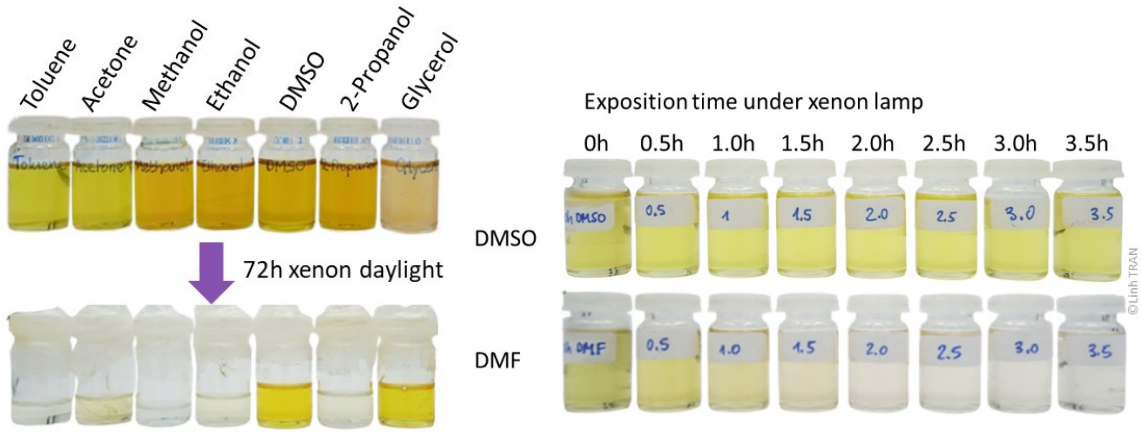


Figure 1. (Left)Accelerated photodegradation after 72h under xenon light (irradiance =50mW/cm^2) of alizarin in different solvents: toluene, acetone, methanol, ethanol, DMSO, 2-propanol, glycerol. (Right) Alizarin in DMSO (radical scavenger) and DMF (which promotes radical formation) under different exposure times under a xenon lamp.

To assess the influence of solvent environment on the photostability of alizarin, we conducted a comparative study using a range of solvents varying in polarity, hydrogen-bonding capacity, and chemical reactivity. Alizarin was dissolved in toluene, acetone, ethanol, 2-propanol, glycerol, dimethyl sulfoxide (DMSO), and each solution was exposed to xenon light for 72 hours. The results are presented in Figure 1 (and S.I. 3) and reveal a clear solvent-dependent behaviour. In all solvents except DMSO and glycerol, photodegradation was observed. Notably, the colour fading was much slower in DMSO and glycerol, even after extended light exposure. This resistance to photodegradation is attributed to their well-known role as radical scavengers, suggesting that free radical pathways, likely involving reactive oxygen species (ROS) or dye-centred radicals, play a major role in the degradation mechanism of alizarin. [265] In contrast, alizarin in DMF exhibited significantly faster degradation rates, indicating that certain solvent environments may facilitate excited-state reactivity or stabilise reactive intermediates. This may arise from solvent-induced stabilisation of reactive excited states or facilitation of electron transfer processes. This behaviour has already been observed in other dyes. [92] Indeed, literature supports the role of DMSO as a photoprotective agent due to its radical scavenging properties, which likely contribute to the suppression of degradation in certain solvents. For example, studies by A. F. Asker *et al.* show that Dimethyl sulfoxide (DMSO) was shown to enhance the photostability of FD&C Red No. 3 dye solutions under UV and fluorescent light, with greater protection observed at higher DMSO concentrations. While pH had no significant effect in the presence of DMSO. Other studies

by S. Bansal *et al.* discuss the high reactivity of DMSO towards reactive oxygen species (ROS), providing a mechanistic basis for the observed solvent-dependent photostability. [266] The study also shows that DMSO significantly suppresses the fluorescence response of the ROS probe DCFH to hypochlorite, likely by scavenging reactive species, whereas DMF allows strong fluorescence activation.

These findings support the hypothesis that photodegradation of alizarin is mediated by reactive radicalar intermediates, and that the solvent can either suppress or promote these pathways depending on its chemical properties.

In cultural heritage studies, water is the most commonly used solvent. However, alizarin has limited solubility in water, which complicates its homogeneous photodegradation analysis in aqueous environments. To address this, we began our investigation using alizarin red S (ARS), a more water-soluble analogue of alizarin. Based on our previous article, the difference of electronic densities during the ground state to first excited state transitions show minimal differences between alizarin and its analogue ARS, suggesting that the sulfonate group does not significantly alter the π-conjugated system involved in degradation.[234] Therefore, the following section focuses on the photodegradation by different lamps of ARS in aqueous media, at different pH.

3.2. Photodegradation in aqueous solution at different pH

Influence of pH

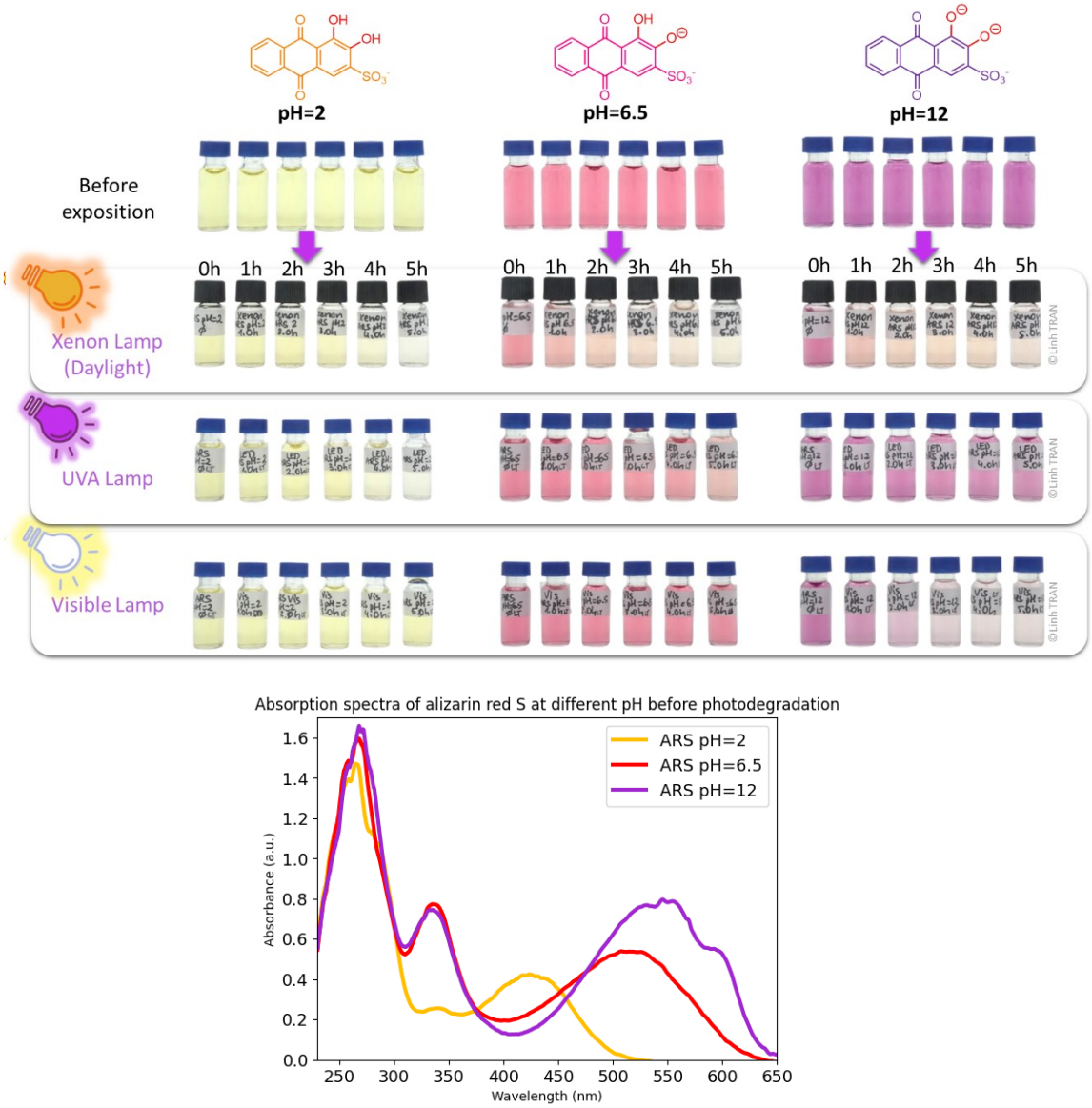


Figure 2. (Top) Photodegradation of alizarin red S (ARS) under different lamps (visible with irradiance= 10mW/cm², UVA with irradiance= 10mW/cm², and xenon daylight lamp with irradiance= 50mW/cm²) and different pH (pH=2, 6.5, and 12). The exposure time varies from 0h to 5h, and a photo is taken every hour. (Bottom) Absorption spectra of ARS under different pH in water, before photodegradation.

Aqueous solutions of alizarin red S (ARS) were exposed to xenon (which closely mimics solar radiation, 300–800 nm), UVA (365 nm), and visible light (400–800 nm) sources (spectral distributions provided in the Supporting Information) at pH values of 2, 6.5, and 12. Photodegradation was monitored over time using UV-VIS spectroscopy by tracking the decrease in the dye's characteristic absorbance peaks at 423 nm for pH=2, at 517 nm for pH=6.5, and at 538-550 nm for pH=12. The results are shown in Figure 2 (details of spectra and kinetics can be found in S.I. 4 to S.I. 13). Under all irradiation conditions and pH, a progressive loss of colour intensity was observed, confirming that ARS undergoes photodegradation in aqueous environments. The absorbance spectra of (ARS) reveal some trends in photodegradation behaviour, influenced by both pH and light wavelength and intensity. Photodegradation was significantly faster under alkaline conditions compared to neutral or acidic environments. This suggests that the di-deprotonated form

of ARS is more photoreactive and may follow a different photodegradation pathway than its protonated form.

At **pH=2**, ARS shows the least photodegradation under xenon and visible light source, so this protonated form is more stable. The absorbance spectra remain relatively stable over time under all kinds of lights, indicating that acidic conditions significantly slow photodegradation. This is likely due to the protonated form of ARS being more photostable or less reactive with photogenerated reactive oxygen species. This is consistent with previous studies in which photobleaching in solution is believed to result from the formation of a tautomeric photoactive triplet state that reacts with oxygen. This can also be explained by the possible Excited-State intramolecular Proton and Electron Transfer (ESIPT) in alizarin when protonated in an acidic solution. [74,258,267]

At **pH=6.5**, a faster photodegradation is observed than at pH=2, especially under xenon light. The absorbance in the UV-VIS region (especially around 250–350 nm and around 510-525 nm) steadily decreases over time, showing the breakdown of ARS chromophores. The visible light source causes only minor changes in absorbance, suggesting that visible light alone has limited effectiveness in inducing degradation at neutral pH.

At **pH=12**, photodegradation is most pronounced under xenon and visible light. The xenon lamp leads to the fastest decrease in absorbance compared to other pH, especially at the main absorption bands (around 250–350 nm and 500–550 nm), suggesting that alkaline conditions favour photodegradation, possibly due to increased formation of reactive anionic species or enhanced susceptibility to photooxidation. UVA light also causes significant spectral changes, though less than xenon. Even visible light, which was relatively ineffective at lower pH, results in noticeable degradation at pH=12, highlighting the increased sensitivity of ARS to light under basic conditions.

At all pH levels, the peak in the visible region of the UV-VIS spectrum is the first to diminish (See S.I. 4). Our previous work has shown that this peak arises from $\pi \rightarrow \pi^*$ transitions. This suggests that the photodegradation process alters the conjugation of ARS first and changes the hydroxy or ketone groups. The peak at 250-350nm also corresponds to a $\pi \rightarrow \pi^*$ transition and decreases later.

Kinetic study

The Figures **S.I. 4-S.I. 13** in S.I. illustrate the kinetic study of photodegradation of ARS under different light sources and across three pH conditions (pH = 2, 6.5, and 12). The relative concentration, $\log(C_t/C_0)$ and $1/C_t$, is plotted with time to assess the order of the kinetic photodegradation reaction under different conditions.

Under Xenon Light

At pH = 2, some degradation is observed, primarily after 150 minutes of irradiation. The degradation curve is initially flat, suggesting an induction period or low early reactivity. The reaction follows a pseudo-zero-order kinetic behaviour. At pH = 6.5, the photodegradation also follows a pseudo-zero-order kinetic pattern. Degradation becomes more pronounced after 150 minutes, though it does not fit neatly into a single kinetic

model throughout the experiment. At pH = 12, a clear and quick decrease in ARS concentration is observed. The $\log(C_t/C_0)$ plot is nearly linear, indicating a pseudo-first-order behaviour, which is further supported by the increasing trend in $1/C_t$.

Under UVA Light

At pH = 2, the degradation shows a delayed onset but then proceeds with a near-linear $\log(C_t/C_0)$ decline, pointing to a delayed pseudo-first-order behaviour. At pH = 6.5, the degradation follows a pseudo-second-order kinetic model. As with xenon, degradation becomes more significant after 150 minutes, but doesn't align perfectly with a single model throughout. At pH = 12, the degradation is strong and consistent, with a near-linear $\log(C_t/C_0)$ curve indicating pseudo-first-order kinetics. This is reinforced by the steady increase in $1/C_t$ over time.

Under Visible Light

At pH = 2, there is almost no photodegradation observed under visible light. At pH = 6.5, the degradation follows a pseudo-first-order kinetic behaviour. However, the reaction lacks consistency across the entire duration, hinting at complex kinetics. At pH = 12, there is a clear, linear decline in $\log(C_t/C_0)$, supporting pseudo-first-order kinetics. The progression of $1/C_t$ also affirms this behaviour.

Photodegradation efficiency of ARS increases with pH. This trend highlights the significance of both the molecular form of ARS (pH-dependent) and the energy/wavelength of the light source in dictating the degradation efficiency and pathway. These findings emphasise the need to consider environmental conditions when evaluating the lightfastness of water-soluble anthraquinone dyes. However, due to numerous external variables (e.g., ionic composition, impurities, and batch-to-batch variability), definitive mechanistic conclusions regarding the photodegradation of alizarin cannot be drawn from these observations alone, nonetheless, the results provide valuable insight into general degradation trends within this family of dyes.

3.3. Alkaline Conditions for Alizarin

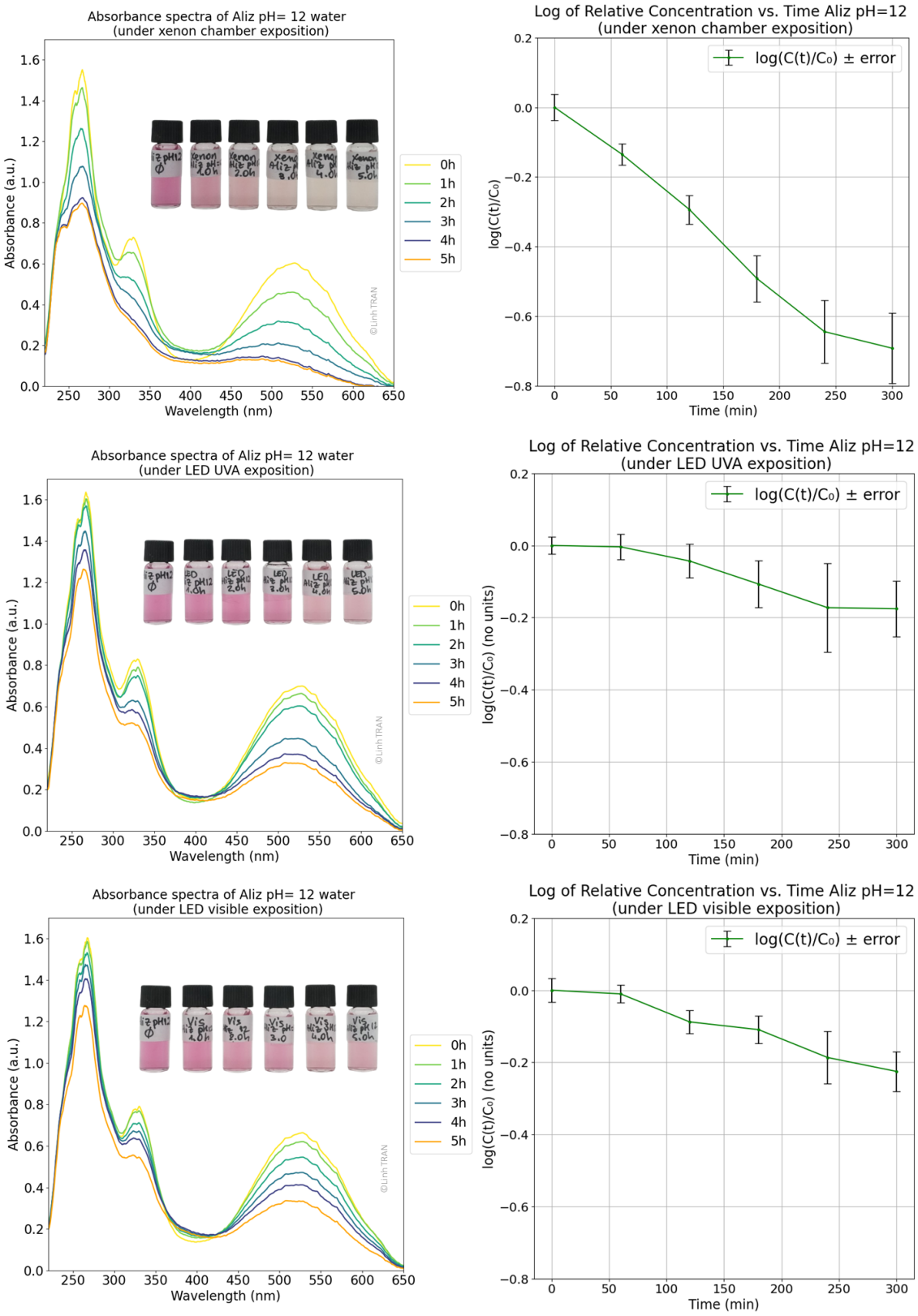


Figure 3. Photodegradation kinetics of alizarin at pH=12 under different light sources and for an increasing exposure time from 0h to 5h. The Absorbance spectra (in a.u.) are shown along with the log(C_t/C_0) to see the kinetic order.

Figure 3 presents the results of the photodegradation kinetic study of alizarin at pH=12 under different light sources (xenon, UVA, and visible lamps). It reveals significant differences in degradation efficiency and reaction kinetics. Under the xenon lamp, which simulates daylight, the relative concentration of alizarin decreases sharply over time, showing a linear trend in the $\log(C_t/C_0)$ plot, indicating pseudo-first-order kinetics. The corresponding $1/C_t$ (See S.I. 14) plot does not show linearity, supporting that the reaction follows first-order rather than second-order kinetics. Under UVA light, degradation occurs more slowly, with a gentler slope in both the relative concentration and $\log(C_t/C_0)$ plots, and a less pronounced change in the inverse concentration, suggesting lower photodegradation efficiency and weaker kinetics. The visible lamp shows even slower degradation, with only a modest decrease in alizarin concentration over time and minimal slope in all plots, indicating minimal photodegradation under visible light. These results suggest that the xenon lamp, with its broader spectrum including UV and visible light, is most effective for the photodegradation of alizarin, followed by visible light, with UVA being the least effective. Some trends are similar to the behaviour of ARS shown previously: about 80% degradation for both alizarin and alizarin red under xenon light, 35% under UVA light but the kinetic of photodegradation of alizarin red S is faster than alizarin. But under visible light, alizarin red S degrades much faster (90%) than alizarin (40%). This highlights the influence of light wavelength and intensity and the nature of the dye in driving photodegradation reactions.

3.4. Products of photodegradation

Alizarin photodegradation in basic aqueous ammonia was studied by HPLC-UV-MS/MS under varying UVA exposure. A UVA lamp was used due to equipment availability, ensuring consistency with RPE analysis, as the use of xenon lamp was technically unfeasible. Results are summarized in Figure 4, which includes a table of relative concentrations for six major identified compounds, along with their proposed molecular structures. Detailed LC-UV and MS/MS spectra are provided in the Supporting Information (S.I.)

	Exposition time	T = 0h	T = 1h	T = 3h	T = 6h
Compounds relative percentage (%)	P1	0.0	0.1	0.4	2.0
	P2	0.0	0.8	2.9	17.2
	P3	0.0	0.1	0.2	0.8
	P4	1.0	1.3	1.6	2.0
	P5	0.6	0.3	0.1	0.1
	P6	98.3	96.7	92.5	69.9
	P7	0.1	0.7	2.3	8

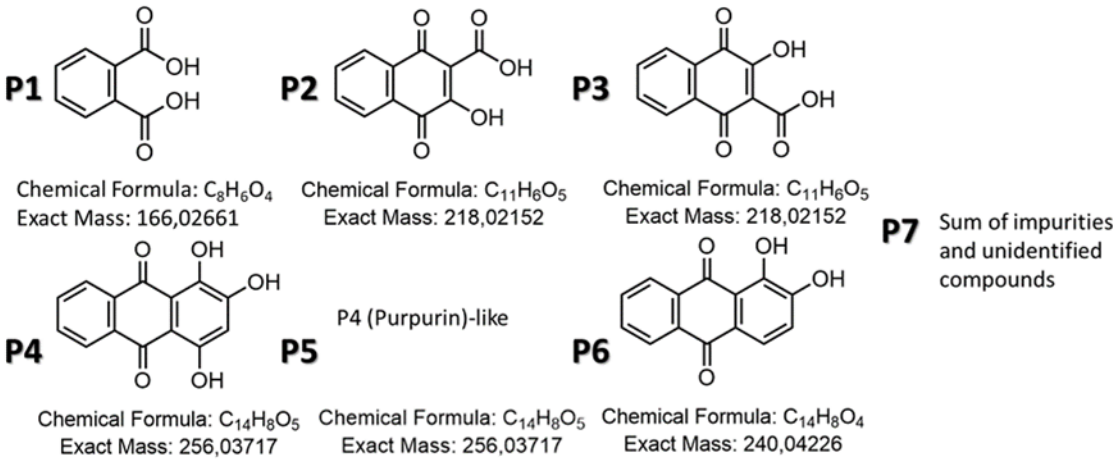


Figure 4. (Top) Integrated LC-UV chromatograms (in relative %) at 260 nm of alizarin irradiated at 0, 1, 3, or 6h in water at pH=12 adjusted with NH4OH. (Bottom) The identified photodegradation products of alizarin under basic conditions with ammonia, and their relative concentration under different exposure time with UVA light. The exact mass are in g/mol.

Initially, alizarin (product P6) accounts for 98.3% of the solution, decreasing to 69.9% after 6 hours of UVA exposure, confirming it as the primary source of degradation products. This aligns with the previous findings, where alizarin absorption intensity decreases to 65% at pH=12 under 5h irradiation under the same UVA light (S.I. 15). The products P4 and P5, identified as purpurin and purpurin-like compounds, respectively, are present as impurities in the initial solution, detectable even before UVA exposure (0 h). An additional group of unidentified minor compounds (P7) constitutes approximately 8% of the solution after 6 hours.

The product P1, identified as a phthalic acid, has been reported in previous studies under xenon light in methanol (neutral conditions) or with alizarin red S in aqueous solution with TiO_2 and visible light. [74,259] However, differences in solvent, pH, and light source compared to this study may lead to an alternative degradation pathway, as the percentages of degraded products are different. In this study, it is also observed under UVA light in basic aqueous conditions, forming slowly and reaching 2% concentration after 6 hours of exposure. The product P2, 3-hydroxy-1,4-dioxo-1,4-dihydronaphthalene-2-carboxylic acid, is the main photodegradation product of alizarin in basic aqueous solution. Its concentration increases from 0.8% at 1 hour to 17.2% after 6 hours. To our knowledge, this compound has not been previously reported in the context of alizarin degradation. The product P3 is an isomer of P2, but forms at a significantly lower level (0.8% after 6 hours), suggesting a regioselective or stereospecific degradation pathway favouring P2 over P3.

3.5. Towards the mechanism: EPR spectra

Light energy can trigger both physical and chemical changes in most organic compounds. Colourants interact with light in diverse ways. When a molecule absorbs a photon and becomes excited, it can release the excess energy through several mechanisms. These include physical processes that return the molecule to its ground state, or photochemical reactions that lead to molecular breakdown and the formation of new compounds.

Photochemical reactions can generally be divided into two main types, depending on which molecule absorbs the photon, as illustrated in Figure 5. [73,268]

(1) Direct photochemical reactions, where the dye molecule absorbs a photon, enters an excited state (dye*), and subsequently undergoes intra- or intermolecular reactions
(2) Indirect photochemical reactions, where a different molecule absorbs the photon, such as a photosensitizer that transfers energy to the dye, or a photocatalyst that activates the dye through oxidation, reduction, or hydrogen atom transfer, thereby initiating the dye's excitation and subsequent reactions

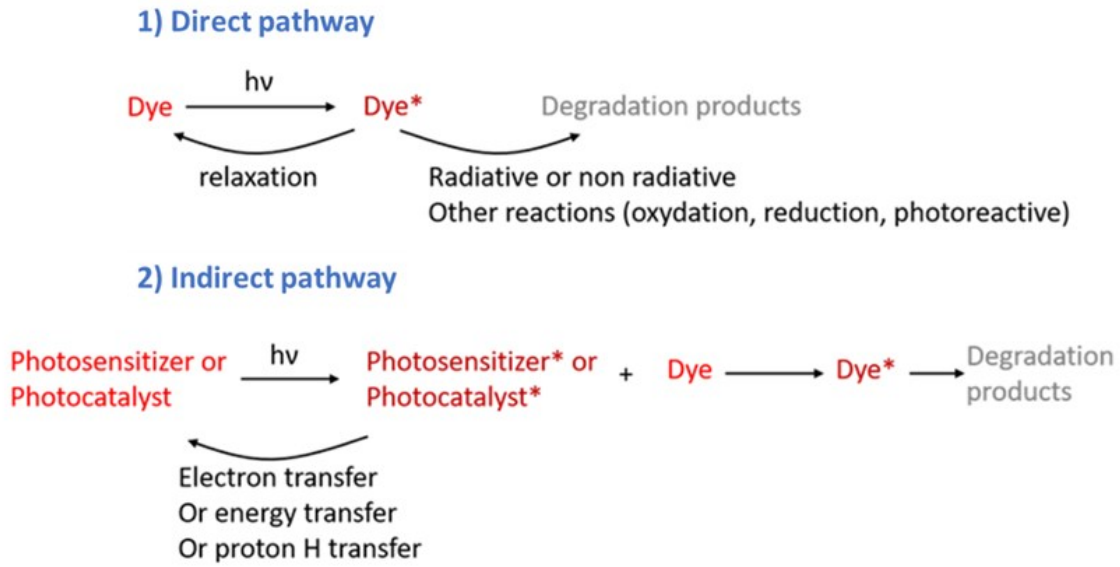


Figure 5. Photodegradation main mechanism: direct and indirect.

To further explore the underlying mechanism of photodegradation, it is crucial to identify the first reactive intermediate formed after light absorption. One plausible pathway involves the generation of radical species. To investigate this possibility, we recorded the EPR spectra of alizarin in aqueous basic media under UVA irradiation. The results are shown in Figure 6. The observed g-value of 2.0061 suggests the formation of an oxygen-centred radical, supporting the involvement of a radical-mediated degradation pathway. The measured hyperfine coupling constants (hfcc) are $a_1 = 4.2$ G and $a_2 = 1.0$ G, suggesting that the unpaired electron interacts with two nearby nuclei with I=1/2, most likely hydrogen atoms. Simulated charges and spin densities are given in Figure 7, (details of computation are given in S.I. 29) suggest that the electron is delocalized over the ring. Notably, carbon atoms 4 and 12, each bonded to a hydrogen atom, may account for the observed coupling. Based on these results, the radical centre is proposed to be located on a deprotonated hydroxy group. Computed atomic charges and spin densities, shown in Figure 7, support this assignment: the hydroxy group at position 2 displays the highest negative charge and the largest spin density among all oxygen atoms.

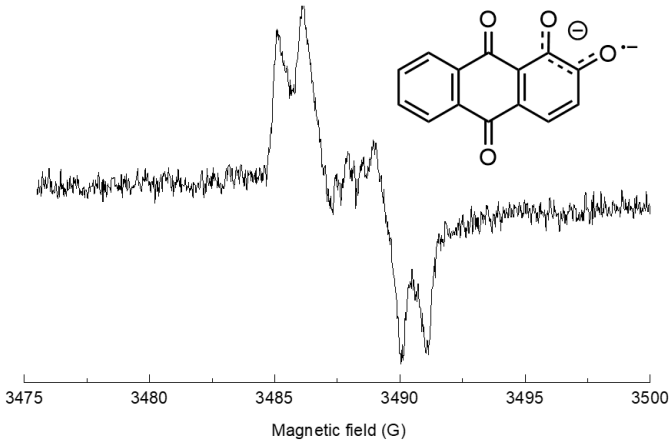


Figure 6. Experimental EPR spectra obtained upon photoexcitation (λ_{max}= 365 nm, irradiance 200 mW/cm^2) of alizarin in water at pH=12 under air. Initial concentration of the alizarin was 0.5 mM (Magnetic field sweep width, SW =25 G; modulation amplitude is 3G. The measured hyperfine coupling constants, hfcc: a_1=4.2G, a_2=1.0G, g = 2.0061 ± 0.0005.

N° ATOM	CHARGE	SPIN
1 C	0.00	0.05
2 C	-0.02	0.09
3 C	-0.15	-0.03
4 C	-0.13	0.10
5 C	-0.10	-0.03
6 C	-0.12	-0.02
7 C	0.03	0.07
8 C	-0.07	0.10
9 C	-0.07	0.08
10 C	0.01	0.11
11 C	-0.08	0.00
12 C	-0.09	0.06
13 C	-0.09	0.03
14 C	-0.08	0.02
18 O	-0.57	0.10
17 O	-0.54	0.06
16 O	-0.50	0.10
15 O	-0.49	0.10
19 H	-0.01	0.00
21 H	0.01	0.00
24 H	0.01	0.00
23 H	0.03	0.00
22 H	0.03	0.01
20 H	-0.01	0.01

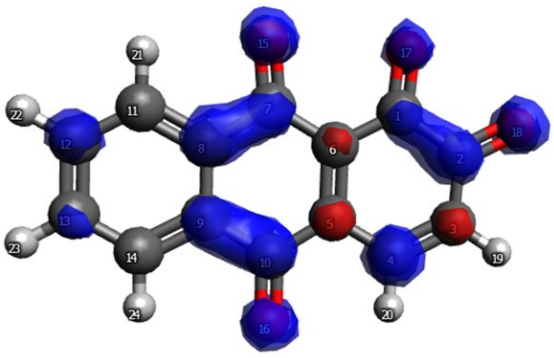


Figure 7. Computed Hirshfeld charges, spin density of alizarin di-deprotonated. The spin density is represented on the right (blue means positive spin and red negative spin density).

Based on the identified degradation products, a tentative mechanistic pathway is proposed in Figure 8. While not yet fully confirmed, this mechanism may be substantiated through future experimental studies. The pathways are hypothesised to proceed via endoperoxide intermediates in the presence of oxygen, consistent with previously

reported photodegradation mechanisms of anthraquinone derivatives in the literature [269–271]. Variations in the position of oxygen incorporation within these pathways give rise to distinct photodegradation products. EPR results indicate the presence of an oxygen-centred radical on carbon 1 and 2. This observation supports the likelihood of Pathway 2 being favoured, as it leads to the formation of degradation product P2, which was identified as the major photoproduct under basic conditions.

4. CONCLUSION

In this study, we explored the photodegradation behaviour of alizarin, a historical dye, in different solutions, in particular in water under basic conditions, to reach a better understanding of the reaction pathways. By employing EPR and HPLC-UV-MS techniques, we identified both radical intermediates and specific degradation products in water at pH=12. Our findings demonstrate that photodegradation is strongly influenced by environmental factors such as pH, solvent, and the light source. Notably, alkaline conditions accelerate degradation, which can be attributed to a radical mechanism involving oxygen radicals. Deprotonation of alizarin enhances its reactivity with oxygen, promoting faster photodegradation through radical pathways.
This study can open ways to improve understanding of degradation processes and help provide recommendations for the preservation of cultural heritage (artworks containing alizarin-based pigments), and offer insight for potential environmental applications (e.g., photodegradation of organic pollutants under controlled conditions). Additionally, these advancements can promote the sustainability of natural dyes, benefiting both cultural heritage conservation and industrial applications, including textiles, cosmetics, and food production.

5. ASSOCIATED CONTENT

Supporting Information

Lamp spectral distribution, kinetic plots, details of HPLC-UV-MS/MS spectra, EPR spectra before and after exposition and with Tempol, absorbance spectra of alizarin in DMSO and DMF, computational protocol.

