

# Quantum chemical replica-exchange umbrella sampling molecular dynamics simulations reveal the formation mechanism of iron phthalocyanine from iron and phthalonitrile

Shingo Ito, Ying Wang, Yuko Okamoto, and Stephan Irle

Citation: *J. Chem. Phys.* **149**, 072332 (2018); doi: 10.1063/1.5026956

View online: <https://doi.org/10.1063/1.5026956>

View Table of Contents: <http://aip.scitation.org/toc/jcp/149/7>

Published by the [American Institute of Physics](#)

---

## Articles you may be interested in

[Transition path theory from biased simulations](#)

The Journal of Chemical Physics **149**, 072336 (2018); 10.1063/1.5027253

[An efficient Bayesian kinetic lumping algorithm to identify metastable conformational states via Gibbs sampling](#)

The Journal of Chemical Physics **149**, 072337 (2018); 10.1063/1.5027001

[Reweighted autoencoded variational Bayes for enhanced sampling \(RAVE\)](#)

The Journal of Chemical Physics **149**, 072301 (2018); 10.1063/1.5025487

[Flexible selection of the solute region in replica exchange with solute tempering: Application to protein-folding simulations](#)

The Journal of Chemical Physics **149**, 072304 (2018); 10.1063/1.5016222

[Molecular exchange Monte Carlo: A generalized method for identity exchanges in grand canonical Monte Carlo simulations](#)

The Journal of Chemical Physics **149**, 072318 (2018); 10.1063/1.5025184

[Redox potential replica exchange molecular dynamics at constant pH in AMBER: Implementation and validation](#)

The Journal of Chemical Physics **149**, 072338 (2018); 10.1063/1.5027379

---

PHYSICS TODAY

WHITEPAPERS

### ADVANCED LIGHT CURE ADHESIVES

Take a closer look at what these environmentally friendly adhesive systems can do

READ NOW

PRESENTED BY  
 MASTERBOND<sup>®</sup>  
ADHESIVES | SEALANTS | COATINGS

# Quantum chemical replica-exchange umbrella sampling molecular dynamics simulations reveal the formation mechanism of iron phthalocyanine from iron and phthalonitrile

Shingo Ito,<sup>1</sup> Ying Wang,<sup>1,a)</sup> Yuko Okamoto,<sup>2,3,4,5,6,b)</sup> and Stephan Irle<sup>1,7,8,b)</sup>

<sup>1</sup>Department of Chemistry, Graduate School of Science, Nagoya University, Nagoya 464-8602, Japan

<sup>2</sup>Department of Physics, Graduate School of Science, Nagoya University, Nagoya 464-8602, Japan

<sup>3</sup>Structural Biology Research Center, Graduate School of Science, Nagoya University, Nagoya 464-8602, Japan

<sup>4</sup>Center for Computational Science, Graduate School of Engineering, Nagoya University, Nagoya 464-8603, Japan

<sup>5</sup>Information Technology Center, Nagoya University, Nagoya 464-8601, Japan

<sup>6</sup>JST-CREST, Nagoya 464-8602, Japan

<sup>7</sup>Institute of Transformative Bio-Molecules (WPI-ITbM), Nagoya University, Nagoya 464-8601, Japan

<sup>8</sup>Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA

(Received 26 February 2018; accepted 15 June 2018; published online 16 July 2018)

Phthalocyanine (Pc) and its metal complexes (MPcs) have been used industrially since their discovery in the early 20th century. The phthalonitrile (PN) method is a well-known synthesis method in which Pc or MPc can be afforded by heating a mixture of PN and metal powders over 280 °C with only moderate yield. However, the formation mechanism of the phthalocyanines and the intermediate stages of this seemingly simple reaction have yet to be fully understood. To study this mechanism computationally, we carried out quantum chemical molecular dynamics (MD) simulations based on the density-functional tight-binding (DFTB) method, applying the replica-exchange umbrella sampling (REUS) method, starting from four PN molecules and one iron atom. The DFTB-REUS-MD simulations successfully yielded FePc, and a metastable structure very similar to FePc but with a reactive nitrene unit was also identified that might explain the incomplete conversion of the reactants into FePc. Analysis of the MD trajectories reveals a three-step FePc formation mechanism for the PN method. *Published by AIP Publishing.* <https://doi.org/10.1063/1.5026956>

## I. INTRODUCTION

One-pot synthetic reaction mechanisms are desirable for large-scale industrial synthesis of chemical materials. Phthalocyanine (Pc) and phthalocyanine metal-complexes (MPcs), shown in Figs. 1(a) and 1(b), respectively, are examples for chemical materials utilized as pigments for decades due to their vivid green or blue color. For example, the copper phthalocyanine complex has a brilliant blue color, making CuPc an ideal photostable dye in pigments.<sup>1</sup> The Pc molecules and complexes possess a cyclic structure, in which all atoms are located in the same plane and are structurally similar to their cousins, the porphyrin molecules and complexes. Remarkably, MPc can be synthesized in a one-pot synthetic reaction, the so-called phthalonitrile (PN) method, albeit only with moderate yield.<sup>2-4</sup> It is a well-known synthesis in which Pc or MPc is obtained from a mixture of PN and metal powder at high temperatures above 280 °C.

Recently, MPcs have been used not only for dyes or pigments but also as organic semiconductors in covalent organic

frameworks,<sup>5,6</sup> optical data storing media, etc. Somewhat surprisingly, Pc and MPc were discovered accidentally in the early 20th century,<sup>7-9</sup> and even more surprisingly, the reaction mechanism underlying the nearly century-old PN method is not fully understood to date.<sup>10</sup>

Examples of hypothetical formation mechanisms for MPc from simple reactants in two different synthetic methods are shown in Fig. 2. In the Wyler method, which has been mainly used industrially due to its higher reaction yield, MPc is synthesized from phthalic anhydride with metal powder or metal salt and urea. The second is the phthalonitrile (PN) method, which is used on laboratory scale. In these methods, PN and metal powder or metal salt are heated up to 280 °C.<sup>2-4</sup>

To this day, the formation mechanism of Pc and MPc complexes has not clearly been established. Moreover, intermediates, transition states, and free energy barriers have not been identified, neither experimentally nor theoretically.

In order to theoretically study the mechanism for forming these molecules by the PN method, we performed quantum chemical molecular dynamics (QM/MD) simulations for the formation of iron phthalocyanine (FePc) from four PNs and one iron atom in the gas phase. Straightforward QM/MD simulations including periodic boundary conditions (PBC) on computationally manageable time scales (nanoseconds) do not

<sup>a)</sup>Present addresses: State Key Laboratory of Rare Earth Resource Utilization, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun 130022, People's Republic of China.

<sup>b)</sup>Authors to whom correspondence should be addressed: okamoto@tb.phys.nagoya-u.ac.jp and irles@ornl.gov.

