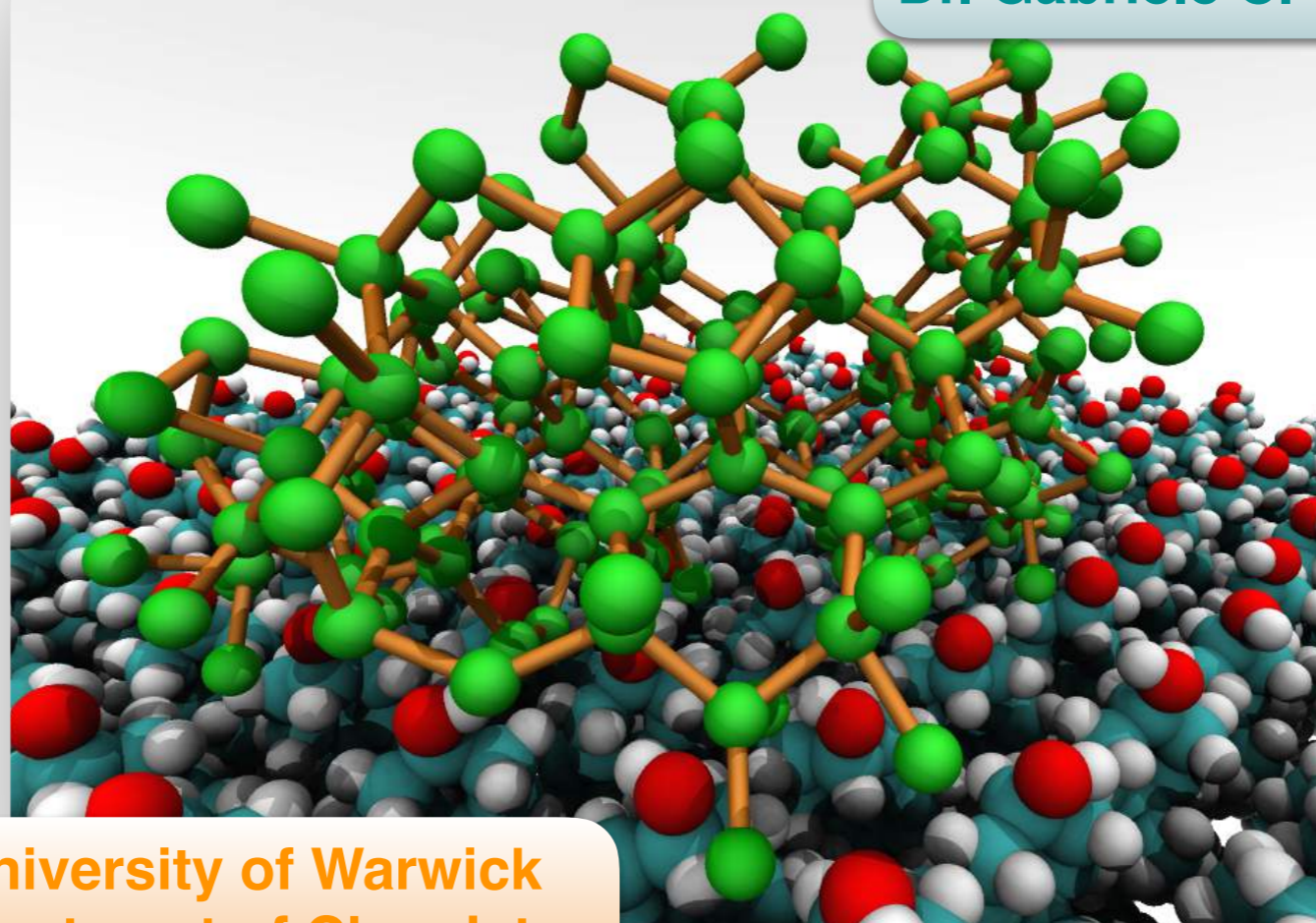# Predicting Crystal Nucleation Rates via Atomistic Simulations

The Case of Ice Nucleation

Dr. Gabriele C. Sosso

**University of Warwick
Department of Chemistry**

G.Sosso@warwick.ac.uk

October 2017

Any questions?

