The Journal of Chemical Physics

Quantitative characterization of exciton from *GW*+Bethe-Salpeter calculation

Cite as: J. Chem. Phys. **146**, 044303 (2017); https://doi.org/10.1063/1.4974320
Submitted: 26 October 2016 . Accepted: 06 January 2017 . Published Online: 23 January 2017

Daichi Hirose, Yoshifumi Noguchi, and Osamu Sugino

AIP Publishing

Quantitative characterization of exciton from *GW*+Bethe-Salpeter calculation

Daichi Hirose,[a] Yoshifumi Noguchi, and Osamu Sugino
Institute for Solid State Physics, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8581, Japan

(Received 26 October 2016; accepted 6 January 2017; published online 23 January 2017)

We propose a method of classifying excitons into local-, Rydberg-, or charge transfer-type as a step toward enabling a data-driven material design of organic solar cells. The classification method is based on the first-principles many-body theory and improves over the conventional method based on state-by-state visualization of the one-electron wave functions. In our method, the exciton wave function is calculated within the level of the *GW*+Bethe-Salpeter equation, which is used to obtain two dimensionless parameters for the automatic classification. We construct criteria for exciton classification from experiences with a model molecule, dipeptide. Then we check the validity of our method using a model β-dipeptide which has a geometry and an excitation spectrum similar to the model dipeptide. In addition, we test the effectiveness of the method using porphyrin molecules, or P_1TA and P_2TA, for which the conventional method is hampered by the strong state hybridization associated with excitation. We find that our method works successfully for P_1TA, but the analysis of P_2TA is hindered by its centrosymmetry. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4974320]

I. INTRODUCTION

Recently, organic solar cells (OSCs) have drawn great interest for their low-cost, flexibility, and high efficiency; current devices have reached 13% for the power conversion efficiency.[1] The key to achieve higher efficiency is probably in a deeper understanding of the processes occurring in OSCs, which start from absorption of photon, followed by formation, migration, and dissociation of excitons, as well as transport and collection of charge. The performance of the solar cell is deteriorated in the exciton migration process when an electron and hole (*e*-*h*) pair recombines before reaching the donor/acceptor interface. Deterioration also occurs in the exciton dissociation process when the binding energy of the pair is large enough to hinder efficient dissociation into free carriers at the interfaces. In this respect, excitons should have a long lifetime and a small binding energy. It is therefore advantageous for the solar cell to utilize a charge-transfer (CT) exciton where the electron and the hole are spatially separated from each other and are weakly bound, but conditions for generating the CT exciton are not well-known.

Theoretically, methods for predicting the excitation spectra have been developed based on density functional theory (DFT),[2–4] wave function-based methods,[5–8] and many-body perturbation methods,[9–18] but relation between the exciton type and the material property has not been established. Moreover, compared to methods to calculate the excited states, methods to analyze the calculated results have been unsophisticated yet; indeed, the analysis has been mostly done on the basis of one-electron theory and the type of excitons has been classified into local-, Rydberg-, or CT-type by visualizing one-electron wave function corresponding to a hole state or an electron state although calculation of the excited state is based on the sophisticated many-body theory. In this context, it is important to advance the analysis method to unambiguously classify the excitons, so that the analysis can be systematically done for a number of organic materials for the purpose of designing a material using a data-driven approach.

Recently, however, advance in the analysis method was made within the scheme of time-dependent DFT (TDDFT).[19–22] The method makes use of the one-electron transition density matrix, which characterizes an excited state in terms of pairs of occupied state and empty state. Taking the density matrix as a weight factor, the average is taken for the overlap of the wave functions[23,24] or for *e*-*h* position operators[25,26] to obtain the *e*-*h* overlapping strength Λ, exciton size, the electron (or hole) delocalization, and the *e*-*h* separation distance. These parameters were calculated for some benchmark molecules[23–26] to show usefulness in characterizing the excitons.

Each of those parameters is, however, not sufficient to classify all the excitons when used individually. To proceed the above-mentioned data-driven research, we consider that it is important to use those parameters in a combined way. In this context, the aim of this paper is to examine the parameters to find a way to classify the excitons.

In doing so, we consider that it is important to calculate the exciton using a reliable computational scheme. For this purpose, TDDFT is not the most appropriate scheme considering that the accuracy in treating the CT and Rydberg excitations is unsatisfactory unless using extremely advanced functionals, and in this regard, we use the two-body Green's function method developed on the basis of the many-body perturbation theory. Our Green's function method is based on the *GW* approximation[9–12] and the all-electron mixed basis, and the developed *GW*+Bethe-Salpeter equation (BSE)[13–18] code was shown to provide sufficiently an accurate description of the dynamical and non-local Coulomb interaction and

[a]hirose@issp.u-tokyo.ac.jp

thus more realistic treatment of the excitonic effect than does TDDFT.[21,27–31]

In this demonstration study, we first apply the exciton analysis method to a model dipeptide[32,33] for which an independent particle approximation (or one-particle picture) is known to work well. By comparing the calculated parameters and the existing visualization method that is based on the wave functions, we investigate how the excitons are classified using the parameters. In addition to the parameters introduced in the previous study, we newly define and examine the exciton binding energy, which has so far been only roughly estimated from a diagonal matrix element of Coulomb attractive interaction. Using the result, we derive two dimensionless parameters appropriate for the characterization of excitons. Then we check the validity of our method using a model β-dipeptide which has a geometry and an excitation spectrum similar to the model dipeptide. The applicability of the method is then demonstrated by applying it to porphyrin molecules (P_1TA and P_2TA),[34] which have been successfully synthesized and have attracted attention because of their light-harvesting function mimicking photosynthesis.[34–40] These molecules are chosen for the demonstration also because the densely populated transition states existing in these systems cause strong state hybridization, which hinders analysis based on one-electron theory. Excitons of P_1TA are successfully classified using our analysis method, but the P_2TA found a typical example where the analysis is hampered by the centrosymmetry. This counterexample prompts further development of the analysis method for the future application.

II. METHODOLOGY

A. *GW*+BSE method

Here, we briefly review the *GW*+BSE method. Our scheme consists of three procedures: (1) DFT calculation within the local density approximation (LDA) to obtain the Kohn-Sham (KS) eigenvalue E^{LDA}, KS eigenvector ψ, and the one-particle Green's function G_0, (2) one-shot *GW* calculation to obtain the quasiparticle (QP) energy E^{GW} and the dynamically screened Coulomb interaction W_0 within the random phase approximation (RPA), where the self-energy operator Σ is approximated as a product of G_0 and W_0 as $\Sigma^{GW} = iG_0W_0$, and (3) BSE calculation within the one-shot *GW* approximation to obtain the excitation energy Ω and the *e-h* amplitude A (see below for the definition), which corresponds to the one-electron transition density matrix of TDDFT. Procedures (1) and (2) are a standard one-shot *GW* calculation employing the Hybertsen-Louie type generalized plasmon pole (GPP) model[10,41] for the dynamical screening effect. The QP energy of the nth state is given as

$$E_n^{GW} = E_n^{\mathrm{LDA}} + Z_n \langle \psi_n | \Sigma^{GW}(E_n^{\mathrm{LDA}}) - \mu_{\mathrm{xc}}^{\mathrm{LDA}} | \psi_n \rangle, \quad (1)$$

$$Z_n = \left[1 - \frac{\partial \Sigma^{GW}(E)}{\partial E} \right]^{-1}_{E=E_n^{\mathrm{LDA}}}, \quad (2)$$

where $\mu_{\mathrm{xc}}^{\mathrm{LDA}}$ is the LDA exchange-correlation potential and Z_n is the renormalization factor. In order to explicitly consider the excitonic effect and to go beyond the independent particle approximation, we solve an eigenvalue problem for BSE,

$$H^{\mathrm{BSE}} A = \Omega A, \quad (3)$$

where H^{BSE} is defined as follows:

$$H^{\mathrm{BSE}}_{eh,e'h'} = (E_e^{GW} - E_h^{GW})\delta_{ee'}\delta_{hh'} + K^{\mathrm{d}}_{eh,e'h'} + 2K^{\mathrm{x}}_{eh,e'h'}, \quad (4)$$

where K^{d} and K^{x} are the direct and the exchange interactions, respectively. We again employ the GPP model to consider the dynamical excitonic effect. Other details are given in Ref. 17. The exciton wavefunction χ_{exc} is given by the *e-h* amplitude A as follows:

$$\chi_{\mathrm{exc}}(\boldsymbol{r}_e, \boldsymbol{r}_h) = \sum_{eh} A_{eh} \psi_e^*(\boldsymbol{r}_e) \psi_h(\boldsymbol{r}_h). \quad (5)$$