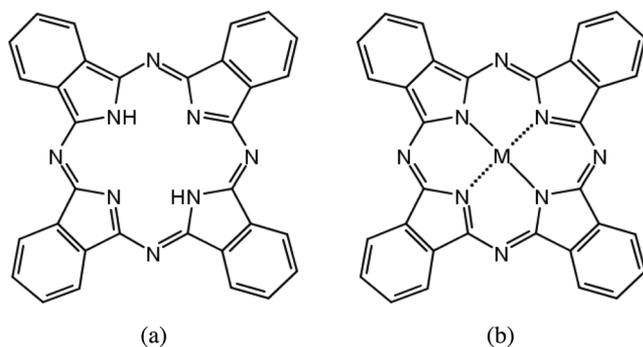


FIG. 1. Structures of (a) Pc and (b) MPc.

yield the desired product, no matter how high a temperature is chosen. On the MD time scale, FePc formation is a rare event because high barriers on the free energy pathway, associated with the anionic cyclic polymerization of the four PN units, must exist. Unfortunately, microsecond time scales are not affordable even for approximate density functional theory (DFT)-based first principles MD simulations. To overcome this difficulty, we employed one of the generalized-ensemble algorithms in this work (for reviews, see, e.g., Refs. 11–15), which greatly enhance conformational sampling. These methods have traditionally been used in combination with classical Monte Carlo (MC) and molecular dynamics (MD) simulations. However, for the study of chemical reactions, some groups reported promising results from QM-MC and QM-MD simulations.<sup>16–18</sup> In the present study, we employed the replica-exchange umbrella sampling (REUS) method,<sup>19</sup> which combines the replica-exchange method (REM)<sup>20,21</sup> and the umbrella sampling (US) method,<sup>22</sup> in combination with the density-functional tight-binding (DFTB)<sup>23</sup> quantum chemical potential. This enhanced reactive MD method has been implemented<sup>24</sup> recently in the freely available DFTB+

program package.<sup>25</sup> The use of multiple umbrella potentials via the REUS method along the reaction coordinate is crucial for the study of the FePc self-assembly mechanism: Since it cannot be expected that the four PN molecules arrange themselves in a concerted single-step reaction around the iron cation, a stepwise bond formation dynamics with intermediate steps demands that multiple regions on the potential energy surface can be visited effectively. For this reason, the REUS with its enhanced sampling capability is required to reveal the intermediate steps of the self-assembly reaction.

This article is organized as follows. In Sec. II, the relevant methods are explained. In Sec. III, computational details of the simulations are given. In Sec. IV, the results of the simulations are presented. Section V presents the conclusions.

## II. METHODS

### A. Umbrella sampling

In the US method,<sup>22</sup> we use an umbrella potential  $V(q)$ , which is defined by

$$V(q) = \frac{1}{2}k\{\xi(q) - d\}^2, \quad (1)$$

where  $\xi(q)$  is a reaction coordinate. Two adjustable parameters, a force constant  $k$  and a reaction coordinate midpoint value  $d$ , are the adjustable parameters in the umbrella potential. The umbrella potential  $V(q)$  is added to the original potential energy  $E_0$ , and the total potential energy is defined by

$$E(q) = E_0(q) + V(q). \quad (2)$$

The umbrella sampling simulation allows the system to explore the region around  $\xi(q) = d$ , which may not be reached during regular MC or MD simulations due to the existence of high energy barriers.

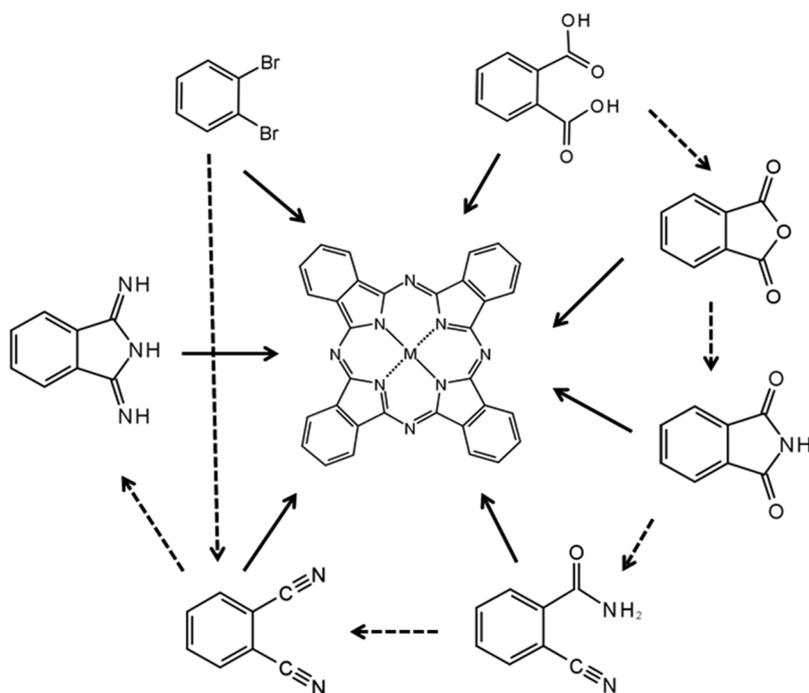


FIG. 2. Hypothetical formation mechanisms for the MPc complex.

In a conventional US simulation,  $M$  independent US simulations at a temperature  $T_0$  are performed with  $M$  different umbrella potentials  $V_m$  ( $m = 1, 2, \dots, M$ ).  $V_m$  ( $m = 1, 2, \dots, M$ ) have midpoint distances  $d_m$  ( $m = 1, 2, \dots, M$ ) and force constants  $k_m$  ( $m = 1, 2, \dots, M$ ) which can depend on  $m$  and are defined by

$$V_m(q) = \frac{1}{2}k_m\{\xi(q) - d_m\}^2. \quad (3)$$

Therefore, the total potential energy is given by

$$E_m(q) = E_0(q) + V_m(q). \quad (4)$$

After  $M$  US simulations with  $M$  different  $E_m$ , the free energy and physical quantities of the original unbiased system, which is based on potential energy  $E_0$  at temperature  $T_0$ , are calculated by using the weighted histogram analysis method (WHAM).<sup>26</sup> By solving the following WHAM equations self-consistently, the probability distribution  $P_{T_0}(\xi)$  and the “dimensionless” Helmholtz free energy  $f_m$  of the original unbiased system can be obtained:<sup>27,28</sup>

$$P_{T_0}(\xi) = \frac{\sum_{m=1}^M N_m(\xi)}{\sum_{m=1}^M n_m e^{f_m - \beta_0 U_m(\xi)}}, \quad (5)$$

$$e^{-f_m} = \sum_{\xi} P_{T_0}(\xi) e^{-\beta_0 U_m(\xi)}, \quad (6)$$

where  $N_m(\xi)$  is the histogram of distributions of reaction coordinate  $\xi$  and  $n_m$  is the total number of samples for the simulation with  $E_m$ . The potential of mean force (PMF),  $F_{T_0}(\xi)$ , which is the Helmholtz free energy as a function of the reaction coordinate of the original unbiased system, is given by

$$F_{T_0}(\xi) = -k_B T_0 \ln P_{T_0}(\xi). \quad (7)$$

The expectation value of a physical quantity  $A$  of the original unbiased system is given by

$$\langle A \rangle = \frac{\sum_{\xi} A(\xi) P_{T_0}(\xi)}{\sum_{\xi} P_{T_0}(\xi)}. \quad (8)$$

## B. Replica-exchange umbrella sampling

The Replica-Exchange Umbrella Sampling (REUS)<sup>19</sup> is one of the generalized-ensemble algorithms that combines REM and US. In REUS, the system consists of  $M$  non-interacting replicas  $i$  at  $M$  different umbrella potentials  $U_m(q)$  ( $i, m = 1, 2, \dots, M$ ). The total potential energy  $E_m(q)$  by the  $m$ th umbrella potential is defined by Eq. (4), where

$$U_m(q) = \sum_{l=1}^L \lambda_m^{(l)} V_m^{(l)}(q), \quad (9)$$

$$V_m^{(l)}(q) = \frac{1}{2}k_m^{(l)}\{\xi(q) - d_m^{(l)}\}^2. \quad (10)$$

Here, the parameters  $\lambda_m^{(l)}$  ( $l = 1, 2, \dots, L; m = 1, 2, \dots, M$ ), usually chosen empirically from short test simulations, are the weighting factors for the corresponding umbrella potentials. Their values are selected to ensure sufficient population overlap between those of neighboring umbrella potentials.

