THE JOURNAL OF PHYSICAL CHEMISTRY B
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCB

Article

Iterative Cup Overlapping: An Efficient Identification Algorithm for Cage Structures of Amorphous Phase Hydrates

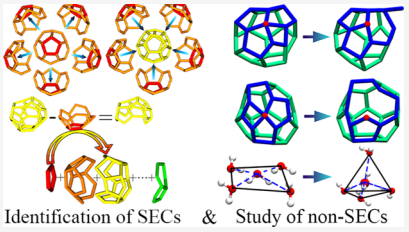
Yongchao Hao, Zhe Xu, Shuai Du, Xuefeng Yang, Tingji Ding, Bowen Wang, Jiafang Xu,* Jun Zhang, and Haiqing Yin

Cite This: *J. Phys. Chem. B* 2021, 125, 1282–1292

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Molecular dynamics studies have revealed that the nucleation pathway of clathrate hydrates involves the evolution from amorphous to crystalline hydrates. In this study, complete cages are further classified into the standard edge-saturated cages (SECs) and nonstandard edge-saturated cages (non-SECs). Centered on studying the structure and evolution of non-SECs and SECs, we propose a novel and efficient algorithm, iterative cup overlapping (ICO), to monitor hydrate nucleation and growth in molecular simulations by identifying SECs and discuss possible causes of the instability of non-SECs. Manipulation of topological information makes it possible for ICO to avoid the repeated searches for identified cages and deduce all SECs with low time costs, improving the efficiency of identification significantly. The accuracy and efficiency of ICO were verified by comparing the identification results with other well-proven algorithms. Furthermore, it was found that non-SECs have short lifetimes and eventually decompose or reorganize into more stable structures. Some evidence suggests that the instability of non-SECs is closely related to the hydrogen-bonding configuration of water-ring aggregations that they contain. The spontaneous evolution of the hydrogen-bonding network into the tetrahedral network may be the main factor that causes the conversion of QWRAs and the evolution of non-SECs.



INTRODUCTION

Gas hydrates are crystalline compounds consisting of polyhedral cage structures of water molecules enveloping guest molecules (e.g., methane, carbon dioxide, etc.). Natural gas hydrates have been regarded as an important natural energy source and a potential future fuel.[1] The energy contained in worldwide gas hydrate accumulations is conservatively estimated to be twice that of currently recoverable fossil fuels.[2] On the other hand, the blockage caused by hydrate formation is one major issue of pipeline flow assurance, posing a serious threat to the safety of oil and gas development.[3] Furthermore, clathrate hydrates are also involved in a variety of scientific fields, such as carbon dioxide sequestration,[4−7] hydrogen and natural gas storage,[8,9] separation of gas mixtures,[10] desalination of seawater,[5,11] and so forth. Although the equilibrium thermodynamic[12−14] and structural properties[15,16] of gas hydrates have been well-understood, the development of relevant scientific research and industrial applications still urgently require further breakthroughs in the mechanism of gas hydrate formation and dissociation. Microlevel understanding can provide a theoretical foundation for the development of technologies for promoting and inhibiting hydrate formation, which is conducive to exploit the value and potential of gas hydrates in industrial fields.

Direct observation of hydrate nucleation and growth by the experiment was challenging, and studies based on molecular simulation provided most evidence for the establishment of microscopic mechanisms.[17] Walsh et al.[18] achieved the first microsecond simulations of spontaneous hydrate nucleation, a pioneering work that fully demonstrated the feasibility of simulating hydrate nucleation, inspiring a series of molecular dynamics studies of hydrate nucleation and growth. Several different hypotheses for hydrate nucleation mechanisms had been proposed, including the labile cluster hypothesis,[19,20] the local structure hypothesis,[21] the cage adsorption hypothesis,[22] and the blob hypothesis.[23] The latter three hypotheses were developed from analyzing hydrate nucleation and growth trajectories of molecular dynamics simulations. Guo et al. proposed the formation of the amorphous intermediate of hydrates and the possibility of the structural transition from the amorphous hydrate phase to the crystal hydrate phase in the cage adsorption hypothesis. After that, Jacobson et al. suggested that the formation of the amorphous intermediate from blobs (clusters of solvent-separated guests) is the first step of the hydrate nucleation pathway.[24] The water molecules surrounding the solvent-separated guests rearrange to form polyhedral cages, which, if the size and lifetime of the cluster are sufficient, results in an amorphous hydrate. The second

Received: October 2, 2020
Revised: January 13, 2021
Published: January 22, 2021

https://dx.doi.org/10.1021/acs.jpcb.0c08964
J. Phys. Chem. B 2021, 125, 1282–1292

step of the hydrate nucleation pathway stated by the blob hypothesis is the evolution from amorphous hydrates to crystalline hydrates, which has been observed in many studies based on molecular dynamics simulation.[24−26] The multistep crystallization mechanism for proteins and colloids is very similar to that of hydrates.[27−29] Moreover, the previous research by Zhang et al.[30] suggests that there may exist multiple pathways for nucleation, including metastable intermediates and the direct formation of the globally stable phase.

The advancement of hydrate formation and dissociation in molecular simulations are commonly characterized by order parameters,[31−33] a quantitative measure of the degree of order in the system. However, without the aid of cage identification, existing order parameters cannot distinguish amorphous from crystalline hydrates or distinguish regions with different clathrate crystal structures.[34] The two most common crystalline hydrate structures[2,17] are structure I containing 5^{12} (pentagonal dodecahedron) and $5^{12}6^2$ (hexagonal truncated trapezohedron) cages and structure II containing 5^{12} and $5^{12}6^4$ (hexadecahedron) cages. Amorphous hydrates are composed of other types of polyhedral cages that do not exist in crystalline hydrates. For the study of hydrate nucleation using molecular simulations, sI, sII, and amorphous hydrates can be distinguished by cage types, and the process of hydrate formation and dissociation can be quantified by the evolution of the number of cages. Therefore, the identification methods of cages are powerful tools for further studying and understanding the microscopic mechanism of hydrate nucleation.

Molinero et al.[35] developed the CHILL+ algorithm which uses the number of staggered and eclipsed water−water bonds to identify water molecules in ice and clathrate hydrate. Jacobson et al.[36] developed a method to identify 5^{12}, $5^{12}6^2$, and $5^{12}6^4$ cages, using the connectivity of water molecules and the topology of the rings. Walsh et al.[37] implemented a unit-cell searching algorithm similar to the method proposed by Jacobson et al. to identify cages, which can identify seven types of cages. Mahmoudinobar and Dias[38] developed an open-source software package called GRADE to identify 5^{12}, $5^{12}6^2$, and $5^{12}6^4$ cages with the same hierarchical approach as Jacobson et al. In addition, GRADE filters water rings by nonplanarity and nonconvexity to speed up the calculation. All the methods mentioned above are able to rapidly identify the main types of cages formed during hydrate nucleation while omitting those cage types minor in number, causing a loss of potentially valuable information. Guo et al.[39,40] developed a method called the face-saturated incomplete cage analysis (FSICA) and identified 1258 types of complete cages and 7015 types of face-saturated incomplete cages (FSICs) from the nucleation trajectories reported by Walsh et al. FSICA can, in principle, identify all complete and face-saturated incomplete cages as long as the diameters of these cages fall between the spacing between neighboring grid points and the cutoff diameters of the search range. Due to the specificity of its identification mechanism, FSICA is highly accurate but also highly time-consuming compared to other methods due to its identification mechanism. In this paper, a new identification algorithm for the cage structures of amorphous phase hydrates, iterative cup overlapping (ICO), is proposed.

In the crystalline hydrates, each water molecule is hydrogen-bonded to four other water molecules, forming a tetrahedrally coordinated structure,[34] and the vertexes of each cage are shared by exactly three edges of that cage. Therefore, we further classified the complete cage defined by Guo et al. into two classes based on the structural characteristics: the standard edge-saturated cages (SECs) in which each vertex is shared by exactly three edges and the nonstandard edge-saturated cages (non-SECs) in which at least one vertex is shared by more than three edges. SECs are structurally similar to the cages in the crystalline hydrates, while non-SECs are structurally different from that. In this work, we focus on the study of the structure and evolution of non-SECs and SECs, propose an algorithm to identify SECs, and discuss possible causes of the instability of non-SECs. The validity, accuracy, and efficiency of ICO were verified by comparing the identification results of ICO, FSICA, and GRADE. The results show that ICO is able to identify all SECs with very low time cost and has the highest estimation efficiency among the three algorithms. Moreover, the results also show that non-SECs are unstable structures that exist for short lifetimes and eventually decompose or reorganize into other structures. Since the most significant difference between the structures of non-SECs and SECs is that SECs only contain ternary water-ring aggregations and non-SECs contain quaternary water-ring aggregations, there must exist the difference of the stability between SECs and non-SECs. In this study, we explore the quaternary water-ring aggregations in non-SECs and their association with the short lifetime of non-SECs. Meanwhile, the evolution of SECs can be applied to monitor the formation and dissociation of hydrates and the approximate composition of cages. The algorithm we proposed in this work provides a new idea for the efficient identification of cage structures of the clathrate hydrates.