Outline

- **The nucleation of crystals**

  - Why do we care?

- **Experiments and Simulations**

  - The irony of the timescale problem

  - Building a bridge: the role of the nucleation rate
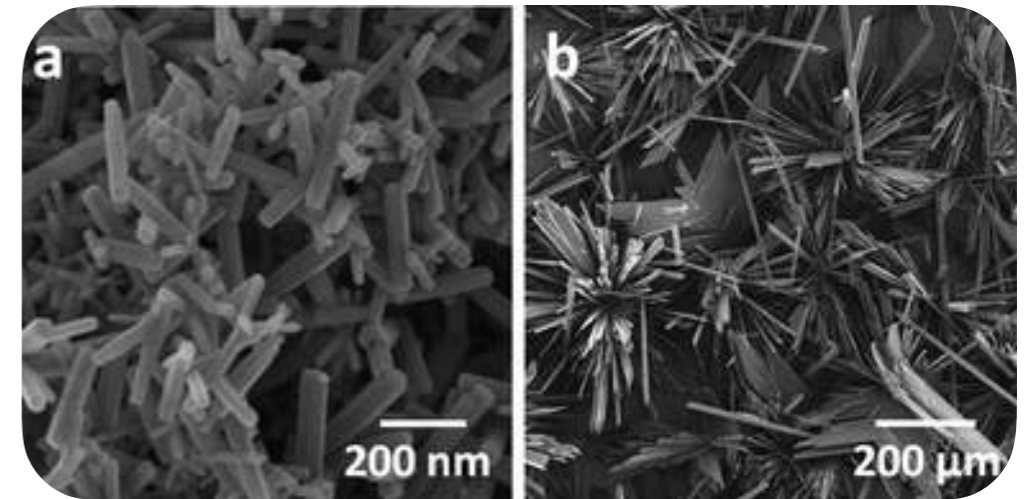
- **The computational options**

  - Free energy-based methods

  - Path sampling methods
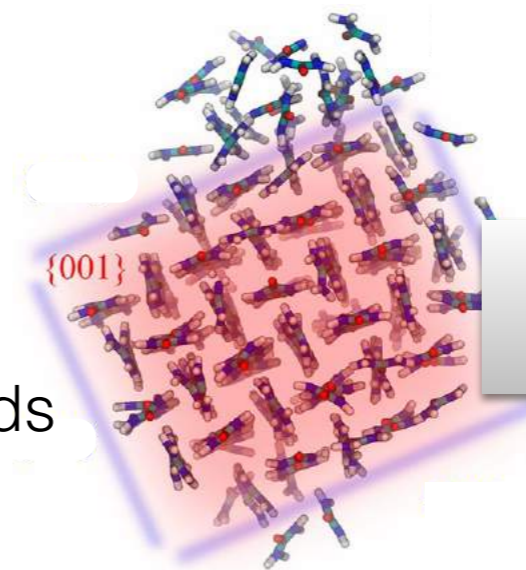
- **Embracing the failure**