The sets of labels for replicas,  $i$ , and umbrella potential,  $m$ , are related by permutation,

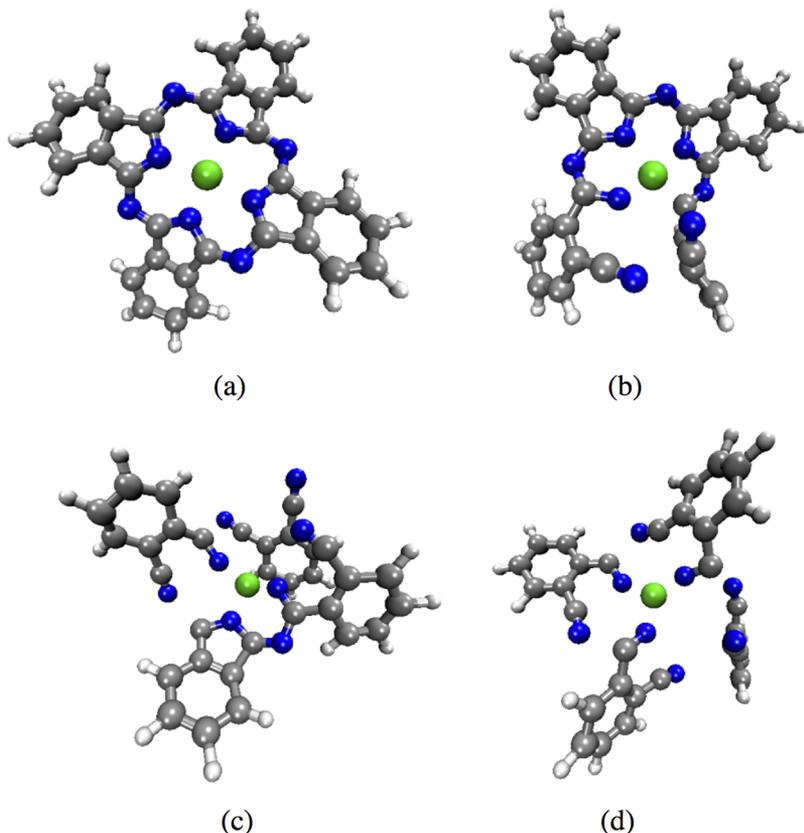


FIG. 3. Four different initial coordinates of the “A set” DFTB-REUS-MD simulations with 16 replicas. The gray colored atoms are carbon, blue colored atoms are nitrogen, white colored atoms are hydrogen, and green colored atoms are iron. (a) Initial coordinates for replicas 1–6. (b) Initial coordinates for replicas 7–9. (c) Initial coordinates for replicas 10 and 11. (d) Initial coordinates for replicas 12–16.

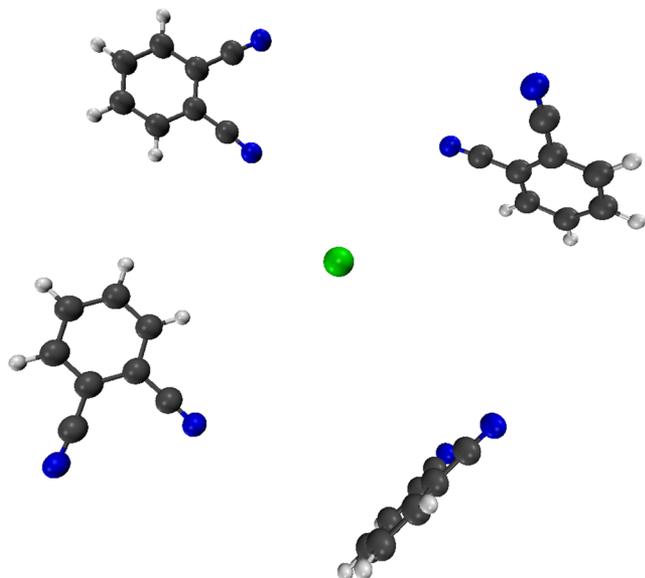


FIG. 4. Initial coordinates of the “B set” DFTB-REUS-MD simulation for each replica. Color coding follows the definition given in Fig. 3.

$$\begin{cases} i(m) = f(m), \\ m(i) = f^{-1}(i), \end{cases} \quad (11)$$

where  $f(m)$  is a permutation function of umbrella potential label  $m$  and  $f^{-1}(m)$  is its inverse.

The state  $X$  in this generalized ensemble can be written as

$$X = \{x_1^{[i(1)]}, \dots, x_M^{[i(M)]}\} = \{x_{m(1)}^{[1]}, \dots, x_{m(M)}^{[M]}\}, \quad (12)$$

where  $x_m^{[i]}$  is specified by the coordinates  $q^{[i]}$  and momenta  $p^{[i]}$  of  $N$  atoms in replica  $i$  at temperature  $T_m$ ,

$$x_m^{[i]} \equiv \{q^{[i]}, p^{[i]}\}_m, \quad (13)$$

$$q^{[i]} = (q_1^{[i]}, q_2^{[i]}, \dots, q_N^{[i]}), \quad (14)$$

$$p^{[i]} = (p_1^{[i]}, p_2^{[i]}, \dots, p_N^{[i]}). \quad (15)$$

Independent MC or MD simulation with  $M$  replicas is carried out simultaneously in the canonical ensemble at temperature  $T_0$ , and every few time steps, replicas  $i$  and  $j$  with neighboring umbrella potentials  $U_m(q)$  and  $U_{m+1}(q)$ , respectively, are exchanged:

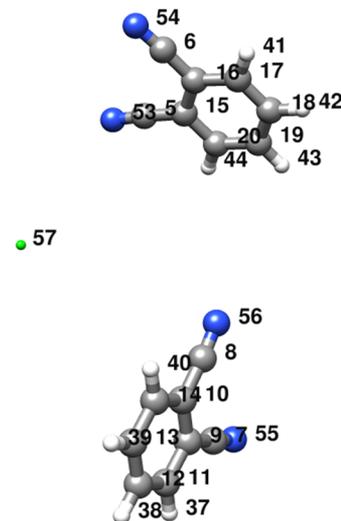
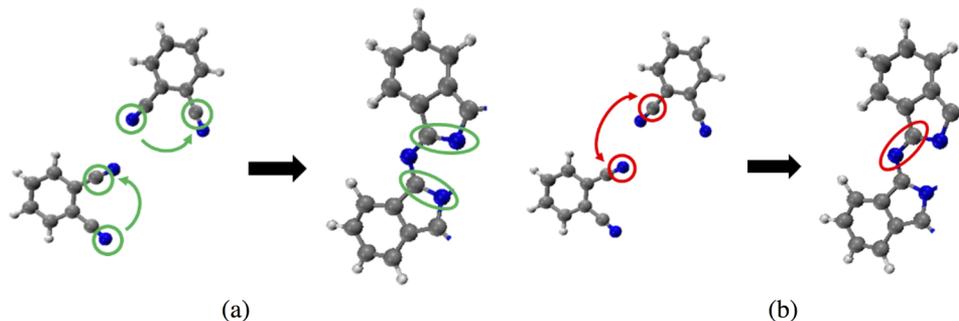


