

Cite this: *Chem. Sci.*, 2017, 8, 1631

# Polydopamine and eumelanin molecular structures investigated with *ab initio* calculations†

Chun-Teh Chen, Francisco J. Martin-Martinez, Gang Seob Jung  
and Markus J. Buehler\*

A set of computational methods that contains a brute-force algorithmic generation of chemical isomers, molecular dynamics (MD) simulations, and density functional theory (DFT) calculations is reported and applied to investigate nearly 3000 probable molecular structures of polydopamine (PDA) and eumelanin. All probable early-polymerized 5,6-dihydroxyindole (DHI) oligomers, ranging from dimers to tetramers, have been systematically analyzed to find the most stable geometry connections as well as to propose a set of molecular models that represents the chemically diverse nature of PDA and eumelanin. Our results indicate that more planar oligomers have a tendency to be more stable. This finding is in good agreement with recent experimental observations, which suggested that PDA and eumelanin are composed of nearly planar oligomers that appear to be stacked together *via*  $\pi$ - $\pi$  interactions to form graphite-like layered aggregates. We also show that there is a group of tetramers notably more stable than the others, implying that even though there is an inherent chemical diversity in PDA and eumelanin, the molecular structures of the majority of the species are quite repetitive. Our results also suggest that larger oligomers are less likely to form. This observation is also consistent with experimental measurements, supporting the existence of small oligomers instead of large polymers as main components of PDA and eumelanin. In summary, this work brings an insight into the controversial structure of PDA and eumelanin, explaining some of the most important structural features, and providing a set of molecular models for more accurate modeling of eumelanin-like materials.

Received 20th October 2016  
Accepted 28th October 2016

DOI: 10.1039/c6sc04692d

www.rsc.org/chemicalscience

## Introduction

Polydopamine (PDA), first reported by Messersmith, Lee, and co-workers in 2007,<sup>1</sup> is a black insoluble material produced by the oxidative polymerization of dopamine under alkaline conditions. PDA shares chemical functionalities (*i.e.*, catechol and nitrogen-containing groups) with mussel adhesive proteins such as *Mytilus edulis* foot protein-5 (Mefp-5).<sup>2</sup> Mussels can attach to various types of surfaces with high bonding strength, even under harsh marine conditions. Such extraordinarily robust adhesion is due to the presence of significant amounts of catechol and amine-rich amino acids in their threads.<sup>3–5</sup> Accordingly, synthetic materials containing these functional groups are the object of growing interest, not only for adhesives but also for many other applications. PDA is a mussel-inspired material since its fundamental building blocks contain both catechol and amine groups. As with mussel threads, PDA can spontaneously deposit on the surface of almost any material regardless of its chemical nature. Therefore, PDA has drawn

extensive interest as a versatile surface functionalization and coating material for a broad range of applications.<sup>1,6,7</sup> Surface coating provides protection of the underlying material against external erosion by agents such as strong oxidants and acids. Furthermore, surface modification can change the surface properties and create additional functionalities of the underlying material. Due to the advantages mentioned above, PDA has entered the scene of materials science in recent years, with applications that go beyond surface coatings,<sup>8,9</sup> into the fields of solar energy,<sup>10</sup> water purification,<sup>11</sup> shape memory polymer,<sup>12</sup> microrobots,<sup>13</sup> biomedicine,<sup>14–16</sup> and nanotechnology.<sup>9,17</sup>

On the other hand, eumelanin, the most common type of melanin, is a natural pigment in the human skin, hair, and eyes. Eumelanin shares many physiochemical properties with PDA due to their similarities in structures and chemical functionalities. Eumelanin is an essential pigment in most organisms and has been studied for decades. X-ray diffraction studies,<sup>18,19</sup> scanning tunneling microscopy (STM) measurements,<sup>20–22</sup> atomic force microscopy (AFM) and transmission electron microscopy (TEM) images<sup>23,24</sup> suggested the existence of small nearly planar oligomers that appear to be stacked together *via*  $\pi$ - $\pi$  interactions to form graphite-like layered aggregates. However, proper explanations for the small molecular size of these oligomers and their detailed molecular

Laboratory for Atomistic and Molecular Mechanics (LAMM), Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, Massachusetts 02139, USA. E-mail: mbuehler@mit.edu

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6sc04692d

structures have not been provided yet. As a result, unlike the vast majority of natural pigments (e.g., chlorophyll and carotenoid), eumelanin cannot be described in terms of a well-defined structure, but rather as a polydisperse mixture of oligomers. The structural investigations of eumelanin are challenging due to its amorphous characteristic as well as insolubility in water and most organic solvents. Therefore, it is currently impossible to provide an accurate picture of eumelanin structure beyond a description of its fundamental building blocks, which are 5,6-dihydroxyindolequinone (DHI) and 5,6-dihydroxyindole-2-carboxylic acid (DHICA).<sup>25,26</sup> In addition, very little information is available regarding the further evolution of these building blocks to eumelanin structure, and its polymerization mechanism is also ambiguous.

The lack of precise knowledge on the molecular structures and polymerization mechanism of eumelanin does not stop the applicability of synthetic eumelanin-like materials, among which PDA is the most popular one. Dopamine, the main constituent of PDA, stands as a natural choice for synthesizing universal coating materials since it is the simplest building block containing the required catechol and amine groups. DHI, the intermediate oxidized product of dopamine, is the precursor of PDA and has been successfully synthesized and characterized in experiments. Consequently, the fundamental building blocks of PDA are DHI and its redox forms, as well as a portion of uncyclized amine-containing units. A recent study suggested that increasing the dopamine concentration while synthesizing PDA leads to higher proportions of uncyclized amine-containing units. On the other hand, lower dopamine concentrations lead to higher levels of DHI.<sup>27</sup> However, unlike eumelanin, DHICA-derived units are absent as building blocks in PDA.<sup>28</sup> As a result, even though PDA is sometimes called synthetic eumelanin in literature,<sup>29</sup> these two materials are not synonymous. In this work, to simplify the problem, we ignore the differences between PDA and eumelanin and consider them as similar materials composed of DHI and its redox forms. As with eumelanin,<sup>30</sup> despite the significant effort on structural investigations of PDA since it was first reported in 2007, the structure of PDA has yet to be determined. Recently, an experimental study<sup>29</sup> using solid state spectroscopic techniques suggested that PDA is a supramolecular aggregate of DHI monomers, which are held together by a combination of charge transfer,  $\pi$ -stacking, and hydrogen bonding interactions. Another experimental study<sup>31</sup> using high-performance liquid chromatography-mass spectrometry (HPLC-MS) reported that a physical, self-assembled trimer of (dopamine)<sub>2</sub>/DHI exists in PDA. However, it is generally accepted in this field that the polymerization of DHI generates a group of covalent oligomers including dimers, trimers, tetramers, and even larger oligomers. From the chemical analysis, it is clear that the existence of various redox forms of DHI monomers, but the way such monomers connect together to form oligomers is still unclear due to the formation of numerous intermediates during the polymerization. This variety of oligomers and redox forms that constitute PDA and eumelanin is referred to as chemical disorder model.<sup>32</sup>

The absent of proper molecular models for PDA and eumelanin implies a fundamental drawback that hampers a better understating, fast development, and property optimization of these materials. Although the science and applications of PDA-based materials have rapidly advanced in recent years, most studies in this field were based on empirical approaches instead of a solid framework of structure–function relationships due to this structural controversy. In this work, a set of computational methods, which includes a brute-force algorithmic generation of chemical isomers, molecular dynamics (MD) simulations, and density functional theory (DFT) calculations, is implemented to investigate molecular structures of PDA and eumelanin. The investigation begins with analyzing the reactivity of DHI and its related radical species that lead to the formation of PDA and eumelanin. The most reactive positions of DHI are identified, and a possible polymerization mechanism of DHI in different redox forms based on spin density calculations, is also provided. In addition, early studies showed that the size of eumelanin protomolecules is around 15–20 Å, suggesting that tetramers and pentamers formed by covalently bonded DHI monomers are the most probable molecular sizes of eumelanin.<sup>18–22</sup> A recent study based on mass spectroscopy results of DHI-melanin also suggested that the majority constituents of the material are tetramers and pentamers of DHI.<sup>32</sup> Consequently, we narrow down our investigation to DHI oligomers up to tetramers. Within the framework of this bottom-up approach and considering the location of the most reactive positions of DHI, we systematically generate and analyze all probable early-polymerized DHI oligomers, ranging from dimers to tetramers, to provide a fundamental explanation of some important structural features of PDA and eumelanin, as well as to propose a set of molecular models that represents the chemically diverse nature of these materials.

