# Unbiased atomistic insight in the competing nucleation mechanisms of methane hydrates

Arjun[a,b], Thom A. Berendsen[a,b], and Peter G. Bolhuis[a,b,1]

[a]Van 't Hoff Institute for Molecular Sciences, University of Amsterdam, 1098 XH Amsterdam, The Netherlands; and [b]Amsterdam Center for Multiscale Modeling, University of Amsterdam, 1090 GD Amsterdam, The Netherlands

Edited by Pablo G. Debenedetti, Princeton University, Princeton, NJ, and approved August 13, 2019 (received for review April 15, 2019)

**Methane hydrates have important industrial and climate implications, yet their formation via homogeneous nucleation under natural, moderate conditions is poorly understood. Obtaining such understanding could lead to improved control of crystallization, as well as insight into polymorph selection in general, but is hampered by limited experimental resolution. Direct molecular dynamics simulations using atomistic force fields could provide such insight, but are not feasible for moderate undercooling, due to the rare event nature of nucleation. Instead, we harvest ensembles of the rare unbiased nucleation trajectories by employing transition path sampling. We find that with decreasing undercooling the mechanism shifts from amorphous to crystalline polymorph formation. At intermediate temperature the 2 mechanisms compete. Reaction coordinate analysis reveals the amount of a specific methane cage type is crucial for crystallization, while irrelevant for amorphous solids. Polymorph selection is thus governed by kinetic accessibility of the correct cage type and, moreover, occurs at precritical nucleus sizes, apparently against Ostwald's step rule. We argue that these results are still in line with classical nucleation theory. Our findings illuminate how selection between competing methane hydrate polymorphs occurs and might generalize to other hydrates and molecular crystal formation.**

gas hydrates | molecular dynamics | polymorph selection | transition path sampling

Methane hydrate is an ice-like solid substance that encages methane inside a hydrogen-bonded water network. While methane gas is hydrophobic under ambient conditions, a mixture of water and methane gas can spontaneously nucleate into hydrates at low temperatures and high pressures (1). Methane hydrates are abundantly present at the ocean floors and have an estimated double the amount of energy stored compared to conventional resources of fossil fuels (2). With the growing demand of natural gas (3), methane hydrates can act as a potential energy resource (4). On the other hand, the formation of hydrates is of concern to the petroleum industry as they can clog oil pipelines (5), causing flow assurance problems. With this background, interest within the scientific community to comprehend the formation mechanism of hydrates has increased tremendously in the past decade.

The primary theoretical framework to understand hydrate formation is based on the widely used classical nucleation theory (CNT), which postulates a growing spherical solid nucleus in a supercooled liquid phase. As an unfavorable interface is created, this nucleus needs to overcome an activation barrier and reach a critical size before it spontaneously further grows into a bulk crystalline phase. The thermodynamically stable phase of methane hydrate is the structure type I (sI) crystal composed of 2 standard cages, $5^{12}6^2$ (12 pentagons and 2 hexagons, Fig. 1*A*) and $5^{12}$ cages in a ratio 3:1 (4). This form is predominant on the Earth's surface. However, methane hydrates could also form via a nonclassical mechanism (6), in which an amorphous metastable intermediate (composed of nonstandard cages and with many fewer $5^{12}6^2$ cages) precedes formation of a crystalline state. Alternatively, the less common crystalline sII phase can form, consisting of $5^{12}6^4$ and $5^{12}$ cages with ratio 1:2 (4).

A conceptual understanding of the nucleation mechanism of methane hydrates is still an open question and, in fact, much debated. Current experimental techniques have limited spatiotemporal resolution and cannot easily provide atomistic details about the mechanism of hydrate nucleation. Molecular dynamics simulations can in principle overcome this limitation and provide such insight. However, under moderate conditions, the free energy barrier associated with nucleation is large, resulting in long induction times before the rare nucleation event occurs. This rare event behavior renders straightforward molecular dynamics simulations impractical (e.g., a barrier of $\sim 10^2 k_B T$ would require a wall clock time exceeding the age of the Universe). As the barrier is much reduced when the driving force is higher, e.g., at stronger undercooling or supersaturation, many simulation studies (1, 6–10) were performed at lower temperatures or supersaturated concentration, to observe nucleation within a feasible simulation time. However, recent experimental work has shown that nucleation at high driving force might differ significantly from the CNT prediction (11). Other approaches have used biasing techniques (12) and/or coarse-grained water models (13) to model the methane hydrate nucleation behavior. While all these approaches aim to tackle the rare event problem, they bias the dynamics, which could severely influence the mechanism. For instance, the most common methane hydrate phase in these simulations tends to be

## Significance

**Mixtures of methane gas and water can form ice-like solid methane hydrates via homogeneous nucleation at natural, moderate conditions. Understanding the formation of the different hydrate polymorphs, required for improved control of crystallization, is hampered by limited experimental resolution. Direct molecular dynamics simulations could provide insight, but it would require computation times beyond the age of the Universe for a single nucleation event to take place. Yet, the nucleation event itself, while rare, occurs fast. Therefore, we sample and examine ensembles of rare unbiased nucleation trajectories. Detailed analysis shows how selection between competing amorphous and crystalline polymorph formation mechanisms takes place. The conclusions shed light onto the formation of clathrate hydrates.**

Author contributions: A. and P.G.B. designed research; A. performed research; A. and T.A.B. analyzed data; and A. and P.G.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.



[1]To whom correspondence may be addressed. Email: p.g.bolhuis@uva.nl.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1906502116/-/DCSupplemental.

First published September 9, 2019.