Author information

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

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

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

Corresponding Author

* Dr. BERRAUD-PACHE Romain

Laboratoire d'Archéologie Moléculaire et Structurale (LAMS)

Sorbonne Université, CNRS, UMR 8220, Paris, 75005, France

Author Contributions

Thanh Huyen Linh Tran: Writing – review & editing, Writing – original draft, Visualization, Software, Investigation, Formal analysis, Data curation: Fabrice Morlet-Savary, Jacques Lalevee: Formal analysis, Data curation. Maguy Jaber: Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Conceptualization. Romain Berraud-Pache: Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Formal analysis, Data curation, Conceptualization.

Acknowledgements

We thank Salomé Poyer for her assistance with the HPLC-UV-MS measurements.

SUPPLEMENTARY DATA

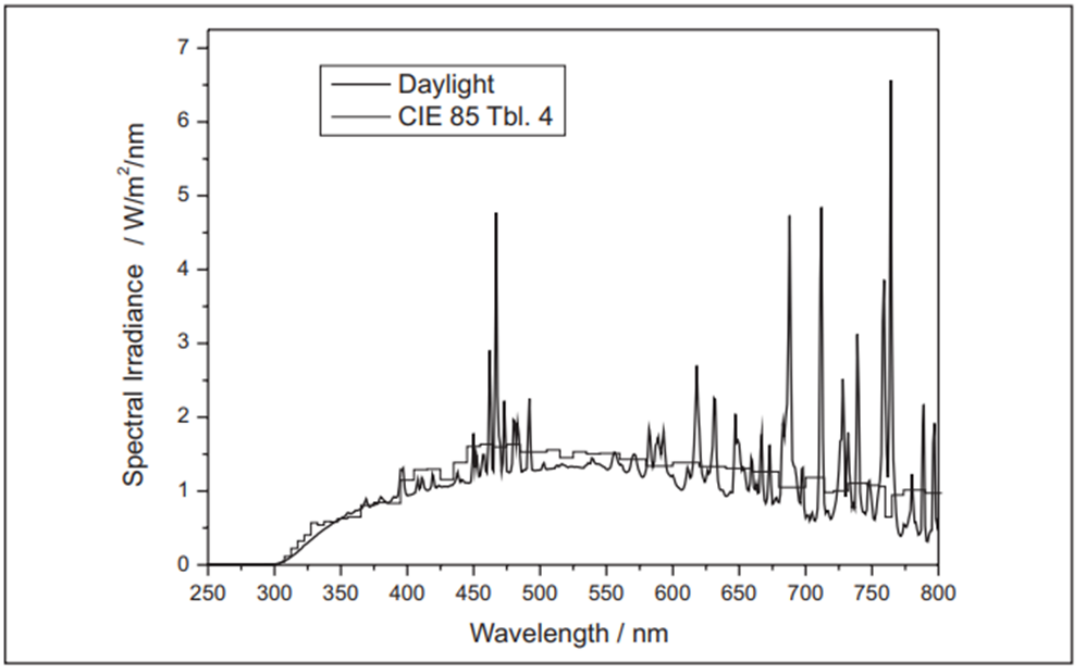
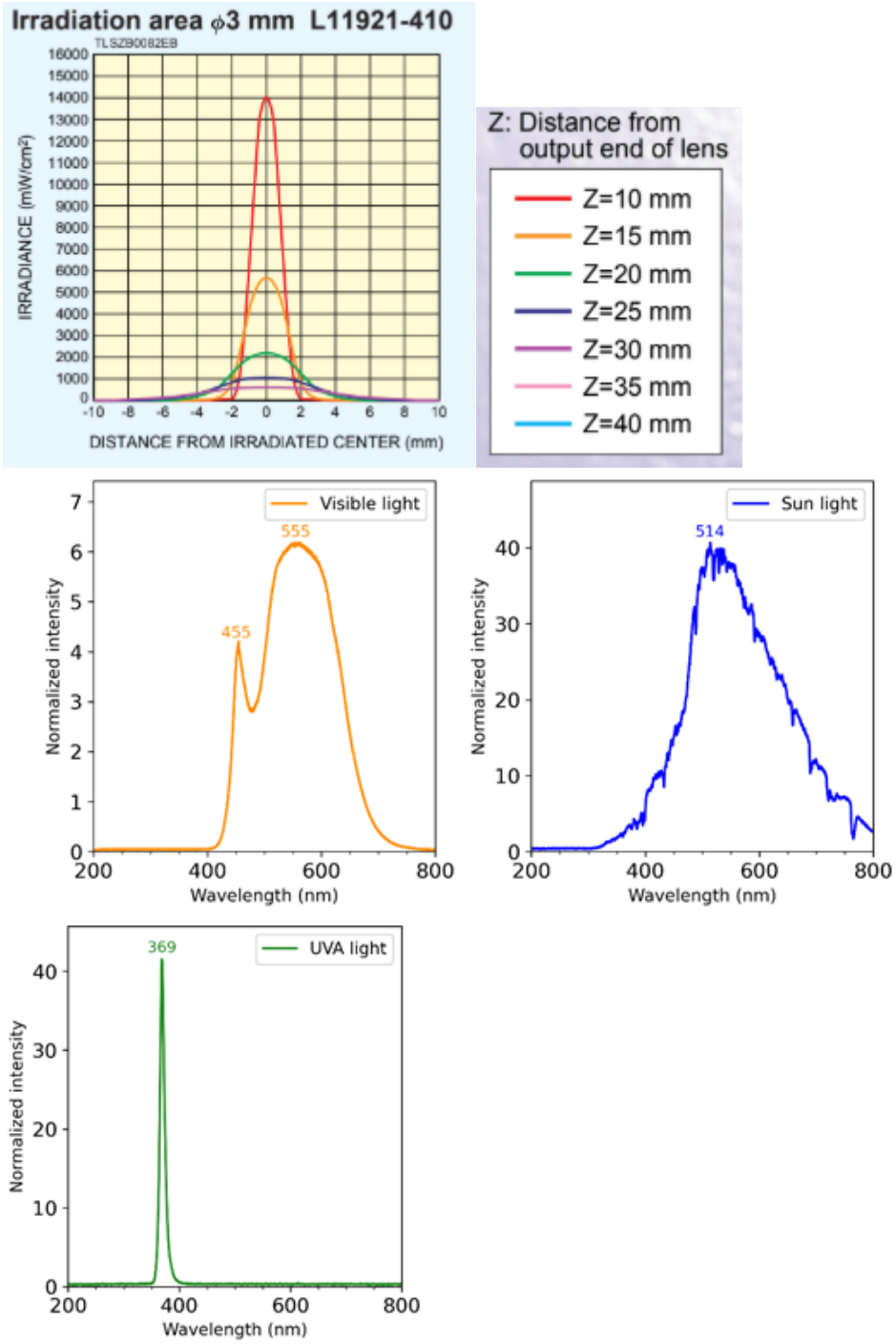


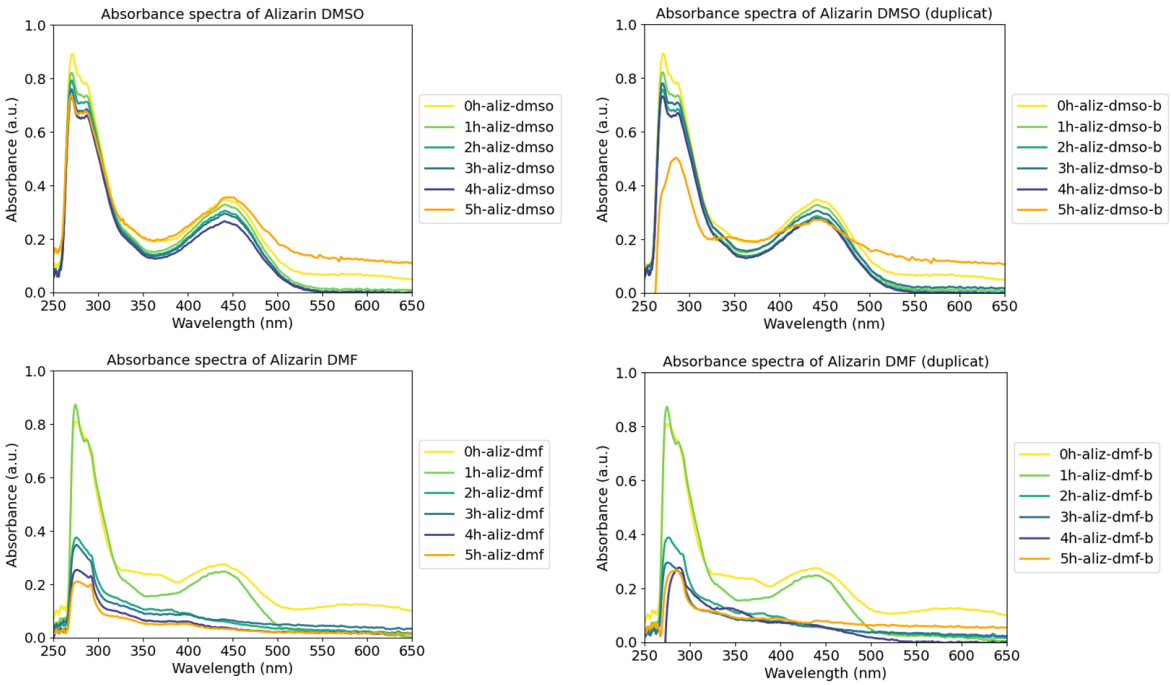
Table 5.1

RSP – Relative Spectral Portions			
Wavelength range	ISO 4892-2	CIE 85 Tbl.4	Daylight filtering
< 290 nm	0.0% – 0.2%	0.0%	0.0%
290 <= <= 320	2.6% – 7.9%	5.4%	3.4%
320 < <= 360	28.2% – 39.8%	38.2%	35.1%
360 < <= 400	54.2% – 67.5%	56.4%	61.6%
290 < <= 400		100%	100%

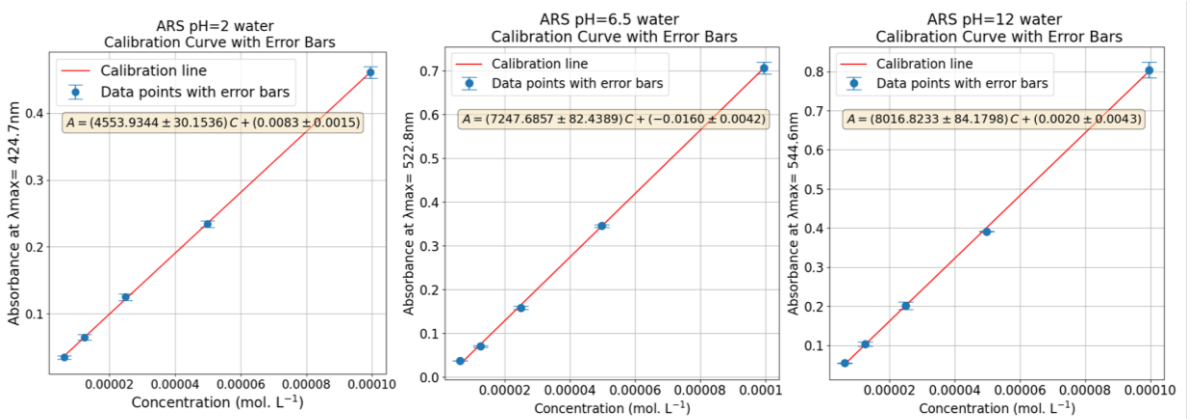
S.I. 1. *Spectral distribution daylight of Xenon Chamber SUNTEST XLS+.*



S.I. 2. *Spectral distribution of UVA lamp and visible lamp used in this study. The sunlight is shown for reference. Spectra are taken from the furnisher or the S.I. of the article https://doi.org/10.1016/j.saa.2024.125371.*

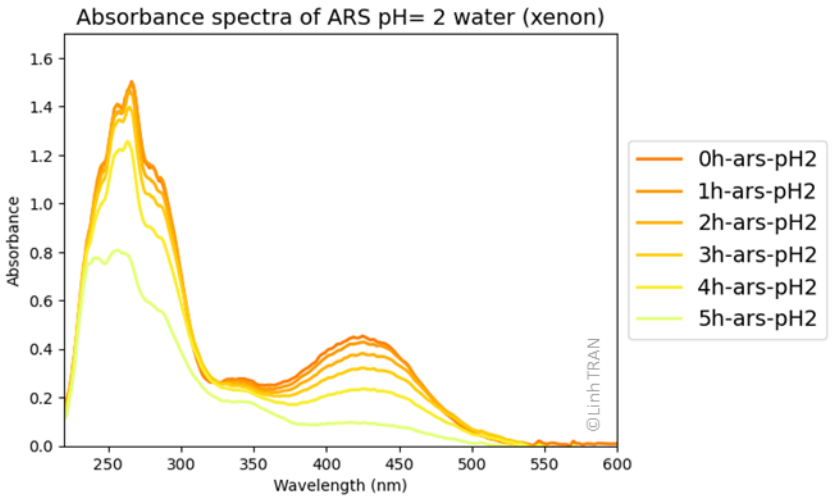


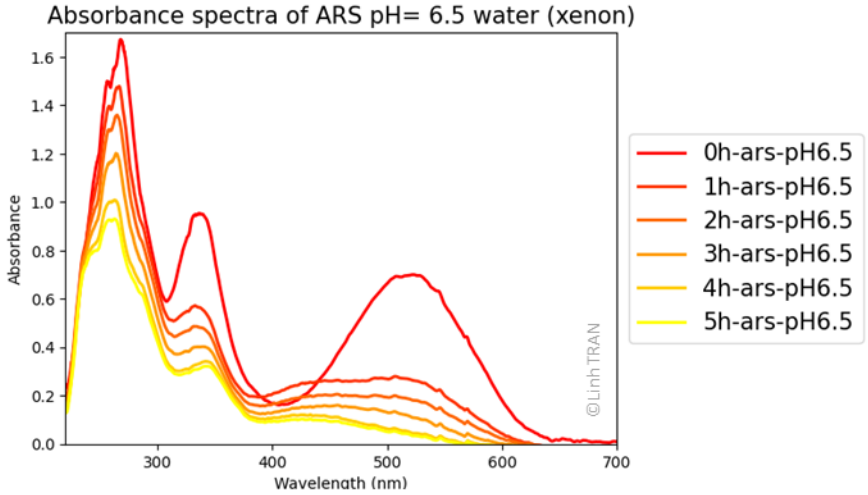
S.I. 3. *Absorbance spectra of alizarin in DMSO (top) and DMF (bottom) after various durations of xenon daylight exposure. Experiments were repeated twice to assess reproducibility.*

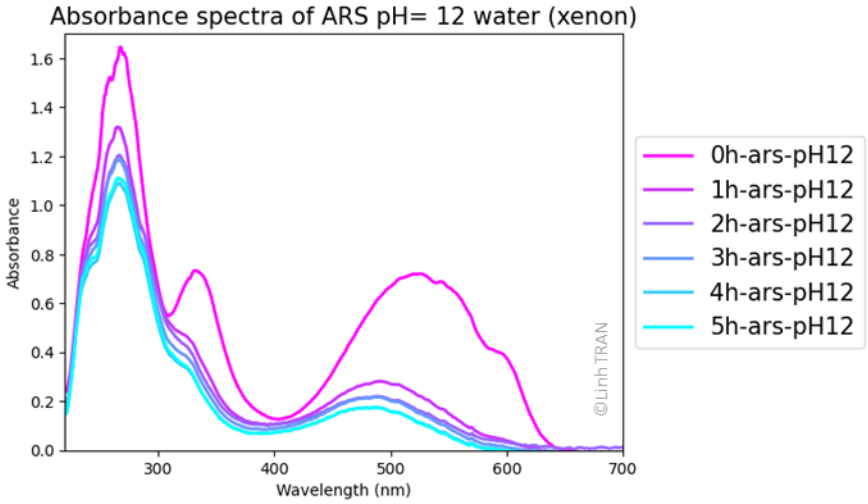


S.I. 4. *Calibration curve with error bars (with triplicats) for pH=2, pH=6.5 and pH=12 in water for alizarin red S. The calibration equation is shown on each plot.*

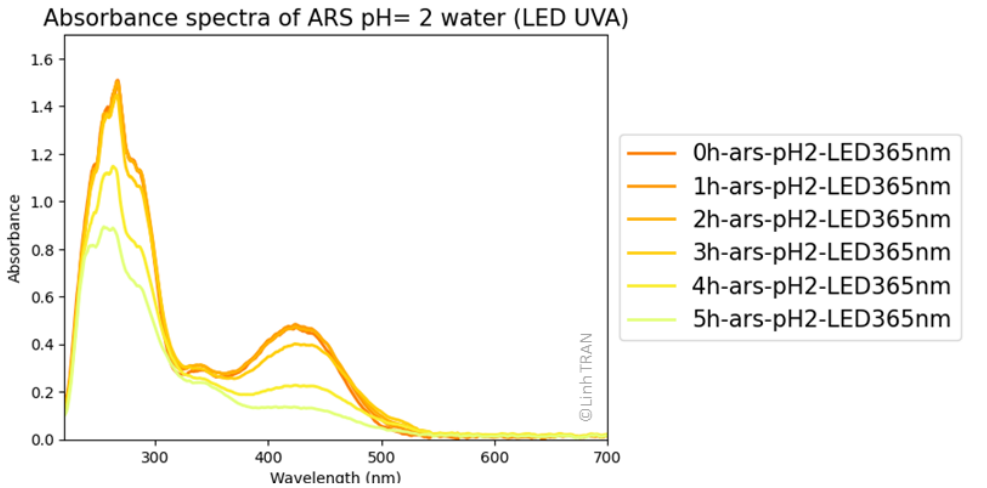
Photodegradation kinetic xenon

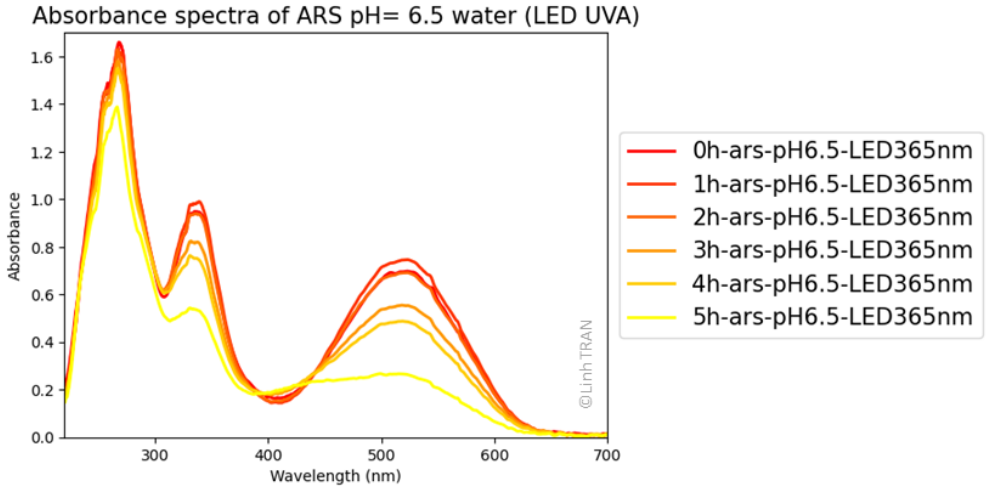


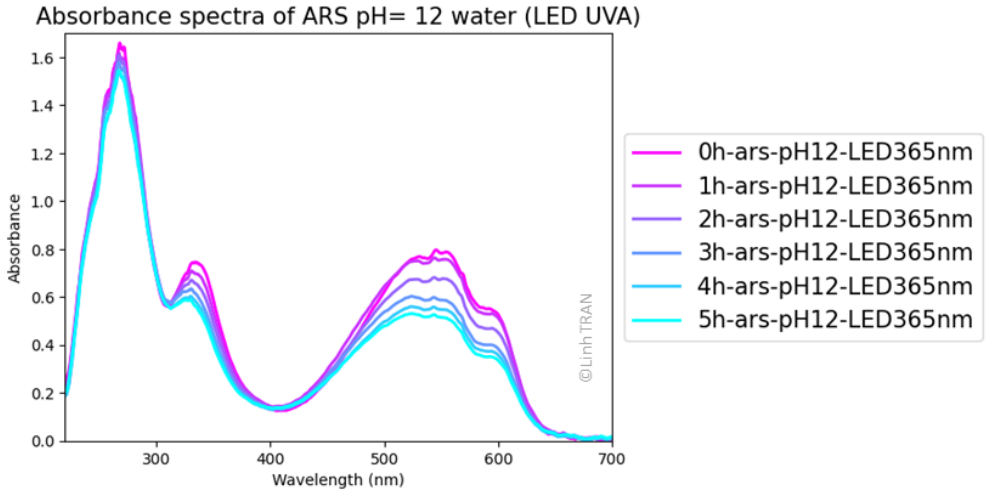


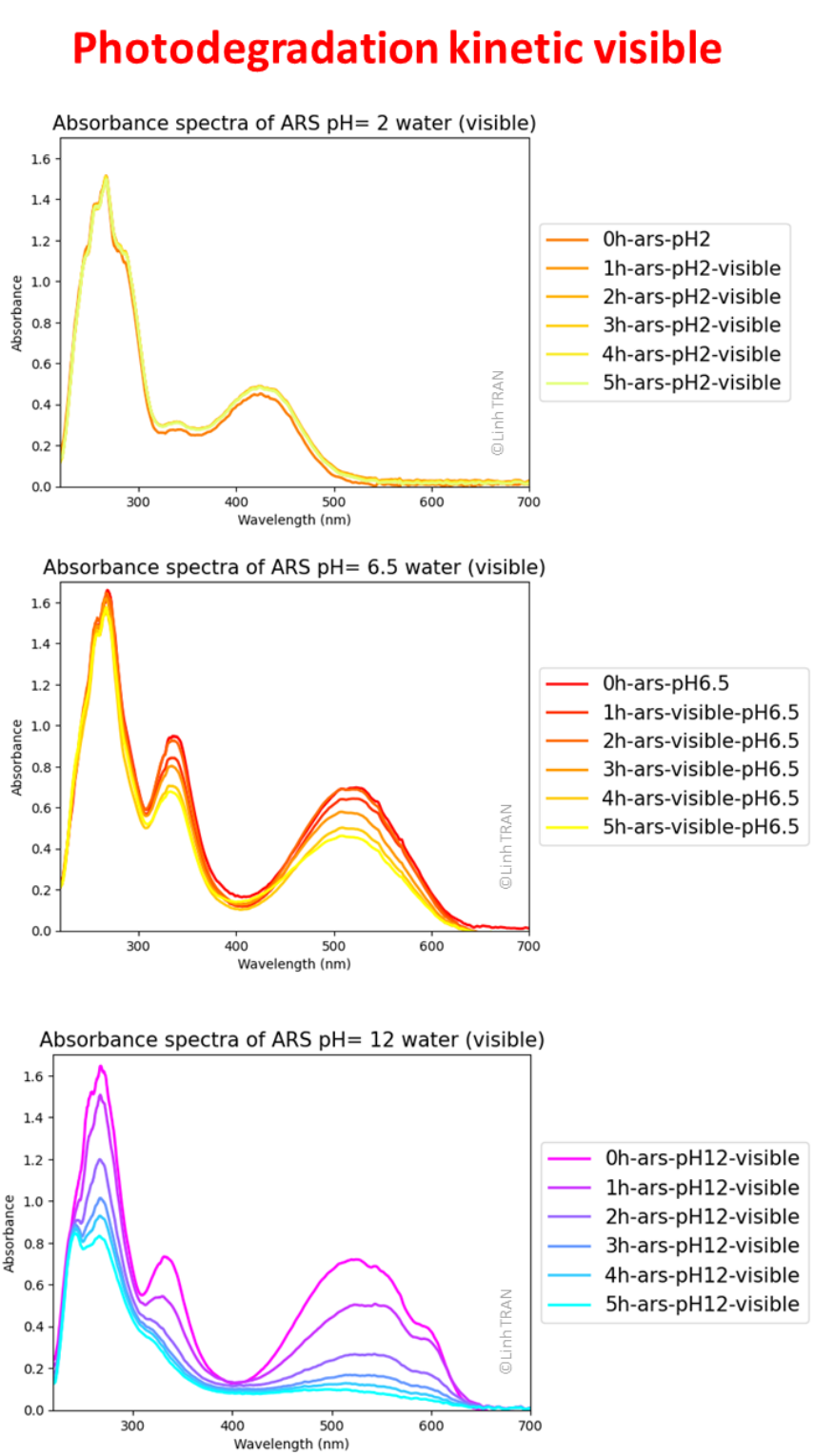


Photodegradation kinetic UVA

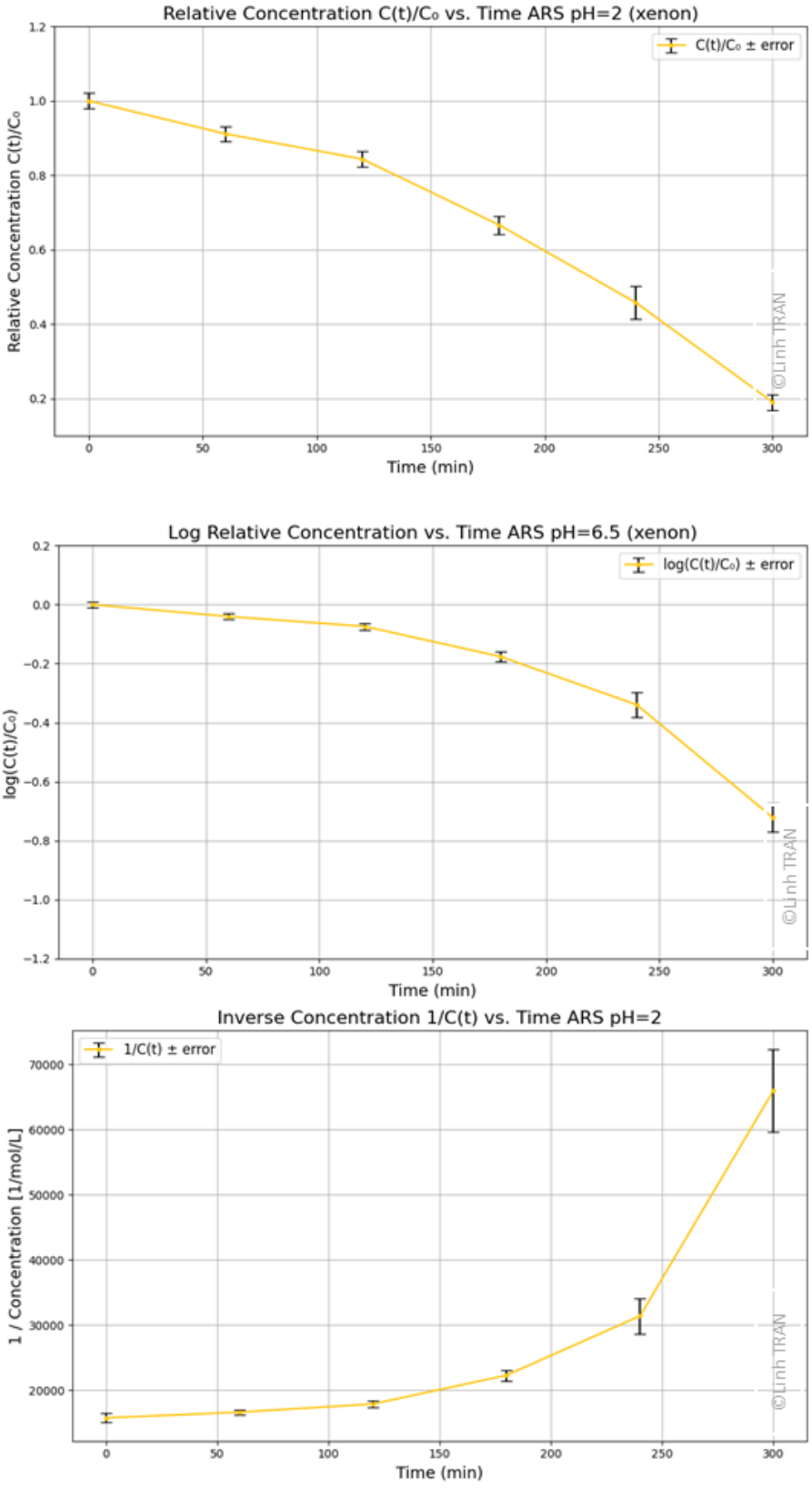




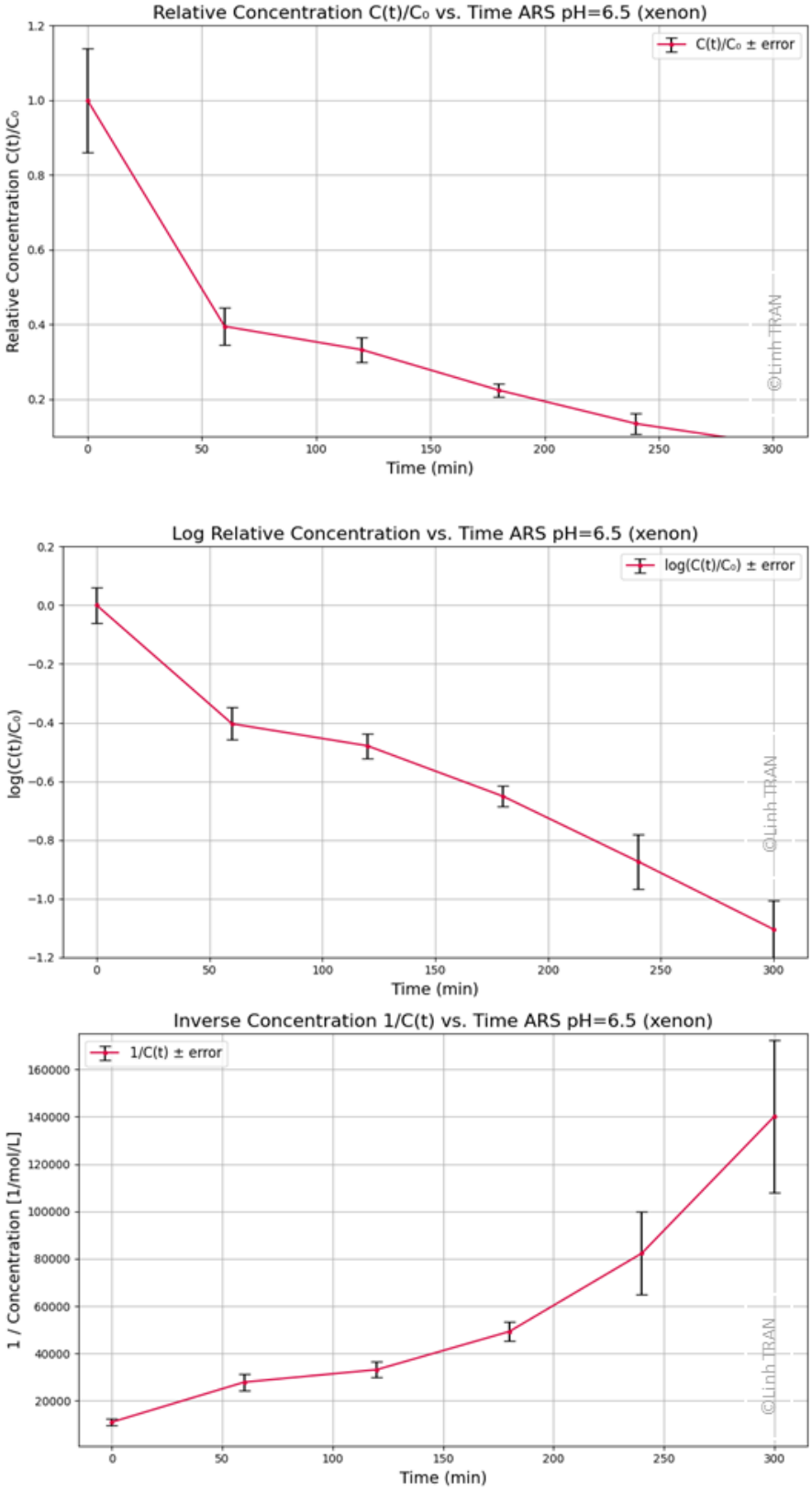




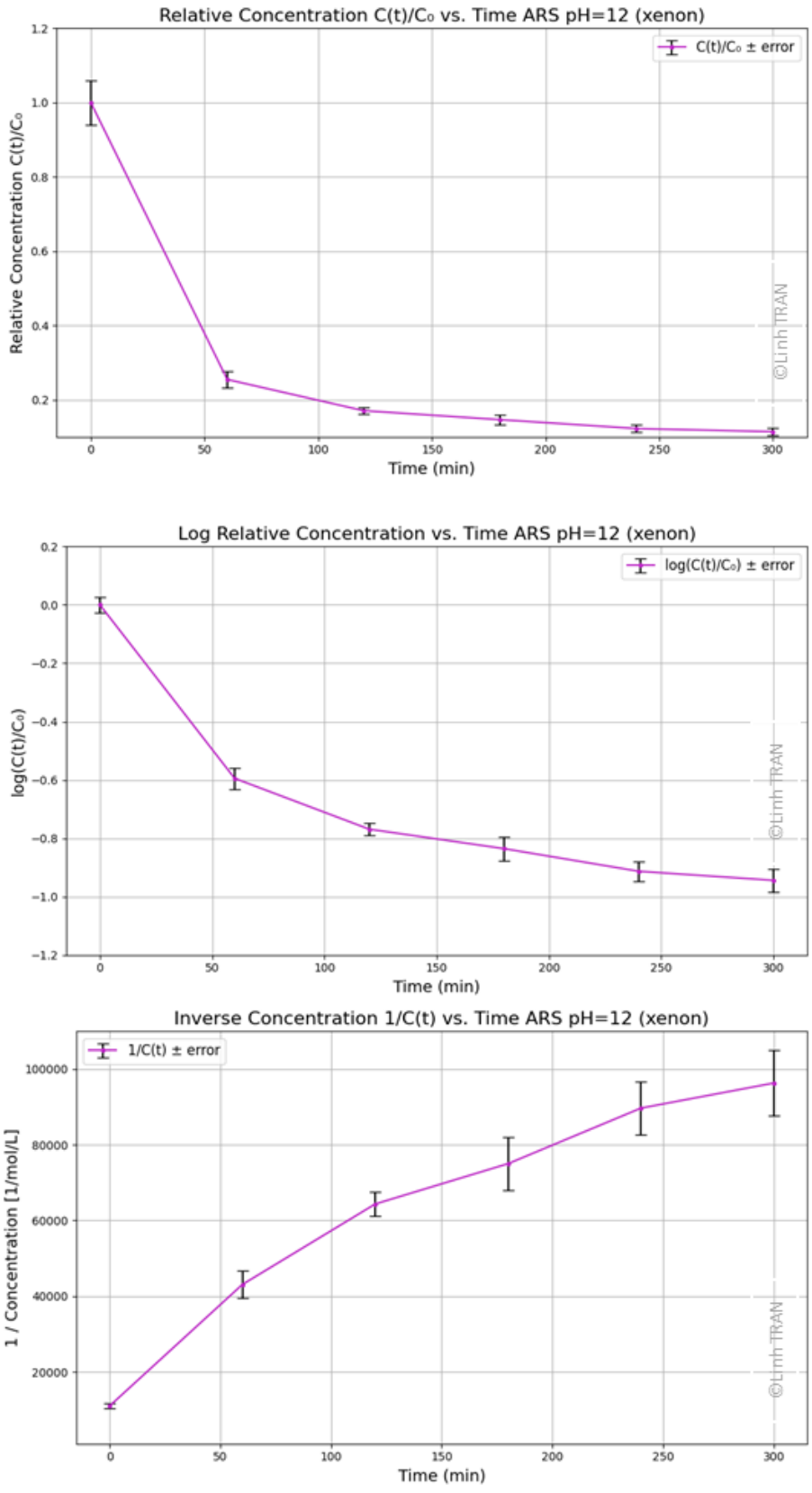
S.I. 5. *UV-VIS absorbance spectra of photodegradation of alizarin red S under different lamps (visible with irradiance= 10mW/cm^2, UVA with irradiance= 10mW/cm^2, and xenon daylight lamp with irradiance= 50mW/cm^2) and different pH (pH=2, 6.5, and 12).*



