

Molecular Modeling

Stability and Optical Absorption of a Comprehensive Virtual Library of Minimal Eumelanin Oligomer Models**

Jun Wang and Lluís Blancafort*

Abstract: Eumelanin is responsible for photoprotection in living organisms. It is made of 5,6-dihydroxyindole (DHI) oligomers. However, lack of detailed structural knowledge limits understanding its function and exploiting its potential in material science. To uncover the relationship between structural stability and optical properties, we have studied a virtual library of 830 DHI dimers. We find a preference for oxidized, polycyclic structures which speaks in favor of graphite-like structures for the larger oligomers, and propose an electrocyclic formation mechanism. Besides widely considered quinone oxidation patterns, also structures with interfragment double bonds and zwitterionic resonance structures are stable. Future theoretical melanin models will have to cover this diversity, and we introduce a new representative set of 49 stable dimers. Some stable oxidized dimers have absorption energies as low as 1.3 eV. They may be present as substructures in the naturally found oligomers and contribute to the absorption spectrum of the biopolymer.

Introduction

Eumelanin is the black pigment found in skin, hair and eyes of living organisms, responsible for photoprotection against harmful light action.^[1] It is one of the three main melanin variants together with pheomelanin and neuromelanin, and it is an insoluble and heterogeneous biopolymer. Eumelanin's photoprotective function is based on a broad absorption spectrum that enables it to absorb all UV and

How to cite: *Angew. Chem. Int. Ed.* **2021**, *60*, 18800–18809
International Edition: doi.org/10.1002/anie.202106289
German Edition: doi.org/10.1002/ange.202106289

visible light of the environment.^[2] It may also have a pathogenic role as generator of reactive oxygen species^[3] and has raised interest in material science because its synthetic analogues, specially polydopamine, are versatile surface functionalization and coating materials combining attractive physical and chemical properties such as robustness, adhesivity, UV absorption, or antioxidant and free radical scavenger activity.^[4] The role of melanin in human health and the potential of this biopolymer in material science make understanding melanin's function and structure-property relationships an important goal.

Melanin is composed of oligomers made of two basic units, 5,6-dihydroxyindole (DHI) and 5,6-dihydroxyindole-2-carboxylic acid (DHICA). In the currently accepted biosynthetic path, tyrosine is transformed to DHI/DHICA in a sequence of enzymatic steps.^[5] DHI and DHICA may be oxidized further by tyrosinase,^[6] dopaquinone or O₂, and oligomerization may involve free radical coupling of *o*-semiquinones.^[7]

Melanin oligomers contain 4 to 8 DHI/DHICA units.^[8] DHI and DHICA are prone to oxidation, i.e. elimination of the pyrrolic or phenolic hydrogen atoms, and the DHI/DHICA fragments may be present in quinone or quinone methide form, where two sites are oxidized on the same fragment (see Scheme 1a). Further, assembly of two fragments with a single oxidized site results in structures with an interfragment CC double bond (Scheme 1b). The fragments

[*] Dr. J. Wang

Jiangsu Key Laboratory for Chemistry of Low-Dimensional Materials, Jiangsu Engineering Laboratory for Environment Functional Materials, School of Chemistry and Chemical Engineering, Huaiyin Normal University
No.111 West Changjiang Road, Huaian 223300, Jiangsu Province (P. R. China)

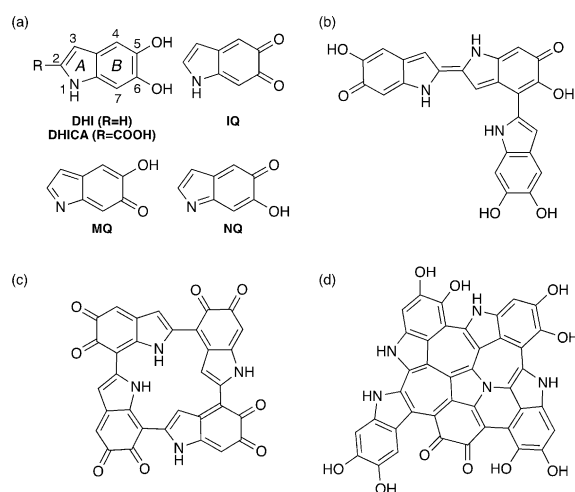
Dr. J. Wang, Dr. L. Blancafort

Institut de Química Computacional i Catàlisi and Departament de Química, Universitat de Girona, Facultat de Ciències
C/M. A. Capmany 69, 17003 Girona (Spain)
E-mail: lluis.blancafort@udg.edu

[**] A previous version of this manuscript has been deposited on a preprint server (<https://doi.org/10.26434/chemrxiv.14528664>).

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:
<https://doi.org/10.1002/anie.202106289>.

© 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.



Scheme 1. Chemical structures of eumelanin components and representative oligomers. (a) DHI/DHICA (with atom numbering and ring labeling) and oxidized indolequinone (IQ), quinone methide (MQ), and quinone imine (NQ) forms; (b) linear trimer with two singly oxidized fragments and an interfragment double bond; (c) macrocyclic tetramer; (d) polycyclic pentamer.

can assemble in linear arrangements (Scheme 1 b), but other conformations have been proposed such as 16-atom macrocycles^[9] (Scheme 1 c) or polycyclic, graphite-like structures^[10] (Scheme 1 d). According to the widely accepted hierarchical structural model,^[4a,11] the oligomers aggregate at different levels, yielding nanoparticles of >100 nm size called melanosomes. However, the structural details are still largely unknown due to the heterogeneous and granular nature of eumelanin and its synthetic analogues, which prevents crystallographic characterization. Issues like the preferred connectivity between DHI/DHICA fragments, the preferred oxidation patterns, the presence of linear and cyclic structures or the relationship between optical properties and oligomer structure are largely unresolved. Because of the difficulties in determining melanin structure with experimental techniques, computational chemistry is destined to play a key role in providing new structural insights, facilitating our understanding of melanin function and increasing our options of exploiting its potential in material science. Following this idea, we use a computational bottom-up approach to understand the principles that govern the formation of DHI oligomers and the optical properties of the basic components.

Computationally, the structural diversity of DHI oligomers has been recently studied in depth in two publications that have considered a total of more than 45 000 structures (dimers up to tetramers) composed by the reduced DHI^[12] and the oxidized indolequinone (IQ) and quinone methide (MQ) fragments.^[13] Some connectivity and oxidation patterns were excluded a priori on the basis of stability arguments, and the large number of compounds remaining even after these constraints are applied gives an idea of the scale of the problem. These studies have yielded a set of representative oligomer models and provided important insights into the relationship between structure and stability, such as the preference for planar structures, the predominance of IQ fragments in the oxidized oligomers, and the favorable formation, from the energetic point of view, of macrocyclic tetramers. However, optical properties were not considered, and a large part of chemical space remained unexplored.

