Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

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

*Jun Wang and Lluís Blancafort**

anie_202106289_sm_miscellaneous_information.pdf
anie_202106289_sm_dhi_dimers.xyz
anie_202106289_sm_dhi_dimer_data.xlsx

Contents

SI1. Classification scheme 2

SI1.1 Resonance/bonding classification 2

SI1.2 Nomenclature 3

SI2. Computational details 5

SI2.1 Choice of electronic structure method 5

SI2.1.1 **Benchmark of ground-state approach** 6

SI2.1.2 **Unstable tautomers** 9

SI2.1.3 **Wave function stability of *z* structures** 9

SI2.1.4 **Benchmark of excited-state approach** 9

SI2.2 Calculation of free energy 9

SI2.3 Thermodynamic scheme to balance stoichiometry 10

SI3. Full histograms of energy distribution 10

SI4. Most stable linear isomers of each connectivity 13

SI5. Distribution plots for S_1-S_3 excitation energy and oscillator strength 13

SI6. Construction of absorption/stability plots 15

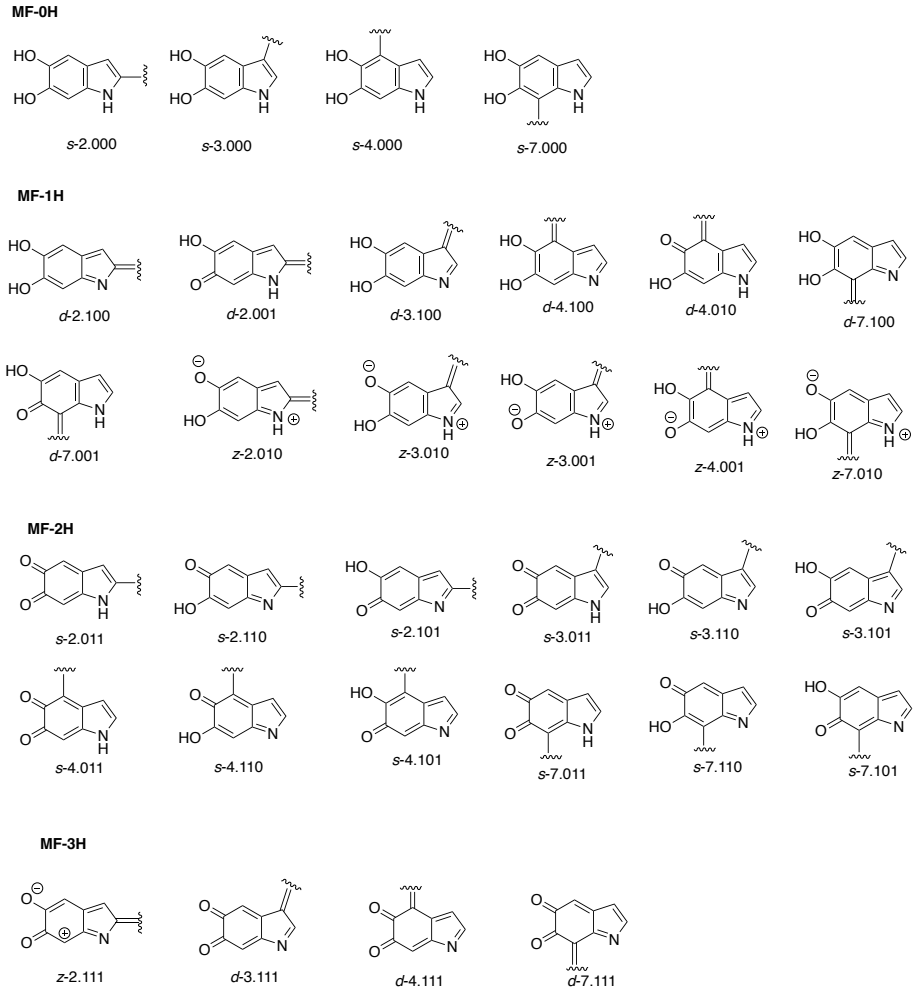
SI7. Analysis of absorptions contributing to highlighted signals in stability *vs* absorption plots 15

SI8. SI References 24

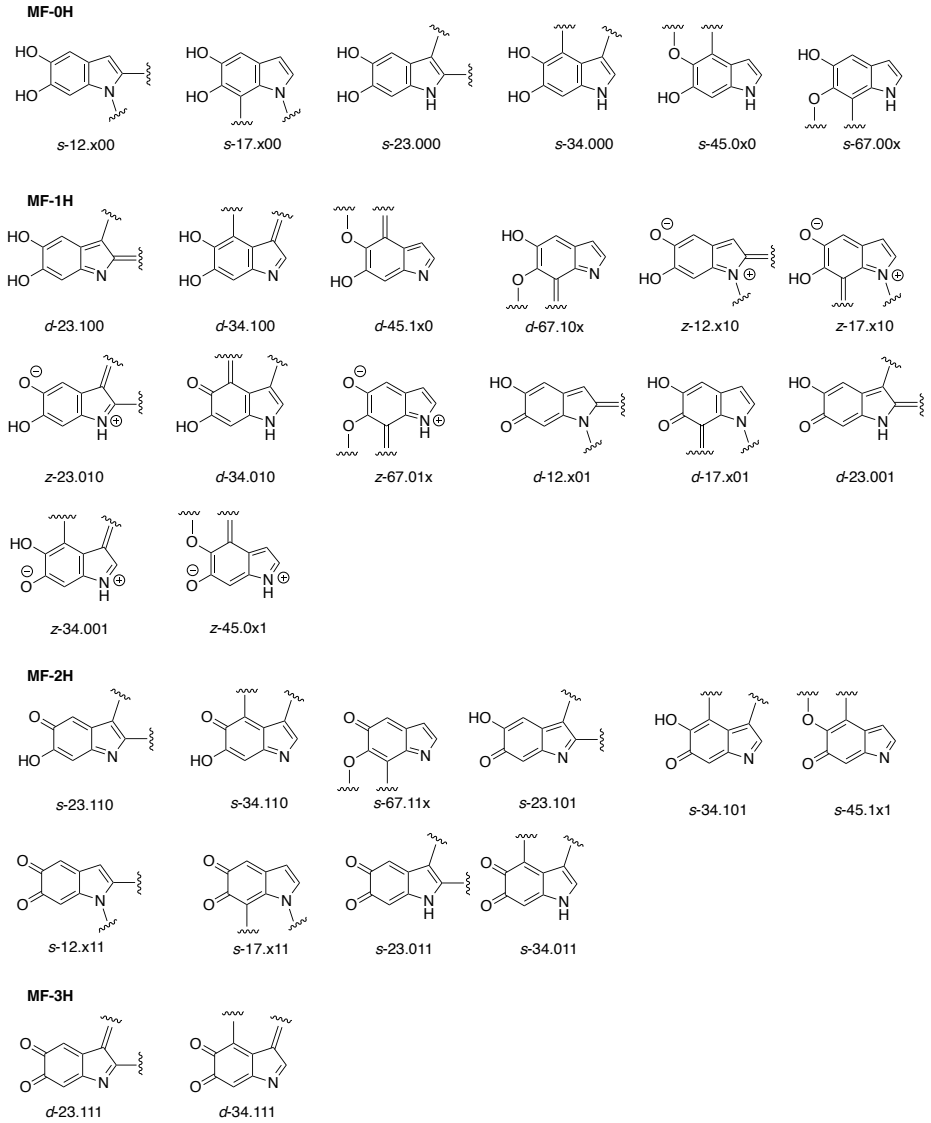
SI1. **Classification scheme**

SI1.1 **Resonance/bonding classification**

The resonance/bonding dimer classification is based on the fragment library shown in Schemes SI1 and SI2. The monomer fragments (MF) are classified as MF-nH, where *n* is the number of oxidized sites, and labelled as *b*-*n*.*ijk*, where *b* is the bond connectivity at the bonding site, *n* the bonding site, and *ijk* codifies whether the N, O_5 or O_6 sites are reduced (digit=0) or oxidized (digit=1). In the cyclic structures where one of the heteroatoms is bonded, the corresponding digit is set to *x*. All linear and cyclic MF-0H and MF-2H fragments can be described by neutral resonance structures with single-bond connectivity at the connecting site(s), and are labeled *s*. In contrast, the MF-1H and MF-3H fragments can be described either by neutral resonance structures with double-bond connectivity at the connecting site(s), which are labeled *d*, or by zwitterionic resonance structures labeled *z*. The rules for the dimers are that two *s* fragments give an *s* dimer; two *d* fragments give a *d* dimer; and two *z* fragments or one *d* and one *z* fragment give a *z* dimer.



