Supporting Information for

Stacked-ring aromaticity from the viewpoint of the effective number of π-electrons

Ryota Sugimori,† Kenji Okada,† Ryohei Kishi, *,†,§,‡,# Yasutaka Kitagawa †,§,‡,#¶

† Department of Materials Engineering Science, Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan

§ Center for Quantum Information and Quantum Biology Division (QIQB), Osaka University, Toyonaka, Osaka 560-8531, Japan

‡ Research Center for Solar Energy Chemistry (RCSEC), Division of Quantum Photochemical Engineering, Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan

[♯] Innovative Catalysis Science Division, Institute for Open and Transdisciplinary Research Initiatives (ICS-OTRI), Osaka University, Suita, Osaka 565-0871, Japan

¶ Spintronics Research Network Division, Institute for Open and Transdisciplinary Research Initiatives

(OTRI-Spin), Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan

*E-mail: kishi.ryohei.es@osaka-u.ac.jp

Contents

S1. Computational details

S2. Analysis of wavefunctions in the diabatic representation

S3. Effects of the number of states to be solved on the results of energies and wavefunctions

S4. Energies and wavefunctions of CBD dimer models at the CASSCF level

S5. S_0 -T₁ energy gap of cyclobutadiene monomers with D_{2h} and D_{4h} symmetries.

S6. Weight of electron configuration in the ground state for Nc(1) and Nc(2) calculated with different sizes of the active space in the SA-CASSCF and QD-

NEVPT2.

S7. Cartesian coordinates for optimized geometries

S1. Detailed procedure of multi-configuration wavefunction analysis based on a diabatic basis.

Here, we explain detailed procedure of multi-configuration wavefunction analysis based on a diabatic basis. The total energies and wavefunctions of the lowest two ${}^{1}A_{g}$ singlet states of these dimer models were evaluated at the quasi-degenerated NEVPT2 $(QD-NEVPT2)^{1,2}$ level using the active orbitals localized on each monomer. There is a discussion on several types of non-invariance problem in strongly-contracted QD-NEVPT2.³ However, the problem due to noninvarience is expected not to change the conclusion of the present study.³ The practical procedure for cyclobutadiene(CBD) dimer models is as follows: First, the canonical molecular orbitals (CMOs) were optimized at the ten-state-averaged(SA10)-CASSCF(8e,8o) level. This active orbital space includes all the valence π electrons and orbitals of the CBD dimer [see Figure S1(a)]. Judging from their spatial distributions and phases, we classified these active CMOs into three subclasses, i.e., (i) CMO25-26, (ii) CMO27-30, and (iii) CMO31-32. We obtained eight localized MOs (LMOs) by performing the Pipek-Mezey localization method⁴ and unitary transformation for each subclass independently [LMO1-8 in Figure S1(b)].

For Ni(II) norcorrole dimer models (**Nc(1)** and **Nc(2)**), eight intermolecular localized orbitals were obtained by the almost same procedure as the CBD dimer models[LMO1-8 in Figure S2].

We obtained the lowest ten singlet states from the CASCI(8e,8o) calculation based on these active LMOs(*without* performing any additional orbital optimizations), followed by the QD-NEVPT2 calculations. We obtained the CASCI-type wavefunctions in the diabatic representation that diagonalize the NEVPT2 effective Hamiltonian. The strongly contracted version of NEVPT2 $(SC-NEVPT2)^{5,6}$ was employed in these calculations. The cc-pVTZ basis set was employed for the CBD dimer, while the madef2-SVP-SDD and auxiliary basis set were employed for the resolution of the identity (RI) approximation for the Norcorrole dimer models. All these calculations were performed using the ORCA 4.2 program package.⁷

Figure S1. The active canonical molecular orbitals (CMOs) for the CBD dimer models obtained at the SA10-CASSCF(8e,8o) calculations (a), and the active localized molecular orbitals (LMOs) obtained by localizing the active CMOs using the Pipek-Mezey method (b).

Figure S2. The active localized molecular orbitals (LMOs) of Ni(II) norcorrole dimer models **Nc(1)** obtained by localizing the active CMOs obtained at the SA10- CASSCF(8e,8o) calculations using the Pipek-Mezey method.

