THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Reaction Coordinate of Incipient Methane Clathrate Hydrate Nucleation

Brian C. Barnes,[†,∥] Brandon C. Knott,[‡,∥] Gregg T. Beckham,*[,‡] David T. Wu,*[,†,§] and Amadeu K. Sum*[,†]

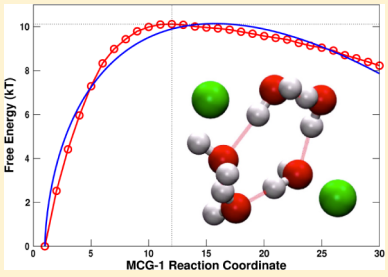
[†]Center for Hydrate Research, Chemical & Biological Engineering Department, Colorado School of Mines, Golden, Colorado 80401, United States

[‡]National Renewable Energy Laboratory, Golden, Colorado 80401, United States

[§]Chemistry and Geochemistry Department, Colorado School of Mines, Golden, Colorado 80401, United States

Supporting Information

ABSTRACT: Nucleation from solution is a ubiquitous phenomenon with relevance to myriad scientific disciplines, including pharmaceuticals, biomineralization, and disease. One prominent example is the nucleation of clathrate hydrates, multicomponent crystalline inclusion compounds relevant to the energy industry where they block pipelines and also constitute a potential vast energy resource. Despite their importance, the molecular mechanism of incipient hydrate formation remains unknown. Herein, we employ advanced molecular simulation tools (p_B histogram, equilibrium path sampling) to provide a statistical-mechanical basis for extracting physical insight into the molecular steps by which clathrates form. Through testing the Mutually Coordinated Guest (MCG) order parameter, we demonstrate that both guest (methane) and host (water) structuring are crucial to accurately describe the nucleation of hydrates and determine a critical nucleus size of MCG-1 = 16 at 255 K and 500 bar. Equipped with a validated (and novel) reaction coordinate, subsequent equilibrium path sampling simulations yield the free energy barrier and nucleation rate. The resulting quantitative nucleation process is described by the MCG clustering mechanism. This constitutes a significant advance in the field of hydrates research, as the fitness of a molecular descriptor has never been statistically verified. More broadly, this work has significance to a wide range of multicomponent nucleation contexts wherein the formation mechanism depends on contributions from both solute and solvent.



INTRODUCTION

Nucleation of solids is a fundamental physical process in numerous fields, such as biomineralization,[1] drug crystallization,[2] the semiconductor industry,[3] self-assembly of colloids,[4] zeolite formation, salt precipitation,[5] protein crystallization, and amyloid fibril formation.[6] More broadly, nucleation is a central problem for studies of vapor condensation, bubble formation in liquids, transitions between solid crystal forms, or polymer melts and metals.[7] Nucleation is an activated process in which a new phase is formed from an existing metastable phase.[8] Knowledge of the molecular mechanism will fill the gap to better influence the onset of nucleation, or its product state, through changes to thermodynamic and kinetic factors. Like other nucleation processes, clathrate hydrate nucleation from solution is difficult to study experimentally due to the length (nanometer) and time (nanoseconds) scales at which nucleation may occur. Activated processes are "rare events", as the time scale waiting for nucleation to occur (which can be as long as weeks or months, experimentally) is orders of magnitude longer than the actual transition time. Nucleation from solution may involve multiple components and might occur in multiple steps, further complicating modeling and simulation efforts. Unbiased molecular dynamics simulation may require significant simulation times to observe a single transition between a reactant and product state, if the transition is accessible at all on a simulation time scale. However, simulation using advanced, rigorous techniques is an excellent tool for studying this type of complex rare event process.

Clathrate hydrates are ice-like inclusion compounds in which one or more types of guest molecules are surrounded by a crystalline lattice of hydrogen-bound water molecules in a polyhedral arrangement.[9] Naturally occurring clathrate hydrates of natural gases, also called gas hydrates, are of wide scientific interest: they are part of the ocean's ecosystem, present in astronomical objects such as comets and other bodies in our solar system, essential components in some catastrophic climate change scenarios, and geologic repositories of energy.[10] Recently, gas production from methane hydrates has been reported in both Alaska and Japan. Gas hydrates formation in pipelines, a challenge ever-present in the flow assurance of oil/gas production, may jeopardize production and cause severe safety concerns.

A molecular-level description of the clathrate hydrate nucleation mechanism is an essential part of advancing our knowledge of gas hydrate kinetics and thermodynamics.[11,12] To date, there have been many studies of hydrate nucleation,[13−15] growth,[16] and stability[17,18] through simulation. However, the reaction coordinate for hydrate nucleation has remained

Received: August 6, 2014
Revised: October 6, 2014
Published: October 27, 2014

   dx.doi.org/10.1021/jp507959q | J. Phys. Chem. B 2014, 118, 13236−13243

elusive, and studies of hydrate formation via direct simulation are challenging, as homogeneous nucleation is a rare event under realistic conditions.[19] In the past two decades, hypotheses for the molecular mechanism have included the labile cluster hypothesis,[9] local structuring hypothesis,[20] and blob hypothesis.[14] We distinguish between a *mechanism*, or conceptual description of events in a reaction, and *reaction coordinate*, or scalar metric quantifying progress along a reaction pathway. Many hydrate order parameters have relied solely on water molecule coordinates, whether locally (a clustering metric) or globally (a system structure metric),[14,21,22] or solely on guest molecule coordinates.[23] None have been demonstrated to be a reaction coordinate (RC) in the sense of accurately quantifying progress along the nucleation pathway. Recently, we introduced the Mutually Coordinated Guest (MCG) order parameter, that quantifies the evolution of hydrate incipient formation and growth while considering the local structure of both water and methane molecules.[24]

To investigate the mechanism and rate of clathrate hydrate nucleation, we study a model system of methane hydrates via molecular dynamics (MD) simulations with both direct simulations and importance sampling. A set of multimicrosecond simulations is used to estimate a critical nucleus size. The MCG parameter is shown to be a good and appropriate reaction coordinate via the p_B histogram test, enabling the calculation of a free energy curve and kinetic prefactor. Nucleation rates using classical nucleation theory and results from the original multimicrosecond simulations are in good agreement. This work demonstrates a successful histogram test coupled with free energy calculations of nucleation from solution. It further provides quantification of the methane hydrate homogeneous nucleation pathway. These findings on hydrate nucleation may find utility in a great variety of multicomponent nucleation systems.