Scheme SI1



Scheme SI2

SI1.2 **Nomenclature**

We introduce an unambiguous, specific nomenclature that describes the structural features in a compact way. For the linear structures, the name has the form *r.lin.u.mn.x.abc-def*. The meaning of each string is as follows:

- *r* (resonance/bonding key): *s* (interfragment single bond); *d* (interfragment double bond) or *z* (zwitterionic)
- *lin*: linear dimer
- *u* (stereochemistry of the interfragment bond): *c* (cis) or *t* (trans)
- *mn*: numbering of the carbons forming the interfragment bond
- *x*: total oxidation state of the dimer
- *abc-def* (oxidation state of N_1, O_5 and O_6 of the first and second fragment, respectively): 0 (reduced) or 1 (oxidized)

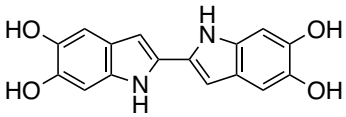
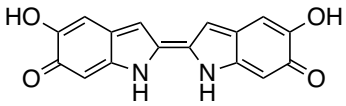
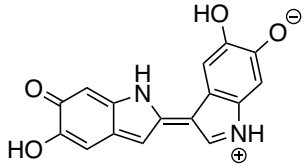
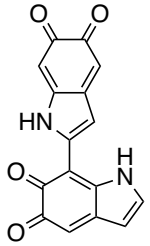
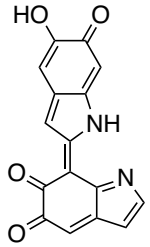
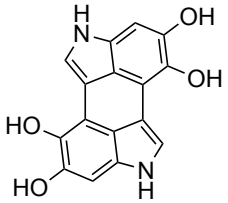
For the cyclic structures, the format is *r.cyc.mn.op.x.abc-def*, where the meaning is:

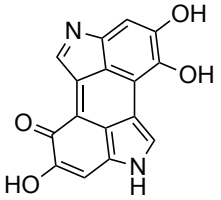
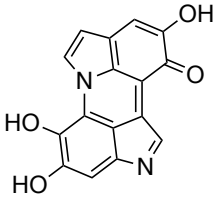
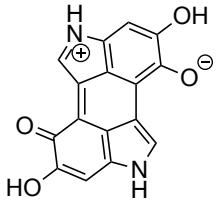
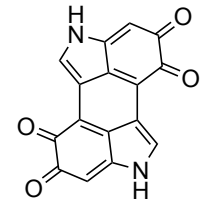
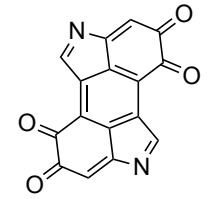
- *r* (resonance/bonding key): *s* (two interfragment single bonds); *d* (one or two interfragment double bond) or *z* (zwitterionic)

- *lin*: cyclic dimer
- *mn.op*: numbering of the carbons forming the two interfragment bonds
- *x*: total oxidation state of the dimer
- *abc-def* (oxidation state of N_1, O_5 and O_6 of the first and second fragment, respectively): 0 (reduced), 1 (oxidized) or X (involved in interfragment bond)

Some examples of nomenclature are given in Table SI2.

Table SI2. Illustration of the nomenclature for relevant structures from Schemes 1 and 2.

Structure	Resonance/ bonding key	Linear/ cyclic	*cis*/ *trans*	Atom connectivity	Total oxidation state	Oxidized sites first ring	Oxidized sites second ring
 s.lin.t.22.0.000-000	s (interfragment single bond)	lin	t	22	0	000	000
 d.lin.c.22.0.001-001	d (interfragment double bond)	lin	c	22	1	001	001
 z.lin.t.23.1.001-001	z (zwitterionic)	lin	t	23	1	001	001
 s.lin.c.27.2.011-011	s (interfragment single bond)	lin	c	27	2	011	011
 d.lin.c.27.001-111	d (interfragment double bond)	lin	c	27	2	001	111
 s.cyc.34.43.0.000-000	s (interfragment single bonds)	cyc	-	34.43	0	000	000

 d.cyc.34.43.1.010-100	d (interfragment double bond)	cyc	-	34.43	1	010	100
 d.cyc.17.43.1.x10-100	d (interfragment double bond)	cyc	-	17.43	1	x10	100
 z.cyc.34.43.1.010-010	z (zwitterionic)	cyc	-	34.43	1	010	010
 s.cyc.34.43.2.011-011	s (interfragment single bonds)	cyc	-	34.43	2	011	011
 d.cyc.34.43.3.111-111	d (interfragment double bond)	cyc	-	34.43	3	111	111

SI2. **Computational details**

SI2.1 **Choice of electronic structure method**

The calculations (ground-state optimizations and excited-state single point calculations) were carried out with density functional theory (DFT), using the CAM-B3LYP functional and the 6-311G(d,p) basis set. This choice is based on the good performance of CAM-B3LYP for the excited-state calculations,[1] since it avoids spurious charge transfer stabilization effects[2] that are likely to occur in the dimers with other functionals such as B3LYP. The performance of CAM-B3LYP for the ground-state optimizations was benchmarked against MP2/cc-pVTZ for the reduced dimers (see SI2.1.1). Solvent effects were taken into account with the SMD continuum model[3] using water as solvent. All calculations were performed with Gaussian 16.[4]

In addition to the 830 reported compounds, we considered six more isomers that were discarded, two because the structure is not stable (optimization led to a different structure after tautomerization, see SI2.1.2) and four because the restricted DFT wave function had an instability indicative of radical character (see SI2.1.3). For every optimized structure, the vertical absorption energy and oscillator strength of the three

lowest excited states is calculated with linear-response time-dependent DFT (TD-DFT). The suitability of TD-CAM-B3LYP was assessed comparing the vertical excitations of the reduced and oxidized monomers with MS-CASTP2 and EOM-EE-CCSD literature values.[5] The differences are approximately 0.2 eV, which lies in the expected precision range for TD-CAM-B3LYP (see SI2.1.4).

SI2.1.1 **Benchmark of ground-state approach**

The validity of our approach to estimate the stability of the different compounds was tested for the set of 20 linear, fully reduced DHI dimers. As a benchmark method we used MP2 with Dunning's correlation-consistent polarized valence triple-z (cc-pVTZ) basis set and SMD solvation. We considered SMD-CAM-B3LYP/6-311G(d,p) with and without the atom-pair wise dispersion D3 correction[6] and SMD-B3LYP/6-311G(d,p) for reference, and compared the relative potential energies of the 22 compounds with the three approaches. The results are shown in Figure SI1, where we display the relative energies following the MP2 energies in decreasing order. In the plots the blue and red bars represent MP2 and DFT energies, respectively. The relative energies span a range of 10 kcal·mol^{-1} with MP2. The CAM-B3LYP energies are somewhat less dispersed, but the range of dispersion is closer to MP2 when the D3 correction is not included (6 *vs* 4 kcal·mol^{-1}, Figures SI1a and SI1b). In addition, the order of the compounds is approximated better with CAM-B3LYP than by CAM-B3LYP-D3. In particular, CAM-B3LYP-D3 clearly underestimates the relative energies of the c.77 and c.47 isomers, which are the less stable ones at the MP2 level and the second and fourth more stable ones with CAM-B3LYP-D3. The root mean square of the difference (RMSD) between the relative energies is 3.1 kcal/mol for CAM-B3LYP and 5.0 kcal/mol for CAM-B3LYP-D3. Overall, the agreement with MP2 is better without the D3 correction. Therefore, the D3 scheme was discarded. Uncorrected B3LYP (see Figure SI1c) has a smaller RMSD for the benchmark set, 1.8 kcal/mol, but given that the difference between CAM-B3LYP and B3LYP is only 1.3 kcal/mol, we have chosen CAM-B3LYP for the optimizations to be consistent with the excited-state calculations. The good agreement between CAM-B3LYP and MP2 is also seen comparing the energy distribution histograms as a function of the type (Figure SI2).

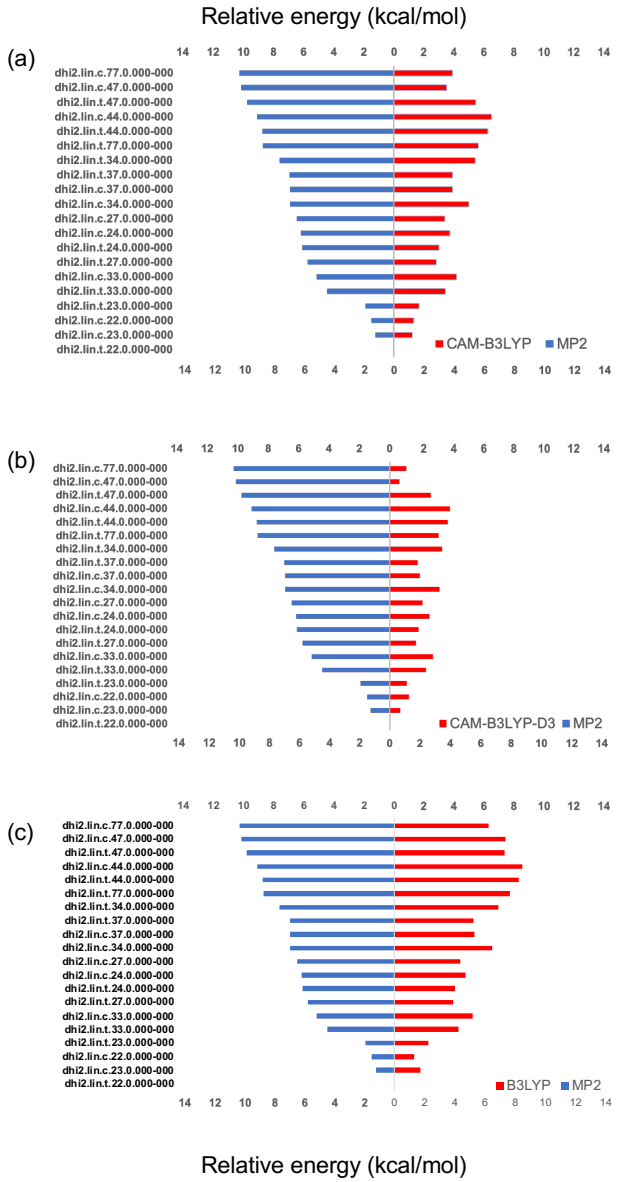


Figure SI1. Relative potential energy of the Ox0-lin compounds calculated at the MP2/cc-pvtz level (blue bars) compared with DFT (red bars): CAM-B3LYP/6-311G** without and with the D3 correction, and B3LYP/6-311G** (panels (a), (b) and (c), respectively).