## Results and discussion

### DHI radical polymerization

A detailed investigation of the reaction mechanisms of DHI that lead to the formation of DHI oligomers, including a complete description of potential energy surfaces, activation energies, and reaction pathways, goes beyond the scope of this work. Despite such complexity, we still can analyze the reactivity of DHI to form radical species within the context of some accepted mechanisms of catechol polymerization, as well as describe the spin density and radical resonant structures that lead to the expected connections between DHI monomers. Therefore, to support the prevalence of certain atomic connections in the formation of DHI oligomers, and to provide some insight into the polymerization process that leads to the molecular structures of eumelanin-like materials, DFT calculations are performed for different redox forms of DHI. Fig. 1a summarizes a possible mechanism for catechol polymerization based on literature data.<sup>33–35</sup> In this mechanism, *o*-quinone groups react *via* aryloxy radical coupling with the unoxidized catechol to form crosslinks. This reaction scheme can be extrapolated to the case of PDA and eumelanin formation, for which the catechol functional group is thought to be the main responsible of



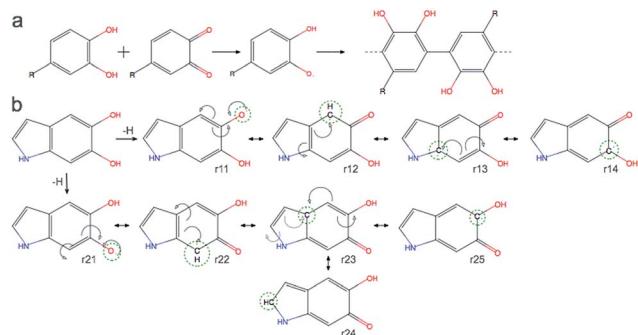


Fig. 1 (a) Simplified possible reaction mechanism for catechol polymerization, based on literature data. Possible pathway for PDA polymerization via reverse dismutation and aryloxy coupling, adapted and modified from ref. 33. (b) Two possible DHI radical species and the resonant structures resulting from the delocalization of the radical electron.

the polymerization. Fig. 1b shows two possible DHI radical species involved in these reactions, and the resonant structures resulting from the delocalization of the radical electron. The dashed circular lines in green highlight the atomic positions where the unpaired electron would delocalize in different resonant structures. Calculating the reactivity of the molecules that lead to these radicals, as well as the spin density of these radical species, is more accurate than depicting all possible resonant structures for different tautomers. Quantum chemical calculations are required to describe the electronic effects involved in the reaction. Here DFT calculations are implemented to compute the so-called Fukui function and spin density of the radical species, from monomers to tetramers. The Fukui function gives an indication of the formation of the radicals while the spin density shows the delocalization of the unpaired electrons that control the radical reaction. Fig. 2a shows the dual descriptor of the Fukui function for DHI and its redox forms. For convenience and consistency with the other species discussed, the reduced form of DHI is denoted as HQ, and the oxidized forms are denoted as quinone-imine (NQ), quinone-methide (MQ) and indolequinone (IQ). Accordingly, four different monomers are considered, namely HQ, NQ, MQ and IQ, where the colored surfaces represent those molecular regions suitable for undergoing either nucleophilic (purple) or electrophilic (green) attack. The results indicate that the reactivity of DHI is mainly concentrated in the catechol functional group, which makes the reaction scheme shown in Fig. 1a a possible reaction scheme for DHI polymerization. Despite the uncertainty in the actual mechanism of the radical generation, an eventual electrophilic attack over the oxygen positions of DHI would be responsible for the formation of the subsequent radical species in the oxygen atoms, while a nucleophilic attack over the carbon positions of the catechol ring would form the subsequent carbon radicals. We are aware that other mechanisms could also be possible, and we only emphasize the localization of the reactivity in the atomic positions consistent with the proposed models. According to this, there are four possible radical species that can be formed in the catechol functional group, coming from hydrogen abstraction over

oxygen atoms (atoms 10 and 11 in Fig. 2f) and over the two external carbon atoms (atoms 4 and 7 in Fig. 2f).

Fig. 1b shows the spin density of these four HQ radical species, where the yellow surfaces indicate the probability density for the unpaired radical electron. We name the radicals where a hydrogen atom is removed from one of the hydroxyl functional groups HQ-o10 and HQ-o11, while HQ-c4 and HQ-c7 are the species where the unpaired electron is generated in one of the carbon atoms of the catechol ring. HQ-o10 and HQ-o11 are highlighted in a red box since these are the species that would initially form, in agreement with reaction scheme shown in Fig. 1a, and thus they might be responsible for the initiation of the radical reaction. From Fig. 2b, it is clear that the atomic positions with higher spin density coincide with those suggested in Fig. 1b by studying the resonant structures. The result shows the potential of using spin density analysis for quickly exploring these radicals. Within this context, HQ-o10 and HQ-o11 would potentially attack HQ, NQ, MQ and IQ to form more radical species through hydrogen abstraction mechanisms,

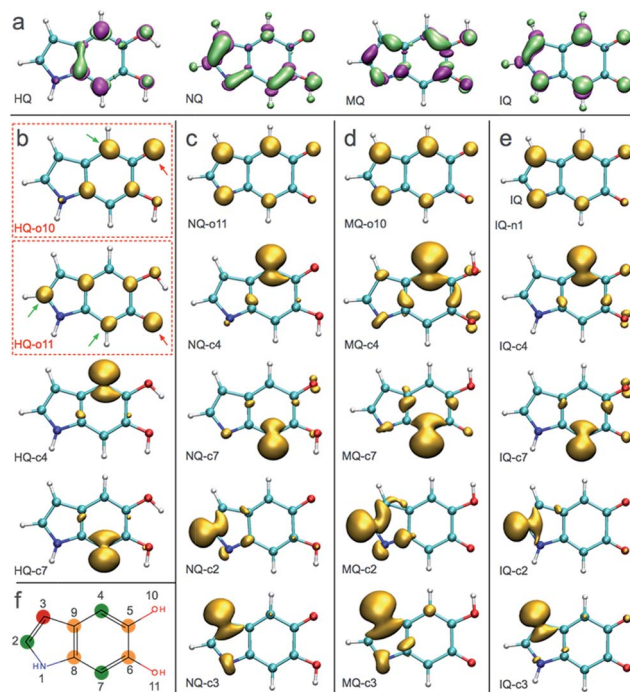


Fig. 2 Reactivity of DHI (HQ) and the spin densities of its radical species. The Fukui function and spin densities for different 2-electron oxidation products of DHI (HQ), namely HQ (non-oxidized), NQ, MQ, and IQ, and possible sites of polymerization. In the Fukui function, green volumes indicate regions that are suitable to undergo electrophilic attack and purple ones indicate regions that are suitable to undergo nucleophilic attack. In the spin density plots, gold indicates regions where the unpaired electron is delocalized. (a) Dual descriptor of the Fukui function for all 2-electron oxidation products of DHI (HQ). (b) Spin densities of the possible radical species derived from HQ. Red dashed box indicates the two main radicals generated by the accepted reaction mechanism. (c) Spin densities of the possible radical species derived from NQ. (d) Spin densities of the possible radical species derived from MQ. (e) Spin densities of the possible radical species derived from IQ. (f) Summary of possible atomic positions for polymerization.





thus generating all radicals shown in Fig. 2c–e. Those species in which the radical is generated by hydrogen abstraction in the hydroxyl group present higher delocalization of the unpaired electron, which is located not only in that oxygen atom but also in some of the carbon atoms. On the contrary, when the radical species are generated in a carbon atom of the catechol, the unpaired electron remains mainly localized in that carbon atom.