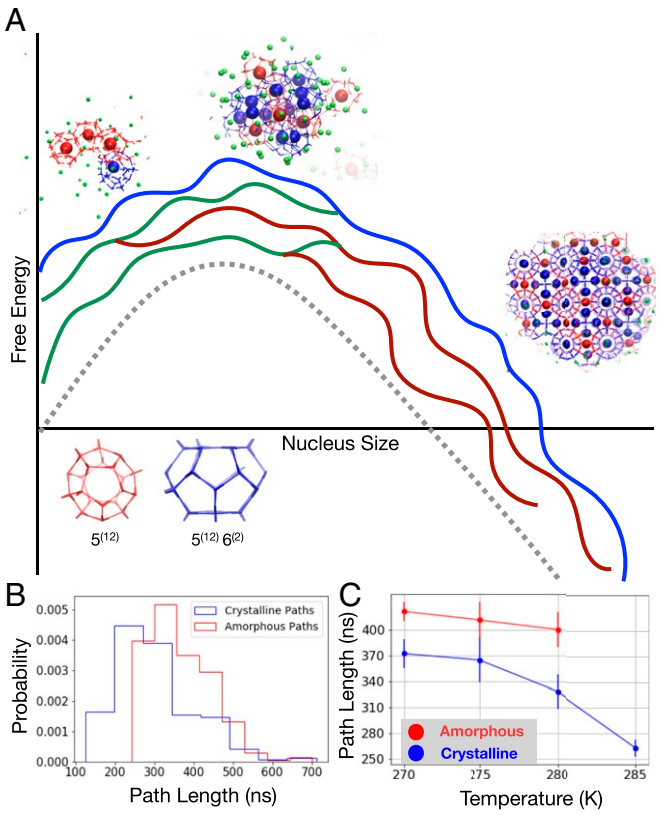


**Fig. 1.** (*A*) Schematic view of TPS 1-way shooting (17). An initial trajectory (blue line) crosses the barrier (dashed curve). Trial trajectories either move backward (green line) to the liquid state or forward (red line) to the solid state. A growing nucleus along a path illustrates the nucleation process and is made up of $5^{12}$ (red) and $5^{12}6^2$ (blue) cages. (*B*) Path length distribution for 280 K separates into amorphous (red) and crystalline (blue) paths, suggesting 2 competing nucleation routes. (*C*) Average path length as a function of temperature. The amorphous paths are ∼50 ns longer than the crystalline paths at any temperature, indicating a different formation mechanism.

amorphous (or sII) and not the experimentally found thermodynamically stable crystalline (sI) phase. This implies that the previously used methods and/or models lack sufficient accuracy, and therefore true, unbiased dynamics using an accurate atomistic force field are required. Moreover, some recent studies (7, 14) indicated that methane hydrate nucleation can follow multiple channels but how this process takes place is yet unexplained.

In this work, we focus on the mechanism of methane hydrate nucleation for moderate undercooling where competition between the formations of crystalline and amorphous solid phases is expected. To avoid all of the abovementioned problems, we employ an accurate atomistic force field (Tip4P/ICE (15)) in combination with the powerful transition path sampling (TPS) (16–18) methodology. TPS harvests an ensemble of unbiased molecular dynamics trajectories connecting the initial reactant and the final product state, by performing a random Monte Carlo walk in trajectory space. By focusing on the rare transition only, the method avoids the induction time involved in the nucleation process, rendering the method exponentially more efficient than straightforward molecular dynamics (MD). It has been stated before (19) that path sampling is necessary to generate a statistically significant set of reactive trajectories to unravel hydrate nucleation. Nevertheless, the cost of even a single trajectory exhibiting a nucleation event is substantial (∼0.5 μs), and while TPS was previously used to study nucleation processes (20–23), these computational costs have prevented use of TPS for atomistic force fields until recently.

The mechanistic aspects of methane hydrate nucleation are studied in a methane/water mixture with a sI stoichiometry, at experimentally relevant temperatures (24) of 270, 275, 280, and 285 K and a pressure of 500 bar. The melting point for the model used is 303 ± 2 K (25). In the metastable liquid state the system is phase separated: A saturated water–methane mixture is in equilibrium with a methane gas reservoir. Starting from an initial path exhibiting the nucleation transition (*SI Appendix*), we performed extensive path sampling simulations, amounting to an aggregate simulation time of more than 1 ms of molecular dynamics, and obtained path ensembles for the barrier crossing from the liquid to the solid phase for different temperatures. Note that we label the product state as solid, because a successful pathway can end up in either an amorphous or a crystalline phase.

Fig. 1 visualizes the scheme for path sampling across a free energy barrier. Starting from an initial path, new pathways are generated which reach either the liquid or the solid state. The individual path length fluctuates as each path can take any route in the free energy landscape and is halted when it reaches the solid or liquid state. Depending on the temperature of sampling, methane hydrate nucleation can occur via 1 of 2 main channels. At low temperatures ($T \leq 275$K), the nucleation process is likely to yield an amorphous solid, whereas at higher temperatures ($T \geq 285$K), the critical nucleus becomes larger, and the preferred nucleating mechanism shifts to the formation of a sI crystal phase. Remarkably, at intermediate temperatures (280 K) both mechanisms are sampled in the TPS run. The path length distribution for 280 K in Fig. 1*B* shows that pathways reaching the amorphous phase are (on average) longer than crystalline paths. Fig. 1*C* shows that the average duration for a path to reach a solid state from the liquid state increases with decreasing temperature as molecules (especially water) slow down at lower temperature, hampering the ordering of the molecules (*SI Appendix*). Moreover, at any particular temperature (e.g., 280 K), transition paths that end in the crystalline state are, on average, shorter (∼325 ns) compared to amorphous paths (∼400 ns) (*SI Appendix*). This faster commitment to the solid phase is also visible as a difference in growth rates (*SI Appendix*, section IIF).

