ARTICLE

Check for updates

Three-body aggregation of guest molecules as a key step in methane hydrate nucleation and growth

Wenfeng Hu[1], Cong Chen [1,2 ✉], Jingyue Sun[1], Ning Zhang[3], Jiafei Zhao[1,2], Yu Liu[1,2], Zheng Ling[1,2], Weizhong Li[1,2], Weiguo Liu[1,2] & Yongchen Song [1,2 ✉]

Gas hydrates have an important role in environmental and astrochemistry, as well as in energy materials research. Although it is widely accepted that gas accumulation is an important and necessary process during hydrate nucleation, how guest molecules aggregate remains largely unknown. Here, we have performed molecular dynamics simulations to clarify the nucleation path of methane hydrate. We demonstrated that methane gather with a three-body aggregate pattern corresponding to the free energy minimum of three-methane hydrophobic interaction. Methane molecules fluctuate around one methane which later becomes the central gas molecule, and when several methanes move into the region within 0.8 nm of the potential central methane, they act as directional methane molecules. Two neighbor directional methanes and the potential central methane form a three-body aggregate as a regular triangle with a distance of ~6.7 Å which is well within the range of typical methane-methane distances in hydrates or in solution. We further showed that hydrate nucleation and growth is inextricably linked to three-body aggregates. By forming one, two, and three three-body aggregates, the possibility of hydrate nucleation at the aggregate increases from 3/6, 5/6 to 6/6. The results show three-body aggregation of guest molecules is a key step in gas hydrate formation.

[1] School of Energy and Power Engineering, Dalian University of Technology, 116024 Dalian, P. R. China. [2] Key Laboratory of Ocean Energy Utilization and Energy Conservation of Ministry of Education, 116024 Dalian, P. R. China. [3] School of Petroleum and Chemical Engineering, Dalian University of Technology, 124221 Panjin, P. R. China. ✉email: congchen@dlut.edu.cn; songyc@dlut.edu.cn

Gas hydrates have an important role in methane hydrate exploitation[1], CO_2 capture and separation[2], water treatment[3], energy storage[4], oil and gas pipelines[5,6], icy satellites[7] and interstellar medium[8,9]. The mechanisms of gas hydrate nucleation and growth is essential for these applications. Understanding the pathways of gas hydrate nucleation and growth are of great scientific significance and will provide guidance for controlling strategies of hydrate formation.

Guest gas exist in the form of a large number of nanobubbles to overcome the solubility barrier when guest gas is provided in hydrate-based technologies or methane hydrate decomposition process[10–12]. The existence of nanobubbles provides a large amount of gas–water interface for hydrate nucleation and have been regarded as the origin of hydrate secondary formation during methane hydrate exploitation[13,14]. However, Although the relationship between hydrate decomposition or formation and nanobubbles has been discussed[15,16], the influence mechanism of nanobubbles on the nucleation of hydrate from the molecular level is still open to questions.

Hypotheses for gas hydrate nucleation mechanism include: classical nucleation theories[17], labile cluster hypothesis[18], local structure hypothesis[19], cage adsorption hypothesis[20], blob hypothesis[21,22], interface nucleation mechanism[23], and hydration layer compression/shedding hypothesis[24]. With classical nucleation theory, the activation barrier of gas hydrate nucleation is predicted combing the free energy change due to creation of a new interface and a more stable phase[17,25]. However, the classical nucleation theory is challenged when handling excess free energy and critical radius of hydrate nucleus. And it does not provide the detailed nucleation pathway as well as the hydrate structure.

The labile cluster hypothesis describes that there are labile rings of water molecules in liquid which can aid formation of the critical nucleus by agglomeration around the dissolved guest molecules[18]. Against from labile cluster hypothesis, the local structuring hypothesis describes the origin of gas hydrate nucleation as the local ordering of guest molecules caused by thermal fluctuation[19]. In cage adsorption hypothesis, gas hydrate formation is triggered by the strong attraction of water cage with guest molecules[20]. The blob hypothesis proposed that the nucleation originates from a guest rich amorphous precursor where the amorphous clathrate cage forms and dissolves until a critical nucleus forms[21]. The interface nucleation mechanism describes that guest molecules diffuse and adsorb by the gas liquid interface where a cage was generated by water molecules around the adsorbed guest molecules[23]. For the hydration layer compression/shedding hypothesis, the water hydration layers of the neighbor methane molecules are compressed to form ternary water-ring aggregations which is recognized as the fundamental structures in gas hydrate nucleation[24]. With these hypotheses, the main debate focuses on whether the gas hydrate nucleation is originally triggered by guest molecules or water molecules. There are several unresolved questions in the research of methane hydrate nucleation and growth[26]. It is well accepted that gas accumulation is an important and necessary process during hydrate nucleation and growth[19,24]. However, how the guest molecules aggregate and their aggregation patterns remains unknown.

Here, we performed a series of molecular dynamics simulations which is a powerful tool to reveal hydrate nucleation and growth pathways[24,27]. Different boxes of methane–water solutions were
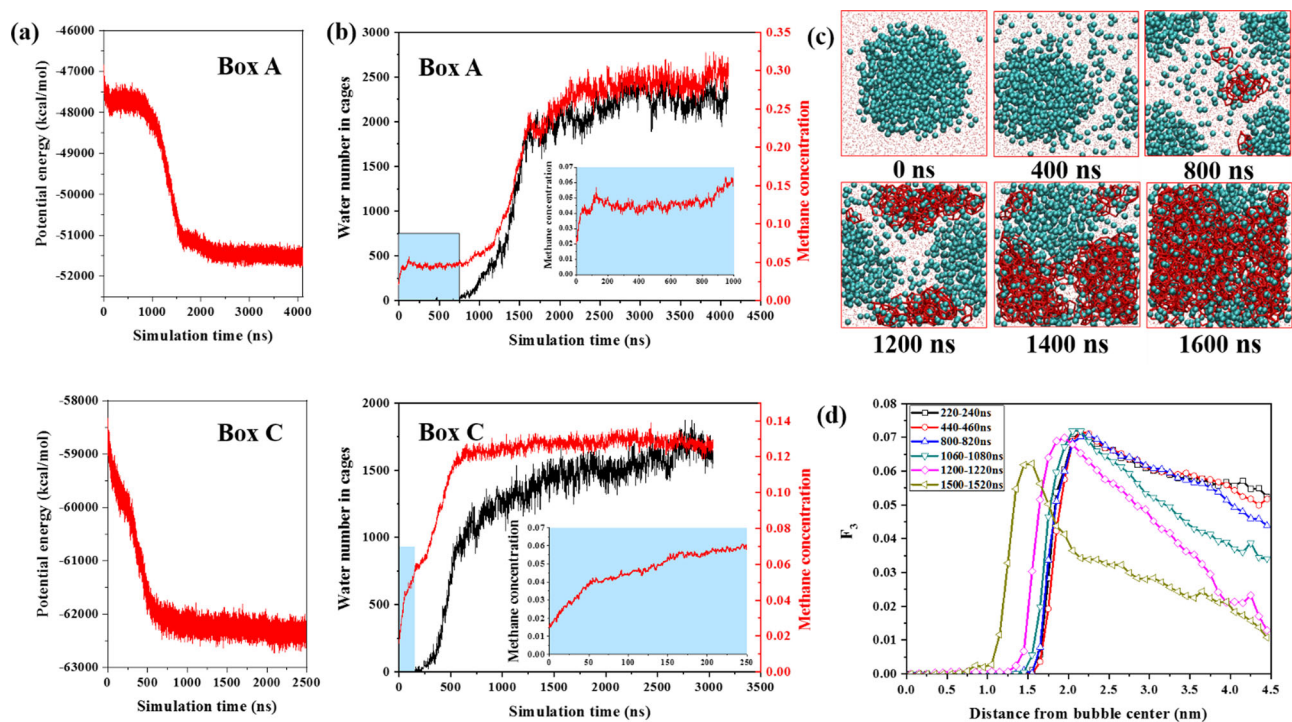