Fig. 2f summarizes the atomic positions that are most likely to polymerize based on the spin density analysis of all possible radical species. The position highlighted in red corresponds to the carbon atom in which the spin density is lower among all different structures, and thus the unpaired electron in this carbon atom is less likely to appear. Accordingly, we expect less number of crosslinking through this position. The positions highlighted in orange (5-, 6-, 8-, and 9-position) present localization of the spin density, but these positions are more difficult to access due to steric effects. On the contrary, the positions highlighted in green (2-, 4-, and 7-position) are easy to access during polymerization and also present higher spin density in most of the cases. Most importantly, in the main radical species generated from HQ, namely HQ-o10 and HQ-o11, these three positions are indeed in agreement with the most probable connections in DHI oligomers. Even though there are four carbon atoms in DHI that are suitable for polymerization, only these three are likely to polymerize in the radical reaction, as we will discuss later. Based on the polymerization sites described for the monomers, we perform a similar analysis for dimer and trimers.

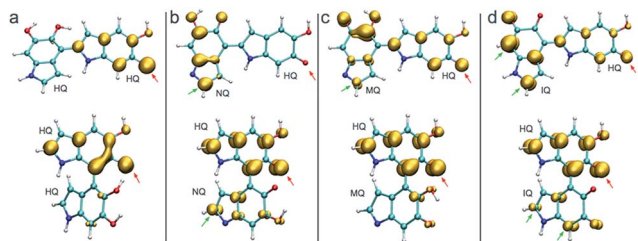
Fig. 3 shows the spin density of some HQ–HQ, HQ–NQ, HQ–MQ and HQ–IQ radical dimers that would be generated from the reaction with HQ-o10 and HQ-o11. The red arrows indicate the position where the hydrogen abstraction occurs, and the green arrows indicate the positions suitable for further functionalization based on the spin density. It is important to remark that HQ–HQ radical presents the spin density mainly localized in the monomer where the hydrogen abstraction occurs. On the contrary, HQ–NQ and HQ–IQ radical dimers present spin density localization in the atomic positions away from the initial radical, and also consistent with the selectivity for 2-, 4-, and 7-position polymerization. This is an important result for the fundamental understanding of the polymerization mechanism of DHI since it supports the need of having oxidized

species of DHI in the reaction media. To further support this result, we extend the spin density analysis to larger oligomers.

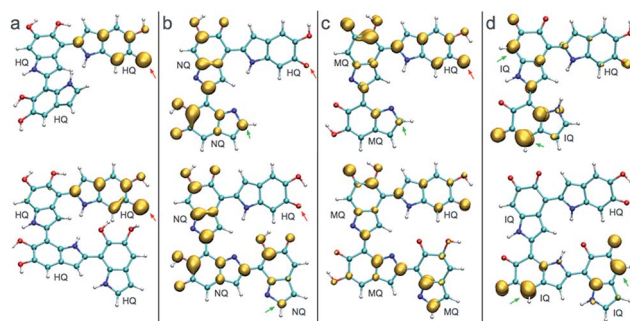
Fig. 4 shows the spin density of some radical trimers and tetramers. As with the previous figure, the red arrows indicate the positions where the hydrogen abstraction occurs, and the green arrows indicate the positions suitable for further functionalization based on the spin density. The results are consistent with the previous analysis for dimers (Fig. 3), and supporting the idea of having oxidized species of DHI in the reaction media. In fact, when the radicals include NQ, MQ, or IQ species, the unpaired electron is delocalized all over the structure, as shown in the spin density plots, and would keep polymerizing through 2-, 4-, and 7-position.

### Proposed molecular structures for dimers

Based on the results discussed in the previous section, the unpaired electron in the radical species, a key player in the polymerization mechanism of DHI, is mainly localized in 2-, 4-, and 7-position. The result implies that the covalent bonds connecting DHI monomers to generate DHI oligomers are more likely to form at these positions. As a result, there are 6 structural unique dimers (constitutional isomers), which are made through 2,2', 2,4', 2,7', 4,4', 4,7', and 7,7'-position. Nevertheless, to evaluate the stability of all possible connectivity, a total of 18 dimers are generated using a brute-force algorithmic generator (see Computational details section). Fig. S1† shows the molecular structures of these 18 dimers. Note that this group of dimers not only covers all 6 structural unique dimers but also includes 12 redundant dimers with different conformations to find the most stable conformation of each structural unique dimer. Table S1† shows the relative energies compared to the most stable dimer as well as the covalent bonding positions of these 18 dimers. The Becke three-parameter Lee–Yang–Parr (B3LYP)<sup>36,37</sup> functional is adopted for optimizing the molecular structures of these 18 dimers, together with the def2-QZVP<sup>38</sup> basis set (see Computational details section). After discarding higher energy conformations, the



**Fig. 3** Spin densities of various trimers and tetramers in different redox forms. Red arrows indicate where the hydrogen abstraction has taken place to produce the radical. Green arrows indicate high spin density regions that enable further polymerization. (a) Spin densities of two possible radicals for HQ–HQ dimer. (c) Spin densities of two possible radicals for HQ–MQ dimer. (d) Spin densities of two possible radicals for HQ–IQ dimer.



**Fig. 4** Spin densities of different trimers and tetramers at different redox forms. Red arrows indicate where the hydrogen abstraction has taken place to produce the radical. (a) Spin densities of possible radicals in a HQ polymerization. (b) Spin densities of possible radicals in a HQ–NQ polymerization. (c) Spin densities of possible radicals in a HQ–MQ polymerization. (d) Spin densities of possible radicals for HQ–IQ polymerization.



most stable dimer is made through 2,2'-position, followed by 2,4'-position (+1.82 kcal mol<sup>-1</sup>), 4,4'-position (+2.19 kcal mol<sup>-1</sup>), 4,7'-position (+2.31 kcal mol<sup>-1</sup>), 2,7'-position (+2.82 kcal mol<sup>-1</sup>), and 7,7'-position (+3.20 kcal mol<sup>-1</sup>).

Fig. 5 shows the molecular structures of the 6 structural unique dimers in the order of their relative energies compared to the most stable one. The dimers made through 2,2'-position (Fig. 5a), 2,4'-position (Fig. 5b), and 2,7'-position (Fig. 5e), have been identified in experiments by oxidation of DHI.<sup>39,40</sup> In addition, the dimer made through 4,7'-position (Fig. 5d) has also been verified in experiments.<sup>41</sup> Although the dimers made through 4,4'-position (Fig. 5c) and 7,7'-position (Fig. 5f) have not been isolated in experiments, these two covalent bonding formations have been identified by oxidation of DHICA.<sup>42</sup> Due to the structural similarity between DHI and DHICA, we expect that these two covalent bonding formations can also form in oxidation of DHI. As a result, the dimers shown in Fig. 5 are all likely to exist in PDA and eumelanin. The activation energies for DHI monomers to form various DHI oligomers are hard to calculate without knowing the comprehensive polymerization mechanism. However, the equal chemical functionalities and similar structural conformations between the different species involved in the polymerization suggest comparable energy barriers be expected. Furthermore, it has been experimentally proven that the polymerization of DHI occurs. Hence, it is reasonable to assume that the formation of oligomers is kinetically possible in most of the cases. Consequently, we assume that the energy of an oligomer determines its prevalence with respect to others. These assumptions define the theoretical framework in this work for evaluating different oligomers that contribute to the overall chemical structure of PDA and eumelanin.