  - Respect the algorithms, blame the force fields

  - The case of ice nucleation

  - Heterogeneous ice nucleation: the going gets tough
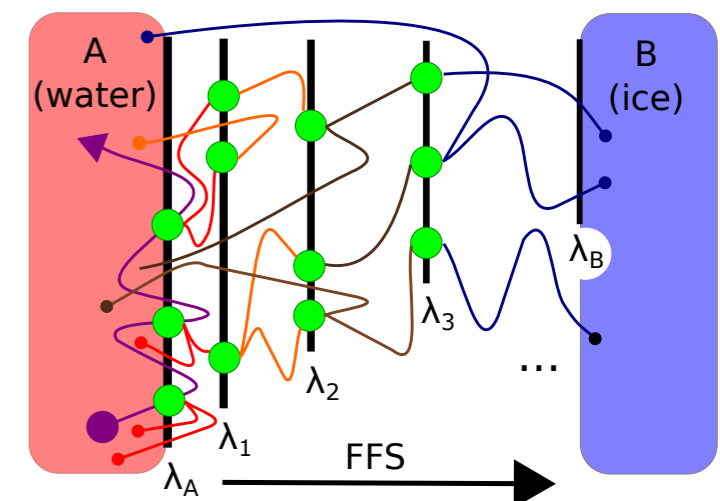
- **Conclusions**

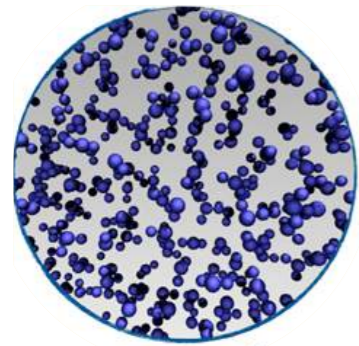 - Where are we going with this?

  - Keep calm & collaborate

Wang, Y.-W., and Meldrum, F.C. (2012). J. Mater. Chem. 22, 22055–22062.

Salvalaglio, M., Mazzotti, M., and Parrinello, M. (2015). Faraday Discuss. 179, 291–307.
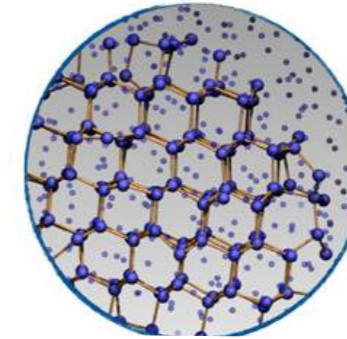
Crystallization:
Ordo ab Chaos

Disorder ➡ Order
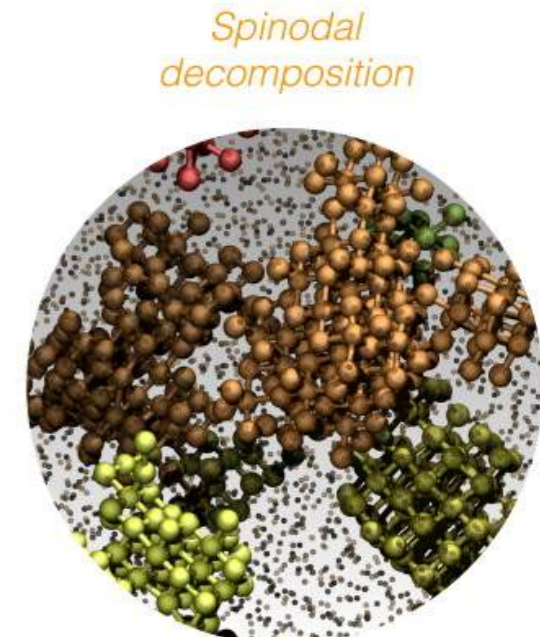
- Supercooled liquid
- Supersaturated solution ➡ Crystal
- Amorphous solid

Crystallization

Crystal nucleation   Crystal growth

$\Delta G^*_\mathcal{N}$

$\Delta G^*_{\mathcal{N},\text{het}}$

Dissolution

Crystal growth

Supercooled Liquid

Crystalline phase

n^*

n^*_het

n

ΔG_N

Homogeneous nucleation

Heterogeneous nucleation

Spinodal decomposition

Sosso, G.C., Chen, J., Cox, S.J., Fitzner, M., Pedevilla, P., Zen, A., and Michaelides, A. (2016). Chem. Rev. 116, 7078–7116.