Fig. 1 Potential energy, water number in cages, methane concentration, F_3 order parameter, and system configuration along typical hydrate growth trajectories. a System potential energy for boxes A and C. **b** The changes of methane concentration as well as water number in cages for boxes A and C. The induction period is denoted as a marked region. The change of methane concentration in the induction period is enlarged. For box A, the concentration increases rapidly in the initial stage and then gradually maintains at a certain value, indicating that the system concentration has reached saturation, and the nucleation is triggered when the system is in a period of supersaturation. While for box C, the concentration of the system continues to increase to ~0.13 without maintaining any concentration value before nucleation starts. **c** Snapshots of a typical hydrate formation process (methane molecules are marked in cygan balls, water cages are marked in red lines). **d** The F_3 order parameters in box A as a function of distance from the center of the nanobubble in different time periods.
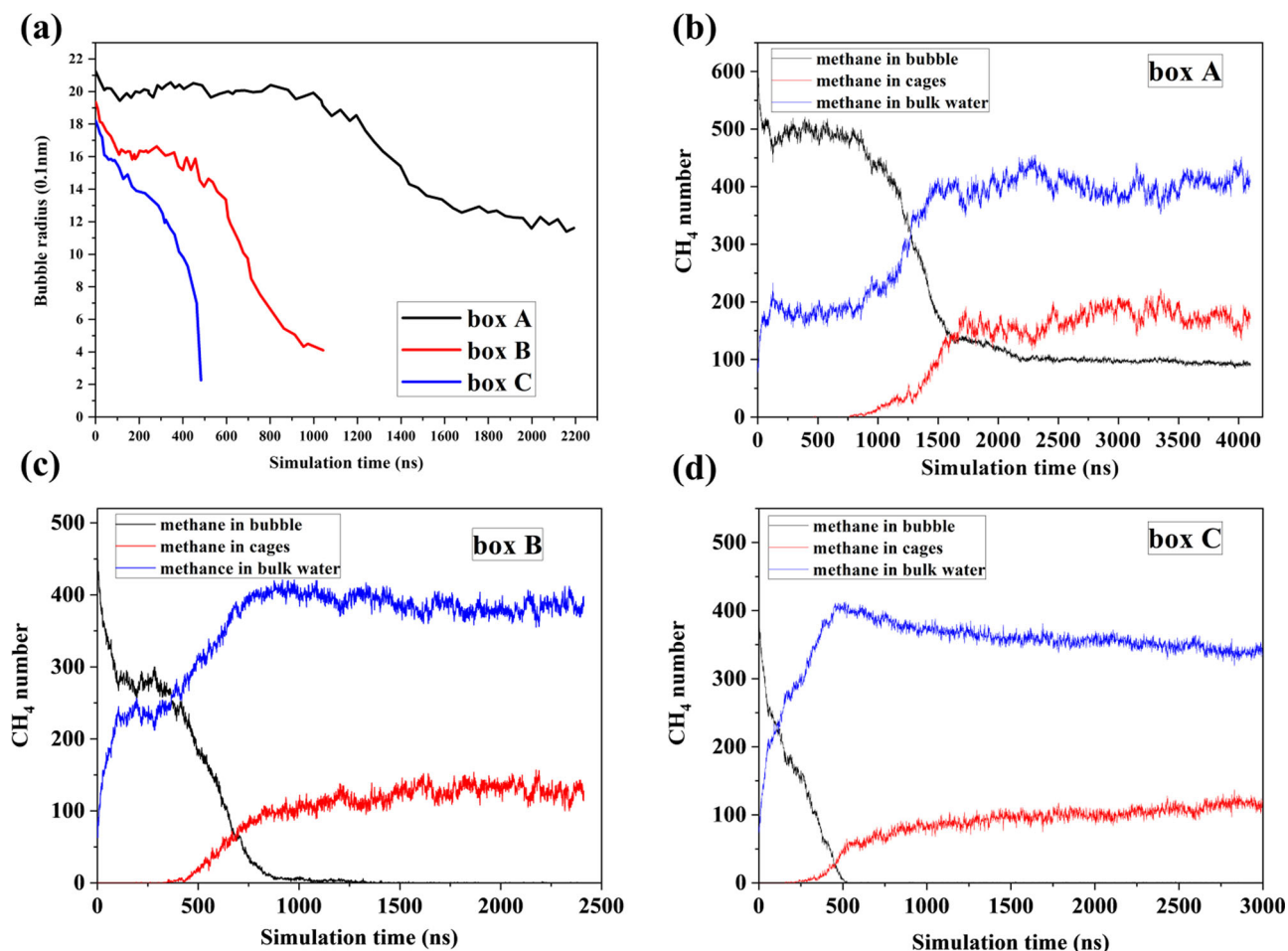
Fig. 2 The evolution of dynamic characteristics in the process of hydrate nucleation for different boxes. a The changes in the bubble radius over time for boxes A–C. **b** The change in the number of methane molecules in different status for box A. **c** The change in the number of methane molecules in different status for box B. **d** The change in the number of methane molecules in different status for box C.

constructed where nanobubbles of methane gas molecules with different sizes were applied as a gas supply (Supplementary Fig. 1a). Simulation boxes were generated using randomly distributed water and methane molecules to avoid the "memory effect"[28]. The initial molar fraction of methane ranges between 0.086 and 0.148, including perfect or insufficient gas supply (Supplementary Fig. 1b). The potential energy, system configuration, hydrate growth features were analyzed using the trajectories of systems with different nanobubble sizes. The effects of nanobubble on gas hydrate nucleation and growth were then investigated. By tracking gas aggregation trajectory, the relationship between hydrate nucleation and methane aggregation patterns has been constructed. Then three-body aggregates of methane molecules were recognized to control hydrate nucleation, which were proved by inserting different numbers of three-body aggregates in a methane–water box to regulate hydrate nucleation sites and growth rates.

Results and discussion
Effects of nanobubble on gas hydrate nucleation and growth.
Figure 1a plots the potential energy along typical hydrate growth trajectories. For box A, the system potential first slightly decreases and then descends sharply after a plateau period. The decrease at the early stage is caused by methane dissolving, and the sharp descending is due to gas hydrate growth as well as accompanying methane dissolving[27]. The plateau indicates a relative equilibrium

between gas bubble and water solution in the induction period. Surprisingly, the potential energy for box C descends continuously. The missing of the plateau is due to the continuously dissolving of methane as huge internal pressure makes small nanobubbles difficult to survive[29]. The different evolution behavior controlled by the initial gas–water ratio is confirmed by the variation of bubble radius (Fig. 2a) and the number of methane with different statuses (Fig. 2b–d) (refer to the "Methods" section for the identification of CH_4 phase). The bubble still exists until hydrate no longer grows in box A, while it disappears for smaller initial gas–water ratios. Similar trends were found from six independent simulations, as illustrated in Supplementary Fig. 2. Figure 1b illustrates the changes in methane concentration as well as water number in cages. After methane dissolution, a supersaturation stage is essential for gas hydrate nucleation[28]. Although not always guarantee nucleation[28], supersaturation is the driving force for gas hydrates crystallization[30,31]. The supersaturation concentration when hydrate nucleates for box A and C is 0.04691 and 0.05583, which are slightly larger than the data of Walsh[32] because we use smaller bubble sizes. The supersaturation concentration proposed in this paper is also greater than critical methane concentration (0.044) which predicted by Guo et al. [20,26]. Hydrate growth rate strongly depends on the increasing rate of methane concentration which agrees well with experiments[33].

Based on the evolution of water number in cages (Fig. 1b) and the cage number (Fig. 3a–c), the hydrate growth experiences three stages as shown in Fig. 3d (the schematic diagram is modified
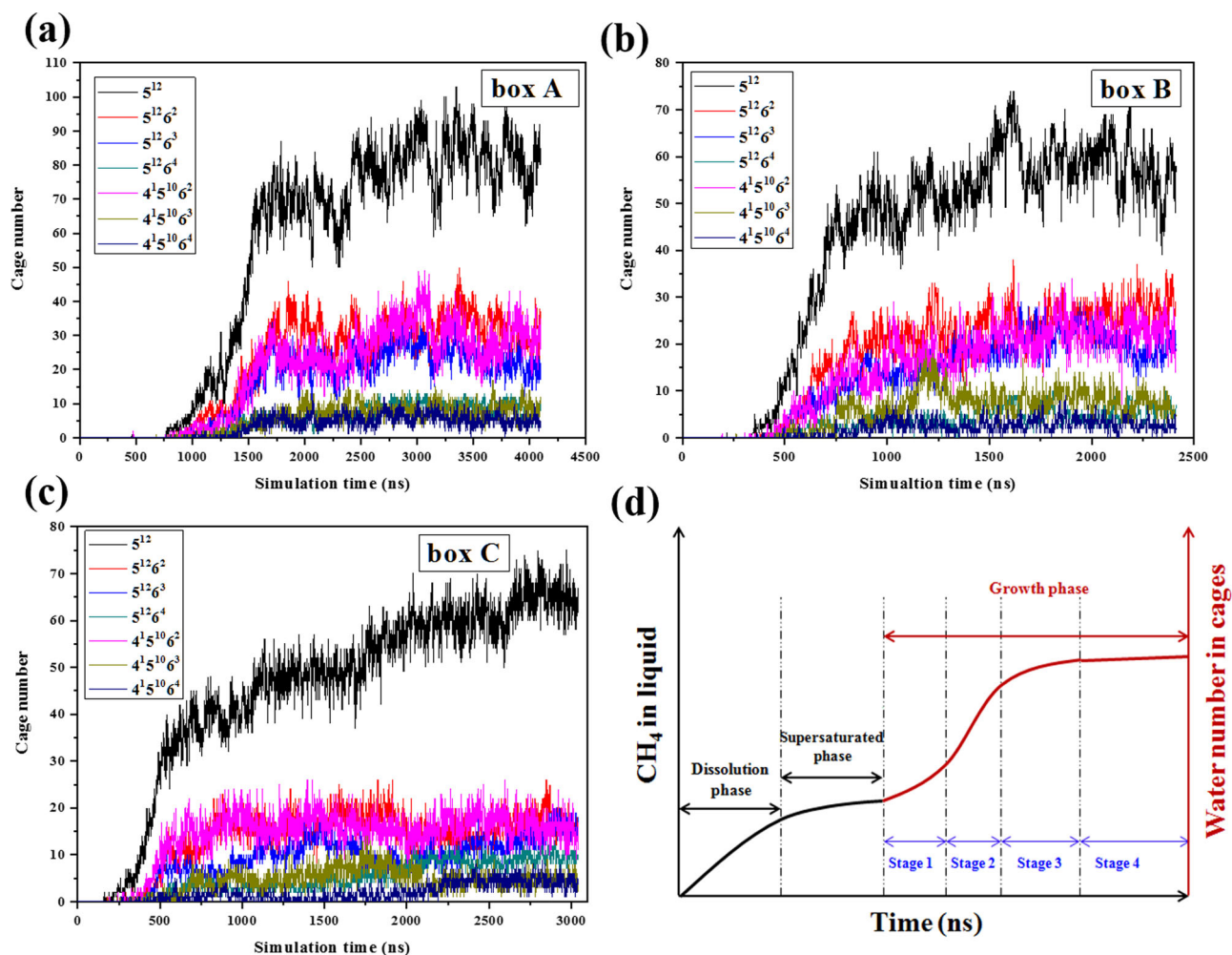
Fig. 3 Hydrate cage number evolution with time and different hydrate growth stages. a–c The change of hydrate cage numbers for simulation boxes A–C. **d** An illustration of different stages for hydrate nucleation and growth. Stage 4 denotes no more hydrate grows.