B. Exciton analysis I: Spatial overlap

It is convenient to quantitatively measure the spatial overlap between an electron and a hole. Following Refs. 23 and 24, we introduce the spatial overlap of the KS eigenvectors as

$$Q_{eh} = \langle |\psi_e| | |\psi_h| \rangle = \int |\psi_e(\boldsymbol{r})| |\psi_h(\boldsymbol{r})| d\boldsymbol{r}. \quad (6)$$

Then we define the quantity Λ as follows:

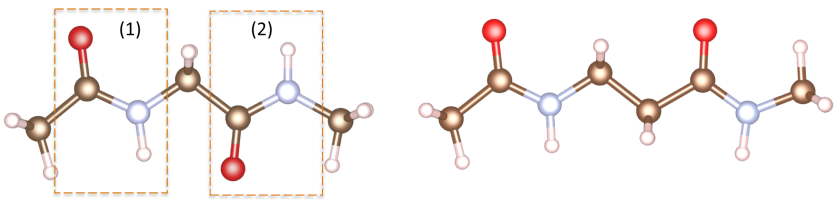
$$\Lambda = \frac{\sum_{eh} |A_{eh}|^2 Q_{eh}}{\sum_{eh} |A_{eh}|^2}, \quad (7)$$

which takes the value $0 \le \Lambda \le 1$. A small value of Λ corresponds to a long-range excitation; a large value corresponds to a short-range excitation.

C. Exciton analysis II: Expectation values

Properties of excitons are characterized by statistical quantities related to the electron and the hole position operators.[25,26] A property of an exciton can be directly calculated as an expectation value,

$$\langle O \rangle = \frac{\langle \chi_{\mathrm{exc}} | O | \chi_{\mathrm{exc}} \rangle}{\langle \chi_{\mathrm{exc}} | \chi_{\mathrm{exc}} \rangle} \quad (8)$$



(a)dipeptide. (b)β-dipeptide

FIG. 1. The atomic geometries optimized for a model dipeptide and β-dipeptide by B3LYP/cc-pVTZ.

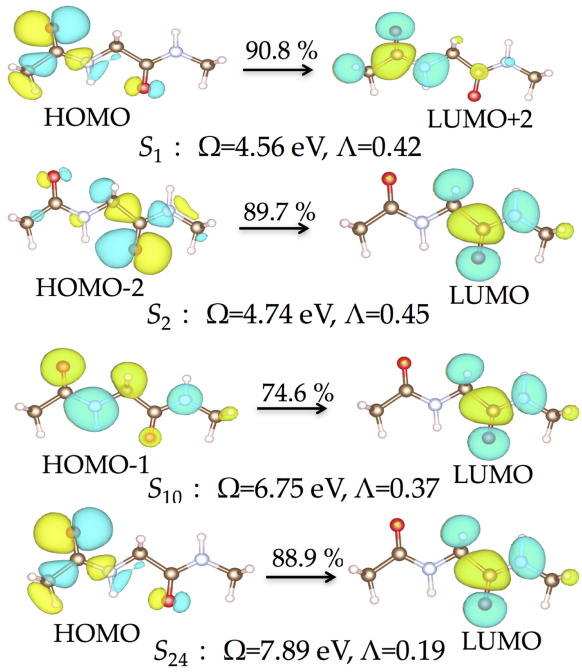


FIG. 2. Exciton profiles of a model dipeptide. S_1, S_2, S_{10}, and S_{24} correspond to W1, W2, CTa, and CTb in Ref. 33, respectively. The most dominant $|A_{eh}|^2$ is shown with %.

for any operator O. Following Refs. 25 and 26, we define the statistical quantities as follows: the e-h separation distance $d_{h\to e}$,

$$d_{h\to e} = |\langle \boldsymbol{r}_e - \boldsymbol{r}_h \rangle| = |\langle \boldsymbol{r}_e \rangle - \langle \boldsymbol{r}_h \rangle|, \tag{9}$$

the exciton size d_{exc},

$$d_{\mathrm{exc}} = \sqrt{\langle |\boldsymbol{r}_e - \boldsymbol{r}_h|^2 \rangle}, \tag{10}$$

the electron and the hole delocalization σ_e and σ_h,

$$\sigma_e = \sqrt{\langle \boldsymbol{r}_e^2 \rangle - \langle \boldsymbol{r}_e \rangle^2}, \tag{11}$$

$$\sigma_h = \sqrt{\langle \boldsymbol{r}_h^2 \rangle - \langle \boldsymbol{r}_h \rangle^2}. \tag{12}$$

Along this line, we newly propose the exciton binding energy E_{b},

$$E_{\mathrm{b}} = -\langle K^{\mathrm{d}} + 2K^{\mathrm{x}} \rangle. \tag{13}$$

Note that one-electron indices (e,h) disappear in the above equation and E_{b} is completely different from the matrix element of Coulomb interaction.

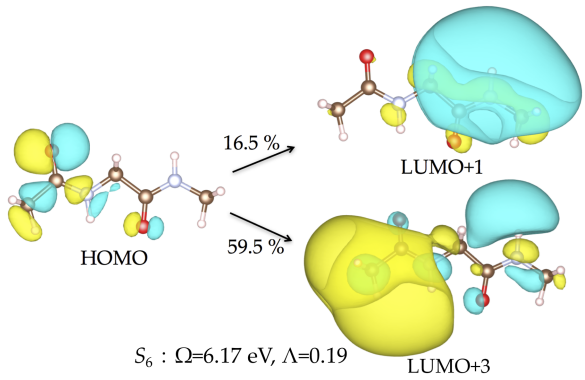


FIG. 3. Exciton profiles of S_6 of a model dipeptide.

D. Calculation setup

Our GW+BSE calculations are executed on our all-electron mixed basis program[42–46] in which the LDA eigenvector is expanded as a linear combination of plane waves (PWs) and numerical atomic orbitals (AOs). It makes possible to accurately describe localized states in the core region and free electron states above the vacuum level. In the GW+BSE calculation, we use an fcc supercell and the Coulomb cut-off technique[47–49] to eliminate spurious interactions with the image cells. The geometry optimization and the TDDFT calculation are, on the contrary, executed using the GAUSSIAN 09 program package: The atomic geometries of the ground state are optimized with use of B3LYP/cc-pVTZ, and the TDDFT calculations are done with use of LDA, B3LYP, and CAM-B3LYP/cc-pVTZ.