The nucleation of crystals
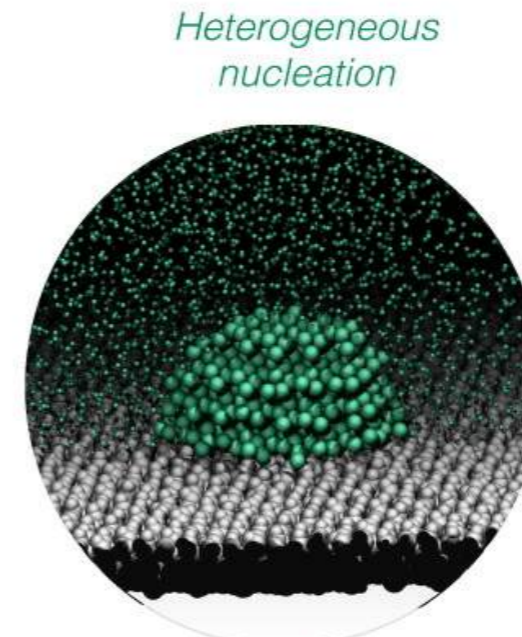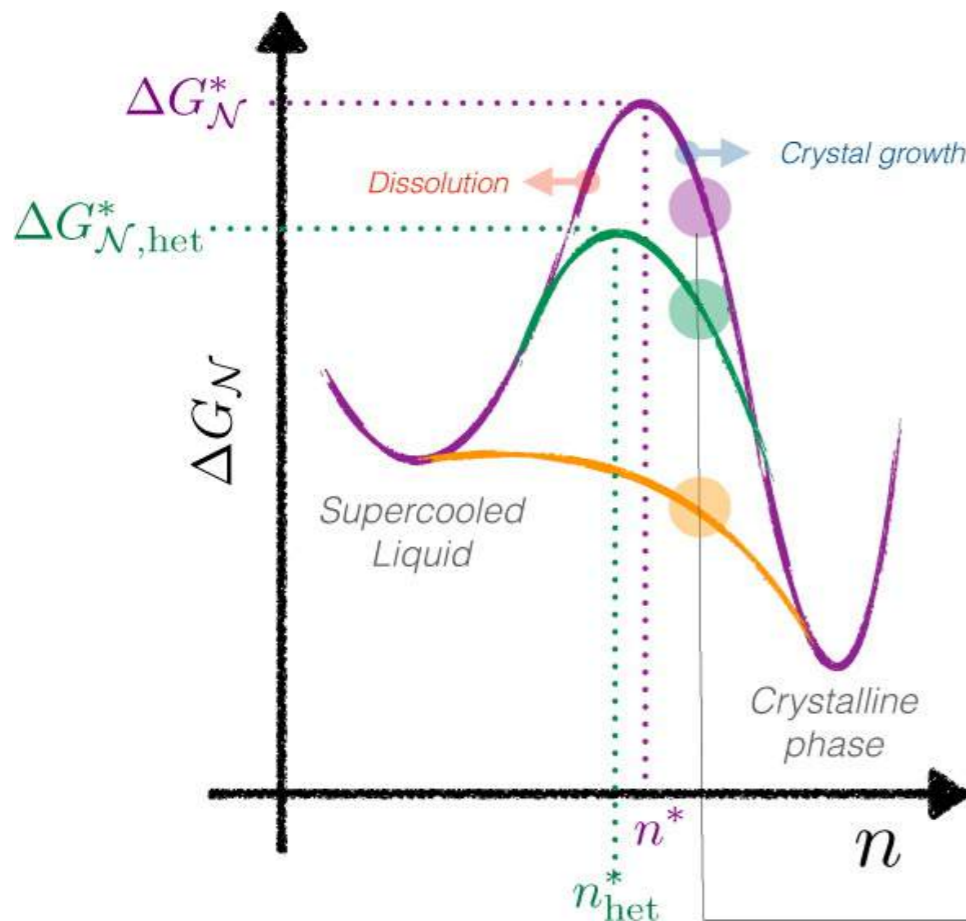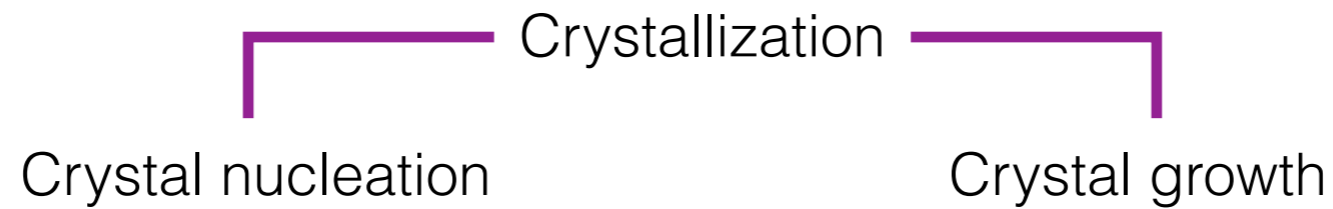
Why do we care?