Turning to the optical properties, the main interest is understanding the origin of the absorption spectrum, which has a characteristic long tail that extends up to 800 nm.^[2] Most absorption studies have been carried out on small compound sets, and the results suggest that the low-energy absorption is due to the presence of oxidized structures. Thus, a study on ten structures (monomers up to pentamers), all in reduced form, gave absorption wave lengths lower than 490 nm,^[14] whereas several studies on linear, macro- and polycyclic oligomers of up to 8 fragments with two to eight oxidized positions gave absorptions in the range 600–800 nm.^[15] Geometric disorder (i.e. aggregation) was shown to affect both position and intensity of the absorptions, and π stacking substantially red-shifted the onset of the absorption spectra.^[14] Further computational studies have considered the radiationless decay mechanisms that may be responsible for the photoprotecting function.^[16] In spite of this significant progress, a systematic study of the optical properties of a larger library of compounds is still missing. Such a study is

particularly important to gain insight into the relationship between oligomer structure, stability and absorption.

To address this issue, we have studied the stability and optical absorption of a comprehensive virtual library of 830 DHI dimers, optimizing the structures and calculating the vertical absorption with density functional theory (DFT). The oligomers present in melanin are composed of 4 to 8 DHI fragments, and the dimers will only be present in small amounts, if any. However, it is important to study the dimers because the oligomer properties will necessarily depend on those of the underlying substructures. In addition, the relatively small number of possible dimers compared to trimers or tetramers, where hundreds of thousands of structures are possible, makes it possible to consider all possible diversity. Therefore, we have considered all plausible connectivity and oxidation patterns (i.e. those with non-radical character), and we include polycyclic structures that are possible constituents of graphitic-like polycycles.^[10] We have determined the relationship between structure and absorption and have also considered the role of stability, which is crucial to understand the relevance of the different fragments for the photoprotecting function. We show that there is a thermodynamic preference for polycyclic and oxidized structures, and some significant optical absorption features of the biopolymer already appear at the level of dimers. Based on a new classification scheme that can be extended to larger oligomers we also propose a representative set of dimers that combines stability and diversity.

Results and Discussion

Library Description and Isomer Classification

Our library contains all plausible linear dimers with CC interfragment connectivity, and polycyclic dimers forming 5- or 6-membered interfragment rings containing up to two nitrogen or oxygen atoms. As plausible we consider those dimers that have a total even number of oxidized sites. An odd number of oxidation sites results in radical dimers which are probably too unstable, although radicals may become relevant for larger oligomers where the radical character can be stabilized by delocalization. For the linear dimers we consider *cis/trans* stereochemistry at the interfragment bond and include both stereoisomers in the set. The resulting set has a total of 538 linear and 292 polycyclic dimers, and the complete library including structures and optical absorption data is available as Supporting Information. From here on we refer to the polycyclic structures as cyclic dimers for brevity.

To understand the relationship between structure, stability and optical properties it is convenient to have a structural classification where similar isomers are grouped in types. For this end we introduce a classification using three criteria: oxidation state, connectivity, and resonance/bonding pattern: (1) *Oxidation state*. We consider all isomers that can be obtained by eliminating up to three pairs of phenolic or pyrrolic hydrogen atoms. The compounds obtained by eliminating 2, 4 or 6 hydrogen atoms are classified as oxidation state 1, 2 and 3 (Ox1, Ox2 and Ox3).

- (2) *Connectivity*. Linear isomers are classified according to the rings forming the interfragment bond as AA, AB or BB (see ring labeling in Scheme 1a), whereas cyclic dimers are classified according to the structure of the interfragment ring as C6, C5 (interfragment ring formed by 6 or 5 carbon atoms) or CX (interfragment ring containing one or two nitrogen or oxygen atoms).
- (3) *Resonance/bonding pattern*. Classifying the oxidation pattern with the monomer-based IQ/MQ/NQ types provides a large number of different types which is not convenient for analysis and still neglects other possible patterns. Therefore we turn to a less detailed, more granular classification based on the number of oxidized sites on each fragment and the resonance structure describing the dimer, in particular the interfragment bond. When the dimer has an even or zero number of oxidized sites on each fragment (i.e. 2 and 0 oxidized sites in Ox1 and 2 and 2 in Ox2), its resonance structure has a single bond between the two fragments. This occurs with 50% of the dimers, which are classified as *s*. When the dimer has an odd number of oxidized sites on each fragment (i.e. 1 oxidized site on each fragment in Ox1, 1 and 3 sites in Ox2 and 3 sites on each fragment in Ox3), there are two more possibilities. In 20% of the dimers there is a neutral resonance structure with a double interfragment bond, and the dimer is classified as *d*. In the remaining 30% of the cases the resonance structure has zwitterionic character (no neutral one is possible), and the dimer is classified as *z*. The rules to classify the dimers as *s*, *d* or *z* can be derived from the fragments that compose the dimer and are explained in the Supporting Information (SI), section SI1.1. As we show below, the *s/d/z* classification is useful to analyze the optical absorption.

Using these criteria, we obtain 26 different linear isomer types and 23 cyclic ones. The number of library components of each type is presented in Table 1.

Thermodynamics and Stability

The relative free energies of the dimers are obtained with DFT using the CAM-B3LYP functional and modeling bulk water solvent. The method was benchmarked against MP2. To discuss the global stability and compare structures with different stoichiometry, we use a thermodynamic Scheme based on the oxidation of the dimers with oxygen (see section SI2 for computational details). The most stable isomers of each type are presented in Schemes 2 and 3. These compounds form a representative set of the dimers that combines chemical diversity with stability. The nomenclature used to refer to individual isomers is explained in SI1.2.

To assess thermodynamic preference for linear vs. cyclic and oxidized vs. reduced dimers, the relative energy distribution of the linear and cyclic dimers is plotted in Figure 1 (panels a and b, respectively). Oxidation states are coded by color. The number of Ox1 and Ox2 isomers in the library is much larger than that of Ox0 and Ox3 ones (see Table 1), which is reflected in the size of the histogram bars. Overall, oxidation is a thermodynamically favorable process for the dimers. Thus, the bars shift progressively towards lower energies passing from fully reduced (Ox0, blue bars) to partially oxidized (Ox1 and Ox2, brown and red bars) and fully oxidized (Ox3, green bars) structures. The most stable linear and cyclic dimers are fully oxidized structures (d.lin.t.33.3.111-111 and d.cyc.34.43.3.111-111 in Schemes 2 and 3). However, the preference for higher oxidation states depends on the connectivity (see Table SI2). In all linear dimers where one of the fragments is connected at C₂, the most stable isomer has oxidation state 2 because the Ox3 dimers with that connectivity all have a *z* resonance structure (e.g. the Ox3-AA-*z* compound in Scheme 2), which is less stable than the neutral resonance structure found for Ox2 fragments. The preference for Ox2 over Ox3 in these fragments is 19–31 kcal mol⁻¹. The situation is reversed for fragments connected at C₃ (except for 2–3 connectivity), where Ox3 is preferred by 12–21 kcal mol⁻¹. For 4-4, 4-7 and 7-7 dimers there is no clear preference because the difference between the most stable Ox2 and Ox3 isomer is 1.5–7.5 kcal mol⁻¹, which probably falls within the error margin