## Mechanistic Interpretation of the Path Ensembles

The obtained ensemble of pathways contains information about the atomistic mechanism and underlying reaction coordinate (RC) of the processes. To obtain further insight, we need to describe the trajectories in terms of meaningful collective variables (CVs). One of the main advantages of TPS is that it allows direct and unbiased extraction of the reaction coordinate (*SI Appendix*). The first CV is the nucleus size, expressed as the number of methanes in the growing solid nucleus, also known as the mutually coordinated guest (MCG) (26). The second CV is the cage ratio (CR) defined as the ratio of $5^{12}6^2$ cages with respect to $5^{12}$ cages. Both CVs are chosen based on extensive reaction coordinate analysis (*SI Appendix*, section IIG) which shows that the size of the nucleus (MCG) and the number of $5^{12}6^2$ cages are key ingredients in the reaction coordinate. The MCG order parameter can distinguish liquid from solid, while the cage ratio distinguishes amorphous (CR $\leq$ 1) from crystal (CR $>$ 1) structures.

To interpret the path ensembles, we present in Fig. 2 all of the sampled (extended; *SI Appendix*) pathways as a 2D path density plot (27), in the plane of the pertinent variables MCG and CR. (Note that we chose CR above the $5^{12}6^2$ cages itself, as the latter would make the plots very skewed.) Paths start in the liquid phase at MCG $\leq$ 2, and undergo nucleation until the

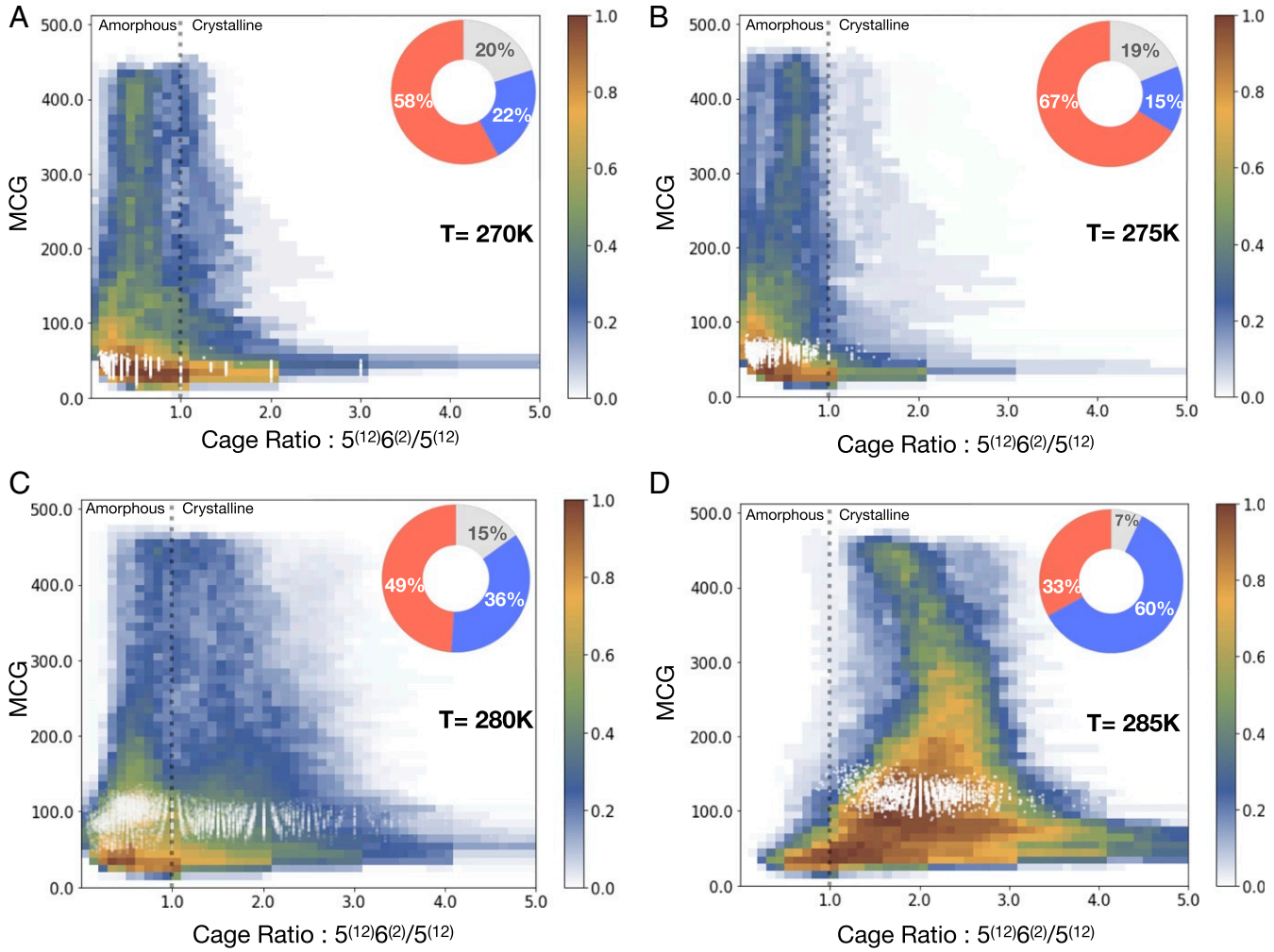


**Fig. 2.** Path densities (PD) from TPS as function of CR and MCG cluster size for the 4 temperatures: (*A*) 270 K, (*B*) 275 K, (*C*) 280 K, and (*D*) 285 K. The white dots overlaying the PD represent the LCP in each path ensemble. The circle diagrams show the overall cage fractions in the LCP, with red color representing $5^{12}$ cages; blue $5^{12}6^2$ cages; and gray color nonstandard $5^{12}6^3$, $5^{12}6^4$, $4^15^{10}6^2$, $4^15^{10}6^3$, and $4^15^{10}6^4$ cages. The black vertical dotted line (at CR = 1) demarcates the crystalline and amorphous regions. Both 270 K and 275 K show dominance of $5^{12}$ cages in the barrier region. At 280 K, the percentage of ($5^{12}6^2$) cages increases and at 285 K, this cage type becomes most abundant. This visible change in the topology of the barrier region reflects the change in mechanism with increasing temperature.