METHODS

Our identification algorithm was inspired by the previous work of Jacobson et al. They proposed a method for identifying the cages of sI and sII hydrates by overlapping specific cups based on specific "merging rules".[36] For example, each 5^{12} cage consists of twelve partial overlapped cups made up of pentagons.

However, thousands of types of cages may occur in amorphous hydrates, tracking only three types of cages is not enough to observe the pathway of hydrate nucleation. On the other hand, it is difficult to establish overlapping rules for each type of cages that may occur.

By studying the structural characteristics of cups and cages, we found that crucial topological information in cups can be used to automate the process of overlapping cups. Based on this finding, we developed a novel algorithm, ICO, for the identification of cages.

Identification of Rings. In this study, water rings are simplified to undirected topological graphs, with the vertexes representing water molecules and edges representing hydrogen bonds connecting water molecules. Two water molecules are regarded as connected by a hydrogen bond[41] when their r_{O-O} (the distance between oxygen atoms) is no more than 3.5 Å and $\angle\alpha$ (the angle between the O−O vector and O−H bond) is no more than 30°, as shown in Figure 1.

A path is an alternating sequence of vertexes and edges, where each vertex and edge occurs only exactly once.[42] If a path starts and ends at the same vertex, it is a ring. The most basic way to search for all the rings in the system is to traverse all possible paths in the depth-first order.[38] Since the main objects to be identified are the common four-membered, five-

https://dx.doi.org/10.1021/acs.jpcb.0c08964
J. Phys. Chem. B 2021, 125, 1282−1292

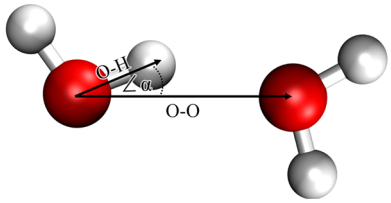


Figure 1. Geometric definition of the hydrogen bond.

membered, and six-membered rings in the cage, the maximum traversal depth is set to seven.

Rings are represented as the sets of vertexes, as shown in eq 1

$$ring_i = \{vertex_1, vertex_2, ..., vertex_n\} \quad (1)$$

where $ring_i$ is an n-membered ring represented as a set of n vertexes and i is the serial number used to distinguish the rings.

Each n-membered ring in the system will be repeatedly searched $2n$ times.[43] If two rings contain the same vertexes, it is considered the same structure. Any two rings sharing one edge are called neighboring rings. After all the rings in the system are identified, a database containing the topology of rings is established.

Identification of Cups. Jacobson et al. defined cups as half-cages formed by sets of five-membered and six-membered rings sharing some of their edges.[36] In this study, we redefined the structure of cups with more details.

Cups are half-cage structures with an n-element ring as the bottom and n other water rings as the lateral surface, as shown in Figure 2. To form a complete cup, an n-membered ring as

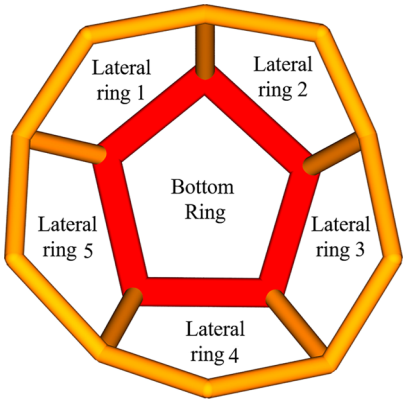


Figure 2. Schematic view of a cup, the red ring is the bottom of the cup, and the orange rings are lateral rings composing the lateral surface of the cup. Any two neighboring lateral ring shares one edge, forming a belt-like closed structure: 1−2−3−4−5−1.

the bottom ring of the cup and its neighboring rings should meet the following three conditions (A), (B), and (C).

(A) The n-membered ring should have at least n neighboring rings to form a cup;
(B) The lateral rings of the cup should form a belt-like closed structure; and
(C) The lateral rings should distribute on the same side of the bottom ring.

Condition A: as can be seen in Figure 2, the number of vertexes contained in the bottom ring is equal to the number of lateral rings, so an n-membered ring should have at least n neighboring rings to form a cup. Depending on condition (A), ICO is able to find all possible bottom rings from the database of rings.

Condition B: we define the belt-like structure as the head-to-tail structure where any two neighboring rings should share exactly one edge, as shown in Figure 2. Figure 3 shows the

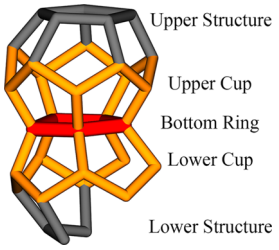


Figure 3. Pair of twin cups. The red ring is the bottom ring, and the orange rings are lateral rings. The bottom ring of a pair of twin cups is usually the contact surface of two structures/cages.

configuration of a pair of twin cups where two cups share the same bottom ring. Twin cups are the result of two structures being in contact. The five-membered ring in Figure 3 has more than five neighboring rings, satisfying condition (A) and forming two belt-like structures. Each belt forms a cup with the bottom ring. According to condition (B), ICO finds all belts from neighboring rings of each possible bottom ring. However, not every belt forms cup with its bottom ring. As can be seen from Figure 4, the n rings distributed on different sides of the n-membered bottom ring can also form a closed structure satisfying the definition of the belt.

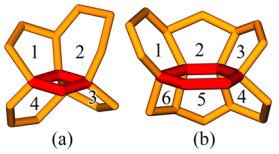


Figure 4. Two examples of pseudo-cups. The red rings are bottom rings, and the orange rings are lateral rings. The closed structure in pseudo-cup (a): 1−2−3−4−1. The closed structure in pseudo-cup (b): 1−2−3−4−5−6−1.

Condition C: Figure 4 gives two examples of pseudo-cups that are not complete cups but still meets the first two conditions. Condition C can avoid the misidentification caused by pseudo-cups. Note that the lateral rings on the same side of the bottom ring will not share the same edge of the bottom ring. For reference, we realize condition C in ICO by requiring that each edge of a cup's bottom ring should be shared with exactly one lateral ring.

By detecting the linking relationships among each ring and its neighboring rings, all the structures meeting the conditions A, B, and C can be found and then a database composed of information about each cup and the rings it contains is established. Since each cup has only one bottom ring, the bottom ring is used to distinguish the cups. Given a bottom ring, up to two corresponding cups can be retrieved from the cup database.

Identification of SECs. The precondition for allowing a cage that can be identified by overlapping cups is that each ring should form a cup with its neighboring rings in the same cage, which is also a vital feature of SECs. Therefore, a $4^a5^b6^c$ cage

https://dx.doi.org/10.1021/acs.jpcb.0c08964
J. Phys. Chem. B 2021, 125, 1282−1292

meeting the standard edge-saturated condition can also be regarded as consisting of $a + b + c$ overlapped cups.

As shown in Figure 5, a $5^{12}6^3$ cage is divided into three belt-like structures. Note that even though layer 3 has only two

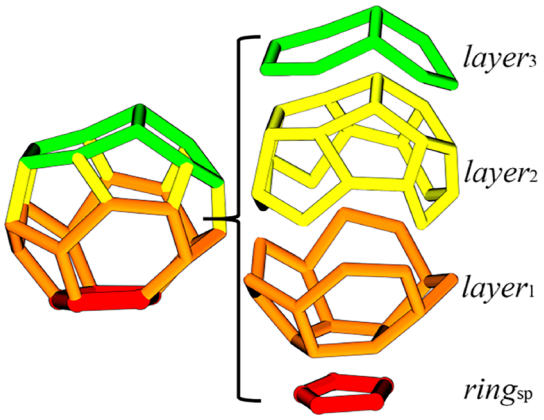


Figure 5. Example of the layered structure of cages.

rings, it still meets the definition of a belt-like structure. The red ring is the starting point for searching the abovementioned $5^{12}6^3$ cage, denoted as $ring_{sp}$. The belt-like structure connecting the starting ring is $layer_1$ (marked orange). The next two layers are $layer_2$ (marked yellow) and $layer_3$ (marked green). The layered structure of SECs is described by eq 2.

$$cage_j = \{ring_{sp}, layer_1, ..., layer_x\} \quad (2)$$

where j is the serial number to distinguish cages and $layer_x$ is a set of rings which form the xth belt-like structure of $cage_j$.