based on Khurana's[28]): initial growth stage 1, where gas bubble collapses slowly; fast growth stage 2, where the shrinking rate of bubbles becomes faster, leading to rapid growth of hydrate; post growth stage 3, where excessive methane concentration in the bulk phase increases the resistance to continuous dissolution of bubbles and reduces the bubble collapse rate. Hydrate grows at a speed of 2–3 times faster in stage 2 compared with stage 1 (Supplementary Table 1). Figure 1c gives the typical snapshots of box A. From macroscopic experiments, hydrate grows from the gas–water interface[33]; however, it nucleates at several nm away from the interface at a molecular level (Supplementary Fig. 3a–c). A water adsorption layer has been found at flat or curved gas–water interface[32,34]. In the adsorption layer, the number density of water rings is large[34], but the local methane concentration is relatively low (Supplementary Fig. 3d–f). As approaching the gas–water interface from 0.8 to 0.4 nm in the water side, methane concentration decreases by 61.7%, 49.1%, and 58.3% for boxes A–C, respectively. The drop of methane concentration seems to be more intense than that near flat gas–water interface[32]. Figure 1d illustrates the radial distribution of F_3 order parameter in different time periods (refer to the "Methos" section for the calculation method of F_3 sequence parameters). Similar to a flat interface[31], a maximum of F_3 order parameter forms at the gas–water interface. However, the F_3 in the water phase gradually decreases as moving away from the interface. At 440-460 ns, the F_3 at ~4.41 nm is the smallest, which agrees well with the location of the first hydrate cage, as shown in

Supplementary Fig. 3a. The average F_3 decreases with time and further away from the interface, it drops faster. However, before 820 ns, F_3 only decreases in the zone where the radial distance from the bubble center is more than 4 nm. After 1080 ns, the interface moves towards the bubble canter as the bubble size tends to decrease (Fig. 2a). After 1200 ns, compared with those in the zone closer to the bubble center, the decreasing rate of F_3 order parameter in the zone at larger distance with the bubble center becomes smaller because of hydrate growth.

Aggregation characteristics of guest molecules during hydrate nucleation. The gas aggregation trajectory around the central methane during hydrate nucleation has been analyzed. The central methane is the one which locates in the center of cage when the first hydrate cage is formed. The first complete hydrate cage was identified by analyzing molecular dynamics simulation trajectories. We track the migration history of the central methane before the hydrate cage is formed. The region within 0.8 nm near the central methane is defined as region A. As shown in Fig. 4a, b, at first, the numbers of methane and water molecules fluctuate; then methane molecules gradually approach to the central methane and at the same time, water molecules move away from the central methane due to adsorption layer compression[24]. Figure 4c, d illustrate that when gas aggregates towards the central methane, the F_3 order parameter of water molecules in region A decreases significantly which directly indicates that the
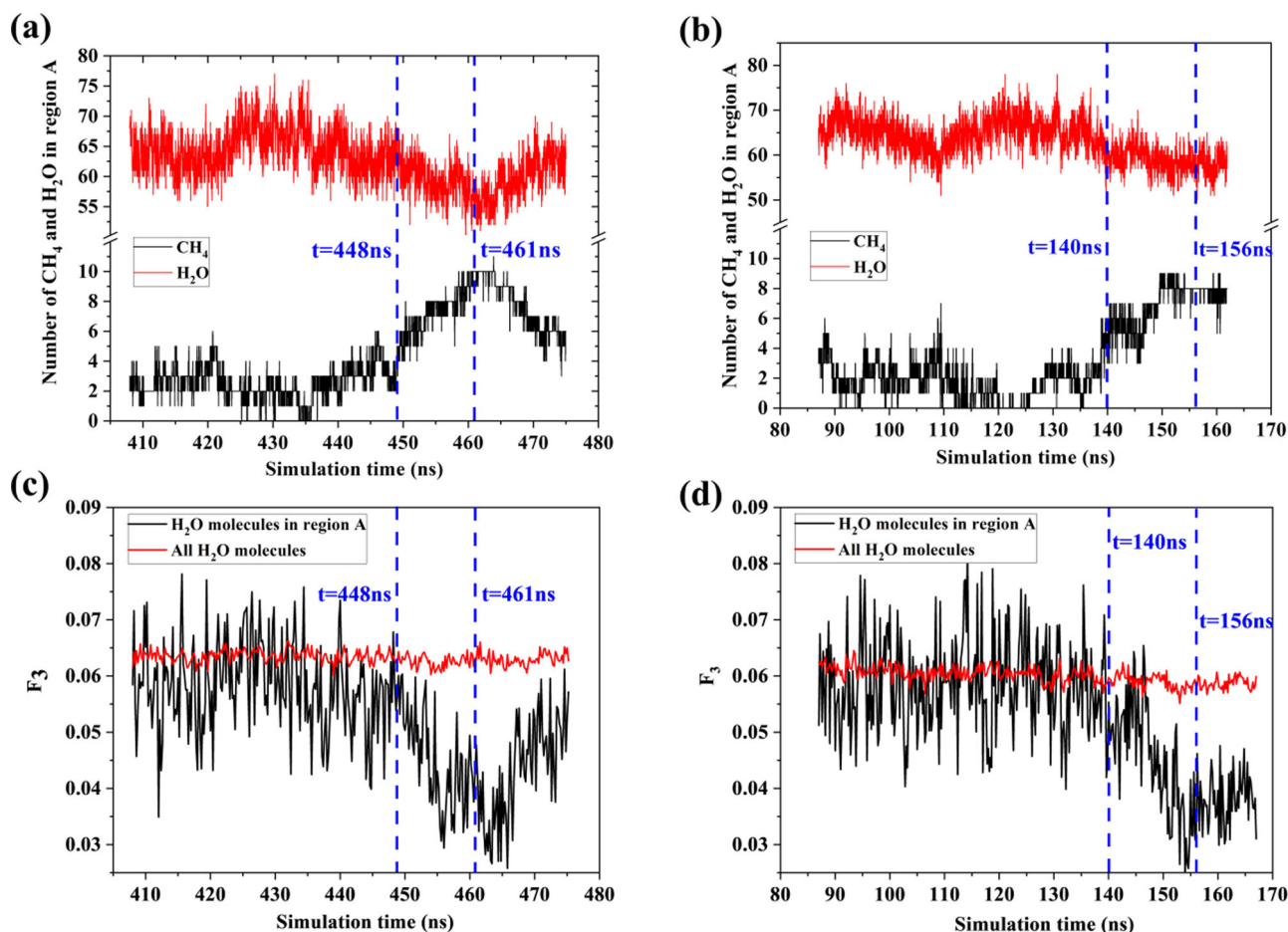
Fig. 4 Methane distribution around central methane in boxes A and C. a, **b** The numbers of methane and water molecules in region A as a function of simulation time for boxes A and C. **c**, **d** Variation of F_3 order parameter for boxes A and C. Both results for all water and water in region A are shown. **e** Change of numbers of methane and water molecules in the L1–L5 layers for box C.