solid nucleus has become postcritical, and will eventually form a system-spanning structure. At low temperature the path density shows 1 dominant channel (Fig. 2*A*). Most paths (80%) end up in the amorphous phase with a cage ratio (CR $\leq 1$). The most common cage type observed throughout the trajectories is the $5^{12}$ for the entire path ensemble, typical for amorphous solid phases. As the system mostly gets trapped in a metastable amorphous structure, the kinetics dominate over thermodynamics at this temperature, and a direct relaxation to the thermodynamically stable state is avoided. This behavior should also be observed for any molecular dynamics simulation of methane hydrate at temperatures below 270 K (28). The amorphous nucleation mechanism might be related to the blob hypothesis (6), which supposes a (metastable) intermediate cluster phase forms before the crystal.

A few trajectories end in a sII hydrate-type structure, characterized by the presence of more $5^{12}6^4$ than $5^{12}6^2$ cages in the last frame of the trajectory. Using Raman spectroscopy, Schicks and Ripmeester (29) showed the sII structure can exist as a kinetic intermediate state before reaching the thermodynamically stable sI type at moderate conditions. We did not observe that rearrangement, due to the long timescales required for the significant reorganization of hydrogen bonds in the solid hydrate.

As the temperature increases, the molecular kinetics become faster (*SI Appendix*), leading to shorter nucleation trajectories and allowing the system to explore more crystalline structures, and the path density shifts to a higher cage ratio. At the highest temperature of 285 K all pathways are dominated by $5^{12}6^2$ cages throughout the transition, with CR $> 1$, and none of the sampled pathways ended in a sII crystal phase. The path density peaks around CR $\approx 2$, very different w.r.t. lower temperatures. While the nucleus grows, the monomers ($5^{12}6^2$ and $5^{12}$ cages) arrange themselves in the sI symmetry, a formation mechanism clearly not in agreement with the blob hypothesis. The fast molecular kinetics at this temperature allow easy arrangement in the most stable hydrate form. Indeed, trajectories are now shorter ($\sim$270 ns) (Fig. 1*C*).

At the intermediate temperature of 280 K, the path density is broad (Fig. 2*C*), suggesting that 2 reaction channels coexist at this temperature, 1 toward the amorphous region and 1 toward the crystalline region. Pathways with initially more $5^{12}$ cages mostly end in an amorphous state while pathways with a higher $5^{12}6^2$ cage fraction solidify toward a sI methane hydrate

(*SI Appendix*, Fig. S5 *D* and *E*). The path sampling switches back and forth between amorphous and crystalline channels, suggesting the free energy barrier for the 2 routes/channels is roughly comparable at this temperature.

The least changed path (LCP) (*SI Appendix* and ref. 17) obtained from the path ensemble is a proxy for the barrier region and hence the critical nucleus. The LCP is plotted as white dots on top of the path densities in Fig. 2. For 270 K the critical nucleus as defined by the LCP has a size between 30 and 40 MCG units, too small to encode much structural information about the solid phase, as there are very few fully formed cages. This can also be inferred from the large degree of discretization in the cage ratio in the LCP, typical for the ratio of small numbers. The size of the critical nucleus grows with temperature to about 120 to 140 MCG units for 285 K.

A structural analysis of the critical nucleus at the different temperatures is given by the circle diagrams in Fig. 2*A*. At 270 and 275 K, the nucleus is predominantly composed of $5^{12}$ cages and thus has a strong amorphous character. The circle diagram in Fig. 2*D* shows that at 285 K the critical nucleus is primarily composed of $5^{12}6^2$ cages, with CR $\approx 2$, indicating that it is already crystalline. Clearly, homogenous nucleation of methane hydrate can directly lead to a thermodynamically stable state, contrary to what was found in previous work (8, 12, 30). At 280 K pathways ending in crystalline and amorphous phases have a similar critical nucleus size, something that has been observed before in the nucleation of ice (31), which indicates that the nucleation barrier for both the amorphous and crystal formations is also located at a similar nucleus size (although at different composition). The LCP switches between $5^{12}$ and $5^{12}6^2$ as the dominant cage type. The cage composition in Fig. 2*C* shows substantially more $5^{12}6^2$ cages compared to 275 K and, in fact, is bimodal.

This bimodality becomes even clearer when we separately plot the amorphous and crystalline pathways (based on their endpoints) for the subcritical region of the nucleation transition in Fig. 3. Both amorphous and crystalline paths have the same kind of cage ratio distribution for MCG $\leq$ 50 (Fig. 3*C*). As the nucleus grows, the amorphous and crystalline paths diverge and the distribution is bimodal in the critical region (Fig. 3*D*). In the postcritical region, the distribution is uniquely separated (Fig. 3*E*). The path densities in Fig. 3 *A* and *B* show that the growing nuclei in the 2 channels branch out in the precritical region. The amorphous paths shift to low CR $\approx 0.2$ values, due to an increasing number of $5^{12}$ cages, before moving back to CR $\approx 0.5$, whereas the crystalline pathways exhibit a steady increase in the amount of $5^{12}6^2$ cages, resulting in a cage ratio above unity in the critical nucleus regime. Clearly, the cage distributions already diverge in the precritical regime. Such selection might be induced by templating, as forming another $5^{12}6^2$ becomes easier if there is already a cage with 2 hexagons present.