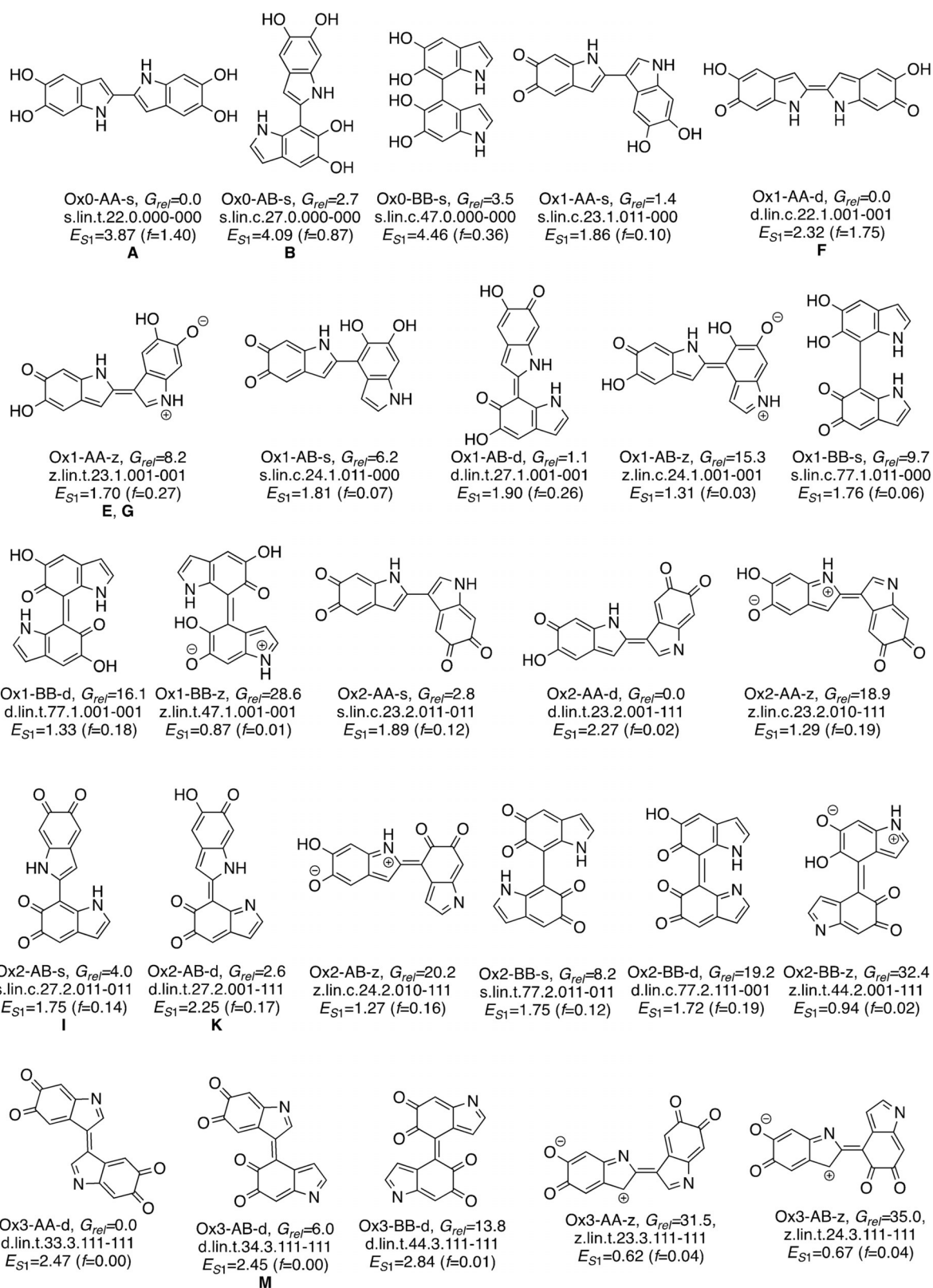
of our calculated thermodynamics. As for the cyclic dimers, the highest oxidation state, Ox3, is preferred for the C6 isomers, and Ox2 for the C5 ones. For the CX isomers, the highest possible oxidation state is Ox2 because the heteroatom that is part of the interfragment ring cannot be oxidized, and this is also the preferred state.

The plot also shows that cyclization is thermodynamically favorable. The difference between the most stable cyclic and linear dimers with the same oxidation state is 50–70 kcal mol⁻¹ (bars of the same color in Figure 1), whereas that between the most stable cyclic and