Figure 6 shows the top view of three cups belonging to the $5^{12}6^3$ cage in Figure 5. As can be seen from Figure 6, each cup

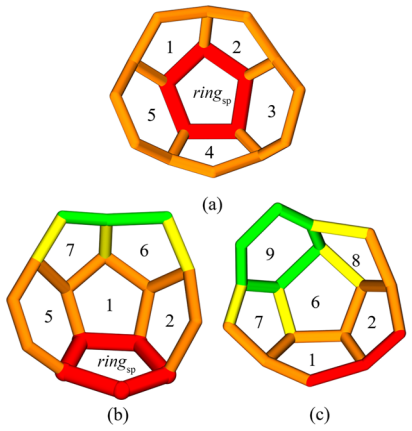


Figure 6. Top view of three cups belonging to the $5^{12}6^3$ cage shown in Figure 5. The numbers in the figure are the serial number of rings. In this figure, $ring_1$, $ring_2$, $ring_3$, $ring_4$, and $ring_5$ are the rings of $layer_1$; $ring_6$, $ring_7$, and $ring_8$ are the rings of $layer_2$; and $ring_9$ is the ring of $layer_3$.

(except the cup with $ring_{sp}$ as the bottom) contains rings belonging to three different layers. For example, Figure 6b shows a cup with $ring_1$ as the bottom ring, which contains $ring_{sp}$ and the rings of $layer_1$ and $layer_2$. The linking relationships among layers contained in the cups are the key to deduce the rings of the next layer based on the rings of the known layers.

ICO deduces all the layered structures with each ring formed in the system as the starting point and further judges whether identified layered structures are cages based on Euler's polyhedron formula. To illustrate how ICO identifies an SEC, the specific identification process is then demonstrated with the $5^{12}6^3$ cage shown in Figure 5 as an example.

The cup with $ring_{sp}$ as the bottom contains all the rings belonging to $layer_1$. Therefore, given $ring_{sp}$, five rings of $layer_1$ are retrieved from the cup database. The union of currently known rings is denoted as U_{0_1}, including $ring_{sp}$ and rings of $layer_1$, as shown in eq 3.

$$U_{0\_1} = \{ring_{sp}, layer_1\} \quad (3)$$

As shown in Figure 7a, if each ring of $layer_1$ as the bottom ring corresponds to a cup, five cups can be retrieved from the cup database. An intermediate structure, denoted as U_{0_2}, is obtained by overlapping the five cups, as shown in Figure 7b.

U_{0_2} contains all the rings of the first two layers and $ring_{sp}$, as shown in eq 4. The relative complement of U_{0_1} in U_{0_2} is the next layer, $layer_2$, as shown in Figure 8.

$$\begin{aligned} U_{0\_2} &= \{cup_1, cup_2, cup_3, cup_4, cup_5\} \\ &= \{ring_{sp}, layer_1, layer_2\} \end{aligned} \quad (4)$$

In fact, if a ring of $layer_1$ is the bottom ring of a pair of twin cups and then it has two corresponding cups in the cup database. Therefore, the intermediate structure obtained by overlapping cups retrieved from the cup database, denoted as U_{0_2}', will also contain cups that are not part of the layered structure currently being searched (hereinafter, referred to as the target structure). U_{0_2}' is described using eq 5

$$U'_{0\_2} = \{cup_1, cup_2^{(1)}, cup_2^{(2)}, cup_3, cup_4, cup_5\} \quad (5)$$

where $cup_2^{(1)}$ and $cup_2^{(2)}$ are a pair of twin cups with $ring_2$ as the bottom ring. U_{0_2}' contains at most the same number of pairs of twin cups as the number of the rings of $layer_1$, and all the cups outside the target structure requires to be eliminated.

Figure 9a shows a pair of twin cups with $ring_2$ as the bottom appeared in the deduction process, where only $cup_2^{(1)}$ (shown in Figure 9b) belongs to the target structure and should be used to deduce rings in the next layer. $cup_2^{(2)}$ (shown in Figure 9c) is outside the target structure and should be eliminated from U_{0_2}'. $cup_2^{(1)}$ shares three rings with $layer_1$ where the bottom ring of twin cups is located, and $cup_2^{(2)}$ only shares its bottom ring with $layer_1$. Therefore, the criterion for determining which cup should be kept is that the cup belonging to the target structure should share exactly three rings with the layer where its bottom ring is located. Comparing each cup of U_{0_2}' with $layer_1$ and eliminating those that do not meet the criterion, a set only containing cups that are part of the cage, U_{0_2}, is obtained.

As shown in Figure 10, the deduction of $layer_{x+1}$ is similar to that of $layer_2$: first, an intermediate structure U_{x-1_x+1}' is obtained by retrieving and overlapping cups based on the rings of known $layer_x$. Then, the cups outside the target structure are eliminated to obtain U_{x-1_x+1}. Finally, all the rings of $layer_{x-1}$ and $layer_x$ are excluded from U_{x-1_x+1} to obtain $layer_{x+1}$. Therefore, the deduction of rings in each layer is an iterative process, which requires the rings in the last two layers as inputs and then outputs the rings in the next layer.

Once a termination condition is triggered, the search process starting from the current $ring_{sp}$ ends, and a layered structure is

https://dx.doi.org/10.1021/acs.jpcb.0c08964
J. Phys. Chem. B 2021, 125, 1282−1292

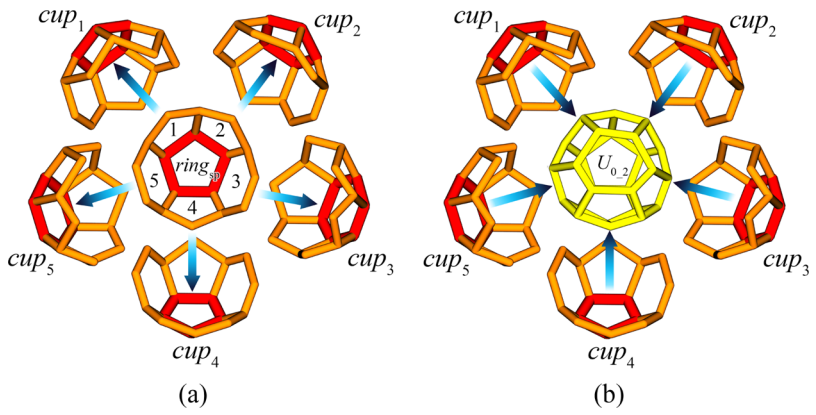


Figure 7. Schematic diagram of retrieving and overlapping cups. (a) Five cups are retrieved from the cup database based on the rings of $layer_1$. (b) U_{0_2} is obtained by overlapping those five cups.

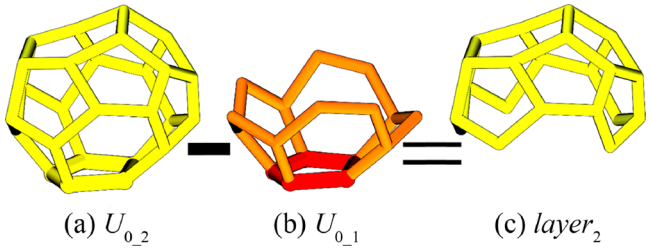


Figure 8. Deduction of the next layer based on the known layers. By removing U_{0_1} from U_{0_2}, the remaining structure (c) is $layer_2$.

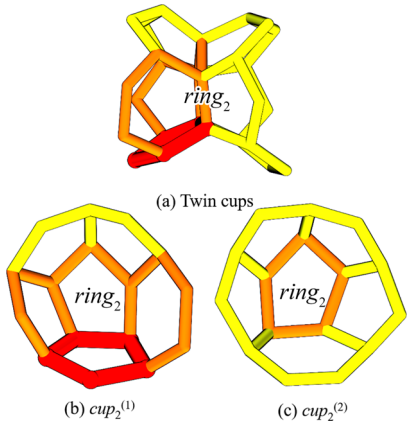


Figure 9. Pair of twin cups with the $ring_2$ of the $5^{12}6^3$ cage (see Figure 5) as the bottom ring.

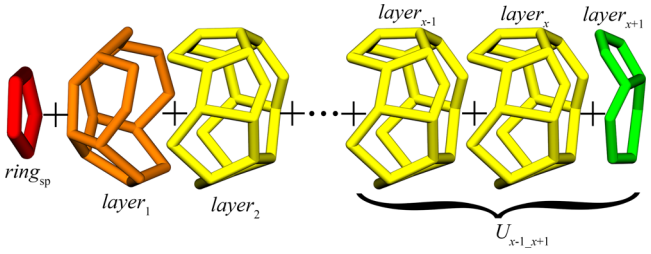


Figure 10. Search for a layered structure starting from the red ring. The last layer of the layered structure is marked green.

found. Based on the triggered termination condition, the searched structures can be classified into incomplete structures and complete cages.