FIG. 6. Atom labels in four phthalonitriles and one iron atom of set B.

$$X = \{\dots, x_m^{[i]}, x_{m+1}^{[j]}, \dots\} \rightarrow X' = \{\dots, x_m^{[j]}, x_{m+1}^{[i]}, \dots\}. \quad (16)$$

The replica exchange is performed by the following probability  $\omega$  of the Metropolis criterion:

$$\omega(X \rightarrow X') = \min\{1, \exp(-\Delta)\}, \quad (17)$$

where

$$\Delta = \beta_0 \{U_m(q^{[j]}) - U_m(q^{[i]}) - U_{m+1}(q^{[i]}) + U_{m+1}(q^{[j]})\}. \quad (18)$$

In a REUS simulation, by replica exchange, a random walk in reaction coordinate  $\xi$  is performed without getting trapped in local-minimum states. After a REUS simulation is performed, the PMF can be obtained as in US using the WHAM [see Eqs. (5)–(7)].

### III. COMPUTATIONAL DETAILS

We employed the density-functional tight-binding (DFTB)-based QM/MD simulation with REUS, which is referred to as DFTB-REUS-MD. All calculations were performed using our modified version of the DFTB+ program.<sup>24,25</sup> The initial system was composed of four PN molecules and a single iron atom in gas phase, with initial coordinates arranged to resemble somewhat the arrangement of the molecular units in an FePc complex. The MD simulation time in these

FIG. 5. Two types of C–N bonds in PNs (a) in “internal” PNs and (b) in “external” PNs.

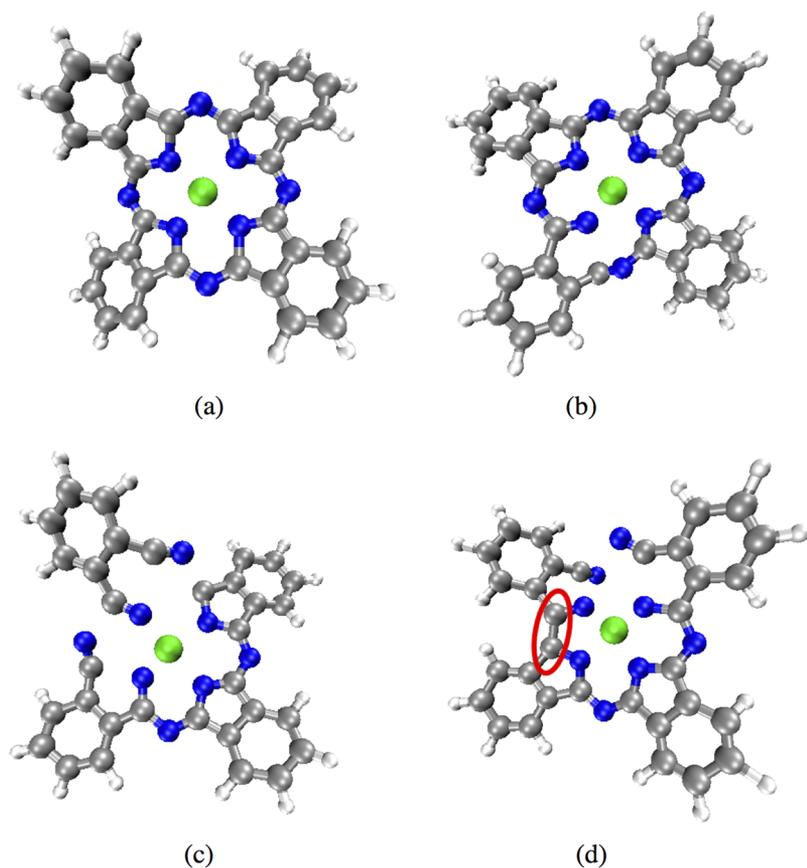


FIG. 7. Typical snapshots from “A set” DFTB-REUS-MD simulation. (a) FePc. (b) PN tetramer. (c) PN trimer. (d) Structure M. The C—C mismatched bond is marked by a red circle.

simulations was 150 ps per replica. To study the dependency of the trajectories on the initial structures, we also carried out a DFTB-REUS-MD simulation of 50 ps per replica by using different initial coordinates where PN molecules and the iron atom were placed more randomly as compared with the former simulation.

Regarding the DFTB, we employed the second-order self-consistent charge (SCC-DFTB, nowadays also referred to as DFTB2) flavor,<sup>23</sup> which can be considered as an approximate Density Functional Theory (DFT) method<sup>29,30</sup> within the Slater-Koster approximation.<sup>31</sup> The SCC-DFTB method has reasonable accuracy and very low computational cost as compared to traditional DFT-based QM methods. The mio-1-1<sup>23</sup> and trans3d-0-1<sup>32</sup> parameter sets were employed throughout this work. A finite electronic temperature of 1000 K was employed in order to obtain convergence of the DFTB electronic density, and the trajectories were therefore run on the DFTB Mermin free energy<sup>33</sup> surfaces.

The simulation of chemical reactions in a DFTB-REUS-MD trajectory is naturally sensitive to the initial placements of the reactant molecules. We therefore decided to perform two sets of simulations. In the first “A set,” we prepared four different sets of initial coordinates, which had four PNs near the iron atom and similar to FePc as shown in Fig. 3. In the second, “B set,” we placed four PNs around one iron atom randomly as shown in Fig. 4, where all replicas started from the same initial coordinates. After the simulation, we performed DFTB geometry optimization for several structures to quench the high-temperature structures.

During the formation of FePc, there were two types of covalent bonds between a carbon atom and a nitrogen atom. One is a C—N bond in “internal” PNs as shown in Fig. 5(a). The other is a C—N bond in “external” PNs, which means that two PNs are bound by a C—N bond as shown in Fig. 5(b). Therefore, we defined a reaction coordinate as a linear combination of eight distances between a carbon atom and a nitrogen atom in internal and external PNs as

$$\xi(\mathbf{r}) = |r_{C_1} - r_{N_{51}}| + |r_{C_2} - r_{N_{49}}| + |r_{C_3} - r_{N_{52}}| + |r_{C_4} - r_{N_{53}}| \\ + |r_{C_5} - r_{N_{54}}| + |r_{C_6} - r_{N_{56}}| + |r_{C_7} - r_{N_{50}}| + |r_{C_8} - r_{N_{55}}|. \quad (19)$$

Subscripts in Eq. (19) correspond to labels of atoms as shown in Fig. 6.

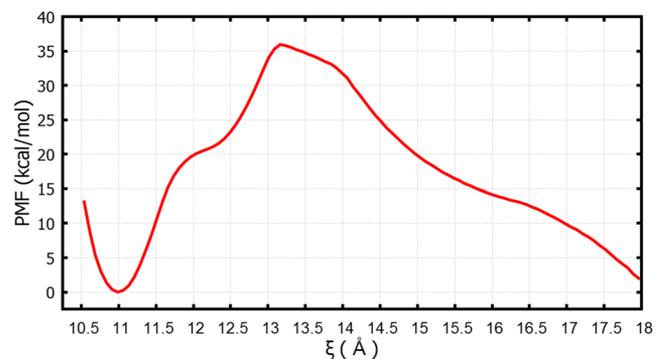


FIG. 8. The PMF calculated from the “A set” DFTB-REUS-MD simulation.