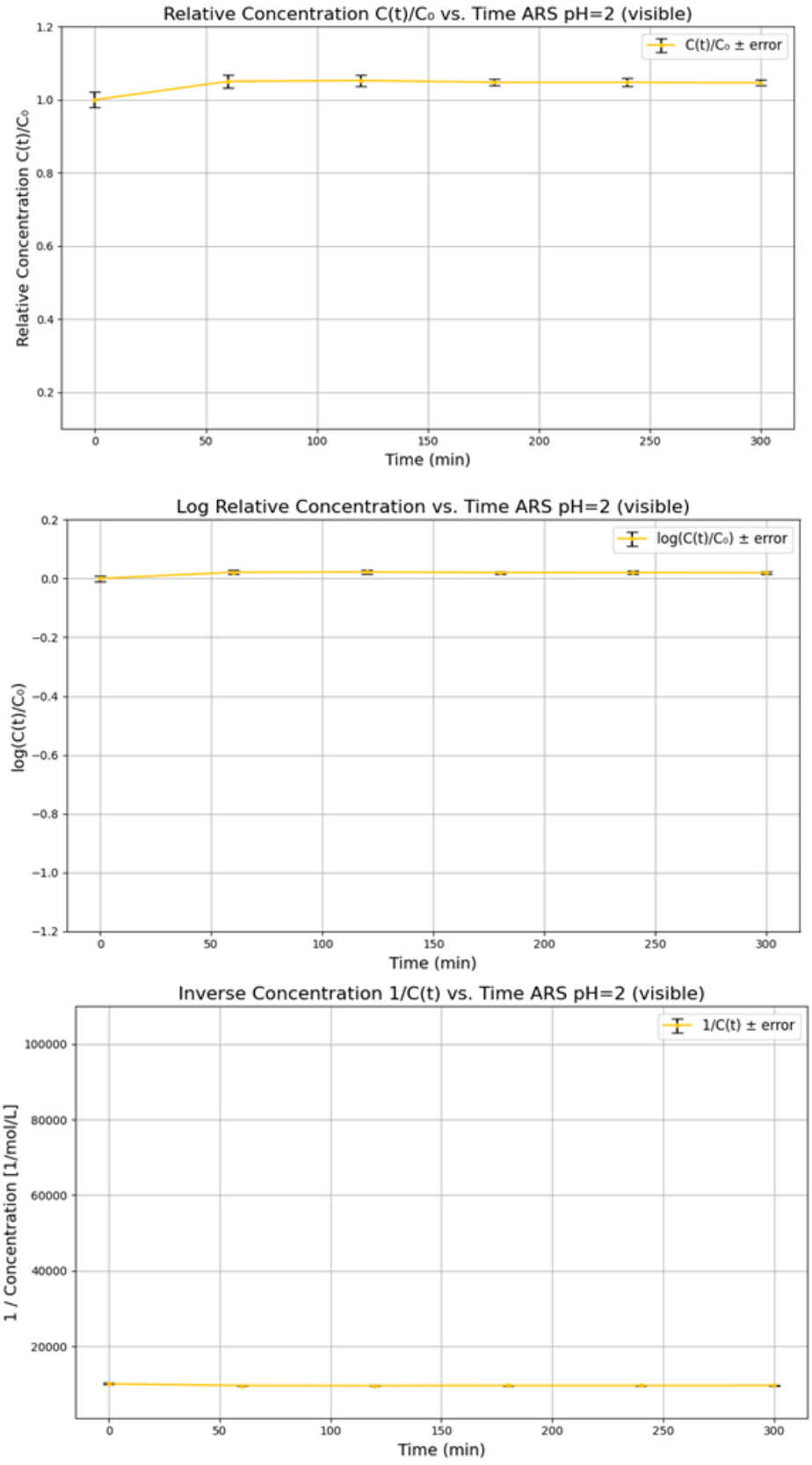
***S.I. 6**. Kinetic study of ARS in water at pH=2 and upon exposition to xenon light (irradiance =50mW/cm^2).*



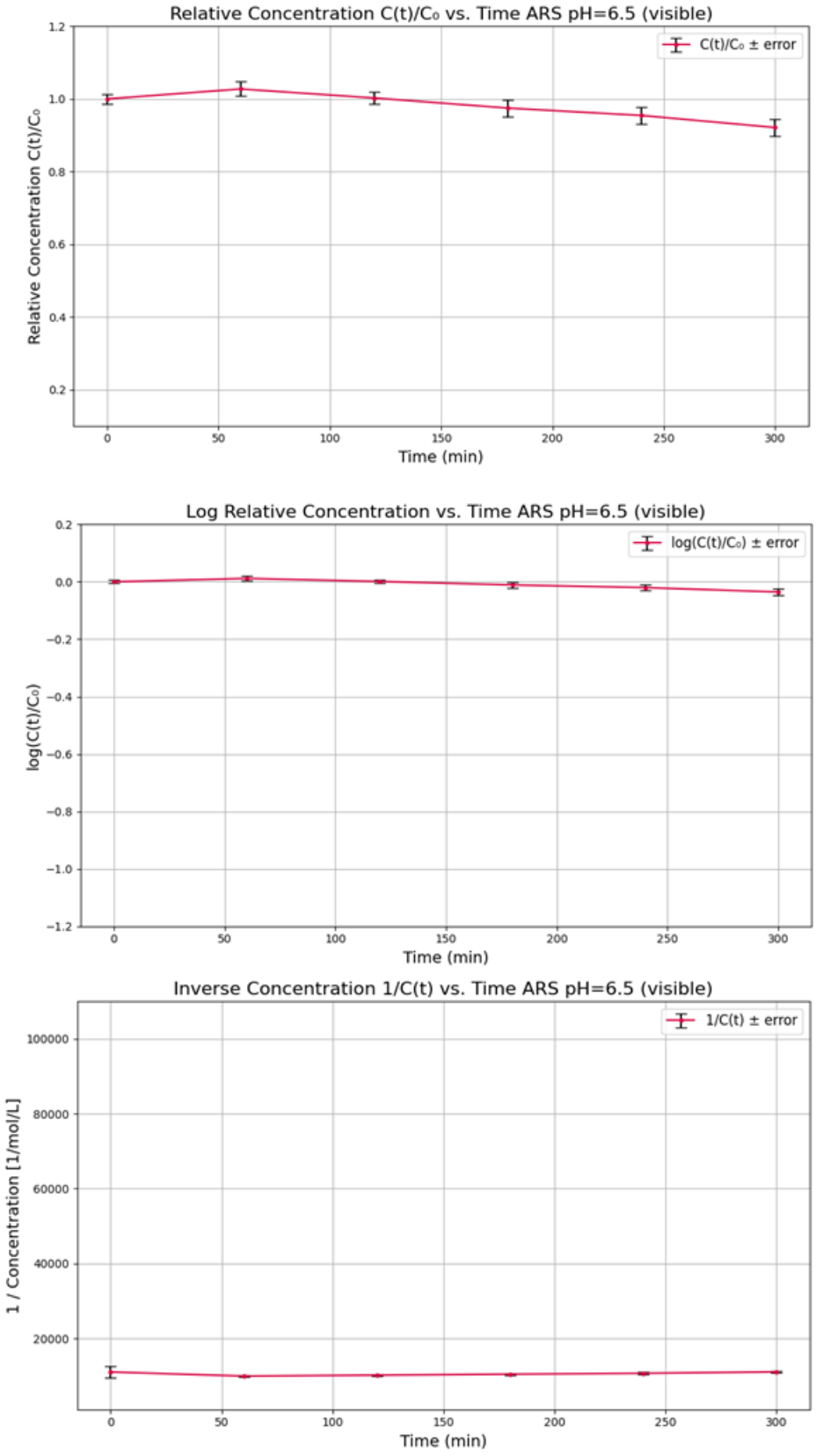
S.I. 7. *Kinetic study of ARS in water at pH=6.5 and upon exposition to xenon light (irradiance =50mW/cm²).*



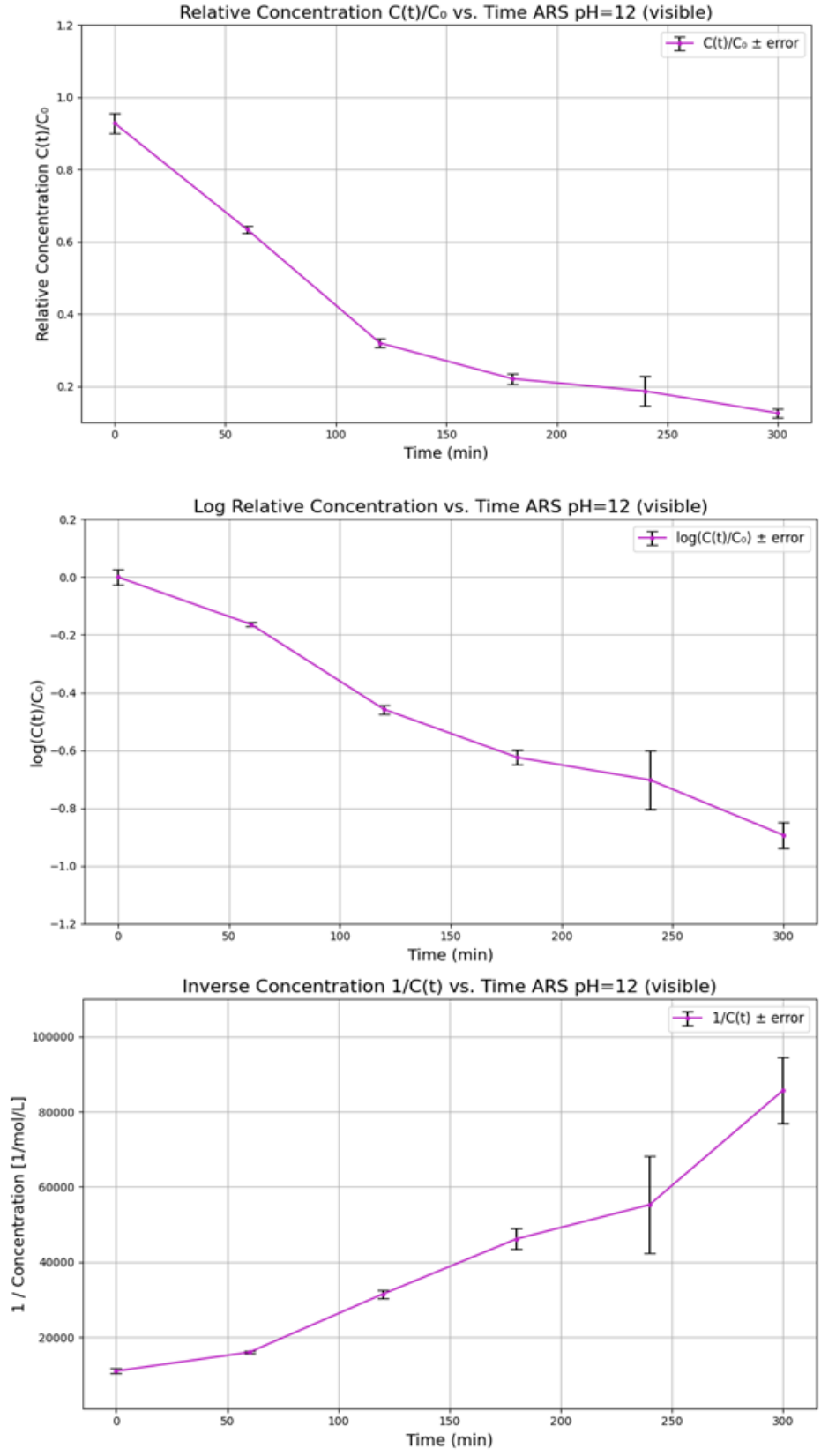
***S.I. 8**. Kinetic study of ARS in water at pH=12 and upon exposition to xenon light (irradiance =50mW/cm^2).*



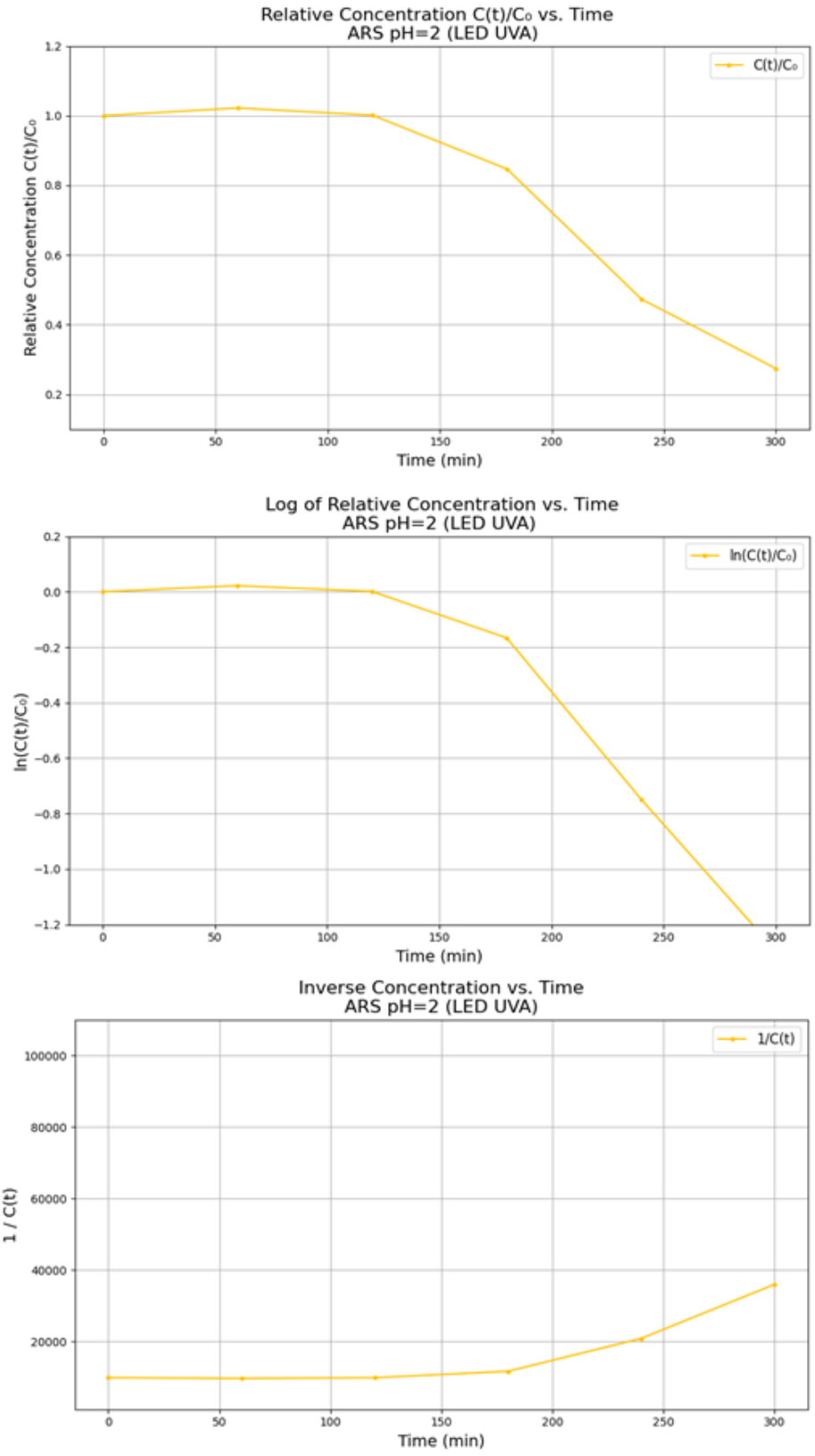
S.I. 9*. Kinetic study of ARS in water at pH=2 and upon exposition to visible light (irradiance =10mW/cm²).*



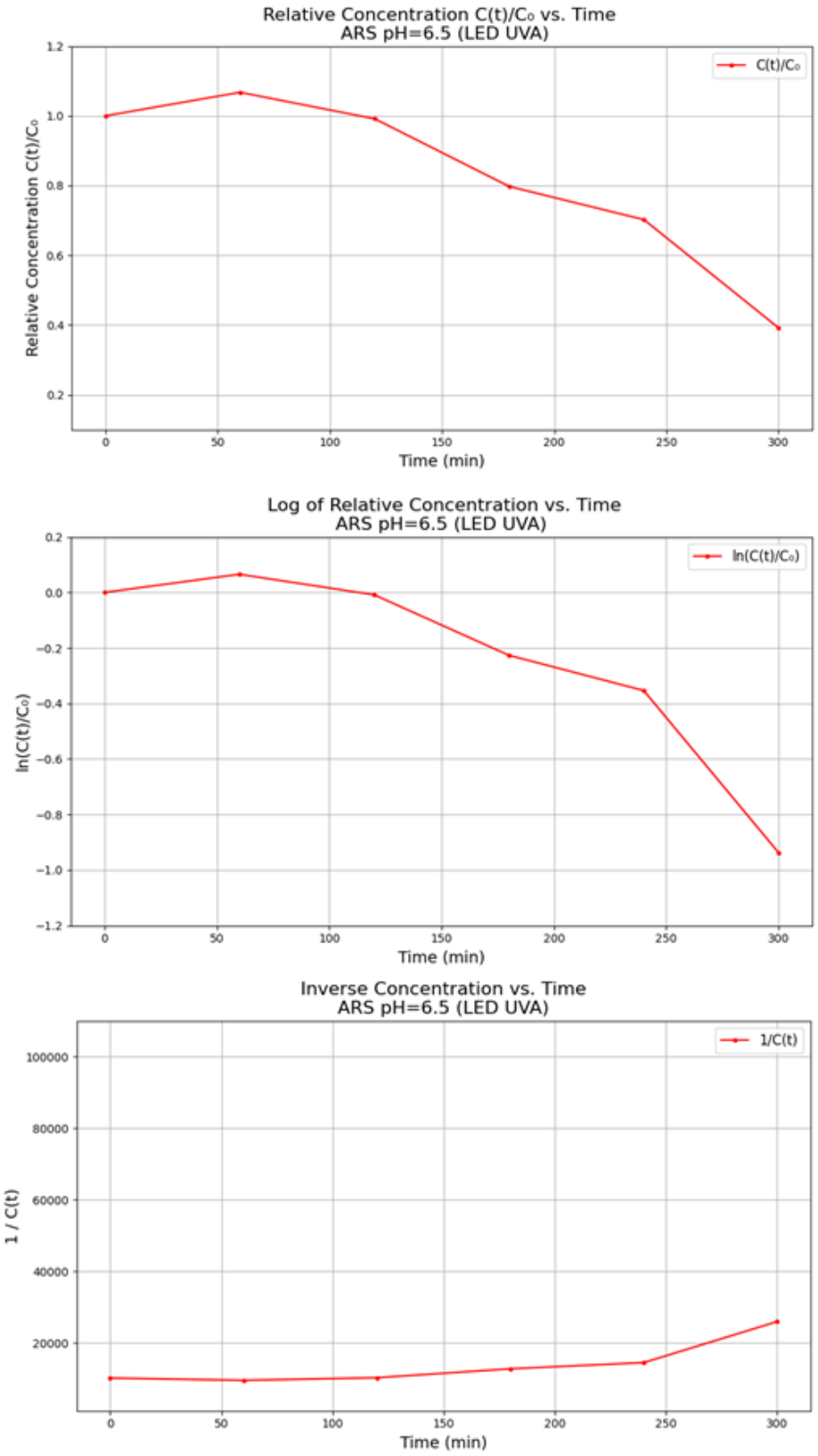
***S.I. 10**. Kinetic study of ARS in water at pH=6.5 and upon exposition to visible light (irradiance =10mW/cm^2).*



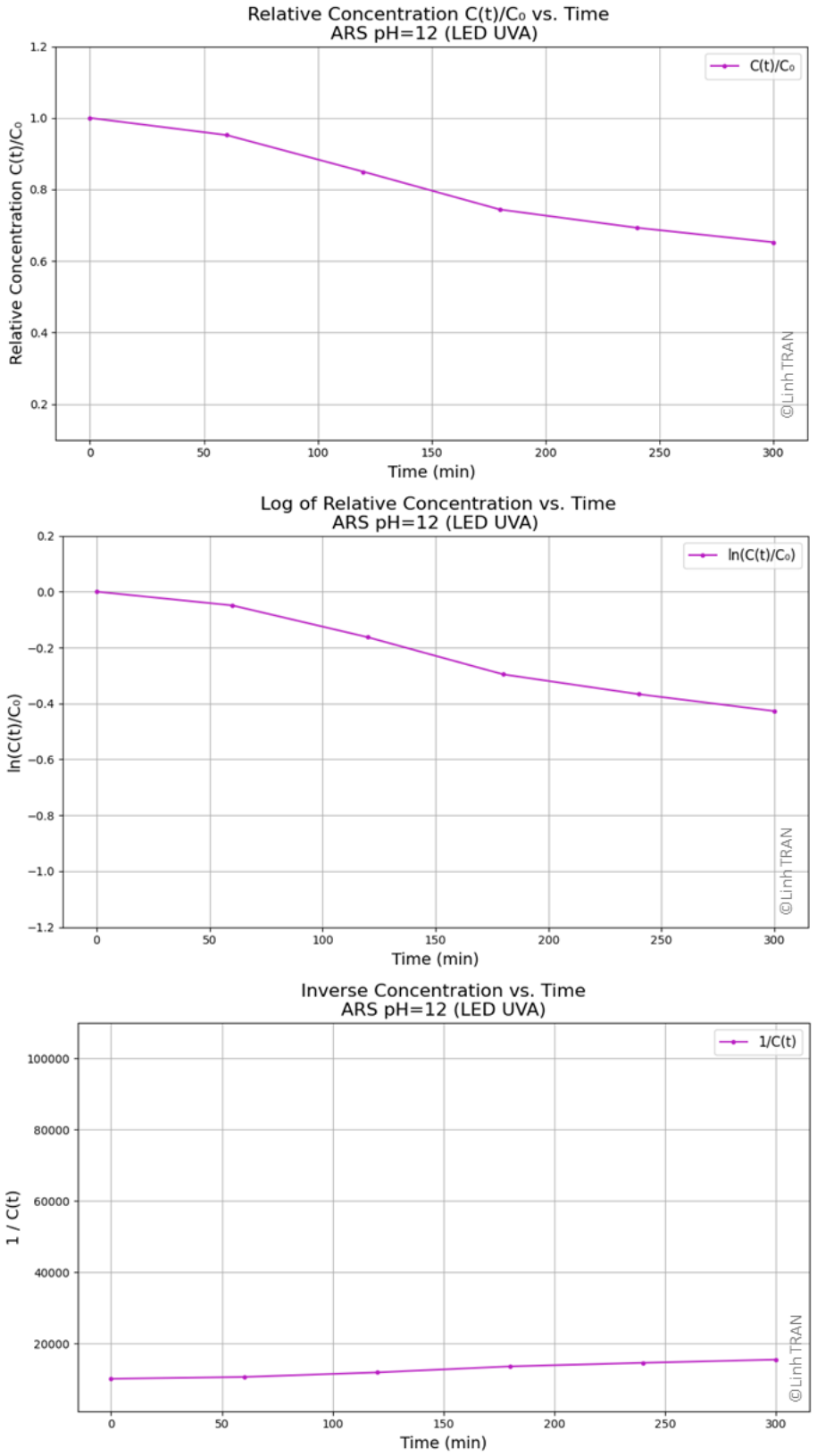
S.I. 11*. Kinetic study of ARS in water at pH=12 and upon exposition to visible light (irradiance =10mW/cm^2).*



S.I. 12*. Kinetic study of ARS in water at pH=2 and upon exposition to UVA light (irradiance =10mW/cm^2).*

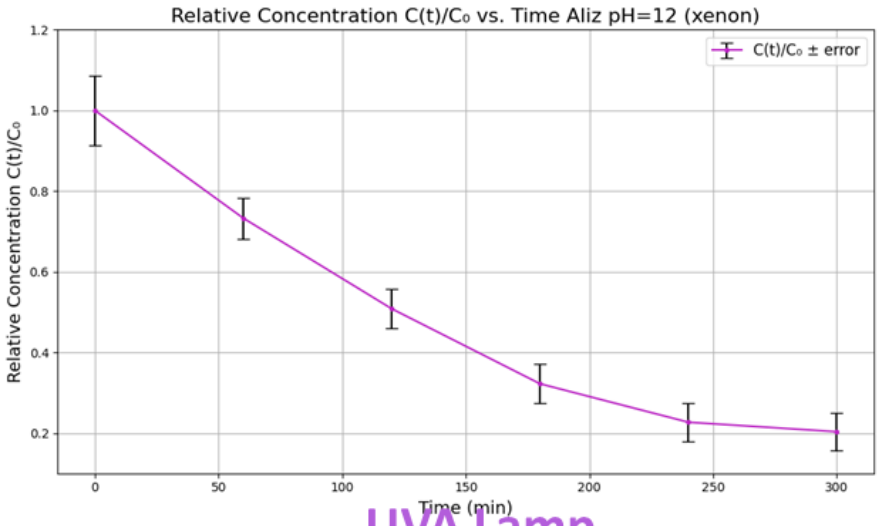


S.I. 13*. Kinetic study of ARS in water at pH=6.5 and upon exposition to UVA light (irradiance =10mW/cm^2).*

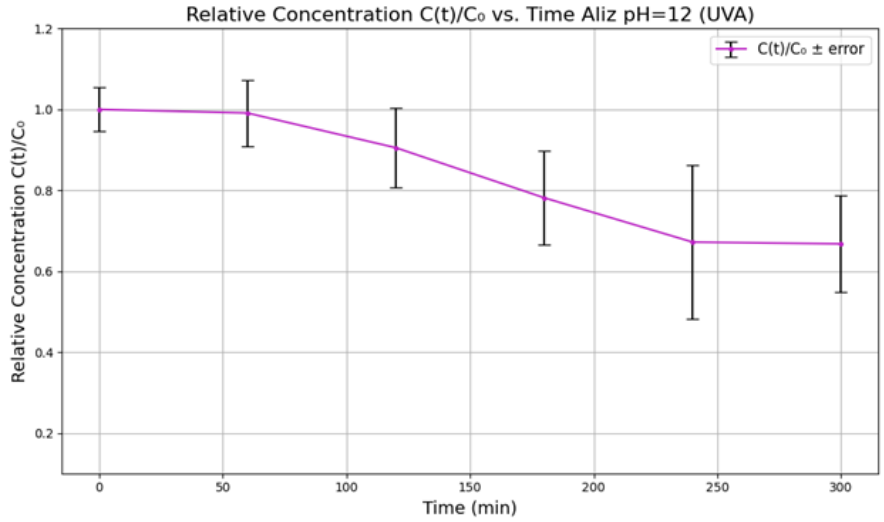


S.I. 14*. Kinetic study of ARS in water at pH=12 and upon exposition to UVA light (irradiance =10mW/cm²).*

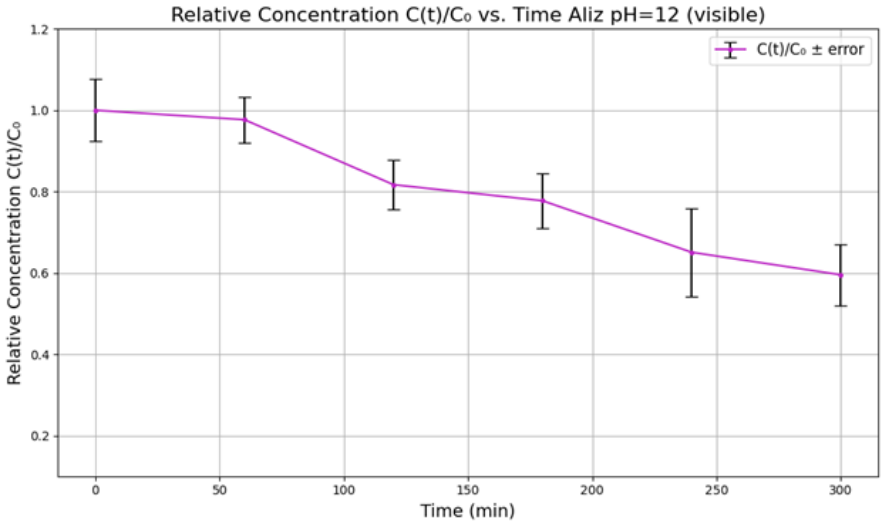
Xenon Lamp (Daylight)