The termination condition for incomplete structures: if a ring with no corresponding cup in the cup database occurred in the deduction process of the target structure, which suggests that the target structure is not an SEC based on the precondition mentioned earlier, the iteration should stop.

The termination condition for complete cages: if the second layer of the target structure has been deduced, then verify after each iteration that the structure conforms to Euler's polyhedron formula (eq 6), where F, V, and E are, respectively, the numbers of faces, vertexes, and edges in the given structure.

$$F + V - E = 2 \tag{6}$$

If the structure conforms to Euler's polyhedron formula, it may be a polyhedron. However, in some cases, unclosed layered structures may also conform to the formula. To ensure the structure is closed, the rings in the last layer should also forms cups with its neighboring rings in the cage.

Characteristics of Non-SECs. Since each vertex in SECs is shared by exactly three edges, three rings sharing a vertex form a ternary water-ring aggregation (TWRA),[44] as shown in Figure 11. On the contrary, non-SECs contain at least one vertex shared by more than four edges and the rings sharing those vertexes form more complex water multiring aggregations, which is the most significant difference between the structures of non-SECs and SECs. Figure 11 gives an example of a non-SEC, where four rings sharing a vertex form a

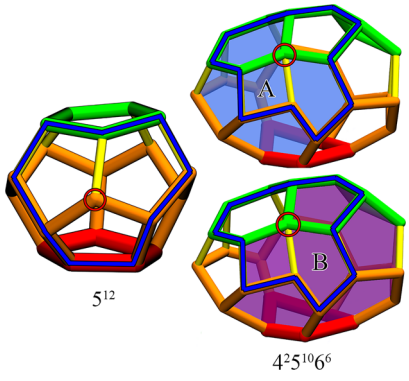


Figure 11. 5^{12} SEC and $4^25^{10}6^6$ non-SEC. The rings filled in blue are the neighboring rings of A, and the rings filled in purple are the neighboring rings of B. The polygons surrounded by blue lines represent the multiring aggregate structures consisting of water rings that share central water molecules which are vertexes in the red circle.

https://dx.doi.org/10.1021/acs.jpcb.0c08964
J. Phys. Chem. B 2021, 125, 1282−1292

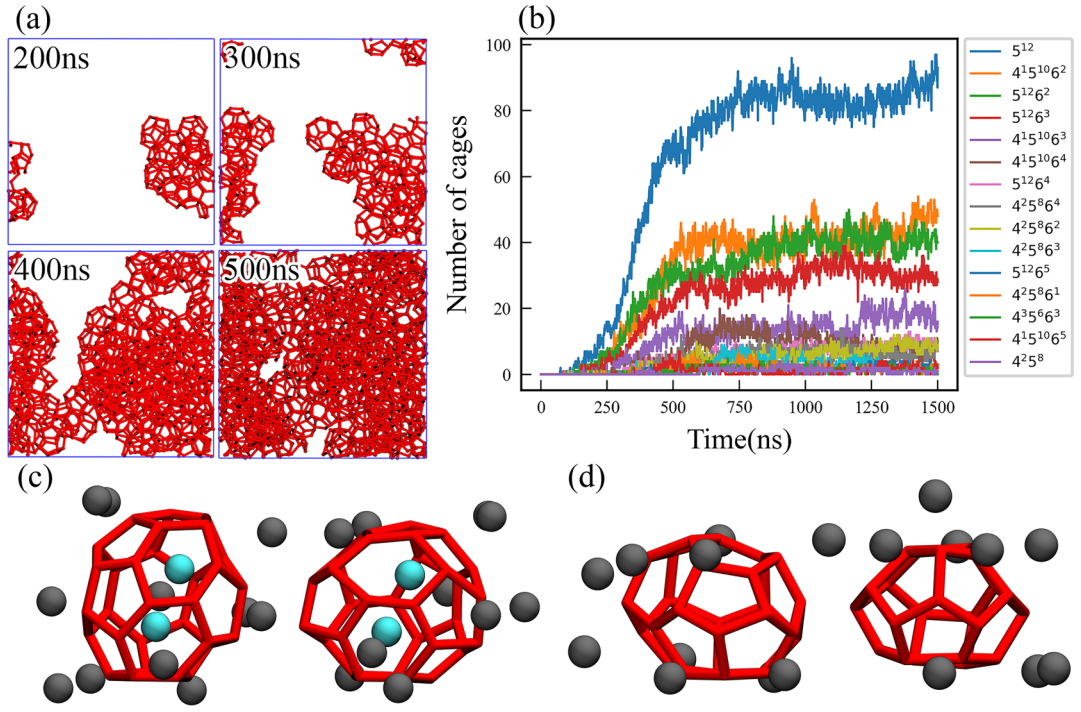


Figure 12. Results of the identification of the cages occurring in the trajectory using ICO. Cyan, gray spheres, and red sticks indicate methane molecules, guest methane molecules, and hydrogen bonds, respectively. (a) Snapshots of cage evolution in the system. (b) Time evolution of 15 most abundant cage types. (c) Doubly occupied cages. (d) Empty cages.

quaternary water-ring aggregation (QWRA). Each $4^2 5^{10} 6^6$ cage has 18 faces, 31 vertexes, and 47 edges, meeting Euler's polyhedron formula. As shown in Figure 11, the vertex (the central water of the QWRA) shared by four edges breaks the continuity of neighboring rings of the four rings in the QWRA. For example, both rings A and B cannot form a complete cup with its neighboring rings. Once the deduction of the algorithm reaches the third layer of the non-SEC, the termination condition for incomplete structures is triggered. Since the structures of non-SECs do not meet the precondition for overlapping cups, ICO cannot identify non-SECs.

In addition, a recent study[44] shows that TWRAs are fundamental structures for characterizing the nucleation pathway, and the lifetime of TWRAs is closely related to their structure. Since SECs only contain TWRAs and non-SECs contain QWRAs or other water multiring aggregations, there must exist the difference of the stability between SECs and non-SECs.

RESULTS AND DISCUSSION

Simulation Details. In this work, a hydrate formation trajectory was analyzed to demonstrate the performance of the ICO algorithm. MD simulation was performed using the LAMMPS package.[45] The TIP4P/Ice model[46] and OPLS-UA model[47] were used for water and methane, respectively. The constructed simulation box is 50.75 × 50.75 × 50.75 Å^3, containing 3487 water molecules and 550 methane molecules, as shown in Figure S1. MD simulation was carried out for 1500 ns at 250 K and 650 bars to study hydrate nucleation.

Identification Results. ICO deduced the coordinates of the cage vertexes for each frame of the trajectory and the obtained new trajectory contained only SECs. Figure 12a includes four snapshots of the trajectory that only contains SECs and shows the evolution of cages in the system. The evolution of the number of 15 most abundant cage types identified by ICO is shown in Figure 12b. Since ICO requires only the coordinates of the water molecules and not that of the guest molecules to identify the cages, double-occupied cages and empty cages can also be identified, as shown in Figure 12c,d.

To verify the validity of the identification algorithm, we analyzed the trajectory of the simulation using ICO, FSICA, and GRADE and evaluated ICO's performance by comparing the time cost and accuracy of the three algorithms. The FSICA program used in this study was written in Perl based on information reported in a previous study[39] by Guo et al. The performance of FSICA is dependent on the spacing between neighboring grid points and the cutoff radius for searching the water molecules surrounding hunting sites. GRADE[38] used in this study is an open-source software written in C++, which can speed up the identification of cages by ignoring deformed rings. Therefore, the performance of GRADE also depends on the setting of filter conditions θ, δ_1, and δ_2. In addition, the ICO algorithm was implemented in Python. Multiple sets of parameters were set for FSICA and GRADE, one of which is default values reported in refs 36, 37, and the others were set to test the performance of the algorithms further.