METHODS

The simulations and calculations performed for this study are presented in three distinct parts, the results of which will follow after the Methods section. The sets of classical molecular dynamics simulations performed consist of unbiased multimicrosecond simulations of nucleation, a statistically rigorous reaction coordinate test, and advanced sampling for free energy curve determination. Analysis in this study uses the MCG-1 form of the "Mutually Coordinated Guest" OP, for which the parameter for neighbor count, N_c, in the MCG OP is set to a value of 1. Physically, the MCG OP seeks to identify guests, in this case methane molecules, that are separated by groups of water molecules that may be well-formed five- or six-member rings in a rapidly fluctuating environment. These structures are elementary parts of the canonical 5^{12}, $5^{12}6^2$, and $5^{12}6^4$ cages that form occupied structure I and structure II hydrate crystals. Neighboring sets of qualifying MCG guests (methane molecules) are then connected via a clustering algorithm in order to identify the size of the largest MCG cluster in a system, and that value used to detect formation of incipient hydrate.

Molecular Models and Systems. Simulation cells consisted of 2944 water and 512 methane molecules, with a cylindrical methane−water interfacial geometry. The initial configuration for the multimicrosecond simulations was an initial configuration used by Walsh[25] for simulations at T = 250 K, P = 500 bar, and was created by melting a 64-unit cell system of sI hydrate at 550 K. Initial configurations for the p_B histogram test were randomly selected from MCG-1 = 16 configurations in the set of multimicrosecond simulations. Initial configurations for the equilibrium path sampling (EPS)[26,27] calculation windows were selected from p_B histogram runs that committed to liquid-like or solid-like basins, using the first occurrence of midpoint window values within 42 individual p_B histogram trajectories (21 pairs, each member of the pair committing to a different basin, and each pair set from a different p_B histogram test initial configuration). Water molecules were modeled with the TIP4P/ice water model,[28] and methane, with the unified atom Lennard-Jones model[29] using Lorentz−Berthelot combining rules for methane−water interactions. Simulation cells were approximately 5 nm per dimension.

Simulation Protocols. MD simulations were performed with Gromacs 4.[30] The Verlet leapfrog algorithm was used to integrate the equations of motion with a time step of 2 fs. The isobaric−isothermal ensemble (constant NPT) was maintained with the Parrinello−Rahman barostat (time constant of 4 ps) and Nosé−Hoover thermostat (time constant of 2 ps), at a state point of P = 500 bar and T = 255 K. The SETTLE algorithm was used to conserve the TIP4P/ice water geometry. Short-range interactions were truncated at 1 nm, and long-range electrostatics were calculated using a particle mesh Ewald algorithm with a Fourier spacing of 0.12 nm. Over 90% of the 200 unbiased simulations were conducted and analyzed to durations of 3 μs or greater. The p_B histogram trajectories were 90 ns each, with data collected from 256 initial configurations (30 trajectories shot from each, for a total of 7680 trajectories). The MCG-1 values for the liquid-like and solid-like basin cutoffs were taken to be 9 and 25, respectively. Details of these choices will be presented in the Results section; no value except the MCG-1 OP is examined in order to determine progress to reactant or product, and the MCG-1 OP is a function of molecular coordinates. In the case where a trajectory did not commit to either basin, it was not included in the p_B calculation for that configuration. Equilibrium path sampling calculations consisted of 1000 strings with 11 nodes each, and each string consisted of 500 ps of total simulation time. EPS requires trajectories to visit a window W that in our case is specified as a range of MCG-1 OP values. Many windows (independent EPS efforts) are used so that the windows overlap, and a contiguous set of MCG-1 values is available from reactant to product. Individual EPS windows had ranges of three units of MCG-1 (e.g., 6−8), with the lowest value window having a MCG-1 range of 2−4 and highest value window having a MCG-1 range of 28−30. Window matching[31] was performed at the integer overlaps between windows. Each of the 14 windows used for the EPS procedure had 21 contributing unique initial configurations, and each of their window sets were run in triplicate. Within any given window, the EPS algorithm ensures that configurations are distributed according to the equilibrium distribution.[26,27] As each configuration was run for 1000 strings, this gives a total of 882 000 individual strings of 500 ps duration each that were performed in order to generate the free energy curve. Convergence of the EPS protocol was tested by examining differences in curvature and the barrier maximum among subsets of the EPS strings (per window), and the final results presented are from the last half of the raw EPS data collected. In EPS curve construction, the expected values for the number of MCG-1 clusters of a given size N, where N is small (6 and below), were taken from non-nucleating multimicrosecond simulations, as the expected values were

greater than 1 and the formal order parameter (OP) definition only tracks the largest MCG cluster in a system. Window matching was performed for MCG-1 size 6 clusters, and results for MCG-1 cluster sizes of 7 and larger were strictly from the EPS windowing procedure. Parameters used for the empirical rate estimate were as in Kashchiev,[32] with the exception of the molar volume of methane (for fugacity calculations), from the NIST WebBook.[33] Equilibrium (P, T) for hydrate formation at their conjugate (P, T), e.g., P_e for T = 255 K, were taken from experimental data.[9] Including some exploratory calculations not presented in this report, the total computer time required to perform the multimicrosecond simulations (179 CPU-core years), p_B histogram test (206 CPU-core years), and EPS calculations (131 CPU core-years) was equivalent to approximately 820 CPU-core years on Intel "Sandy Bridge" class processors.