### Proposed molecular structures for trimers

Following the same strategy applied for dimers, a total of 216 trimers are generated using the brute-force algorithmic generator. Fig. S2† shows the molecular structures of these 216 trimers. This group of trimers covers all 27 structurally unique trimers and also includes 189 redundant trimers with different

conformations to find the most stable conformation of each structural unique trimer. Note that the B3LYP functional with the def2-QZVP basis set adopted for optimizing the molecular structures of the dimers is too expensive for the trimers. Consequently, the Becke–Lee–Yang–Parr (BLYP)<sup>43,44</sup> functional is adopted for optimizing the molecular structures of the trimers, together with the def2-SVP<sup>38</sup> basis set (see Computational details section). After discarding those structures with higher energy conformations, the top 12 stable trimers are identified, which are made through 2,4'- & 2,2'-position, 7,4'- & 2,2'-position (+0.78 kcal mol<sup>-1</sup>), 2,2'- & 4,4'-position (+0.98 kcal mol<sup>-1</sup>), 2,2'- & 7,4'-position (+1.08 kcal mol<sup>-1</sup>), 2,2'- & 7,2'-position (+1.43 kcal mol<sup>-1</sup>), 7,7'- & 2,2'-position (+2.31 kcal mol<sup>-1</sup>), 2,4'- & 7,4'-position (+2.41 kcal mol<sup>-1</sup>), 4,2'- & 4,4'-position (+2.51 kcal mol<sup>-1</sup>), 4,2'- & 7,4'-position (+2.64 kcal mol<sup>-1</sup>), 7,2'- & 4,2'-position (+2.65 kcal mol<sup>-1</sup>), 2,4'- & 7,2'-position (+2.71 kcal mol<sup>-1</sup>), and 4,2'- & 4,7'-position (+2.73 kcal mol<sup>-1</sup>). Fig. 6 shows the molecular structures of these top 12 stable trimers in the order of their relative energies. Note that the DFT calculations for the trimers are implemented using the small def2-SVP basic set instead of the large def2-QZVP basic set due to computational limitations. The basic set error in the calculations could be larger than the energy differences between the trimers (see Computational details section), and thus the rank in Fig. 6 might not be exactly correct. However, the group of trimers shown in Fig. 6 should be at least more stable than the other trimers. Consequently, instead of providing the exact rank for the trimers, here we can only propose a group of

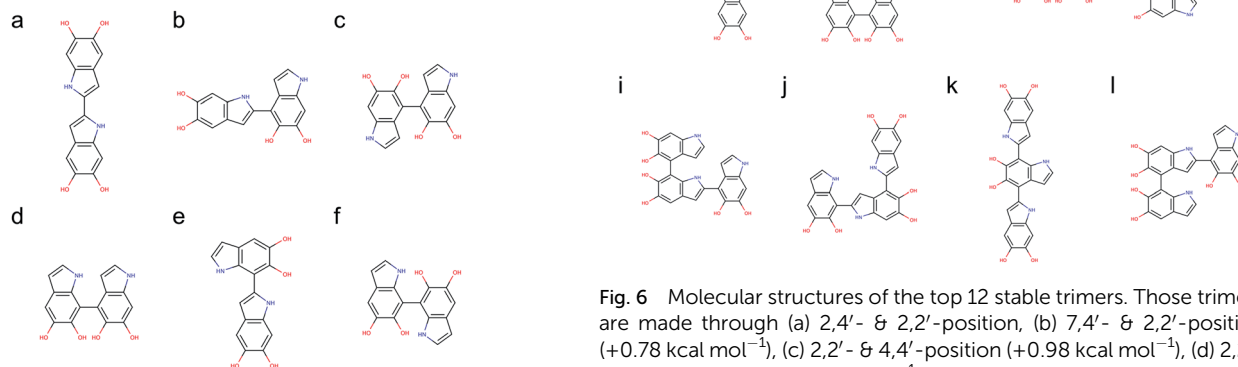


Fig. 5 Molecular structures of 6 structural unique dimers. The most stable dimer is made through (a) 2,2'-position, followed by (b) 2,4'-position (+1.82 kcal mol<sup>-1</sup>), (c) 4,4'-position (+2.19 kcal mol<sup>-1</sup>), (d) 4,7'-position (+2.31 kcal mol<sup>-1</sup>), (e) 2,7'-position (+2.82 kcal mol<sup>-1</sup>), and (f) 7,7'-position (+3.20 kcal mol<sup>-1</sup>).

Fig. 6 Molecular structures of the top 12 stable trimers. Those trimers are made through (a) 2,4'- & 2,2'-position (+0.78 kcal mol<sup>-1</sup>), (b) 7,4'- & 2,2'-position (+0.98 kcal mol<sup>-1</sup>), (c) 2,2'- & 4,4'-position (+1.08 kcal mol<sup>-1</sup>), (d) 2,2'- & 7,4'-position (+1.43 kcal mol<sup>-1</sup>), (e) 2,2'- & 7,2'-position (+2.31 kcal mol<sup>-1</sup>), (f) 7,7'- & 2,2'-position (+2.41 kcal mol<sup>-1</sup>), (g) 2,4'- & 7,4'-position (+2.51 kcal mol<sup>-1</sup>), (h) 4,2'- & 4,4'-position (+2.64 kcal mol<sup>-1</sup>), (i) 4,2'- & 7,4'-position (+2.65 kcal mol<sup>-1</sup>), (j) 7,2'- & 4,2'-position (+2.71 kcal mol<sup>-1</sup>), (k) 2,4'- & 7,2'-position (+2.73 kcal mol<sup>-1</sup>), and (l) 4,2'- & 4,7'-position (+2.73 kcal mol<sup>-1</sup>).



trimers that is more energetic favorable. In fact, the trimer made through 2,4'- & 2,7'-position (Fig. 6j) has been identified by oxidation of DHI.<sup>39,40</sup>

### Proposed molecular structures for tetramers

The number of DHI oligomers that need to be investigated increases exponentially as the size increases. A total of 2592 tetramers are generated using the brute-force algorithmic generator. This group of tetramers covers all 162 structural unique tetramers as well as 2430 redundant tetramers that account for different conformations to find the most stable conformation of each structural unique tetramer. As with the calculations for the trimers, the BLYP<sup>43,44</sup> functional is adopted for optimizing the molecular structures of the tetramers, together with the def2-SVP<sup>38</sup> basis (see Computational details section). After removing the results of higher energy conformations, the systematic search for the most stable molecular structures of the tetramers is completed. Following the same strategy, the top 16 stable tetramers are identified, which are made through 4,2'- & 7,4'- & 2,2'-position, 2,2'- & 4,4'- & 2,2'-position (+0.59 kcal mol<sup>-1</sup>), 2,2'- & 4,2'- & 4,4'-position (+1.03 kcal mol<sup>-1</sup>), 2,2'- & 4,7'- & 2,2'-position (+1.08 kcal mol<sup>-1</sup>), 2,2'- & 7,2'- & 4,4'-position (+1.18 kcal mol<sup>-1</sup>), 2,2'- & 4,2'- & 7,4'-position (+1.20 kcal mol<sup>-1</sup>), 2,4'- & 2,4'- & 2,2'-position (+1.33 kcal mol<sup>-1</sup>), 2,4'- & 2,2'- & 4,2'-position (+1.50 kcal mol<sup>-1</sup>), 2,2'- & 7,7'- & 2,4'-position (+1.56 kcal mol<sup>-1</sup>), 2,2'- & 7,4'- & 2,7'-position (+1.91 kcal mol<sup>-1</sup>), 2,2'- & 4,2'- & 4,7'-position (+2.03 kcal mol<sup>-1</sup>), 7,2'- & 7,4'- & 2,2'-position (+2.06 kcal mol<sup>-1</sup>), 2,2'- & 4,2'- & 7,7'-position (+2.11 kcal mol<sup>-1</sup>), 2,2'- & 4,7'- & 4,2'-position (+2.22 kcal mol<sup>-1</sup>), 2,4'- & 2,7'- & 2,2'-position (+2.24 kcal mol<sup>-1</sup>), and 2,4'- & 2,2'- & 4,7'-position (+2.33 kcal mol<sup>-1</sup>). Fig. 7 shows the molecular structures of these top 16 stable tetramers in the order of their relative energies. As discussed before, the rank shown in Fig. 7 might not be exactly correct due to the accuracy in the calculations. However, the group of tetramers shown in Fig. 7 should be at least more energetic favorable than the other tetramers.