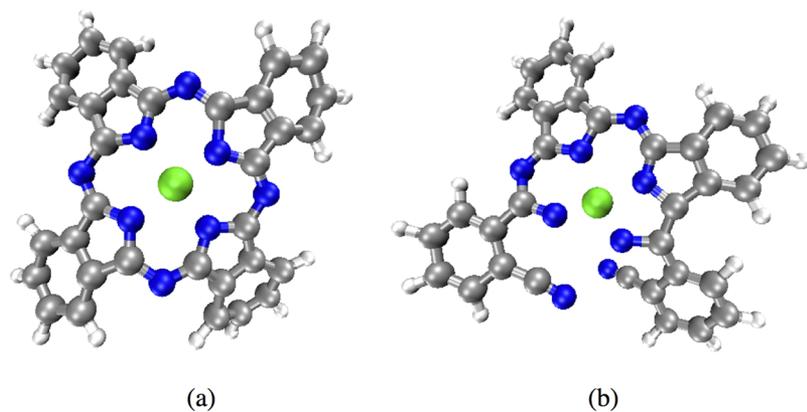


FIG. 9. Snapshots from the “B set” DFTB-REUS-MD simulation. (a) FePc. (b) Structure M.

We used  $M = 16$  replicas for the DFTB-REUS-MD simulations, which have different  $k$  and  $d$  values in each umbrella potential. Each replica had the same initial coordinates. In the “A set” DFTB-REUS-MD simulation, we set  $k$  values

to be 1.0, 50.0, 100.0, 100.0, 100.0, 75.0, 50.0, 50.0, 50.0, 20.0, 20.0, 20.0, 20.0, 20.0, 20.0, and 20.0 kcal/mol  $\text{\AA}^2$  and  $d_m$  values were chosen to be 11.0, 11.5, 12.0, 12.25, 12.5, 12.75, 13.0, 13.25, 13.5, 14.0, 14.5, 15.0, 15.5, 16.0, 16.5, and

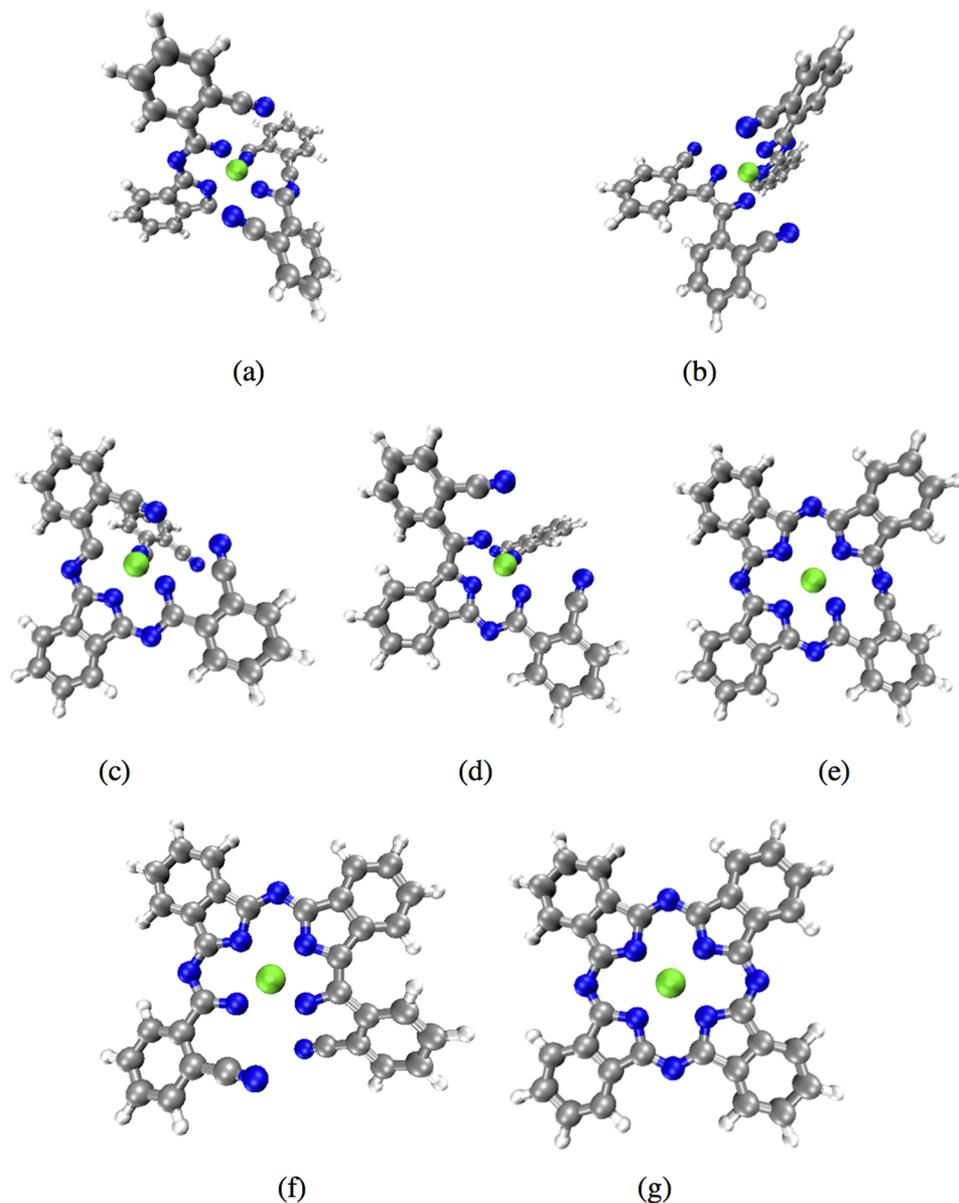


FIG. 10. Optimized DFTB structures. (a) PN dimer and two isolated PNs. (b) C—C dimer and two isolated PNs. (c) PN trimer and one isolated PN. (d) C—C trimer and one isolated PN. (e) PN tetramer. (f) Structure M. (g) FePc.

17.0 Å. In the “B set” DFTB-REUS-MD simulation, we set  $k = 1.5$  kcal/mol Å<sup>2</sup> for all replicas and  $d_m$  values were chosen to be 11.0, 11.5, 12.0, 12.25, 12.5, 12.75, 13.0, 13.25, 13.5, 14.0, 14.5, 15.0, 15.5, 16.0, 16.5, and 17.0 Å. Both MD nuclear temperature and electronic temperature were set to 1000 K and the former was controlled by the No se-Hoover chain method<sup>34</sup> in the NVT ensemble. We set the MD time step to be 0.5 fs and each pair of replicas was exchanged every 40 MD steps for the former simulation and 20 MD steps for the latter simulation. Therefore, the total numbers of replica-exchange trials were 2500 and 5000 in these REUS simulations. After the simulations, we computed the PMF according to Eq. (7).