Due to the specificity of the identification method, FSICA can identify all cages (including non-SECs, SECs, and FSICs) with a diameter smaller than the cutoff diameter and greater than the resolution of grid points. The accuracy of the ICO was evaluated by taking the identification results of FSICA as a benchmark. Table 1 reports the identification results of three algorithms for the MD trajectory. The results show that ICO has an excellent accuracy in identifying SECs that include cages in crystalline clathrates. However, ICO cannot identify non-SECs due to the mechanism of the identification. At present, FSICA is the only method able to identify non-SECs. From

https://dx.doi.org/10.1021/acs.jpcb.0c08964
J. Phys. Chem. B 2021, 125, 1282−1292

Table 1. Identification Results of Three Algorithms for the Last 50 Frames of MD Trajectory[a]

algorithms	average time costs (s)	cages in sI, sII	SECs	non-SECs
ICO	691	6974	13,695	0
FSICA[b]	376,254	6974	13,695	73
FSICA[c]	35,397	6974	13,663	73
FSICA[d]	5918	6965	13,583	69
GRADE[e]	424	4233	4233	0
GRADE[f]	449	4410	4410	0

[a]Each test was run on a single thread. [b]The cutoff radius was set at 6 Å. [c]The cutoff radius was set at 12 Å, and the spacing between neighboring grid points was set at 1 Å (default). [d]The spacing between neighboring grid points was set at 2 Å, and the cutoff radius was set at 6 Å. [e]δ_1 and δ_2 were set at 1.8 and 2.6 Å, respectively (default). [f]δ_1 and δ_2 were set at 3 Å (maximum value shown in ref 36).

Table 1, it can be seen that when the cutoff radius decreases or the spacing of grid points increases, the speed of FSICA will rapidly improve while the accuracy of FSICA will slightly decrease, resulting in a significant improvement of identification efficiency. Therefore, optimization of parameter selection may allow FSICA to find a balance between accuracy and speed. To achieve the highest accuracy among the three algorithms and to observe any potentially valuable information, FSICA needs to process intensive calculations with high time costs, which is more suitable for studying the hydrate nucleation mechanism. The results also show that GRADE has remarkably low time costs for identifying cages due to filtering deformed rings. However, the rings of the cages that form during hydrate nucleation have less planarity and nonconvexity than those that form during hydrate growth, which requires increasing values of the three filter conditions in GRADE beyond their default values for those distorted cages to be identified. Considering the performance of GRADE, it should be suitable for monitoring hydrate growth. Compared with GRADE and FSICA, we evaluate the efficiency of ICO as "middle to high" due to the ability of ICO to identify most types of cages with middle to high accuracy and relatively low time costs. ICO can be applied as an efficient identification method for monitoring hydrate nucleation. To demonstrate each algorithm's performance comprehensively, we also tested two segments of the MD trajectory in the early (151−200 ns) and middle (351−400 ns) stages of hydrate nucleation, as shown in Table S1. Briefly, these three algorithms should be applied to different scientific questions.

Significant Advantage. Manipulation of topological information to deduce cages makes it possible for ICO to

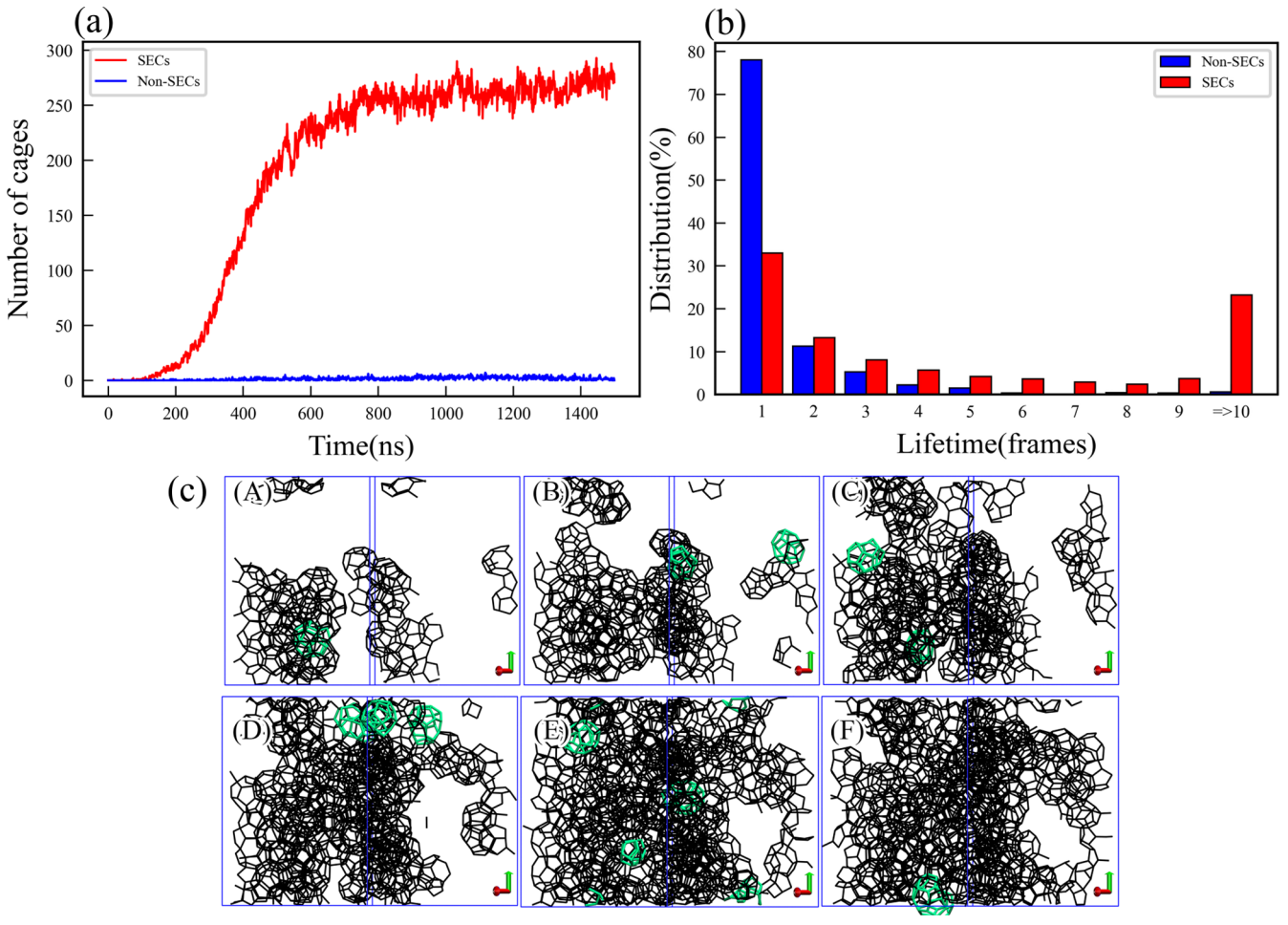


Figure 13. (a) Time evolution of number of non-SECs and SECs in the system. (b) Lifetime distribution of non-SECs and SECs. (c) Snapshots of non-SECs and SECs in the system: (A) 300, (B) 350, (C) 400, (D) 450, (E) 500, and (F) 550 ns. Green and black structures represent non-SECs and SECs, respectively. Blue lines are the boundary of the simulation box. Due to periodic boundary conditions employed, the seemingly incomplete structures shown in the snapshots are in fact complete cages.

https://dx.doi.org/10.1021/acs.jpcb.0c08964
J. Phys. Chem. B 2021, 125, 1282−1292