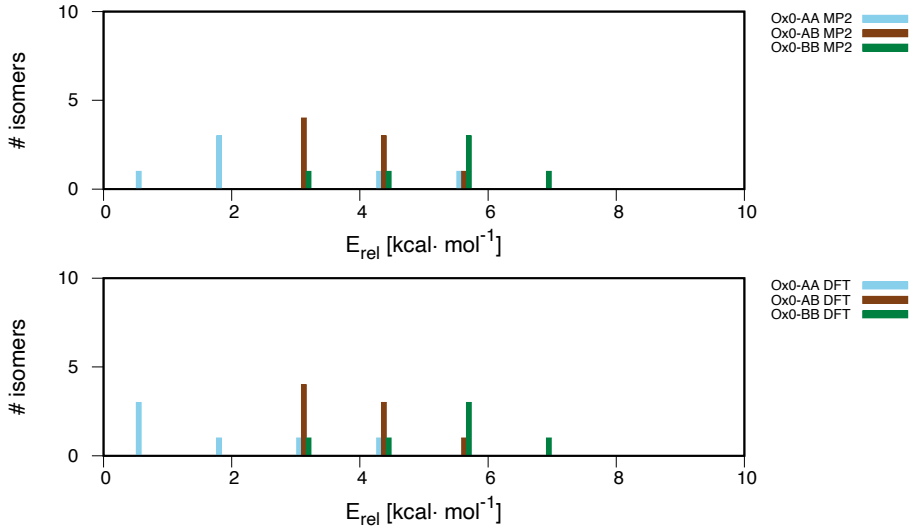


Figure SI2. Energy distribution of Ox0-lin isomers as a function of their type calculated with MP2/cc-pvtz and CAM-B3LYP/6-311G** (upper and lower panels, respectively).

SI2.1.2 **Unstable tautomers**

Two structures (z.lin.t.47.1.010-010 and z.lin.c.77.1.100-010) were not stable towards interfragment tautomerization. Upon optimization, they yielded structures z.lin.t.47.1.000-110 and z.lin.c.77.000-110, respectively. The unstable isomers were not included in the library.

SI2.1.3 **Wave function stability of *z* structures**

The structures with *z* resonance/bonding character may be prone to have a diradical character that makes them unsuitable for our theoretical approach. This possibility was examined reoptimizing all *z* structures with unrestricted CAM-B3LYP and comparing the results with the restricted approach. Four structures (z.lin.c.37.1.001-010, z.lin.c.37.1.010-001, z.lin.t.37.1.100-010 and z.lin.t.44.1.010-001) give a lower energy minimum with unrestricted CAM-B3LYP, which is indicative of instability of the restricted wave function and possible radical bicharacter. The structures were therefore discarded.

SI2.1.4 **Benchmark of excited-state approach**

The suitability of TD-CAM-B3LYP for the vertical excitations is assessed comparing the vertical excitations of the DHI, IQ and NQ monomers in the gas phase with high-level EOM-EE-CCSD and MS-CASPT2 calculations from Ref. [7] (Table SI1). TD-DFT excitations in solution are given in the fifth column. The differences of gas-phase TD-DFT with respect to EOM-EE-CCSD are 0.16 - 0.23 eV, and those with respect to MS-CASPT2 0-16 - 0.28 eV, *ie* overall there is a difference of approximately 0.2 eV which lies within the expected precision range of TD-CAM-B3LYP. Note that inclusion of solvation effects has a substantial effect on the excitation energies of IQ and NQ (decrease by approximately 0.2 eV) because of an increase in the dipole moment upon excitation.

Table SI1. Vertical S_1 excitation energy [eV] of the DHI monomer and the oxidized IQ and NQ forms calculated at different levels of theory.

	EOM-EE-CCSD/ 6-311++G(d,p)[a]	MS-CASPT2/ 6-311++G(d,p)[a]	TD-CAM-B3LYP/ 6-311G(d,p)[b]	TD-CAM-B3LYP/ 6-311G(d,p)[c]
DHI	4.55	4.58	4.78	4.79
IQ	2.33	2.43	2.15	1.92
NQ	2.37	2.05	2.21	2.04

[a]Gas phase, Ref. [7]. [b]Gas phase, this work. [c]Solution (SMD), this work.

SI2.2 **Calculation of free energy**

The free energy *G* of the different compounds is the free energy in solution at 298 K including translational, rotational, and vibrational contributions, and the electronic contribution for O_2. To save computational effort, the vibrational contribution of the dimers has been obtained from the SMD calculated frequencies for representative structures, namely the most stable cyclic and linear isomer of all oxidation states, and the same contribution has been used for all cyclic or linear isomers of the same oxidation state. Similarly, for all dimers the rotational entropy is approximated using the moments of inertia of the representative structures mentioned above. Following Eq. 6a of Ref. [8], for the Gibbs translational free energy correction $T \cdot S_{tr}$, the translational entropy is calculated using the free solvent volume of water with an intermediate value of 7 for

the C_{free} constant. For the dimers and water we assume standard conditions (a_X = 1), while for oxygen we assume an activity of 2.5e-4, which corresponds to the solubility of oxygen in water at 298 K (8 mg·L^{-1})[9]. The latter term is included as RT·ln(a_X) in $G(O_2)$.

SI2.3 Thermodynamic scheme to balance stoichiometry

To compare the energy of dimers with different stoichiometry, we consider that the oxidation proceeds with ground-state (triplet) oxygen, as in one of the common methods of dopa polymerization.[10] In this case, it is possible to compare the energy of isomers of different stoichiometry with the following equation, where the oxidation is balanced with O_2 and H_2O molecules (Equation S1):

$$C_{16}H_{12}N_2O_4 + m/2\ O_2 \longrightarrow C_{16}H_{12-2m}N_2O_4 + m\ H_2O \qquad [S1]$$

In Equation S1, $C_{16}H_{12}N_2O_4$ is the stochiometric formula of the fully reduced linear dimers (oxidation state 0). Using Equation S1, the energy of dimers with different composition can be compared calculating the following energy, which we term G^* (Equation S2):

$$G^*_{C_{16}H_{12-2m}N_2O_4} = G_{C_{16}H_{12-2m}N_2O_4} + mG_{H_2O} - (m/2)G_{O_2} \qquad [S2]$$

The linear dimers with oxidation state 0, 1, 2 and 3 correspond to m = 0, 1, 2 and 3. As for the cyclic dimers, they have 2 hydrogen atoms less than the linear counterparts. Therefore, cyclic dimers with oxidation state 0 - 3 have m = 1 - 4.

SI3. Full histograms of energy distribution

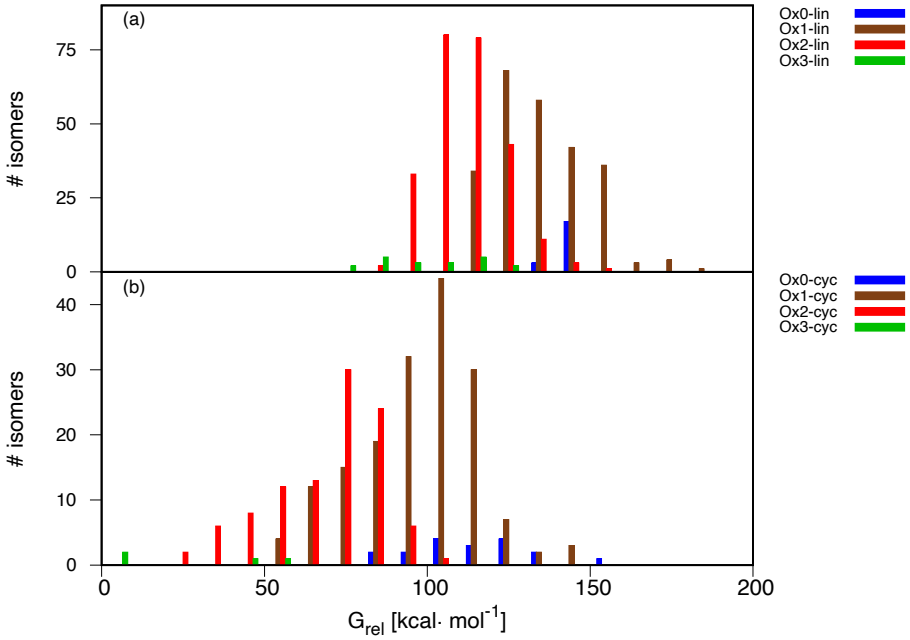


Figure SI3. Free energy distribution of (a) linear and (b) cyclic dimers as a function of their oxidation state, relative to the most stable library dimer (*cf* Figure 1).