Selected DFTB-optimized structures were reoptimized using the Tao, Perdew, Staroverov, and Scuseria (TPSS)<sup>34</sup> density functional theory (DFT) with Ahlrich’s split-valence all-electron basis set,<sup>35</sup> with polarization functions on all non-hydrogen atoms. The resolution-of-identity (RI) approach<sup>36</sup> was employed as implemented in the program package TURBOMOLE.<sup>37</sup> With the exception of the iron atom, which was computed in its quintet spin state, we employed the singlet spin state in all DFT calculations.

## IV. RESULTS

### A. The “A set” REUS simulation

We performed the DFTB-REUS-MD simulation for 150 ps per replica. We obtained several structures as shown in Fig. 7. After the simulation, we calculated the PMF along the reaction coordinate as shown in Fig. 8. We could obtain the PMF only in the vicinity of the FePc structure in the simulation “set A.”

In the PMF in Fig. 8, the lowest free energy is at  $\xi = 11.0$  Å and only FePc was found as shown in Fig. 7(a) in this region.

This means that FePc was the most stable structure with a reaction coordinate value between 10.5 Å and 18.0 Å. There are shoulders around  $\xi = 12.0$  Å and 16.0 Å. In the former shoulder, almost complete FePc structures, which only failed to create the last C–N bond in the internal PN, were formed as shown in Fig. 7(b). In the latter shoulder, which is located between  $\xi = 15.0$  Å and 17.0 Å, there were some PN trimers and these trimers had very similar structures as shown in Fig. 7(c).

Around  $\xi = 13.5$  Å, there is a very high barrier and this barrier separates the FePc and a trimer similar to FePc. Therefore, we propose that the addition of the last PN to the oligomerized trimer is a bottleneck in the FePc formation mechanism. We also found a specific tetramer, shown in Fig. 7(b), which is very similar to FePc in this region. As we mentioned above, internal and external bonds between N and C atoms were made by the process of forming FePc. By contrast, an external bond between C and C, which should not be made by the process of forming FePc, appeared in this structure as shown in Fig. 7(d), and the mismatched bond pattern results in this structure having a reactive nitrene group that is prone to further chemical reactions. In the following, we will refer to this structure as “Structure M.” Structure M had the reaction coordinate value between  $\xi = 15.0$  Å and 17.0 Å.

### B. The “B set” REUS simulation

We also performed the DFTB-REUS-MD simulation for 50 ps per replica using the initial coordinates in Fig. 4, which set the four PNs and iron atoms more randomly than the initial structures in the “A set” simulations. In this simulation, we again succeeded in the formation of FePc, which had not been possible using conventional DFTB-MD simulations without

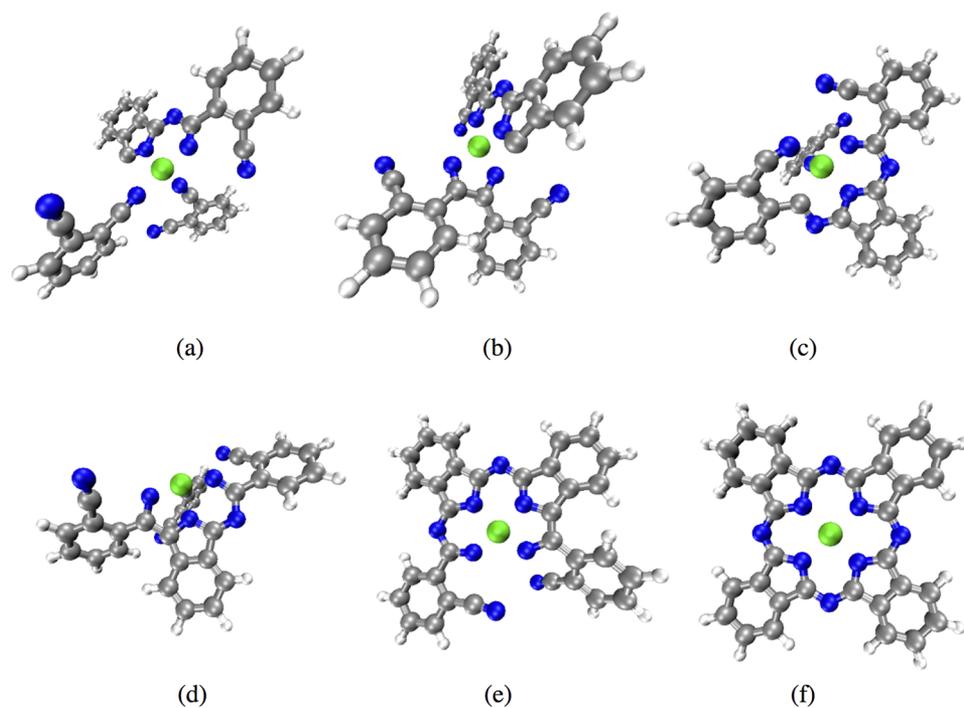


FIG. 11. DFT structures, reoptimized from DFTB structures. (a) PN dimer and two isolated PNs. (b) C–C dimer and two isolated PNs. (c) PN trimer and one isolated PN. (d) C–C trimer and one isolated PN. (e) Structure M. (f) FePc.

TABLE I. The total energy after geometry optimizations by DFTB.

	Mermin free energy DFTB $\Delta E$ (kcal/mol)	DFT $\Delta E$ (kcal/mol)
4 PNs + 1 Fe	259.8	307.6
PN dimer	103.5	179.2
C—C dimer	69.7	211.3
PN trimer	63.4	167.8
C—C trimer	59.0	132.2
PN tetramer	30.7	N/A <sup>a</sup>
Structure M	20.1	94.8
FePc	0.0	0.0

<sup>a</sup>In DFT optimization, the PN tetramer formed the FePc.

REUS. We also found Structure M again, with its characteristic C—C bond and nitrene reactive center. These two structures are shown in Fig. 9. Furthermore, we found dimer and trimer structures, which had a bond between C and C like Structure M, and these dimers and trimers formed Structure M finally. We call these structures the “C—C dimer” and “C—C trimer.” The movies of forming FePc in one of the replicas (Replica 5) and Structure M in one of the replicas (Replica 2) by the “B set” REUS simulation are given in the [supplementary material](#).

In order to study the stability of the structures, we performed DFTB optimizations of several structures, which are the initial molecular structure, PN dimer, C—C dimer, PN trimer, C—C trimer, PN tetramer, Structure M, and FePc as observed in the “B Set” DFTB-REUS-MD simulation. After eight structures were fully optimized by DFTB, we re-optimized them using the RI-TPSS-D3/def-SV(P)<sup>35–37</sup> method. The optimized structures from both DFTB and DFT are shown in Figs. 10 and 11 with the chemical structure files (xyz) of these optimized structures as given in the

[supplementary material](#). We also list the  $\Delta E$  values relative to the energy of the FePc optimized structure at the respective level of theory in Table I for both DFTB and DFT methods.

In Table I, FePc has the lowest energy and is the most stable structure. Surprisingly, by DFTB optimization, Structure M, which has a C—C bond, is more stable than the PN tetramer. This means that this C—C bond structure is one of the metastable states in the process of forming FePc. The overall exothermicity of the reaction, starting from 4 PN molecules and one iron atom to FePc, is 260 kcal/mol at the DFTB level of theory and 308 kcal/mol at the DFT level of theory. Intermediate structures are difficult to describe at the DFT level, as the ground state is an open-shell singlet, and hence the close-shell singlet DFT energies reported here are much higher than the Mermin free energies of DFTB. Nevertheless, the DFT geometry optimizations confirm the existence of the intermediate structures, except for the PN tetramer structure, which rearranged into the FePc lower energy structure.