## Connection to Classical Nucleation Theory

The critical nucleus size at each temperature can also be estimated by averaging the MCG value of the shooting points in each TPS simulation (Fig. 4*A* and *SI Appendix*, Fig. S4). Using CNT we can then estimate the behavior of the nucleation free energy barrier for the formation of the crystalline and amorphous phases. In *SI Appendix* we show how the driving force $\Delta\mu$ can be estimated from the overpressure of methane in the gas phase and can be related to the critical nucleus size in Fig. 4*A*. Together with the previously established surface tension of methane hydrates (32, 33), this in turn determines the CNT free energy barriers, shown in Fig. 4*B*. Note that at 280 K the barriers for amorphous and crystalline nucleation are indistinguishable. Note also that the CNT barriers are generally lower than previous estimates (12), possibly due to the different water

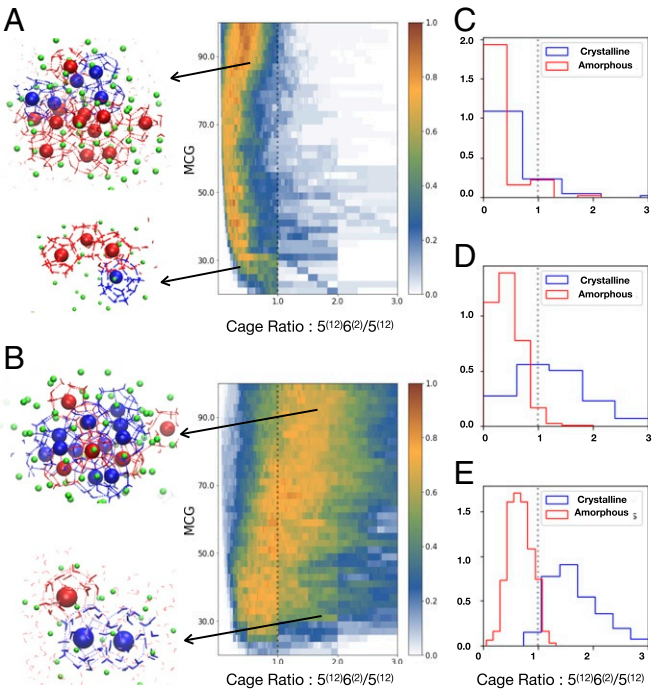


**Fig. 3.** (*A* and *B*) Path densities for amorphous (*A*) and crystalline (*B*) paths obtained at 280 K. The 2 routes already diverge in the precritical region. (*A* and *B*, *Left*) Typical nuclei along the path are depicted, with $5^{12}$ cages in red, $5^{12}6^2$ cages in blue, and methanes in the MCG cluster with incomplete cages as green spheres. (*C*–*E*) Cage ratio distributions for specific nucleus size ranges for the amorphous and crystalline trajectories in the 280-K path ensemble. (*C*) In the precritical domain, MCG $\leq$ 50, both the crystalline and amorphous distributions overlap, indicating that 2 channels are merged in this region. (*D*) In the critical region, 80 $\leq$ MCG $\leq$ 100, the amorphous and crystalline pathways show unimodal distribution peaking at different values of cage ratio with large overlap. (*E*) The 2 distributions are separated in the postcritical region, 250 $\leq$ MCG $\leq$ 300, with minor overlap, showing 2 separate channels.

model used or the higher driving force induced by the spherical methane gas reservoir.

Fig. 4*C* combines the above findings in a schematic CNT-based free energy landscape. For low temperature conditions the free energy landscape is tilted toward the amorphous side. When the temperature increases, the free energy landscape tilts to make the barrier at higher crystallinity the preferred one. While we stress that this is only a qualitative analysis and mechanistic details will change with the guest molecule, this switching mechanism may be common in other gas hydrates and water nucleation. Further research using enhanced sampling methods can give a more accurate quantitative estimate.

## Discussion

In summary, using extensive transition path sampling simulations (exceeding 1 ms simulation time) we have shown that for decreasing undercooling, homogeneous methane hydrate nucleation switches from a mechanism in which an amorphous solid is formed to a mechanism in which a crystalline solid is formed. In line with CNT, the free energy landscape has 2 underlying channels with 2 nucleation barriers that depend on the temperature. At low temperatures, where the molecular kinetics are slow, the system nucleates into an amorphous intermediate through a rough energy landscape, while at higher temperatures, kinetics is faster and pathways are more guided by thermodynamic driving forces and reach the stable sI structure. This agrees with simulations performed in the NVE ensemble (14), which found