RESULTS

Critical Size Estimate. We performed 200 multimicrosecond MD simulations in the isobaric−isothermal ensemble with P = 500 bar and T = 255 K, as detailed above. This corresponds to a point roughly 59.5 K below liquid−hydrate−vapor coexistence for P = 500 bar and the molecular model used.[34] Only 46 of the 200 simulations nucleated and grew hydrate, indicating that we were operating at the edge of current computationally tractable conditions for unbiased, spontaneous nucleation with this model system. The critical nucleus size was estimated in terms of the MCG-1 OP using a basin commitment approach that mimics a proper committor test. Results for MCG-1 OP calculations on the multimicrosecond trajectories are presented in Figure S1 (Supporting Information). In order to estimate a critical nucleus size, clusters having various MCG-1 values were tracked for their progress to a liquid-like basin or solid-like basin within a certain Δt for relaxation of a cluster. Figure 1 (inset) provides results for Δt ranging from 20 to 100 ns. The choice of a MCG-1 value of 25 for the solid-like cutoff was made through inspection of the MD trajectories and a MCG-1 histogram to observe an OP value above which nucleation almost always proceeds, while the liquid-like cutoff of 9 was chosen to be a value near the high end of the distribution of liquid-like basin MCG-1 values, capturing its fluctuations. Figure 1 (inset) demonstrates these results, with the probability of a solid-like basin outcome denoted as P_{prod} in order to differentiate this result from a rigorous p_B histogram test. Figure 1 is a P_{prod} = 0.5 contour of the inset surface. The plateau in the bottom graph near a MCG-1 value of 16 in the 60−100 ns range identifies this as a likely critical nucleus size. Importantly, this is how we arrived at the trial MCG-1 value to use in the p_B histogram test.

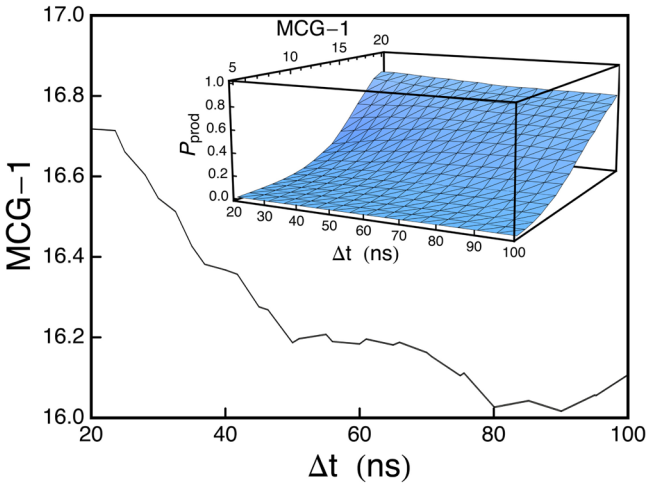


Figure 1. Inset: Probability of a MCG-1 cluster committing to a solid product outcome after a given Δt, as a function of the test cluster MCG-1 value. The MCG-1 values for the liquid-like and solid-like basin cutoffs were 9 and 25. Main: P_{prod} = 0.5 contour of the inset surface.

p_B Histogram Test. The quality of a proposed reaction coordinate can be evaluated via the p_B histogram test, in which the results of many p_B estimates shot from trial transition state configurations are compiled into a histogram.[35−37] p_B estimates are defined as the probability of a trial configuration committing to the product basin over many independent MD trajectories differing only by initial velocity randomization. A good RC will give a binomially shaped histogram (often fit by Gaussian or Beta distributions), and if the trial transition state value is near the true transition state value, the histogram will be peaked near 0.5. Figure 2 presents the histogram test results from 256 trial configurations with a MCG-1 value of 16, using 30 independent trajectories for each configuration. Figure 3a gives an example snapshot of a starting configuration, showing a MCG cluster, liquid phase, and gas phase. Parts b−d of Figure 3 are representative examples of size 16 clusters found in the initial configurations for the p_B histogram test. This histogram has a mean of μ = 0.50 ± 1.37% and a standard deviation of σ = 0.22 ± 4.97%. Relative uncertainties were calculated using a p_B histogram specific method.[38]

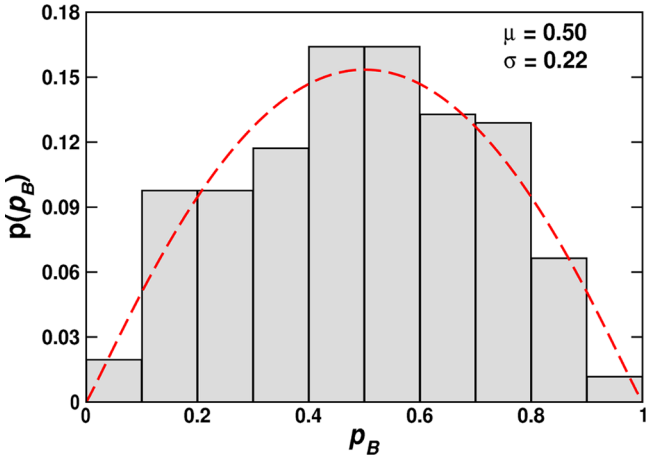


Figure 2. p_B histogram test result for MCG-1 size 16 clusters. This histogram has a mean of μ = 0.50 ± 1.37% and a standard deviation of σ = 0.22 ± 4.97%. Beta probability density function (dashed red) for a distribution with μ = 0.50 and σ = 0.22.

Free Energy Curve and Nucleation Rate. The proven suitability of MCG-1 as a reaction coordinate enables its use to compute a kinetically meaningful free energy curve and nucleation rate. Using the MCG-1 OP as a RC, we performed equilibrium path sampling calculations to map the free energy curve for methane hydrate nucleation at P = 500 bar and T = 255 K. EPS results are presented in Figure 4. The barrier maximum is at a MCG-1 value of 12 with a value of 10.1 kT. In these results, the free energies for MCG-1 values of 12 and 16 differ by only 0.2 kT. In Figure 4, a classical nucleation theory fit is shown in addition to the EPS results. The classical