III. RESULTS AND DISCUSSION

A. Intuitive understanding of parameters

To get an insight into our exciton analysis method, we apply it to a simple molecular system, or a model dipeptide shown in Fig. 1(a), and check how the excitons are characterized by the parameters introduced above. This model dipeptide is expected as one of the simplest molecules to show a CT excitation at a low photoenergy because the wave functions at the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) localize at separate areas (1) and (2) shown in Fig. 1(a), respectively. Because of this property, this molecule has been studied as a benchmark molecule, and the four excitons labeled as W1, W2, CTa,

TABLE I. Parameters characterizing excitons of a model dipeptide. Shown are the excitation energy (Ω), parameters of exciton (Λ, d_{exc}, σ_h, σ_e, $d_{h\to e}$, and E_{b}), and the classified type.

	Ω (eV)	Λ	d_{exc} (Å)	σ_h (Å)	σ_e (Å)	$d_{h\to e}$ (Å)	$d_{h\to e}/d_{\mathrm{exc}}$	E_{b} (eV)	Type
S_1	4.56	0.42	2.35	1.26	1.80	0.84	0.36	7.52	L
S_2	4.74	0.45	2.37	1.30	1.72	0.98	0.42	7.55	L
S_6	6.17	0.19	3.76	1.55	3.01	1.74	0.46	4.18	R
S_{10}	6.75	0.37	3.53	1.93	2.04	2.11	0.60	4.17	L
S_{24}	7.89	0.19	4.56	1.39	1.73	3.98	0.87	3.95	CT

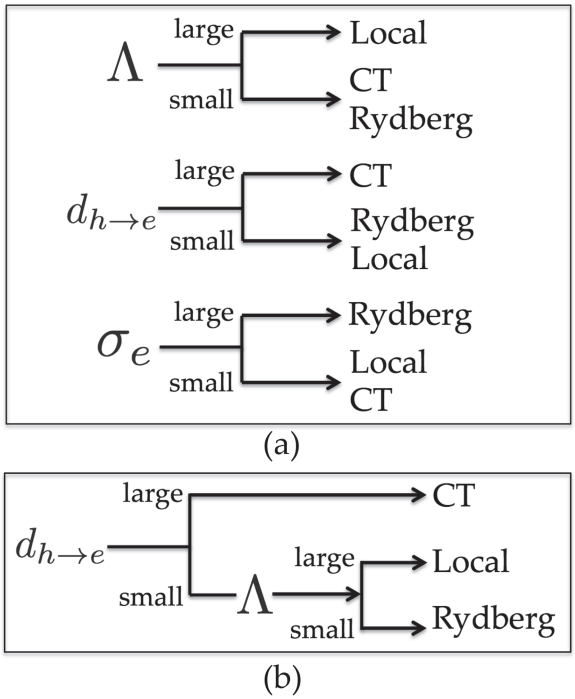


FIG. 4. (a) Classification of excitons by respective parameters. (b) Proposed flowchart of excitons using two parameters.

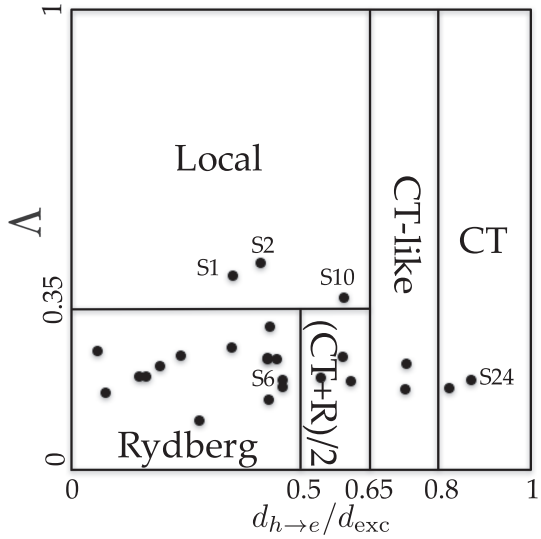


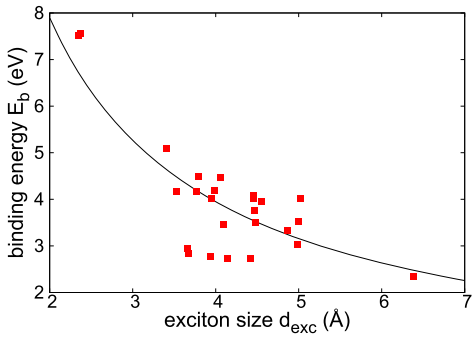
FIG. 5. The Λ-$d_{h\to e}/d_{\mathrm{exc}}$ plot of a model dipeptide for S_n (n = 1-25); S_1, S_2, S_6, S_{10}, and S_{24} are explicitly labeled. All the excitons are classified into local, Rydberg, and CT types.

and CTb have been analyzed intensively. According to the previous studies (Refs. 32 and 33), the first two excitons (W1 and W2) are categorized as the local excitons where the overlap between the hole and electron wave functions is large, and the last two excitons (CTa and CTb) are categorized as the CT excitons where the overlap is almost zero. In our BSE calculation, these excitons correspond to S_1, S_2, S_{10}, and S_{24}, respectively. In addition, we also discuss the S_6 exciton as an example of the Rydberg type. In this section, we show that these four excitons are classified by the parameters in accordance with the previous study while S_6 is not.

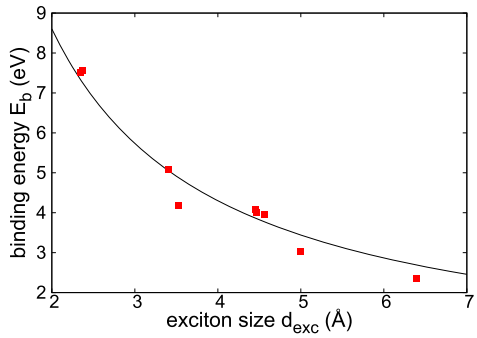
As shown in Fig. 2, S_1, S_2, S_{10}, and S_{24} have a single dominant transition, so that we can expect that the results obtained in our exciton analysis can be intuitively understood on the basis of the one-electron theory. We evaluate the value of Λ for S_1 and S_2 as 0.4–0.5 and the value of Λ for S_{24} as 0.19, suggesting that S_1 and S_2 are the local excitons while S_{24} is the CT exciton. The value for S_{10} is 0.37 and is only slightly smaller than the value for S_1 and S_2, suggesting that this exciton to be classified as the local exciton although S_{10} was classified as the CT exciton in the previous study.[32,33] This contradicting result requires a careful check on the wave function. The wave functions for the electron state are commonly LUMO for S_2, S_{10}, and S_{24} while those for the hole state are HOMO-2, HOMO-1, and HOMO, respectively. The electron and hole wave functions have a large overlap for S_2, but have a negligibly small overlap for S_{24}. The overlap of S_{10} is smaller than that of S_2 but is obviously much larger than that of S_{24}. Therefore, we conclude that S_{10} should be classified as the local exciton. Note that the supplementary material in Ref. 23 supports our results in which the Λ values of S_1, S_2, and S_{10} are large and that of S_{24} is small.

For S_6, however, we find that Λ is not necessarily a good descriptor. S_6 is involved with two major transitions where both occur from the HOMO to Rydberg levels (see Fig. 3). The value of Λ is 0.19 and is the same as that of the CT exciton (=S_{24}). This indicates that the type of excitons cannot be classified by Λ alone.

In this context, we investigate how the value of other parameters can be combined to make the most appropriate descriptor set. Table I shows the calculated parameters: Λ, the exciton size d_{exc}, the electron (or hole) delocalization



(a)S_1-S_{25}.



(b)R and (CT+R)/2 are removed.