S2. Analysis of wavefunctions in the diabatic representation

Here, we explain the crucial electron configurations of the CASCI wavefunctions in the diabatic representation. Although we performed calculations based on the CAS(8e,8o) space, we focused on the configurations involved in the CAS(4e,4o) space, i.e., those composed of HOMOs and LUMOs of CBD monomers (LMO3-6). All the electron configurations involved in the CAS(4e,4o) spaces that contribute to the singlet state wavefunctions are shown in Figure S-3. Here, **G**, **D**, **T,** and **E** represent the ground singlet, doubly excited singlet, and (spin-adapted) singly excited triplet, (spin-adapted) singly excited singlet configurations (configuration state functions, CSFs) of each monomer (thus, three types of $S = 1$ configurations with $M_S = -1$, 0, and 1 are included in **T**). Thus, for example, **GG** corresponds to the configuration where HOMO is doubly occupied in both monomers. **GD** and **DG** represent the configurations where LUMO is doubly occupied in one of the monomers. **TT** represents the ${}^{1}(T_{1}T_{1})$ configuration. Note that since monomeric **T** and **E** are the spin-adapted CSFs, **TT** and **EE** were constructed by taking care of the monomeric and dimeric spin-symmetries. Namely, **TT** and **EE** are represented by the following equation using the HOMOs ($h_A = LMO3$, $h_B =$ LMO4) and LUMOs $(l_A = LMO5, l_B = LMO6)$ of the monomers.

$$
\mathbf{TT} = \frac{1}{\sqrt{3}} \left(\left(|h_A l_A \bar{h}_B \bar{l}_B| + |\bar{h}_A \bar{l}_A h_B l_B| \right) - \frac{1}{2} \left(|h_A \bar{l}_A h_B \bar{l}_B| + |h_A \bar{l}_A \bar{h}_B l_B| + |\bar{h}_A l_A h_B \bar{l}_B| + |\bar{h}_A l_A \bar{h}_B l_B| \right) \right)
$$

$$
\mathbf{EE} = \frac{1}{2} (|h_A \bar{l}_A h_B \bar{l}_B| - |h_A \bar{l}_A \bar{h}_B l_B| - |\bar{h}_A l_A h_B \bar{l}_B| + |\bar{h}_A l_A \bar{h}_B l_B|)
$$

CT represents the charge-transfer configurations and there are ten types of (spin-adapted) **CT** configurations (**CT1-10**). These **CT** configurations are further classified into three subclasses, **CTA**, **CTB**, and **CTC**. **CTA** relates to **TT** by the one-electron transfer process from the HOMO (or LUMO) of one monomer to the HOMO (or LUMO) of the other monomer. **CTB** relates to **TT** by the one-electron transfer process from the HOMO (or LUMO) of one monomer to the LUMO (or HOMO) of the other monomer. **CTC** relates to **TT** by the two-electron transfer process. The weights of electrotonic configurations and energies for the $1¹$ Ag and $2¹$ Ag states given in Figure 3 and Figure 4

were obtained by diagonalizing the NEVPT2 effective Hamiltonian matrix (*H*d) represented in these configuration bases (Figure S4).

Figure S-3. All the electron configurations involved in the CAS(4e,4o) spaces that contribute to the singlet state wavefunctions.

Figure S-4. Schematic representation of Hamiltonian matrix(H_d) of CAS(4e,4o) in the diabatic representation and Hamiltonian matrix (*Hd*') partially diagonalized within the matrix block in H_d composed of GG, GD, DG, and DD configurations.

Suppose that $|\psi_G(A)\rangle$ and $|\psi_G(B)\rangle$ are the singlet ground state wavefunctions of the isolated monomer A and B, respectively. They can be represented by the superposition of the **G** and **D** configurations (C_G , $C_D \ge 0$),

$$
|\psi_G(A)\rangle = C_G \mathbf{G} - C_D \mathbf{D}
$$

$$
|\psi_G(B)\rangle = C_G \mathbf{G} - C_D \mathbf{D}
$$

where $C_G > C_D$ for the D_{2h} monomer, and $C_G = C_D$ for the D_{4h} monomer. To understand the interactions between these configurations step-by-step, we first consider a partially diagonalized form (*Hd*') of the Hamiltonian matrix for the matrix block composed of **GG**, **GD**, **DG**, and **DD** (Figure S4, right). The lowest eigenfunction within this subspace, **OS1**, is represented by a linear combination of the direct product of $|\psi_{G}(A)\rangle$ and $|\psi_G(B)\rangle$.

$$
\boldsymbol{OS1} = |\psi_G(\mathbf{A})\rangle \otimes |\psi_G(\mathbf{B})\rangle = C_G^2 \boldsymbol{GG} - C_G C_D (\boldsymbol{GD} + \boldsymbol{DG}) + C_D^2 \boldsymbol{DD}
$$

Since $C_G = C_D$ is satisfied in the D_{4h} monomers, the weights of **GG**, **GD**, **DG**, and **DD** in **OS1** are equal in the D_{4h} dimer. In the case of D_{2h} dimer, C_G is much larger than C_D , and thus, **OS1** is mainly described by **GG**.