We are aware of the evidence of porphyrin-like tetramers that has been found in natural eumelanin pigment as well as in PDA by using electrochemical fingerprinting.<sup>45</sup> These tetramers are made through 2,7'- & 2,7'- & 2,7'-position, creating an interior ring where all the nitrogen atoms are in the center, in an arrangement similar to porphyrin. Such arrangement was first proposed by Kaxiras<sup>25,26</sup> and have been adopted to study the structural, mechanical, and optical properties of PDA and eumelanin in our previous computational work.<sup>46,47</sup> However, this particular kind of molecular structures, in which an additional covalent bond is formed between the first and fourth monomers, is not considered in this work. The reason is that two hydrogen atoms need to be removed when an additional covalent bond. Therefore, the number of atoms in porphyrin-like tetramers is different from that in the tetramers considered in this work. Consequently, the energy comparison between tetramers with different numbers of atoms will be problematic. Despite PDA and eumelanin might contain porphyrin-like tetramers in certain concentrations, it is impractical to assume

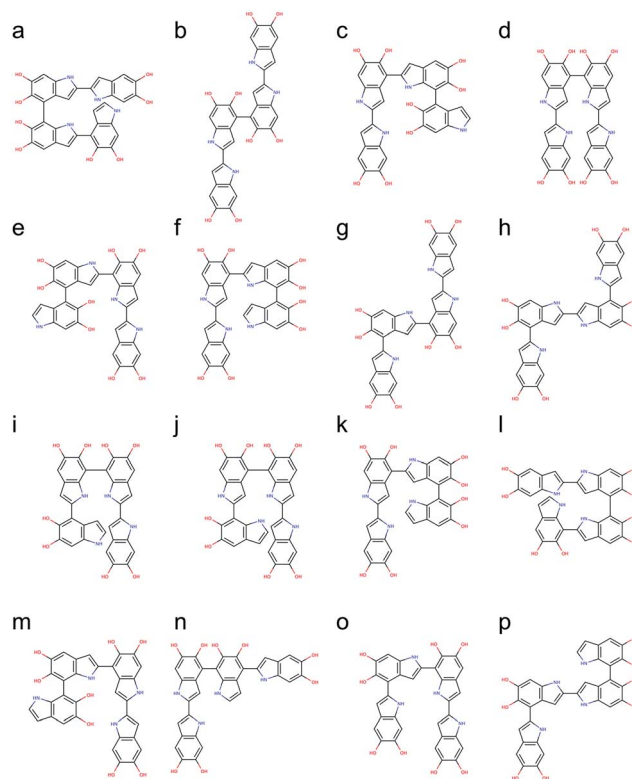


Fig. 7 Molecular structures of the top 16 stable tetramers. Those tetramers are made through (a) 4,2'- & 7,4'- & 2,2'-position, (b) 2,2'- & 4,4'- & 2,2'-position (+0.59 kcal mol<sup>-1</sup>), (c) 2,2'- & 4,2'- & 4,4'-position (+1.03 kcal mol<sup>-1</sup>), (d) 2,2'- & 4,7'- & 2,2'-position (+1.08 kcal mol<sup>-1</sup>), (e) 2,2'- & 7,2'- & 4,4'-position (+1.18 kcal mol<sup>-1</sup>), (f) 2,2'- & 4,2'- & 7,4'-position (+1.20 kcal mol<sup>-1</sup>), (g) 2,4'- & 2,4'- & 2,2'-position (+1.33 kcal mol<sup>-1</sup>), (h) 2,4'- & 2,2'- & 4,2'-position (+1.50 kcal mol<sup>-1</sup>), (i) 2,2'- & 7,7'- & 2,4'-position (+1.56 kcal mol<sup>-1</sup>), (j) 2,2'- & 7,4'- & 2,7'-position (+1.91 kcal mol<sup>-1</sup>), (k) 2,2'- & 4,2'- & 4,7'-position (+2.03 kcal mol<sup>-1</sup>), (l) 7,2'- & 7,4'- & 2,2'-position (+2.06 kcal mol<sup>-1</sup>), (m) 2,2'- & 4,2'- & 7,7'-position (+2.11 kcal mol<sup>-1</sup>), (n) 2,2'- & 4,7'- & 4,2'-position (+2.22 kcal mol<sup>-1</sup>), (o) 2,4'- & 2,7'- & 2,2'-position (+2.24 kcal mol<sup>-1</sup>), and (p) 2,4'- & 2,2'- & 4,7'-position (+2.33 kcal mol<sup>-1</sup>).

all tetramers are in the porphyrin-like arrangement. Therefore, considering other species of probable molecular structures that represent the chemically diverse nature of these materials, is important to perform more accurate simulations.

### Statistical characteristics of data for trimers and tetramers

Fig. 8a shows the relative energies and ranks of the 216 trimers, from the most to least stable one. Note that the relative energy increases dramatically as the rank goes up. Fig. 8b plots the projection products against the relative energies of all trimers. The projection products quantify the planarity of DHI oligomers, ranging from zero to one, being the latter a completely planar molecular structure (see ESI†). The distribution of points in the cloud shows a trend that the projection product decreases as the relative energy increases. The result implies that those trimers with more planar molecular structures present lower relative energies, and thus they are more stable. Fig. 8c shows the probability of finding certain trimers in a molecular mixture





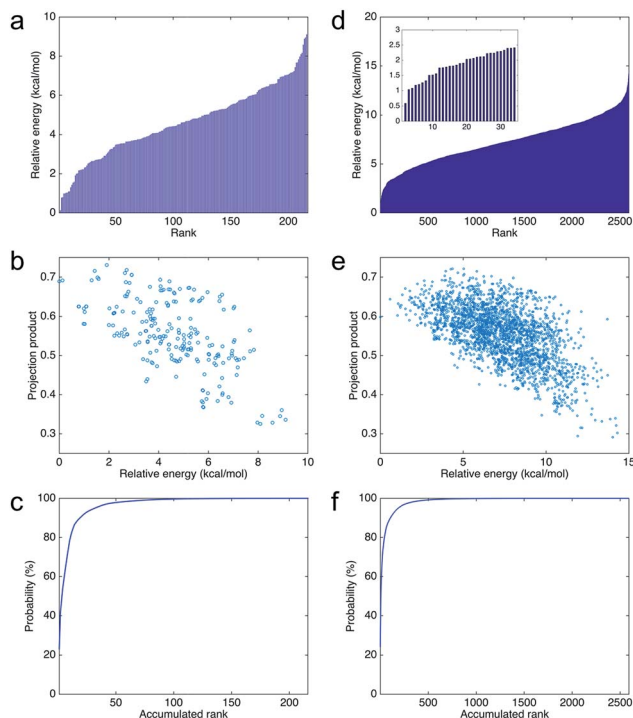


Fig. 8 Statistical characteristics of data for trimers and tetramers. (a) The relative energies and ranks of the 216 trimers, depending on their relative energies, from the most to least stable one. (b) The projection products against the relative energies of all trimers. (c) The cumulative Boltzmann distribution function for all trimers. (d) The relative energies and ranks of the 2592 tetramers, depending on their relative energies, from the most to least stable one. The inner figure zooms in the relative energies of the top 34 ranked tetramers. (e) The projection products against the relative energies of all tetramers. (f) The cumulative Boltzmann distribution function for all tetramers.

of all trimers based on the Boltzmann distribution. The top 12 stable trimers shown in Fig. 6 represent 97% of the population of all trimers. The curve is calculated using the Boltzmann distribution function as:

$$p_n = \frac{\sum_{i=1}^n e^{-\varepsilon_i/kT}}{\sum_{i=1}^m e^{-\varepsilon_i/kT}} \quad (1)$$

where  $p_n$  is the probability of finding the molecular structures that are made of the top  $n$  (e.g., from 1 to 216 for the trimers) stable molecules including their higher energy conformations,  $m$  is the total number of molecules (e.g., 216 for the trimers),  $\varepsilon_i$  is the energy of molecule- $i$  (here we use its relative energy),  $k$  is the Boltzmann constant, and  $T$  is the temperature (here we use 300 K). Fig. 8d shows the relative energies and ranks of the 2592 tetramers. Note that the relative energy increases even more dramatically in the beginning compared to that in the case of trimers (Fig. 8a). The inner figure in Fig. 8d zooms in the relative energies of the top 34 tetramers. These 34 tetramers include the top 16 stable tetramers shown in Fig. 7 and their higher energy conformations. Note that these tetramers are significantly more stable (with lower energy) than the other tetramers.