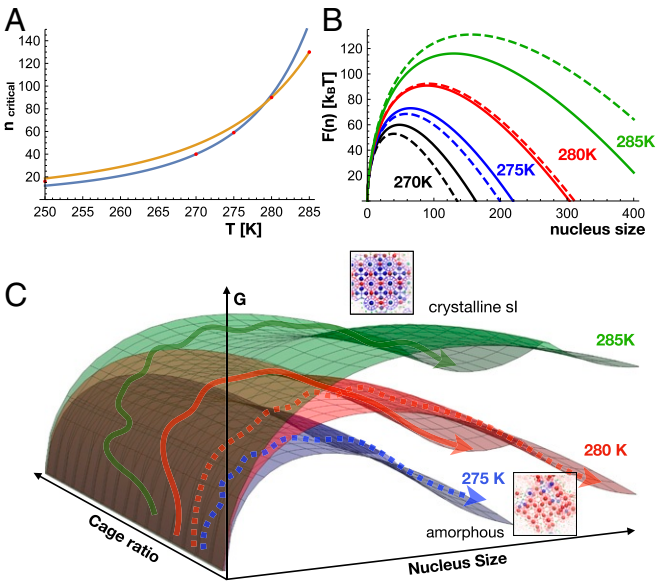


**Fig. 4.** (*A*) The temperature dependence of the critical nucleus size. Red points are the critical sizes based on the TPS results. The fit to the amorphous phase (blue curve) and the fit to the crystalline phase (orange curve) cross at 280 K. (*B*) The CNT free energy barriers for amorphous and crystalline nucleation for 270, 275, 280, and 285 K. Note that the barrier height for both amorphous and crystal phases is equal around 280 K, as expected. (*C*) The findings can be summarized in an idealized CNT free energy surface as a function of size and cage ratio for 275 K (blue), 280 K (red), and 285 K (green). The arrows indicate schematically pathways traversing from liquid to solid (dashed arrows, to amorphous phase; solid arrows, to crystalline phase). While at low temperature (e.g., 275 K), the free energy barrier to nucleate the amorphous solid is lowest, the trend is reversed at higher temperature (e.g., 285 K), where sampled pathways mostly end up in the crystalline phase. At 280 K both mechanisms are accessible.

enhancement of crystallinity in the nucleus, most likely due to the release of latent heat.

At intermediate temperatures the 2 mechanisms compete, with 1 channel leading to the thermodynamically stable sI hydrate, while the other leads to the amorphous phase or to metastable sII structures. This 2-channel picture is further confirmed by reaction coordinate analysis, which shows that at high temperature, the size and the crystal composition are important, while at low temperature only the size is important. For the high-temperature regime the reaction coordinate close to the dividing surface is characterized by both the nucleus size and the $5^{12}6^2$ cage type. In contrast, the low-temperature regime shows no dependence on cage type, but only on the nucleus size. Moreover, the optimal dividing surface for the crystallization shows an anticorrelation between the number of large cages and the cluster size. This finding corroborates with the earlier prediction for simple Lennard-Jones systems (20, 23) that crystal quality is a determinant in critical nuclei: Critical nuclei can be small but well formed or large and of lesser quality. The similar trend found here suggests that this is a more general principle in crystal nucleation.

One might think that the slow kinetics at low temperature is the reason that prevents the cage ratio from becoming sufficiently high, but even at low temperature the cage ratios can fluctuate quite a bit in the precritical regime. It is already in this regime that the selection for the final structure is made. Recent experiments on protein crystallization (34) also indicate that polymorph selection can occur precritical. This behavior seems not to follow Ostwald's step rule (35), which states that a metastable state is formed before the thermodynamically stable state is formed. However, one might argue that all paths, including the crystalline paths, begin in the metastable amorphous state before they shift to the crystalline phase, which would be in line with the step rule.

Our results are consistent with several previously reported hypotheses for methane hydrate nucleation: the labile cluster hypothesis, local structuring, cage adsorption, the blob hypothesis, and the funnel picture. All these hypotheses are aiming at describing the nucleation from different perspectives. Guo et al. (36) proposed the cage adsorption hypothesis in which methane molecules are strongly attracted to hydrate-like cages. As small $5^{12}$ cages attract methane molecules by a force similar to hydrophobic interaction, these cages should be the first to appear. Indeed, in our path density plots all pathways start in the lower left region. The labile cluster hypothesis (37) states that hydrated cages form and then associate into larger nuclei. The local structuring mechanism (38) pointed out that the guest molecules would have to come together to locally structure the surrounding water. Indeed, we observe only fully formed cages in the core of the growing nucleus. The blob hypothesis (6) is a combination of both these concepts, proposing an initial formation of amorphous aggregate that lacks long-range order of molecules, eventually growing into an amorphous clathrate, which then would transform into a crystalline SI clathrate. We observe this hydrate nucleation mechanism at 270 K and 275 K. However, for 280 K and certainly at 285 K, the amorphous blob is replaced by a more crystalline nucleus even at the initial (precritical) stage (as shown in path density plots). The formation of a blob and subsequently an amorphous phase can be related to kinetic trapping in the rugged nucleation free energy landscape. Kusalik and coworkers (7) have interpreted this behavior in terms of a potential energy funnel model as a conceptual description of molecular crystallization. It is possible to interpret our results in this light, with 2 underlying energy funnels guiding the system in either the amorphous or the crystalline phases.

Our findings contribute to answering the long-pending question about how methane hydrates nucleate at moderate conditions and give a clearer view on polymorph selection in hydrate nucleation. This insight may assist future studies that aim to improve materials synthesis and dissolving strategies of hydrates. Another result of this work is that the common notion of 2-step nucleation (such as the blob hypothesis) is not always applicable to hydrate systems. While the detailed mechanisms depend on the guest molecules, these systems can exhibit a 2-channel nucleation free energy landscape, in which the low temperature 2-step nucleation mechanism is replaced by a direct 1-step crystallization mechanism at higher temperature.

## Materials and Methods

**System Setup and Force Field.** Water was represented by the TIP4P/Ice model (15) and methane was modeled using united atom Lennard-Jones interactions ($\epsilon$ = 1.22927 kJ/mol and $\sigma$ = 3.700 Å). This combination was shown previously to mimic experimentally determined properties very well (25). Most of the MD simulations were performed using OpenMM 7.1.1 (39). As an initial configuration, a stable $4 \times 4 \times 4$ methane hydrate sI structure was generated from the unit cell definition (40) and equilibrated at a pressure of 500 bar (for details see *SI Appendix*).