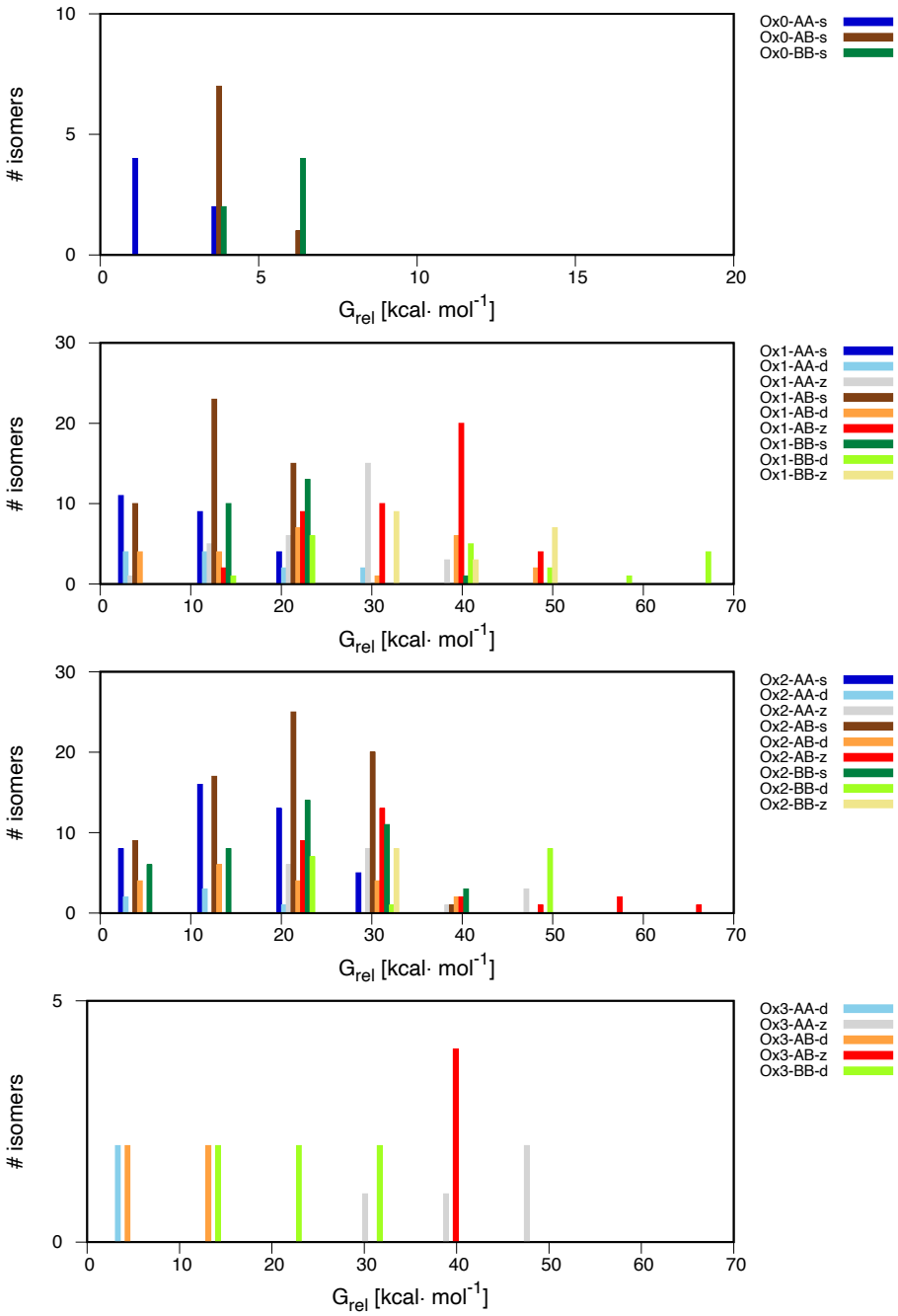


Figure SI4. Free energy distribution of linear isomers as a function of type. Compounds are grouped by oxidation state, and energies are relative to the most stable dimer of each group (*cf* Figure 2).

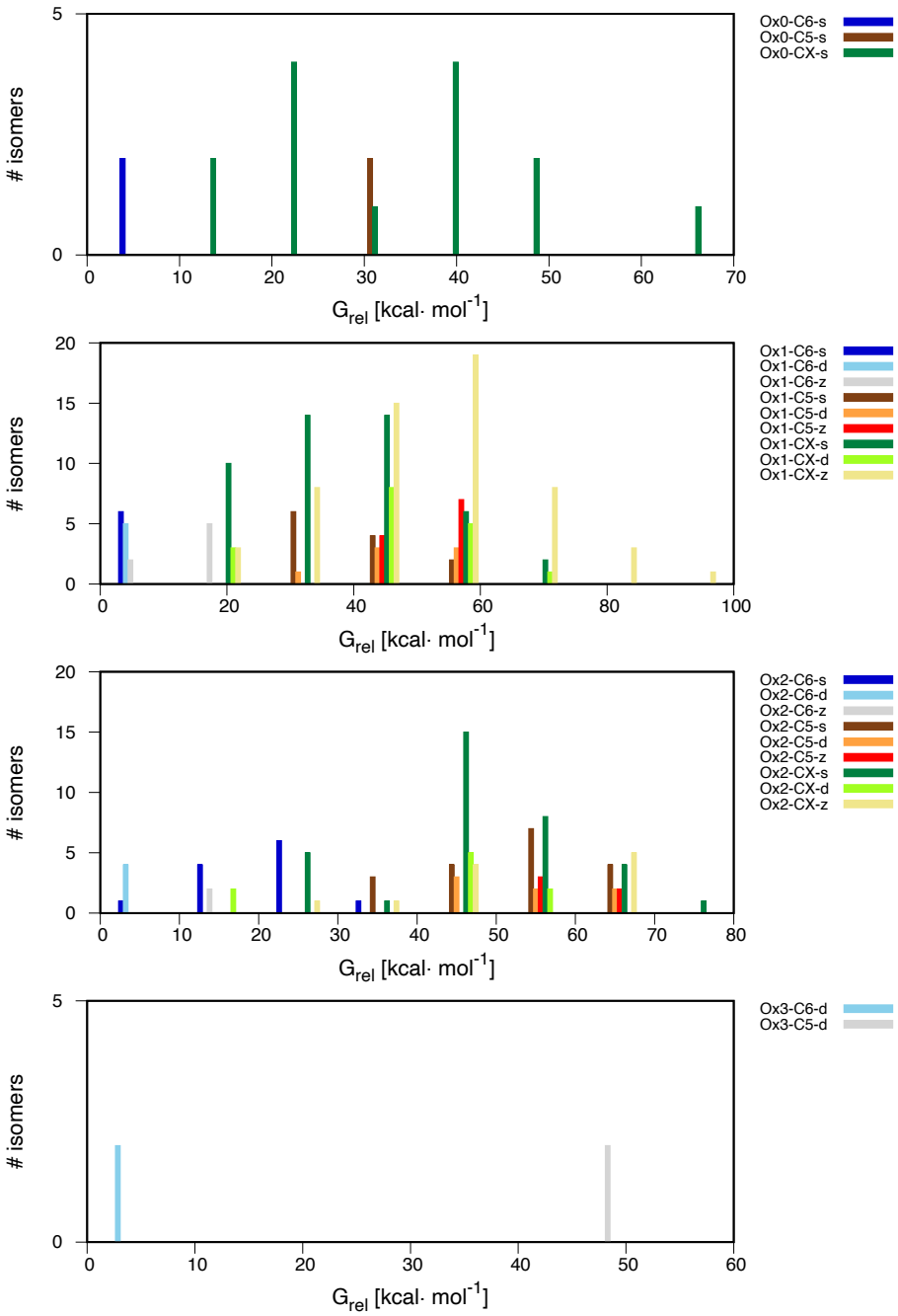


Figure SI5. Free energy distribution of cyclic isomers as a function of type. Compounds are grouped by oxidation state, and energies are relative to the most stable dimer of each group (*cf* Figure 3).