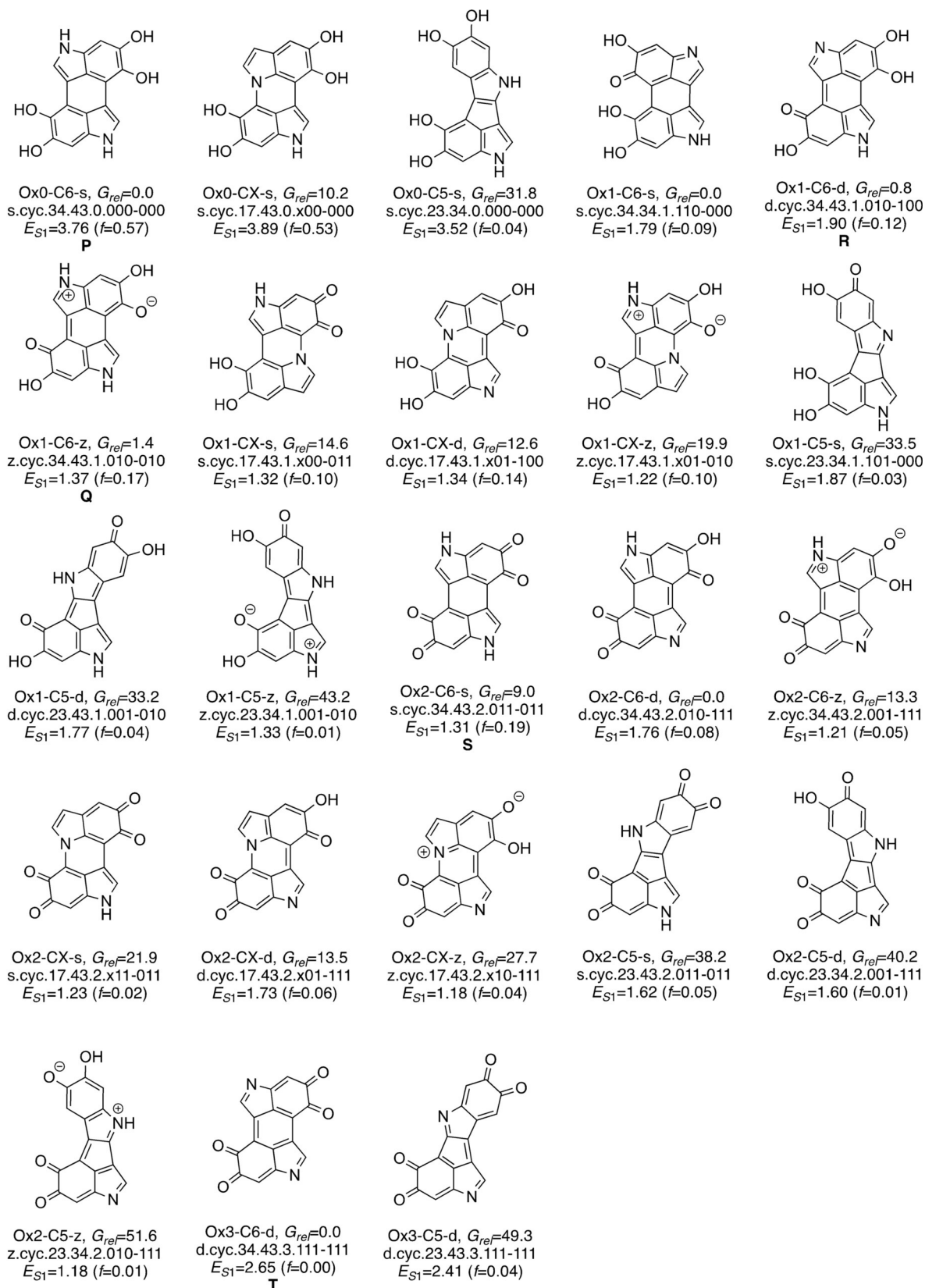
Table 1: Number of dimers of each type.

	Linear isomers									
	AAAs	AAAd	AAAz	ABs	ABd	ABz	BBs	BBd	BBz	Total
Ox0	6	–	–	8	–	–	6	–	–	20
Ox1	24	12	30	48	24	45 ^[a]	24	20	19 ^[a]	246
Ox2	42	6	18	72	20	28	42	16	8	252
Ox3	–	2	4	–	4	4	–	6	–	20
Total	72	20	52	128	48	80	72	42	30	538
	Cyclic isomers									
	C6s	C6d	C6z	C5s	C5d	C5z	CXs	CXd	CXz	Total
Ox0	2	–	–	2	–	–	14	–	–	18
Ox1	6	5	7	12	7	11	46	17	57	168
Ox2	12	4	2	18	7	5	34	9	11	102
Ox3	–	2	–	–	2	–	–	–	–	4
Total	20	11	9	32	16	16	94	26	68	292

[a] Six other possible Ox1-AB-*z* and Ox1-BB-*z* isomers were discarded because of tautomerization or wave function instability (see SI 2.1.2 and SI 2.1.3).



Scheme 2. Most stable dimers for the 26 linear types. Labelling code: isomer type and free energy in kcal mol⁻¹ (relative to the most stable linear isomer of every oxidation state); full name; S_1 vertical absorption in eV and oscillator strength in brackets. Letters **A–M** refer to the signals of Figure 5.



Scheme 3. Most stable dimers for the 23 cyclic types. Labelling code: isomer type and free energy in kcal mol^{-1} (relative to the most stable cyclic isomer of every oxidation state); full name; S_1 vertical absorption in eV and oscillator strength in brackets. Letters P–T refer to the signals of Figure 5.

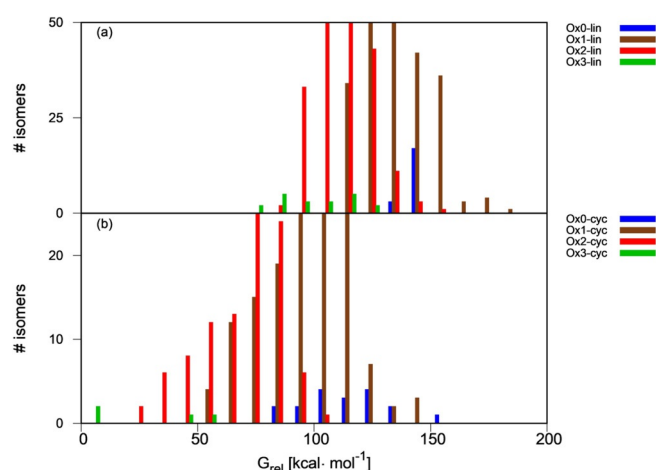
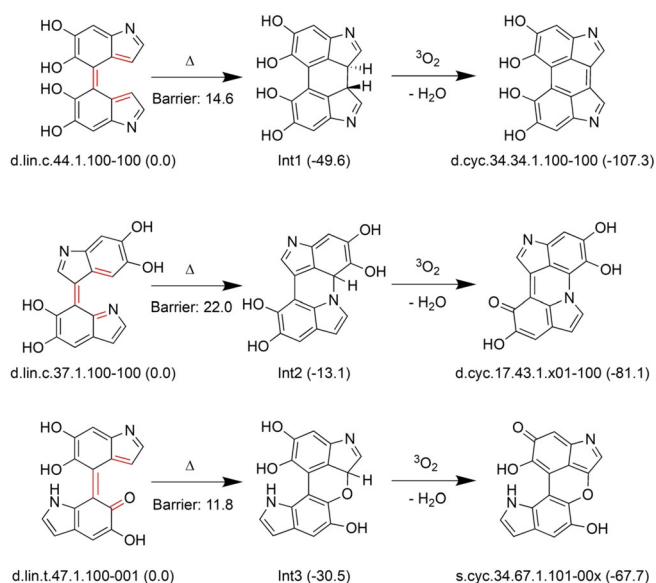


Figure 1. Free energy distribution of (a) linear and (b) cyclic dimers as a function of their oxidation state, relative to the most stable library dimer (y -axis cut at 50 and 25 isomers for clarity, see Figure S13 for full plots).

linear dimers with the same stoichiometry (e.g. most stable linear Ox1 vs. cyclic Ox0) is 20–50 kcal mol⁻¹.

As a possible route for formation of the cyclic structures we have investigated a two-step path that involves an electrocyclization, a variant of the hexatriene to cyclohexadiene rearrangement (see Scheme 4). This route is possible when the dimer has the *cis*-like triene arrangement highlighted in red. The cyclization step leads to an intermediate, followed by oxidation (H₂ elimination) to the cyclic isomers. This route was suggested by optimizations of some linear structures that lead spontaneously to the cyclized intermediates. It has been studied in detail for three representative examples leading to an all-carbon interfragment ring (Ox1-C6 type product), and two containing an N and an O atom (OX1-CX type). The electrocyclization steps have low barriers of



Scheme 4. Reaction scheme for formation of cyclic isomers via electrocyclization. Relative free energies in kcal mol⁻¹ in brackets.