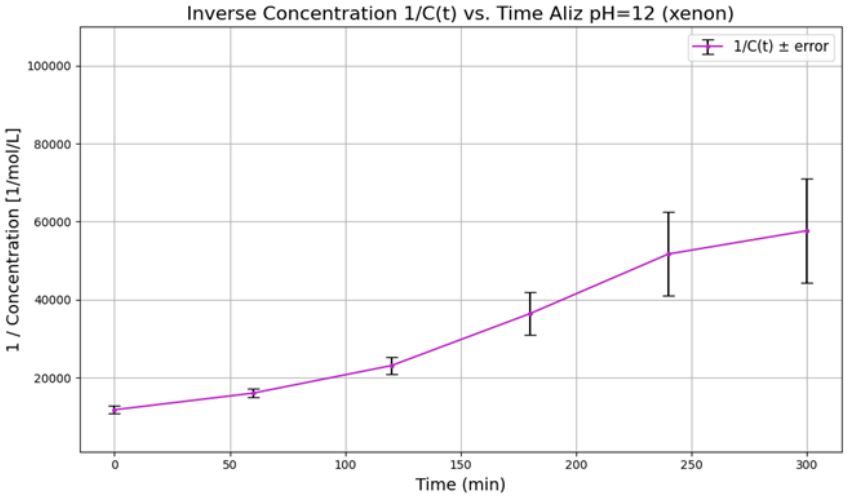
UVA Lamp



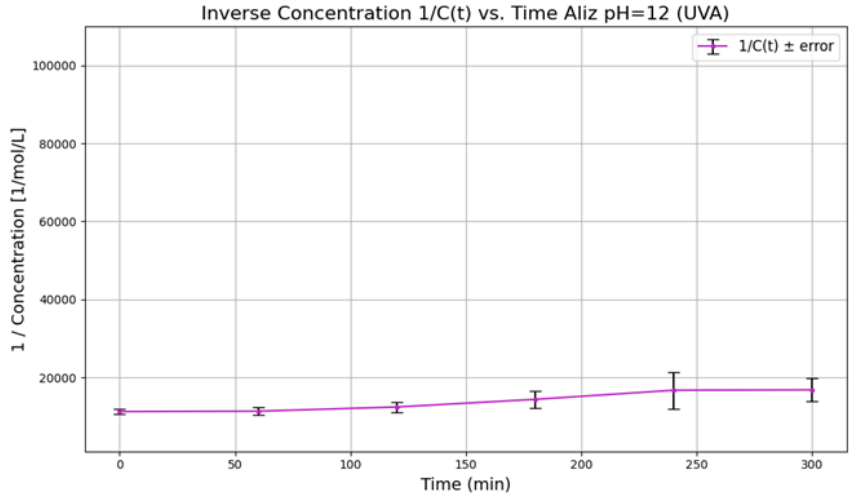
Visible Lamp



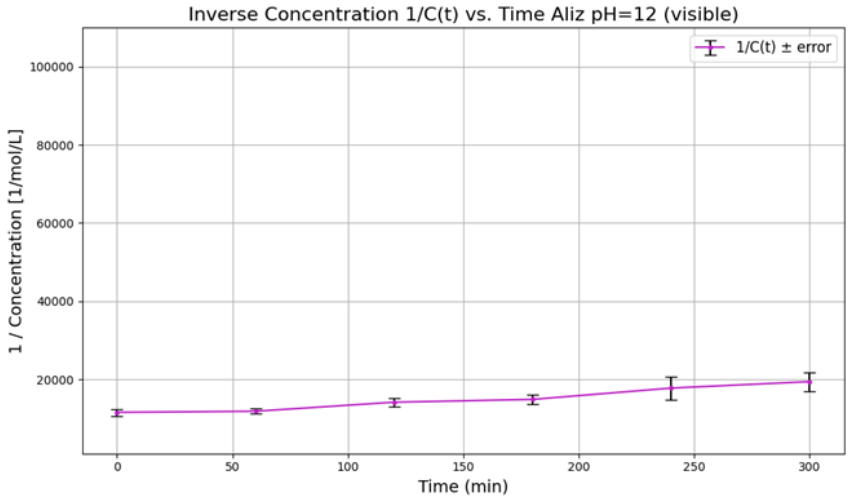
Xenon Lamp (Daylight)



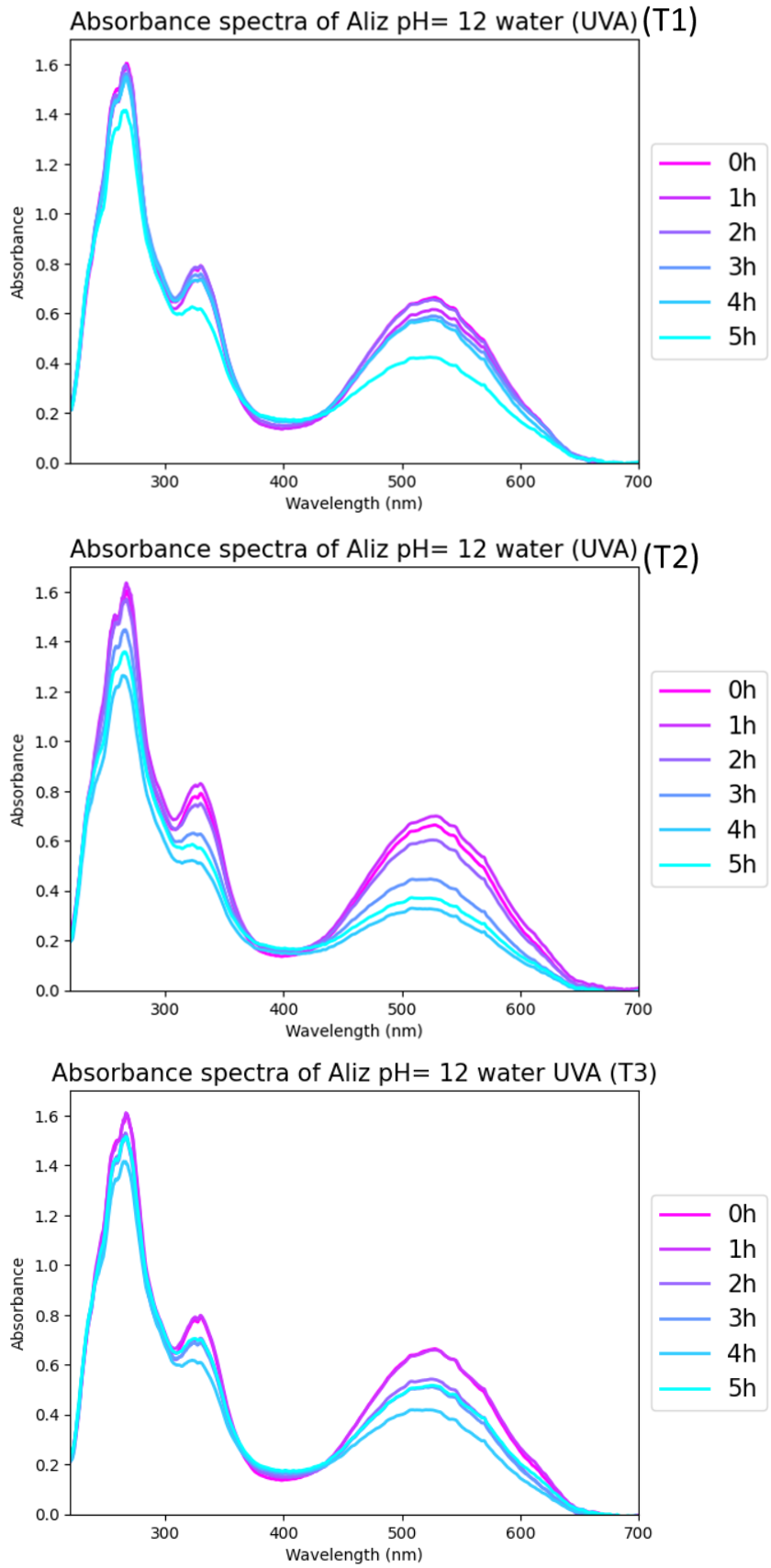
UVA Lamp



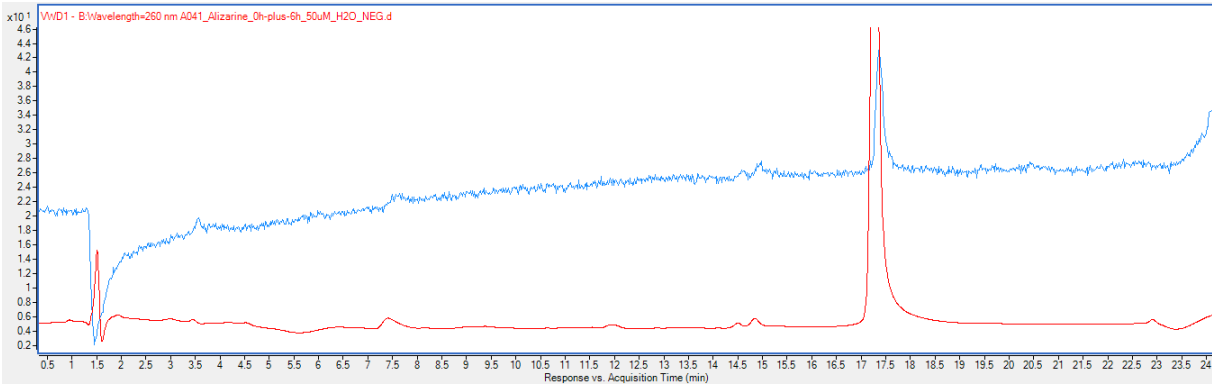
Visible Lamp



S.I. 15. *Photodegradation of alizarin at pH=12 in water under different lamps (xenon, visiblen, UVA). The plots shows concentration C(t) over exposition time in minutes and 1/C(t).*



S.I. 16. *Alizarin photodegradation under UVA light (triplicats: T1, T2 and T3).*



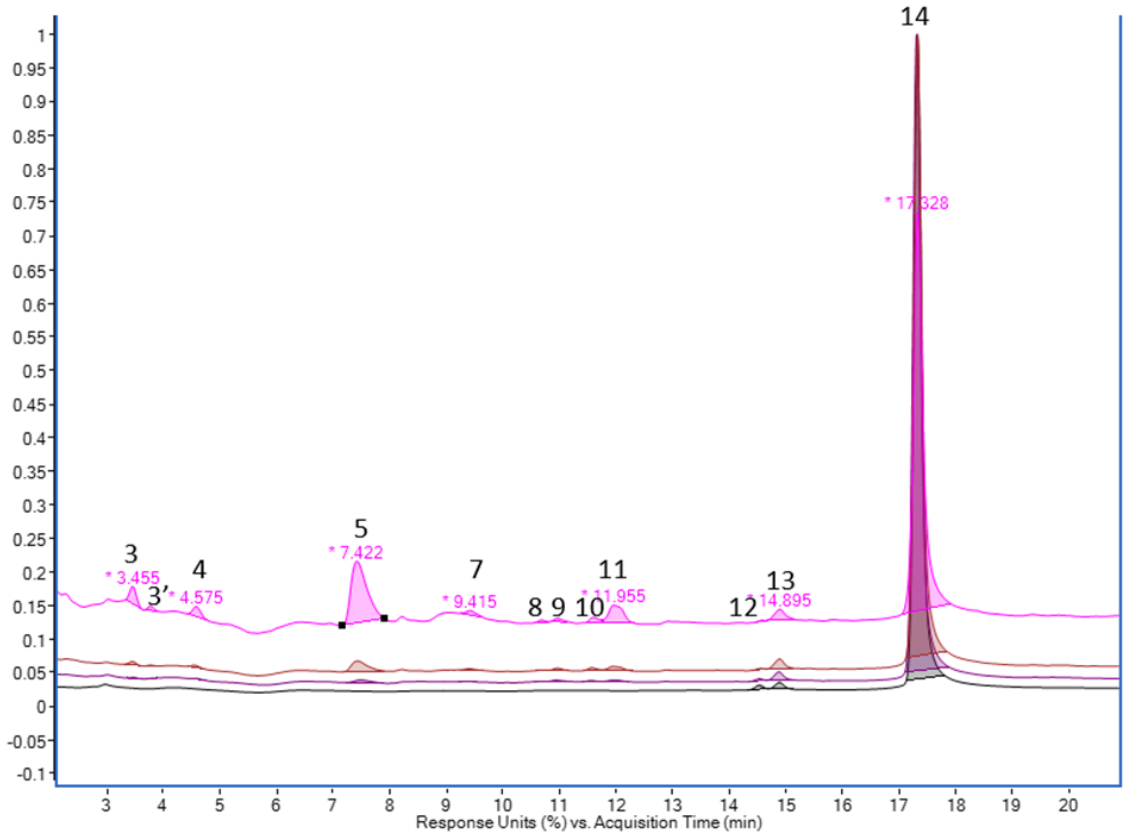
S.I. 17. *Overlay of MS signal (negative ion mode) in blue and UV signal at 260 nm in red for the mixture of alizarin before and after irradiation of 6h in water at pH=12 adjusted with* NH_4OH.

S.I. Table 1. *Peak identification from LC-UV-MS experiments in negative and positive ionization modes. For ND m/z, it corresponds to a UV signal without peaks in the MS dimension, and for ND in the chemical formula, it corresponds to compounds that do not contain the COH proportion to fit alizarin-like degradation products.*

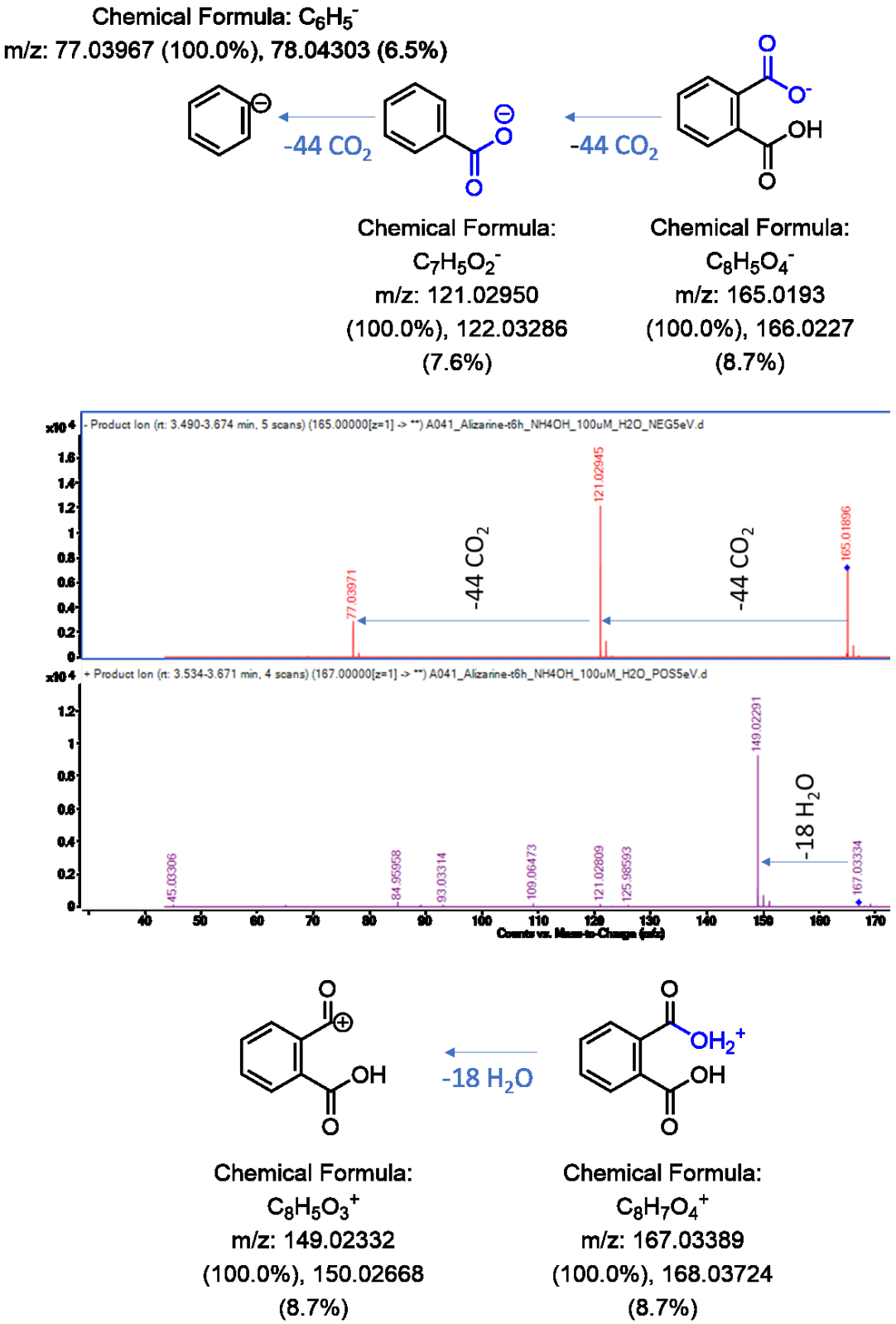
Peak	RT	m/z		Chemical formula (neutral)
		neg	pos	
1	2.0-2.2	164.0353	166.0502	C8H7NO3
2	3.0	ND	ND	
3	3.6-3.9	165.0192	167.0338	C8H6O4
3b	4.0	217.0617	219.0768	ND
4	4.7-5.3-8.2	224.0350	ND	C10H9O6
5	7.5-9.4	243.0298 217.0145	219.0292	C13H8O5 C11H6O5
6	8.3	ND	ND	
7	9.4	217.0142	ND	C11H6O5
8a	10.7	215.0461	217.0610	ND
8	10.8	240.0297		C10H9O7
9	10.9	212.0348	218.2116 (ND)	C9H9O6
10	12.1	ND	218.0449	C11H7NO4
11	12.3	ND	234.2066	ND
12	14.5	255.0296	257.0445	C14H8O5
13	14.9	255.0296	257.0445	C14H8O5
14	17.4	239.0351	241.0495	Alizarin C14H8O4

S.I. Table 2*. Integrated LC-UV chromatograms (in relative %). at 260 nm of alizarin irradiated at 0, 1, 3 or 6h in water at pH=12 adjusted with NH_4OH.*

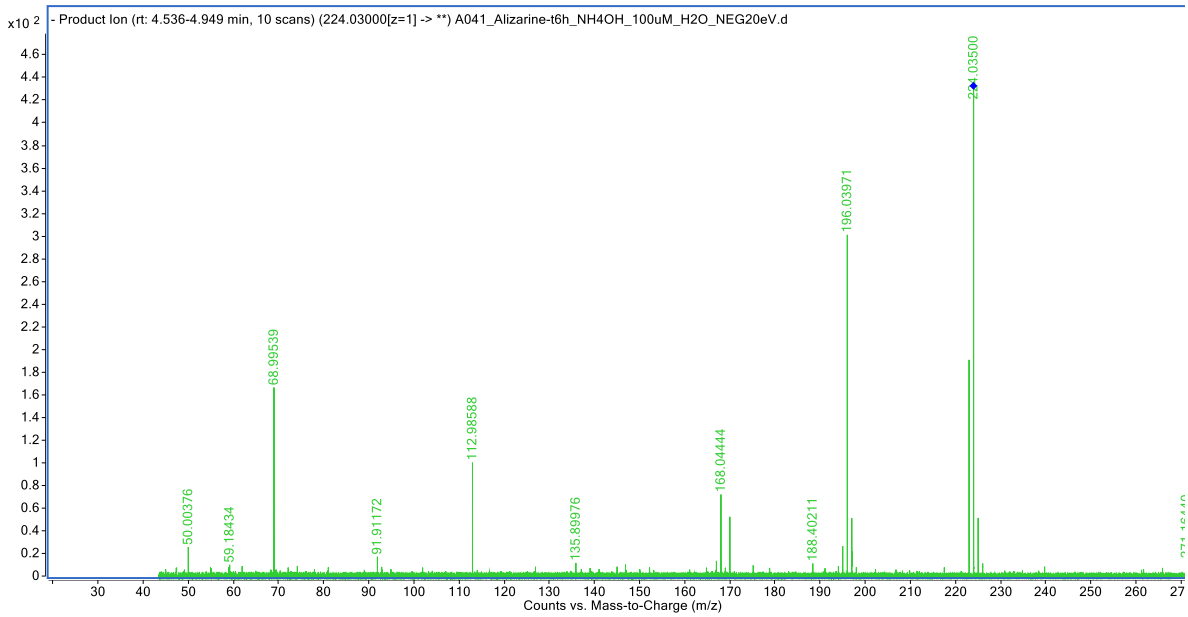
Compound	T = 0h	T = 1h	T = 3h	T = 6h
3	0.0	0.1	0.4	2.0
3'	0.0	0.0	0.1	0.5
4	0.0	0.1	0.3	1.4
5	0.0	0.8	2.9	17.2
7	0.0	0.1	0.2	0.8
8	0.0	0.0	0.0	0.3
9	0.0	0.2	0.3	0.5
10	0.0	0.2	0.5	0.8
11	0.0	0.3	1.1	4.5
12	0.6	0.3	0.1	0.1
13	1.0	1.3	1.6	2.0
14	98.3	96.7	92.5	69.9



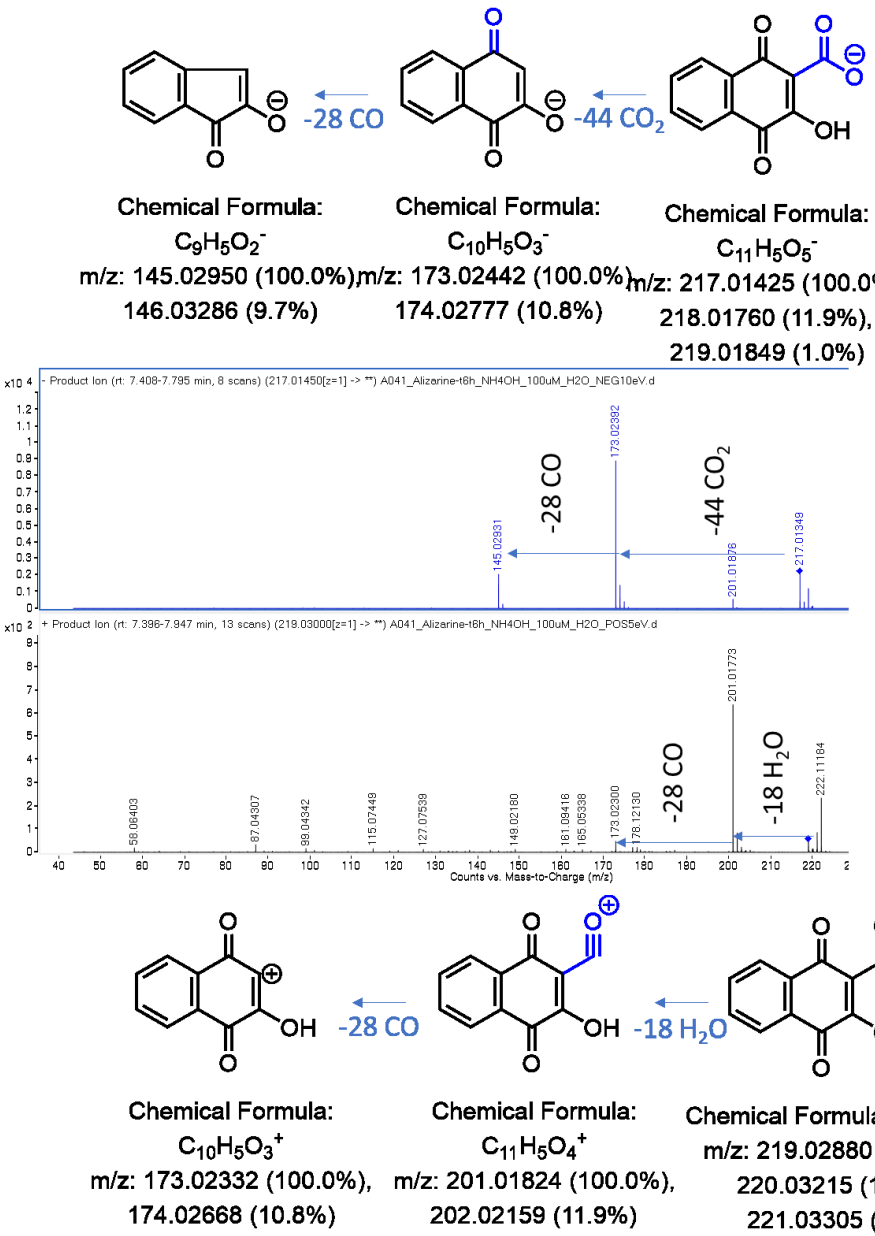
***S.I. 18**. LC-UV chromatograms at 260 nm of alizarin irradiated at 0 (black), 1h (purple), 3h (red), or 6h (pink) in water at pH=12 adjusted with NH_4OH.*



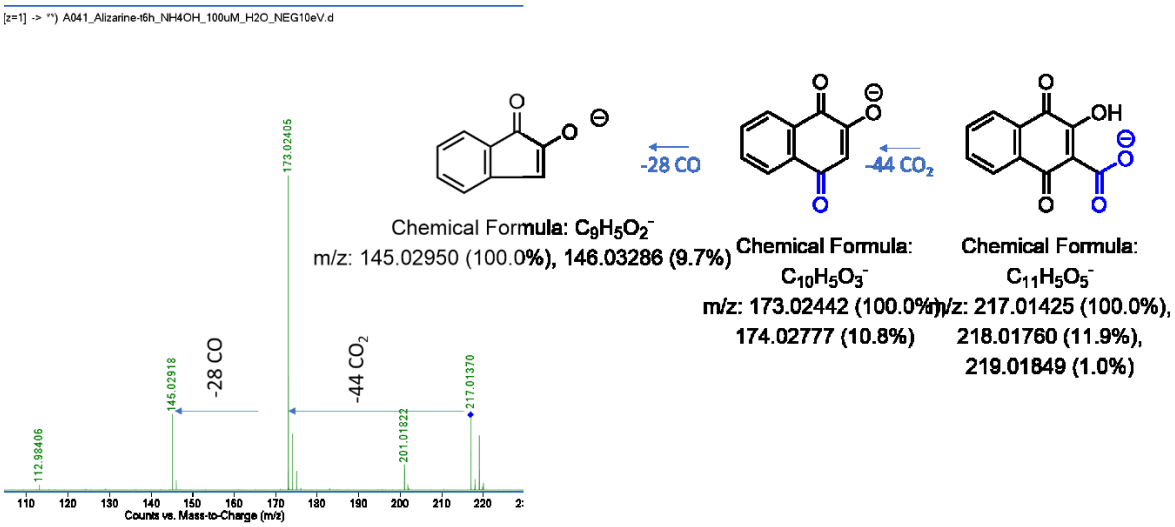
S.I. 19*. MS/MS spectra of compound 3 in negative mode (top) and positive (bottom). (molecular structure noted as P1 in main article).*



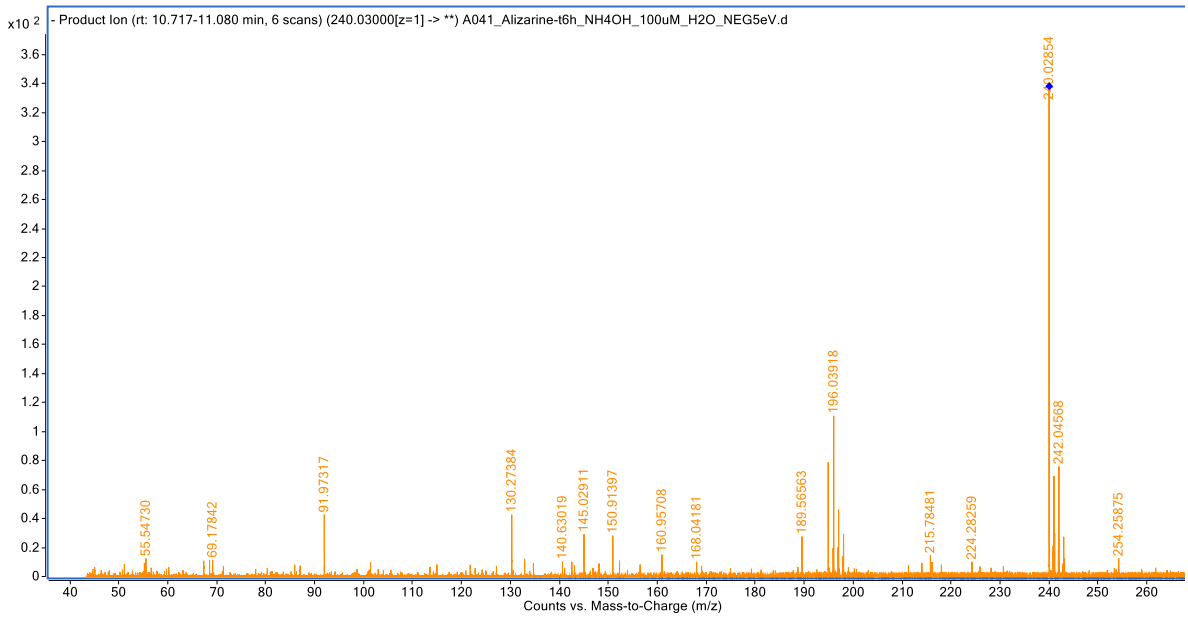
***S.I. 20**. MS/MS spectrum of compound 4 in negative mode.*



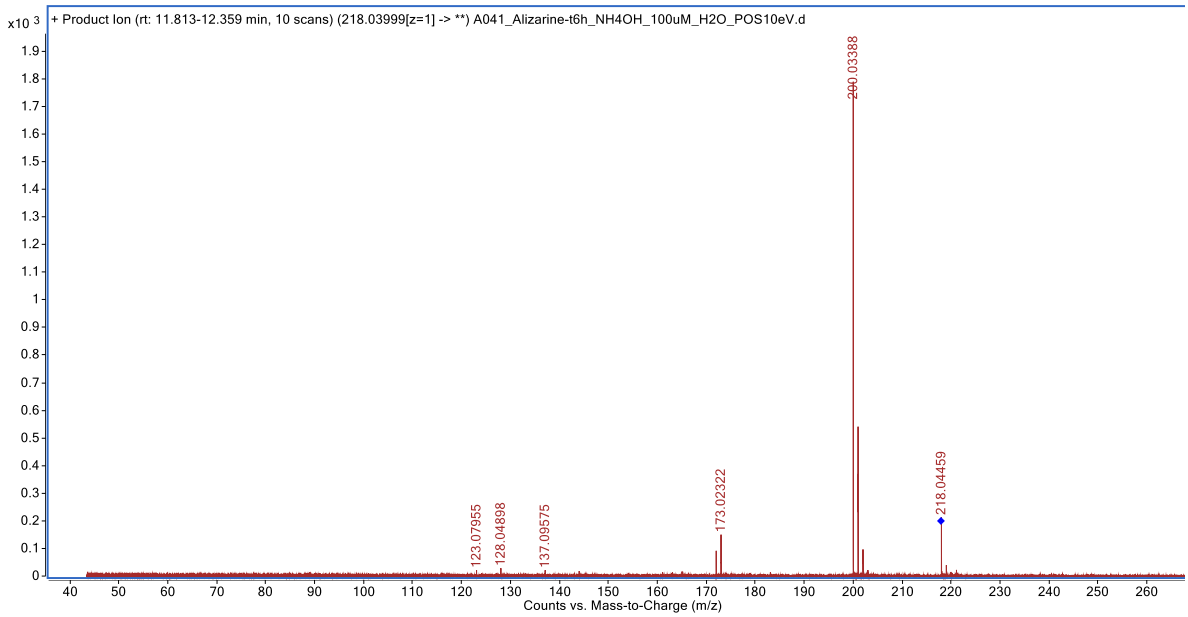
S.I. 21*. MS/MS spectra of compound 5 in negative mode (top) and positive (bottom). (molecular structure noted as P2 in main article).*



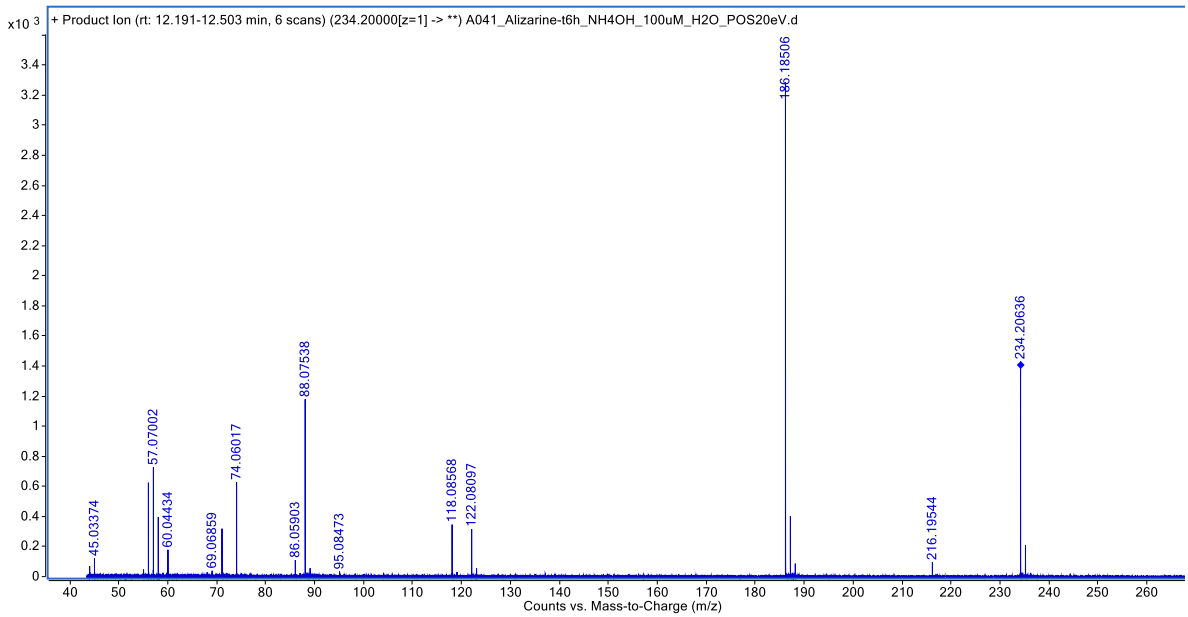
***S.I. 22**. MS/MS spectrum of compound 7 in negative mode. (molecular structure noted as P3 in main article).*