SI4. **Most stable linear isomers of each connectivity**

Table SI2. Most stable linear isomers of each connectivity.

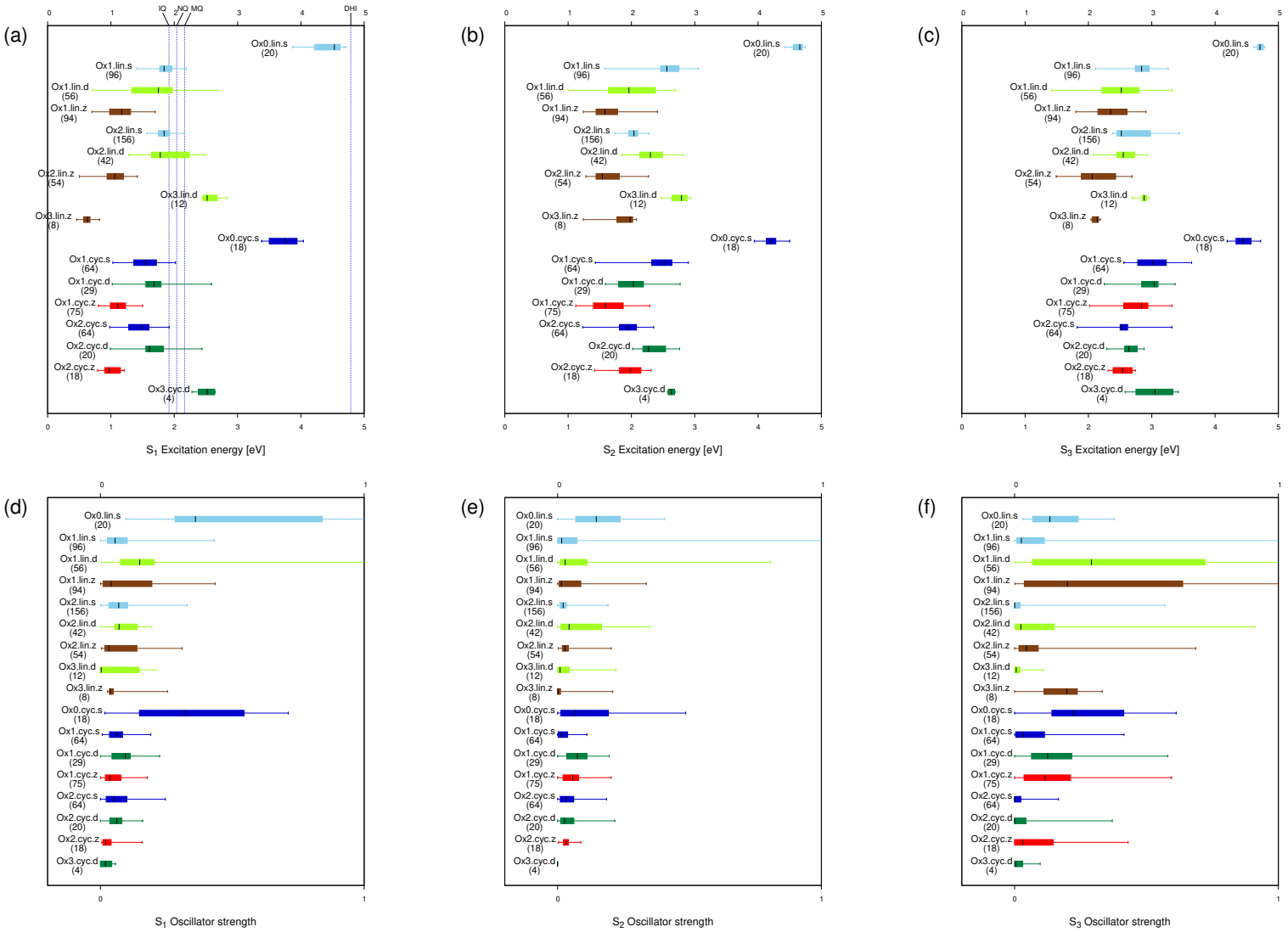
Connectivity	Most stable isomer		ΔG [kcal·mol^{-1}][a]
	Name	Oxidation state	
22	s.lin.t.22.2.011-011	2	-30.7
23	d.lin.t.23.2.001-111	2	-18.5
24	d.lin.c.24.2.001-111	2	-19.2
27	d.lin.t.27.2.001-111	2	-26.0
33	d.lin.t.33.3.111-111	3	-20.6
34	d.lin.t.34.3.111-111	3	-14.8
37	d.lin.c.37.3.111-111	3	-12.0
44	d.lin.t.44.3.111-111	3	-7.5
47	d.lin.c.47.3.111-111	3	-1.5
77	s.lin.t.77.2.011-011	2	-5.2

[a]Energy relative to the lowest energy isomer with Ox3 (22, 23, 24, 27 and 77 connectivity) or Ox2 (33, 34, 37, 44, 47) oxidation state.

SI5. **Distribution plots for S_1-S_3 excitation energy and oscillator strength**

See Figure SI6 in the next page. Caption:

Figure SI6. Distribution of S_1-S_3 excitation energy (a-c, respectively) and oscillator strength (d-f) for the dimers grouped by oxidation state and resonance/bonding, separating linear from cyclic. The filled boxes cover the second and third quartile of each group, and the whiskers the full range from the lowest to the highest absorption. The label includes the number of compounds from each type. Color coding: Linear and cyclic compounds of the same resonance/bonding type are plotted with the same color, *eg* all linear *s* compounds in light blue and all cyclic *s* compounds in dark blue. The vertical dashed lines in (a) are, for reference, the absorption of the DHI monomer and the oxidized IQ, NQ and MQ monomers.


(a)
(b)
(c)
(d)
(e)
(f)
IQ
NQ
MQ
DHI
Ox0.lin.s (20)
Ox1.lin.s (96)
Ox1.lin.d (56)
Ox1.lin.z (94)
Ox2.lin.s (156)
Ox2.lin.d (42)
Ox2.lin.z (54)
Ox3.lin.d (12)
Ox3.lin.z (8)
Ox0.cyc.s (18)
Ox1.cyc.s (64)
Ox1.cyc.d (29)
Ox1.cyc.z (75)
Ox2.cyc.s (64)
Ox2.cyc.d (20)
Ox2.cyc.z (18)
Ox3.cyc.d (4)
S1 Excitation energy [eV]
S2 Excitation energy [eV]
S3 Excitation energy [eV]
S1 Oscillator strength
S2 Oscillator strength
S3 Oscillator strength

SI6. **Construction of absorption/stability plots**

To construct the absorption/stability plots (Figure 5), for every group of compounds a two-dimensional grid of bins is created with a spacing of 0.125 eV (x axis) and 5 kcal·mol^{-1} (y axis) (*eg* 41 x 5 for Figure 5a). The contents of every bin, z_{ij}, is the sum of the oscillator strengths assigned to that bin according to the following formula:

$$z_{ij} = \sum_{p,q} f_{pq} \, \forall \left(G_p \in \left[G_j \right], E_{pq} \in \left[E_i \right] \right) \qquad [S3]$$

where the sum runs over p, the number of compounds in the group, and q=1,3, the number of vertical excitations calculated for every compound. The sum is performed when the energy of compound p, G_p, and the excitation energy E_{pq} fall into the bin relative energy and excitation ranges (G_j and E_i, respectively). After the sum is performed, the values of z are normalized to 1 dividing by the maximum of z_{ij}.

SI7. **Analysis of absorptions contributing to highlighted signals in stability *vs* absorption plots**