structural order of water molecules is inseparable from the proximity of methane molecules. To further investigate how methane molecules gather, region A is divided into several layers and Fig. 5a illustrates the numbers of methane and water molecules in these layers. Before nucleation, there is one or two methane within 0.4 nm of the central methane. After 448 ns (the moment when F_3 of water in region A starts to decrease, Fig. 4c), the methane in L1 and L2 layers gradually move out, and at the meantime, the number of methane in L4 layer (0.6–0.7 nm) continues to increase. During nucleation, the numbers of water in L2 and L4 layers fluctuate but no significant variation has been found. More water molecules move into L1 layer while the number of water molecules in L3 is significantly reduced. At 461 ns, the numbers of water in L1 and L2 are ~12 and ~8, respectively, which agrees well with the structure of sI hydrate cage[35]. The numbers of water and methane in L5 slightly change but it is not significant. The maximum number of methane molecules reduced in L1 and L2 is 4 while the maximum number of methane molecules increased in L3 and L4 is 8, indicating that the increased methane molecules in L3 and L4 not only come from L1 and L2, but also from outer space. Similar characteristics have been revealed and the results for box C are illustrated in Supplementary Fig. 4.

Gas aggregation pattern of the first hydrate cage during nucleation has been further investigated. In all simulation runs, 5^{12} cages are always formed prior to $5^{12}6^2$ and $5^{12}6^4$, so only the aggregation around the 5^{12} hydrate cage is analyzed. The neighbor CH$_4$ molecules which are located within 0.8 nm from

the central methane when the hydrate cage is formed, were marked and the result for box A is shown in Fig. 5b. Ten methane were found around the hydrate cage. Each of them is located at the opposite side of one five-membered ring corresponding to the central methane and they will be referred to as "directional methane" of the five-membered rings. A three-body pattern is formed by the central methane and the two directional methane of two neighbor five-membered rings. The aggregation geometry of these three-body methane was analyzed using $L_{\text{three-body}}$ and $\theta_{\text{three-body}}$ where $L_{\text{three-body}}$ denotes the distance between the central methane and one directional methane and $\theta_{\text{three-body}}$ represents the angle formed by the three methane molecules (centered at the central methane). Figure 5c displays $L_{\text{three-body}}$ of the 10 directional methane as a function of time. From 448 ns, the directional methane molecules aggregate towards the central methane and the final distances are ~6.7 Å which equals the peak distance of radial distribution functions between methane and methane molecules in pure methane hydrate (Supplementary Fig. 5) and the solvent-separated methane–methane pair in water[36,37]. Two methane molecules (with id 198 and 90) have been close to the central methane at ~3.9 Å (the distance of a contact methane–methane pair in water[37]) for several ns before nucleation. Although the nature is unclear, it is well known that the hydrophobic interaction of methane in gas hydrate is governed by solvent separated pairs[38]. They tend to be solvated by water upon nucleation from contact to solvent separation interaction with the central methane. Figure 5d shows the variation of $\theta_{\text{three-body}}$ with time. Although the angles frequently
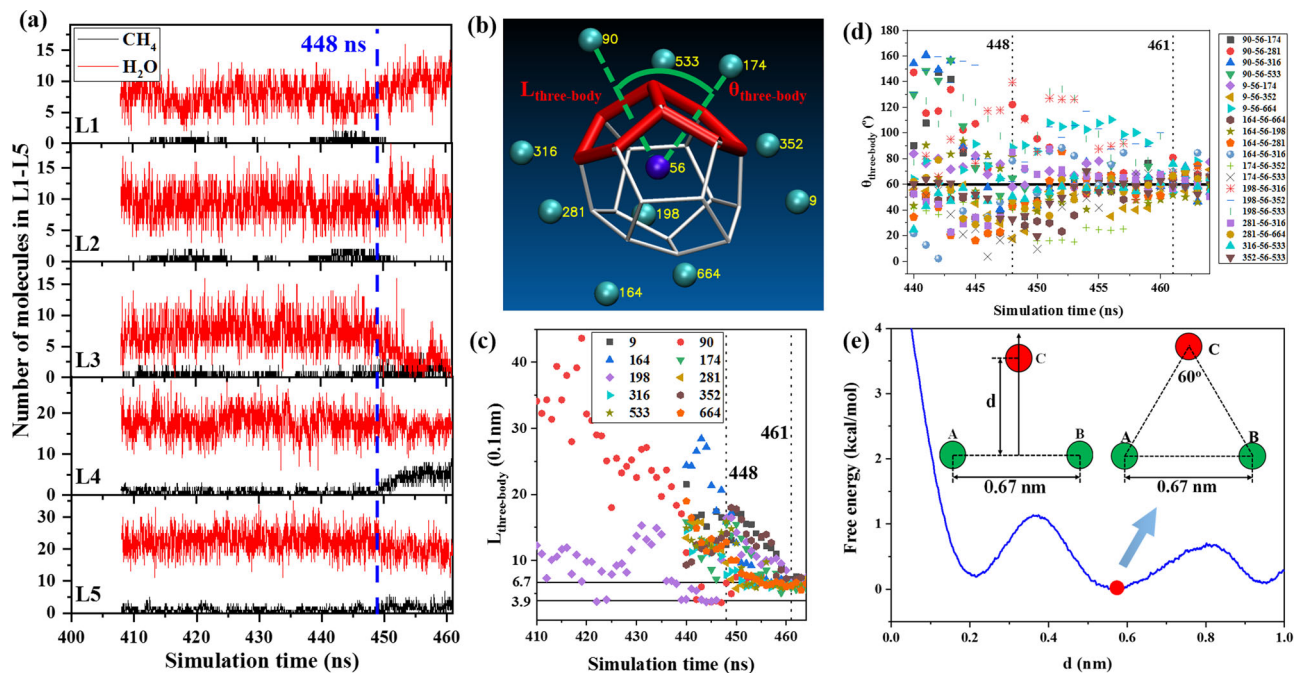
Fig. 5 Gas aggregation characteristics around the central methane during hydrate nucleation (box A). a Change of numbers of methane and water molecules in the L1–L5 layers. The layers are distributed by the distance to central methane. L1 ≤ 0.4 nm, 0.4 nm < L2 ≤ 0.5 nm, 0.5 nm < L3 ≤ 0.6 nm, 0.6 nm < L4 ≤ 0.7 nm, 0.7 nm < L5 ≤ 0.8 nm. **b** A typical formed methane hydrate 5^{12} cage and the neighbor methane in region A. Balls represent different methane molecules (the central methane: blue; others: cyan) and their ids are recorded in yellow. A five-membered ring as well as its neighbor with one common ring are marked in red. **c** Change of $L_{three-body}$ during hydrate nucleation. Two methane with id 90 and 198 are shown from 410 ns and the rest are shown from 440 ns for clarity. **d** The angle $\theta_{three-body}$ varying with time during hydrate nucleation. **e** Free energy of the three-body interactions for methane in water at 50 MPa and 250 K. Methane A, B are frozen and methane C locates at the perpendicular bisector of methane A and B. d is a reaction coordinate which is the vertical distance of methane C with A and B. The red ball indicates the minimum of free energy curve.

fluctuate, they favor being ~60° when the directional methane molecules aggregate towards the central methane. For a 5^{12} cage in the crystalline sI methane hydrate which has the full 12 directional methane, the corresponding $L_{three-body}$ ranges from 6.6 to 7.0 Å with an average 6.7 (0.2) Å and the corresponding $\theta_{three-body}$ is 49.8°–67.9° with an average 63.7 (5.4)°. The difference of $\theta_{three-body}$ in the nucleation process and the crystalline state is not surprising. During nucleation the directional methane molecules have not been captured by hydrate cages and other cages will continue to form which causes $L_{three-body}$ and $\theta_{three-body}$ to slightly fluctuate. Similar evolutionary characteristics of $L_{three-body}$ and $\theta_{three-body}$ for box C have been found and the results are illustrated in Supplementary Fig. 6. By fixing the separation of two methane at 6.7 Å, the free energy of the three-body interaction of methane molecules was calculated as illustrated in Fig. 5e (refer to the "Methods" section for the calculation of free energy). It has been shown that for three-body interaction the optimally packed cluster can only form when the third methane is placed symmetry relative to the methane dimer[37]. As a result, we select the distance between the center of two frozen methane and the third methane as a reaction coordinate. The free energy minimum locates at $d = 0.58$ nm showing that the three methane molecules favor forming a regular triangle, which agrees well with the gas aggregate pattern in Fig. 5c, d. This implies that during nucleation, methane molecules aggregate with a three-body solvent separated pattern. The three methane molecules which aggregate with this pattern are referred to as a "three-body aggregate".