**Order Parameters.** For analysis we track various CVs or order parameters (OPs). The nucleus size is measured by the MCG parameter (26), which counts methane molecules involved in the largest solid nucleus in the system, using a home-written code. In addition, we determine several other size-related OPs, such as the number of waters in the nucleus (see *SI Appendix* for details). The structure of the growing nucleus is characterized by its cage types. We analyze the cage type for each methane in the MCG-based cluster with an algorithm similar to that in ref. 41, using a home-written code. We identified 7 main types of cage structure, $5^{12}$, $5^{12}6^2$, $5^{12}6^3$, $5^{12}6^4$, $4^15^{10}6^2$, $4^15^{10}6^3$, and $4^15^{10}6^4$, where the superscript indicates the number of the

polygons made by the hydrogen-bonded water in the cage facet. The base number gives the type of polygon (4, square; 5, pentagon; 6, hexagon). The ratio of the number of $5^{12}6^2$ and $5^{12}$ cages, denoted the cage ratio (CR), is used as an indicator of sI crystallinity and has been employed previously (10, 12, 14, 42). This cage ratio is $CR = 3$ for a perfect sI, and lower than unity, $CR \leq 1$, for an amorphous or sII structure.

**Transition Path Sampling.** We perform TPS (16), employing the spring shooting algorithm (17) to generate trial trajectories. We define the liquid stable state by the absence of any cluster larger than a dimer, MCG $\leq 2$. The liquid is in equilibrium with a spherical methane reservoir, formed by methanes that do not dissolve. We define the solid stable state by MCG $\geq 420$. For efficiency we halt trajectories at MCG $\geq 200$, where all paths are postcritical and fully committed toward the solid phase. To determine the final fate of these trajectories for MCG $> 200$ we extended each successful pathway by regular MD until MCG $\geq 420$, to solidify the entire system. A trajectory is classified as amorphous (or sII), if the transition path in the $250 <$ MCG $<$ 300 range has more frames with CR $\leq$ 1 than CR $>$ 1, and crystalline otherwise.

An initial trajectory was generated by melting the $4 \times 4 \times 4$ sI crystal at 400 K and time reversing the trajectory. This unphysical trajectory was brought to the required temperature by shooting trajectories from selected intermediate frames. Two trajectories from the same frame that reached the solid and liquid, respectively, were glued together and used as an initial trajectory for TPS. This path was relaxed subsequently (*SI Appendix*). All TPS runs were executed using the OpenPathSampling (18) package. All trial paths were stored for later analysis.

**Reaction Coordinate Analysis by Likelihood Maximization.** One of the major advantages of TPS is that it allows an unbiased evaluation of the reaction coordinate. We employed the likelihood maximization (LM) approach of Peters and Trout (43) to identify pertinent ingredients in the reaction coordinate. Interpreting the TPS shooting point data as instances of a committor calculation, the LM method predicts the best model to represent those data. Starting from a predefined set of candidate order parameters, the LM method finds linear combinations of these OPs with the highest likelihood to reproduce the shooting point data. For more information see *SI Appendix*.

**ACKNOWLEDGMENTS.** We acknowledge funding from the Industrial Partnership Program "Computational Sciences for Energy Research" (Grant 15CSER051), which is financially supported by the Netherlands Organization for Scientific Research and cofinanced by Shell Global Solutions International B.V. A. thanks Sander Roet for help with the cage analysis code. We thank Daniel Bonn, Noushine Shahidzadeh, and Bernd Ensing for a critical reading of the manuscript.