We further analyzed the structures, which were obtained by the “set B” DFTB-REUS-MD simulation, and found specific processes, basically following similar pathways as for “set A.” When PNs and the iron atom formed a FePc, each molecule was bound on specific pathways as shown in Fig. 12. As the first step, four PNs moved toward the iron atom and then PNs were bound with the iron atom [see Fig. 12(a)]. In the next step, carbon and nitrogen atoms in one PN, which was bound with the iron atom, created internal bonds. At the same time, this chemically altered PN, was bound with another PN and formed a “PN dimer” [see Fig. 12(b)]. In the same way, this dimer made an internal N—C bond and then formed “PN trimer” in turn by binding with the other PN [see Fig. 12(c)]. In the end, the last PN was bound with the PN trimer and a FePc was formed [see Fig. 12(d)].

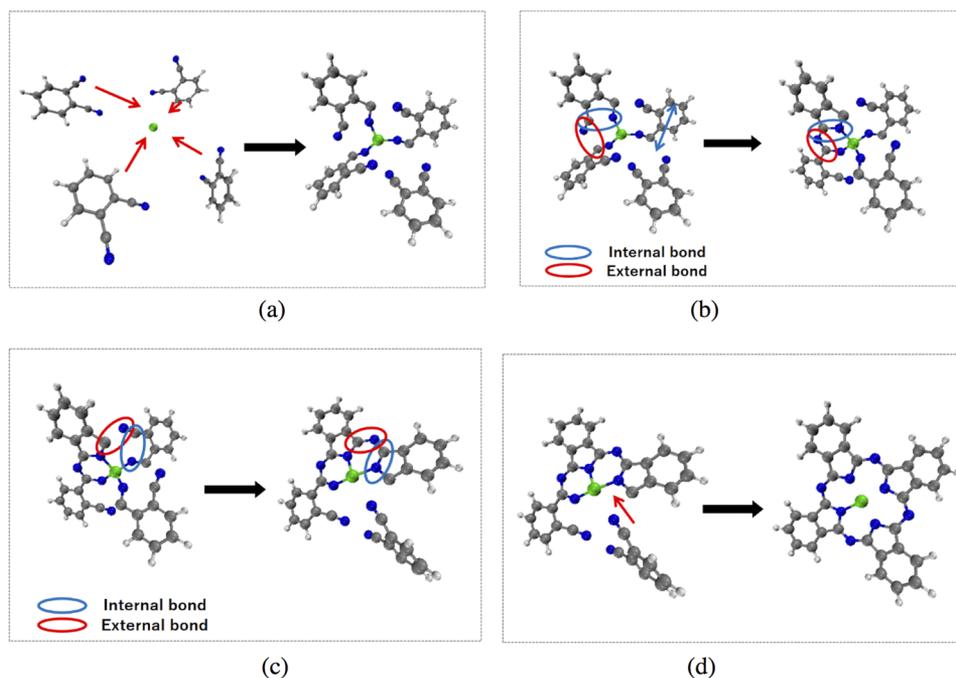


FIG. 12. Specific pathway of forming a FePc. (a) Four PNs moved into the iron atom. (b) Two PNs formed a PN dimer. (c) One of the two PNs was bound to the PN dimer and formed a PN trimer. (d) The last PN was bound to the PN trimer and formed a FePc.

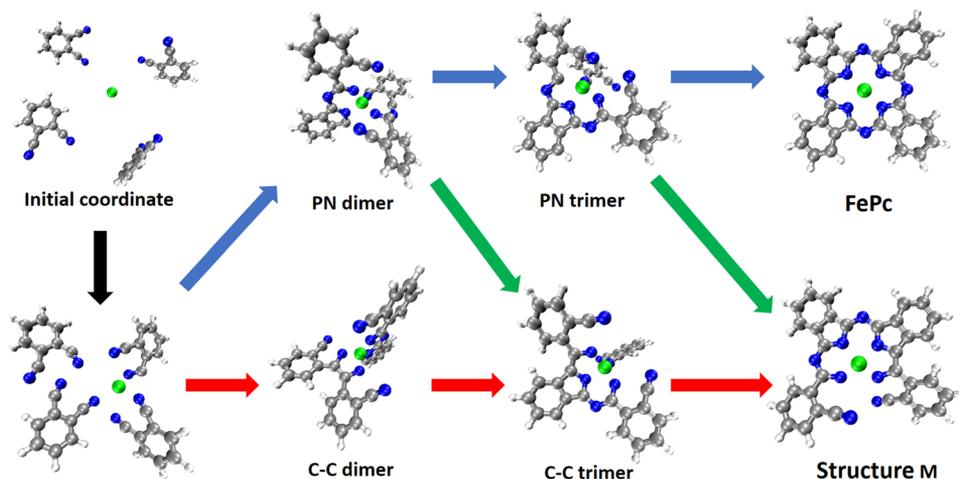


FIG. 13. The stepwise formation mechanism pathway of FePc from Fe and four PN molecules.

## V. CONCLUSIONS

We succeeded in observing the formation of FePc from one iron atom and four PNs in DFTB-REUS-MD simulations, irrespective of the initial geometries, at 1000 K. We calculated the PMF for the reaction coordinate used for the umbrella sampling and obtained interesting structures besides the thermodynamically most stable FePc, most prominently Structure M, which is a metastable structure featuring a reactive nitrene unit. Its prevalence indicates that such misbonded structures are frequently and irreversibly formed, preventing the formation of FePc and thereby causing the experimentally observed incomplete reaction with substantial byproduct in the one-pot PN method.

Moreover, by analyzing the structures and PMF, which were obtained by two sets of simulations differing only in initial geometries, we found specific pathways for the formation of a FePc from four PN and one iron atom. Based on our observations, we propose a “step-wise” pathway of forming FePc in this system and this stepwise pathway proceeds as shown in Fig. 13.

In the first step, four PNs approach the iron atom and are bound with the iron atom. In the second step, each PN connects with a neighboring PN stepwise but only three of the four PNs form a trimer. In these two steps, the reaction of forming FePc proceeds smoothly. However, the last step requires more time since the final fourth PN molecule needs to rotate and “squeeze” into the same plane as that of the PN trimer to finally form the prominent FePc structure. On the other hand, during the delicate cyclization reaction, if a bond between carbon and carbon is made, Structure M is formed. Once Structure M is formed, it is very difficult to break the strong C—C bond. Experimentally, it is known that the yield of FePc from PNs and iron or iron metal is not very high. Therefore, we also propose that Structure M is one of the final products not forming FePc.