Figure 6a illustrates the gas aggregation process in region A. The formation of cages is significantly related to the proximity of methane. Figure 6b shows the numbers of methane, five-membered rings and three-body aggregates in region A.

As methane aggregates, the number of five-membered rings gradually increases. However, the number of five-membered rings is not directly proportional to the degree of methane accumulation. From 454 to 459 ns, there are nine methane molecules around the central methane; however, the number of five-membered rings changes from 9 to 17. The variation of the number of five-membered rings is caused by the changing of gas aggregation pattern, as shown in Fig. 5c, d. The number of five-membered rings is highly correlated with the number of three-body aggregates. As more methane molecules are involved in the three-body aggregates, the number of five-membered rings increases accordingly. Figure 6c, d demonstrate the number of water in cages and the three-body aggregate number as a function of simulation time (refer to the "Methods" section for the identification of three-body aggregates). It is found that the evolution of the number of water in cages is consistent with that of three-body aggregate. The good consistency is also obtained for six independent runs of box A as summarized in Supplementary Fig. 7. It shows that the growth of hydrate is inextricably linked to the three-body aggregates. It should be noted that the aggregation characteristics of guest molecules during hydrate nucleation were found in six independent simulations for each simulation box except for the typical time related with different induction times (Supplementary Table 2).

Effects of three-body aggregates on gas hydrate nucleation and growth. To reveal the effect of three-body aggregate on hydrate nucleation and growth, systems with different numbers of three-body aggregates were constructed. Figure 7 displays the effects of three-body aggregates on hydrate nucleation sites. Three-body aggregation of methane molecules is a key step for hydrate
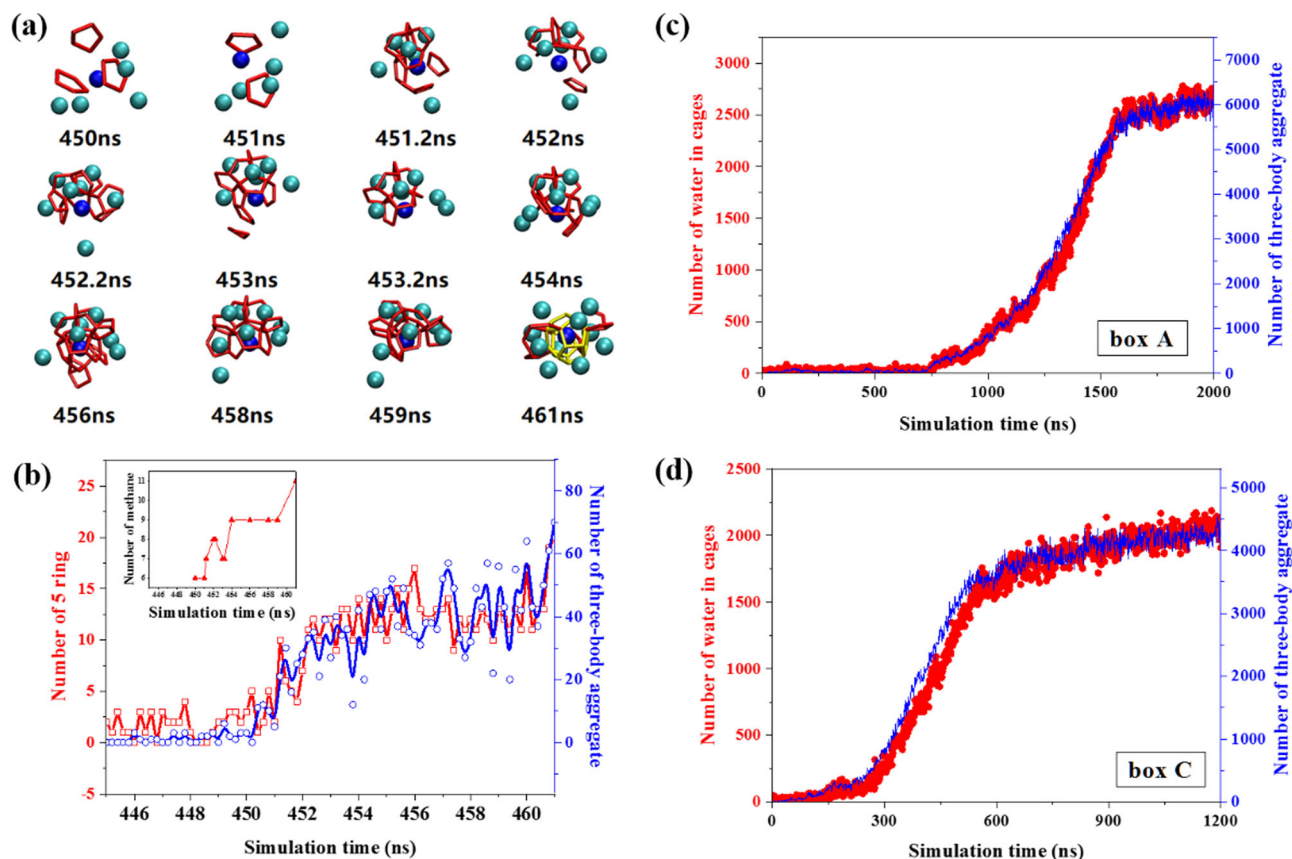
Fig. 6 Snapshots of gas aggregation and evolution of the number of three-body aggregates. a Gas aggregation process in region A which is within 0.8 nm of the central methane. The central methane is represented by the blue ball. The red lines indicate the five-membered rings and the yellow lines represent a complete 5^{12} cage. **b** The number of five-membered rings and the three-body aggregates structure in region A during nucleation. The number of methane molecules in the same region was drawn in the inner picture. **c**, **d** Number of water in cages as well as the number of three-body aggregates during hydrate nucleation and growth for boxes A and C, respectively.

nucleation; however, it is not the sufficient condition. More methane molecules are required to gather around to form a hydrate cage. It is still possible for those methane molecules at other locations to form three-body aggregates by themselves. If only one three-body aggregate is initially provided, hydrate nucleates at the three-body aggregate in 3/6 (runs 2, 4, 5) of the independent runs. In the case of two three-body aggregates, hydrate nucleation was triggered by the three-body aggregates for 5/6 (runs 2–6) of the performed runs. In all the six independent runs, hydrates were found to nucleate at the three-body aggregate if there are initially three three-body aggregates in the system. It takes a long time for gas molecules to spontaneously accumulate into "three-body aggregates" as shown in Fig. 6c, d. By placing a three-body aggregate, the possibility of other methane molecules to aggregate with a three-body solvent separated pattern was enhanced. The possibility is further strengthened by enlarging the number of initial three-body aggregates. Two of the aggregates both trigger hydrate nucleation in run 3 for the system with two aggregates and runs 3, 4 and 6 for the system with three aggregates. However, nucleation at three aggregates simultaneously was not found in these systems.

Figure 8a–c demonstrates the cage numbers as a function of simulation time for different systems with one, two, and three three-body aggregates. Due to the spontaneous nature of hydrate nucleation, the induction time and growth speed show deviations for different runs even the initial systems are the same[24]. However, the acceleration of three-body aggregates on hydrate nucleation and growth is still clear. The sum of cage numbers for

six runs is illustrated in Fig. 8d. It is rather obvious that three-body aggregates enhance hydrate nucleation and growth. As the number of three-body aggregates increases in the initial system, the hydrate nucleation induction time becomes shorter and the hydrate grows at a faster rate. At 443 ns, the average numbers of hydrate cages for systems with initially two and three three-body aggregates are ~4.3 and ~10.3 times of that for the system with only one three-body aggregate, respectively. The ratios decrease to ~2.6 and ~5.4 at 543 ns. It takes ~633 ns for the system with one three-body aggregate to form 50 cages. For the systems with two and three three-body aggregates, the period has been shortened by ~16% and ~38%, respectively. For systems containing zero, one, two and three three-body aggregates, the nucleation rates of the hydrate in the four systems are 7.91×10^{24}, 7.91×10^{24}, 9.57×10^{24}, and 1.18×10^{25} nuclei $cm^{-3}\,s^{-1}$, respectively, which agree well with the predicted data by Walsh et al. [32]. The nucleation rate is nearly affected by placing only one three-body aggregate. However, it seems that by placing two and three three-body aggregates, the nucleation rate increases by 21% and 49%, respectively, compared with the system without any three-body aggregate.