dx.doi.org/10.1021/jp507959q | J. Phys. Chem. B 2014, 118, 13236−13243

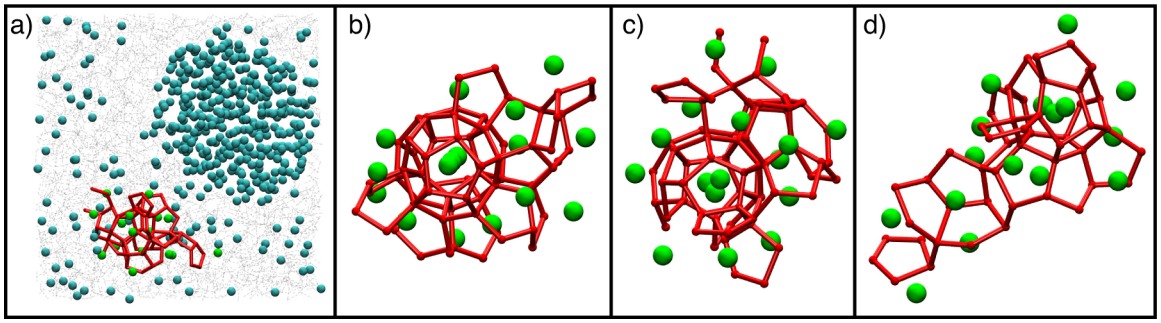


Figure 3. (a) System snapshot for the initial configuration of a representative p_B trajectory. Methane molecules are shown in cyan, except for methane molecules participating in the largest MCG cluster which are shown in green. Dynamic bonds between MCG cluster water oxygen atoms are shown in red (with a 3.5 Å cutoff) in order to illustrate cage and partial cage structure. Hydrogen bonds between other water molecules are shown in gray. (b–d) Representative structures of size 16 MCG-1 clusters used in histogram shooting. Colors as in part a.

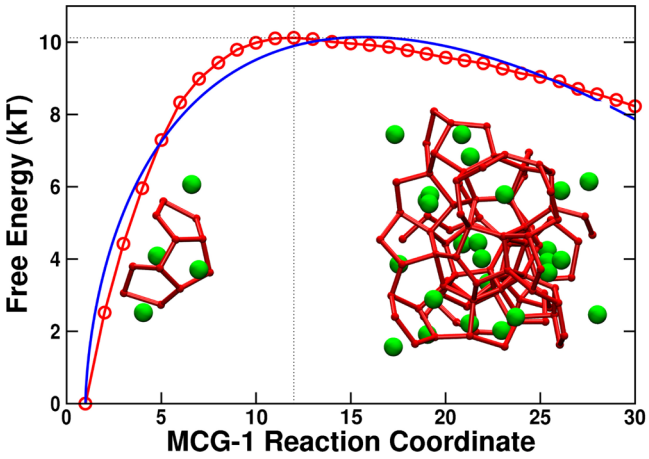


Figure 4. Free energy curve as a function of MCG-1 from equilibrium path sampling calculations (red). Best fit of the classical nucleation theory (blue line) equation, $F = a_0(n - 1)^{2/3} - a_1(n - 1)$. Light dashed lines correspond to the value of the MCG and free energy at the maximum. Images of MCG-1 clusters, size 4 and size 25, are shown as insets.

nucleation theory (CNT) form for free energy is $F = a_0(n - 1)^{2/3} - a_1(n - 1)$, where n is the value of the MCG-1 OP (this allows the general CNT form to have a value of zero for a MCG-1 "monomer" that is simply dissolved methane in water). The fit parameters are $a_0 = 5.0614\ kT$ and $a_1 = 1.3762\ kT$ with a correlation coefficient of 0.99 to the fitted EPS data. Following an expression by Kashchiev,[32] the hydrate–liquid surface energy may be estimated from the fit. In that work, the surface term a_0 is equal to $c\nu_h^{2/3}\sigma_{eff}$, where c is a shape factor (assumed to be spherical), ν_h is the per-guest volume of a hydrate building unit, and σ_{eff} is the effective specific surface energy. Our volume of a hydrate building unit is calculated as the per-guest average convex hull volume (constructed from oxygen and guest coordinates) of MCG-1 = 16 clusters used in p_B histogram shooting, and is calculated to be 0.116 nm^3. Using this expression, the surface energy is estimated to be 15.5 mJ/m^2. The expression for the classical nucleation theory rate J_{CNT} is

$$J_{CNT} = \rho Z D \exp(-\Delta F(n^\dagger)/kT) \quad (1)$$

where ρ is the monomer number density, Z is the Zeldovich factor, D is the diffusivity along the reaction coordinate, $\Delta F(n^\dagger)$ is the free energy at the barrier maximum corresponding to the $n^\dagger$ value of the OP, k is the Boltzmann constant, and T is the temperature. The parameter values used in rate calculations and results are presented in Table 1. ρ is estimated from a non-nucleating simulation, as the average density of methane in the liquid phase not participating in MCG clusters. Z is estimated from a parabolic fit to the free energy curve between MCG-1 values of 7–17. The first nanosecond of p_B histogram shots was used to calculate D, as demonstrated in Figure S2 (Supporting Information).

Table 1. Nucleation Rate Expression Quantities and Values

monomer number density, ρ (cm^{-3})	7.95×10^{20}
Zeldovich factor, Z	0.047
diffusivity, D (MCG-1^2/ns)	1.57
nucleation barrier, $F(n^\dagger)/kT$	10.1
nuc. rate from CNT, J_{CNT} (nuclei cm^{-3} s^{-1})	2.35×10^{24}
nucleating trajectories/total trajectories, N_R	46/200
average liquid phase volume, V_{liq} (nm^3)	84.9
sum of simulation time for non-nuc. traj. (ns)	597 737
nuc. rate from sim., J_{sim} (nuclei cm^{-3} s^{-1})	9.07×10^{23}

We also calculated a rate from the direct MD simulations for comparison to prior results of Walsh et al.[25] The maximum likelihood estimate for the rate is

$$J_{sim} = \frac{N_R}{(\sum_{i=1}^{N_R} \tau_i + \sum_{j=1}^{N_N} \tau_j) \cdot V_{liq}} \quad (2)$$

where N_R is the number of reactive (nucleating) trajectories, N_N is the number of remaining trajectories, τ represents the induction or total simulation time for the specific reactive or non-nucleating trajectory within the sums, and V_{liq} the average volume of the liquid phase. The calculated rate using classical nucleation theory, $J_{CNT} = 2.35 \times 10^{24}$ nuclei cm^{-3} s^{-1}, compares favorably to the rate calculated from the maximum likelihood estimate (9.07×10^{23} nuclei cm^{-3} s^{-1}). The rates differ by only a factor of 2.6.