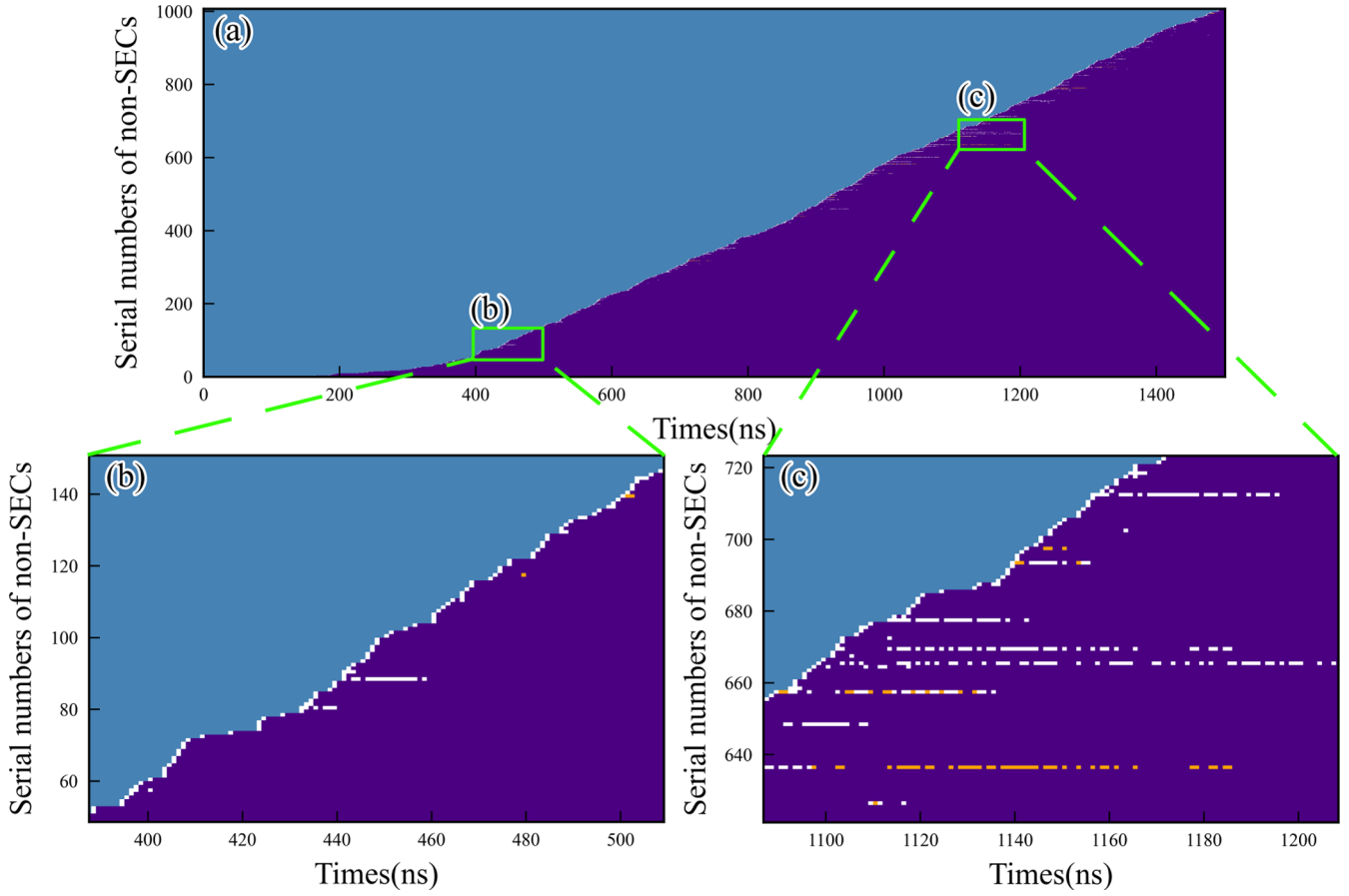


Figure 14. Time evolution of each non-SEC. Each row of pixels represents the structure evolution of a non-SEC. For each non-SEC, blue represents that the cage has not formed, white represents that the cage has formed, purple means that cage disappears with the conversion of QWRAs, and orange means that cage disappears, but QWRAs remain original structures.

avoid the repeated searches for cages that have already been identified, which is the most significant advantage of ICO and makes our algorithm more efficient and rapid.

Since cages are undirected graphs, searches starting from different starting points may find the same cage. Although repeated searches do not affect the accuracy of the results, they reduce the efficiency of identification. For example, by searching for all water molecules within 6 Å surrounding hunting sites, the FSICA can identify cages from the local hydrogen-bonding network formed by these water molecules. The hunting sites of FSICA include dissolved guest molecules and spatial grids in the simulation system. To avoid missing any of the cages that appear in the simulation system, the spacing of the grid points needs to be smaller than the minimum diameter of cages.[39] Therefore, each cage in the system will contain multiple hunting sites and inevitably be repeatedly searched multiple times according to the identification principle of the FSICA.

In the case of cage identification by the ICO, the starting point of the searching for a cage can be any ring contained in that cage, which indicates that a $4^a5^b6^c$ cage will be repeatedly identified $a + b + c$ times. Therefore, the algorithm also has the problem of repeated searches, but it can be avoided. ICO can determine whether a layered structure is an identified cage at the beginning of the deduction.

The contact surface of two convex polyhedral cages in hydrates should be a ring rather than a three-dimensional cup structure since two convex polyhedrons' contact surface is a plane. If the contact area of two cages is a cup, their configuration should be the nested structure in which the larger cage encloses the smaller cage, and both share the same cup, as shown in Figure S2. However, this nested structure does not conform to the tetrahedrally coordinated structure of water molecules in the hydrogen-bonding network, so it is impossible to exist in clathrate hydrates. Therefore, the relationship between cups and cages should be the many-to-one correspondence, where each cup can only belong to exactly one cage. To avoid repeated searching of identified cages, after a cage has been identified, all the cups it contains should be removed from the cup database. Therefore, ICO will no longer search for identified cages again.

Study on Instability of Non-SECs. By comparing the identification results of ICO and FSICA, 1007 non-SECs were found in the trajectory. Figure 14a shows the numbers of SECs identified by ICO and non-SECs identified by FSICA over time. Figure 14a indicates that most cages meet the standard edge-saturation condition, and therefore, the evolution of SECs is sufficient to monitor the hydrate formation and dissociation. Compared with SECs, the number of non-SECs is low and has no significant change over time. Besides, the crystallization of clathrate hydrates is the process of converting other cages into 5^{12}, $5^{12}6^2$, and $5^{12}6^4$ cages, so the only way for non-SECs to participate in crystallization is to convert to SECs. Since the calculation of crystallinity OP[34] only requires the vertexes of 5^{12}, $5^{12}6^2$, and $5^{12}6^4$ cages, the advance of crystallization of the system can be quantified by SECs.

https://dx.doi.org/10.1021/acs.jpcb.0c08964
J. Phys. Chem. B 2021, 125, 1282−1292

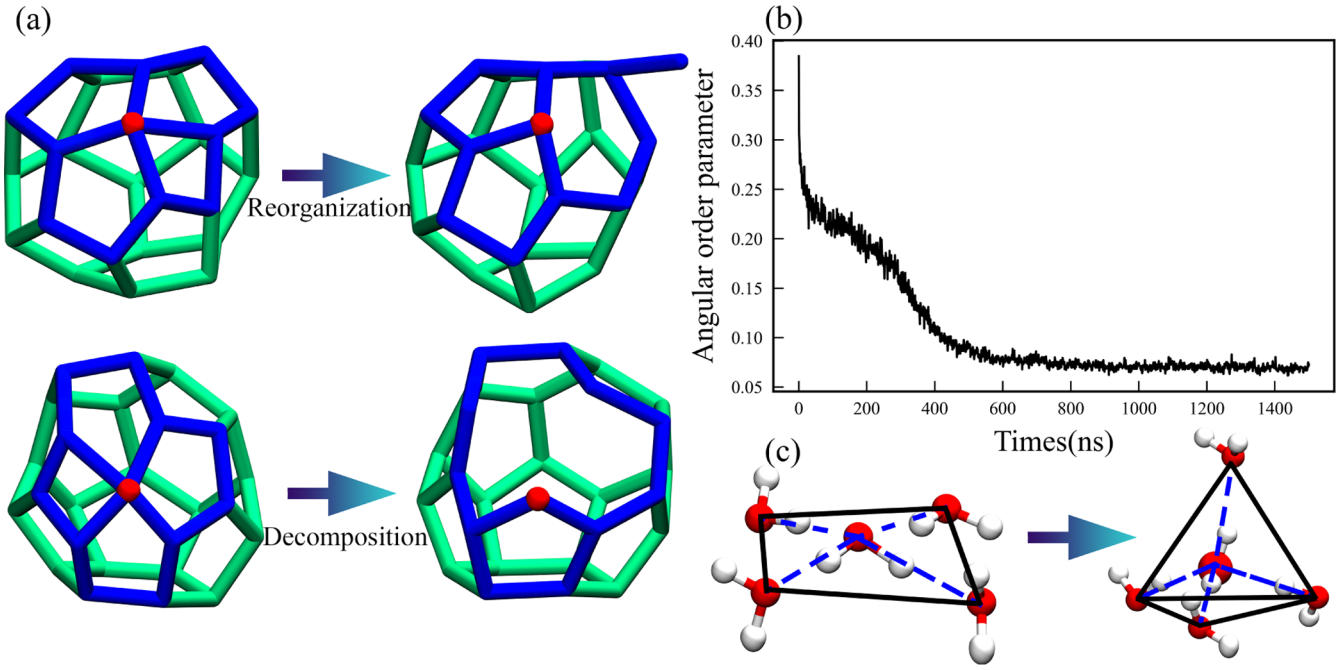


Figure 15. (a) Examples of conversion of QWRAs and non-SECs due to hydrogen bond-breaking of central water molecules. Red spheres represent central water molecules of QWRAs. Blue and green sticks represent the edges (hydrogen bonds) of QWRAs and TWRAs, respectively. (b) Time evolution of the average angular order parameter of water molecules in the system. (c) Hydrogen-bonding configuration of central water. Blue dashed lines represent hydrogen bonds.