Therefore, even though there are 162 structural unique tetramers, the majority species of tetramers in PDA and eumelanin could be only a few. A possible explanation is that as oligomers become larger there are more constraints for adding a new monomer. For example, when two monomers are trying to form a dimer, these two monomers can freely rotate against each other, and thus they can adjust their relative positions to the lowest energy configuration. On the other hand, when a new monomer is trying to form a tetramer with a trimer, the monomer is not only interacting with the terminal monomer to which it is attaching but also interacting with the other two monomers *via* non-covalent interactions such as the van der Waals and charge interactions. As a result, the new monomer has to compromise on a higher energy configuration when it attaches to an oligomer. In other words, the degree of freedom for adding a new monomer to an oligomer decreases as the number of DHI units in the oligomer increases due to overlapping electron clouds (steric effects). This result explains the existence of relatively small degrees of polymerization in PDA and eumelanin.

Fig. 8e shows the projection products and relative energies of all tetramers. In the figure, there is a clear trend showing that more planar tetramers present lower relative energies, and thus they are more likely to exist. This finding agrees with the fact that PDA and eumelanin are composed of nearly planar oligomers. Note that the planarity in the molecular structures of PDA and eumelanin is essential to explain the physiochemical properties of eumelanin-like materials. Only if the oligomers are nearly planar, layered aggregates made of stacked oligomers *via*  $\pi$ - $\pi$  interactions can be observed in experiments.<sup>24,46</sup> Furthermore, only if the oligomers are able to stack together closely to form so-called secondary structures, the excitonic effects among the oligomers can be strong enough to produce the broadband absorption spectrum of PDA and eumelanin.<sup>47</sup> Fig. 8f shows the probability of finding certain tetramers in a molecular mixture of all tetramers. Based on the Boltzmann distribution, the top 16 stable tetramers shown in Fig. 7 represent 77% of the population of all tetramers. This result suggests that the majority species of tetramers is quite repetitive even though the number of probable tetramers is large.

### Self-assembly mechanism

Using nearly planar molecular models to simulate the self-assembly mechanism of PDA and eumelanin is important. Fig. 9a shows the DFT optimized geometry of the most planar tetramer (Fig. 7o) among the top 16 stable tetramers (Fig. 7). This tetramer not only is one of the most stable tetramers identified in this work but also has a very high projection product of 0.68 (Fig. 8e). While eight of the tetramers are placed in a simulation box, they quickly stack together to form a layered aggregate shown in Fig. 9b in the simulation (see Computational details section). This type of layered aggregates is one of the most important structural features of eumelanin-like materials. Note that only nearly planar molecular models are able to form such aggregate structures. For example, Fig. 9c shows the DFT optimized geometry of the tetramer that has the



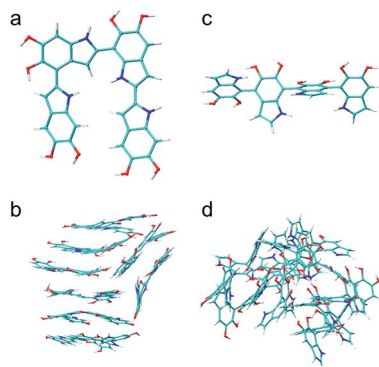


Fig. 9 Molecular and aggregated structures of the most and least stable tetramers. (a) The molecular structure of the most stable tetramer. (b) The layered aggregate formed by 8 of the most stable tetramers shown in (a). (c) The molecular structure of the least stable tetramer. (d) The amorphous aggregate formed by 8 of the least tetramers shown in (c).

lowest projection product of 0.29 (Fig. 8e). Interestingly, this tetramer is also one of the least stable tetramers, which ranks 2589 of 2592. With the same simulation setup, as eight of the tetramers are placed in a simulation box, they also quickly aggregate in the simulation. However, they do not stack together to form a layered aggregate, instead, they form an amorphous structure shown in Fig. 9d. This result clearly demonstrates that choosing more realistic molecular models of PDA and eumelanin for simulations is crucial to get meaningful results.

## Conclusions

In this work, we discuss possible reaction mechanisms of DHI and identify its most reactive positions. For the first time, a first-principle approach is applied to provide a quantum chemical explanation for the existence of certain preferential connections between DHI monomers. Using reactivity descriptors such as the Fukui function and spin densities calculations, we support some of the assumed mechanisms of the DHI polymerization. These results are then adopted as an input for the brute-force algorithmic generator that generates all probable chemical isomers for small DHI oligomers. From a total number of nearly 3000 isomers, the most stable dimers, trimers, and tetramers of DHI are selected to propose molecular models that will help more accurate modeling of PDA and eumelanin. The relative energies between molecular models reported in this work can be applied to estimate a proper ratio of these molecules in simulations. We also provide fundamental explanations for some important structural features of PDA and eumelanin, such as the structural planarity and small size of oligomers. It is important to remark that we have evaluated DHI oligomers only up to tetramers and have not considered different redox forms due to computational limitations. Note that PDA and eumelanin contain a portion of oligomers larger than tetramers, and oxidation has a profound impact on the molecular structures. Consequently, this work cannot completely elucidate the

controversial molecular structures of PDA and eumelanin. However, this work provides some new possible findings and explanations to the structural properties of eumelanin-like materials. Most importantly, the theoretical framework and brute-force algorithmic generator in this work can be easily applied to more realistic situations such as to explore larger oligomers (*e.g.*, pentamers and hexamers), to consider different redox forms (*e.g.*, NQ, MQ, and IQ), and to include DHICA, in future studies when more computational resources are available.

## Computational details

### Reactivity and spin density

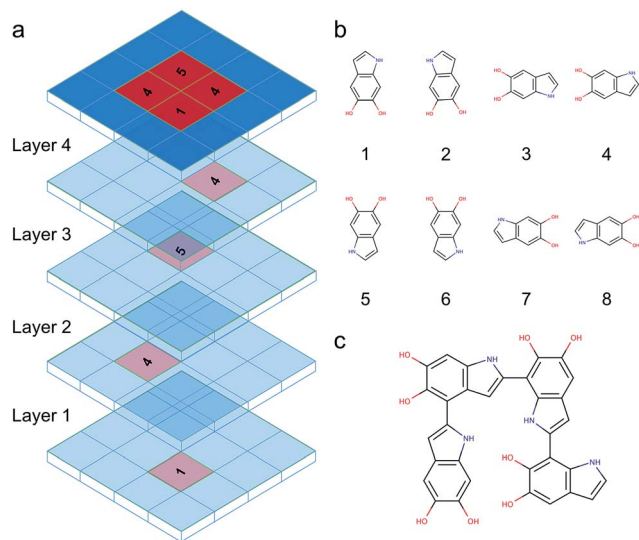
The reactivity of various DHI monomers and oligomers has been characterized using the Fukui function dual descriptor within the framework of conceptual DFT.<sup>48</sup> The Fukui function describes how the electron density changes in response to an increase or a decrease in the number of electrons. Accordingly, it is employed to predict the preferred site for either a nucleophilic or electrophilic attacks. The population of unpaired electrons for different sites in the radicals has been quantified with electron spin densities, for which unrestricted open shell DFT calculations are performed.