FIG. 6. The relation between the exciton size d_{exc} and the binding energy E_{b} for a model dipeptide. A solid line is obtained by fitting the data to $E_{\mathrm{b}} = a/d_{\mathrm{exc}}$. (a) S_1-S_{25}, (b) R and (CT+R)/2 are removed from S_1-S_{25}.

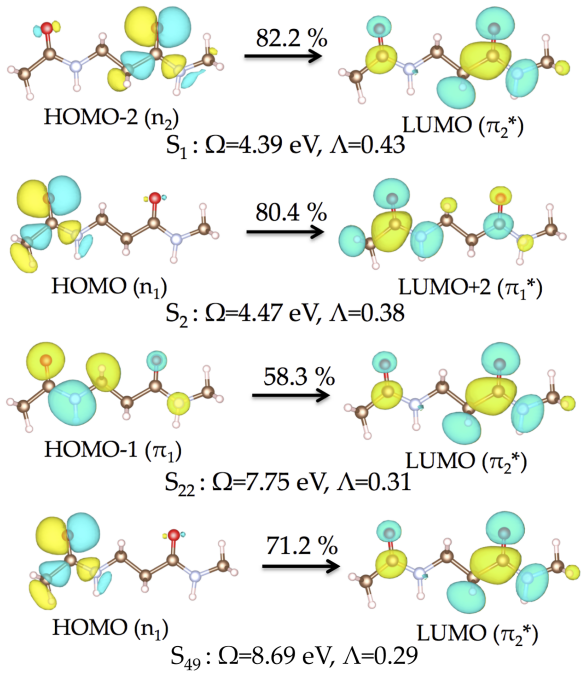


FIG. 7. Exciton profiles of a model β-dipeptide. The most dominant $|A_{eh}|^2$ is shown with %.

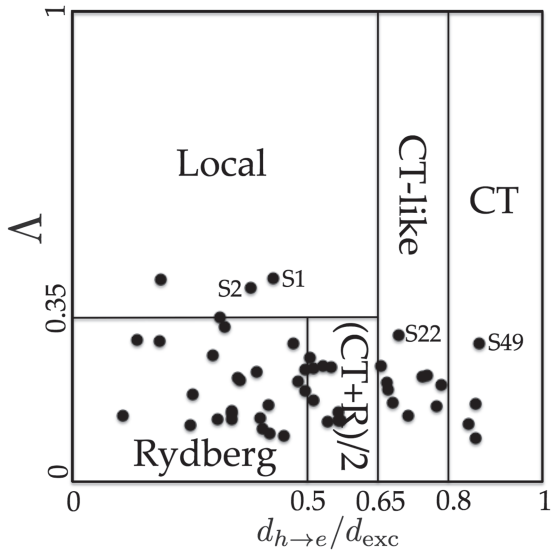


FIG. 8. The Λ-$d_{h\to e}/d_{\rm exc}$ plot of a model β-dipeptide for S_n (n = 1-50); S_1, S_2, S_{22}, and S_{49} are explicitly labeled. All the excitons are classified into local, Rydberg, and CT types.

σ_e (σ_h), the e-h separation distance $d_{h\to e}$, and the exciton binding energy E_b as given in Eqs. (7)–(13). S_1 and S_2 have common features that σ_e and σ_h are about 1.3-1.8 Å and $d_{h\to e}$ is about 0.8-0.9 Å, indicating that the localization radius is larger than the separation distance and thus should be classified as the local exciton, in consistent with the spatial distribution of the wave functions shown in Fig. 2. It is also found that $E_{\rm b}$ is quite large (>7.5 eV) in consistent with the local nature of the excitation. These results show that S_1 and S_2 are well characterized by the parameters σ_e, σ_h, and $d_{h\to e}$.

For S_6, not only σ_e (3.0 Å) but also $d_{\rm exc}$ (3.8 Å) are large compared with $d_{h\to e}$ (1.7 Å), indicating that the delocalization length and the exciton size are much larger than the e-h separation distance in consistent with the Rydberg nature of the excitation. For S_{10}, the e-h separation distance $d_{h\to e}$ is slightly larger than that of S_1 and S_2, and the electron delocalization length is almost the same as that of S_1 and S_2. This result is a characteristic of a local exciton. For S_{24}, the e-h separation distance $d_{h\to e}$ (=4.0 Å) is large while the overlap Λ (<0.2) is small, faithfully representing a CT type of this excitation. This result is also consistent with the large exciton size $d_{\rm exc}$ (=4.6 Å) and the small exciton binding energy $E_{\rm b}$ (=4.0 eV).

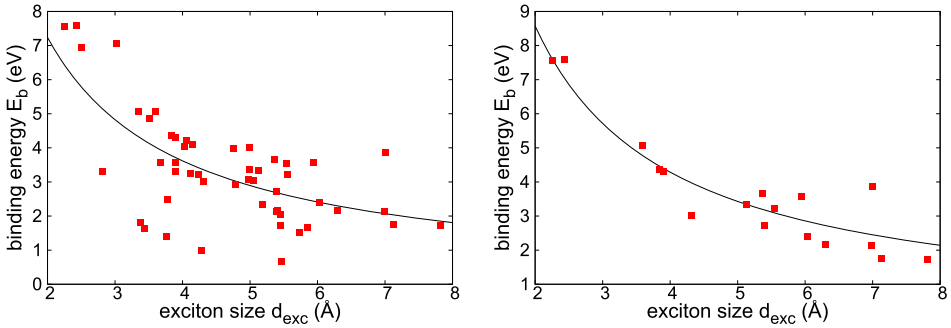
All the results listed in Table I are thus reasonably consistent with the intuitive classification of the exciton and therefore should show a good performance of the parameter-based classification. From the above results, we have learned that we need multiple parameters for the classification; one of the possible choices would be to use Λ and the e-h separation distance $d_{h\to e}$ to distinguish CT from others, and then to use the electron delocalization for the distinction of the Rydberg. Before investigating for the most appropriate parameter set, here we discuss the correlation of the exciton binding energy with the exciton size for a deeper insight into the parameters.

B. Classification method for the type of excitons

We now construct a method of classifying excitons by generalizing the above discussion. First, we will use the three observations obtained in Sec. III A as summarized in Fig. 4(a): (1) Those exciton having large Λ are local, (2) those with large $d_{h\to e}$ are of the CT type, and (3) those with large σ_e are of the Rydberg type. Second, single parameter is not sufficient to classify all the excitons; so we will use two parameters, such as $d_{h\to e}$ and Λ out of the three parameters introduced above (see Fig. 4(b)). To construct the method in this way, we consider that it is preferable to set the border line for the classification by using a dimensionless parameter that

TABLE II. Parameters characterizing excitons of the model β-dipeptide.

	Ω (eV)	Λ	$d_{\rm exc}$ (Å)	σ_h (Å)	σ_e (Å)	$d_{h\to e}$ (Å)	$d_{h\to e}/d_{\rm exc}$	$E_{\rm b}$ (eV)	Type
$S_1(n_2\to\pi_2^*)$	4.39	0.43	2.44	1.32	1.8	1.04	0.43	7.60	L
$S_2(n_1\to\pi_1^*)$	4.47	0.41	2.25	1.31	1.67	0.86	0.38	7.56	L
$S_{22}(\pi_1\to\pi_2^*)$	7.75	0.31	4.31	2.01	2.26	3.00	0.69	3.01	CT-like
$S_{49}(n_1\to\pi_2^*)$	8.69	0.29	5.39	2.00	1.81	4.67	0.87	2.72	CT



(a)S_1-S_{50}. (b)R and (CT+R)/2 are removed.