DISCUSSION

The general methodology and insights gained from our study advance our understanding in a wide range of contexts involving nucleation from solution using system-dependent reaction coordinates. Modeling nucleation from solution presents at least two significant challenges: long induction times (rare event problem) and the necessity to maintain a constant driving force for nucleation. Advanced sampling techniques are thus essential to the study of nucleation from solution via simulation.

dx.doi.org/10.1021/jp507959q | J. Phys. Chem. B 2014, 118, 13236–13243

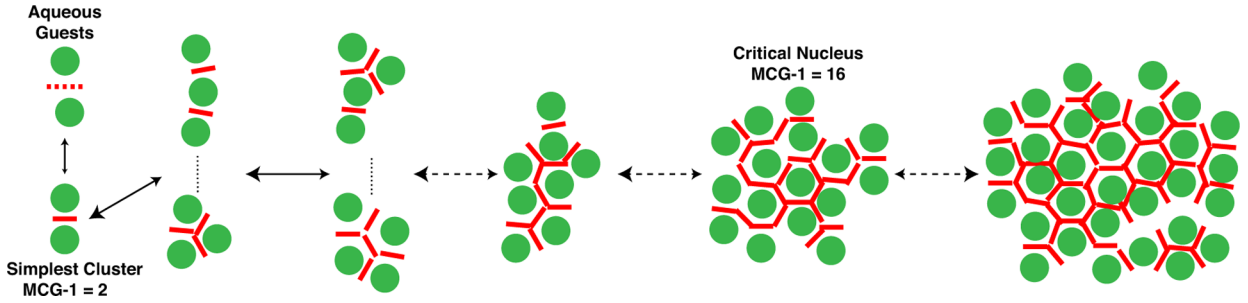


Figure 5. Conceptual picture for the mechanism of methane hydrate nucleation. Guests (green) are mutually coordinated to water clusters (red), and these structures grow and rearrange in a stepwise, nondeterministic fashion. Arrowhead size depicts the likelihood of clusters preferentially shrinking or growing. Horizontal dashed lines in the cluster growth stages represent that several intermediate steps are needed to grow from a size 4 cluster to a (critical) size 16 cluster, or from that critical nucleus to a size 32 cluster. Vertical dotted lines are used to indicate that clusters of a specific size may exist with a variety of connectivities. Clusters are not necessarily crystalline or spherical, as demonstrated by the disordered guest positions and branching connectivities. Representations are for illustrative purposes and not to scale.

In prior simulation studies, for example of Lennard-Jones systems, both the size and structure of the incipient nucleus were found to be important in describing nucleation.[39,40] Those important advances influenced the choice of order parameter used in this study, as size and structure are not always incorporated into a single order parameter. Our MCG-1 order parameter is designed to incorporate both size (number of methane molecules) and structure (geometry of solute and solvent in the cluster) in quantifying nucleation. It thus captures the addition or removal of methane to the nucleus in concert with the cooperative local rearrangement of water molecules. The OP does not presume a crystalline structure, so it serves as a RC even when disordered structures are formed in the earliest stages of nucleation—as actually observed in MD trajectories. Our results show that the critical nucleus is small (due to the high driving force for formation); thus, it is unsurprising that disordered structures initially form. The shape of the free energy curve differs from that given by classical nucleation theory (CNT) both at small sizes and in being flatter for large sizes. These differences might be attributed to the nonspherical shapes of our small nuclei, which are far from a continuum description, and curvature corrections to the interfacial energy. The flux of molecules at the nucleating interface and the "surface cloud" of a nucleus have been found to be important for nucleation in other systems.[41] By design, the MCG-1 reaction coordinate includes contributions from guest molecules that are exposed on the surface of a cluster and may not even be part of a clathrate-like structure, as long as coordinating water molecules connect the guest to the rest of the cluster. An order parameter such as a cage count metric would not capture this surface behavior.

As a culmination of this study, Figure 5 presents our conceptual interpretation for the incipient methane clathrate hydrate nucleation mechanism, denoted as Mutually Coordinated Guest clustering. The key idea of this mechanism is that hydrate nuclei grow in a stepwise manner, using water rings with adsorbed guest molecules on either side as building blocks for the nuclei and subsequent growth. Hydrate cages may be formed through this process and clusters may have significant spherical character, but branched structures are also possible and noncrystalline structures are observed on the pathway for critical nucleus formation and growth. The guests may be highly connected through mutual solvation of water rings, but only one connected neighbor is required to be considered part of the nucleating cluster. The mechanism is stochastic in nature, with clusters larger than the MCG-1 value at the free energy barrier maximum ($n^{\dagger}$) still having a significant probability to dissolve or grow. After the adsorption of a large number of mutually coordinated guests, a nucleus above the critical size is sustained and may eventually anneal to a crystalline hydrate structure. It is important to highlight that there is a multitude of pathways (molecular arrangements of a cluster) leading to the critical nucleus but with a single value of the critical nucleus size. As shown in Figure 3, the critical nucleus size comprises molecules forming about one full cage, that is, a cluster much smaller than a unit cell for structure I methane hydrate. As such, a crystalline structure cannot be defined at the critical nucleus size.