1. M. R. Walsh, C. A. Koh, E. D. Sloan, A. K. Sum, D. T. Wu, Microsecond simulations of spontaneous methane hydrate nucleation and growth. *Science* **326**, 1095–1098 (2009).
2. E. D. Sloan Jr, C. Koh, *Clathrate Hydrates of Natural Gases* (CRC Press, 2007).
3. US Energy Information Administration, Annual energy outlook 2018 (US Department of Energy, Washington, DC, 2018).
4. E. D. Sloan Jr, Fundamental principles and applications of natural gas hydrates. *Nature* **426**, 353–359 (2003).
5. C. A. Koh, E. D. Sloan, A. K. Sum, D. T. Wu, Fundamentals and applications of gas hydrates. *Annu. Rev. Chem. Biomol. Eng.* **2**, 237–257 (2011).
6. L. C. Jacobson, W. Hujo, V. Molinero, Amorphous precursors in the nucleation of clathrate hydrates. *J. Am. Chem. Soc.* **132**, 11806–11811 (2010).
7. K. W. Hall, S. Carpendale, P. G. Kusalik, Evidence from mixed hydrate nucleation for a funnel model of crystallization. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12041–12046 (2016).
8. Y. Bi, T. Li, Probing methane hydrate nucleation through the forward flux sampling method. *J. Phys. Chem. B* **118**, 13324–13332 (2014).
9. R. S. DeFever, S. Sarupria, Nucleation mechanism of clathrate hydrates of water-soluble guest molecules. *J. Chem. Phys.* **147**, 204503 (2017).
10. F. Jimenez-Angeles, A. Firoozabadi, Nucleation of methane hydrates at moderate subcooling by molecular dynamics simulations. *J. Phys. Chem.* **118**, 11310–11318 (2014).
11. P. J. Smeets *et al.*, A classical view on nonclassical nucleation. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E7882–E7890 (2017).
12. M. Lauricella, S. Meloni, N. J. English, B. Peters, G. Ciccotti, Methane clathrate hydrate nucleation mechanism by advanced molecular simulations. *J. Phys. Chem. C* **118**, 22847–22857 (2014).
13. Y. Bi, A. Porras, T. Li, Free energy landscape and molecular pathways of gas hydrate nucleation. *J. Chem. Phys.* **145**, 211909 (2016).
14. Z. Zhang, M. R. Walsh, G. J. Guo, 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.* **17**, 8870–8876 (2015).
15. J. Abascal, E. Sanz, R. García Fernández, C. Vega, A potential model for the study of ices and amorphous water: Tip4p/ICE. *J. Chem. Phys.* **122**, 234511 (2005).
16. P. G. Bolhuis, D. Chandler, C. Dellago, P. L. Geissler, Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Annu. Rev. Phys. Chem.* **53**, 291–318 (2002).
17. Z. F. Brotzakis, P. G. Bolhuis, A one-way shooting algorithm for transition path sampling of asymmetric barriers. *J. Chem. Phys.* **145**, 164112 (2016).
18. D. W. H. Swenson, J. H. Prinz, F. Noe, J. D. Chodera, P. G. Bolhuis, OpenPathSampling: A python framework for path sampling simulations. *J. Chem. Theory Comput.* **15**, 813–836 (2018).
19. P. G. Debenedetti, S. Sarupria, Hydrate molecular ballet. *Science* **326**, 1070–1071 (2009).
20. D. Moroni, P. R. Ten Wolde, P. G. Bolhuis, Interplay between structure and size in a critical crystal nucleus. *Phys. Rev. Lett.* **94**, 235703 (2005).
21. F. Trudu, D. Donadio, M. Parrinello, Freezing of a Lennard-Jones fluid: From nucleation to spinodal regime. *Phys. Rev. Lett.* **97**, 105701 (2006).
22. W. Lechner, C. Dellago, P. G. Bolhuis, Role of the prestructured surface cloud in crystal nucleation. *Phys. Rev. Lett.* **106**, 085701 (2011).
23. G. T. Beckham, B. Peters, Optimizing nucleus size metrics for liquid–solid nucleation from transition paths of near-nanosecond duration. *J. Chem. Phys. Lett.* **2**, 1133–1138 (2011).
24. P. Zhang, Q. Wu, C. Mu, Influence of temperature on methane hydrate formation. *Sci. Rep.* **7**, 7904 (2017).
25. M. Conde, C. Vega, Determining the three-phase coexistence line in methane hydrates using computer simulations. *J. Chem. Phys.* **133**, 064507 (2010).
26. B. C. Barnes, G. T. Beckham, D. T. Wu, A. K. Sum, Two-component order parameter for quantifying clathrate hydrate nucleation and growth. *J. Chem. Pnys.* **140**, 164506 (2014).
27. J. Juraszek, P. G. Bolhuis, Rate constant and reaction coordinate of trp-cage folding in explicit water. *Biophys. J.* **95**, 4246–4257 (2008).
28. M. R. Walsh *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).
29. J. M. Schicks, J. A. Ripmeester, The coexistence of two different methane hydrate phases under moderate pressure and temperature conditions: Kinetic versus thermodynamic products. *Angew. Chem. Int. Ed.* **43**, 3310–3313 (2004).
30. J. Vatamanu, P. G. Kusalik, Molecular insights into the heterogeneous crystal growth of sI methane hydrate. *J. Phys. Chem. B* **110**, 15896–15904 (2006).
31. A. Zaragoza *et al.*, Competition between ices Ih and Ic in homogeneous water freezing. *J. Chem. Phys.* **143**, 134504 (2015).
32. L. C. Jacobson, V. Molinero, Can amorphous nuclei grow crystalline clathrates? The size and crystallinity of critical clathrate nuclei. *J. Am. Chem. Soc.* **133**, 6458–6463 (2011).
33. R. Anderson, M. Llamedo, B. Tohidi, R. W. Burgass, Experimental measurement of methane and carbon dioxide clathrate hydrate equilibria in mesoporous silica. *J. Phys. Chem. B* **107**, 3507–3514 (2003).
34. A. E. Van Driessche *et al.*, Molecular nucleation mechanisms and control strategies for crystal polymorph selection. *Nature* **556**, 89–94 (2018).
35. W. Ostwald, Studies on the formation and conversion of solid cokes. *J. Phys. Chem.* **22**, 289–330 (1897).
36. G. J. Guo, M. Li, Y. G. Zhang, C. H. Wu, 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).
37. E. Sloan, F. Fleyfel, A molecular mechanism for gas hydrate nucleation from ice. *AIChE J.* **37**, 1281–1292 (1991).
38. R. Radhakrishnan, B. L. Trout, A new approach for studying nucleation phenomena using molecular simulations: Application to CO2 hydrate clathrates. *J. Chem. Phys.* **117**, 1786–1796 (2002).
39. P. Eastman *et al.*, Openmm 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput. Biol.* **13**, e1005659 (2017).
40. F. Takeuchi *et al.*, Water proton configurations in structures I, II, and H clathrate hydrate unit cells. *J. Chem. Phys.* **138**, 124504 (2013).
41. L. C. Jacobson, W. Hujo, V. Molinero, Thermodynamic stability and growth of guest-free clathrate hydrates: A low-density crystal phase of water. *J. Phys. Chem. B* **113**, 10298–10307 (2009).
42. S. Sarupria, P. G. Debenedetti, Homogeneous nucleation of methane hydrate in microsecond molecular dynamics simulations. *J. Phys. Chem. Lett.* **3**, 2942–2947 (2012).
43. B. Peters, B. L. Trout, Obtaining reaction coordinates by likelihood maximization. *J. Chem. Phys.* **125**, 054108 (2006).