The lifetime distribution is calculated by counting the number of frames that each cage occurs, as shown in Figure 13b. It can be seen from the figure that about 30% of the SECs only occur in one frame and about 25% of the SECs can persist more than ten frames. A recent study[44] suggests that TWRAs will evolve toward specific stable structures during hydrate nucleation. As mentioned earlier, SECs are composed of TWRAs. Thus, SECs containing unstable TWRAs will spontaneously convert to more stable SECs, which explain why about 30% of the SECs are unstable. As shown in Figure 13c, the snapshots of SECs and non-SECs in the system show that the number and location of non-SECs are different in each snapshot, while the SEC nuclei are growing and expanding. Furthermore, the lifetime distribution of non-SECs indicates that most non-SECs only occur in one frame. Since the interval between the frames of trajectory is 1 ns, the lifetime of most non-SECs is less than 1 ns, which indicates that non-SECs are unstable.

The results of the identification of water multiring aggregations show that only QWRAs and TWRAs exist in the system. The identification method for water multiring aggregations is shown in the Supporting Information. Therefore, the most significant difference between the structures of non-SECs and SECs is that non-SECs contain QWRAs. The main factor contributing to the short lifetime of non-SECs should be the conversion of the structure of QWRAs. To verify this assumption, we traced all non-SECs and QWRAs they contain. Figure 14a shows the evolution of the 1007 non-SECs found in the trajectory, where non-SECs are numbered in the order in which they appear in the trajectory. It is clear that the white lines of the non-SECs in Figure 14b are very short and close to one pixel. Since white pixels represent cages' presence, those short lines suggest that the non-SECs in the early stage of nucleation convert into other structures immediately after they occur, while the non-SECs in Figure 14c have longer but discontinuous white lines interrupted by the purple or orange pixels, which indicate that the non-SECs in the later stage of nucleation have longer lifetimes and their structures undergo several changes caused by the conversion of QWRAs or TWRAs. These non-SECs correspond to the few non-SECs in Figure 13b persisting for more than 10 ns. When a non-SEC disappears, we use purple pixels to indicate that its QWRAs' structures undergo conversions and orange pixels to indicate that its QWRAs maintain their original structures. As can be seen in the figure, only in some rare events, the white lines end in orange pixels, and almost all of them end in purple pixels, indicating that the conversion of QWRAs is closely related to the disappearance of the vast majority of non-SECs. Besides, the full-size Figure 14a is shown in Figure S3.

This evidence suggests that non-SECs exist for a short lifetime after occurence and eventually decompose or reorganize into more stable structures, which is the most likely explanation for the continuously low number of non-SECs, as shown in Figure 13a.

Figure 15a shows two examples of conversion of QWRAs due to hydrogen bond-breaking of central water molecules, which were observed by tracing the evolution of non-SECs. The QWRA was reorganized into a TWRA with one hydrogen bonds of the central water breaking, which may be the main reason for the evolution of non-SECs into SECs. As two hydrogen bonds of the central water breaks, the QWRA cracks and plays a role of decomposition of the non-SEC. It was found that each central water molecule and its four neighboring water molecules distribute in a pyramidal configuration, which is different from the tetrahedrally coordinated structure in ice and hydrates.[31,34,43] Generally, this pyramidal configuration is not stable. The deviation of the hydrogen-bonding configuration of a water molecule from the perfect tetrahedral structure can be quantified by the angular order parameter (AOP).[31] The average AOP of water in the

https://dx.doi.org/10.1021/acs.jpcb.0c08964
J. Phys. Chem. B 2021, 125, 1282−1292

system shows that the hydrogen-bonding network in the system spontaneously develops into the tetrahedral network under the driving force, as shown in Figure 15b. Since the hydrogen-bonding configuration of QWRAs' central water molecules is pyramidal, the tendency for hydrogen-bonding configuration to convert into tetrahedral may drive the breaking of hydrogen bonds of central water, explaining the essence of the conversion of QWRAs, as shown in 15c.

CONCLUSIONS

In this work, we present a novel identification method for the cage structures of clathrate hydrates, which can identify all the SECs in the system to monitor and analyze the hydrate nucleation and growth in molecular dynamics simulation. ICO ingeniously uses topological information to identify SECs and avoids repeated searches for the identified cages, significantly improving identification efficiency. By comparing with other proven algorithms, the accuracy and efficiency of ICO were verified. ICO can be implemented in any programming language to meet the research needs for monitoring hydrate nucleation and growth. ICO provides a new approach to identify polyhedrons from complex networks, with a potential value in other scientific areas. Furthermore, the results show that non-SECs exist for short lifetimes and eventually decompose or reorganize into more stable structures. Since the hydrogen-bonding configuration of QWRAs' central water molecules is pyramidal, the spontaneous evolution of the hydrogen-bonding network into the tetrahedral network may be the main factor that causes the conversion of QWRAs in non-SECs.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.0c08964.

Identification of water multiring aggregations (PDF)

AUTHOR INFORMATION

Corresponding Author

Jiafang Xu − *School of Petroleum Engineering, China University of Petroleum (East China), Qingdao 266580, P. R. China; Key Laboratory of Unconventional Oil & Gas Development (China University of Petroleum (East China)), Ministry of Education, Qingdao 266580, P. R. China*; Email: xjiafang@upc.edu.cn

Authors

Yongchao Hao − *School of Petroleum Engineering, China University of Petroleum (East China), Qingdao 266580, P. R. China*; orcid.org/0000-0001-6655-5795

Zhe Xu − *School of Petroleum Engineering, China University of Petroleum (East China), Qingdao 266580, P. R. China*

Shuai Du − *School of Petroleum Engineering, China University of Petroleum (East China), Qingdao 266580, P. R. China*; orcid.org/0000-0002-8941-7130

Xuefeng Yang − *School of Petroleum Engineering, China University of Petroleum (East China), Qingdao 266580, P. R. China*

Tingji Ding − *School of Petroleum Engineering, China University of Petroleum (East China), Qingdao 266580, P. R. China*; orcid.org/0000-0002-4261-5703

Bowen Wang − *School of Petroleum Engineering, China University of Petroleum (East China), Qingdao 266580, P. R. China*

Jun Zhang − *School of Material Science & Engineering, China University of Petroleum (East China), Qingdao 266580, P. R. China*; orcid.org/0000-0001-7786-4825

Haiqing Yin − *School of Science, China University of Petroleum (East China), Qingdao 266580, P. R. China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.0c08964

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (U1762212, 51874343, and 51874332), "PCSIRT" (IRT_14R58), and Shandong Provincial Key Laboratory of Oilfield Chemistry.

REFERENCES

(1) Macdonald, G. J. The Future of Methane as an Energy Resource. *Annu. Rev. Energy* **1990**, *15*, 53.

(2) Sloan, E. D., Jr; Koh, C. A.; Koh, C. *Clathrate Hydrates of Natural Gases*; CRC Press, 2007.

(3) Sum, A. K.; Koh, C. A.; Sloan, E. D. Clathrate Hydrates: From Laboratory Science to Engineering Practice. *Ind. Eng. Chem. Res.* **2009**, *48*, 7457−7465.

(4) Park, Y.; Kim, D.-Y.; Lee, J.-W.; Huh, D.-G.; Park, K.-P.; Lee, J.; Lee, H. Sequestering Carbon Dioxide into Complex Structures of Naturally Occurring Gas Hydrates. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 12690−12694.

(5) Lee, J.; Kim, K.-S.; Seo, Y. Thermodynamic, structural, and kinetic studies of cyclopentane + CO2 hydrates: Applications for desalination and CO2 capture. *Chem. Eng. J.* **2019**, *375*, 121974.

(6) Dashti, H.; Thomas, D.; Amiri, A. Modeling of Hydrate-Based CO2 Capture with Nucleation Stage and Induction Time Prediction Capability. *J. Clean. Prod.* **2019**, *231*, 805−816.

(7) Kim, S. M.; Lee, J. D.; Lee, H. J.; Lee, E. K.; Kim, Y. Gas Hydrate Formation Method to Capture the Carbon Dioxide for Pre-Combustion Process in IGCC Plant. *Int. J. Hydrogen Energy* **2011**, *36*, 1115−1121.

(8) Mao, W. L.; Mao, H. k.; Goncharov, A. F.; Struzhkin, V. V.; Guo, Q.; Hu, J.; Hu, J.; Hemley, R. J.; Somayazulu, M.; Zhao, Y. Hydrogen Clusters in Clathrate Hydrate. *Science* **2002**, *297*, 2247−2249.

(9) Sloan, E. D. Fundamental Principles and Applications of Natural Gas Hydrates. *Nature* **2003**, *426*, 353−359.

(10) Vorotyntsev, V. M.; Malyshev, V. M.; Taraburov, P. G.; Mochalov, G. M. Separation of Gas Mixtures by Continuous Gas Hydrate Crystallization. *Theor. Found. Chem. Eng.* **2001**, *35*, 513−515.