In order to describe reaction paths for rare events, an accurate RC is needed, and the committor analysis in the p_B histogram test provides a means to determine an accurate RC for clathrate hydrate nucleation. The standard deviation in a histogram test is a measure of RC accuracy.[38] With a mean μ of 0.50 and a relatively small standard deviation σ of 0.22, the test indicates that the MCG-1 OP is a valid RC. The mean indicates that the critical nucleus size is at a MCG-1 value of 16, validating the estimate from the multimicrosecond MD data set in Figure 1. The quality of the histogram we obtain is equal to or better than other published work on complex systems,[42−44] as measured by the standard deviations reported in those studies, that ranged up to σ = 0.25 (about 13% larger). This RC, used in equilibrium path sampling calculations, allows for a kinetically meaningful free energy curve to be constructed. Free energy curves for multicomponent nucleation from solution for any species are exceedingly rare but have been constructed for hard sphere and Lennard-Jones systems.[45,46] The current work determines a free energy curve for clathrate hydrate nucleation, a system of much greater structural complexity. While only one order parameter is tested in this work, we have reason to believe that, under these conditions, a p_B histogram produced using a single component OP may be of lesser quality. For example, a cage count metric may be unsuitable because the disordered clusters may often not contain a complete cage or be above the critical size when only one cage is formed. A single-component guest metric, such as the LC-SSG, may have fluctuations in the liquid basin on the order of the same size as the nucleus at the onset of nucleation.[24] Therefore, a p_B histogram may be expected to have a significantly higher variance compared to the MCG-1 RC p_B histogram. Transition path sampling[47,48] has been applied before to salt nucleation from a supersaturated solution, but the supersaturation varied along each nucleation trajectory.[49] Importantly, the inclusion of

a methane gas bubble in the current work naturally helps to maintain the solution supersaturation during the nucleation process. The solution is supersaturated with respect to the hydrate phase, as in our simulation three phases may coexist (liquid, hydrate, gas). The liquid and gas are at a metastable coexistence, and there is a driving force to form hydrate. As hydrate forms, methane will be driven out of the gas phase to replace the methane in solution phase that was consumed in hydrate formation. This continues until either the gas phase is depleted or the hydrate phase spans the simulation cell and can no longer grow. The asymmetry in the free energy curve and stochastic dynamics explain why the maximum (MCG-1 value of 12) in the free energy curve is at a different value for the RC than the critical nucleus size (MCG-1 value of 16) as determined by the p_B histogram test. If one imagines a random walk along the free energy curve, biased by its slope, then taking a few steps "down" to a smaller cluster size from the maximum of 12 quickly commits a walker to the liquid-like basin. However, taking a few steps "up" to a larger cluster size results in the walker still being on a relatively flat part of the free energy surface with the possibility of walking "back" to the barrier maximum and liquid-like basin. Therefore, the free energy maximum and p_B histogram determination of the critical nucleus size differ.

The free energy curve is reasonably fit with the CNT model, as shown in Figure 4. The empirical surface energy used by Kashchiev for hydrate nucleation (10−20 mJ/m^2 depending on the curvature of a heterogeneous interface or homogeneous nucleation case) is in good agreement with the surface energy calculated from our CNT fit (15.5 mJ/m^2). The free energy maximum (10.1 kT) was a component in the rate estimation performed via CNT. The agreement between the rate estimates from CNT and that from direct simulation is exceptional, with both on the order of 10^{24} nuclei cm^{-3} s^{-1}. These are qualitatively consistent with prior work[25] reporting a rate on the order of 10^{25} nuclei cm^{-3} s^{-1} given the higher driving forces used in that study (250 K, 500 bar), and with recent work using forward flux sampling and an entirely different model system and conditions that calculated a rate of 6×10^{21} nuclei cm^{-3} s^{-1}.[50]

Clathrate hydrate nucleation rates, inferred from experimentally measured induction time distributions, have been reported in the range of 10^{-7}−10^{-3} nuclei cm^{-3} s^{-1} corresponding to milder conditions than those utilized in the current study.[51,52] Kashchiev[53] calculated nucleation rates for a range of pressures and temperatures, with a model inspired by CNT but parametrized with experimental data. Using their model for the state point in this study, the nucleation rate is estimated to be 1.6×10^{26} nuclei cm^{-3} s^{-1} under isobaric conditions. This connects our results to predictions derived from theory, simulation, and experiment. While our calculated rates are significantly higher than those from experimental estimates, it has generally been difficult to reconcile nucleation rates from theory and experiment at similar conditions for simpler systems.[54,55] The agreement between our numerically estimated rate and the rate from the CNT model should give confidence in applying this approach to a state point at more moderate conditions. The Kashchiev model predicts an even higher rate for homogeneous nucleation at this state point, indicating that its empirical parameters may need to be revisited at high driving force. At higher temperatures or lower pressures, brute force simulation of clathrate hydrate nucleation is no longer computationally tractable with atomistic models. At more moderate conditions, it is possible that the nucleation pathway may be different, and those future studies would be of great fundamental and practical interest.

CONCLUSION

The MCG-1 RC is identified as an accurate quantitative descriptor for the molecular mechanism for methane hydrate nucleation at high driving force. We observe a pathway in which initial structures leading to the critical nucleus contain both partial and complete hydrate cages but are not defined by a crystalline structure. The fundamental building block for nucleation and growth comes in the form of water clusters that have coordinated guest molecules on each side of the cluster; this solvent/solute interaction provides a combination of size and structure metrics. These building blocks form amorphous clusters that rearrange to, and grow, clathrate cages and from which long-range order later emerges. This process is captured in the Mutually Coordinated Guest clustering mechanism for incipient hydrate nucleation, and an example of a two-step nucleation pathway. The coordination of a guest to a water cluster is through weak (hydrophobic/hydrophilic) interactions, on a small noncrystalline and potentially aspherical nucleus. The process is diffusive with a slightly sloped free energy curve near the barrier. This study provides a step forward in a long-standing problem: multicomponent nucleation from solution, using complex and realistic models with a rigorous advanced sampling methodology.

ASSOCIATED CONTENT

Supporting Information

Figures showing MCG-1 results for multimicrosecond MD trajectories and diffusivity along the MCG-1 reaction coordinate. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: Gregg.Beckham@nrel.gov.
*E-mail: dwu@mines.edu.
*E-mail: asum@mines.edu.

Author Contributions

‖B.C.B., B.C.K.: Contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This project was funded by the U.S. National Science Foundation (CHE-1125235). B.C.K. and G.T.B. thank the NREL Director's Fellowship Program for Funding. High-performance computing resources were provided by the National Renewable Energy Laboratory, Golden Energy Computing Organization (Colorado School of Mines), and Sandia National Laboratories.