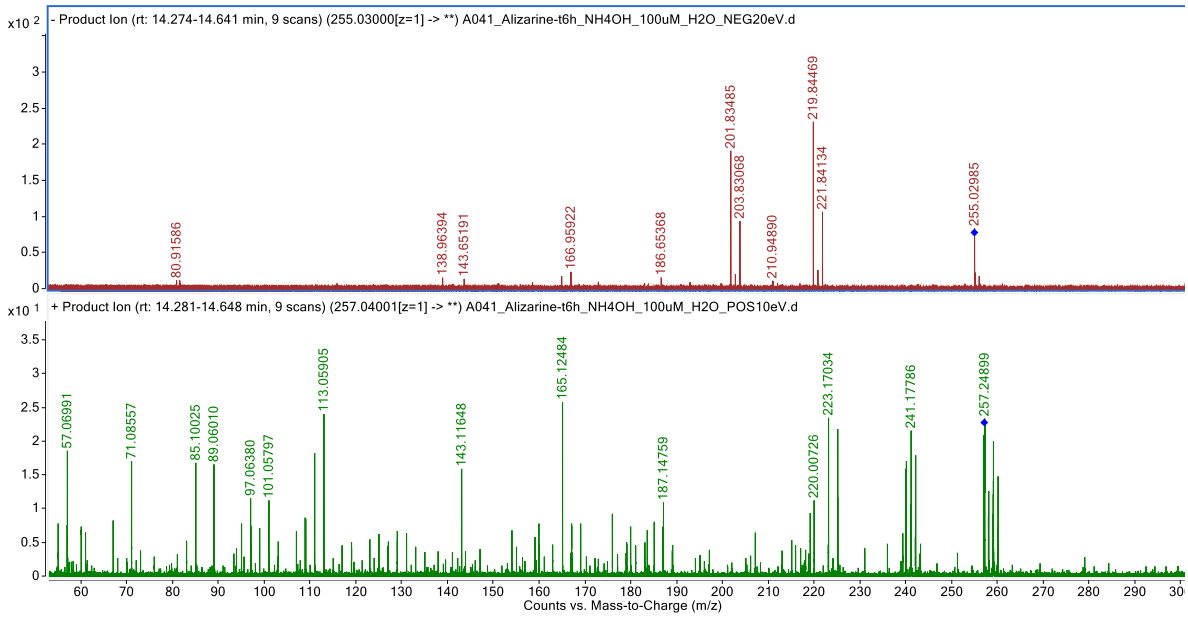
***S.I. 23**. MS/MS spectrum of compound 8 in negative mode.*



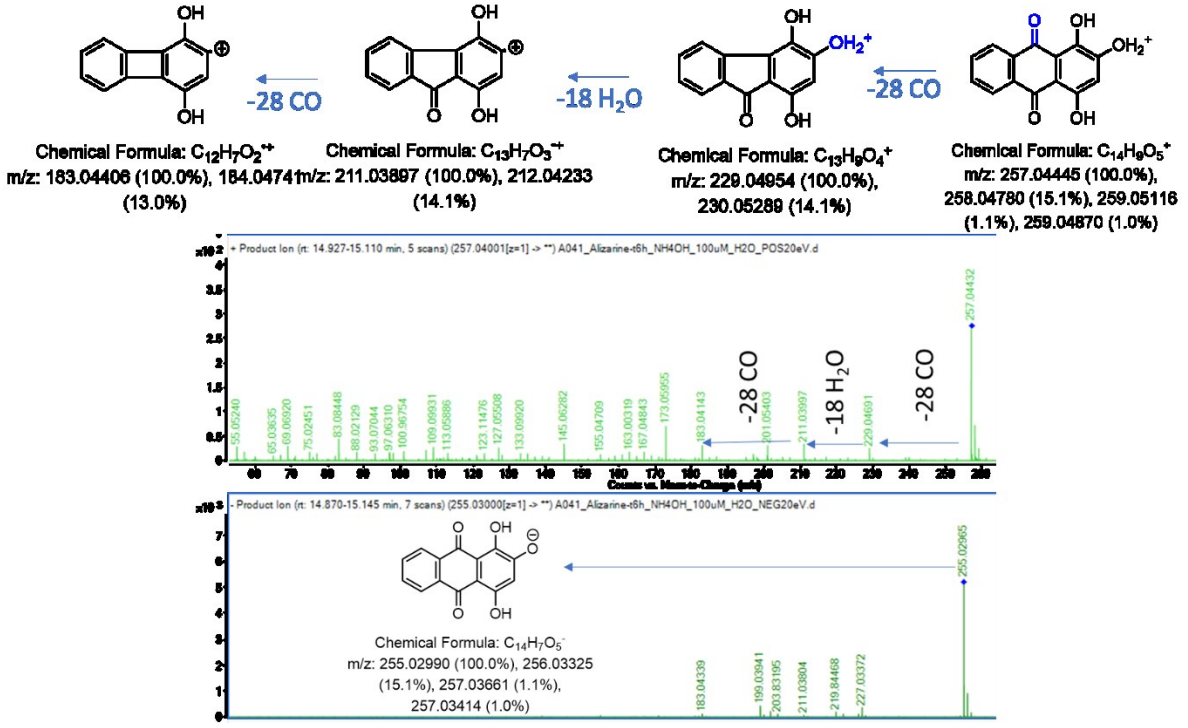
***S.I. 24**. MS/MS spectrum of compound 10 in positive mode. (molecular structure noted as P5 in main article).*



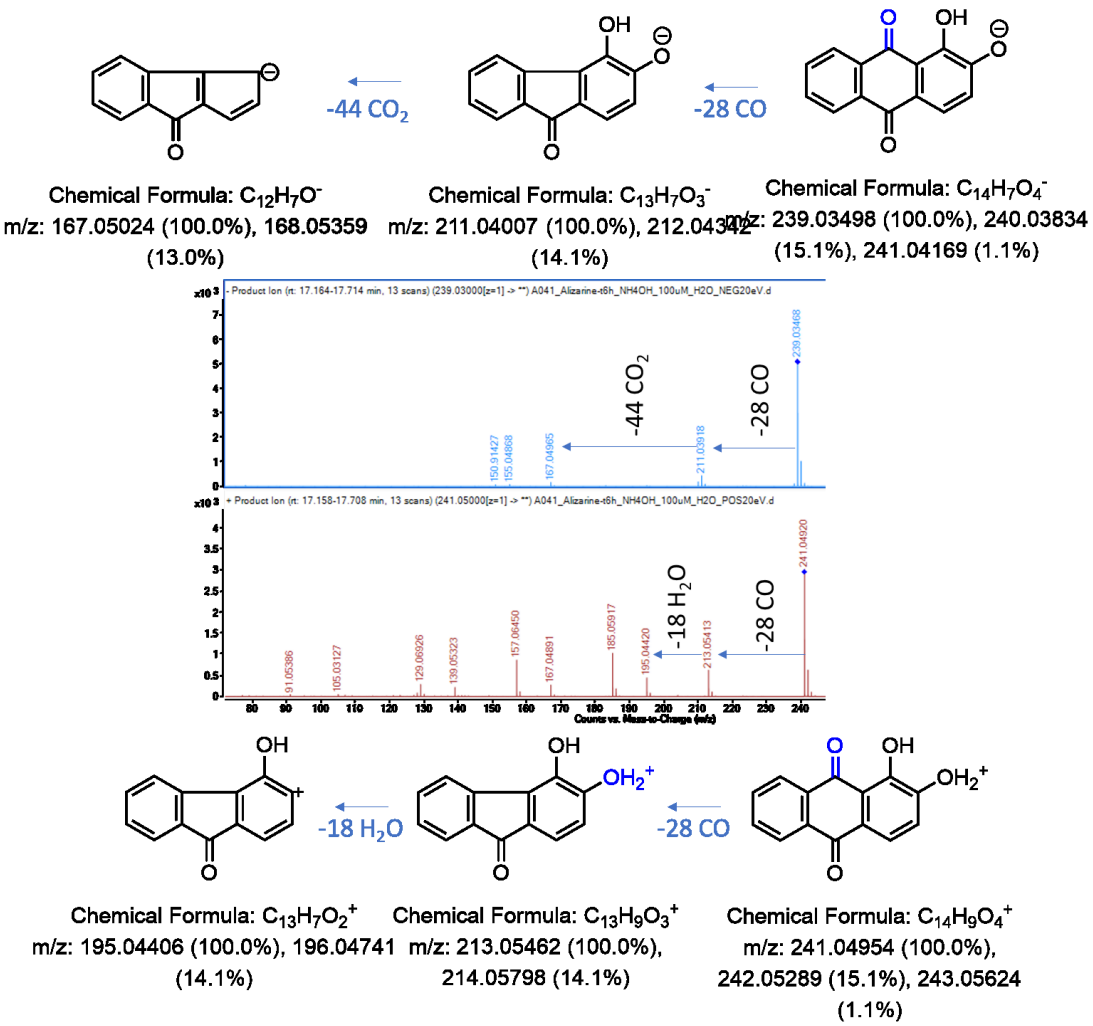
***S.I. 25**. MS/MS spectrum of compound 11 in positive mode.*



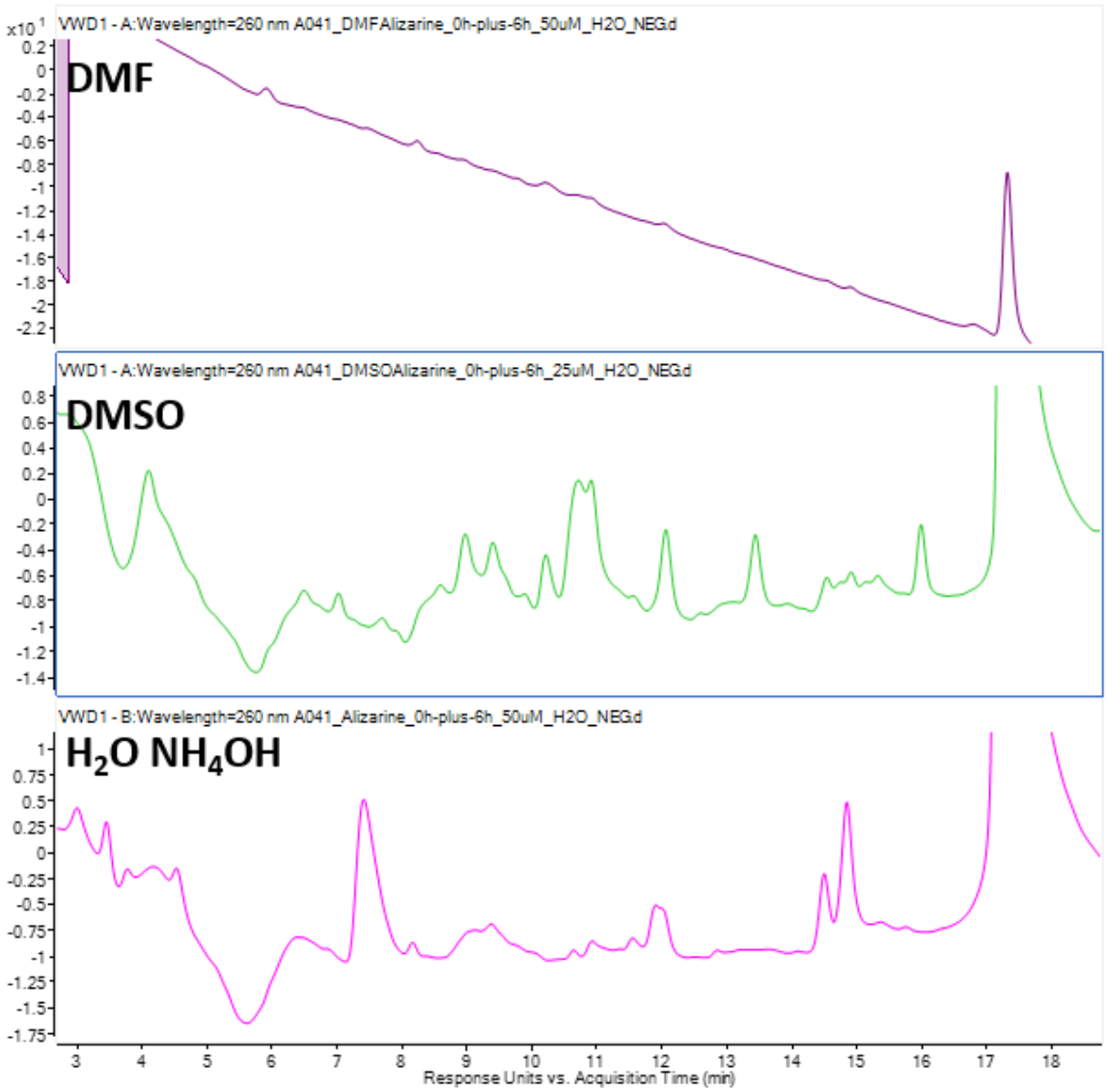
***S.I. 26**. MS/MS spectra of compound 12 in negative mode (top) and positive mode (bottom). (molecular structure noted as P5 in main article).*



***S.I. 27**. MS/MS spectra of compound 13 in positive mode (top) and negative mode (bottom). (molecular structure noted as P4 in main article).*



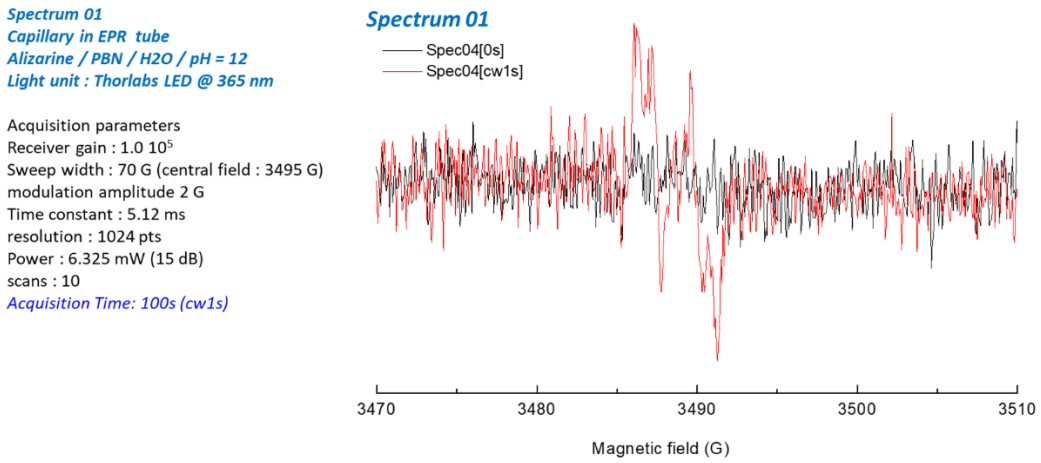
S.I. 28*. MS/MS spectra of compound 14 in negative mode (top) and positive mode (bottom). (molecular structure noted as P6 in main article).*



***S.I. 29**. The experiments were performed after a dilution by 4 in water (250µL Alizarine in DMF/DMSO without irradiation + 250 µl alizarine irradiated in DMF/DMSO + 500 µL of water). The LC-UV spectra at 260 nm shows differences in the degradation product. While DMF creates a UV signal suppression in UV, too high concentrations of DMSO leads to excessive clogging of the MS source. If we want to continue the structural investigations in these solvents, dilution will be required.*

All calculations were carried out using the version 6.0.0 of the ORCA quantum chemistry program package. The geometries were optimized using the PBE0 functional complemented by the D3(BJ) dispersion correction. Tight thresholds were employed for the self-consistent field (SCF) and optimization (tightSCF and tightopt keywords in ORCA nomenclature) . The triple zeta def2-TZVP basis set, coupled with matching auxiliary basis sets, was consistently employed throughout all the computations. The implicit water solvation using the linear response Conductor-like Polarizable Continuum Model (C-PCM) water was choosen. The charge is computed using orca keyworld print [P_Hirshfeld]. The spin densities were plotted using orca_plot and isovalue is set to 0.003 a.u.

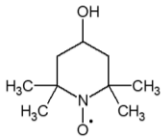
***S.I. 30**. Computation protocol.*



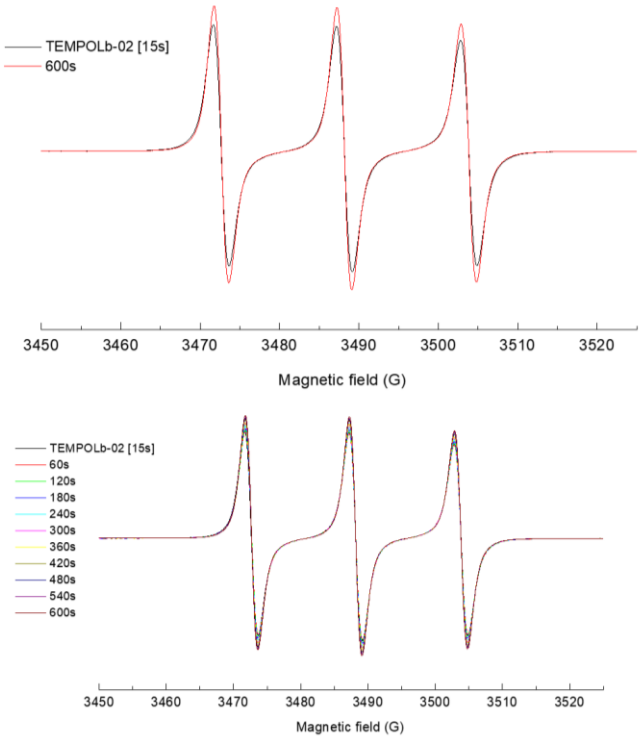
***S.I. 31**. Experimental EPR spectra obtained before (black line) and upon photoexcitation (red line, λ_{max}= 365 nm, irradiance 200 mW/cm^2) of alizarin in water at pH=12 under air. Initial concentration of the hydroxyanthraquinones was 0.5 mM (Magnetic field sweep width, SW =70 G; modulation amplitude is 2G. The measured hyperfine coupling constants, hfcc: a_1=4.2G, a_2=1.0G, g = 2.0061 ± 0.0005. [Alizarin]=0.5 mM*

Spectrum TEMPOLb-01 (02)
Capillary in EPR tube
Alizarine / TEMPOL / DMSO
Light unit : Thorlabs LED @ 365 nm

Acquisition parameters
Receiver gain : 1.0 10^3
Sweep width : 80 G (central field : 3487 G)
modulation amplitude 1 G
Time constant : 2.56 ms
resolution : 1024 pts
Power : 0.2 mW (30 dB)
scans : 3
Sweep time : 5.24s



TEMPOL : 4-Hydroxy-TEMPO, 4-Hydroxy-2,2,6,6-tetramethylpiperidine-1-oxyl
Numéro CAS: 2226-96-2
Poids moléculaire : 172.24



***S.I. 32**. Experimental EPR spectra obtained upon different photoexcitation time (λ_{max}= 365 nm) of alizarin in water at pH=12 under air, with TEMPOL. [Alizarin]=0.5 mM, [Tempol]= 0.05 mM*

$$E_{photon} = \frac{hc}{\lambda} = \frac{6.626 \times 10^{-34}\ J{\cdot}s \times 3.0 \times 10^{8}\ m/s}{520 \times 10^{-9}\ m} = 3.825 \times 10^{-19}\ J$$

$$E_{total} = 10\ \text{mW/cm}^2 = 0.01\ \text{W/cm}^2$$

$$N_{photons} = \frac{E_{total}}{E_{photon}} = \frac{0.01\ \text{W}}{3.825 \times 10^{-19}\ \text{J}} \approx 2{,}61 \times 10^{16}\ \text{photons/s}$$

$$N_{photons}\ for\ 1h\ exposition: \approx 2{,}61 \times 10^{16}\ \text{photons/s} * 3600\text{s} = 9{,}40 * 10^{19}\ \text{photons/cm}^2$$

***S.I. 33**. Example of calculation of number of photons for 10mW/cm² at λ=550 nm.*

SUMMARY

Article 5: "A computational insight into the photodegradation mechanism of alizarin "

(The Article 5 is currently under preparation and is not included in this thesis submission.)

Studying photodegradation has long engaged chemists, biologist, physicists, and curators.[40,68] The photodegradation is known to be a photochemical reaction leading to molecular changes. It is known to be initiated by an electronic excitation, which is triggered by absorption of a photon in the ultraviolet (200nm-400nm) or visible part (400nm-760nm) of the electromagnetic spectrum. [268,272,273] After photoexcitation, the molecule dissipates energy through different processes, which can be radiative with light emission (such as fluorescence, phosphorescence) or radiationless without light emission. These include photophysical processes like internal conversion and intersystem crossing, or photochemical processes leading to molecular changes (such as oxidation, reduction, photoisomerisation, etc.), as presented in Figure 1.

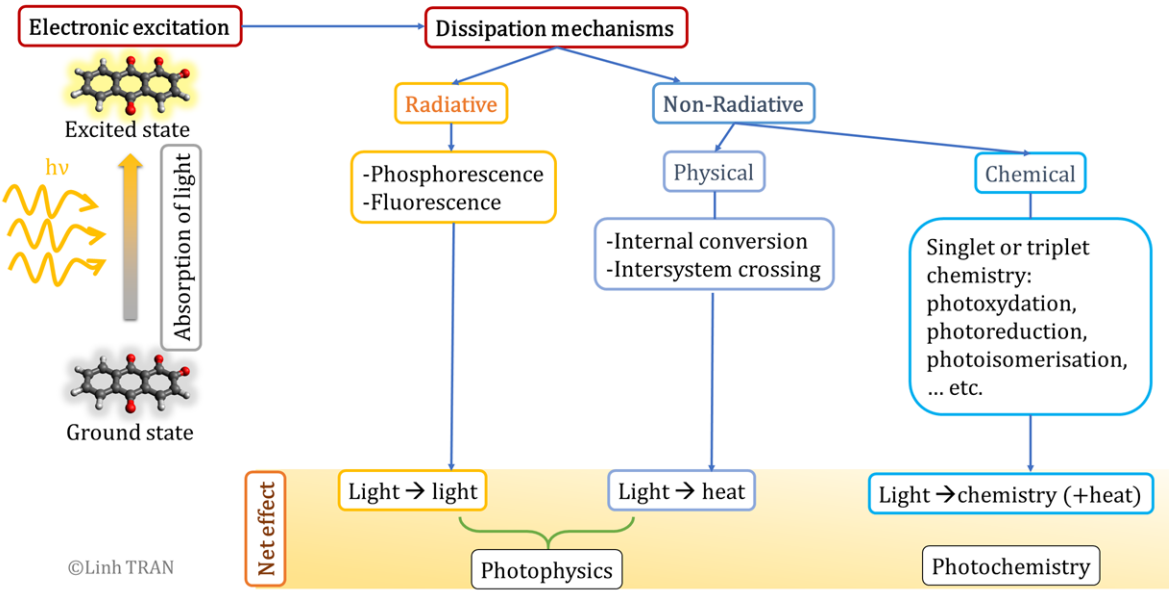


Figure 1. *Schematic representation of the main photophysical and photochemical processes that may occur following molecular excitation by light irradiation. The gained electronic energy can be dissipated by various pathways, including radiative and non-radiative processes. Radiative processes emit light such as fluorescence or phosphorescence. Non-radiative processes don't emit light, like internal conversion (IC), intersystem crossing (ISC), and photochemical pathways leading to radical formation and bond breaking.* [272]

Photochemical processes are responsible for colour changes in colourants, with the specific pathways varying depending on the type of colourant and environmental

conditions. [68] The mechanisms underlying those processes remain highly complex and only partially understood, as both theoretical approaches and experimental methods have inherent limitations in capturing the full scope of these processes. [68,272,274] Indeed, despite the availability of advanced technologies, gaining a detailed understanding of these dynamics remains complex and tedious. [272] From a theoretical perspective, a major challenge is the inherent approximations in computational methods, which become more significant as the system size increases and when dealing with heterogeneous materials such as paintings. In particular, the mechanistic study of the photodegradation of madder is scarce, although madder is known to be one of the first natural sources of red colours and has been used for thousands of years for dyeing and painting. This last article gives a theoretical insight into the photodegradation mechanism of alizarin, which photoproducts in basic media were previously identified in the article 4.

As seen in our previous work (article 1), the UV–VIS spectrum of di-deprotonated alizarin exhibits an electronic transition in the visible region, with a maximum absorption at 570 nm, mainly attributed to a $\pi \rightarrow \pi^*$ transition . [234] Two additional shoulders at 530 nm and 610 nm are observed, corresponding to vibronic coupling. As seen in the article 4, the experimental photodegradation of alizarin in solution under xenon lamp exposure (300–800 nm) was monitored by UV–VIS spectroscopy, and the resulting spectra are presented in Figure 2. Notably, the absorption features in the 530–610 nm range decrease more rapidly than the other peaks at the UV region of the spectra. This suggests the disruption of the conjugated system responsible for the $\pi \rightarrow \pi^*$ transition. Based on previously identified photoproducts (article 4) and the fact that photodegradation is influenced by oxygen [74], we propose a photodegradation mechanism with an endoperoxide intermediate, which contains a peroxide C-O-O-C bond. A review by E. L. Clennan provides a comprehensive overview of endoperoxide chemistry, including their formation mechanisms and reactivity. [275] Based on that literature, we suggest that oxygen may lead to the formation of an endoperoxide intermediate that then further decomposes to yield the photoproducts. This mechanism is further examined using computational methods to assess its reactivity, identify transition states, and estimate the associated energy barriers.

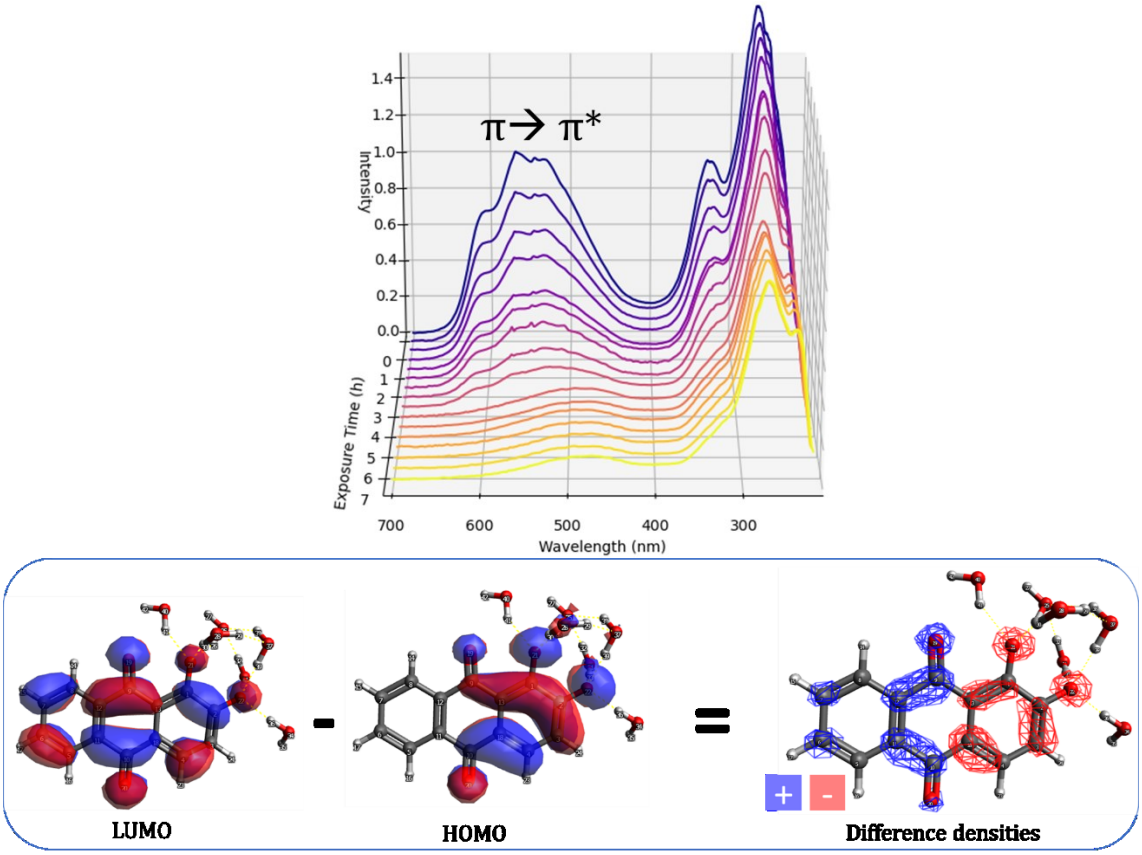


Figure 2*. Photodegradation of alizarin at pH=12.5 under xenon light exposure, monitored by UV–VIS absorbance spectra over time. Concentration is 10^{-4} mol.L^{-1}. Irradiance is 50mW/cm^2. The corresponding HOMO and LUMO and the difference densities is also depicted for the first state corresponding to the experimental transition at 570nm. Blue means positive, and red means negative. Isovalue is 0.02 a.u.*

To this end, potential energy surfaces (PES) were scanned both in the ground state and in the excited states. The minimum energy path connecting reactants to products, including key transition states, was investigated using the Nudged Elastic Band (NEB) method. NEB calculations were performed for the ground state (S0) and the first three singlet excited states (S1 to S3), using TD-DFT with nroots=4 and iroot=1 (in ORCA nomenclature), to track the S1 pathway. These results are currently under ongoing analysis. A representative result is shown in Figure 2, which displays the computed NEB pathways for both S0 and S1 to S3, alongside selected alizarin geometries along the S0 and S1 trajectories. The analysis focuses on the first singlet excited state (S1). At the initial point, the di-deprotonated alizarin molecule is not yet interacting with singlet oxygen $^1O_2$ (image 1). Progressing along the reaction path, the first transition state appears at image 5, with an energy barrier of 40 kcal/mol in the ground state and 22 kcal/mol in the S1 state. This step leads to the formation of an endoperoxide species featuring a C–O–O–C peroxide bond. Subsequently, the right-hand ring formed by carbon atoms {C13, C1–C4, C18} bends due to steric repulsion, forming an intermediate structure (image 8). Further along the NEB path, a significant energetic event is observed at image 15. It corresponds to the concerted breaking of three bonds: the O–O peroxide, the C1–C2, and the C18–C4 bonds. This geometry lies 42 kcal/mol above image 8 and marks a key multi-bond rearrangement in the photoproduct formation pathway. Notably, the energy profiles of

the S0 and S1 states converge at image 16 without touching, suggesting the possible presence of an avoided crossing, which is currently under investigation. The energy barrier for the process is consistently lower in the excited state S1 than in S0, indicating that the reaction is more likely to proceed via the photochemically accessible S1 state. This study is still in progress, with further analysis underway. A complete version of the work is expected to be submitted by the end of September.