FIG. 9. The relation between the exciton size d_{exc} and the binding energy E_b for a model β-dipeptide. A solid line is obtained by fitting the data to $E_b = a/d_{exc}$.

is independent of the molecular size. Although Λ is dimensionless, $d_{h\to e}$ and σ_e have a dimension of length. With this in mind, we propose to use the ratio $d_{h\to e}/d_{exc}$ or σ_e/d_{exc} together with Λ. It is more convenient to use the former ($d_{h\to e}/d_{exc}$) than the latter (σ_e/d_{exc}) because the former is normalized to 1, namely, distributed in the range between 0 and 1.

On the basis of the above discussion, we propose the classification method as shown in Fig. 5 using two dimensionless parameters Λ and $d_{h\to e}/d_{exc}$, where the border lines are determined to adjust to the experience with the model dipeptide. We note that there is a region in which the excitons cannot be simply classified to CT or Rydberg but are a mix of them. We add such a region in Fig. 5 and is denoted as (CT+R)/2. We also add CT-like which is interpreted as a mix of local and CT in view of $d_{h\to e}/d_{exc}$.[50,51] Λ and $d_{h\to e}/d_{exc}$ are not independent so that if one is larger, the other should be smaller. Thus, there should be no excitation corresponding to the top right part of the Λ-$d_{h\to e}/d_{exc}$ plot, but we defined for simplicity as in Fig. 5. Note that although our classification is based on the *GW*+BSE calculation, we believe that it is applicable to TDDFT calculation as well.

In Fig. 6(a), the relation between the exciton size d_{exc} and the binding energy E_b is shown for S_1-S_{25}. Considering that the excitonic effect in the *GW*+BSE method is taken into account through the direct dynamical screened Coulomb term (K^d) and the exchange bare Coulomb term (K^x) in Eq. (4), the exciton binding energy is expected to be inversely proportional to the distance between electron and hole. Such inversely proportional correlation can be roughly seen in Fig. 6(a), and its root mean squared error (RMSE) to the fitting curve ($E_b = a/d_{exc}$) is 0.76. We consider that the origin of large variation is due to the Rydberg states which variate the screening effect. Fig. 6(b) shows the plot in which the R and (CT+R)/2 excitons are removed from S_1-S_{25}. In Fig. 6(b), the inversely proportional correlation can be reasonably well seen, and its RMSE is reduced to be 0.34 from 0.76.

C. β-dipeptide

We check the validity of our method using the model β-dipeptide shown in Fig. 1(b), which has a longer chain compared to the model dipeptide but a similar spatial distribution of the exciton. In our calculation, S_1, S_2, S_{22}, and S_{49} shown in Fig. 7 correspond to $n_2 \to \pi_2^*, n_1 \to \pi_1^*, \pi_1 \to \pi_2^*$, and $n_1 \to \pi_2^*$ in Refs. 23 and 32, respectively. In previous studies,[23,32] the first two excitons are categorized as the local excitons, while the last two excitons are categorized as the CT excitons. In comparison with Fig. 2 or the model dipeptide, similar characters of the excitations can be seen except that S_1 and S_2 are interchanged.

Table II shows calculated parameters and Fig. 8 shows the Λ-$d_{h\to e}/d_{exc}$ plot. In our classification method, S_1 and S_2 are classified into the local excitons, while S_{22} and S_{49} are the CT-like and CT exciton, respectively. These classifications are consistent with intuitive ones, while classification of S_{22} and S_{49} needs to be checked with care. The wave functions for the electron state are commonly LUMO for S_{22} and S_{49}, while those for the hole state are HOMO-1 and HOMO, respectively. The HOMO is localized at left side of the molecule, but the HOMO-1 is more delocalized. From this, we can intuitively understand the classification that S_{22} and S_{49} correspond to CT-like and CT, respectively. Our

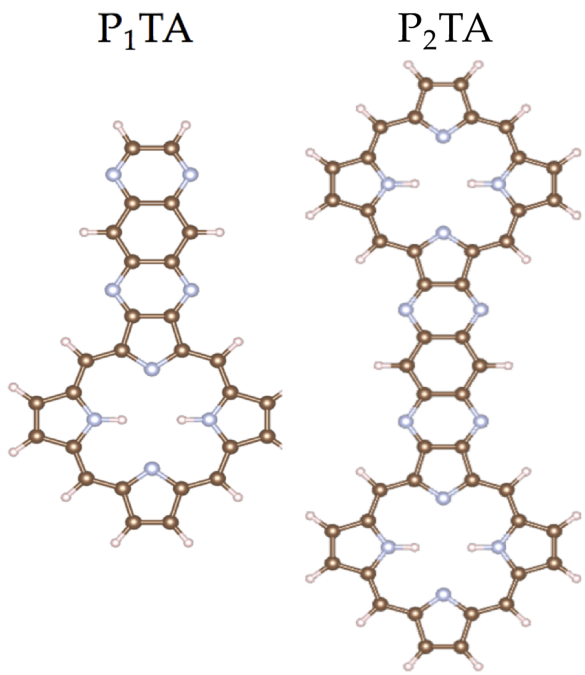


FIG. 10. The atomic geometries optimized for P_1TA and P_2TA by B3LYP/cc-pVTZ.

TABLE III. Parameters characterizing excitons of P_1TA.

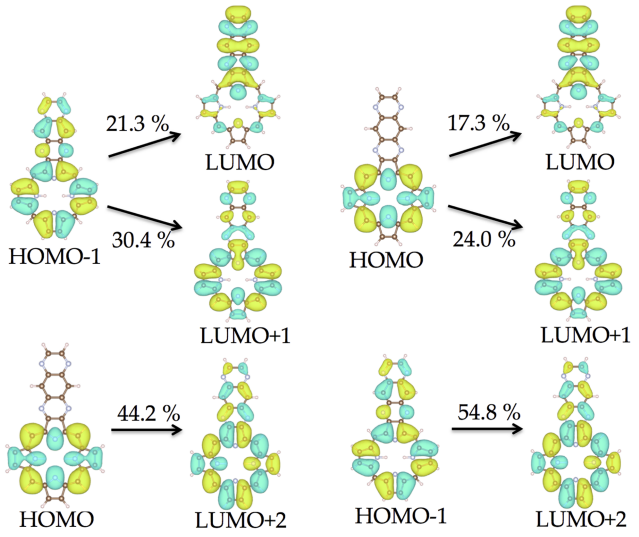
	Ω (eV)	Λ	d_{exc} (Å)	σ_h (Å)	σ_e (Å)	$d_{h\to e}$ (Å)	$d_{h\to e}/d_{exc}$	E_b (eV)	Type
S_1	2.06	0.67	5.23	3.77	4.07	0.23	0.04	3.20	L
S_2	2.32	0.68	5.43	3.84	4.11	0.19	0.04	2.93	L
S_3	2.42	0.61	7.06	4.56	3.95	4.31	0.61	2.58	L
S_4	2.67	0.40	3.35	2.46	2.94	0.48	0.14	5.38	L
S_5	2.72	0.44	7.64	3.37	3.75	5.73	0.75	2.59	CT-like
S_6	2.72	0.45	3.57	2.58	3.13	0.45	0.13	4.84	L
S_7	3.16	0.70	5.25	4.30	3.63	1.47	0.28	2.85	L
S_8	3.22	0.42	5.71	4.17	4.91	1.08	0.19	3.03	L
S_9	3.24	0.65	4.95	3.41	4.19	0.47	0.10	3.48	L
S_{10}	3.36	0.54	5.21	3.32	3.99	1.66	0.32	3.23	L
S_{11}	3.40	0.69	5.80	5.84	5.10	0.18	0.03	2.66	L
S_{12}	3.54	0.28	6.41	4.18	5.37	0.06	0.01	2.77	R
S_{13}	3.64	0.30	6.05	3.31	5.39	0.07	0.01	2.88	R
S_{14}	3.74	0.60	5.64	5.09	4.85	0.80	0.14	2.33	L
S_{15}	3.78	0.56	6.56	4.53	5.22	1.30	0.20	2.63	L
S_{16}	3.80	0.25	5.71	3.31	4.98	1.96	0.34	3.37	R
S_{17}	3.82	0.22	6.41	4.13	5.77	1.42	0.22	3.48	R
S_{18}	3.82	0.51	5.92	5.15	4.46	0.52	0.09	3.12	L
S_{19}	3.92	0.13	6.82	3.33	6.37	0.78	0.11	2.39	R
S_{20}	4.00	0.54	6.94	4.64	4.76	2.29	0.33	2.78	L