Then, we discuss the weights of electron configurations in the $1¹A_g$ and $2¹A_g$ states of the *D*2h and *D*4h-dimer models as a function of stacking distance *d*, based on the configuration interactions in H_d ^t that can be interpreted by the electron transfers between the monomers. For large *d*, such intermonomer interactions are negligible and the lowest singlet state is more stable than the lowest triplet state in both the *D*2h and *D*4h

monomers. Therefore, at a sufficiently large d, we can approximate the wave functions of the $1¹A_g$ state of the CBD dimer models by that of **OS1**. Indeed, as shown in Figures 3(b) and Figure 4(b), the 1^1A_g states of the D_{2h} and D_{4h} dimer in large *d* regions are composed of **GG**, **GD**, **DG**, and **DD** configurations. As we expected, the **GG**, **GD**, **DG**, and **DD** almost equally contribute to the 1^1A_g state for the D_{4h} dimer. In contrast, the 21 Ag state wavefunction in the large *d* region is primarily represented by **TT**, since **T** is the lowest excited state of the isolated monomer.

As *d* decreases, the interactions between the configurations become considerable: The weight of **TT** increases at a small d in the $1¹A_g$ states (see Figures 3(b) and Figure 4(b)). **TT** relates to **GD** or **DG** by the HOMO→HOMO/LUMO→LUMO, to **GG** by the HOMO→LUMO, and to **DD** by the LUMO→HOMO intermonomer two-electron transfer processes. However, in eclipsed dimer models, the intermonomer HOMO→ LUMO transfer integrals are expected to vanish due to the symmetry. Thus, it is expected that the weight of TT in the $1¹A_g$ states increases when the direct coupling between **TT** and **GD** or **DG** in **OS1** increases. The weight of **TT** in the 1^1A_g state begins to increase in the D_{4h} dimer model around $d = 3.4$ -4.2 Å where the intermolecular orbital overlaps are still small, meaning that the direct coupling between **TT** and **GD** or **DG** in **OS1** is considerable around this region. This tendency probably originates from to the small energy gap between the **TT** and **OS1** configurations. However, according to the recent computational results for the S_0-T_1 energy gap of the isolated D_{4h} monomer by Monino et al., $⁸$ we speculate that the energy gap can be underestimated to some extent at</sup> the present QD-NEVPT2 level. Thus, the weight of **TT** in the $1¹A_g$ state of the D_{4h} dimer model around $d = 3.4$ -4.2 Å would be overestimated at the present level of approximation.

On the other hand, in the D_{2h} dimer model around $d = 3.4{\text -}4.2$ Å, the weight of **TT** in the 1^1A_g state remained small and increased around $d \leq 3$ Å where the intermolecular orbital overlaps are considerable. From this result, the direct coupling between **TT** and **GD** or **DG** in **OS1** can be small in the D_{2h} dimer model, probably due to the larger energy gap between these configurations. Thus, **CT** configurations would play crucial roles in the mixing of TT in the 1^1A_g state at small *d*. Indeed, CTA relates not only to TT but also to **GD** or **DG** via HOMO→HOMO or LUMO→LUMO one-electron transfer process. **CTA** relates to **TT** and **GD** or **DG** via HOMO→HOMO or LUMO→LUMO two-electron transfer process. In general, the coupling strengths via the one-electron transfer processes are much larger than those via the two-electron transfer ones. **CTB** and **CTC**

relate to **TT**, **GG**, and **DD** configurations via the HOMO→LUMO or LUMO→HOMO electron transfer process, but the corresponding transfer integrals vanish due to the symmetry of the dimer models. Therefore, we expect that **CTA** can efficiently help mix **TT** in the 1^1A_g state at small *d*. In other words, lowering the lowest **CTA** energy level with decreasing *d*, taking along with the **TT** contributions via **CTA**-**TT** electronic coupling, is a key to understand the energy stabilization mechanism.