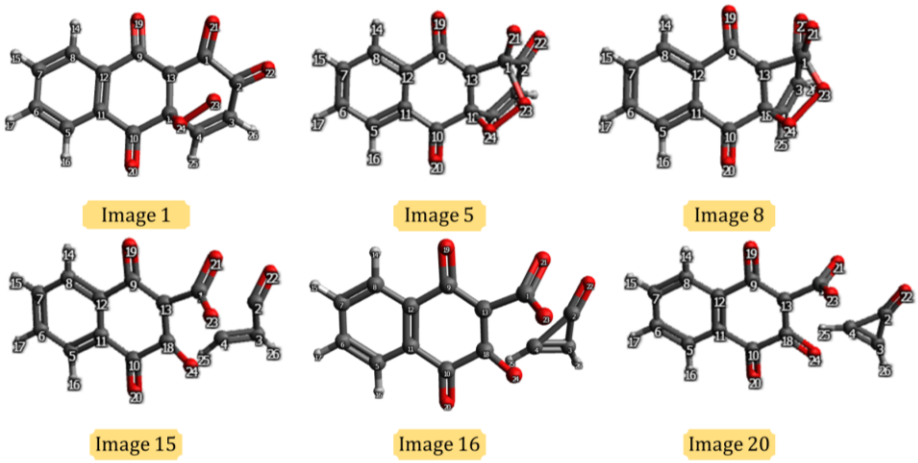
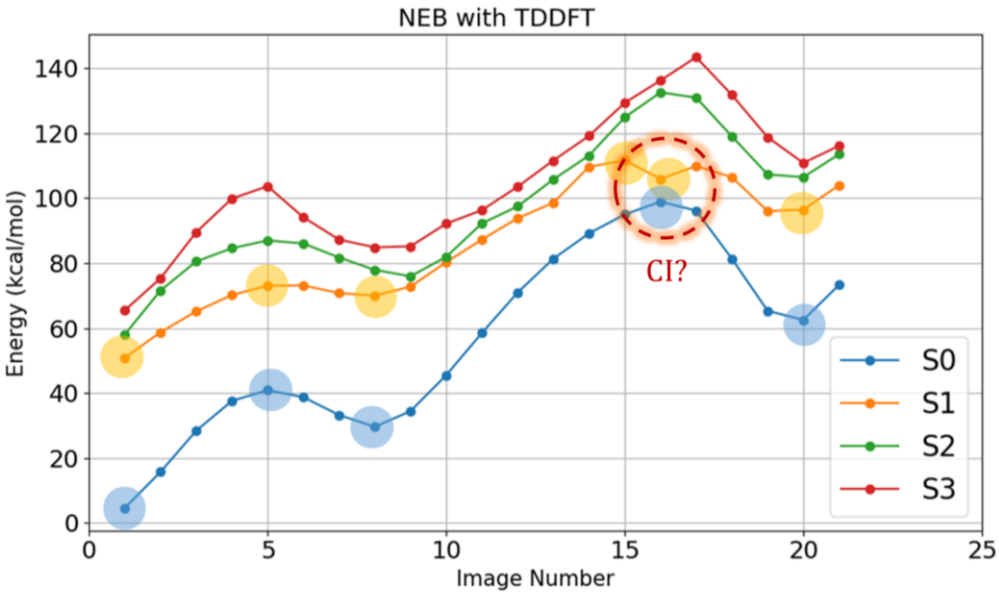


Image 1
Image 5
Image 8
Image 15
Image 16
Image 20


NEB with TDDFT
Energy (kcal/mol)
Image Number
CI?
S0
S1
S2
S3

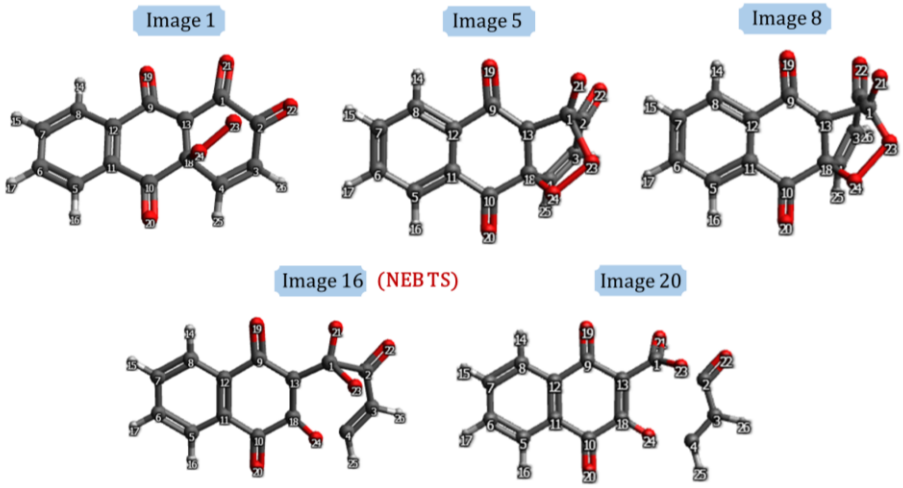


Image 1
Image 5
Image 8
Image 16 (NEB TS)
Image 20

	Relative energy in kcal/mol			
N° image	S0	S1	S2	S3
0	0.00	42.90	56.26	58.52
1	4.64	50.85	58.01	65.45
2	15.73	58.71	71.54	75.30
3	28.40	65.12	80.49	89.47
4	37.57	70.14	84.57	99.83
5	40.88	73.05	86.95	103.64
6	38.72	73.12	86.01	94.08
7	33.23	70.73	81.71	87.22
8	29.61	69.95	77.90	84.82
9	34.43	72.76	75.92	85.16
10	45.49	80.25	81.91	92.12
11	58.50	87.24	92.22	96.28
12	70.94	93.79	97.43	103.49
13	81.32	98.55	105.74	111.58
14	89.14	109.62	113.10	119.14
15	94.95	111.58	124.85	129.32
16	98.88	105.91	132.51	136.19
17	96.14	109.88	130.91	143.36
18	81.31	106.41	118.98	131.71
19	65.29	95.98	107.26	118.71
20	62.51	96.47	106.44	110.82
21	73.27	103.93	113.49	116.07

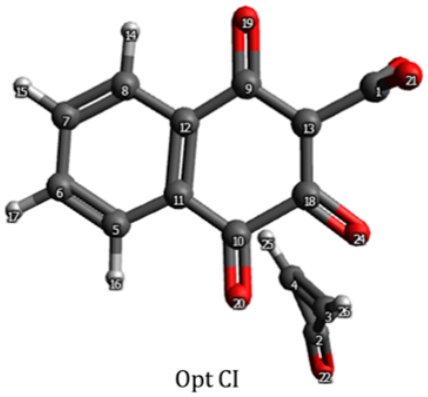
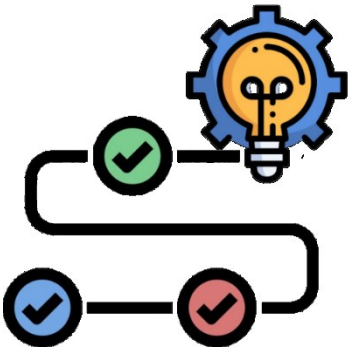


***Figure 3**. TDDFT-based NEB pathway computed with nroots = 4, considering either the ground state (iroot = 0, S0). Molecular geometries of extremums along the path are depicted for the ground state S0(blue) and first excited state S1 (yellow). The final geometry of the path was further optimised using a transition state search (OptTS), and the resulting structure is shown along with the vibrational mode corresponding to the single imaginary frequency, confirming its nature as a transition state. The table below shows the energies relative to the S0. A Conical intersection (CI) geometry optimisation was done and the resulted molecule is shown.*

CHAPTER 6

CONCLUSION AND PERSPECTIVES



CHAPTER 6 | CONCLUSION AND PERSPECTIVES

This thesis has focused on the exploration, understanding, and modelling of the optical properties and photodegradation pathways of dyes derived from madder roots (*Rubia tinctorum* L.), a historically significant complex natural dye. Used since antiquity across various artistic media, textiles, manuscripts, paintings, and decorative objects, madder roots can be considered as one of the most used dye because of its fascinating red shades. Madder has been the subject of extensive research from many points of view until today. However, its susceptibility to light-induced fading poses ongoing challenges, in particular in the preservation of cultural heritage.

In this context, our research sought to build missing puzzle pieces of madder dye photodegradation mechanisms in order to develop better conservation strategies. Like many natural plant-based colourants, madder represents a highly complex mixture composed of multiple dye molecules. Therefore, we need to first identify and build a spectroscopic database of those molecules. The challenges are that some of them are unstable, not commercially available, or challenging to synthesise, resulting in a significant lack of reference spectra for proper identification and characterisation. Furthermore, its composition varies with species, age, and processing methods. Yet, to fully comprehend how photodegradation is influenced by the irradiation wavelength, the main factor is the absorption spectrum. Indeed, some colourants experience higher degrees of photodegradation when exposed to light whose irradiance spectrum matches their absorption spectrum better. Moreover, the photodegradation process is influenced by a complex interplay of both internal and diverse external factors. These parameters are often interdependent and difficult to isolate experimentally.

To remove these gaps, the first phase of this research was therefore dedicated to establishing a computational protocol capable of predicting the optical and spectroscopic properties of individual madder dyes. A robust protocol was optimised to predict the spectroscopic (UV-Visible, NMR) and colourimetric (CIELAB colour coordinates) properties of the main madder dyes, both isolated and in complex mixtures. This protocol made it possible not only to validate predictions with accuracy comparable to experimental data, but also to access properties that are inaccessible in the laboratory for unstable, non-commercialised, or difficult-to-isolate molecules. This enabled us to build a more complete database of colours and absorbance spectra for 31 madder dyes present in 6 different species for two different protonation states. This led to the prediction of spectra of a complex mixture of dyes. This comprehensive dataset serves as a foundation for both diagnostic applications and degradation modelling, enabling researchers to better differentiate between dye components in mixtures and to infer degradation processes based on spectral evolution.

The final stage is the study of key degradation parameters, particularly the roles of pH, solvent environment, light, and oxygen on the photodegradation mechanism. Focusing on alizarin, one of the principal components of madder, and its less-studied di-deprotonated (basic) form, the research revealed novel proofs about radical degradation pathways involving a radical oxygen, and the identification of previously unreported

photodegradation products. These findings represent a significant contribution to our understanding of the molecular processes driving the deterioration of madder-based pigments in artworks.

Ultimately, the work presented in this thesis demonstrates the power and relevance of combining advanced molecular modelling with experimental methodologies in the context of heritage science. It underscores its potential not only as a predictive tool but also as a means to support the interpretation of complex optical behaviours and to overcome experimental limitations. This research paves the way for future studies on other natural or historical or natural dyes and reinforces the role of interdisciplinary strategies in the ongoing effort to preserve the colours of the past. This research contributes to the revival of natural dyes, supporting their renewed use in sustainable and ecological art and industrial applications such as textiles, cosmetics, dyeing, pollutant discharges, art, pharmaceutics, etc.

The work carried out in this thesis opens numerous promising avenues for future exploration, both computational and experimental.

PERSPECTIVES

A natural continuation would be to deepen the understanding of photodegradation mechanisms by modelling excited states and reaction pathways using advanced quantum chemical methods. Tools such as Intrinsic Reaction Coordinate (IRC) analysis, the Nudged Elastic Band (NEB) method, and Molecular Dynamics (MD) simulations, particularly in combination with ORCA or other DFT-based software, could be used to explore energy landscapes, locate conical intersections, and determine whether degradation occurs via singlet or triplet states. Spin-orbit coupling might have to be studied. Such studies (Figure 28) would provide insight into the mechanistic nature of light-induced degradation: What are the energy barriers? What reactive intermediates are formed? Is there a preferred photoreaction channel, and under which conditions?

Those studies can be completed with experimental detection of endoperoxides and their quantification. Other experiments with pulsed laser photolysis for example could make it possible to reveal the transient states of reactive intermediates regenerated on the nanosecond timescale.

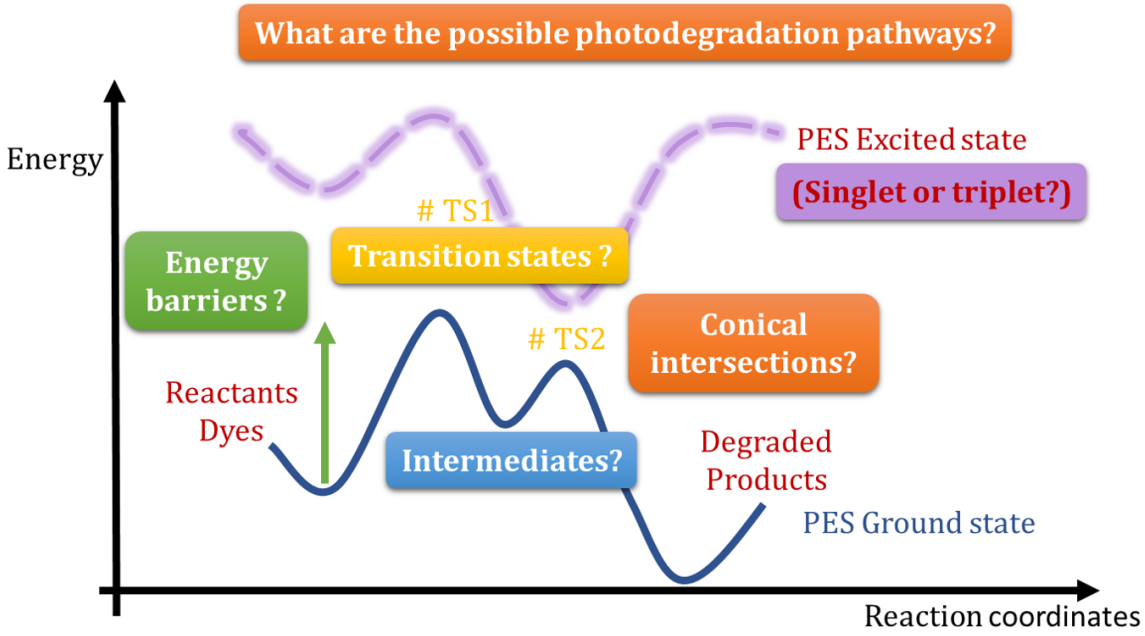


Figure 28. Some of missing questions about photodegradation mechanisms.

In parallel, the photostability of other madder-derived dyes, such as purpurin, or glycosylated compounds like ruberythric acid and lucidin primveroside, could be investigated under varying environmental conditions. By combining photodegradation kinetics with RPE and HPLC-UV-MS, the structures most susceptible to fading could be identified, expanding the understanding of dye stability and vulnerability in cultural heritage contexts.

The approach developed in this work for mixtures of dyes could also be reversed: by comparing experimental spectra with theoretically calculated spectra, it may be possible to deconvolute complex dye mixtures and estimate the relative concentrations of each species. Applying this method to mixtures containing both native dyes and their photodegradation products could help identify degradation patterns over time. This approach would provide valuable information about the chemical evolution of historical pigment formulations and offer insights into fading mechanisms and the degree of chemical transformation in heritage objects.

A major and essential direction for future research lies in the study of solid-state systems, particularly madder lake pigments, complex assemblies formed through the coordination of dyes with metal cations (mordants), as represented in Figure 29. A two-step computational strategy is envisioned: first, simulations in solution to evaluate dye-metal binding stability; then, solid-state to elucidate pigment structure and optical properties. The initial solution-phase calculations could identify the most stable configurations of madder dyes complexed with various cations, beyond alum (Al^{3+}), cations like Cu^{2+}, Fe^{2+}, Fe^{3+}, and Mg^{2+} could be explored. Different coordination geometries (monodentate, bidentate), dye-to-cation ratios, and dye types (e.g., alizarin, purpurin) could be tested. This would help clarify whether lake pigments form via complexation followed by

precipitation or direct co-precipitation, and how pH and mordant type influence the final structure. Understanding these formation mechanisms is vital to predicting the long-term stability of these pigments. It could also help identify the dyeing mechanism on wool: how does the dye complex with mordant on wools? How does it affect photodegradation?

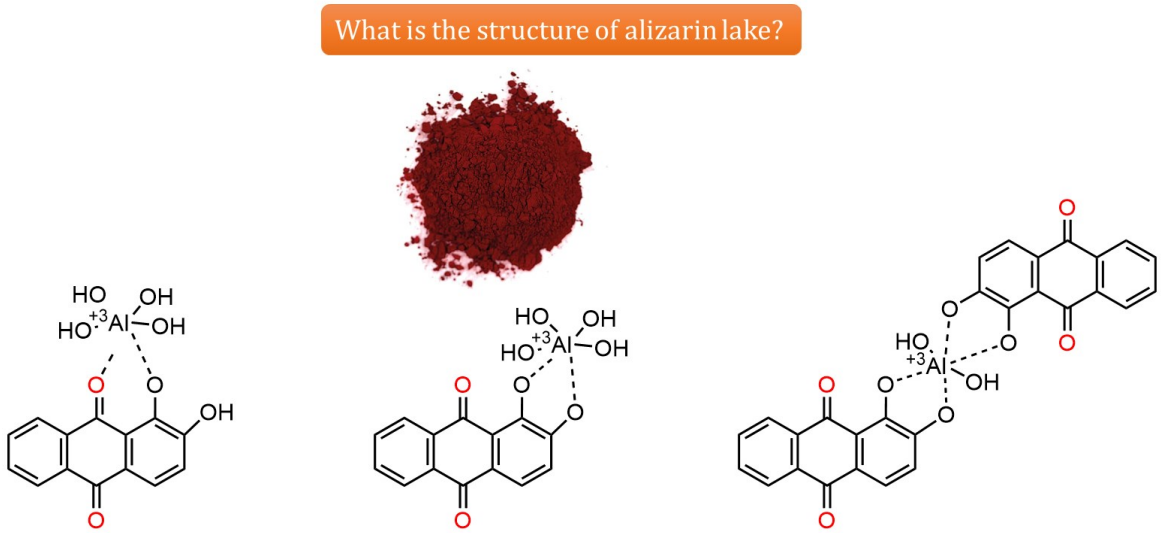


Figure 29. What is the structure of madder lake?

Once the pigment structure is proposed, it could be used to simulate key properties such as colour reflectance spectra, IR and Raman spectra, and even solid-state NMR. To model amorphous or semi-ordered solids, a combination of Monte Carlo simulations, periodic DFT, and molecular dynamics could be employed using software such as CP2K, VASP, or ORCA. These simulations would be instrumental in interpreting experimental data and locating chemically fragile bonds likely to be affected during photodegradation. Experimental validation could involve synthesising model pigments, exposing them to controlled lighting, and recording their spectra over time. A current challenge remains the heterogeneous nature of photodegradation in solids, which complicates the isolation of degraded regions. Developing strategies to spatially resolve degradation within samples will be crucial.

In parallel, the spectral database developed in this work could be expanded and shared with the research community, serving as a reference for the non-invasive identification of natural dyes in historical artworks.

Another key perspective involves the reaction mechanisms in solid phases, now made possible by the identification of degradation products. This approach could be extended to other dye systems, both organic and inorganic, broadening its impact beyond madder.

From an interdisciplinary standpoint, the computational-experimental protocol developed here could be applied to other domains where bio-based dyes are used, such as cosmetics, textiles, the food industry, and pharmaceuticals. It offers a framework for

sustainable development, reduction of experimental testing, and rational formulation of stable, eco-friendly colourants.

Finally, this research contributes to the revival and valorisation of dye plants. In a time when ancestral know-how is being rediscovered, such tools can guide the selection of species and processing methods that ensure both aesthetic appeal and chemical resilience, while also honouring their cultural heritage.



REFERENCES

(1) Pastoureau, M. *Red: The History of a Color*; Princeton University Press, 2017.
(2) Williams, T. N. Mad about Madder. *Nature Chemistry* **2022**, *14*, 1–1. https://doi.org/10.1038/s41557-022-01015-x.
(3) *The Chemistry and Application of Dyes*; Waring, D. R., Hallas, G., Eds.; Springer US: Boston, MA, 1990. https://doi.org/10.1007/978-1-4684-7715-3.
(4) Pastoureau, M. *The Colours of Our Memories*; John Wiley & Sons, 2020.
(5) Chenciner, R. *Madder Red: A History of Luxury and Trade*; Routledge, 2003.
(6) Silva, L. C.; Otero, V.; Melo, M. J.; Cabrita, E. J.; Mafra, L. What's the Madder? Characterization of Old Fashioned Alizarin/Aluminum Red Pigments Using Liquid and Solid-State NMR. *Colorants* **2023**, *2* (4), 601–617. https://doi.org/10.3390/colorants2040031.
(7) Chavanne, C.; Verney, A.; Paquier-Berthelot, C.; Bostal, M.; Buléon, P.; Walter, P. Bayeux Tapestry: First Use of Early Synthetic Dyes for the Restoration of a Masterpiece. *Dyes and Pigments* **2023**, *208*, 110798. https://doi.org/10.1016/j.dyepig.2022.110798.
(8) Taburet-Delahaye, É.; Déjean, R.; Reyer, D. de; Nowik, W. La Dame à la licorne, sa conservation et l'évaluation colorimétrique du nettoyage. *Technè. La science au service de l'histoire de l'art et de la préservation des biens culturels* **2015**, No. 41, 86–93. https://doi.org/10.4000/techne.4758.
(9) *Madder Game Piece | The Medieval Garden Enclosed | The Metropolitan Museum of Art, New York*. https://blog.metmuseum.org/cloistersgardens/2013/03/08/madder-red/game-piece1/ (accessed 2025-06-18).
(10) Potts, D. T. On the History of Madder (Rubia Peregrina L., and Rubia Tinctorum L.) in Pre-Modern Iran and the Caucasus. *Asiatische Studien - Études Asiatiques* **2022**, *76* (4), 785–819. https://doi.org/10.1515/asia-2021-0039.
(11) Kakoulli, I.; Radpour, R.; Lin, Y.; Svoboda, M.; Fischer, C. Application of Forensic Photography for the Detection and Mapping of Egyptian Blue and Madder Lake in Hellenistic Polychrome Terracottas Based on Their Photophysical Properties. *Dyes and Pigments* **2017**, *136*, 104–115. https://doi.org/10.1016/j.dyepig.2016.08.030.
(12) *Madder | Description, Species, & Uses | Britannica*. https://www.britannica.com/plant/madder (accessed 2025-03-13).
(13) Humans, I. W. G. on the E. of C. R. to. Rubia Tinctorum, Morinda Officinalis and Anthraquinones. In *Some Traditional Herbal Medicines, Some Mycotoxins, Naphthalene and Styrene*; International Agency for Research on Cancer, 2002.
(14) Schorr, S.; Aviad, I.; Laufer, A. Vital Staining with Alizarin in Clinical Malignant Conditions of Bone. *Radiology* **1959**, *73* (3), 410–417. https://doi.org/10.1148/73.3.410.
(15) Ford, L. L. Chemical Analysis and Elucidation of Anthraquinone and Flavonoid Type Compounds with Applications to Historical Artefacts and Sustainability, University of Leeds, University of Leeds, 2017.
(16) A Colourful Past. *Nature* **2006**, *440* (7083), 384–384. https://doi.org/10.1038/440384a.
(17) Dutta, P.; Mahjebin, S.; Sufian, M. A.; Razaya Rabbi, Md.; Chowdhury, S.; Imran, I. H. Impacts of Natural and Synthetic Mordants on Cotton Knit Fabric Dyed with Natural

Dye from Onion Skin in Perspective of Eco-Friendly Textile Process. *Materials Today: Proceedings* **2021**, *47*, 2633–2640. https://doi.org/10.1016/j.matpr.2021.05.229.
(18) Ackacha, M. A.; Połeć-Pawlak, K.; Jarosz, M. Identification of Anthraquinone Coloring Matters in Natural Red Dyestuffs by High Performance Liquid Chromatography with Ultraviolet and Electrospray Mass Spectrometric Detection. *Journal of Separation Science* **2003**, *26* (11), 1028–1034. https://doi.org/10.1002/jssc.200301484.
(19) Blackburn, R. S. Natural Dyes in Madder (*Rubia* Spp.) and Their Extraction and Analysis in Historical Textiles. *Coloration Technol* **2017**, *133* (6), 449–462. https://doi.org/10.1111/cote.12308.
(20) Derksen, G. C. H. Red, Redder, Madder : Analysis and Isolation of Anthraquinones from Madder Roots (Rubia Tinctorum), Wageningen Universiteit, University in Wageningen, Netherlands, 2001. https://research.wur.nl/en/publications/red-redder-madder-analysis-and-isolation-of-anthraquinones-from-m (accessed 2024-02-29).
(21) Lajkó, E.; Bányai, P.; Zámbó, Z.; Kursinszki, L.; Szőke, É.; Kőhidai, L. Targeted Tumor Therapy by Rubia Tinctorum L.: Analytical Characterization of Hydroxyanthraquinones and Investigation of Their Selective Cytotoxic, Adhesion and Migration Modulator Effects on Melanoma Cell Lines (A2058 and HT168-M1). *Cancer Cell International* **2015**, *15* (1), 119. https://doi.org/10.1186/s12935-015-0271-4.
(22) Baroni, A.; Comite, V.; Guglielmi, V.; Casanova, M.; Redegalli, P.; Fermo, P. Identification of Natural Dyes in Historical Tapestries, a LC-MS/MS Approach. *J. Phys.: Conf. Ser.* **2022**, *2204* (1), 012065. https://doi.org/10.1088/1742-6596/2204/1/012065.
(23) Wenger, S. R.; D'Alessandro, D. M. Aqueous Electrochemical Direct Air Capture Using Alizarin Red S. *ChemSusChem* **2025**, *18* (3), e202401315. https://doi.org/10.1002/cssc.202401315.
(24) Zinck, N.; Franz-Odendaal, T. A. Accurate Whole-Mount Bone and Cartilage Staining Requires Acid-Free Conditions. *The Anatomical Record* **2021**, *304* (5), 958–960. https://doi.org/10.1002/ar.24526.
(25) Aslan, F. New Natural Dyes Extracted by Ultrasonic and Soxhlet Method: Effect on Dye-Sensitized Solar Cell Photovoltaic Performance. *Opt Quant Electron* **2024**, *56* (4), 645. https://doi.org/10.1007/s11082-024-06294-x.
(26) Gürses, A.; Açıkyıldız, M.; Güneş, K.; Gürses, M. S. Classification of Dye and Pigments. In *Dyes and Pigments*; SpringerBriefs in Molecular Science; Springer International Publishing: Cham, 2016; pp 31–45. https://doi.org/10.1007/978-3-319-33892-7_3.
(27) texte, B. l'Anglais (12-1272) A. du; Corbichon, J. T.; Enlumineur, E. d'Espinque (14-14 ?). Barthélémy l'Anglais, Livre des propriétés des choses, traduit du latin par Jean Corbechon, 1479. https://gallica.bnf.fr/ark:/12148/btv1b10532588f (accessed 2025-06-19).
(28) Hellot, J. (1685-1766) A. du texte. *L'art de la teinture des laines et des étoffes de laine en grand et petit teint , avec une instruction sur les déboüillis, par M. Hellot,...*; 1750.
(29) Sanyova, J. CONTRIBUTION À L'ÉTUDE DE LA STRUCTURE ET DES PROPRIÉTÉS DES LAQUES DE GARANCE. Thesis, UNIVERSITÉ LIBRE DE BRUXELLES FACULTÉ DES SCIENCES APPLIQUÉES SERVICE DE CHIMIE ORGANIQUE, Bruxelles, 2001.
(30) Veiga, T.; Moro, A. J.; Nabais, P.; Vilarigues, M.; Otero, V. A First Approach to the Study of Winsor & Newton's 19th-Century Manufacture of Madder Red Lake Pigments. *Heritage* **2023**, *6* (4), 3606–3621. https://doi.org/10.3390/heritage6040192.

(31) Carlyle, L. *The Artist's Assistant: Oil Painting Instruction Manuals and Handbooks in Britain 1800-1900, with Reference to Selected Eighteenth-Century Sources*; Archetype Publications: London, 2001.
(32) Derksen, G. C. H.; van Holthoon, F. L.; Willemen, H. M.; Krul, C. A. M.; Franssen, M. C. R.; van Beek, T. A. Development of a Process for Obtaining Non-Mutagenic Madder Root (*Rubia Tinctorum*) Extract for Textile Dyeing. *Industrial Crops and Products* **2021**, *164*, 113344. https://doi.org/10.1016/j.indcrop.2021.113344.
(33) Blackburn, R. S. Natural Dyes in Madder (Rubia Spp.) and Their Extraction and Analysis in Historical Textiles. *Coloration Technology* **2017**, *133* (6), 449–462. https://doi.org/10.1111/cote.12308.
(34) Ford, L.; Rayner, C. M.; Blackburn, R. S. Degradation of Lucidin: New Insights into the Fate of This Natural Pigment Present in Dyer's Madder (Rubia Tinctorum L.) during the Extraction of Textile Artefacts. *Dyes and Pigments* **2018**, *154*, 290–295. https://doi.org/10.1016/j.dyepig.2018.03.023.
(35) Mouri, C.; Laursen, R. Identification of Anthraquinone Markers for Distinguishing Rubia Species in Madder-Dyed Textiles by HPLC. *Microchim Acta* **2012**, *179* (1–2), 105–113. https://doi.org/10.1007/s00604-012-0868-4.
(36) Karadag, R.; Torgan, E.; Yurdun, T. Formation and HPLC Analysis of the Natural Lake Pigment Obtained from Madder (Rubia Tinctorum L.). *Reviews in Analytical Chemistry* **2010**, *29* (1), 1–12. https://doi.org/10.1515/REVAC.2010.29.1.1.
(37) Rafaelly, L.; Heron, S.; Nowik, W.; Tchapla, A. Optimisation of ESI-MS Detection for the HPLC of Anthraquinone Dyes. *Dyes and Pigments* **2008**, *77* (1), 191–203. https://doi.org/10.1016/j.dyepig.2007.05.007.
(38) Degano, I.; Tognotti, P.; Kunzelman, D.; Modugno, F. HPLC-DAD and HPLC-ESI-Q-ToF Characterisation of Early 20th Century Lake and Organic Pigments from Lefranc Archives. *Herit Sci* **2017**, *5* (1), 7. https://doi.org/10.1186/s40494-017-0120-y.
(39) Chavanne, C.; Troalen, L. G.; Fronty, I. B.; Buléon, P.; Walter, P. Noninvasive Characterization and Quantification of Anthraquinones in Dyed Woolen Threads by Visible Diffuse Reflectance Spectroscopy. *Anal. Chem.* **2022**, *94* (21), 7674–7682. https://doi.org/10.1021/acs.analchem.2c01073.
(40) Miliani, C.; Monico, L.; Melo, M. J.; Fantacci, S.; Angelin, E. M.; Romani, A.; Janssens, K. Photochemistry of Artists' Dyes and Pigments: Towards Better Understanding and Prevention of Colour Change in Works of Art. *Angew. Chem. Int. Ed.* **2018**, *57* (25), 7324–7334. https://doi.org/10.1002/anie.201802801.
(41) Vanmeert, F.; Van der Snickt, G.; Janssens, K. Plumbonacrite Identified by X-Ray Powder Diffraction Tomography as a Missing Link during Degradation of Red Lead in a Van Gogh Painting. *Angew. Chem. Int. Ed.* **2015**, *54* (12), 3607–3610. https://doi.org/10.1002/anie.201411691.
(42) Cruz Mojica, K.; León Santiago, M.; Varela García, E.; Reyes Lezama, M.; Tapia Mendoza, E. New Progress in the Extraction and Identification of Mexican Cochineal (Dactylopius Coccus) in Fresh Dyed Wool and Dyed Wool with Artificially Accelerated Aging and Color Description. *Color Research & Application* **2023**, *48* (1), 151–164. https://doi.org/10.1002/col.22828.
(43) Sultana, N.; Gunning, S.; Furst, S. J.; Garrard, K. P.; Dow, T. A.; Vinueza, N. R. Direct Analysis of Textile Dyes from Trace Fibers by Automated Microfluidics Extraction System Coupled with Q-TOF Mass Spectrometer for Forensic Applications. *Forensic Science International* **2018**, *289*, 67–74. https://doi.org/10.1016/j.forsciint.2018.05.020.

(44) Ford, L.; Henderson, R. L.; Rayner, C. M.; Blackburn, R. S. Mild Extraction Methods Using Aqueous Glucose Solution for the Analysis of Natural Dyes in Textile Artefacts Dyed with Dyer's Madder (Rubia Tinctorum L.). *Journal of Chromatography A* **2017**, *1487*, 36–46. https://doi.org/10.1016/j.chroma.2017.01.053.
(45) De Viguerie, L.; Michelin, A.; Radepont, M.; Pottier, F.; Alfeld, M.; Walter, P.; Glanville, H. Imagerie hyperspectrale pour l'analyse d'œuvres peintes. In *Instrumentation portable*; Benech, C., Cantin, N., Languille, M.-A., Mazuy, A., Robinet, L., Zazzo, A., Eds.; Editions des archives contemporaines, 2019; pp 43–65. https://doi.org/10.17184/eac.2457.
(46) Zhuang, G.; Pedetti, S.; Bourlier, Y.; Jonnard, P.; Méthivier, C.; Walter, P.; Pradier, C.-M.; Jaber, M. New Insights into the Structure and Degradation of Alizarin Lake Pigments: Input of the Surface Study Approach. *J. Phys. Chem. C* **2020**, *124* (23), 12370–12380. https://doi.org/10.1021/acs.jpcc.0c00746.
(47) Amat, A.; Miliani, C.; Romani, A.; Fantacci, S. DFT/TDDFT Investigation on the UV-Vis Absorption and Fluorescence Properties of Alizarin Dye. *Phys. Chem. Chem. Phys.* **2015**, *17* (9), 6374–6382. https://doi.org/10.1039/C4CP04728A.
(48) Tissier, R.-C.; Rigaud, B.; Thureau, P.; Huix-Rotllant, M.; Jaber, M.; Ferré, N. Stressing the Differences in Alizarin and Purpurin Dyes through UV-Visible Light Absorption and ^{1}H-NMR Spectroscopies. *Phys. Chem. Chem. Phys.* **2022**, *24* (32), 19452–19462. https://doi.org/10.1039/D2CP00520D.
(49) Pagliai, M.; Osticioli, I.; Nevin, A.; Siano, S.; Cardini, G.; Schettino, V. DFT Calculations of the IR and Raman Spectra of Anthraquinone Dyes and Lakes. *Journal of Raman Spectroscopy* **2018**, *49* (4), 668–683. https://doi.org/10.1002/jrs.5334.
(50) Lofrumento, C.; Platania, E.; Ricci, M.; Becucci, M.; Castellucci, E. M. SERS Spectra of Alizarin Anion–Ag_n ($n = 2, 4, 14$) Systems: TDDFT Calculation and Comparison with Experiment. *J. Phys. Chem. C* **2016**, *120* (22), 12234–12241. https://doi.org/10.1021/acs.jpcc.5b12321.
(51) Mannino, M. R.; Orecchio, S.; Gennaro, G. Microanalytical Method for Studying Paintings by Use of Fluorescence Spectroscopy Combined with Principal Component Analysis. *Microchemical Journal* **2013**, *110*, 407–416. https://doi.org/10.1016/j.microc.2013.05.014.
(52) Rambaldi, D. C.; Pozzi, F.; Shibayama, N.; Leona, M.; Preusser, F. D. Surface-enhanced Raman Spectroscopy of Various Madder Species on Wool Fibers: The Role of Pseudopurpurin in the Interpretation of the Spectra. *J Raman Spectroscopy* **2015**, *46* (11), 1073–1081. https://doi.org/10.1002/jrs.4726.
(53) Whitney, A. V.; Van Duyne, R. P.; Casadio, F. An Innovative Surface-Enhanced Raman Spectroscopy (SERS) Method for the Identification of Six Historical Red Lakes and Dyestuffs. *J. Raman Spectrosc.* **2006**, *37* (10), 993–1002. https://doi.org/10.1002/jrs.1576.
(54) Casadio, F.; Leona, M.; Lombardi, J. R.; Van Duyne, R. Identification of Organic Colorants in Fibers, Paints, and Glazes by Surface Enhanced Raman Spectroscopy. *Acc. Chem. Res.* **2010**, *43* (6), 782–791. https://doi.org/10.1021/ar100019q.
(55) Lombardi, L.; Serafini, I.; Guiso, M.; Sciubba, F.; Bianco, A. A New Approach to the Mild Extraction of Madder Dyes from Lake and Textile. *Microchemical Journal* **2016**, *126*, 373–380. https://doi.org/10.1016/j.microc.2015.12.021.
(56) Henderson, R. L.; Rayner, C. M.; Blackburn, R. S. Isolation and Extraction of Lucidin Primeveroside from Rubia Tinctorum L. and Crystal Structure Elucidation. *Phytochemistry* **2013**, *95*, 105–108. https://doi.org/10.1016/j.phytochem.2013.07.001.