12–22 kcal mol⁻¹ and are thermodynamically favorable by 13–50 kcal mol⁻¹. The hydrogen elimination step, whose thermodynamics has been estimated in the same fashion as the oxidation steps, is also highly favorable due to aromaticity gain in the final products. These examples show that an alternative route to radical reaction^[7] is possible for the formation of the cyclic structures.

To identify which are the thermodynamically most favorable types in each oxidation state, we present free energy distribution plots of the linear and cyclic isomer types in Figures 2 and 3. In the histogram bars, the isomers of each type are grouped according to their energy relative to the most stable isomer of every oxidation state. Due to structural diversity, the compounds of same oxidation state (*ie* same stoichiometry) cover a broad energy range, up to 70 kcal mol⁻¹ for Ox1 and Ox2. In Figures 2 and 3 we center on the most stable isomers and plot the most stable energy range of 0–20 kcal mol⁻¹. Histograms for the complete sets are provided in the SI (Figures S14, S15).

For the linear isomers (Figure 2), all Ox0 compounds fall into a narrow energy range of < 10 kcal mol⁻¹. Some stability trends are visible. First, for all oxidation states there is a preference for AA ring connectivity and, to a lesser extent, AB, over BB. In the three oxidation states the most stable isomer has AA connectivity, and the most stable BB structures lie 8–10 kcal mol⁻¹ higher. This is related with the coplanarity of the DHI fragments. Coplanarity is favored in the AA case because the connection involves five-membered rings and there is less steric contact between the rings than in

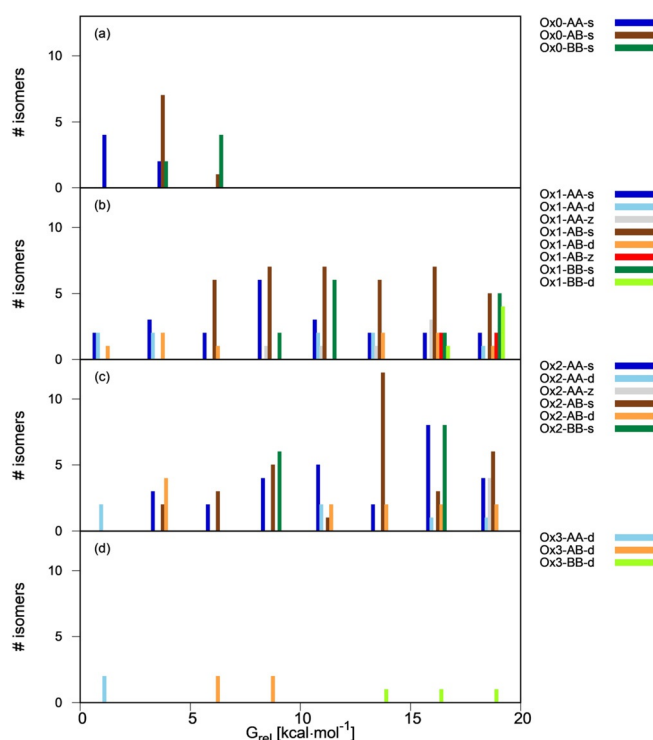


Figure 2. Free energy distribution of linear isomers as a function of their type (0–20 kcal mol⁻¹ energy range, see Figure S14 for the full range). Panels (a)–(d) display Ox0–Ox3 dimers, and energies are relative to the most stable dimer of each group.

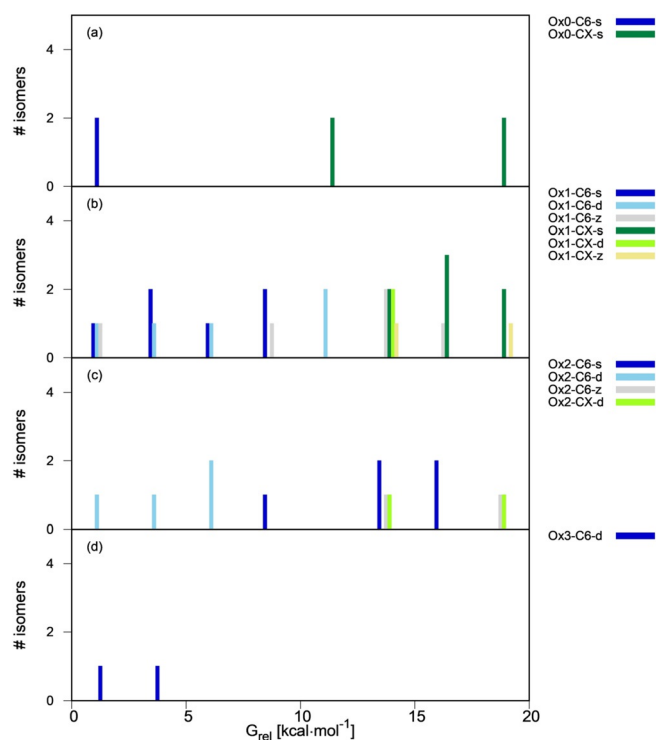


Figure 3. Free energy distribution of cyclic isomers as a function of their type (0–20 kcal mol⁻¹ energy range, see Figure S15 for the full range). Panels (a)–(d) display Ox0–Ox3 dimers, and energies are relative to the most stable dimer of each group.

the BB case, where the connection involves two six-membered rings. More coplanarity results in an energetic stabilization thanks to an increase in electronic delocalization, and therefore the AA compounds are more stable. This preference for coplanar structures was also observed in previous studies of linear oligomer libraries.^[12,13]

The preference for AA and AB connectivity is also consistent with experimental studies on dimers isolated from oxidative DHI oligomerization, where the preferred connectivity was 2-2, 2-4 and 2-7.^[17] However there are some differences. First, according to our studies 2-2 is the thermodynamically most stable connectivity for Ox0 and Ox1, but these dimers were only identified experimentally when the polymerization was carried out in the presence of transition metal ions. In addition, some of our most stable dimers have connectivity at C₃, which was not found experimentally. This suggests that other factors than the product stability, such as the thermodynamics of the radical precursors or the dimerization kinetics, are also important.