## SUPPLEMENTARY MATERIAL

See [supplementary material](#) for Movie 1: The movie of FePc formation from the initial coordinate by the “B set” DFTB-REUS-MD simulation in Replica 5. The frame rate of

this movie was set to be 1 frame per 200 MD steps. Movie 2: The movie of Structure M formation from the initial coordinate by “B set” DFTB-REUS-MD simulation in Replica 2. The frame rate of this movie was set to be 1 frame per 200 MD steps. Also included are chemical structure files: PN\_dimer\_DFTB (the optimized Cartesian coordinate file of PN dimer of “B set” simulation by DFTB optimization); CeC\_dimer\_DFTB (the optimized Cartesian coordinate file of C-C dimer of “B set” simulation by DFTB optimization); PN\_trimer\_DFTB (the optimized Cartesian coordinate file of the PN trimer of “B set” simulation by DFTB optimization); CeC\_trimer\_DFTB (the optimized Cartesian coordinate file of the C—C trimer of “B set” simulation by DFTB optimization); PN\_tetramer\_DFTB: (the optimized Cartesian coordinate file of the PN tetramer of “B set” simulation by DFTB optimization); Structure\_M: (the optimized Cartesian coordinate file of the C—C tetramer of “B set” simulation by DFTB optimization); FePc\_DFTB: (the optimized Cartesian coordinate file of FePc of “B set” simulation by DFTB optimization); PN\_dimer\_DFT: (the optimized Cartesian coordinate file of the PN dimer of “B set” simulation by DFT optimization); CeC\_dimer\_DFT: (the optimized Cartesian coordinate file of the C—C dimer of “B set” simulation by DFT optimization); PN\_trimer\_DFT: (the optimized Cartesian coordinate file of the PN trimer of “B set” simulation by DFT optimization); CeC\_trimer\_DFT: (the optimized Cartesian coordinate file of the C—C trimer of “B set” simulation by DFT optimization); Structure\_M: (the optimized Cartesian coordinate file of the C—C tetramer of “B set” simulation by DFT optimization); FePc\_DFT: (the optimized Cartesian coordinate file of FePc of “B set” simulation by DFT optimization).

## ACKNOWLEDGMENTS

Some of the computations were performed on the supercomputers at the Institute for Molecular Science and at the Information Technology Center, Nagoya University. This work was supported, in part, by the Grants-in-Aid for Scientific Research (S) (No. JP16H06353), (A) (No. JP25247071), and the Grants-in-Aid for Scientific Research on Innovative Areas “Dynamical Ordering and Integrated Functions” (No. JP25102009). S. Irlé was supported in part by the

Laboratory Directed Research and Development (LDRD) Program of Oak Ridge National Laboratory. ORNL is managed by UT-Battelle, LLC, 651 for DOE under Contract No. DE-AC05-00OR22725.

- <sup>1</sup>A. R. Hedges, *Chem. Rev.* **98**, 2035 (1998).
- <sup>2</sup>P. A. Barrett, C. E. Dent, and R. P. Linstead, *J. Chem. Soc.* **1936**, 1719.
- <sup>3</sup>R. M. Christie and D. D. Deans, *J. Chem. Soc., Perkin Trans. 2* **1989**, 193.
- <sup>4</sup>B. N. McKeown, *Phthalocyanine Materials: Synthesis, Structure and Function* (Cambridge University Press, Cambridge, 1998).
- <sup>5</sup>P. Gregory, *High-Technology Applications of Organic Colorants* (Springer, USA, Boston, MA, 1991).
- <sup>6</sup>*Chemistry and Technology of Printing and Imaging Systems*, edited by P. Gregory (Springer, The Netherlands, Dordrecht, 1996).
- <sup>7</sup>A. Braun and J. Tcherniac, *Ber. Dtsch. Chem. Ges.* **40**, 2709 (1907).
- <sup>8</sup>H. de Diesbach and E. von der Weid, *Helv. Chim. Acta* **10**, 886 (1927).
- <sup>9</sup>R. P. Linstead and A. R. Lowe, *J. Chem. Soc.* **1934**, 1022.
- <sup>10</sup>S. W. Oliver and T. D. Smith, *J. Chem. Soc., Perkin Trans. 2* **1987**, 1597.
- <sup>11</sup>U. H. E. Hansmann and Y. Okamoto, *Curr. Opin. Struct. Biol.* **9**, 177 (1999).
- <sup>12</sup>A. Mitsutake, Y. Sugita, and Y. Okamoto, *Biopolymers* **60**, 96 (2001).
- <sup>13</sup>Y. Sugita and Y. Okamoto, in *Lecture Notes in Computational Science and Engineering, Computational Methods for Macromolecules: Challenges and Applications*, edited by T. Schlick and H. H. Gan (Springer-Verlag, Berlin, 2002), pp. 304–332.
- <sup>14</sup>Y. Okamoto, *J. Mol. Graphics Modell.* **22**, 425 (2004).
- <sup>15</sup>A. Mitsutake, Y. Mori, and Y. Okamoto, in *Biomolecular Simulations*, edited by L. Monticelli and E. Salonen (Humana Press, New York, 2013), pp. 153–195.
- <sup>16</sup>Y. Ishikawa, Y. Sugita, T. Nishikawa, and Y. Okamoto, *Chem. Phys. Lett.* **333**, 199 (2001).
- <sup>17</sup>T. H. Choi, *Chem. Phys. Lett.* **543**, 45 (2012).
- <sup>18</sup>Y. Mori and Y. Okamoto, *Phys. Rev. E* **87**, 023301 (2013).
- <sup>19</sup>Y. Sugita, A. Kitao, and Y. Okamoto, *J. Chem. Phys.* **113**, 6042 (2000).
- <sup>20</sup>K. Hukushima and K. Nemoto, *J. Phys. Soc. Jpn.* **65**, 1604 (1996).
- <sup>21</sup>Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.* **314**, 141 (1999).
- <sup>22</sup>G. M. Torrie and J. P. Valleau, *J. Comput. Phys.* **23**, 187 (1977).
- <sup>23</sup>M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert, *Phys. Rev. B* **58**, 7260 (1998).
- <sup>24</sup>S. Ito, S. Irle, and Y. Okamoto, *Comput. Phys. Commun.* **204**, 1 (2016).
- <sup>25</sup>B. Aradi, B. Hourahine, and T. Frauenheim, *J. Phys. Chem. A* **111**, 5678 (2007).
- <sup>26</sup>S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, *J. Comput. Chem.* **13**, 1011 (1992).
- <sup>27</sup>E. M. Boczko and C. L. Brooks, *J. Phys. Chem.* **97**, 4509 (1993).
- <sup>28</sup>B. Roux, *Comput. Phys. Commun.* **91**, 275 (1995).
- <sup>29</sup>P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).
- <sup>30</sup>W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- <sup>31</sup>G. F. Koster and J. C. Slater, *Phys. Rev.* **95**, 1167 (1954).
- <sup>32</sup>G. Zheng, H. A. Witek, P. Bobadova-Parvanova, S. Irle, D. G. Musaev, R. Prabhakar, K. Morokuma, M. Lundberg, M. Elstner, C. Köhler, and T. Frauenheim, *J. Chem. Theory Comput.* **3**, 1349 (2007).
- <sup>33</sup>M. Weinert and J. W. Davenport, *Phys. Rev. B* **45**, 13709 (1992).
- <sup>34</sup>G. J. Martyna, M. L. Klein, and M. Tuckerman, *J. Chem. Phys.* **97**, 2635 (1992).
- <sup>35</sup>K. Eichkorn, F. Weigend, O. Treutler, and R. Ahlrichs, *Theor. Chem. Acc.: Theor. Comput. Model. (Theor. Chim. Acta)* **97**, 119 (1997).
- <sup>36</sup>J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, *Phys. Rev. Lett.* **91**, 146401 (2003).
- <sup>37</sup>K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs, *Chem. Phys. Lett.* **242**, 652 (1995).