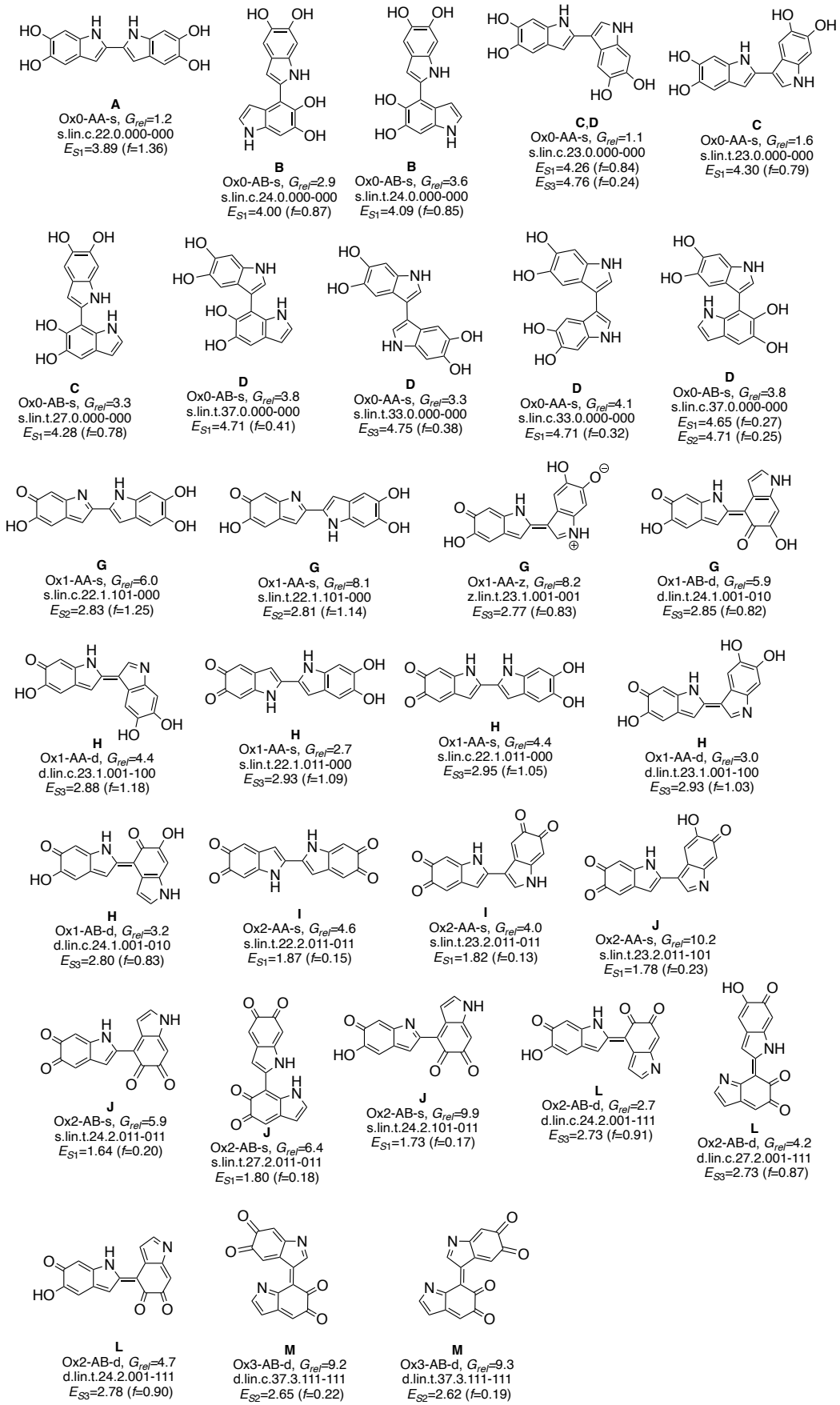
For every highlighted signal of Figure 5, the compounds that contribute most to the absorption (largest oscillator strength) are listed in Tables SI3 and SI4 (linear and cyclic dimers, respectively). The structures are displayed in Schemes 2, 3 and SI3. Orbital plots are provided in Figure SI7.

Table SI3. Most significant contributions to the highlighted absorption signals of the stability *vs* absorption plots, linear dimers (Figure 5a-d).

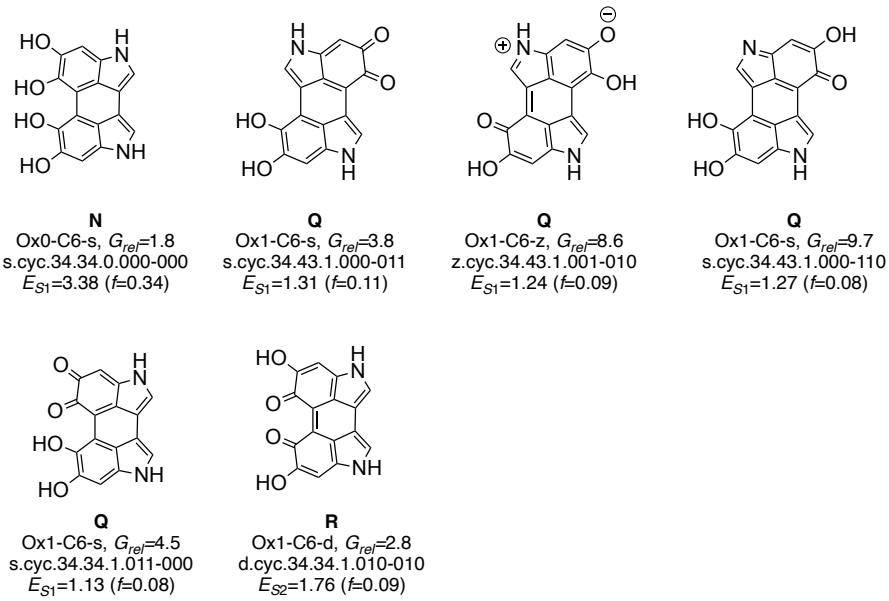
Group	Signal	G_{rel} [kcal·mol^{-1}]	E_{exc} [eV]	f	State	Compound name	Type	Scheme
Ox0-lin	A	0.0	3.87	1.4002	S_1	s.lin.t.22.0.000-000	Ox0-AA-s	2
		1.2	3.89	1.3618	S_1	s.lin.c.22.0.000-000	Ox0-AA-s	SI3
	B	2.7	4.09	0.8725	S_1	s.lin.c.27.0.000-000	Ox0-AB-s	2
		2.9	4.00	0.8724	S_1	s.lin.c.24.0.000-000	Ox0-AB-s	SI3
		3.6	4.09	0.8481	S_1	s.lin.t.24.0.000-000	Ox0-AB-s	SI3
	C	1.1	4.26	0.8383	S_1	s.lin.c.23.0.000-000	Ox0-AA-s	SI3
		1.6	4.30	0.7923	S_1	s.lin.t.23.0.000-000	Ox0-AA-s	SI3
		3.3	4.28	0.7777	S_1	s.lin.t.27.0.000-000	Ox0-AB-s	SI3
	D	3.8	4.71	0.4052	S_2	s.lin.t.37.0.000-000	Ox0-AB-s	SI3
		3.3	4.75	0.3787	S_3	s.lin.t.33.0.000-000	Ox0-AA-s	SI3
		4.1	4.71	0.3159	S_1	s.lin.c.33.0.000-000	Ox0-AA-s	SI3
		3.8	4.65	0.2735	S_1	s.lin.c.37.0.000-000	Ox0-AB-s	SI3
		3.8	4.71	0.2504	S_2	s.lin.c.37.0.000-000	Ox0-AB-s	SI3
		1.1	4.76	0.2414	S_3	s.lin.c.23.0.000-000	Ox0-AA-s	SI3
Ox1-lin	E	8.2	1.70	0.2704	S_1	z.lin.t.23.1.001-001	Ox1-AA-z	2
	F	0.0	2.32	1.7509	S_1	d.lin.c.22.1.001-001	Ox1-AA-d	2
	G	6.0	2.83	1.2455	S_2	s.lin.c.22.1.101-000	Ox1-AA-s	SI3
		8.1	2.81	1.1386	S_2	s.lin.t.22.1.101-000	Ox1-AA-s	SI3
		8.2	2.77	0.8346	S_3	z.lin.t.23.1.001-001	Ox1-AA-z	SI3
		5.9	2.85	0.8222	S_3	d.lin.t.24.1.001-010	Ox1-AB-d	SI3
	H	4.4	2.88	1.1785	S_3	d.lin.c.23.1.001-100	Ox1-AA-d	SI3
		2.7	2.93	1.0876	S_3	s.lin.t.22.1.011-000	Ox1-AA-s	SI3
		4.4	2.95	1.0451	S_3	s.lin.c.22.1.011-000	Ox1-AA-s	SI3
		3.0	2.93	1.0347	S_3	d.lin.t.23.1.001-100	Ox1-AA-d	SI3
		3.2	2.80	0.8264	S_3	d.lin.c.24.1.001-010	Ox1-AB-d	SI3
Ox2-lin	I	4.6	1.87	0.1505	S_1	s.lin.t.22.2.011-011	Ox2-AA-s	SI3
		3.9	1.75	0.1401	S_1	s.lin.c.27.2.011-011	Ox2-AB-s	2
		4.0	1.82	0.1267	S_1	s.lin.t.23.2.011-011	Ox2-AA-s	SI3
	J	10.2	1.78	0.2333	S_1	s.lin.t.23.2.011-101	Ox2-AA-s	SI3
		5.9	1.64	0.2046	S_1	s.lin.t.24.2.011-011	Ox2-AB-s	SI3
		6.4	1.80	0.1776	S_1	s.lin.t.27.2.011-011	Ox2-AB-s	SI3
		9.9	1.73	0.1735	S_1	s.lin.t.24.2.101-011	Ox2-AB-s	SI3
	K	2.6	2.25	0.1723	S_1	d.lin.t.27.2.001-111	Ox2-AB-d	2
	L	2.7	2.73	0.9111	S_3	d.lin.c.24.2.001-111	Ox2-AB-d	SI3
		4.2	2.73	0.8692	S_3	d.lin.c.27.2.001-111	Ox2-AB-d	SI3
		4.7	2.78	0.9011	S_3	d.lin.t.24.2.001-111	Ox2-AB-d	SI3
Ox3-lin	M	9.2	2.65	0.2221	S_2	d.lin.t.37.3.111-111	Ox3-AB-d	SI3
		9.3	2.62	0.1924	S_2	d.lin.c.37.3.111-111	Ox3-AB-d	SI3

Table SI4. Most significant contributions to the highlighted absorption signals of the stability *vs* absorption plots, cyclic dimers (Figure 5e-h).