Turning to the oxidation and bonding pattern, the *d* pattern is somewhat preferred over *s* for AA and AB compounds of Ox1 and Ox2 (*s* isomers: dark blue bars in Figure 2b,c; *d* isomers: light blue and orange bars; see also Scheme 2, where the Ox1-AA-*d* dimer is more stable than the Ox1-AA-*s* one, and the same happens for Ox1-AB, Ox2-AA, Ox2-AB). The preference for an interfragment double bond may be due to a favorable electronic delocalization among the two fragments, which is favored by coplanarity in the AA and AB compounds. In contrast to this, the Ox1-BB and Ox2-BB

compounds cannot adopt coplanarity and *s* connectivity is preferred (dark vs. light green bars). Among the *s* compounds, those with IQ fragments are preferred, in line with previous results.^[13] Finally, the *z* compounds (grey and red bars in Figure 2b,c) are less stable than their *s* and *d* counterparts.

In the cyclic isomers (Figure 3), there is a clear preference for the C6 isomers (light and dark blue and grey bars) in all oxidation states, with the most stable CX and C5 isomers lying 10–15 and 30–35 kcal mol⁻¹ higher in energy, respectively (see Figure S15 for the higher-energy range). In the CX isomers, the presence of a nitrogen atom in the ring is more favorable energetically than that of an oxygen. The most stable O-containing CX isomer is about 10 kcal mol⁻¹ less stable than its N counterpart in Ox1 compounds, and 27 kcal mol⁻¹ in the Ox2 set.

Looking at the oxidation and bonding pattern, the *d* and *z* patterns with odd number of oxidized sites on every fragment are also favorable for the cyclic compounds. For the Ox1-C6 compounds, the most stable group includes one isomer of each resonance/bonding type (Figure 3b, see also Scheme 3), and the same happens with the Ox1-C5 compounds (Figure S15b). For Ox2 (Figure 3c), the most favored connectivity is *d*, and the *s* and *z* isomers lie within 15 kcal mol⁻¹. Overall, *z* resonance is comparatively more favored in the cyclic isomers, which may imply that the zwitterionic charges can be stabilized better than in the linear ones.

Optical Properties

Vertical absorption energy and oscillator strength of the dimers have been obtained with time-dependent DFT. We focus on the lowest-energy transitions and consider the S₁-S₀ states, although more states would be necessary to cover the higher-energy part of the spectrum. The relationship between structure and S₁ absorption is illustrated in Figure 4 for the linear dimers with oxidation state 0 and 1. In this plot, every horizontal line represents the absorption range of a group of compounds; the filled boxes represent the 50% of compounds in the group that absorb closest to the median

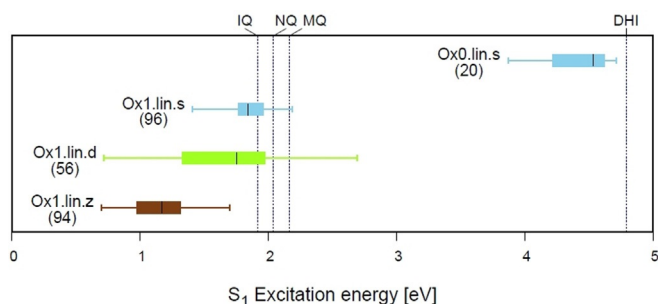


Figure 4. Distribution of S₁ excitation energy for the linear Ox0 and Ox1 dimers. The Ox1 dimers are grouped by resonance/bonding in Ox1.lin.s, Ox1.lin.d and Ox1.lin.z type. The lines between the whiskers cover the full range of absorption of each type, and the filled boxes the second and third quartile. The label includes the number of compounds from each type. The vertical dashed lines give, for reference, the absorption of the DHI monomer and the oxidized IQ, NQ and MQ monomers.

absorption (second and third quartile), and the whiskers cover the full range of absorption for that group. The median is shown as a black line, and the label identifies the group and the number of compounds in it. The vertical dashed lines give the S_1 absorptions of the DHI, IQ, MQ and NQ monomers for reference, calculated at the same level of theory. Several trends can be recognized. The fully reduced linear dimers (Ox0.lin.s, stick in the upper right part of the panel) absorb between 3.9 and 4.7 eV, which is in line with previous calculations giving 3.7–4.3 eV for several dimers.^[14] Compared to the DHI excitation of 4.8 eV, the effect of dimerization is lowering the excitation energy by up to 0.9 eV. In line with previous studies,^[15] the Ox1 dimers absorb at lower energies, between 0.7 and 2.7 eV. Compared to the oxidized monomers, dimerization can either increase or lower the excitation energy.

Turning to the relationship with bonding and resonance, if one compares the Ox1.lin.s, Ox1.lin.d and Ox1.lin.z sticks, the median absorption of the Ox1.lin.s and Ox1.lin.d compounds is similar, 1.8 eV, but the Ox1.lin.d compounds cover a broader range (0.7–2.7 eV) compared to Ox1.lin.s (1.4–2.2 eV). In turn, the Ox1.lin.z compounds absorb at lower energies, 0.7–1.7 eV. This shows that some compounds with *d* and *z* resonance/bonding absorb at lower energy than the *s* group and may be particularly important for the lower energy absorption.

Plots for the excitation energy distribution of all compound types are provided in the Supporting Information, Figure SI6a-c (S_1 - S_3 states, respectively). We summarize the main trends here. For the S_1 absorption, the trend both for linear and cyclic compounds is that the partially oxidized dimers absorb in the lower energy region, with small differences between oxidation states 1 and 2 and lower energy for the *d* and *z* types. However, complete oxidation leads to an increase in the absorption energy for the Ox3 compounds with *d* resonance/bonding character. This is probably due to the fact that these structures recover aromaticity compared to the Ox1 and Ox2 ones, which increases the HOMO–LUMO gap and the excitation energy (see compound structures in Schemes 2 and 3). As for the S_2 and S_3 states, they show similar trends to S_1 .

Turning to the absorption intensity, plots for the oscillator strength distribution among the different compound types are provided in Figures SI6d-f (S_1 - S_3). The main trends are that the more reduced Ox0 and Ox1 compounds have higher S_1 and S_2 oscillator strengths than Ox2 and Ox3, and in general S_3 absorptions have stronger oscillator strength than S_1 and S_2 . The lower oscillator strength found for the low energy states is consistent with the shape of the spectrum of the biopolymer, which shows a gradually decreasing tail towards lower energy.^[2]

To understand better how the different dimers may contribute to the absorption spectrum, it is necessary to consider their relative stability. This is illustrated in Figure 5. The eight panels correspond to the linear and cyclic isomers of Ox0–Ox3, and every signal on the plot arises from one or more compounds that absorb at the energy indicated on the *x* axis and have the relative free energy (G_{rel}) given on the *y* axis. The plot includes absorption of the three lowest states,