classifications are also consistent with previous studies,[23,32] when regarding CT-like as one of the CTs. From the above discussion, we can confirm the validity of our classification method.

Fig. 9(a) shows the relation between the exciton size d_{exc} and the binding energy E_b for S_1-S_{50}. The variation of this plot is very large; RMSE of the fitting amounts to be 1.11. We consider that this is because the state contains large amount of Rydberg excitons. Fig. 9(b) shows the plot in which the R and (CT+R)/2 excitons are removed from S_1-S_{50}. In comparison with Fig. 9(a), the variation is significantly reduced and the RMSE is 0.57. Therefore, after removing the R and (CT+R)/2 excitons, E_b approximately correlates with $1/d_{exc}$ as does in the model dipeptide.

D. P_1TA: Strong hybridization system

We now apply our method to two porphyrin molecules, which have attracted attention for the solar cell application. The first one is P_1TA (Fig. 10). Table III shows the calculated parameters of P_1TA for S_1-S_{20}. Note that, because of strong state hybridization associated with excitation as shown in Fig. 11, it is not appropriate to use the wave function visualization method for the analysis. All the excitons are shown in the Λ-$d_{h\to e}/d_{exc}$ plot as can be



(a)S_1: Ω = 2.06 eV, Λ = 0.67. (b)S_2: Ω = 2.32 eV, Λ = 0.68.

FIG. 11. Exciton profiles of P_1TA. Lowest two transition components are shown.

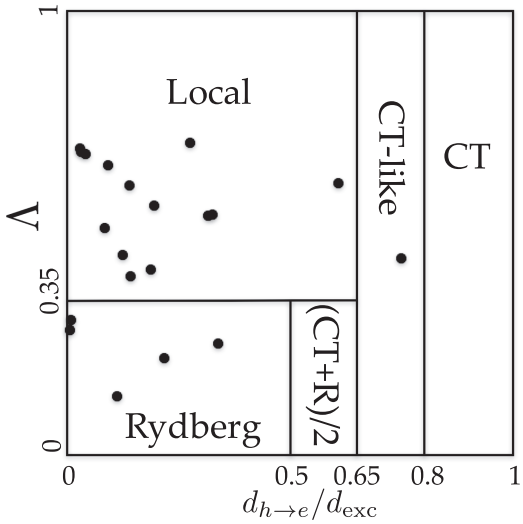
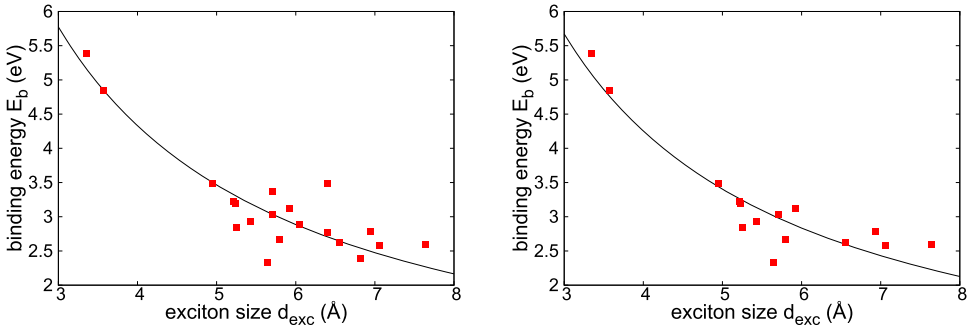


FIG. 12. The Λ-$d_{h\to e}/d_{exc}$ plot of P_1TA for S_n (n = 1-20).



(a)S_1-S_{20}. (b)R and (CT+R)/2 are removed.

FIG. 13. The relation between the exciton size d_{exc} and the binding energy E_{b} for P_1TA.

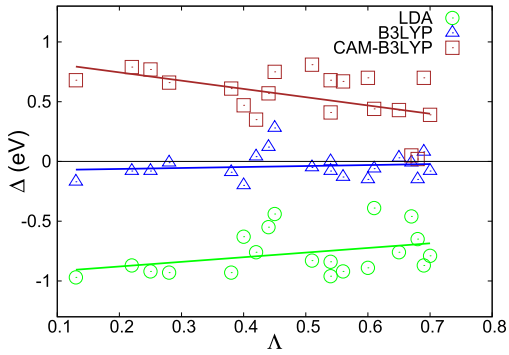


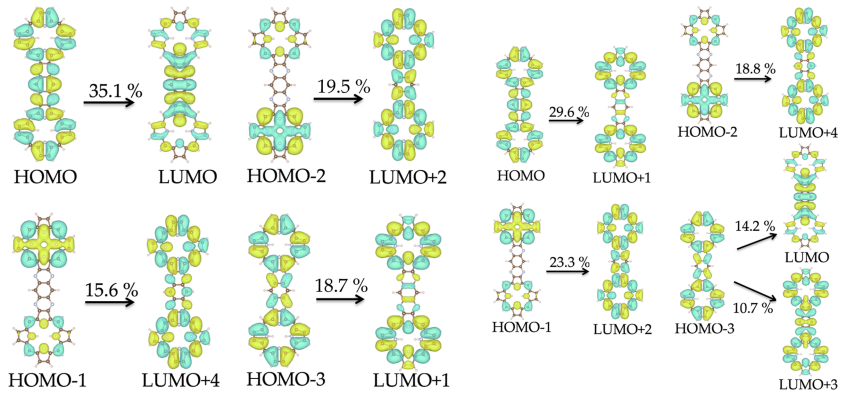
FIG. 14. Λ dependence of optical gaps of P_1TA. We plot the difference between optical gaps calculated by TDDFT and the *GW*+BSE method ($\Delta = \Omega_n^{\mathrm{TDDFT}} - \Omega_n^{\mathrm{BSE}}$). Circles: LDA values; triangles: B3LYP values; squares: CAM-B3LYP values.

seen in Fig. 12 and are accordingly classified as shown in Table III.

The first two excitons (S_1 and S_2) of P_1TA have approximately the same value for Λ (=0.67-0.68) and $d_{h\to e}/d_{\mathrm{exc}}$ (=0.04) are located in the region of a local exciton as shown in Fig. 12. While the two excitons are both classified as local, the wave function is slightly different from each other as can be found by looking at the transition components (Fig. 11); S_1 does not have a component of CT while 17.3% of S_2 is composed of a HOMO to LUMO transition of the CT type. S_2 is, however, more importantly composed of HOMO-1 to LUMO+2 (=54.8%) and HOMO to LUMO+1 (=24.0%), both of which are of the local type. This way, we can understand that it is not so easy to see how an exciton is more CT-like or not only from the wave function. Note that the Λ values are larger than that of a model dipeptide because of the π electron character of the porphyrin system.