REFERENCES

(1) Navrotsky, A. Energetic Clues to Pathways in Biomineralization: Precursors, Clusters and Nanoparticles. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12096−12101.

(2) Blagden, N.; Davey, R. J. Polymorph Selection: Challenges for the Future? *Cryst. Growth Des.* **2003**, *3*, 873−885.

(3) Yamauchi, N.; Reif, R. Polycrystalline Silicon Thin Films Processed with Silicon Ion Implantation and Subsequent Solid-Phase

Crystallization: Theory, Experiments, and Thin-Film Transistor Applications. *J. Appl. Phys.* **1994**, *75*, 3235−3257.
(4) Auer, S.; Frenkel, D. Prediction of Absolute Crystal-Nucleation Rate in Hard-Sphere Colloids. *Nature* **2001**, *409*, 1020−1023.
(5) Gebauer, D.; Völkel, A.; Cölfen, H. Stable Prenucleation Calcium Carbonate Clusters. *Science* **2008**, *322*, 1819−1822.
(6) Kashchiev, D.; Cabriolu, R.; Auer, S. Confounding the Paradigm: Peculiarities of Amyloid Fibril Nucleation. *J. Am. Chem. Soc.* **2013**, *135*, 1531−1539.
(7) Agarwal, V.; Peters, B. In *Advances in Chemical Physics*, Vol. *135*; Rice, S. A., Dinner, A. R., Eds.; John Wiley and Sons, Inc.: Hoboken, New Jersey, 2014; Chapter "Solute precipitate nucleation: a review of theory and simulation advances".
(8) Gebauer, D.; Kellermeier, M.; Gale, J. D.; Bergström, L.; Cölfen, H. Pre-Nucleation Clusters As Solute Precursors in Crystallisation. *Chem. Soc. Rev.* **2014**, *43*, 2348−2371.
(9) Sloan, E. D.; Koh, C. A. *Clathrate Hydrates of Natural Gases*, 3rd ed.; CRC Press: Boca Raton, FL, 2008; pp 45−111.
(10) Koh, C. A.; Sloan, E. D.; Sum, A. K.; Wu, D. T. Fundamental and Applications of Gas Hydrates. *Annu. Rev. Chem. Biomol. Eng.* **2011**, *2*, 237−57.
(11) Ribero, C. P.; Lage, P. L. C. Modeling of Hydrate Formation Kinetics: State-Of-The-Art and Future Directions. *Chem. Eng. Sci.* **2008**, *63*, 2007−2034.
(12) Barnes, B. C.; Sum, A. K. Advances in Molecular Simulations of Clathrate Hydrates. *Curr. Opin. Chem. Eng.* **2013**, *2*, 184−190.
(13) 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.
(14) Jacobson, L. C.; Matsumoto, M.; Molinero, V. Order Parameters for the Multistep Crystallization of Clathrate Hydrates. *J. Chem. Phys.* **2011**, *135*, 074501.
(15) Sarupria, S.; Debenedetti, P. G. Homogeneous Nucleation of Methane Hydrate in Microsecond Molecular Dynamics Simulations. *J. Phys. Chem. Lett.* **2012**, *3*, 2942−2947.
(16) Liang, S.; Rozmanov, D.; Kusalik, P. G. Crystal Growth Simulations of Methane Hydrates in the Presence of Silica Surfaces. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19856−19864.
(17) English, N. J.; Phelan, G. M. Molecular Dynamics Study of Thermal-Driven Methane Hydrate Dissociation. *J. Chem. Phys.* **2009**, *131*, 074704.
(18) Bagherzadeh, S. A.; Englezos, P.; Alavi, S.; Ripmeester, J. A. Molecular Modeling of the Dissociation of Methane Hydrate in Contact with a Silica Surface. *J. Phys. Chem. B* **2012**, *116*, 3188−3197.
(19) Knott, B.; Molinero, V.; Doherty, M. F.; Peters, B. Homogeneous Nucleation of Methane Hydrates: Unrealistic Under Realistic Conditions. *J. Am. Chem. Soc.* **2012**, *134*, 19544−19547.
(20) Radhakrishnan, R.; Trout, B. L. A New Approach for Studying Nucleation Phenomena Using Molecular Simulations: Application to CO_2 Hydrate Clathrates. *J. Chem. Phys.* **2002**, *117*, 1786−1796.
(21) Hawtin, R. W.; Quigley, D.; Rodger, P. M. Gas Hydrate Nucleation and Cage Formation at a Water/Methane Interface. *Phys. Chem. Chem. Phys.* **2008**, *10*, 4853−4864.
(22) Matsumoto, M.; Baba, A.; Ohmine, I. Topological Building Blocks of Hydrogen Bond Network in Water. *J. Chem. Phys.* **2007**, *127*, 134504.
(23) 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.
(24) Barnes, B. C.; Beckham, G. T.; Wu, D. T.; Sum, A. K. Two-Component Order Parameter for Quantifying Clathrate Hydrate Nucleation and Growth. *J. Chem. Phys.* **2014**, *140*, 164506.
(25) Walsh, M. R.; Beckham, G. T.; Koh, C. A.; Sloan, E. D.; Wu, D. T.; Sum, A. K. Methane Hydrate Nucleation Rates from Molecular Dynamics Simulations: Effects of Aqueous Methane Concentration, Interfacial Curvature and System Size. *J. Phys. Chem. C* **2011**, *115*, 21241−21248.
(26) Radhakrishnan, R.; Schlick, T. Biomolecular Free Energy Profiles by a Shooting/Umbrella Sampling Protocol, "BOLAS". *J. Chem. Phys.* **2004**, *121*, 2436−2444.
(27) Peters, B.; Zimmermann, N. E. R.; Beckham, G. T.; Tester, J. W.; Trout, B. L. Path Sampling Calculation of Methane Diffusivity in Natural Gas Hydrates from a Water−Vacancy Assisted Mechanism. *J. Am. Chem. Soc.* **2008**, *130*, 17342−17350.
(28) 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.* **2005**, *122*, 234511.
(29) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. Optimized Intermolecular Potential Functions for Liquid Hydrocarbons. *J. Am. Chem. Soc.* **1984**, *106*, 6638−6646.
(30) Hess, B.; Kutznera, C.; van der Spoel, D.; Lindahl, E. Gromacs 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(31) Lynden-Bell, R. M.; van Duijneveldt, J. S.; Frenkel, D. Free Energy Changes on Freezing and Melting Ductile Metals. *Mol. Phys.* **1993**, *80*, 801−814.
(32) Kashchiev, D.; Firoozabadi, A. Driving Force for Crystallization of Gas Hydrates. *J. Cryst. Growth* **2002**, *241*, 220−230.
(33) *NIST Chemistry WebBook*, NIST Standard Reference Database Number 69; Linstrom, P. J., Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg, MD, 2014; http://webbook.nist.gov (retrieved May 30, 2014).
(34) Jensen, L.; Thomsen, K.; von Solms, N.; Wierzchowski, S.; Walsh, M. R.; Koh, C. A.; Sloan, E. D.; Wu, D. T.; Sum, A. K. Calculation of Liquid Water-Hydrate-Methane Vapor Phase Equilibria from Molecular Simulations. *J. Phys. Chem. B* **2010**, *114*, 5775−5782.
(35) Du, R.; Pande, V. S.; Grosberg, A. Y.; Tanaka, T.; Shakhnovich, E. S. On the Transition Coordinate for Protein Folding. *J. Chem. Phys.* **1998**, *108*, 334−350.
(36) Geissler, P. L.; Dellago, C.; Chandler, D. Kinetic Pathways of Ion Pair Dissociation in Water. *J. Phys. Chem. B* **1999**, *103*, 3706−3710.
(37) Bolhuis, P. G.; Dellago, C.; Chandler, D. Reaction Coordinates of Biomolecular Isomerization. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5877−5882.
(38) Peters, B. Using the Histogram Test to Quantify Reaction Coordinate Error. *J. Chem. Phys.* **2006**, *125*, 241101.
(39) Moroni, D.; ten Wolde, P. R.; Bolhuis, P. G. Interplay Between Structure and Size in a Critical Crystal Nucleus. *Phys. Rev. Lett.* **2005**, *94*, 235703.
(40) Beckham, G. T.; Peters, B. Optimizing Nucleus Size Metrics for Liquid-Solid Nucleation from Transition Paths of Near-Nanosecond Duration. *J. Phys. Chem. Lett.* **2011**, *2*, 1133−1138.
(41) Lechner, W.; Dellago, C.; Bolhuis, P. G. Role of the Prestructured Surface Cloud in Crystal Nucleation. *Phys. Rev. Lett.* **2011**, *106*, 085701.
(42) Mullen, R. G.; Shea, J.-E.; Peters, B. Transmission Coefficients, Committors, and Solvent Coordinates in Ion-Pair Dissociation. *J. Chem. Theory Comput.* **2014**, *10*, 659−667.
(43) Knott, B.; Momeni, M. H.; Crowley, M. F.; Mackenzie, L. F.; Götz, A. W.; Sandgren, M.; Withers, S. G.; Ståhlberg, J.; Beckham, G. T. The Mechanism of Cellulose Hydrolysis by a Two-Step, Retaining Cellobiohydrolase Elucidated by Structural and Transition Path Sampling Studies. *J. Am. Chem. Soc.* **2014**, *136*, 321−329.
(44) Lechner, W.; Dellago, C.; Bolhuis, P. G. Reaction Coordinates for the Crystal Nucleation of Colloidal Suspensions Extracted from the Reweighted Path Ensemble. *J. Chem. Phys.* **2011**, *135*, 154110.
(45) Punnathanam, S.; Monson, P. Crystal Nucleation in Binary Hard Sphere Mixtures: A Monte Carlo Simulation Study. *J. Chem. Phys.* **2006**, *125*, 024508.
(46) Jungblut, S.; Dellago, C. Crystallization of a Binary Lennard-Jones Mixture. *J. Chem. Phys.* **2011**, *134*, 104501.
(47) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. TRANSITION PATH SAMPLING: Throwing Ropes over Rough Mountain Passes, in the Dark. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291−318.