Conclusion

In summary, we constructed different methane–water boxes to investigate the effects of nanobubble on hydrate nucleation and growth. The evolution behavior was controlled by the initial gas–water ratio. At larger gas–water ratios, a plateau was found
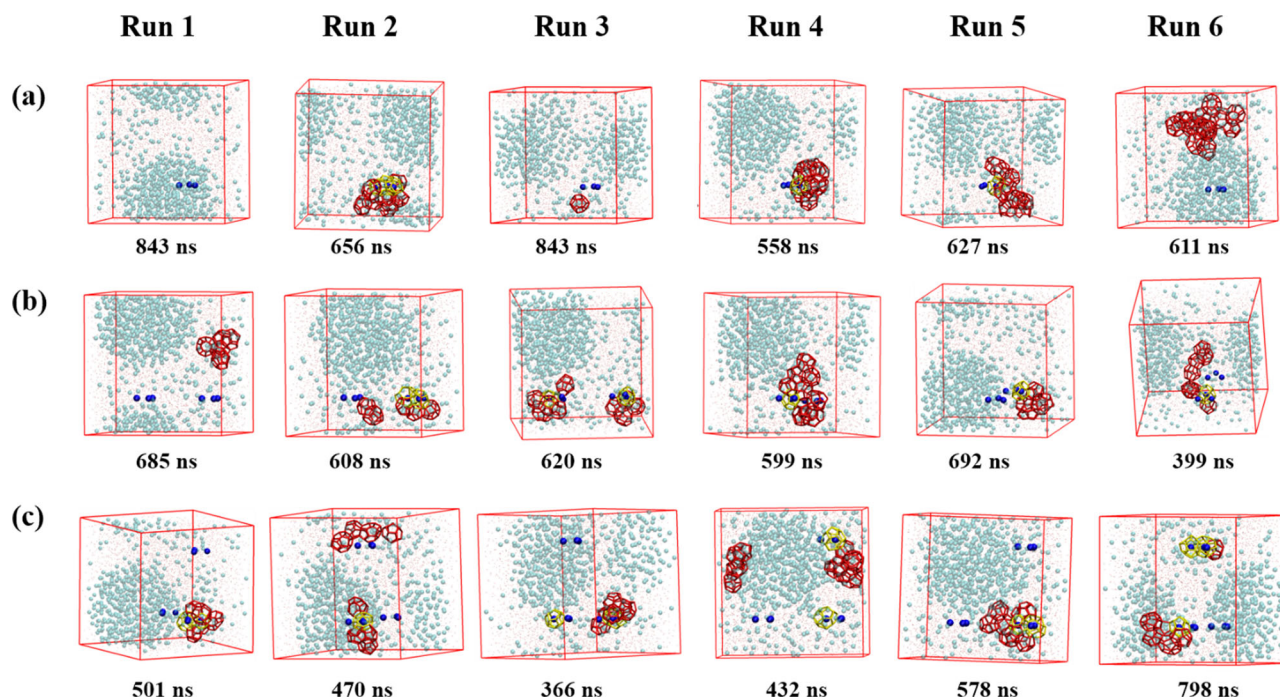
Fig. 7 Effects of three-body aggregates on hydrate nucleation sites. Different number of three-body aggregates were added into the initial configuration and the corresponding numbers of methane molecules were removed to ensure the consistency of the number of molecules in the system. The methane molecules in the three-body aggregates were fixed during simulation to keep their aggregation patterns. For each system, six independent runs were performed. **a–c** display the results for systems with one, two and three three-body aggregates, respectively. The snapshots when the number of hydrate cages becomes larger than 10 are present. If the number of hydrate cages is always smaller than 10, the snapshot at final status is provided. The blue balls represent methane molecules in the three-body aggregates. The simulation boxes are rotated for clarity. The hydrate cages which central methane is a part of the three-body aggregate is marked in yellow; otherwise, the cages are drawn in red.

for curves of potential energy, methane concentration, and water number in cages, indicating a relative equilibrium between gas bubble and water solution in the induction period. Gas hydrate nucleates at several nm away from the gas–water interface at a molecular level as the methane concentration is low and F_3 order parameter is high at the interface.

The gas aggregation trajectory was tracked and evolution of methane in different layers around the central methane show distinct features. During nucleation, methane in layers (≤ 0.5 nm) gradually move out, while the number of methane in an outer layer (0.6–0.7 nm) continues to increase. Gas aggregation pattern analysis demonstrates that a regular triangle pattern is formed by the central methane and the two directional methane of two neighbor five-membered rings. The length of the triangle is ~6.7 Å which equals the peak distance of C–C radial distribution functions in methane hydrate and the solvent separated methane–methane pair in water. The regular triangle pattern is confirmed as the free energy minimum for three body interaction of methane molecules in water when two of them are solvent separated. The variation of the three-body aggregates of guest molecules is well consistent with that of five-membered rings and number of water in cages. By providing different numbers of three-body aggregates, the nucleation sites and hydrate growth rate have been significantly regulated. If one, two, and three three-body aggregates are initially provided, hydrate nucleation triggers at these three-body aggregates in 3/6, 5/6, and 6/6 of the independent runs, and in some cases, more than one aggregates trigger hydrate nucleation. As the number of three-body aggregates increases in the initial system, the hydrate nucleation induction time becomes shorter and the hydrate grows at a faster rate. The results demonstrate three-body aggregation of guest molecules as a key step in gas hydrate nucleation and growth.

However, what causes the methane molecules aggregate into the found pattern is unclear. We did search the trajectory of these methane molecules, however, the structure of methane–water solutions is rather complicated under high pressure and low temperature conditions. As shown by Guo et al. [39], there are thousands of cage types occurring in methane aqueous solution and in hydrate nucleation processes. Even we identify all the cage types, it is hard to distinguish what "cages" are the main cause of the three-body aggregate. As the cages and water molecules around the methane molecules vary frequently with time and the existence of cage affects the formation of other cages. Among the theories related with methane–water interaction, the cage adsorption hypothesis is most probably the reason to cause three-body aggregate. Future studies are required to investigate what cages and how to adsorb methane molecules to form the three-body aggregate and thus to control the gas hydrate nucleation and growth.

Methods

MD simulation. A system shown in Supplementary Fig. 1a was built. All molecular dynamics simulations were performed by nanoscale molecular dynamics (NAMD)[40]. The CHARMM format potential energy function was used to calculate the total potential energy of the system. For CH_4, we adopt the united-atom model TraPPE[41], and for H_2O, we adopted the four-point rigid model TIP4P/ice[42]. The detailed force field parameters are listed in Supplementary Table 3. TraPPE–TIP4P/ice model combination was proved to predict the hydrate phase diagram[43] successfully and has been applied by several authors[15,44,45]. In all simulations, the periodic boundary conditions were applied in the three directions of x, y, and z to eliminate boundary effects. The initial velocities of particles satisfied the Maxwell–Boltzmann distribution. The cut-off radius is set to be 1.2 nm and the time step was set to be 2.0 fs. The particle mesh Ewald[46] (PME) method was used to calculate the Coulombic interactions for full system and the desired relation error in force set to be 10^{-6}. The non-iterative SETTLE algorithm is used to keep the rigidity of water[47]. The pressure of the system was maintained
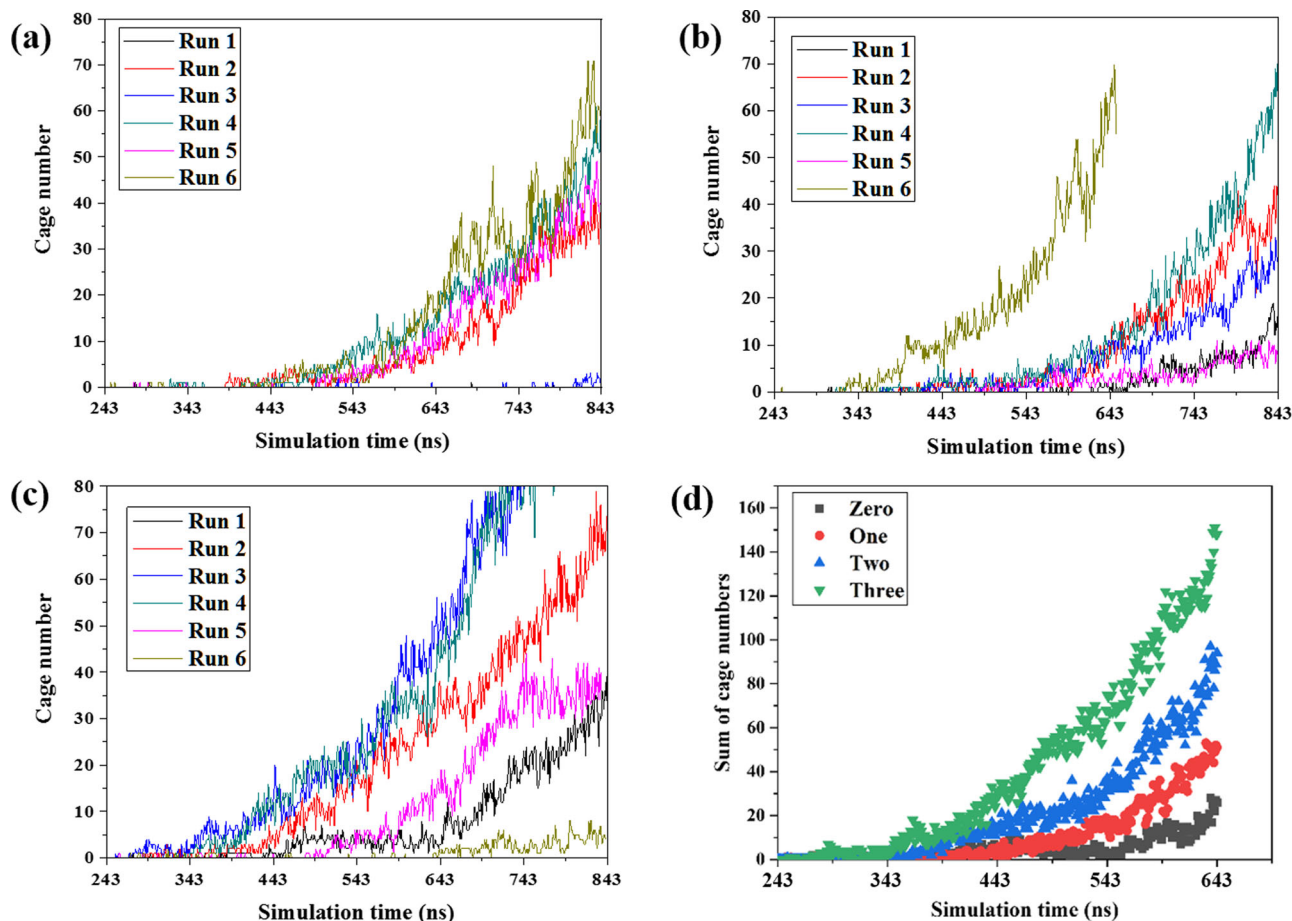
Fig. 8 Acceleration of three-body aggregates on hydrate nucleation and growth. The total cage numbers were calculated and the results for one, two, and three three-body aggregates systems are illustrated in (**a–c**), respectively. **d** The cage numbers for the six independent runs were summed and compared for systems with zero, one, two, and three aggregates.