(57) Cuoco, G.; Mathe, C.; Archier, P.; Chemat, F.; Vieillescazes, C. A Multivariate Study of the Performance of an Ultrasound-Assisted Madder Dyes Extraction and Characterization by Liquid Chromatography-Photodiode Array Detection. *Ultrasonics Sonochemistry* **2009**, *16* (1), 75–82. https://doi.org/10.1016/j.ultsonch.2008.05.014.
(58) Boldizsár, I.; Szűcs, Z.; Füzfai, Zs.; Molnár-Perl, I. Identification and Quantification of the Constituents of Madder Root by Gas Chromatography and High-Performance Liquid Chromatography. *Journal of Chromatography A* **2006**, *1133* (1–2), 259–274. https://doi.org/10.1016/j.chroma.2006.08.021.
(59) Cuoco, G.; Mathe, C.; Archier, P.; Vieillescazes, C. Characterization of Madder and Garancine in Historic French Red Materials by Liquid Chromatography-Photodiode Array Detection. *Journal of Cultural Heritage* **2011**, *12* (1), 98–104. https://doi.org/10.1016/j.culher.2010.05.005.
(60) Blackburn, R. S. Natural Dyes in Madder (Rubia Spp.) and Their Extraction and Analysis in Historical Textiles. *Coloration Technology* **2017**, *133* (6), 449–462. https://doi.org/10.1111/cote.12308.
(61) Murthy, H. N.; Joseph, K. S.; Paek, K. Y.; Park, S. Y. Anthraquinone Production from Cell and Organ Cultures of Rubia Species: An Overview. *Metabolites* **2022**, *13* (1), 39. https://doi.org/10.3390/metabo13010039.
(62) *Van Gogh's Fading Colors Inspire Scientific Inquiry*. Chemical & Engineering News. https://cen.acs.org/articles/94/i5/Van-Goghs-Fading-Colors-Inspire.html (accessed 2023-07-05).
(63) Homer, W. *For to Be a Farmer's Boy*; 1887. https://www.artic.edu/artworks/93433/for-to-be-a-farmer-s-boy (accessed 2025-03-10).
(64) Derrick, M.; Newman ,Richard; and Wright, J. Characterization of Yellow and Red Natural Organic Colorants on Japanese Woodblock Prints by EEM Fluorescence Spectroscopy. *Journal of the American Institute for Conservation* **2017**, *56* (3–4), 171–193. https://doi.org/10.1080/01971360.2016.1275438.
(65) Hendriks, E.; Brokerhof, A. W. Valuing Van Gogh's Colours: From the Past to the Future. **2017**.
(66) Khan, S.; Noor, T.; Iqbal, N.; Yaqoob, L. Photocatalytic Dye Degradation from Textile Wastewater: A Review. *ACS Omega* **2024**, *9* (20), 21751–21767. https://doi.org/10.1021/acsomega.4c00887.
(67) BEEK, H. C. A. V.; and HEERTJES, P. M. Fading by Light of Organic Dyes on Textiles and Other Materials. *Studies in Conservation* **1966**, *11* (3), 123–132. https://doi.org/10.1179/sic.1966.016.
(68) Groeneveld, I.; Kanelli, M.; Ariese, F.; Van Bommel, M. R. Parameters That Affect the Photodegradation of Dyes and Pigments in Solution and on Substrate – An Overview. *Dyes and Pigments* **2023**, *210*, 110999. https://doi.org/10.1016/j.dyepig.2022.110999.
(69) Aslam, M.; Ismail, I. M. I.; Chandrasekaran, S.; Qari, H. A.; Hameed, A. How the Dyes Are Degraded/Mineralized in a Photocatalytic System? The Possible Role of Auxochromes. *Water Air Soil Pollut* **2015**, *226* (3), 70. https://doi.org/10.1007/s11270-015-2301-1.
(70) Delaire, J. *Photophysique et photochimie: Des fondements aux applications*; EDP sciences, 2016.
(71) Fading by Light of Organic Dyes on Textiles and Other Materials. *STUDIES IN CONSERVATION* **1966**, No. 3.

(72) Giles, C. H. The Lightfastness of Dyes : A Review. **1963**.
(73) Tan, S. N.; Yuen, M. L.; Ramli, R. A. Photocatalysis of Dyes: Operational Parameters, Mechanisms, and Degradation Pathway. *Green Analytical Chemistry* **2025**, *12*, 100230. https://doi.org/10.1016/j.greeac.2025.100230.
(74) Jiang, H.-Y.; Hu, X.-D.; Zhu, J.-J.; Wan, J.; Yao, J.-B. Studies on the Photofading of Alizarin, the Main Component of Madder. *Dyes and Pigments* **2021**, *185*, 108940. https://doi.org/10.1016/j.dyepig.2020.108940.
(75) del Hoyo-Meĺendez, J. M.; and Mecklenburg, M. F. An Investigation of the Reciprocity Principle of Light Exposures Using Microfading Spectrometry. *Spectroscopy Letters* **2011**, *44* (1), 52–62. https://doi.org/10.1080/00387010903508572.
(76) Lachman, L.; Swartz, C. J.; Urbanyi, T.; Cooper, J. Influence of Light Intensity on the Fading of Several Water-Soluble Dyes. *Journal of the American Pharmaceutical Association (Scientific ed.)* **1960**, *49* (3), 165–169. https://doi.org/10.1002/jps.3030490312.
(77) Seixas De Melo, J.; Moura, A. P.; Melo, M. J. Photophysical and Spectroscopic Studies of Indigo Derivatives in Their Keto and Leuco Forms. *J. Phys. Chem. A* **2004**, *108* (34), 6975–6981. https://doi.org/10.1021/jp049076y.
(78) Jespersen, L.; Str�mdahl, L. D.; Olsen, K.; Skibsted, L. H. Heat and Light Stability of Three Natural Blue Colorants for Use in Confectionery and Beverages. *Eur Food Res Technol* **2005**, *220* (3–4), 261–266. https://doi.org/10.1007/s00217-004-1062-7.
(79) Jørgensen, K.; Skibsted, L. H. Light Sensitivity of Cochineal. Quantum Yields for Photodegradation of Carminic Acid and Conjugate Bases in Aqueous Solution. *Food Chemistry* **1991**, *40* (1), 25–34. https://doi.org/10.1016/0308-8146(91)90016-H.
(80) Batchelor, S. N.; Carr, D.; Coleman, C. E.; Fairclough, L.; Jarvis, A. The Photofading Mechanism of Commercial Reactive Dyes on Cotton. *Dyes and Pigments* **2003**, *59* (3), 269–275. https://doi.org/10.1016/S0143-7208(03)00118-9.
(81) Weyermann, C.; Kirsch, D.; Costa-Vera, C.; Spengler, B. Photofading of Ballpoint Dyes Studied on Paper by LDI and MALDI MS. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (3), 297–306. https://doi.org/10.1016/j.jasms.2005.11.010.
(82) Forster, A. L.; Bitter, J. L.; Rosenthal, S.; Brooks, S.; Watson, S. S. Photofading in Cotton Fibers Dyed Using Red, Yellow, and Blue Direct Dyes during Examination with Microspectrophotometry (MSP). *Forensic Chem* **2017**, *5*, 72–78. https://doi.org/10.1016/j.forc.2017.06.006.
(83) Saito, M.; Minemura, C.; Nanashima, N.; Kashiwagi, M. Color Fading Behavior of Anthraquinone Dyes Due to Environmental Conditions. *Textile Research Journal* **1988**, *58* (8), 450–454. https://doi.org/10.1177/004051758805800804.
(84) Schwen, G.; Schmidt, G. Some Experiments on the Effect of Dye, Fibre, and Atmosphere on Light Fastness. *Journal of the Society of Dyers and Colourists* **1959**, *75* (2), 101–105. https://doi.org/10.1111/j.1478-4408.1959.tb02314.x.
(85) Arney, J. S.; Jacobs, A. J.; Newman, R. The Influence of Oxygen on the Fading of Organic Colorants. *Journal of the American Institute for Conservation* **1979**, *18* (2), 108–117. https://doi.org/10.2307/3179370.
(86) Karlsson, J. K. G.; Woodford, O. J.; Al-Aqar, R.; Harriman, A. Effects of Temperature and Concentration on the Rate of Photobleaching of Erythrosine in Water. *J. Phys. Chem. A* **2017**, *121* (45), 8569–8576. https://doi.org/10.1021/acs.jpca.7b06440.
(87) McLaren, K. The Importance of Temperature and Relative Humidity in Light-Fastness Testing: A Correction. *Journal of the Society of Dyers and Colourists* **1962**, *78* (1), 34–36. https://doi.org/10.1111/j.1478-4408.1962.tb02457.x.

(88) Giles, C. H.; Baxter, G.; Rahman, S. M. K. Studies of High Fastness to Light in Coloring Matters in Hydrophilic Substrates. *Textile Research Journal* **1961**, *31* (10), 831–844. https://doi.org/10.1177/004051756103101001.
(89) Manhita, A.; Ferreira, V.; Vargas, H.; Ribeiro, I.; Candeias, A.; Teixeira, D.; Ferreira, T.; Dias, C. B. Enlightening the Influence of Mordant, Dyeing Technique and Photodegradation on the Colour Hue of Textiles Dyed with Madder – A Chromatographic and Spectrometric Approach. *Microchemical Journal* **2011**, *98* (1), 82–90. https://doi.org/10.1016/j.microc.2010.12.002.
(90) Cox Crews, P. The Influence of Mordant on the Lightfastness of Yellow Natural Dyes: Journal of the American Institute for Conservation: Vol 21, No 2.
(91) Beltran, V.; Marchetti, A.; De Meyer, S.; Nuyts, G.; De Wael, K. Geranium Lake Pigments: The Role of the Synthesis on the Structure and Composition. *Dyes and Pigments* **2021**, *189*, 109260. https://doi.org/10.1016/j.dyepig.2021.109260.
(92) Sousa, M. M.; Miguel, C.; Rodrigues, I.; Parola, A. J.; Pina, F.; Seixas de Melo, J. S.; Melo, M. J. A Photochemical Study on the Blue Dye Indigo: From Solution to Ancient Andean Textiles. *Photochem Photobiol Sci* **2008**, *7* (11), 1353–1359. https://doi.org/10.1039/b809578g.
(93) Confortin, D.; Neevel, H.; Brustolon, M.; Franco, L.; Kettelarij, A. J.; Williams, R. M.; Bommel, M. R. van. Crystal Violet: Study of the Photo-Fading of an Early Synthetic Dye in Aqueous Solution and on Paper with HPLC-PDA, LC-MS and FORS. *J. Phys.: Conf. Ser.* **2010**, *231* (1), 012011. https://doi.org/10.1088/1742-6596/231/1/012011.
(94) Chavanne, C. Les couleurs de la Tapisserie de Bayeux. Thesis, Sorbonne University, Paris, 2022.
(95) Hui-Yu, J. *Studies on the photofading of alizarin, the main component of madder | Elsevier Enhanced Reader*. https://doi.org/10.1016/j.dyepig.2020.108940.
(96) Machatová, Z.; Barbieriková, Z.; Poliak, P.; Jančovičová, V.; Lukeš, V.; Brezová, V. Study of Natural Anthraquinone Colorants by EPR and UV/Vis Spectroscopy. *Dyes and Pigments* **2016**, *132*, 79–93. https://doi.org/10.1016/j.dyepig.2016.04.046.
(97) Osbaldeston, T. A. THE HERBAL OF DIOSCORIDES THE GREEK. *De Materia Medica*.
(98) *"De Materia Medica" by Dioscorides.* Library of Congress, Washington, D.C. 20540 USA. https://www.loc.gov/item/2021666851/ (accessed 2024-05-22).
(99) Parks, L. R. The Chemistry of Turkey-Red Dyeing. *J. Phys. Chem.* **1931**, *35* (2), 488–510. https://doi.org/10.1021/j150320a008.
(100) Ford, L.; L. Henderson, R.; M. Rayner, C.; S. Blackburn, R. Mild Extraction Methods Using Aqueous Glucose Solution for the Analysis of Natural Dyes in Textile Artefacts Dyed with Dyer's Madder (Rubia Tinctorum L.). *Journal of Chromatography A* **2017**, *1487*, 36–46. https://doi.org/10.1016/j.chroma.2017.01.053.
(101) Carta, L.; Biczysko, M.; Bloino, J.; Licari, D.; Barone, V. Environmental and Complexation Effects on the Structures and Spectroscopic Signatures of Organic Pigments Relevant to Cultural Heritage: The Case of Alizarin and Alizarin–Mg(Ii)/Al(Iii) Complexes. *Phys. Chem. Chem. Phys.* **2014**, *16* (7), 2897. https://doi.org/10.1039/c3cp50499a.
(102) Fantacci, S.; Amat, A.; Sgamellotti, A. *Computational Chemistry Meets Cultural Heritage: Challenges and Perspectives*. ACS Publications. https://doi.org/10.1021/ar100012b.
(103) Barone, V.; Alessandrini, S.; Biczysko, M.; Cheeseman, J. R.; Clary, D. C.; McCoy, A. B.; DiRisio, R. J.; Neese, F.; Melosso, M.; Puzzarini, C. Computational Molecular

Spectroscopy. *Nat Rev Methods Primers* **2021**, *1* (1), 38. https://doi.org/10.1038/s43586-021-00034-1.
(104) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA Quantum Chemistry Program Package. *The Journal of Chemical Physics* **2020**, *152* (22), 224108. https://doi.org/10.1063/5.0004608.
(105) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angewandte Chemie International Edition* **2009**, *48* (7), 1198–1229. https://doi.org/10.1002/anie.200802019.
(106) Laurent, A. D.; Jacquemin, D. TD-DFT Benchmarks: A Review. *International Journal of Quantum Chemistry* **2013**, *113* (17), 2019–2039. https://doi.org/10.1002/qua.24438.
(107) Jacquemin, D.; Brémond, E.; Planchat, A.; Ciofini, I.; Adamo, C. TD-DFT Vibronic Couplings in Anthraquinones: From Basis Set and Functional Benchmarks to Applications for Industrial Dyes. *J. Chem. Theory Comput.* **2011**, *7* (6), 1882–1892. https://doi.org/10.1021/ct200259k.
(108) Jacquemin, D.; Wathelet, V.; Perpète, E. A.; Adamo, C. Extensive TD-DFT Benchmark: Singlet-Excited States of Organic Molecules. *J. Chem. Theory Comput.* **2009**, *5* (9), 2420–2435. https://doi.org/10.1021/ct900298e.
(109) Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. Benchmarks of Electronically Excited States: Basis Set Effects on CASPT2 Results. *The Journal of Chemical Physics* **2010**, *133* (17), 174318. https://doi.org/10.1063/1.3499598.
(110) Jacquemin, D.; Mennucci, B.; Adamo, C. Excited-State Calculations with TD-DFT: From Benchmarks to Simulations in Complex Environments. *Phys. Chem. Chem. Phys.* **2011**, *13* (38), 16987. https://doi.org/10.1039/c1cp22144b.
(111) Goerigk, L.; Grimme, S. Assessment of TD-DFT Methods and of Various Spin Scaled CIS(D) and CC2 Versions for the Treatment of Low-Lying Valence Excitations of Large Organic Dyes. *The Journal of Chemical Physics* **2010**, *132* (18), 184103. https://doi.org/10.1063/1.3418614.
(112) Laurent, A. D.; Adamo, C.; Jacquemin, D. Dye Chemistry with Time-Dependent Density Functional Theory. *Phys. Chem. Chem. Phys.* **2014**, *16* (28), 14334–14356. https://doi.org/10.1039/C3CP55336A.
(113) Dierksen, M.; Grimme, S. The Vibronic Structure of Electronic Absorption Spectra of Large Molecules: A Time-Dependent Density Functional Study on the Influence of "Exact" Hartree—Fock Exchange. *J. Phys. Chem. A* **2004**, *108* (46), 10225–10237. https://doi.org/10.1021/jp047289h.
(114) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *The Journal of Chemical Physics* **1999**, *110* (13), 6158–6170. https://doi.org/10.1063/1.478522.
(115) Tapavicza, E.; Furche, F.; Sundholm, D. Importance of Vibronic Effects in the UV–Vis Spectrum of the 7,7,8,8-Tetracyanoquinodimethane Anion. *J. Chem. Theory Comput.* **2016**, *12* (10), 5058–5066. https://doi.org/10.1021/acs.jctc.6b00720.
(116) Tirri, B.; Turelli, M.; Boissonnat, G.; Ciofini, I.; Adamo, C. Protocols for the In-Silico Screening of the Perceived Color of Industrial Dyes: Anthraquinones and Indigos as Study Cases. *Dyes and Pigments* **2023**, *208*, 110826. https://doi.org/10.1016/j.dyepig.2022.110826.
(117) Jacquemin, D.; Brémond, E.; Ciofini, I.; Adamo, C. Impact of Vibronic Couplings on Perceived Colors: Two Anthraquinones as a Working Example. *J. Phys. Chem. Lett.* **2012**, *3* (4), 468–471. https://doi.org/10.1021/jz201552x.

(118) Brémond, E.; Bahers, T. L.; Ricci, G.; Ciofini, I.; Adamo, C. In Silico Assessment of the HPLC–UV Response Coefficients. *Computational and Theoretical Chemistry* **2014**, *1040–1041*, 1–5.
(119) Benkyi, I.; Tapavicza, E.; Fliegl, H.; Sundholm, D. Calculation of Vibrationally Resolved Absorption Spectra of Acenes and Pyrene. *Phys. Chem. Chem. Phys.* **2019**, *21* (37), 21094–21103. https://doi.org/10.1039/C9CP04178H.
(120) De Mitri, N.; Monti, S.; Prampolini, G.; Barone, V. Absorption and Emission Spectra of a Flexible Dye in Solution: A Computational Time-Dependent Approach. *J. Chem. Theory Comput.* **2013**, *9* (10), 4507–4516. https://doi.org/10.1021/ct4005799.
(121) Raucci, U.; Perrella, F.; Donati, G.; Zoppi, M.; Petrone, A.; Rega, N. Ab-Initio Molecular Dynamics and Hybrid Explicit-Implicit Solvation Model for Aqueous and Nonaqueous Solvents: GFP Chromophore in Water and Methanol Solution as Case Study. *Journal of Computational Chemistry* **2020**, *41* (26), 2228–2239. https://doi.org/10.1002/jcc.26384.
(122) Tirri, B.; Mazzone, G.; Ottochian, A.; Gomar, J.; Raucci, U.; Adamo, C.; Ciofini, I. A Combined Monte Carlo/DFT Approach to Simulate UV-Vis Spectra of Molecules and Aggregates: Merocyanine Dyes as a Case Study. *Journal of Computational Chemistry* **2021**, *42* (15), 1054–1063. https://doi.org/10.1002/jcc.26505.
(123) Haverkort, F.; Stradomska, A.; de Vries, A. H.; Knoester, J. Investigating the Structure of Aggregates of an Amphiphilic Cyanine Dye with Molecular Dynamics Simulations. *J. Phys. Chem. B* **2013**, *117* (19), 5857–5867. https://doi.org/10.1021/jp4005696.
(124) Branduardi, D.; Bussi, G.; Parrinello, M. Metadynamics with Adaptive Gaussians. *J Chem Theory Comput* **2012**, *8* (7), 2247–2254. https://doi.org/10.1021/ct3002464.
(125) Zhang, Y.; Voth, G. A. A Combined Metadynamics and Umbrella Sampling Method for the Calculation of Ion Permeation Free Energy Profiles. *J Chem Theory Comput* **2011**, *7* (7), 2277–2283. https://doi.org/10.1021/ct200100e.
(126) Le Person, A.; Cornard, J.-P.; Say-Liang-Fat, S. Studies of the Tautomeric Forms of Alizarin in the Ground State by Electronic Spectroscopy Combined with Quantum Chemical Calculations. *Chemical Physics Letters* **2011**, *517* (1–3), 41–45. https://doi.org/10.1016/j.cplett.2011.10.015.
(127) Douma, D. H.; M'Passi-Mabiala, B.; Gebauer, R. Optical Properties of an Organic Dye from Time-Dependent Density Functional Theory with Explicit Solvent: The Case of Alizarin. *The Journal of Chemical Physics* **2012**, *137* (15), 154314. https://doi.org/10.1063/1.4758877.
(128) Keeler, J.; Wothers, P. *Chemical Structure and Reactivity: An Integrated Approach*; Oxford University Press: Oxford, 2014.
(129) Olivo, H. F. A Complete Introduction to Modern NMR Spectroscopy By Roger S. Macomber. Wiley Interscience, New York, NY. 1998. Xvii + 382 Pp. 21.5 × 28 Cm. ISBN 0-471-15736-8. $49.95. *J. Med. Chem.* **1998**, *41* (19), 3758–3758. https://doi.org/10.1021/jm980304k.
(130) Ardrey, R. E. *Liquid Chromatography – Mass Spectrometry: An Introduction*, 1st ed.; Analytical Techniques in the Sciences; Wiley, 2003. https://doi.org/10.1002/0470867299.
(131) Picollo, M.; Aceto, M.; Vitorino, T. UV-Vis Spectroscopy. *Physical Sciences Reviews* **2019**, *4* (4), 20180008. https://doi.org/10.1515/psr-2018-0008.
(132) O'Connor, Z. Traditional Colour Theory: A Review. *Color Research & Application* **2021**, *46* (4), 838–847. https://doi.org/10.1002/col.22609.

(133) Newton, I. (1642-1727) A. du texte. *Opticks or, a Treatise of the reflexions, refractions, inflexions and colours of light . Also two treatises of the species and magnitude of curvilinear figures*; 1704.
(134) Johnston-Feller, R. Color Science in the Examination of Museum Objects: Nondestructive Procedures. *Color Research & Application* **2002**, *27* (6), 456–457. https://doi.org/10.1002/col.10107.
(135) Tran, T. H. L.; Berraud-Pache, R.; Jaber, M. In-Silico Color Prediction Process for Natural Dyes in Madder. *Dyes and Pigments* **2025**, *237*, 112701. https://doi.org/10.1016/j.dyepig.2025.112701.
(136) Wyszecki, G. *Color Science : Concepts and Methods, Quantitative Data and Formulae*; New York : Wiley, 1982.
(137) Johnston-Feller, R. Color Science in the Examination of Museum Objects: Nondestructive Procedures. *Color Research & Application* **2002**, *27* (6), 456–457. https://doi.org/10.1002/col.10107.
(138) Mokrzycki, W.; Tatol, M. Color Difference Delta E - A Survey. *Machine Graphics and Vision* **2011**, *20*, 383–411.
(139) Hill, B.; Roger, Th.; Vorhagen, F. W. Comparative Analysis of the Quantization of Color Spaces on the Basis of the CIELAB Color-Difference Formula. *ACM Trans. Graph.* **1997**, *16* (2), 109–154. https://doi.org/10.1145/248210.248212.
(140) Habekost, M. Which Color Differencing Equation Should Be Used? **2013**, No. 6.
(141) Sharma, G.; Wu, W.; Dalal, E. N. The CIEDE2000 Color-Difference Formula: Implementation Notes, Supplementary Test Data, and Mathematical Observations. *Color Res. Appl.* **2005**, *30* (1), 21–30. https://doi.org/10.1002/col.20070.
(142) Cardon, D.; Brémaud, I.; Quye, A.; Balfour Paul, J. Exploring Colors from the Past: In the Steps of Eighteenth-Century Dyers from France and England. *The Textile Museum Journal* **2020**, *47* (1), 9–27.
(143) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Challenges for Density Functional Theory. *Chem. Rev.* **2012**, *112* (1), 289–320. https://doi.org/10.1021/cr200107z.
(144) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications Inc.: Mineola, New York, 2012.
(145) Atkins, P. W.; Friedman, R. S.; Atkins, P. W.; Friedman, R. S. *Molecular Quantum Mechanics*, Fifth Edition, Fifth Edition.; Oxford University Press: Oxford, New York, 2010.
(146) Neese, F. Software Update: The ORCA Program System—Version 5.0. *WIREs Computational Molecular Science* **2022**, *12* (5), e1606. https://doi.org/10.1002/wcms.1606.
(147) Bursch, M.; Mewes, J.-M.; Hansen, A.; Grimme, S. Best-Practice DFT Protocols for Basic Molecular Computational Chemistry**. *Angewandte Chemie* **2022**, *134* (42), e202205735. https://doi.org/10.1002/ange.202205735.
(148) Slater, J. C. The Theory of Complex Spectra. *Phys. Rev.* **1929**, *34* (10), 1293–1322. https://doi.org/10.1103/PhysRev.34.1293.
(149) Roothaan, C. C. J. New Developments in Molecular Orbital Theory. *Rev. Mod. Phys.* **1951**, *23* (2), 69–89. https://doi.org/10.1103/RevModPhys.23.69.
(150) Handy, N. C.; Pople, J. A.; Shavitt, I. Samuel Francis Boys. *J. Phys. Chem.* **1996**, *100* (15), 6007–6016. https://doi.org/10.1021/jp963465d.
(151) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297–3305. https://doi.org/10.1039/B508541A.

(152) Jensen, F. Segmented Contracted Basis Sets Optimized for Nuclear Magnetic Shielding. *J. Chem. Theory Comput.* **2015**, *11* (1), 132–138. https://doi.org/10.1021/ct5009526.
(153) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *Journal of Computational Chemistry 32* (7).
(154) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.*
(155) Neese, F.; Olbrich, G. Efficient Use of the Resolution of the Identity Approximation in Time-Dependent Density Functional Calculations with Hybrid Density Functionals. *Chemical Physics Letters* **2002**, *362* (1), 170–178. https://doi.org/10.1016/S0009-2614(02)01053-9.
(156) Skylaris, C.-K.; Gagliardi, L.; Handy, N. C.; Ioannou, A. G.; Spencer, S.; Willetts, A. On the Resolution of Identity Coulomb Energy Approximation in Density Functional Theory. *Journal of Molecular Structure: THEOCHEM* **2000**, *501–502*, 229–239. https://doi.org/10.1016/S0166-1280(99)00434-0.
(157) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136* (3B), B864–B871. https://doi.org/10.1103/PhysRev.136.B864.
(158) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140* (4A), A1133–A1138. https://doi.org/10.1103/PhysRev.140.A1133.
(159) de Oliveira, M. T.; Alves, J. M. A.; Braga, A. A. C.; Wilson, D. J. D.; Barboza, C. A. Do Double-Hybrid Exchange–Correlation Functionals Provide Accurate Chemical Shifts? A Benchmark Assessment for Proton NMR. *J. Chem. Theory Comput.* **2021**, *17* (11), 6876–6885. https://doi.org/10.1021/acs.jctc.1c00604.
(160) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865.
(161) Perdew, J. P.; Schmidt, K. Jacob's Ladder of Density Functional Approximations for the Exchange-Correlation Energy. *AIP Conference Proceedings* **2001**, *577* (1), 1–20. https://doi.org/10.1063/1.1390175.
(162) Adamo, C.; Cossi, M.; Barone, V. An Accurate Density Functional Method for the Study of Magnetic Properties: The PBE0 Model. *Journal of Molecular Structure: THEOCHEM* **1999**, *493* (1), 145–157. https://doi.org/10.1016/S0166-1280(99)00235-3.
(163) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *The Journal of Chemical Physics* **1999**, *110* (13), 6158–6170. https://doi.org/10.1063/1.478522.
(164) Ernzerhof, M.; Scuseria, G. E. Assessment of the Perdew–Burke–Ernzerhof Exchange-Correlation Functional. *The Journal of Chemical Physics* **1999**, *110* (11), 5029–5036. https://doi.org/10.1063/1.478401.
(165) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange–Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chemical Physics Letters* **2004**, *393* (1), 51–57. https://doi.org/10.1016/j.cplett.2004.06.011.
(166) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *The Journal of Chemical Physics* **1993**, *98* (7), 5648–5652. https://doi.org/10.1063/1.464913.
(167) Parr, R. G.; Weitao, Y. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press, 1994.