Group	Signal	G_{rel} [kcal·mol^{-1}]	E_{exc} [eV]	f	State	Compound name	Type	Scheme
Ox0-cyc	N	1.8	3.38	0.3404	S_1	s.cyc.34.34.0.000-000	Ox0-C6-s	SI4
	P	0.0	3.76	0.5672	S_1	s.cyc.34.43.0.000-000	Ox0-C6-s	3
Ox1-cyc	Q	1.3	1.37	0.1701	S_1	z.cyc.34.43.1.010-010	Ox1-C6-z	3
		3.8	1.31	0.1064	S_1	s.cyc.34.43.1.000-011	Ox1-C6-s	SI4
		8.6	1.24	0.0889	S_1	z.cyc.34.43.1.001-010	Ox1-C6-z	SI4
		9.7	1.27	0.0874	S_1	s.cyc.34.43.1.000-110	Ox1-C6-s	SI4
		4.5	1.13	0.081	S_1	s.cyc.34.34.1.011-000	Ox1-C6-s	SI4
	R	0.7	1.90	0.1198	S_1	d.cyc.34.43.1.010-100	Ox1-C6-d	3
		2.8	1.76	0.0924	S_2	d.cyc.34.34.1.010-010	Ox1-C6-d	SI4
		0.0	1.79	0.0897	S_1	s.cyc.34.34.1.110-000	Ox1-C6-s	3
Ox2-cyc	S	9.1	1.31	0.1900	S_1	s.cyc.34.43.2.011-011	Ox2-C6-s	3
Ox3-cyc	T	0.0	3.30	0.0973	S_3	d.cyc.34.43.3.111-111	Ox3-C6-d	3



Scheme SI3. Additional compounds highlighted in Figure 5 (linear dimers). Labelling follows the convention of Scheme 2 and 3 in the main text.



Scheme SI4. Additional compounds highlighted in Figure 5 (cyclic dimers). Labelling follows the convention of Schemes 2 and 3 in the main text.

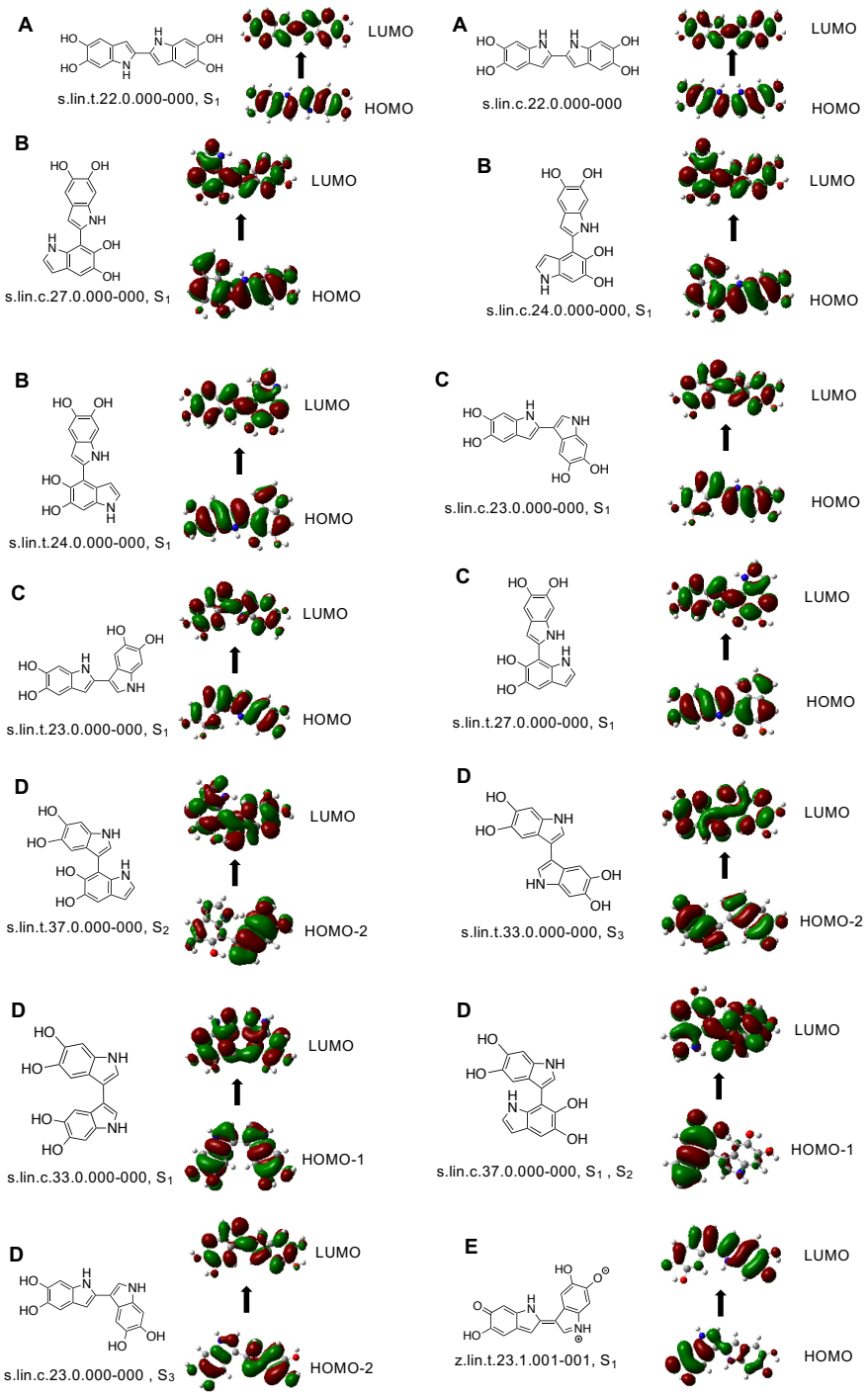


Figure SI7. Orbital plots with the main contribution to the absorptions corresponding to signals A-E in Figure 5.

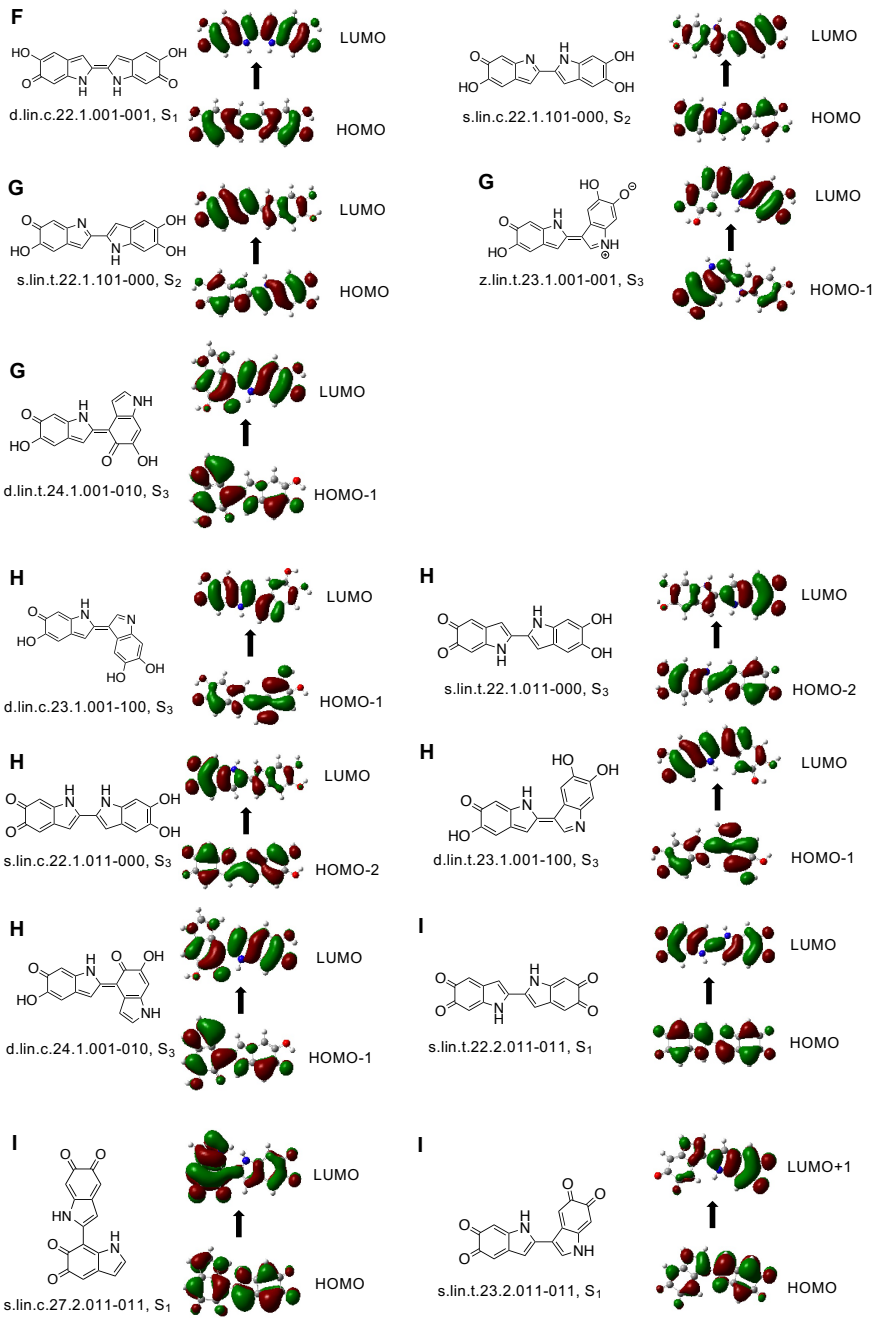


Figure SI8. Orbital plots with the main contribution to the absorptions corresponding to signals F-I in Figure 5.