using the Langevin position[48] method with a Langevin piston period of 100 time steps and a Langevin piston decay of 50 time steps. The temperature of the system was controlled by Langevin dynamic[49] method with a damping coefficient of 5 ps. The system runs thousands of nanoseconds at 50 MPa and 250 K to output data. The in-house codes written by our team were used to analyze the data and VMD software was used to visualize the simulation results[50]. Hydrate cages were identified according to the computational code of Sum[51]. It should be noted that a thermostat has been applied which may affect the transport coefficients in liquids, however, the gas aggregation patterns found in this manuscript are not influenced.

Identification of three-body aggregates. The three-body aggregate is composed of three methane molecules in the liquid phases, and these three methane have a certain positional relationship. With methane 1 as the center, methane 2 and methane 3 are at a distance of 0.55–0.75 nm from methane 1, and the angle between them is 50°–70° (with methane 1 as the vertex).

Identification of CH4 phase. The number of solvent molecules around methane in gas and liquid is different. By calculating the RDF of CH_4–H_2O in methane–water dilute solution, it was found the number of water molecules in the first solvation shell of methane is about 20 within 5.5 Å. Therefore, a methane molecule was defined as gas phase if the number of water molecules in the range of 5.5 Å of it is <11. A methane molecule with more than 10 water-like water molecules in the first hydration shell was considered to be in the solution phase.

Identification of H2O phase. The phase of water molecules can be identified by calculating the angular order parameters of water molecules and counting the number of water rings that water molecules participate in. AOP is an angular order parameter consisting of three water molecule configurations that describe the degree of deviation of a tetrahedron formed by a central oxygen atom and other

oxygen atoms from the regular tetrahedron, which can be defined as[52]

$$AOP = \sum_{j=1}^{n_i-1} \sum_{k=j+1}^{n_i} \left(|\cos\theta_{jik}|\cos\theta_{jik} + \left(\cos 109.47°\right)^2 \right)^2$$

where θ_{jik} donates the angle between the oxygen atoms of water molecule j, i and k; $109.47°$ indicates the angle between the center of the regular tetrahedron and the vertices. The cut-off distance between the oxygen atom and its nearest neighbors is 3.5 Å, corresponding to the first minimum in the water oxygen–oxygen radial distribution function of the hydrate phase. If a water molecule with AOP < 0.4 and participates in 4 or 5 or 6 five-membered rings, the water molecule was defined as hydrate phase. After distinguishing water molecules into hydrate phase or liquid phase using the above method, it is necessary to perform a second identification. If a water molecule in hydrate has three or more adjacent water molecules in liquid (the adjacent distance is 3.5 Å), the water molecule in hydrate will be revised as liquid water; if a water molecule in liquid has three or more adjacent water molecules in hydrate (the adjacent distance is 3.5 Å), the water molecule in the liquid will be revised as hydrate water.

F3 Order parameter. The F_3 order parameter is also an angular order parameter to describe the degree of deviation of a tetrahedron formed by a central oxygen atom and other oxygen atoms from the regular tetrahedron. The F_3 order parameter has been commonly used to characterize hydrate structures in the previous research. Unlike AOP, the calculation of F_3 order parameter involves the averages of values, which can be defined as[13,27,53]

$$F_3 = \frac{AOP}{n_i(n_i-1)/2}$$

where n_i represents the nearest neighbors of water molecule i. For water molecules in liquid, $F_3 = 0.09$; for water molecules in hydrate, $F_3 = 0.01$.

Free energy calculation. The free energy of the three-body interactions for methane was calculated by the extended-system adaptive biasing forces (eABF)[54],

which is a variant of the adaptive biasing forces (ABF) method[55], which can overcome the limitations of the traditional ABF method and improve the convergence speed of the simulation. As shown in Fig. 5e, in the calculation process, methane A and B are fixed, and methane C is located at the perpendicular bisector of methane A and B. The used colvar (collective variables) component is distance. The reaction pathway d is defined as the distance from methane C to the center of methane A and B. The lower boundary of d is 0, the upper boundary of d is 10, and the calculated width is 0.01. The units of lower boundary, upper boundary and width were Å. The lower and upper wall constant are set to be 100.0, and the unit is kcal mol^{-1}. The movement of methane C in the y and z directions is restricted.

Data availability

The initial and final configurations of molecular dynamics trajectories were provided in ScienceDB with the access link as: https://www.scidb.cn/s/6R32Uv. Other data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability

The code for analyzing three-body aggregates that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Sloan, E. D. & Koh, C. A. *Clathrate Hydrates of Natural Gases*, 3rd edn (CRC Press, 2008).
2. Lim, J., Choi, W., Mok, J. & Seo, Y. Kinetic CO2 selectivity in clathrate-based CO2 capture for upgrading CO2-rich natural gas and biogas. *Chem. Eng. J.* **369**, 686–693 (2019).
3. Sahu, P., Krishnaswamy, S., Ponnani, K. & Pande, N. K. A thermodynamic approach to selection of suitable hydrate formers for seawater desalination. *Desalination* **436**, 144–151 (2018).
4. Yu, C. et al. Hydrogen and chemical energy storage in gas hydrate at mild conditions. *Int. J. Hydrog. Energy* **45**, 14915–14921 (2020).
5. Gbaruko, B. C., Igwe, J. C., Gbaruko, P. N. & Nwokeoma, R. C. Gas hydrates and clathrates: flow assurance, environmental and economic perspectives and the Nigerian liquified natural gas project. *J. Pet. Sci. Eng.* **56**, 192–198 (2007).
6. Hammerschmidt, E. G. Formation of gas hydrates in natural gas transmission lines. *Ind. Eng. Chem.* **26**, 855 (1934).
7. Kamata, S., Nimmo, F., Sekine, Y., Kuramoto, K. & Tani, A. Pluto's ocean is capped and insulated by gas hydrates. *Nat. Geosci.* **12**, 407–410 (2019).
8. Ghosh, J. et al. Clathrate hydrates in interstellar environment. *Proc. Natl Acad. Sci. USA* **116**, 1526–1531 (2019).
9. Luspay-Kuti, A. et al. The presence of clathrates in comet 67P/Churyumov-Gerasimenko. *Sci. Adv.* **2**, https://doi.org/10.1126/sciadv.1501781 (2016).
10. Sujith, K. S. & Ramachandran, C. N. Carbon dioxide induced bubble formation in a CH4–CO2–H2O ternary system: a molecular dynamics simulation study. *Phys. Chem. Chem. Phys.* **18**, 3746–3754 (2016).
11. Uddin, M. & Coombe, D. Kinetics of CH4 and CO2 hydrate dissociation and gas bubble evolution via MD simulation. *J. Phys. Chem. A* **118**, 1971–1988 (2014).
12. Yang, L., Falenty, A., Chaouachi, M., Haberthur, D. & Kuhs, W. F. Synchrotron X-ray computed microtomography study on gas hydrate decomposition in a sedimentary matrix. *Geochem. Geophys. Geosyst.* **17**, 3717–3732 (2016).
13. Bagherzadeh, S. A., Alavi, S., Ripmeester, J. & Englezos, P. Formation of methane nano-bubbles during hydrate decomposition and their effect on hydrate growth. *J. Chem. Phys.* **142**, 214701 (2015).
14. Uchida, T., Yamazaki, K. & Gohara, K. Gas nanobubbles as nucleation acceleration in the gas-hydrate memory effect. *J. Phys. Chem. C* **120**, 26620–26629 (2016).
15. Yagasaki, T., Matsumoto, M., Andoh, Y., Okazaki, S. & Tanaka, H. Effect of bubble formation on the dissociation of methane hydrate in water: a molecular dynamics study. *J. Phys. Chem. B* **118**, 1900–1906 (2014).
16. Huang, X. et al. Effect of micro- and nanobubbles on the crystallization of THF hydrate based on the observation by atomic force microscopy. *J. Phys. Chem. C* **124**, 13966–13975 (2020).
17. Sloan, E. D. Fundamental principles and applications of natural gas hydrates. *Nature* **426**, 353–359 (2003).
18. Christiansen, R. L. & Sloan, E. D. In *International Conference on Natural Gas Hydrates Vol. 715 Annals of the New York Academy of Sciences* (eds. Sloan, E. D., Happel, J. & Hnatow, M. A.) 283–305 (The New York Academy of Sciences, 1994).
19. Radhakrishnan, R. & Trout, B. L. A new approach for studying nucleation phenomena using molecular simulations: application to CO2 hydrate clathrates. *J. Chem. Phys.* **117**, 1786–1796 (2002).
20. 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.* **11**, 10427–10437 (2009).
21. Vatamanu, J. & Kusalik, P. G. Observation of two-step nucleation in methane hydrates. *Phys. Chem. Chem. Phys.* **12**, 15065–15072 (2010).
22. Jacobson, L. C., Hujo, W. & Molinero, V. Amorphous precursors in the nucleation of clathrate hydrates. *J. Am. Chem. Soc.* **132**, 11806–11811 (2010).
23. Kvamme, B. & Progep, N. G. H. *A New Theory for the Kinetics of Hydrate Formation. Proceedings of the 2nd International Conference on Natural Gas Hydrates*, 139–146 (Toulouse, France, 1996).
24. Li, L. et al. Unraveling nucleation pathway in methane clathrate formation. *Proc. Natl Acad. Sci. USA* **117**, 24701–24708 (2020).
25. Bishnoi, P. R. & Natarajan, V. Formation and decomposition of gas hydrates. *Fluid Phase Equilib.* **117**, 168–177 (1996).
26. Guo, G.-J. & Zhang, Z. Open questions on methane hydrate nucleation. *Commun. Chem.* **4**, https://doi.org/10.1038/s42004-021-00539-6 (2021).
27. 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* **326**, 1095–1098 (2009).
28. Khurana, M., Yin, Z. & Linga, P. A review of clathrate hydrate nucleation. *ACS Sustain. Chem. Eng.* **5**, 11176–11203 (2017).
29. Tan, B. H., An, H. & Ohl, C.-D. Surface nanobubbles are stabilized by hydrophobic attraction. *Phys. Rev. Lett.* **120**, 164502, https://doi.org/10.1103/PhysRevLett.120.164502 (2018).
30. Kashchiev, D. & Firoozabadi, A. Driving force for crystallization of gas hydrates. *J. Cryst. Growth* **241**, 220–230 (2002).
31. Reed, S. K. & Westacott, R. E. The interface between water and a hydrophobic gas. *Phys. Chem. Chem. Phys.* **10**, 4614–4622 (2008).
32. Walsh, M. R. et al. Methane hydrate nucleation rates from molecular dynamics simulations: effects of aqueous methane concentration, interfacial curvature, and system size. *J. Phys. Chem. C* **115**, 21241–21248 (2011).
33. Maini, B. B. & Bishnoi, P. R. Experimental investigation of hydrate formation behavior of a natural-gas bubble in a simulated deep-sea environment. *Chem. Eng. Sci.* **36**, 183–189 (1981).
34. Kaur, S. P., Sujith, K. S. & Ramachandran, C. N. Formation of a nanobubble and its effect on the structural ordering of water in a CH4-N-2-CO2-H2O mixture. *Phys. Chem. Chem. Phys.* **20**, 9157–9166 (2018).
35. Yousuf, M. et al. Novel results on structural investigations of natural minerals of clathrate hydrates. *Appl. Phys. A* **78**, 925–939 (2004).
36. Li, J.-L., Car, R., Tang, C. & Wingreen, N. S. Hydrophobic interaction and hydrogen-bond network for a methane pair in liquid water. *Proc. Natl Acad. Sci. USA* **104**, 2626–2630 (2007).
37. Moghaddam, M. S., Shimizu, S. & Chan, H. S. Temperature dependence of three-body hydrophobic interactions: potential of mean force, enthalpy, entropy, heat capacity, and nonadditivity. *J. Am. Chem. Soc.* **127**, 303–316 (2005).
38. Matsumoto, M. Four-body cooperativity in hydrophonic association of methane. *J. Phys. Chem. Lett.* **1**, 1552–1556 (2010).
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.* **13**, 12048–12057 (2011).
40. Phillips, J. C. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
41. Jorgensen, W. L., Maxwell, D. S. & TiradoRives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118**, 11225–11236 (1996).
42. Abascal, J. L. F., Sanz, E., Fernandez, R. G. & Vega, C. A potential model for the study of ices and amorphous water: TIP4P/Ice. *J. Chem. Phys.* **122**, 234511 https://doi.org/10.1063/1.1931662 (2005).
43. Conde, M. M. & Vega, C. Determining the three-phase coexistence line in methane hydrates using computer simulations. *J. Chem. Phys.* **133**, 064507 https://doi.org/10.1063/1.3466751 (2010).
44. Bagherzadeh, S. A., Alavi, S., Ripmeester, J. A. & Englezos, P. Evolution of methane during gas hydrate dissociation. *Fluid Phase Equilib.* **358**, 114–120 (2013).
45. Bagherzadeh, S. A., Englezos, P., Alavi, S. & Ripmeester, J. A. Molecular simulation of non-equilibrium methane hydrate decomposition process. *J. Chem. Thermodyn.* **44**, 13–19 (2012).
46. Darden, T., York, D. & Pedersen, L. Particle Mesh Ewald—an N.Log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
47. Miyamoto, S. & Kollman, P. A. Settle—an analytical version of the shake and rattle algorithm for rigid water models. *J. Comput. Chem.* **13**, 952–962 (1992).
48. Feller, S. E., Zhang, Y. H., Pastor, R. W. & Brooks, B. R. Constant-pressure molecular-dynamics simulation - the Langevin piston method. *J. Chem. Phys.* **103**, 4613–4621 (1995).

49. Brunger, A. *X-PLOR:version 3.1: A System for X-ray Crystallography and NMR* (Yale University Press, 1992).

50. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph. Model.* **14**, 33–38 (1996).

51. Walsh, M. R. et al. The cages, dynamics, and structuring of incipient methane clathrate hydrates. *Phys. Chem. Chem. Phys.* **13**, 19951–19959 (2011).

52. Uttormark, M. J., Thompson, M. O., Baez, L. A. & Clancy, P. Solid–liquid cluster recognition in heterogeneous systems. *Mol. Simul.* **11**, 121–144 (1993).

53. Sarupria, S. & Debenedetti, P. G. Molecular dynamics study of carbon dioxide hydrate dissociation. *J. Phys. Chem. A* **115**, 6102–6111 (2011).

54. Lesage, A., Lelievre, T., Stoltz, G. & Henin, J. Smoothed biasing forces yield unbiased free energies with the extended-system adaptive biasing force method. *J. Phys. Chem. B* **121**, 3676–3685 (2017).

55. Darve, E., Rodriguez-Gomez, D. & Pohorille, A. Adaptive biasing force method for scalar and vector free energy calculations. *J. Chem. Phys.* **128**, https://doi.org/10.1063/1.2829861 (2008).

Acknowledgements

Author contributions

W.H. and C.C. wrote this paper. W.H. performed simulations and analyzed results. C.C. and Y.S. supervised the research. W.H., C.C., J.S., and N.Z. produced the pictures. J.Z., Y.L., Z.L., W. Li and W. Liu helped to revise the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42004-022-00652-0.

Correspondence and requests for materials should be addressed to Cong Chen or Yongchen Song.

Peer review information *Communications Chemistry* thanks Guang-Jun Guo and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.