### Brute-force algorithmic generator

DHI oligomers are composed of various numbers of DHI units with different configurations. In this work, we use multi-layers of checkerboard plates to represent different polymerization degrees of DHI oligomers. For example, a tetramer can be generated with four layers of checkerboard plates (Fig. 10a). We consider eight different orientations of a DHI monomer (Fig. 10b) to generate DHI oligomers. Orientation 1 is the default orientation. Orientation 2 is a flipped structure of orientation 1. Orientation 3 is a 90 degree clockwise-rotated structure of orientation 1. Orientation 4 is a flipped structure of orientation 3. Similarly, orientation 5 is a 180 degree clockwise-rotated structure of orientation 1. Orientation 6 is a flipped structure of orientation 5. Orientation 7 is a 270 degree clockwise-rotated structure of orientation 1. Orientation 8 is a flipped structure of orientation 7. With these eight different orientations, we can generate all probable DHI oligomers by assembling two or more DHI monomers with different orientations on checkerboard plates. Take dimers for an example, the first monomer in the brute-force algorithmic generator is the default orientation, orientation 1, which is placed on the first layer of a checkerboard plate. Note that it makes no difference which orientation of a DHI monomer is chosen as the first monomer. The brute-force algorithmic generator can generate all probable DHI oligomers no matter which orientation of a DHI monomer is chosen as the starting point. According to our results, the most reactive positions of DHI are the 2-, 4-, and 7-position, and thus we assume that there are three reactive positions in DHI. The first reactive position considered in the brute-force algorithmic generator is the 7-position. There are six orientations of a DHI monomer (*i.e.*, orientation 1, 2, 5, 6, 7, and 8) can form







**Fig. 10** Brute-force algorithmic generator using multi-layers of checkerboard plates to generate different DHI oligomers. (a) Four layers of checkerboard plates for generating a DHI tetramer. The red blocks represent the elements that are occupied for generating a DHI tetramer and the blue blocks represent available elements. The numbers on the red blocks represent the orientation of DHI monomers shown in (b). Eight different orientations of a DHI monomer are considered in the algorithm. (c) The molecular structure of a tetramer that is generated according to the four layers of checkerboard plates shown in (a).

a covalent bond with orientation 1 at the 7-position. The second reactive position considered is the 2-position. There are also six orientations of a DHI monomer (*i.e.*, orientation 3, 4, 5, 6, 7, and 8) can form a covalent bond with orientation 1 at the 2-position. The third reaction position considered is the 4-position. Similarly, there are also six orientations of a DHI monomer (*i.e.*, orientation 1, 2, 3, 4, 5, and 6) can form a covalent bond with orientation 1 at the 4-position. Consequently, a total of 18 combinations of positions and orientations can be generated on the second layer of checkerboard plates. Combining the first and second layers of checkerboard plates, a total of 18 dimers can be generated. Fig. S3<sup>†</sup> shows the checkerboard representations of these 18 dimers. The numbers on the checkerboard plates indicate the orientations of DHI monomers, where the corresponding molecular structures are shown in Fig. S1.<sup>†</sup> While generating trimers, the third layer of checkerboard plates is added to connect the third DHI monomer with the second DHI monomer. Since the second DHI monomer is already formed a covalent bond with the first DHI monomer, there are only two reactive positions available to form a covalent bond with the third DHI monomer. As a result, a total of 12 combinations of positions and orientations can be generated on the third layer of checkerboard plates for an individual dimer. Combining the first, second, and third layers of checkerboard plates, a total of 216 trimers can be generated. The molecular structures of these 216 trimers are shown in Fig. S2.<sup>†</sup> Similarly, a total of 2592 tetramers can be generated in the brute-force algorithmic generator. Fig. 10c shows the molecular structure of a tetramer generated according to the four layers of

checkerboard plates shown in Fig. 10a. Finding the lowest energy conformation of an oligomer is critical for comparing its energy with other oligomers. Note that the brute-force algorithmic generator not only can generate all probable structural unique oligomers but also can generate all different conformations, which oligomers can adopt by rotating about covalent (sigma) bonds. As a result, the conformational analysis is done at the same time while we evaluate the energies of different conformations of oligomers.

### Atomistic modeling and equilibration

Full atomistic MD simulations are performed to relax and equilibrate the initial molecular structures of DHI oligomers generated using the brute-force algorithmic generator. The large-scale atomic/molecular massively parallel simulator (LAMMPS)<sup>49</sup> with the consistent valence force field (CVFF) are adopted in this work. The CVFF has been widely applied in modeling organic molecules with aromatic rings such as benzene, polydopamine,<sup>46,47,50</sup> and caffeine co-crystals.<sup>51</sup> Energy minimization using the conjugate gradient (CG) algorithm is performed to relax the initial molecular structures. After energy minimization, MD simulations are performed to further equilibrate the molecular structures. The integration time step is 1.0 fs and the nonbonding interactions (12–6 Lennard-Jones and coulombic interactions) are computed with a cutoff of 12 Å. The MD equilibration includes 10 iterations of equilibrations in order to find the most stable geometry of each molecule. Each iteration contains four steps of the NVT ensembles. In the first step, a molecule is equilibrated with the NVT ensemble at a temperature starting from 1.0 K and increasing linearly to 300 K in 1.0 ps. In the second step, the molecule is equilibrated with the NVT ensemble at 300 K for 1.0 ps. In the third step, the molecule is equilibrated with the NVT ensemble at a temperature starting from 300 K and decreasing linearly to 1.0 K in 1.0 ps. In the final step, the molecule is equilibrated with the NVT ensemble at 1.0 K for 1.0 ps. A Langevin thermostat is performed to control the temperature. When the fourth step of the NVT ensembles is finished, the molecule is then relaxed with energy minimization again to calculate its energy. After that, its geometry and the corresponding energy are recorded. This set of the NVT ensembles is repeated 10 times. As a result, a total of 10 geometries and the corresponding energies are recorded for each molecule. The geometry with the lowest energy is then selected for further geometry optimizations with DFT. The MD equilibration scheme is shown in Fig. S4.<sup>†</sup>

### Geometry optimization

The geometries obtained from the MD equilibration are further optimized with DFT using ORCA quantum chemistry package.<sup>52</sup> The dispersion correction DFT-D3 with Becke–Johnson damping (D3BJ)<sup>53</sup> is implemented. In addition, PDA and eumelanin are both synthesized in water solution. Therefore, the SMD model,<sup>54</sup> a continuum solvation model based on the quantum mechanical charge density of a solute molecule interacting with a continuum description of the solvent, is adopted to take water effects into account. The B3LYP<sup>36,37</sup> functional is adopted for



optimizing the molecular structures of the dimers, together with the def2-QZVP<sup>38</sup> basis set. However, the B3LYP/def2-QZVP is too expensive for optimizing the molecular structures of the trimers and tetramers. As a result, the BLYP<sup>43,44</sup> functional together with the def2-SVP<sup>38</sup> basis set is adopted for optimizing the trimers and tetramers. A benchmark study is conducted to evaluate the accuracy of using the BLYP/def2-SVP for optimizing the molecular structures of DHI oligomers. The 18 dimers (Fig. S1†) are used for the benchmark. Fig. S5† shows the relative energies of the 18 dimers compared to the dimer-9 (with lowest energy) using the B3LYP/def2-QZVP and BLYP/def2-SVP. The reference values are calculated using the B3LYP/def2-QZVP. The large def2-QZVP basis set minimizes the basis set error and provides the results (blue bars) near the basis set limit for the B3LYP. On the other hand, the BLYP/def2-SVP yields very different results (yellow bars) compared to the reference values. The difference mainly comes from the so-called basis set superposition error (BSSE) when applying small basis sets.<sup>55</sup> In this work, we adopt the recently developed geometrical counterpoise correction (gCP)<sup>56</sup> to circumvent the BSSE. The improvement of using the BSSE-correction gCP is significant. The results (green bars) calculated using the BLYP/def2-SVP/gCP are quite close to the reference values (blue bars). The energy differences between using the B3LYP/def2-QZVP and BLYP/def2-SVP/gCP are within 1 kcal mol<sup>-1</sup>. Since the BLYP/def2-SVP/gCP can provide acceptable results and it is at least 500 times faster than the more accurate B3LYP/def2-QZVP, the BLYP/def2-SVP/gCP is adopted for the geometry optimizations of the trimers and tetramers. The energies obtained in the DFT calculations are used to benchmark the molecular structures of DHI oligomers.

### Self-assembly modeling

Two tetrameric models are used to study the self-assembly mechanism. For each model, eight tetramers are separated by a distance larger than 25 Å in the initial configuration, to ensure that there are no intermolecular interactions. After energy minimization, the system is equilibrated with the NVT ensemble at 300 K for 1.0 ns. The snapshots in Fig. 9b and d are captured after the simulations are finished.

## Acknowledgements

C. T. C. and M. J. B. are supported by CRP Henri Tudor in the framework of the BioNanotechnology project. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575.