For large *d*, the lowest **CTA** energy level is higher than those of **TT** and **OS1**, and as *d* decreases, the **CTA** level becomes lower and start to couple to **TT** that constitutes the 2^1A_g state. Indeed, from Figure 3(b) for the D_{2h} dimer, the weight of $CT(CTA)$ starts to increase in the 2^1A_g state $(2.8 \text{ Å} < d < 4.0 \text{ Å})$. Owing to the **CTA-TT** coupling, the energy of **CTA-TT**-like 2^1A_g state becomes lower with decreasing *d*. As a result, the energy-level crossing occurs. Even though **CTA** can also couple to **OS1**, **CTA-TT** coupling is expected to be larger than **CTA**-**OS1** coupling (since **CTA** can couple to **OS1** only via **DG** and **GD** in **OS1**, not via **GG** and **DD**). Thus, after the crossing point, **TT** and **CT**(**CTA**) strongly contributed to the 11 Ag state via **CTA-TT** coupling; **GD** and **DG** also contributed to the 1^1A_g state to some extent via CTA-OS1 coupling; 2^1A_g state is described by **GG**.

S3. State number dependencies of energy and weight of electron configurations as a function of stacking distance.

In this section, we discuss the state number dependencies of QD-NEVPT2 results of the CBD dimer models since the reference SA-CASSCF (and QD-NEVPT2) results often depend strongly on the number of states to be averaged. Figures S-5, S-6, and S-7 show the results of the *D*2h and *D*4h dimer models at the QD-NEVPT2 level based on the SA5- CASSCF(8e,8o), SA10-CASSCF(8e,8o), and SA20-CASSCF(8e,8o) references. From these results, SA10- and SA20-CASSCF results are almost close.

At $d \le 2.2$ Å, the QD-NEVPT2 with the SA5-CASSCF(8e,8o) reference cannot describe the singlet excited state of interest for the CBD *D*4h dimer model due to its significant destabilization (see Figure S-7). In addition, at $d \le 2.2$ Å, the QD-NEVPT2 with the SA20-CASSCF(8e,8o) reference failed to give smooth variations in the energy and weight of electron configurations for the CBD *D*4h dimer models. These results suggest that taking full the π -valence orbitals is no longer sufficient for describing the low-lying singlet states at such a small *d* region. On the other hand, for 2.2 $\mathring{A} \leq d \leq 5.0$ Å, the QD-NEVPT2 calculations with the SA10-CASSCF(8e, 8o) reference are expected to describe the electronic structures of low-lying singlet states qualitatively and quantitatively.

Figure S-5. The energy of the CBD *D*2h and *D*4h dimer models as a function of *d* calculated at the QD-NEVPT2 level of theory based on the SA5-CASSCF(8e, 8o), SA10- CASSCF(8e, 8o), SA20-CASSCF(8e, 8o) references.

Figure S-6. The weights of electron configurations of the CBD *D*2h dimer models as a function of *d* calculated at the QD-NEVPT2 level of theory based on the SA5- CASSCF(8e, 8o), SA10-CASSCF(8e, 8o), and SA20-CASSCF(8e, 8o) references.

Figure S-7. The weights of electron configurations of the CBD *D*4h dimer models as a function of *d* calculated at the QD-NEVPT2 level of theory based on the SA5- CASSCF(8e, 8o), SA10-CASSCF(8e, 8o), SA20-CASSCF(8e, 8o) references.

S4. Energies and wavefunctions of CBD dimer models at the CASSCF level

Figure S-8. Energies of 1^1A_g and 2^1A_g states of the D_{2h} -dimer model as a function of stacking distance *d* (a) and the weights of electron configurations (b) calculated using CASSCF(8e, 8o)/cc-pVTZ level of theory.

Figure S-9. Energies of 1^1A_g and 2^1A_g states of the D_{4h}-dimer model as a function of stacking distance *d* (a) and the weights of electron configurations (b) calculated using CASSCF(8e, 8o)/cc-pVTZ level of theory.

S5. S_0 -T₁ energy gap of cyclobutadiene monomer with D_{2h} and D_{4h} symmetries.

Table S1. Total energies of the lowest singlet $(E(S_0))$ and triplet $(E(T_1))$ states and the energy difference $E(S_0) - E(T_1)$ for CBD monomer with the D_{2h} and D_{4h} symmetries calculated at SA-4-QD-NEVPT2(4,4)/cc-pVTZ level.[#]

	D_{2h} structures	D_{4h} structures
$E(S_0)$ [Hartree]	-154.3338905	-154.3176726
$E(T_1)$ [Hartree]	-154.2838901	-154.316722
$E(S_0) - E(T_1)$ [kcal/mol]	-31.4	-0.60

The state-averaging (SA-4) was considered for the three singlet and one triplet states.