(168) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys Rev B Condens Matter* **1988**, *37* (2), 785–789. https://doi.org/10.1103/physrevb.37.785.
(169) Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation. *The Journal of Chemical Physics* **2006**, *124* (3), 034108. https://doi.org/10.1063/1.2148954.
(170) Runge, E.; Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* **1984**, *52* (12), 997–1000. https://doi.org/10.1103/PhysRevLett.52.997.
(171) Tirri, B. Détermination d'un Protocole de Calcul Pour La Prédiction de Spectres UV-Vis de Molécules En Solution.
(172) de Souza, B.; Farias, G.; Neese, F.; Izsák, R. Predicting Phosphorescence Rates of Light Organic Molecules Using Time-Dependent Density Functional Theory and the Path Integral Approach to Dynamics. *J. Chem. Theory Comput.* **2019**, *15* (3), 1896–1904. https://doi.org/10.1021/acs.jctc.8b00841.
(173) Schatz, G. C. *Quantum Mechanics in Chemistry*; Mineola, N.Y. : Dover Publications, 2002.
(174) Strickler, S. J.; Berg, R. A. Relationship between Absorption Intensity and Fluorescence Lifetime of Molecules. *The Journal of Chemical Physics* **1962**, *37* (4), 814–822. https://doi.org/10.1063/1.1733166.
(175) de Souza, B.; Neese, F.; Izsák, R. On the Theoretical Prediction of Fluorescence Rates from First Principles Using the Path Integral Approach. *The Journal of Chemical Physics* **2018**, *148* (3), 034104. https://doi.org/10.1063/1.5010895.
(176) Orlandi, G.; Siebrand, W. Theory of Vibronic Intensity Borrowing. Comparison of Herzberg-Teller and Born-Oppenheimer Coupling. *The Journal of Chemical Physics* **2003**, *58* (10), 4513–4523. https://doi.org/10.1063/1.1679014.
(177) Avila Ferrer, F. J.; Santoro, F. Comparison of Vertical and Adiabatic Harmonic Approaches for the Calculation of the Vibrational Structure of Electronic Spectra. *Phys. Chem. Chem. Phys.* **2012**, *14* (39), 13549. https://doi.org/10.1039/c2cp41169e.
(178) Small, G. J. Herzberg–Teller Vibronic Coupling and the Duschinsky Effect. *The Journal of Chemical Physics* **2003**, *54* (8), 3300–3306. https://doi.org/10.1063/1.1675343.
(179) Cancès, E.; Mennucci, B.; Tomasi, J. A New Integral Equation Formalism for the Polarizable Continuum Model: Theoretical Background and Applications to Isotropic and Anisotropic Dielectrics. *J. Chem. Phys.* **1997**, *107* (8), 3032–3041. https://doi.org/10.1063/1.474659.
(180) Mennucci, B.; Cancès, E.; Tomasi, J. Evaluation of Solvent Effects in Isotropic and Anisotropic Dielectrics and in Ionic Solutions with a Unified Integral Equation Method: Theoretical Bases, Computational Implementation, and Numerical Applications. *J. Phys. Chem. B* **1997**, *101* (49), 10506–10517. https://doi.org/10.1021/jp971959k.
(181) Baroni, A.; Comite, V.; Guglielmi, V.; Casanova, M.; Redegalli, P.; Fermo, P. Identification of Natural Dyes in Historical Tapestries, a LC-MS/MS Approach. *J. Phys.: Conf. Ser.* **2022**, *2204* (1), 012065. https://doi.org/10.1088/1742-6596/2204/1/012065.
(182) Fonseca, B.; Schmidt Patterson, C.; Ganio, M.; MacLennan, D.; Trentelman, K. Seeing Red: Towards an Improved Protocol for the Identification of Madder- and Cochineal-

Based Pigments by Fiber Optics Reflectance Spectroscopy (FORS). *Herit Sci* **2019**, *7* (1), 92. https://doi.org/10.1186/s40494-019-0335-1.
(183) Brommelle, N. Material for a History of Conservation. The 1850 and 1853 Reports on the National Gallery. *Studies in Conservation* **1956**, *2* (4), 176. https://doi.org/10.2307/1504963.
(184) Pronti, L.; Mazzitelli, J.-B.; Bracciale, M. P.; Massini Rosati, L.; Vieillescazes, C.; Santarelli, M. L.; Felici, A. C. Multi-Technique Characterisation of Commercial Alizarin-Based Lakes. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2018**, *200*, 10–19. https://doi.org/10.1016/j.saa.2018.04.008.
(185) Janssens, K.; Van Der Snickt, G.; Vanmeert, F.; Legrand, S.; Nuyts, G.; Alfeld, M.; Monico, L.; Anaf, W.; De Nolf, W.; Vermeulen, M.; Verbeeck, J.; De Wael, K. Non-Invasive and Non-Destructive Examination of Artistic Pigments, Paints, and Paintings by Means of X-Ray Methods. *Top Curr Chem (Z)* **2016**, *374* (6), 81. https://doi.org/10.1007/s41061-016-0079-2.
(186) Creagh, D. C. The Characterization of Artefacts of Cultural Heritage Significance Using Physical Techniques. *Radiation Physics and Chemistry* **2005**, *74* (6), 426–442. https://doi.org/10.1016/j.radphyschem.2005.08.011.
(187) Krizsán, K.; Szókán, Gy.; Toth, Z. A.; Hollósy, F.; László, M.; Khlafulla, A. HPLC Analysis of Anthraquinone Derivatives in Madder Root (Rubia Tinctorum) and Its Cell Cultures. *Journal of Liquid Chromatography & Related Technologies* **1996**, *19* (14), 2295–2314. https://doi.org/10.1080/10826079608017158.
(188) Degano, I.; Tognotti, P.; Kunzelman, D.; Modugno, F. HPLC-DAD and HPLC-ESI-Q-ToF Characterisation of Early 20th Century Lake and Organic Pigments from Lefranc Archives. *Herit Sci* **2017**, *5* (1), 7. https://doi.org/10.1186/s40494-017-0120-y.
(189) Cosentino, A. FORS Spectral Database of Historical Pigments in Different Binders. *e-cons* **2014**, 54–65. https://doi.org/10.18236/econs2.201410.
(190) Whitney, A. V.; Casadio, F.; Van Duyne, R. P. Identification and Characterization of Artists' Red Dyes and Their Mixtures by Surface-Enhanced Raman Spectroscopy. *Appl Spectrosc* **2007**, *61* (9), 994–1000. https://doi.org/10.1366/000370207781745838.
(191) Tian, G.; Cui, R.; Liu, C.; Hu, X. Natural Dyeing with Madder: Exploring Traditional Techniques and Color Characteristics. *Fibres & Textiles in Eastern Europe* **2023**, *31* (4), 10–21. https://doi.org/10.2478/ftee-2023-0031.
(192) Justino, L. L. G.; Braz, S.; Ramos, M. L. Spectroscopic and DFT Study of Alizarin Red S Complexes of Ga(III) in Semi-Aqueous Solution. *Photochem* **2023**, *3* (1), 61–81. https://doi.org/10.3390/photochem3010005.
(193) Amat, A.; Miliani, C.; Romani, A.; Fantacci, S. DFT/TDDFT Investigation on the UV-Vis Absorption and Fluorescence Properties of Alizarin Dye. *Phys. Chem. Chem. Phys.* **2015**, *17* (9), 6374–6382. https://doi.org/10.1039/C4CP04728A.
(194) Mech, J.; Grela, M. A.; Szaciłowski, K. Ground and Excited State Properties of Alizarin and Its Isomers. *Dyes and Pigments* **2014**, *103*, 202–213. https://doi.org/10.1016/j.dyepig.2013.12.009.
(195) Cysewski, P.; Jeliński, T.; Przybyłek, M.; Shyichuk, A. Color Prediction from First Principle Quantum Chemistry Computations: A Case of Alizarin Dissolved in Methanol. *New J. Chem.* **2012**, *36* (9), 1836. https://doi.org/10.1039/c2nj40327g.
(196) Anouar, E. H.; Osman, C. P.; Weber, J.-F. F.; Ismail, N. H. UV/Visible Spectra of a Series of Natural and Synthesised Anthraquinones: Experimental and Quantum Chemical Approaches. *SpringerPlus* **2014**, *3* (1), 233. https://doi.org/10.1186/2193-1801-3-233.

(197) Berenbeim, J. A.; Boldissar, S.; Owens, S.; Haggmark, M. R.; Gate, G.; Siouri, F. M.; Cohen, T.; Rode, M. F.; Patterson, C. S.; de Vries, M. S. Excited State Intramolecular Proton Transfer in Hydroxyanthraquinones: Toward Predicting Fading of Organic Red Colorants in Art. *Sci Adv* **2019**, *5* (9), eaaw5227. https://doi.org/10.1126/sciadv.aaw5227.
(198) Tissier, R.-C.; Rigaud, B.; Thureau, P.; Jaber, M.; Ferré, N. Stressing the Differences in Alizarin and Purpurin Dyes through UV- Visible Light Absorption and 1H-NMR Spectroscopies. 10.
(199) Di Tommaso, S.; Bousquet, D.; Moulin, D.; Baltenneck, F.; Riva, P.; David, H.; Fadli, A.; Gomar, J.; Ciofini, I.; Adamo, C. Theoretical Approaches for Predicting the Color of Rigid Dyes in Solution. *J. Comput. Chem.* **2017**, *38* (13), 998–1004. https://doi.org/10.1002/jcc.24774.
(200) Tirri, B.; Turelli, M.; Boissonnat, G.; Ciofini, I.; Adamo, C. Protocols for the In-Silico Screening of the Perceived Color of Industrial Dyes: Anthraquinones and Indigos as Study Cases. *Dyes and Pigments* **2023**, *208*, 110826. https://doi.org/10.1016/j.dyepig.2022.110826.
(201) Pedone, A.; Biczysko, M.; Barone, V. Environmental Effects in Computational Spectroscopy: Accuracy and Interpretation. *ChemPhysChem* **2010**, *11* (9), 1812–1832. https://doi.org/10.1002/cphc.200900976.
(202) Miliani, C.; Romani, A.; Favaro, G. Acidichromic effects in 1,2-di- and 1,2,4-tri-hydroxyanthraquinones. A spectrophotometric and fluorimetric study. *Journal of Physical Organic Chemistry* **2000**, *13* (3), 141–150. https://doi.org/10.1002/(SICI)1099-1395(200003)13:3<141::AID-POC220>3.0.CO;2-J.
(203) Shalaby, A. A.; Mohamed, A. A. Determination of Acid Dissociation Constants of Alizarin Red S, Methyl Orange, Bromothymol Blue and Bromophenol Blue Using a Digital Camera. *RSC Adv.* **2020**, *10* (19), 11311–11316. https://doi.org/10.1039/C9RA10568A.
(204) *Software update: The ORCA program system—Version 5.0 - Neese - 2022 - WIREs Computational Molecular Science - Wiley Online Library*. https://wires.onlinelibrary.wiley.com/doi/full/10.1002/wcms.1606 (accessed 2023-08-24).
(205) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15* (3), 1652–1671. https://doi.org/10.1021/acs.jctc.8b01176.
(206) Marques, M. A. L.; Maitra, N. T.; Nogueira, F. M. S.; Gross, E. K. U.; Rubio, A. Fundamentals of Time-Dependent Density Functional Theory. In *Fundamentals of Time-Dependent Density Functional Theory*; 2012; pp 1–559.
(207) *Effect of the damping function in dispersion corrected density functional theory - Grimme - 2011 - Journal of Computational Chemistry - Wiley Online Library*. https://onlinelibrary.wiley.com/doi/10.1002/jcc.21759 (accessed 2023-12-03).
(208) Becke, A. D.; Johnson, E. R. A Density-Functional Model of the Dispersion Interaction. *J Chem Phys* **2005**, *123* (15), 154101. https://doi.org/10.1063/1.2065267.
(209) *A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu | The Journal of Chemical Physics | AIP Publishing*. https://pubs.aip.org/aip/jcp/article-

abstract/132/15/154104/926936/A-consistent-and-accurate-ab-initio?redirectedFrom=fulltext (accessed 2023-12-03).
(210) Pracht, P.; Bohle, F.; Grimme, S. Automated Exploration of the Low-Energy Chemical Space with Fast Quantum Chemical Methods. *Phys. Chem. Chem. Phys.* **2020**, *22* (14), 7169–7192. https://doi.org/10.1039/C9CP06869D.
(211) Liakos, D. G.; Guo, Y.; Neese, F. Comprehensive Benchmark Results for the Domain Based Local Pair Natural Orbital Coupled Cluster Method (DLPNO-CCSD(T)) for Closed- and Open-Shell Systems. *J. Phys. Chem. A* **2020**, *124* (1), 90–100. https://doi.org/10.1021/acs.jpca.9b05734.
(212) Garcia-Ratés, M.; Neese, F. Efficient Implementation of the Analytical Second Derivatives of Hartree–Fock and Hybrid DFT Energies within the Framework of the Conductor-like Polarizable Continuum Model. *Journal of Computational Chemistry* **2019**, *40* (20), 1816–1828. https://doi.org/10.1002/jcc.25833.
(213) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *Journal of Molecular Biology* **1976**, *103* (2), 227–249. https://doi.org/10.1016/0022-2836(76)90311-9.
(214) Groenhof, G. Introduction to QM/MM Simulations. In *Biomolecular Simulations: Methods and Protocols*; Monticelli, L., Salonen, E., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2013; pp 43–66. https://doi.org/10.1007/978-1-62703-017-5_3.
(215) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. Explicit Reversible Integrators for Extended Systems Dynamics. *Molecular Physics* **1996**, *87* (5), 1117–1157. https://doi.org/10.1080/00268979600100761.
(216) Sussmann, R.; Neuhauser, R.; Neusser, H. J. High-Resolution UV Spectroscopy of Vibronic Bands in p-Difluorobenzene and p-Difluorobenzene–Ar: The Role of Herzberg–Teller Coupling. *Can. J. Phys.* **1994**, *72* (11–12), 1179–1186. https://doi.org/10.1139/p94-151.
(217) Santoro, F.; Jacquemin, D. Going beyond the Vertical Approximation with Time-Dependent Density Functional Theory. *WIREs Computational Molecular Science* **2016**, *6* (5), 460–486. https://doi.org/10.1002/wcms.1260.
(218) *Colorimetry — Part 4: CIE 1976 L*a*b* colour space | CIE*. https://cie.co.at/publications/colorimetry-part-4-cie-1976-lab-colour-space-1 (accessed 2024-01-29).
(219) Karma, I. G. M. Determination and Measurement of Color Dissimilarity. *International Journal of Engineering and Emerging Technology* **2020**, *5*, 67. https://doi.org/10.24843/IJEET.2020.v05.i01.p13.
(220) Santra, G.; Martin, J. M. L. Pure and Hybrid SCAN, rSCAN, and r2SCAN: Which One Is Preferred in KS- and HF-DFT Calculations, and How Does D4 Dispersion Correction Affect This Ranking? *Molecules* **2022**, *27* (1), 141. https://doi.org/10.3390/molecules27010141.
(221) Preat, J.; Laurent, A. D.; Michaux, C.; Perpète, E. A.; Jacquemin, D. Impact of Tautomers on the Absorption Spectra of Neutral and Anionic Alizarin and Quinizarin Dyes. *Journal of Molecular Structure: THEOCHEM* **2009**, *901* (1–3), 24–30. https://doi.org/10.1016/j.theochem.2008.12.032.
(222) Giovannini, T.; Egidi, F.; Cappelli, C. Molecular Spectroscopy of Aqueous Solutions: A Theoretical Perspective. *Chem. Soc. Rev.* **2020**, *49* (16), 5664–5677. https://doi.org/10.1039/C9CS00464E.

(223) Ford, L.; Rayner, C. M.; Blackburn, R. S. Degradation of Lucidin: New Insights into the Fate of This Natural Pigment Present in Dyer's Madder (Rubia Tinctorum L.) during the Extraction of Textile Artefacts. *Dyes and Pigments* **2018**, *154*, 290–295. https://doi.org/10.1016/j.dyepig.2018.03.023.
(224) Singh, R.; Geetanjali; Chauhan, S. M. S. 9,10-Anthraquinones and Other Biologically Active Compounds from the Genus Rubia. *Chem Biodivers* **2004**, *1* (9), 1241–1264. https://doi.org/10.1002/cbdv.200490088.
(225) Ferreira, E. S. B.; Hulme, A. N.; McNab, H.; Quye, A. The Natural Constituents of Historical Textile Dyes. *Chem. Soc. Rev.* **2004**, *33* (6), 329. https://doi.org/10.1039/b305697j.
(226) Anouar, E. H.; Osman, C. P.; Weber, J.-F. F.; Ismail, N. H. UV/Visible Spectra of a Series of Natural and Synthesised Anthraquinones: Experimental and Quantum Chemical Approaches. *Springerplus* **2014**, *3*, 233. https://doi.org/10.1186/2193-1801-3-233.
(227) Jacquemin, D.; Wathelet, V.; Preat, J.; Perpète, E. A. Ab Initio Tools for the Accurate Prediction of the Visible Spectra of Anthraquinones. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2007**, *67* (2), 334–341. https://doi.org/10.1016/j.saa.2006.07.023.
(228) Jacquemin, D.; Assfeld, X.; Preat, J.; Perpète, E. A. Comparison of Theoretical Approaches for Predicting the UV/Vis Spectra of Anthraquinones. *Molecular Physics* **2007**, *105* (2–3), 325–331. https://doi.org/10.1080/00268970601140974.
(229) Fehér, P. P.; Madarász, Á.; Stirling, A. A Practice-Oriented Benchmark Strategy to Predict the UV-Vis Spectra of Organic Photocatalysts**. *Chemistry–Methods* **2023**, *3* (5), e202200069. https://doi.org/10.1002/cmtd.202200069.
(230) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98* (45), 11623–11627. https://doi.org/10.1021/j100096a001.
(231) Perpète, E. A.; Wathelet, V.; Preat, J.; Lambert, C.; Jacquemin, D. Toward a Theoretical Quantitative Estimation of the Λmax of Anthraquinones-Based Dyes. *J. Chem. Theory Comput.* **2006**, *2* (2), 434–440. https://doi.org/10.1021/ct050287w.
(232) Dulo, B.; Phan, K.; Githaiga, J.; Raes, K.; De Meester, S. Natural Quinone Dyes: A Review on Structure, Extraction Techniques, Analysis and Application Potential. *Waste Biomass Valor* **2021**, *12* (12), 6339–6374. https://doi.org/10.1007/s12649-021-01443-9.
(233) Ford, L.; Rayner, C. M.; Blackburn, R. S. Degradation of Lucidin: New Insights into the Fate of This Natural Pigment Present in Dyer's Madder (Rubia Tinctorum L.) during the Extraction of Textile Artefacts. *Dyes and Pigments* **2018**, *154*, 290–295. https://doi.org/10.1016/j.dyepig.2018.03.023.
(234) Tran, T. H. L.; Rigaud, B.; Jaber, M.; Berraud-Pache, R. Rapid and Complete Prediction of Alizarin in Solution by Combining Experimental Data with Computational Methods. *Dyes and Pigments* **2024**, *228*, 112242. https://doi.org/10.1016/j.dyepig.2024.112242.
(235) Takeda, T.; Kasahara, Y.; Akutagawa, T. Color-Tunable Arylaminoanthraquinone Dyes through Hydrogen-Bond-Assisted Charge Transfer Interaction. *RSC Adv.* **2021**, *11* (39), 24217–24231. https://doi.org/10.1039/D1RA03985G.
(236) *Home page - Medicinal Plant Names Services*. https://mpns.science.kew.org/ (accessed 2024-05-23).

(237) Ahn, C.; Zeng, X.; Obendorf, S. K. High-Performance Liquid Chromatography-Diode Array Detector-Mass Selective Detector Analysis of Major Natural Dyes with the Application of H_2O_2/Ultraviolet Treatment as a Way to Simulate Burial Degradation of Textiles. *Textile Research Journal* **2015**, *85* (3), 238–250. https://doi.org/10.1177/0040517514545258.
(238) *Degradation and Color Fading of Cotton Fabrics Dyed with Natural Dyes and Mordants*; Cardamone, J. M., Baker, M. T., Eds.; ACS Symposium Series; American Chemical Society: Washington, DC, 2000; Vol. 779. https://doi.org/10.1021/bk-2001-0779.
(239) Cuoco, G.; Mathe, C.; Archier, P.; Vieillescazes, C. Characterization of Madder and Garancine in Historic French Red Materials by Liquid Chromatography-Photodiode Array Detection. *Journal of Cultural Heritage* **2011**, *12* (1), 98–104. https://doi.org/10.1016/j.culher.2010.05.005.
(240) Kunzelman, D.; Picollo, M.; Bacci, M.; Bisulca, C. UV-Vis-NIR Reflectance Spectroscopy of Red Lakes in Paintings.
(241) Bouř, P. Computational Spectroscopy. Methods, Experiments and Applications. Edited by Jörg Grunenberg. *Angew Chem Int Ed* **2011**, *50* (25), 5611–5612. https://doi.org/10.1002/anie.201101367.
(242) Bestwick, J. S.; Jones, D. J.; Jones, H. E.; Kalomenopoulos, P. G.; Szabla, R.; Lawrence, A. L. Total Synthesis and Prediction of Ulodione Natural Products Guided by DFT Calculations. *Angewandte Chemie International Edition* **2022**, *61* (32), e202207004. https://doi.org/10.1002/anie.202207004.
(243) Jacquemin, D.; Wathelet, V.; Perpète, E. A.; Adamo, C. Extensive TD-DFT Benchmark: Singlet-Excited States of Organic Molecules. *J. Chem. Theory Comput.* **2009**, *5* (9), 2420–2435. https://doi.org/10.1021/ct900298e.
(244) Brémond, É. A. G.; Kieffer, J.; Adamo, C. A Reliable Method for Fitting TD-DFT Transitions to Experimental UV–Visible Spectra. *Journal of Molecular Structure: THEOCHEM* **2010**, *954* (1–3), 52–56. https://doi.org/10.1016/j.theochem.2010.04.038.
(245) Muniz-Miranda, F.; Pedone, A.; Battistelli, G.; Montalti, M.; Bloino, J.; Barone, V. Benchmarking TD-DFT against Vibrationally Resolved Absorption Spectra at Room Temperature: 7-Aminocoumarins as Test Cases. *J. Chem. Theory Comput.* **2015**, *11* (11), 5371–5384. https://doi.org/10.1021/acs.jctc.5b00750.
(246) Derksen, G. C. H.; Naayer, M.; Van Beek, T. A.; Capelle, A.; Haaksman, I. K.; Van Doren, H. A.; De Groot, Æ. Chemical and Enzymatic Hydrolysis of Anthraquinone Glycosides from Madder Roots. *Phytochemical Analysis* **2003**, *14* (3), 137–144. https://doi.org/10.1002/pca.694.
(247) Kaspar, F.; Crameri, F. Coloring Chemistry—How Mindful Color Choices Improve Chemical Communication. *Angewandte Chemie International Edition* **2022**, *61* (16), e202114910. https://doi.org/10.1002/anie.202114910.
(248) Absalan, G.; Bananejad, A.; Ghaemi, M. Removal of Alizarin Red and Purpurin from Aqueous Solutions Using Fe3O4 Magnetic Nanoparticles. *Analytical and Bioanalytical Chemistry Research* **2017**, *4* (1), 65–77. https://doi.org/10.22036/abcr.2017.41099.
(249) Alegbe, E. O.; Uthman, T. O. A Review of History, Properties, Classification, Applications and Challenges of Natural and Synthetic Dyes. *Heliyon* **2024**, *10* (13), e33646. https://doi.org/10.1016/j.heliyon.2024.e33646.
(250) *Handbook of Natural Colorants*, Repr.; Bechtold, T., Ed.; Wiley series in renewable resources; Wiley: Chichester, 2010.

(251) Shahid, M.; Shahid-ul-Islam; Mohammad, F. Recent Advancements in Natural Dye Applications: A Review. *Journal of Cleaner Production* **2013**, *53*, 310–331. https://doi.org/10.1016/j.jclepro.2013.03.031.
(252) Sealy, C. Looking over the Artist's Shoulder. *Materials Today* **2008**, *11* (11), 40–44. https://doi.org/10.1016/S1369-7021(08)70237-2.
(253) Melo, M. J.; Claro, A. Bright Light: Microspectrofluorimetry for the Characterization of Lake Pigments and Dyes in Works of Art. *Acc. Chem. Res.* **2010**, *43* (6), 857–866. https://doi.org/10.1021/ar9001894.
(254) Burnstock, A.; Hendriks, E.; Klaas Jan, van den B.; Kirby, J. Comparison of the Fading and Surface Deterioration of Red Lake Pigments in Six Paintings by Vincent van Gogh with Artificially Aged Paint Reconstructions. In *ICOM Committee for Conservation 14th Triennial Meeting The Hague 12-16 September 2005*; Preprints of the 14th Triennial Meeting of the ICOM Committee for Conservation. 1. 459-466.; ICOM COMMITTEE FOR CONSERVATION: The Hague, 2005; Vol. 1, pp 460–466.
(255) BROMMELLE, N. S. The Russell and Abney Report on the Action of Light on Water Colours. *Studies in Conservation* **1964**, *9* (4), 140–152. https://doi.org/10.1179/sic.1964.024.
(256) Grazia, C.; Clementi, C.; Miliani, C.; Romani, A. Photophysical Properties of Alizarin and Purpurin Al(III) Complexes in Solution and in Solid State. *Photochem Photobiol Sci* **2011**, *10* (7), 1249–1254. https://doi.org/10.1039/c1pp05039g.
(257) Berrie, B. H.; Strumfels, Y. Change Is Permanent: Thoughts on the Fading of Cochineal-Based Watercolor Pigments. *Herit Sci* **2017**, *5* (1), 30. https://doi.org/10.1186/s40494-017-0143-4.
(258) Tan, J. A.; Garakyaraghi, S.; Tagami, K. A.; Frano, K. A.; Crockett, H. M.; Ogata, A. F.; Patterson, J. D.; Wustholz, K. L. Contributions from Excited-State Proton and Electron Transfer to the Blinking and Photobleaching Dynamics of Alizarin and Purpurin. *J. Phys. Chem. C* **2017**, *121* (1), 97–106. https://doi.org/10.1021/acs.jpcc.6b09818.
(259) Liu, G.; Wu, T.; Zhao, J.; Hidaka, H.; Serpone, N. Photoassisted Degradation of Dye Pollutants. 8. Irreversible Degradation of Alizarin Red under Visible Light Radiation in Air-Equilibrated Aqueous TiO 2 Dispersions. *Environ. Sci. Technol.* **1999**, *33* (12), 2081–2087. https://doi.org/10.1021/es9807643.
(260) Del Hoyo-Meléndez, J. M.; Mecklenburg, M. F. An Investigation of the Reciprocity Principle of Light Exposures Using Microfading Spectrometry. *Spectroscopy Letters* **2011**, *44* (1), 52–62. https://doi.org/10.1080/00387010903508572.
(261) *ISO 105-P01:1993(en), Textiles — Tests for colour fastness — Part P01: Colour fastness to dry heat (excluding pressing)*. https://www.iso.org/obp/ui/en/#iso:std:iso:105:-P01:ed-2:v1:en (accessed 2025-07-02).
(262) Selvius DeRoo, C.; Armitage, R. A. Direct Identification of Dyes in Textiles by Direct Analysis in Real Time-Time of Flight Mass Spectrometry. *Anal. Chem.* **2011**, *83* (18), 6924–6928. https://doi.org/10.1021/ac201747s.
(263) De Santis, D.; Moresi, M. Production of Alizarin Extracts from *Rubia Tinctorum* and Assessment of Their Dyeing Properties. *Industrial Crops and Products* **2007**, *26* (2), 151–162. https://doi.org/10.1016/j.indcrop.2007.02.002.
(264) Shen, C.; Shen, Y.; Wen, Y.; Wang, H.; Liu, W. Fast and Highly Efficient Removal of Dyes under Alkaline Conditions Using Magnetic Chitosan-Fe(III) Hydrogel. *Water Research* **2011**, *45* (16), 5200–5210. https://doi.org/10.1016/j.watres.2011.07.018.
(265) Rosenblum, W. I.; El-Sabban, F. Dimethyl Sulfoxide (DMSO) and Glycerol, Hydroxyl Radical Scavengers, Impair Platelet Aggregation within and Eliminate the

Accompanying Vasodilation of, Injured Mouse Pial Arterioles. *Stroke* **1982**, *13* (1), 35–39. https://doi.org/10.1161/01.STR.13.1.35.
(266) Bansal, S.; Wang, B. A Critical Factor in Reactive Oxygen Species (ROS) Studies: The Need to Understand the Chemistry of the Solvent Used: The Case of DMSO. *Chem. Sci.* **2024**, *15* (43), 17843–17851. https://doi.org/10.1039/D4SC05038J.
(267) Thomas, J.; Townsend, J. H.; Hackney, S.; Strlič, M. A Chemiluminescence Study of Madder Lakes on Paper. *Polymer Degradation and Stability* **2010**, *95* (12), 2343–2349. https://doi.org/10.1016/j.polymdegradstab.2010.08.024.
(268) Meallet-Renault, R.; Piard, J.; Delaire, J.; Clavier, G. *Photophysique et Photochimie: Des Fondements Aux Applications*; De Boeck, 2016.
(269) Bauch, M.; Klaper, M.; Linker, T. Intermediates in the Cleavage of Endoperoxides. *Journal of Physical Organic Chemistry* **2017**, *30* (4), e3607. https://doi.org/10.1002/poc.3607.
(270) Donkers, R. L.; Workentin, M. S. Elucidation of the Electron Transfer Reduction Mechanism of Anthracene Endoperoxides. *J. Am. Chem. Soc.* **2004**, *126* (6), 1688–1698. https://doi.org/10.1021/ja035828a.
(271) Lauer, A.; Dobryakov, A. L.; Kovalenko, S. A.; Fidder, H.; Heyne, K. Dual Photochemistry of Anthracene-9,10-Endoperoxide Studied by Femtosecond Spectroscopy. *Phys. Chem. Chem. Phys.* **2011**, *13* (19), 8723. https://doi.org/10.1039/c0cp02218g.
(272) Schalk, O.; Tapavicza, E. *Photochemistry*, American Chemical Society.; iftenenv; American Chemical Society, 2021. https://doi.org/10.1021/acs.infocus.7e4009.
(273) Turro, N. J. *Modern Molecular Photochemistry*, Wiley-VCH.; ChemPhysChem; University Science Books, 1991; Vol. 12.
(274) Olivucci, M. *Computational Photochemistry*; Elsevier Science Ltd: Amsterdam ; Boston, 2005.
(275) Clennan, E. L. Aromatic Endoperoxides. *Photochemistry and Photobiology* **2023**, *99* (2), 204–220. https://doi.org/10.1111/php.13674.

ABBREVIATIONS

HAQ: Hydroxyanthraquinones

AQ: Anthraquinones

DFT: Density Functional Theory

TD: time dependent

(TD)-DFT: (time-dependent) density functional theory

HF: Hartree-Fock

HPLC: High-Performance Liquid Chromatography

MS: Mass Spectrometry

UV-VIS: UV-visible

EPR: Electron Paramagnetic Resonance

NMR: Nuclear Magnetic Resonance

FTIR: Fourier Transform Infrared Spectroscopy

XRD: X-ray Diffraction

IR: Infrared

Raman SERS: Surface-Enhanced Raman Scattering

AH: Adiabatic Hessian

VG: Vertical Gradient

FC: Franck-Condon

HT: Herzberg-Teller

DMSO: DimethylSulfoxide

DMF: Dimethylformamide

ESIPT: Excited-State Intramolecular Proton Transfer

ADC: algebraic diagrammatic construction

CC: coupled-cluster

CCSD(T): coupled-cluster theory with single and double excitations and perturbative treatment for triple excitations

COSMO: COnductor like Screening MOdel

ESR: electron spin resonance

MM: molecular mechanics

MP2: Møller–Plesset theory

PCM: polarizable continuum model

PES: potential energy surface

QM: quantum mechanics (QM′: a different QM level)

CC3: approximate coupled cluster singles, doubles, and triples model

EOM: equation of motion

FC/HT: Franck-Condon/Herzberg-Teller

GGA: generalized gradient approximation



ABSTRACT

La garance (*Rubia tinctorum* L.) est l'une des plus anciennes sources naturelles de colorants, prisée pour ses teintes rouges et utilisée depuis des millénaires dans les textiles, la peinture, et présente dans de nombreuses œuvres du patrimoine culturel. Ses racines renferment plus d'une trentaine de colorants de type anthraquinone, dont la composition varie selon l'espèce et le procédé de préparation. Toutefois, ces colorants sont sensibles à la photodégradation, un phénomène dont les mécanismes soulèvent encore plusieurs questions non résolues. Les avancées dans ce domaine sont limitées par l'instabilité de certains composés, les difficultés de synthèse, ainsi que les lacunes dans les bases de données spectroscopiques et colorimétriques. La modélisation théorique est également freinée par les limites computationnelles liées à la taille des systèmes étudiés. Pour répondre à ces enjeux, nous avons combiné des approches expérimentales et théoriques, en utilisant un modèle en phase solution, afin d'examiner les propriétés spectrales, colorimétriques et la stabilité des colorants de la garance. Un protocole de calcul basé sur la DFT/TD-DFT, intégrant les effets du solvant, du pH, des conformères et des vibrations, a été validé sur l'alizarine (l'un des principaux colorants de la garance), puis étendu à une trentaine de colorants issus de différentes espèces. Cela a permis de constituer une base de données spectro-colorimétrique, incluant des molécules inaccessibles expérimentalement.

Nous avons également modélisé les spectres UV-VIS de mélanges de colorants, permettant de reconstituer avec succès les spectres et les couleurs d'extraits naturels, ouvrant ainsi la voie à l'identification non invasive des composants dans les objets patrimoniaux. Par ailleurs, la photodégradation de l'alizarine a été étudiée dans diverses conditions (solvant, pH, lumière, oxygène), révélant un mécanisme radicalaire, avec formation probable d'endoperoxydes en milieu basique riche en oxygène. Ces travaux améliorent la compréhension de la dégradation des colorants de la garance, enrichissent les bases de données essentielles à la conservation du patrimoine, et présentent des applications potentielles dans les domaines de la cosmétique, de l'agroalimentaire, du textile, de la pharmacie, des capteurs et de la photochimie.

The madder plant (*Rubia tinctorum* L.) is one of the oldest natural sources of dye, valued for its red hues and has been used for millennia in textiles, painting, and can be found in many cultural heritage artworks. Its roots contain over thirty anthraquinone-based dyes, varying by species and preparation process. However, these dyes are prone to photodegradation, whose mechanism still raises several unsolved questions. Progress is limited by the instability of some compounds, their synthesis difficulties, and gaps in spectroscopic and colourimetric databases. Theoretical modelling is also challenged by computational limits for large systems. To address this, we combined experimental and theoretical approaches using a solution-phase model to investigate the spectral, colourimetric, and stability properties of madder dyes. A DFT/TD-DFT-based computational protocol, incorporating solvent, pH, conformers, and vibrational effects, was validated on alizarin (one of the main colourants in madder) and extended to about thirty madder dyes from various species. This enabled the creation of a spectro-colourimetric database, including experimentally inaccessible molecules. We also modelled UV-VIS spectra of dye mixtures, successfully reconstructing natural extract spectra and colour, thus supporting non-invasive identification of components in heritage objects. Additionally, we studied the photodegradation of alizarin under various conditions (solvent, pH, light, oxygen), revealing a radical-based mechanism and probable formation of endoperoxides in oxygen-rich, basic environments. This research advances the understanding of madder dye degradation, contributes to essential databases for heritage conservation, and has broader applications in cosmetics, food, textiles, pharmaceuticals, sensors, and photochemistry.