## Notes and references

- H. Lee, S. M. Dellatore, W. M. Miller and P. B. Messersmith, *Science*, 2007, **318**, 426–430.
- C. Zhong, T. Gurry, A. A. Cheng, J. Downey, Z. Deng, C. M. Stultz and T. K. Lu, *Nat. Nanotechnol.*, 2014, **9**, 858–866.
- Z. Qin and M. J. Buehler, *Nat. Commun.*, 2013, **4**, 2187.
- Z. Qin and M. J. Buehler, *J. Mech. Phys. Solids*, 2014, **62**, 19–30.
- Z. Qin and M. Buehler, *Appl. Phys. Lett.*, 2012, **101**, 083702.
- Y. Liu, K. Ai and L. Lu, *Chem. Rev.*, 2014, **114**, 5057–5115.
- M. d'Ischia, A. Napolitano, V. Ball, C. T. Chen and M. J. Buehler, *Acc. Chem. Res.*, 2014, **47**, 3541–3550.
- W. Sheng, B. Li, X. Wang, B. Dai, B. Yu, X. Jia and F. Zhou, *Chem. Sci.*, 2015, **6**, 2068–2073.
- V. Ball, D. Del Frari, M. Michel, M. J. Buehler, V. Toniazio, M. K. Singh, J. Gracio and D. Ruch, *BioNanoScience*, 2012, **2**, 16–34.
- S. H. Ku, J. S. Lee and C. B. Park, *Langmuir*, 2010, **26**, 15104–15108.
- H. Gao, Y. Sun, J. Zhou, R. Xu and H. Duan, *ACS Appl. Mater. Interfaces*, 2013, **5**, 425–432.
- Z. Li, X. Zhang, S. Wang, Y. Yang, B. Qin, K. Wang, T. Xie, Y. Wei and Y. Ji, *Chem. Sci.*, 2016, **7**, 4741.
- J. Mu, C. Hou, H. Wang, Y. Li, Q. Zhang and M. Zhu, *Sci. Adv.*, 2015, **1**, e1500533.
- M. E. Lynge, R. van der Westen, A. Postma and B. Stadler, *Nanoscale*, 2011, **3**, 4916–4928.
- W. Qiang, W. Li, X. Li, X. Chen and D. Xu, *Chem. Sci.*, 2014, **5**, 3018–3024.
- B. Wang, G. Wang, B. Zhao, J. Chen, X. Zhang and R. Tang, *Chem. Sci.*, 2014, **5**, 3463–3468.
- S. M. Kang, S. Park, D. Kim, S. Y. Park, R. S. Ruoff and H. Lee, *Adv. Funct. Mater.*, 2011, **21**, 108–112.
- J. Cheng, S. C. Moss, M. Eisner and P. Zschack, *Pigm. Cell Res.*, 1994, **7**, 255–262.
- J. Cheng, S. C. Moss and M. Eisner, *Pigm. Cell Res.*, 1994, **7**, 263–273.
- P. Díaz, Y. Gimeno, P. Carro, S. González, P. L. Schilardi, G. Benítez, R. C. Salvarezza and A. H. Creus, *Langmuir*, 2005, **21**, 5924–5930.
- G. Zajac, J. Gallas, J. Cheng, M. Eisner, S. Moss and A. Alvarado-Swaigood, *Biochim. Biophys. Acta, Gen. Subj.*, 1994, **1199**, 271–278.
- G. W. Zajac, J. M. Gallas and A. E. Alvarado-Swaigood, *J. Vac. Sci. Technol., B: Microelectron. Nanometer Struct.–Process., Meas., Phenom.*, 1994, **12**, 1512–1516.
- C. M. Clancy, J. B. Nofsinger, R. K. Hanks and J. D. Simon, *J. Phys. Chem. B*, 2000, **104**, 7871–7873.
- A. A. Watt, J. P. Bothma and P. Meredith, *Soft Matter*, 2009, **5**, 3754–3760.
- E. Kaxiras, A. Tsolakidis, G. Zonios and S. Meng, *Phys. Rev. Lett.*, 2006, **97**, 218102.
- S. Meng and E. Kaxiras, *Biophys. J.*, 2008, **94**, 2095–2105.
- N. F. Della Vecchia, R. Avolio, M. Alfè, M. E. Errico, A. Napolitano and M. d'Ischia, *Adv. Funct. Mater.*, 2013, **23**, 1331–1340.
- M. d'Ischia, K. Wakamatsu, A. Napolitano, S. Briganti, J. C. Garcia-Borron, D. Kovacs, P. Meredith, A. Pezzella, M. Picardo and T. Sarna, *Pigm. Cell Melanoma Res.*, 2013, **26**, 616–633.
- D. R. Dreyer, D. J. Miller, B. D. Freeman, D. R. Paul and C. W. Bielawski, *Langmuir*, 2012, **28**, 6428–6435.
- M. L. Tran, B. J. Powell and P. Meredith, *Biophys. J.*, 2006, **90**, 743–752.



- 31 S. Hong, Y. S. Na, S. Choi, I. T. Song, W. Y. Kim and H. Lee, *Adv. Funct. Mater.*, 2012, **22**, 4711–4717.
- 32 H. Okuda, K. Yoshino, K. Wakamatsu, S. Ito and T. Sota, *Pigm. Cell Melanoma Res.*, 2014, **27**, 664–667.
- 33 J. Yang, M. A. Cohen Stuart and M. Kamperman, *Chem. Soc. Rev.*, 2014, **43**, 8271–8298.
- 34 E. J. Land, C. A. Ramsden and P. A. Riley, *Acc. Chem. Res.*, 2003, **36**, 300–308.
- 35 B. Kalyanaraman, C. Felix and R. Sealy, *Environ. Health Perspect.*, 1985, **64**, 185.
- 36 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372–1377.
- 37 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
- 38 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.
- 39 L. Panzella, A. Pezzella, A. Napolitano and M. d'Ischia, *Org. Lett.*, 2007, **9**, 1411–1414.
- 40 M. d'Ischia, A. Napolitano, A. Pezzella, E. J. Land, C. A. Ramsden and P. A. Riley, in *Advances in Heterocyclic Chemistry*, ed. R. K. Alan, Academic Press, 2005, vol. 89, pp. 1–63.
- 41 J. r. Liebscher, R. Mrówczyński, H. A. Scheidt, C. Filip, N. D. Hādade, R. Turcu, A. Bende and S. Beck, *Langmuir*, 2013, **29**, 10539–10548.
- 42 A. Pezzella, A. Napolitano, M. d'Ischia and G. Prota, *Tetrahedron*, 1996, **52**, 7913–7920.
- 43 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785.
- 44 A. D. Becke, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1988, **38**, 3098.
- 45 Y. J. Kim, A. Khetan, W. Wu, S. E. Chun, V. Viswanathan, J. F. Whitacre and C. J. Bettinger, *Adv. Mater.*, 2016, **28**, 3173.
- 46 C. T. Chen, V. Ball, J. J. D. Gracio, M. K. Singh, V. Toniazio, D. Ruch and M. J. Buehler, *ACS Nano*, 2013, **7**, 1524–1532.
- 47 C. T. Chen, C. Chuang, J. S. Cao, V. Ball, D. Ruch and M. J. Buehler, *Nat. Commun.*, 2014, **5**, 3859.
- 48 P. Geerlings, F. De Proft and W. Langenaeker, *Chem. Rev.*, 2003, **103**, 1793–1874.
- 49 S. Plimpton, *J. Comput. Phys.*, 1995, **117**, 1–19.
- 50 S. C. Lin, C. T. Chen, I. Bdikin, V. Ball, J. Gracio and M. J. Buehler, *Soft Matter*, 2014, **10**, 457–464.
- 51 C. T. Chen, S. Ghosh, C. M. Reddy and M. J. Buehler, *Phys. Chem. Chem. Phys.*, 2014, **16**, 13165–13171.
- 52 F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2012, **2**, 73–78.
- 53 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.
- 54 A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B*, 2009, **113**, 6378–6396.
- 55 H. Kruse, L. Goerigk and S. Grimme, *J. Org. Chem.*, 2012, **77**, 10824–10834.
- 56 H. Kruse and S. Grimme, *J. Chem. Phys.*, 2012, **136**, 154101.