(11) Han, S.; Rhee, Y.-W.; Kang, S.-P. Investigation of Salt Removal Using Cyclopentane Hydrate Formation and Washing Treatment for Seawater Desalination. *Desalination* **2017**, *404*, 132−137.

(12) Klauda, J. B.; Sandler, S. I. A Fugacity Model for Gas Hydrate Phase Equilibria. *Ind. Eng. Chem. Res.* **2000**, *39*, 3377−3386.

(13) Englezos, P.; Bishnoi, P. R. Prediction of Gas Hydrate Formation Conditions in Aqueous Electrolyte Solutions. *AIChE J.* **1988**, *34*, 1718−1721.

(14) Parrish, W. R.; Prausnitz, J. M. Dissociation Pressures of Gas Hydrates Formed by Gas Mixtures. *Ind. Eng. Chem. Process Des. Dev.* **1972**, *11*, 26−35.

(15) McMullan, R. K.; Kvick, Å. Neutron Diffraction Study of the Structure II Clathrate Hydrate: 3.5Xe.8CCl4.136D2O at 13 and 100 K. *Acta Crystallogr., Sect. B: Struct. Sci.* **1990**, *46*, 390−399.
(16) Hollander, F.; Jeffrey, G. A. Neutron diffraction study of the crystal structure of ethylene oxide deuterohydrate at 80°K. *J. Chem. Phys.* **1977**, *66*, 4699−4705.
(17) Koh, C. A.; Westacott, R. E.; Zhang, W.; Hirachand, K.; Creek, J. L.; Soper, A. K. Mechanisms of Gas Hydrate Formation and Inhibition. *Fluid Phase Equilib.* **2002**, *194−197*, 143−151.
(18) Walsh, M. R.; Koh, C. A.; Sloan, E. D.; Sum, A. K.; Wu, D. T. Microsecond Simulations of Spontaneous Methane Hydrate Nucleation and Growth. *Science* **2009**, *326*, 1095−1098.
(19) Sloan, E. D., Jr; Fleyfel, F. A Molecular Mechanism for Gas Hydrate Nucleation from Ice. *AIChE J.* **1991**, *37*, 1281−1292.
(20) Christiansen, R. L.; Sloan, E. D. Mechanisms and Kinetics of Hydrate Formation. *Ann. N.Y. Acad. Sci.* **1994**, *715*, 283−305.
(21) Radhakrishnan, R.; Trout, B. L. A New Approach for Studying Nucleation Phenomena Using Molecular Simulations: Application to CO2 Hydrate Clathrates. *J. Chem. Phys.* **2002**, *117*, 1786−1796.
(22) Guo, G.-J.; Li, M.; Zhang, Y.-G.; Wu, C.-H. Why Can Water Cages Adsorb Aqueous Methane? A Potential of Mean Force Calculation on Hydrate Nucleation Mechanisms. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10427−10437.
(23) Jacobson, L. C.; Hujo, W.; Molinero, V. Amorphous Precursors in the Nucleation of Clathrate Hydrates. *J. Am. Chem. Soc.* **2010**, *132*, 11806−11811.
(24) Jacobson, L. C.; Molinero, V. Can Amorphous Nuclei Grow Crystalline Clathrates? The Size and Crystallinity of Critical Clathrate Nuclei. *J. Am. Chem. Soc.* **2011**, *133*, 6458−6463.
(25) Liang, S.; Kusalik, P. G. Explorations of Gas Hydrate Crystal Growth by Molecular Simulations. *Chem. Phys. Lett.* **2010**, *494*, 123−133.
(26) Vatamanu, J.; Kusalik, P. G. Observation of Two-Step Nucleation in Methane Hydrates. *Phys. Chem. Chem. Phys.* **2010**, *12*, 15065−15072.
(27) Liu, H.; Kumar, S. K.; Douglas, J. F. Self-Assembly-Induced Protein Crystallization. *Phys. Rev. Lett.* **2009**, *103*, 018101.
(28) Ten Wolde, P. R.; Frenkel, D. Enhancement of Protein Crystal Nucleation by Critical Density Fluctuations. *Science* **1997**, *277*, 1975−1978.
(29) Savage, J. R.; Dinsmore, A. D. Experimental Evidence for Two-Step Nucleation in Colloidal Crystallization. *Phys. Rev. Lett.* **2009**, *102*, 198302.
(30) Zhang, Z.; Walsh, M. R.; Guo, G.-J. Microcanonical Molecular Simulations of Methane Hydrate Nucleation and Growth: Evidence That Direct Nucleation to SI Hydrate Is among the Multiple Nucleation Pathways. *Phys. Chem. Chem. Phys.* **2015**, *17*, 8870−8876.
(31) Báez, L. A.; Clancy, P. Computer Simulation of the Crystal Growth and Dissolution of Natural Gas Hydrates. *Ann. N.Y. Acad. Sci.* **1994**, *715*, 177−186.
(32) Rodger, P. M.; Forester, T. R.; Smith, W. Simulations of the methane hydrate/methane gas interface near hydrate forming conditions conditions. *Fluid Phase Equilib.* **1996**, *116*, 326−332.
(33) Chakraborty, S. N.; Grzelak, E. M.; Barnes, B. C.; Wu, D. T.; Sum, A. K. Voronoi Tessellation Analysis of Clathrate Hydrates. *J. Phys. Chem. C* **2012**, *116*, 20040−20046.
(34) Jacobson, L. C.; Matsumoto, M.; Molinero, V. Order Parameters for the Multistep Crystallization of Clathrate Hydrates. *J. Chem. Phys.* **2011**, *135*, 074501.
(35) Nguyen, A. H.; Molinero, V. Identification of Clathrate Hydrates, Hexagonal Ice, Cubic Ice, and Liquid Water in Simulations: The CHILL+ Algorithm. *J. Phys. Chem. B* **2015**, *119*, 9369−9376.
(36) Jacobson, L. C.; Hujo, W.; Molinero, V. Thermodynamic Stability and Growth of Guest-Free Clathrate Hydrates: A Low-Density Crystal Phase of Water. *J. Phys. Chem. B* **2009**, *113*, 10298−10307.
(37) Walsh, M. R.; Rainey, J. D.; Lafond, P. G.; Park, D.-H.; Beckham, G. T.; Jones, M. D.; Lee, K.-H.; Koh, C. A.; Sloan, E. D.; Wu, D. T.; Sum, A. K. The Cages, Dynamics, and Structuring of Incipient Methane Clathrate Hydrates. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19951−19959.
(38) Mahmoudinobar, F.; Dias, C. L. GRADE: A Code to Determine Clathrate Hydrate Structures. *Comput. Phys. Commun.* **2019**, *244*, 385−391.
(39) Guo, G.-J.; Zhang, Y.-G.; Liu, C.-J.; Li, K.-H. Using the Face-Saturated Incomplete Cage Analysis to Quantify the Cage Compositions and Cage Linking Structures of Amorphous Phase Hydrates. *Phys. Chem. Chem. Phys.* **2011**, *13*, 12048−12057.
(40) Guo, G. J.; Zhang, Y. G.; Li, M.; Wu, C. H. Can the Dodecahedral Water Cluster Naturally Form in Methane Aqueous Solutions? A Molecular Dynamics Study on the Hydrate Nucleation Mechanisms. *J. Chem. Phys.* **2008**, *128*, 194504.
(41) Laage, D.; Hynes, J. T. A Molecular Jump Mechanism of Water Reorientation. *Science* **2006**, *311*, 832−835.
(42) Downs, G. M.; Gillet, V. J.; Holliday, J. D.; Lynch, M. F. Review of Ring Perception Algorithms for Chemical Graphs. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 172−187.
(43) Matsumoto, M.; Baba, A.; Ohmine, I. Topological Building Blocks of Hydrogen Bond Network in Water. *J. Chem. Phys.* **2007**, *127*, 134504.
(44) Li, L.; Zhong, J.; Yan, Y.; Zhang, J.; Xu, J.; Francisco, J. S.; Zeng, X. C. Unraveling Nucleation Pathway in Methane Clathrate Formation. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 24701−24708.
(45) Plimpton, S. Short-Range Molecular Dynamics. *J. Comput. Phys.* **1997**, *117*, 1−42.
(46) Abascal, J. L. F.; Sanz, E.; Fernández, R. G.; Vega, C. A Potential Model for the Study of Ices and Amorphous Water: TIP4P/Ice. *J. Chem. Phys.* **2005**, *122*, 234511.
(47) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. Optimized Intermolecular Potential Functions for Liquid Hydrocarbons. *J. Am. Chem. Soc.* **1984**, *106*, 6638−6646.

https://dx.doi.org/10.1021/acs.jpcb.0c08964
J. Phys. Chem. B 2021, 125, 1282−1292