Supporting Information

MD simulations were carried out using LAMMPS.[1] The velocity Verlet algorithm was used to integrate the equations of motion. As previously noted,[2] the absence of fast rotational motions allows for integration of the equations of motion with a 10 fs timestep. Simulations were performed in the *NpT* ensemble using a Nose-Hoover thermostat and barostat with damping constants of 5 ps and 25 ps, respectively. Periodic boundary conditions were applied in all directions. Solution/gas coexistence simulations were performed with 8,000 molecules. The cross-sectional area of the simulation box is constrained to 2.5 nm by 2.5 nm and the barostat controls the pressure by varying the box length. Solution/hydrate coexistence simulations consist of 6,017 molecules, and the pressure is maintained by allowing each box dimension to expand/contract independently. These simulations begin in the completely demixed state and consist of 200 ns of equilibration, followed by 400 ns (solution/gas) or 350 ns (solution/hydrate) of data collection. Data collection consists of calculating the mole fraction of methane in the bulk solution. We define "bulk" solution as being at least 1.5 nm away from either interface.

We have slightly reparameterized the one-site model for methane originally presented by Jacobson and Molinero.[2] The intermolecular interactions have the form of the Stillinger-Weber potential (note that ϕ and γ here are distinct from the shape factor and surface energy used in the main article):

$$E=\sum_{i}\sum_{j>i}\phi_2(r_{ij})+\sum_{i}\sum_{j\neq i}\sum_{k>j}\phi_3(r_{ij},r_{ik},\theta_{ijk})$$

$$\phi_2(r_{ij})=A\varepsilon\left[B\left(\frac{\sigma}{r_{ij}}\right)^4-1\right]\exp\left(\frac{\sigma}{r_{ij}-a\sigma}\right)$$

$$\phi_3(r_{ij},r_{ik},\theta_{ijk})=\lambda\varepsilon[\cos\theta_{ijk}-\cos\theta_0]^2\exp\left(\frac{\gamma\sigma}{r_{ij}-a\sigma}\right)\exp\left(\frac{\gamma\sigma}{r_{ik}-a\sigma}\right) \quad (1)$$

where r_{ij} is the distance between particles i and j and θ_{ijk} is the angle subtended by the vectors between the positions of the $i-j$ and $i-k$ pairs of particles. The constants A = 7.049556277, B = 0.602245584, γ = 1.2, a = 1.8, and θ_0 = 109.5° are the same as for mW water.[3] The interactions for the methane model are as follows:[2]

Table 1. Interaction parameters for the coarse-grained model

	ε (kcal/mol)	σ (Å)	λ
water[a]	6.189	2.3925	23.15
methane (previous)[b]	0.340	4.08	0
methane (current)	0.140	4.08	0
water-methane[b]	0.180	4.00	0

[a]Reference 3. [b]Reference 2.

While this model has been shown to accurately reproduce the solubility of methane in water (in coexistence with methane gas) at 178 atm and 313 K,[2] we found that at high pressures (up to 1000 atm), this quantity was lower than expected. Thus, we have made the methane-methane interaction less attractive, which tends to "push" more methane out of the gas phase, increasing its solubility. To determine the magnitude of this alteration, we match Handa's calculations[4] at 500 atm and 273 K. The new ε for methane-methane interactions is 0.14 kcal/mol (in place of the previous value of 0.34 kcal/mol).

We determine the melting (dissociation) temperature T_m of this reparameterized model at 100 atm and 900 atm using the direct coexistence method.[2] Three phases coexist in these simulations: aqueous solution saturated with methane, hydrate crystal, and methane gas. At constant pressure, a series of simulations are performed at various temperatures. The temperature at which the phases coexist in equilibrium (without significant growth or dissolution of any phase) is the T_m for that pressure. At 100 atm, T_m = 292+-3 K; at 900 atm, T_m = 298+-2 K.

Due to the lack of hydrogen atoms, the mW water model diffuses faster than true water (by a factor of 4.4 at 273K and 2.8 at 298K).[3] Except for the rate of monomer attachment (which likely differs from experiments by the same factor as the diffusion coefficient), all other nucleation-relevant quantities are equilibrium quantities. The mW model has been shown to faithfully represent the thermodynamics of water crystallization (see the Supporting Information of Ref. 5). As for the dynamics, the time scale of water reorientation is orders-of-magnitude faster than the diffusion along the reaction coordinate (and the hydrate growth rate), therefore it cannot be a limiting factor for the prediction of the crystallization rates. In addition, the ratio of water/methane diffusion coefficients in the coarse-grained model is the same as in experiments.[2] The main source of error for our calculation in the crystallization rate would be due to *D,* which matches within an order of magnitude. Thus, the computed rate is probably an upper bound on experiments, and we expect the experimental value to be within an order of magnitude of our prediction.

We utilize hybrid MC/MD[6,7] to estimate the single-cage population by umbrella sampling along a coordinate defined as the number of 5- and 6- membered rings of water molecules within a 4.5 Angstrom cutoff of a methane molecule. Equilibration was not a problem for this calculation as it does not involve methane-water partitioning.

For seeded simulations, we prepare nuclei of five different sizes (consisting of about 200, 300, 400, 550, and 700 cages). For each size, we immerse seeds of hydrate prepared from the sI unit cell into solution in equilibrium with a methane bubble (the total number of particles is on the order of 70,000). From each seed four replicates are initiated. These four replicates are decorrelated from one another by approximately 20 ns of total MD run time utilizing MC/MD with a nucleus size constraint. This allowed for some equilibration/decorrelation of the nuclei while maintaining the nucleus size near the desired starting value. These were the starting configurations for our unconstrained simulations.

The diffusion coefficient along n was determined from near-critical nuclei consisting of about 300 cages. This procedure entails: 1) generating fifteen decorrelated configurations from each of the four configurations with about 300 cages by running approximately 5 ns of total MD run time utilizing MC/MD with a nucleus size constraint, 2) initiating fifty trajectories with randomized momenta from each of these fifteen configurations, , 3) running each of these fifty trajectories with unconstrained *NpT* MD

for 500 ps, calculating the mean squared displacement (MSD) $<(n(t) - <n(t)>)^2>$ every 5 ps, and 4) averaging all of these trajectories and plotting them in Figure 2.

SUPPORTING INFORMATION REFERENCES

(1) Plimpton, S. *J. Comput. Phys.* **1995**, *117*, 1-19.
(2) Jacobson, L. C.; Molinero, V. *J. Phys. Chem. B* **2010**, *114*, 7302-7311.
(3) Molinero, V.; Moore, E. B. *J. Phys. Chem. B* **2009**, *113*, 4008-4016.
(4) Handa, Y. P. *J. Phys. Chem.* **1990**, *94*, 2652-2657.
(5) Moore, E. B.; Molinero, V. *Nature* **2011**, *479*, 506-U226.
(6) Duane, S.; Kennedy, A. D.; Pendleton, B. J.; Roweth, D. *Phys. Lett. B* **1987**, *195*, 216-222.
(7) Forrest, B. M.; Suter, U. W. *J. Chem. Phys.* **1994**, *101*, 2616-2629.