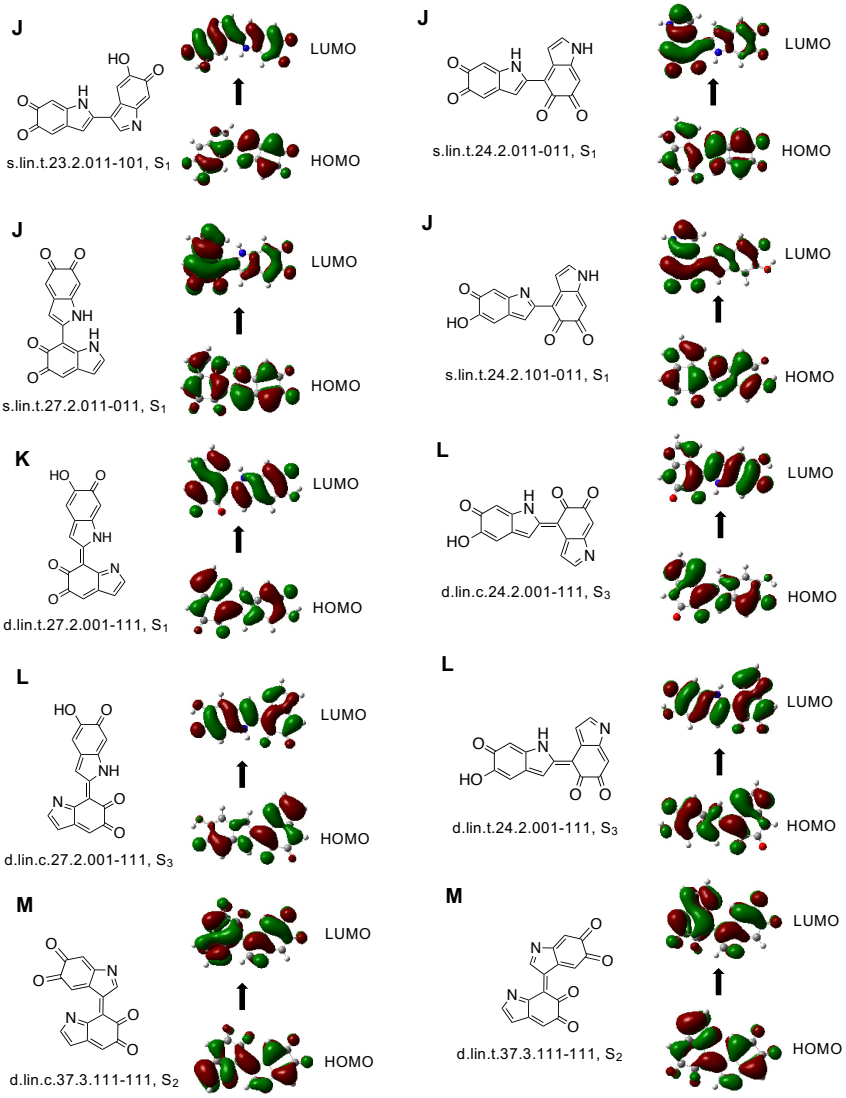


Figure SI9. Orbital plots with the main contribution to the absorptions corresponding to signals J-M in Figure 5.

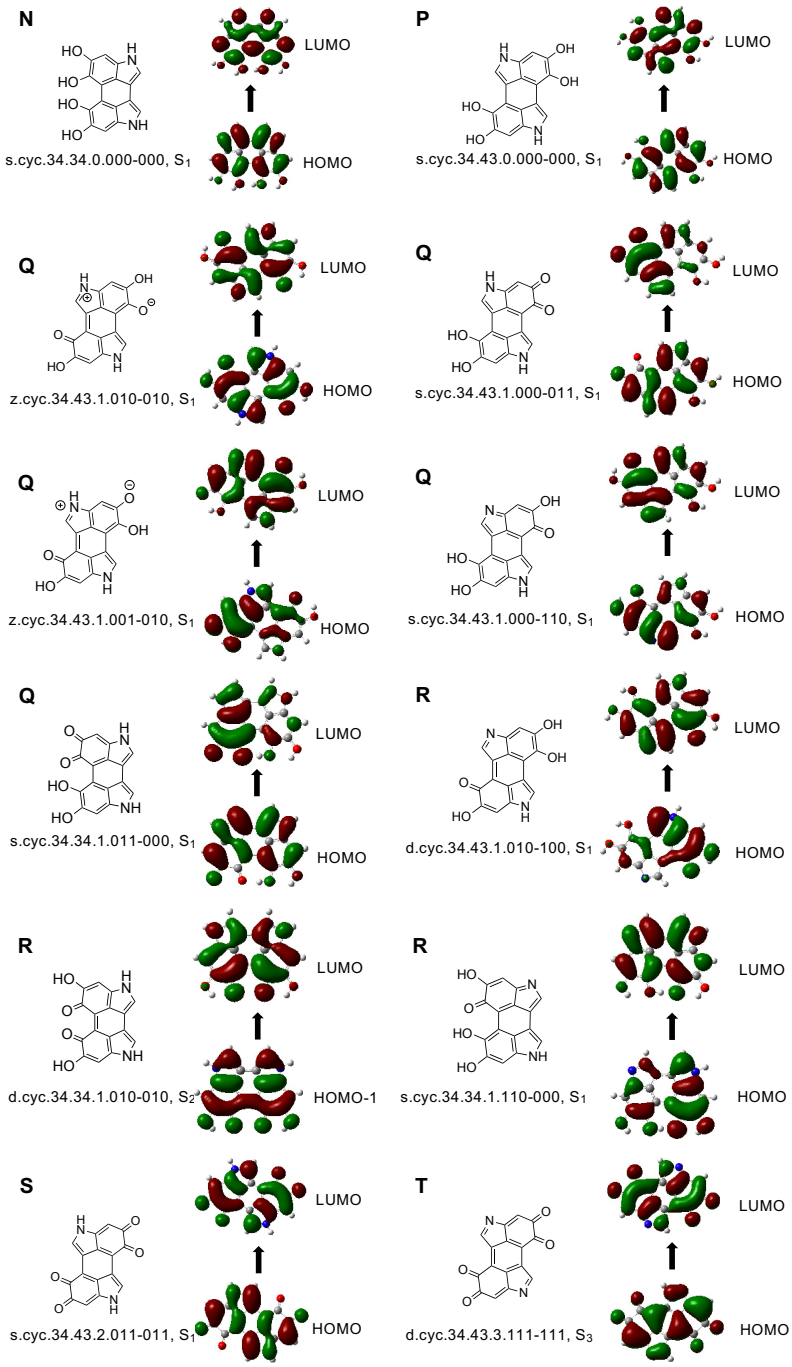


Figure SI10. Orbital plots with the main contribution to the absorptions corresponding to signals N-T in Figure 5.

SI8. **SI References**

[1] a) A. D. Laurent, D. Jacquemin, *Int. J. Quantum Chem.* **2013**, *113*, 2019-2039; b) C. Suellen, R. G. Freitas, P.-F. Loos, D. Jacquemin, *J. Chem. Theory Comput.* **2019**, *15*, 4581-4590.

[2] A. Dreuw, M. Head-Gordon, *Chem. Rev.* **2005**, *105*, 4009-4037.

[3] A. V. Marenich, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. B* **2009**, *113*, 6378-6396.

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

[5] P. Ghosh, D. Ghosh, *J. Phys. Chem. B* **2017**, *121*, 5988-5994.

[6] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*.

[7] D. Ghosh, *WIREs Comput. Mol. Sci.* **2021**, *11*.

[8] M. Mammen, E. I. Shakhnovich, J. M. Deutch, G. M. Whitesides, *J. Org. Chem.* **1998**, *63*, 3821-3830.

[9] M. N. Kutty*, Vol. 2021*, **1987**.

[10] 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.