Because of the large hybridization, it is not possible to justify our analysis method by comparing with the intuitive method, but we can check consistency between the exciton binding energy and the exciton size as follows. Fig. 13(a) shows the relation between the exciton size d_{exc} and the binding energy E_{b} for S_1-S_{20}, and Fig. 13(b) shows the plot from which the R and (CT+R)/2 excitons are removed. In this case, we found that E_{b} is approximately proportional to $1/d_{\mathrm{exc}}$ even without removal. Correspondingly, the RMSE is only slightly reduced from 0.31 to 0.28 by the removal.

Before investigating another porphyrin molecule, here we compare our *GW*+BSE analysis with the previous TDDFT analysis regarding the Λ values. The previous analysis[23,24] showed that the accuracy of the excitation energy correlates with the Λ value such that the systems with smaller Λ are more affected by inaccuracy of the exchange-correlation functional. This is due to the well-known fact that the CT and Rydberg transitions are difficult to treat within TDDFT. Such correlation between the accuracy of TDDFT and the value of Λ can be found also by plotting the difference in the



(a)S_1: Ω = 2.10 eV, Λ = 0.69. (b)S_2: Ω = 2.13 eV, Λ = 0.67.

FIG. 15. Exciton profiles of P_2TA. Lowest two transition components are shown.

TABLE IV. Parameters characterizing excitons of P_2TA.

	Ω (eV)	Λ	d_{exc} (Å)	E_b (eV)	Type
S_1	2.10	0.69	5.72	2.99	L
S_2	2.13	0.67	5.39	3.08	L
S_3	2.31	0.71	5.79	2.84	L
S_4	2.38	0.71	5.38	2.87	L
S_5	2.44	0.69	6.96	2.30	L
S_6	2.72	0.57	8.19	2.50	L
S_7	2.88	0.51	8.16	2.34	L
S_8	2.89	0.52	8.12	2.40	L
S_9	3.15	0.32	6.55	2.65	R
S_{10}	3.16	0.69	7.32	3.06	L
S_{11}	3.17	0.31	6.49	2.67	R
S_{12}	3.25	0.52	3.79	4.40	L
S_{13}	3.26	0.51	3.85	4.40	L
S_{14}	3.28	0.37	5.90	3.11	L
S_{15}	3.28	0.37	5.85	3.20	L
S_{16}	3.34	0.70	7.06	2.79	L
S_{17}	3.38	0.62	8.94	2.03	L
S_{18}	3.42	0.70	6.36	2.14	L
S_{19}	3.46	0.62	6.83	2.48	L
S_{20}	3.50	0.29	7.52	2.55	R

excitation energy between TDDFT and *GW*+BSE, against Λ for n = 1-20 as shown in Fig. 14. In addition to this, it is seen that the difference is smaller in magnitude for the bare B3LYP than the CAM-B3LYP where the long-range part of the Coulomb interaction is corrected using the attenuated method. It is surprising that the bare B3LYP provides a much smaller difference (<0.3 eV) from *GW*+BSE although CAM-B3LYP should be more accurate in treating the long-ranged Coulomb interaction.

E. P_2TA: Strong hybridization and centrosymmetric system

The second porphyrin molecule, P_2TA, also shows strong hybridization as can be seen in Fig. 15, and thus, the wave function visualization analysis is not effective. P_2TA has a centrosymmetric geometry and, because of this particular reason, parameters $\langle \boldsymbol{r}_e \rangle$ and $\langle \boldsymbol{r}_h \rangle$ vanish identically. This invalidates the parameter $d_{h\to e}$ and thus makes it impossible to distinguish CT from Rydberg using the parameters introduced to our classification method. This fact indicates incompleteness

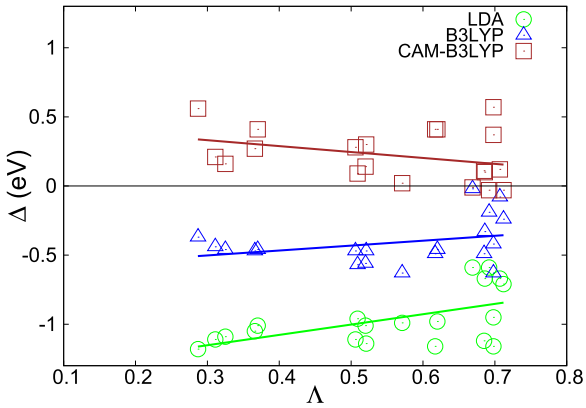
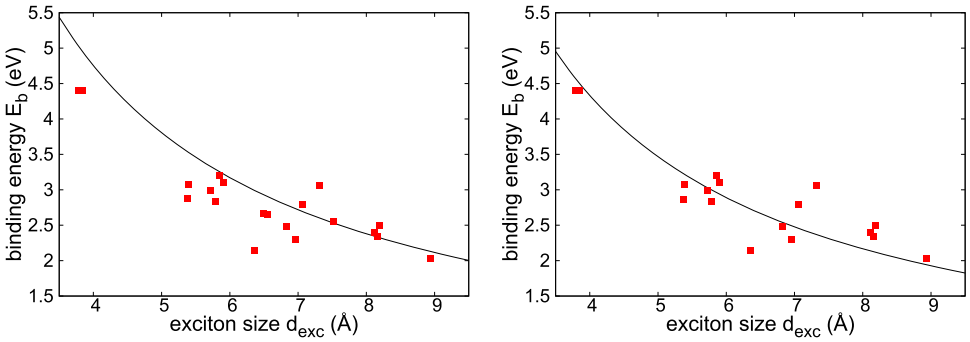


FIG. 16. Λ dependence of optical gaps of P_2TA. We plot the difference between optical gaps calculated by TDDFT and the *GW*+BSE method ($\Delta = \Omega_n^{TDDFT} - \Omega_n^{BSE}$). Circles: LDA values; triangles: B3LYP values; squares: CAM-B3LYP values.

of our analysis method in classifying all the excitons. Leaving construction of such complete analysis method as a subject of future study, we will continue the study on the excitons using the parameters Λ, d_{exc}, and E_b shown in Table IV.

From the value of Λ, we find that all the excitons are local except for S_9, S_{11}, and S_{20}, which are either CT, Rydberg, or (CT+R)/2. By looking at the wave functions, we can intuitively assign these three excitons as Rydberg excitons. It is noted, however, that the assignment would be ambiguous if the excitons were located near a border line.

Finally, we compare the *GW*+BSE method and TDDFT using the Λ values as did in Fig. 16: In Fig. 16, we show the plot for S_1-S_{20} of P_2TA. Similar to P_1TA, the LDA gap and the B3LYP gap are smaller and the CAM-B3LYP gap is larger than the *GW*+BSE gap, and the difference becomes larger in magnitude as Λ is reduced. Contrary to the case of P_2TA, however, CAM-B3LYP provides excitation energy most closely to the *GW*+BSE method. Fig. 17(a) shows the relation between the exciton size d_{exc} and the binding energy E_b for S_1-S_{20}, and Fig. 17(b) shows the plot from which the R and (CT+R)/2 excitons are removed. In this case, we found that E_b is approximately proportional to $1/d_{exc}$ even without the removal. Correspondingly, the RMSE is only slightly reduced from 0.39 to 0.32 removal.



(a)S_1-S_{20}. (b)R and (CT+R)/2 are removed.

FIG. 17. The relation between the exciton size d_{exc} and the binding energy E_b for P_2TA.