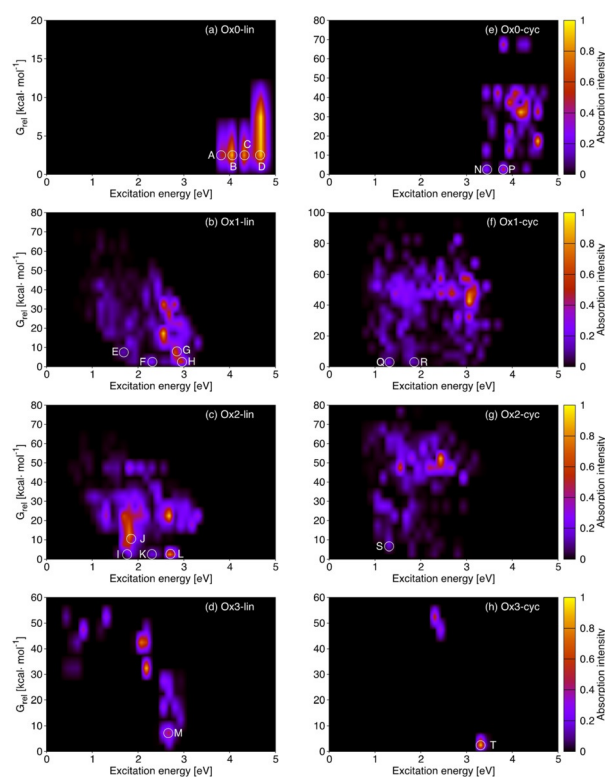


Figure 5. S_1 - S_3 absorption properties (excitation energy, *x* axis, and oscillator strength, color bars) vs. stability (energy relative to the most stable compound of every group) of (a)–(d) linear and (e)–(h) cyclic library compounds grouped by oxidation state. The letters A–T identify the most relevant features of stable compounds (low G_{rel}).

S_1 - S_3 (see SI5 for details). The absorption intensity is coded by the color bars on the right and is proportional to the calculated oscillator strength. The lower parts of the plots show how the most stable compounds of each group may contribute to the spectrum of a hypothetical compound mixture. The most relevant absorption features are highlighted with a white circle and labeled with the letters A–T.

The upper part of Figure 5 (panels a and e) shows the absorption of the reduced (Ox0) linear and cyclic dimers. The absorption starts at 3.8 and 3.4 eV, respectively, which is consistent with the distribution of Figure 4. The linear dimers show higher absorption intensity than the cyclic ones, especially for the stable compounds (compare features A–D with N and P). Some of the cyclic Ox0 dimers have more intense absorption, but their G_{rel} is high, so they will probably be less relevant. Turning to the Ox1 and Ox2 compounds, the linear dimers with low G_{rel} (bottom of panels b and c, signals E–H and I–L) absorb between 1.5 and 3 eV. The dimers absorbing at lower energies in these groups (1–1.5 eV) are relatively unstable, with $G_{rel} \geq 15$ kcal mol⁻¹. For the cyclic Ox1 and Ox2 dimers (panels f and g), the absorptions of the stable compounds (signals Q, R and S) are less intense than those found for their linear counterparts, but they appear at somewhat lower energy (1.3–1.4 eV, signals Q and S). As in the Ox0 case, the cyclic dimers with more intense absorption are relatively unstable. Finally, the fully oxidized Ox3 compounds absorb at significantly higher energy than the

Ox1 and Ox2 analogues (panels d and h, signals M and T). This may partly be due to a gain in aromaticity, as discussed above, and it is reinforced by the fact that the lowest states have n,π^* character, so that the signals correspond to higher-energy states (S_2 and S_3 for M and T, respectively). Some Ox3 compounds absorb between 1 and 2 eV, but they are unstable ($G_{\text{rel}} > 40 \text{ kcal mol}^{-1}$) and probably not relevant.

The compounds responsible for the highlighted signals in Figure 5 are identified in Table SI4. For each signal we present the most significant contributions (absorptions with higher oscillator strength), and the corresponding structures are displayed in Schemes 2, 3 and SI3. To understand the relevance of the different structural types for the absorption spectrum, we analyze the resonance/bonding pattern. We center on the Ox1 and Ox2 groups because in the Ox0 and Ox3 cases all dimers have *s* and *d* resonance/bonding type, respectively. In the Ox1-lin group, the two low-energy signals (E and F) arise from S_1 absorption of a *z* and a *d* compound, whereas the higher-energy ones (G and H) correspond to S_2 and S_3 absorptions from a larger group of compounds including *s*, *d* and *z* contributions. In contrast to this, the lowest-energy absorptions in the Ox2-lin group (I and J) arise from *s* compounds formed by two IQ fragments, whereas the higher-energy ones (K and L) come from *d* compounds (S_3 absorptions in the L case). Turning to the cyclic dimers, the Q and R signals come from a *z* compound and a group of *s* and *d* compounds, respectively, and signal S from an *s* compound formed by two IQ fragments. The orbitals with the main contribution to the absorptions corresponding to signals A–T are plotted in Figures SI7–SI10. In most cases the orbitals are delocalized between the two fragments, suggesting delocalization of the excitation. In future work it will be interesting to analyze the excitation character in depth to compare the relative importance of fragment delocalized, localized and charge transfer states.

To summarize, the absorption vs. stability plots show that some of the thermodynamically stable Ox1 and Ox2 dimers have absorptions as low as 1.3 eV, with the absorptions of the cyclic compounds being weaker and lower in energy than the linear ones. Compounds with the three possible resonance/bonding patterns, *s*, *d* and *z*, are responsible for the different signals. In particular, linear and cyclic dimers formed by two IQ fragments have an important contribution to the low energy absorptions (signals I, J, and S). We also find that some of the dimers have strong S_2 and S_3 absorptions below 3 eV. Population of these states may have implications for the excited-state dynamics and the non-radiative decay of the polymer.

Conclusion

Our work is the first comprehensive study of the stability and optical properties of a group of melanin components, the dimers, and their relationship with structure. With our classification in terms of oxidation state, connectivity and resonance/bonding, we have generated a representative set of the most stable 49 structures covering the chemical diversity of the dimers. Previous works have introduced representative

sets of dimers up to the tetramers based solely on stability,^[12,13] but these are found to be quite repetitive from the structural point of view. Our approach is complementary and compensates this by combining stability with diversity criteria.

From the structural point of view, we find a thermodynamic preference for oxidized structures which is consistent with the fact that DHI undergoes oxidative polymerization under mild conditions. Moreover cyclic structures are favored compared to linear ones. This speaks in favor of the polycyclic graphite-like structures,^[10] which have been also favored by a recent Raman spectroscopic study of dopa melanin.^[18] We have also found an energetically accessible electrocyclic path for the formation of cyclic dimers which may complement the usually assumed radical oligomerization mechanism.

Turning to optical absorption, strongly red shifted absorptions down to about 1.3 eV ($\lambda > 900 \text{ nm}$) can already be seen for some dimers that are among the most stable ones. These fragments may have a relevant contribution, as substructures of larger oligomers, to the absorption of melanin samples. Importantly, these compounds have diverse resonance/bonding pattern including interfragment double bonds and zwitterionic resonance structures. Future theoretical melanin models will have to cover this structural diversity.