S6. Weight of each electron configuration in the ground state for Nc(1) and Nc(2) calculated with different sizes of the active space in the SA-CASSCF and QD-NEVPT2.

Table S2. Weight [%] of each electron configuration in the ground state for **Nc(1)** calculated with different sizes of the active space in the SA-CASSCF and QD-NEVPT2.

	CASSCF(4,4) ^b	CASSCF(8,8) ^b	CASECF(12,12) ^b	QD-NEVPT2 $(4,4)^b$	QD-NEVPT2 $(8,8)^b$
GG	0.7	3.6	0.1	1.9	0.1
GD	5.1	7.9	3.6	9.8	9.1
DG	5.1	8.6	3.6	9.8	9.1
DD	0.1	0.4	0.0	0.2	0.0
TT	51.4	37.9	41.7	30.4	37.2
CT	25.2	24.4	24.6	25.6	26.7
else	6.2	14.9	18.7	8.6	14.3
$CT + TT$ ^a	76.6	62.3	66.3	55.9	63.9

^aThe sum of the weights of TT and CT configurations.

^b The ten-state-averaged (SA10-)CASSCF with *ma*-def2-SVP-SDD basis were employed in these calculations.

	CASSCF(4,4) ^b	$CASSCF(8,8)^b$	CASECF(12,12) ^b	QD-NEVPT2 $(4,4)^b$	QD-NEVPT2 $(8,8)^b$
GG	95.9	73.5	83.5	87.5	86.6
GD	0.6	0.2	1.2	2.0	1.1
DG	1.1	1.3	0.8	2.8	0.7
DD	0.0	0.0	0.0	0.0	0.0
TT	0.0	0.0	0.0	0.0	0.0
CT	0.0	0.0	0.0	0.0	0.0
else	2.4	24.8	14.4	7.6	11.5
$CT + TT$ ^a	0.0	0.0	0.0	0.0	0.0

Table S3. Weight [%] of each electron configuration in the ground state for **Nc(2)** calculated with different sizes of the active space in the SA-CASSCF and QD-NEVPT2.

^aThe sum of the weights of TT and CT configurations.

^b The ten-state-averaged (SA10-)CASSCF with *ma*-def2-SVP-SDD basis were employed in these calculations.

S7. Cartesian coordinates for optimized geometries

Optimized geometry of cyclobutadiene with the *D***4h symmetry in the singlet state at the CASSCF(4e, 4o) NEVPT2/cc-pVTZ level.**

Optimized geometry of cyclobutadiene with the *D***2h symmetry in the singlet state at the CASSCF(4e, 4o) NEVPT2/cc-pVTZ level.**

Cartesian coordinates for Nc(1).

Cartesian coordinates for Nc(2).

REFERENCES

- 1. Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. Introduction of n-electron valence states for multireference perturbation theory. *Journal of Chemical Physics* **2001**, 114, 10252-10264.
- 2. Angeli, C.; Borini, S.; Cestari, M.; Cimiraglia, R. A Quasidegenerate Formulation of the Second Order N Electron Valence State Perturbation Theory Approach. *J. Chem. Phys*. **2004**, 121, 4043-4049.
- 3. Lang, L; Sivalingam, K.; Neese F., *J. Chem. Phys.* **2020**, *152*, 014109.
- 4. Pipek, J.; Mezey, P. G. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys*. **1989**, 90, 4916- 4926.
- 5. Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. N-Electron Valence State Perturbation Theory: A Fast Implementation of the Strongly Contracted Variant. *Chem. Phys. Lett*. **2001**, 350, 297–305.
- 6. Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. N-Electron Valence State Perturbation Theory: A Spinless Formulation and an Efficient Implementation of the Strongly Contracted and of the Partially Contracted Variants. *J. Chem. Phys*. **2002**, 117, 9138- 915.
- 7. Neese, F. Software Update: The ORCA Program System, Version 4.0. *WIREs Comput. Mol. Sci*. **2018**, 8, e1327.
- 8. Monino, E.; Boggio-Pasqua, M.; Scemama, A.; Jacquemin, D.; Loos, P.-F. Reference Energies for Cyclobutadiene: Automerization and Excited States. J. Phys. Chem. A 2022, 126, 4664−4679.