IV. CONCLUSION

We have constructed a quantitative method to classify excitons in terms of two dimensionless parameters (Λ and $d_{h\to e}/d_{\mathrm{exc}}$) determined within the many-body theory without relying on the one-electron theory. This improves over the conventional analysis method to visualize the wave functions state by state, which is cumbersome and ambiguous, and thus is an important step toward realization of an automatic classification required for a data-driven material design of organic solar cells. Our analysis method was constructed using the first-principles *GW*+BSE calculation of a model dipeptide and was applied to some porphyrin molecules, which have strong hybridization and are thus generally difficult to classify using the conventional scheme. Our method provided a reasonable classification in general, but failed to work for centrosymmetric molecules like P_2TA, indicating that the method should be improved for a realization of complete classification.

ACKNOWLEDGMENTS

The present calculations are performed with supercomputers from the Institute for Solid State Physics, University of Tokyo. Y.N. was supported by Grant-in-Aid for Young Scientist (B) (No. 23740288) and Grant-in-Aid for Scientific Research (C) (No. 26400383) from Japan Society for The Promotion of Science (JSPS).

This work was supported by the Strategic Programs for Innovative Research (SPIRE), MEXT, and the Computational Materials Science Initiative (CMSI), Japan.

[1] S. M. Ryno, M. K. Ravva, X. Chen, H. Li, and J.-L. Brédas, Adv. Energy Mater. **6**, 1601370 (2016).
[2] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[3] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[4] W. Kohn, Rev. Mod. Phys. **71**, 1253 (1999).
[5] J. F. Stanton and R. J. Bartlett, J. Chem. Phys. **98**, 7029 (1993).
[6] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).
[7] N. Nakatsuji, Chem. Phys. Lett. **59**, 362 (1978).
[8] H. Sekino and R. J. Bartlett, Int. J. Quantum Chem. **26**, 255 (1984).
[9] L. Hedin, Phys. Rev. **139**, A796 (1965).
[10] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[11] G. Strinati, H. J. Mattausch, and W. Hanke, Phys. Rev. Lett. **45**, 290 (1980).
[12] G. Strinati, H. J. Mattausch, and W. Hanke, Phys. Rev. B **25**, 2867 (1982).
[13] G. Strinati, Phys. Rev. B **29**, 5718 (1984).
[14] G. Onida, L. Reining, R. W. Godby, R. Del Sole, and W. Andreoni, Phys. Rev. Lett. **75**, 818 (1995).
[15] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).
[16] M. Rohlfing and S. G. Louie, Phys. Rev. Lett. **81**, 2312 (1998).
[17] M. Rohlfing and S. G. Louie, Phys. Rev. B **62**, 4927 (2000).
[18] G. Strinati, La Riv. del Nuovo Cimento **11**, 1 (1988).
[19] E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984).
[20] E. K. U. Gross and W. Kohn, Phys. Rev. Lett. **55**, 2850 (1985).
[21] M. E. Casida, C. Jamorski, K. C. Casida, and D. R. Salahub, J. Chem. Phys. **108**, 4439 (1998).
[22] M. E. Casida, *Time-Dependent Density Functional Response Theory for Molecules*, edited by D. P. Chong (World Scientific, Singapore, 1995), Vol. 1.
[23] M. J. G. Peach, P. Benfield, T. Helgaker, and D. J. Tozer, J. Chem. Phys. **128**, 044118 (2008).
[24] P. Dev, S. Agrawal, and N. J. English, J. Chem. Phys. **136**, 224301 (2012).
[25] F. Plasser, B. Thomitzni, S. A. Bäppler, J. Wenzel, D. R. Rehn, M. Wormit, and A. Dreuw, J. Comput. Chem. **36**, 1609 (2015).
[26] S. A. Mewes, F. Plasser, and A. Dreuw, J. Chem. Phys. **143**, 171101 (2015).
[27] D. J. Tozer and N. C. Handy, J. Chem. Phys. **109**, 10180 (1998).
[28] Z.-L. Cai, K. Sendt, and J. R. Reimers, J. Chem. Phys. **117**, 5543 (2002).
[29] A. Dreuw, J. L. Weisman, and M. Head-Gordon, J. Chem. Phys. **119**, 2943 (2003).
[30] D. J. Tozer, J. Chem. Phys. **119**, 12697 (2003).
[31] O. Gritsenko and E. J. Baerends, J. Chem. Phys. **121**, 655 (2004).
[32] L. Serrano-Andrés and M. P. Fülscher, J. Am. Chem. Soc. **120**, 10912 (1998).
[33] C. Faber, P. Boulanger, I. Duchemin, C. Attaccalite, and X. Blase, J. Chem. Phys. **139**, 194308 (2013).
[34] K. Sendt, L. A. Johnston, W. A. Hough, M. J. Crossley, N. S. Hush, and J. R. Reimers, J. Am. Chem. Soc. **124**, 9299 (2002).
[35] M. Gouterman, *The Porphyrins*, edited by D. Dolphin (Academic, New York, 1978), Vol. 3.
[36] C. Li, J. Ly, B. Lei, W. Fan, D. Zhang, J. Han, M. Meyyappan, M. Thompson, and C. Zhou, J. Phys. Chem. B **108**, 9646 (2004).
[37] G. Sedghi, V. M. García-Suárez, L. J. Esdaile, H. L. Anderson, C. J. Lambert, S. Martín, D. Bethell, S. J. Higgins, M. Elliott, N. Bennett *et al.*, Nat. Nanotechnol. **6**, 517 (2011).
[38] L.-L. Li and E. W.-G. Diau, Chem. Soc. Rev. **42**, 291 (2013).
[39] S. Mathew, A. Yella, P. Gao, R. Humphry-Baker, B. F. Curchod, N. Ashari-Astani, I. Tavernelli, U. Rothlisberger, M. K. Nazeeruddin, and M. Grätzel, Nat. Chem. **6**, 242 (2014).
[40] A. Yella, H.-W. Lee, H. N. Tsao, C. Yi, A. K. Chandiran, M. K. Nazeeruddin, E. W.-G. Diau, C.-Y. Yeh, S. M. Zakeeruddin, and M. Grätzel, Science **334**, 629 (2011).
[41] W. G. Aulbur, L. Jönsson, and J. W. Wilkins, Solid State Phys. **54**, 1 (2000).
[42] S. Ishii, K. Ohno, Y. Kawazoe, and S. G. Louie, Phys. Rev. B **63**, 155104 (2001).
[43] Y. Noguchi, S. Ishii, K. Ohno, and T. Sasaki, J. Chem. Phys. **129**, 104104 (2008).
[44] Y. Noguchi and K. Ohno, Phys. Rev. A **81**, 045201 (2010).
[45] Y. Noguchi, M. Hiyama, H. Akiyama, and N. Koga, J. Chem. Phys. **141**, 044309 (2014).
[46] D. Hirose, Y. Noguchi, and O. Sugino, Phys. Rev. B **91**, 205111 (2015).
[47] C. A. Rozzi, D. Varsano, A. Marini, E. K. U. Gross, and A. Rubio, Phys. Rev. B **73**, 205119 (2006).
[48] J. Spencer and A. Alavi, Phys. Rev. B **77**, 193110 (2008).
[49] F. Hüser, T. Olsen, and K. S. Thygesen, Phys. Rev. B **87**, 235132 (2013).
[50] S. Sharifzadeh, P. Darancet, L. Kronik, and J. B. Neaton, J. Phys. Chem. Lett. **4**, 2197 (2013).
[51] B. Moore, H. Sun, N. Govind, K. Kowalski, and J. Autschbach, J. Chem. Theory Comput. **11**, 3305 (2015).