As an outlook, we plan to extend our approach to the DHICA dimers. In this case the preference for AA and AB connectivity found for the DHI dimers is likely to be reversed in favor of BB (i.e. connectivity at the C_4 and C_7 positions).^[17] Since this is related with the preferred resonance/bonding character and the optical absorption, it will be interesting to see how the absorption of the DHICA dimers differs from the DHI ones, and also whether the preference for oxidized and cyclic structures is maintained. Our approach will also be extended systematically to larger oligomers and their aggregates, since an accurate modeling of the optical properties of melanin certainly needs to address this level of complexity. One question here is whether the optical properties of the linear oligomers are dominated by those of the smaller fragments, or whether extended conjugation effects are dominant. It will also be interesting to determine whether cyclic and oxidized structures are also preferred for the larger oligomers, and in the case of the cyclic compounds we will include structures with seven-membered interfragment rings in our sets that cannot be formed at the dimer level. Given the complexity of the oligomerization scheme,^[7] it will probably also be necessary to go beyond thermodynamics and include kinetic aspects in the construction of representative theoretical melanin models.

Acknowledgements

Financial support from the European Commission, project 844230 (MSCA fellowship for J. W.), and Ministerio de Ciencia, Innovación y Universidades (Spain), project PID-2019-104654GB-I00, and computational time at Red Española de Supercomputación, projects QSB-2019-3-0006, QSB-2020-1-006, QSB-2020-2-0011, and QSB-2020-3-0015, is gratefully acknowledged.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: melanin · optical properties · structure-property relationships · virtual library

-
- [1] a) M. d'Ischia, K. Wakamatsu, F. Cicoira, E. Di Mauro, J. C. Garcia-Borron, S. Commo, I. Galvan, G. Ghanem, K. Kenzo, P. Meredith, A. Pezzella, C. Santato, T. Sarna, J. D. Simon, L. Zecca, F. A. Zucca, A. Napolitano, S. Ito, *Pigment Cell Melanoma Res.* **2015**, *28*, 520–544; b) P. Meredith, T. Sarna, *Pigment Cell Res.* **2006**, *19*, 572–594; c) M. d'Ischia, A. Napolitano, V. Ball, C. T. Chen, M. J. Buehler, *Acc. Chem. Res.* **2014**, *47*, 3541–3550; d) M. d'Ischia, K. Wakamatsu, A. Napolitano, S. Briganti, J.-C. Garcia-Borron, D. Kovacs, P. Meredith, A. Pezzella, M. Picardo, T. Sarna, J. D. Simon, S. Ito, *Pigment Cell Melanoma Res.* **2013**, *26*, 616–633.
- [2] S. P. Nighswander-Rempel, J. Riesz, J. Gilmore, P. Meredith, *J. Chem. Phys.* **2005**, *123*, 194901.
- [3] D. E. Brash, L. C. P. Goncalves, E. J. H. Bechara, *Trends Mol. Med.* **2018**, *24*, 527–541.
- [4] a) M. d'Ischia, A. Napolitano, A. Pezzella, P. Meredith, T. Sarna, *Angew. Chem. Int. Ed.* **2009**, *48*, 3914–3921; *Angew. Chem.* **2009**, *121*, 3972–3979; b) E. Di Mauro, R. Xu, G. Soliveri, C. Santato, *MRS Commun.* **2017**, *7*, 141–151; c) M. d'Ischia, A. Napolitano, A. Pezzella, P. Meredith, M. Buehler, *Angew. Chem. Int. Ed.* **2020**, *59*, 11196–11205; *Angew. Chem.* **2020**, *132*, 11292–11301.
- [5] S. Ito, K. Wakamatsu, *Photochem. Photobiol.* **2008**, *84*, 582–592.
- [6] B. Roulier, B. Peres, R. Haudecoeur, *J. Med. Chem.* **2020**, *63*, 13428–13443.
- [7] A. Pezzella, O. Crescenzi, L. Panzella, A. Napolitano, E. J. Land, V. Barone, M. d'Ischia, *J. Am. Chem. Soc.* **2013**, *135*, 12142–12149.
- [8] a) H. Okuda, K. Yoshino, K. Wakamatsu, S. Ito, T. Sota, *Pigment Cell Melanoma Res.* **2014**, *27*, 664–667; b) A. Pezzella, A. Napolitano, M. d'Ischia, G. Protta, R. Seraglia, P. Traldi, *Rapid Commun. Mass Spectrom.* **1997**, *11*, 368–372.
- [9] E. Kaxiras, A. Tsolakidis, G. Zonios, S. Meng, *Phys. Rev. Lett.* **2006**, *97*, 218102.
- [10] J. Cheng, S. C. Moss, M. Eisner, *Pigment Cell Res.* **1994**, *7*, 263–273.
- [11] a) K. Y. Ju, J. Kang, J. H. Chang, J. K. Lee, *Biomacromolecules* **2016**, *17*, 2860–2872; b) A. A. R. Watt, J. P. Bothma, P. Meredith, *Soft Matter* **2009**, *5*, 3754–3760.
- [12] C. T. Chen, F. J. Martin-Martinez, G. S. Jung, M. J. Buehler, *Chem. Sci.* **2017**, *8*, 1631–1641.
- [13] C. T. Chen, M. J. Buehler, *Phys. Chem. Chem. Phys.* **2018**, *20*, 28135–28143.
- [14] D. Tuna, A. Udvarhelyi, A. L. Sobolewski, W. Domcke, T. Domratcheva, *J. Phys. Chem. B* **2016**, *120*, 3493–3502.
- [15] a) C. T. Chen, C. Chuang, J. S. Cao, V. Ball, D. Ruch, M. J. Buehler, *Nat. Commun.* **2014**, *5*, 3859; b) B. Marchetti, T. N. V. Karsili, M. N. R. Ashfold, W. Domcke, *Phys. Chem. Chem. Phys.* **2016**, *18*, 20007–20027.
- [16] a) D. Ghosh, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2021**, *11*, 1505; b) J. J. Nogueira, A. Corani, A. El Nahhas, A. Pezzella, M. d'Ischia, L. Gonzalez, V. Sundstrom, *J. Phys. Chem. Lett.* **2017**, *8*, 1004–1008.
- [17] M. D'Ischia, A. Napolitano, A. Pezzella, E. J. Land, C. A. Ramsden, P. A. Riley, *Advances in Heterocyclic Chemistry; Publisher: Elsevier Academic Press, Vol. 89* (Ed.: A. R. Katritzky), **2005**, pp. 1–63.
- [18] C. Grieco, F. R. Kohl, A. T. Hanes, B. Kohler, *Nat. Commun.* **2020**, *11*, 4569.

Manuscript received: May 10, 2021

Revised manuscript received: June 9, 2021

Accepted manuscript online: June 11, 2021

Version of record online: July 16, 2021