(48) Peters, B. Recent Advances in Transition Path Sampling: Accurate Reaction Coordinates, Likelihood Maximisation and Diffusive Barrier-Crossing Dynamics. *Mol. Simul.* **2010**, *36*, 1265−1281.
(49) Zahn, D. Atomistic Mechanism of NaCl Nucleation from an Aqueous Solution. *Phys. Rev. Lett.* **2004**, *92*, 040801.
(50) Bi, Y.; Li, T. Probing Methane Hydrate Nucleation Through Forward Flux Sampling Method. *J. Phys. Chem. B* [Online early access]. DOI: 10.1021/jp503000u. Published Online: May 21, **2014**. http://pubs.acs.org/doi/full/10.1021/jp503000u (accessed May 21, 2014).
(51) Devarakonda, S.; Groysman, A.; Myerson, A. S. THF-Water Hydrate Crystallization: An Experimental Investigation. *J. Cryst. Growth* **1999**, *204*, 525−538.
(52) Abay, H. K.; Svartaas, T. M. Multicomponent Gas Hydrate Nucleation: The Effect of Cooling Rate and Composition. *Energy Fuels* **2011**, *25*, 42−51.
(53) Kashchiev, D.; Firoozabadi, A. Nucleation of Gas Hydrates. *J. Cryst. Growth* **2002**, *243*, 476−489.
(54) Filion, L.; Ni, R.; Frenkel, D.; Dijkstra, M. Simulation of Nucleation in Almost Hard-Sphere Colloids: The Discrepancy Between Experiment and Simulation Persists. *J. Chem. Phys.* **2011**, *134*, 134901.
(55) Ramsteiner, I. B.; Jensen, K. E.; Weitz, D. A.; Spaepen, F. Experimental Observation of the Crystallization of Hard-Sphere Colloidal Particles by Sedimentation onto Flat and Patterned Surfaces. *Phys. Rev. E* **2009**, *79*, 011403.