- **The formation of ice**
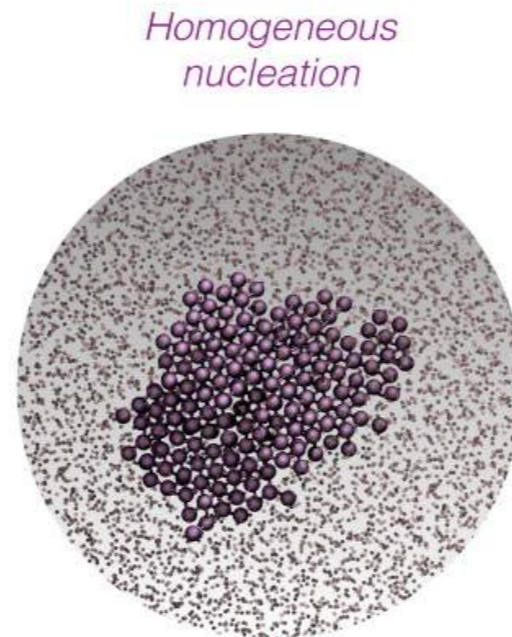
  - Atmospheric science

  - Cryobiology

- **Hydrocarbon clathrates**
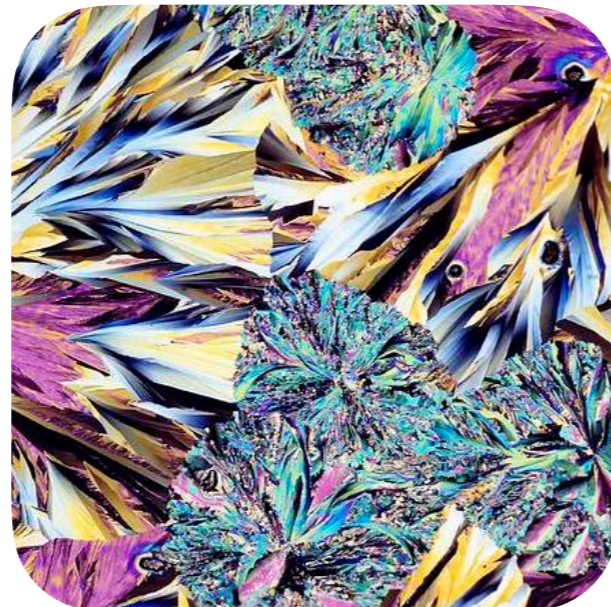
  - Oil industry

- **Molecular crystals from solutions**

  - Pharmaceuticals

  - Drug design

- **Living things**

  - Biomineralization

  - Alzheimer disease

WARWICK
THE UNIVERSITY OF WARWICK

Classical Nucleation Theory

$$\Delta G_{\mathcal{N}} = \underbrace{4\pi r^2 \gamma_{\mathcal{S}}}_{\text{surface term}} - \underbrace{\frac{4\pi}{3} r^3 \Delta\mu_{\mathcal{V}}}_{\text{volume term}}$$
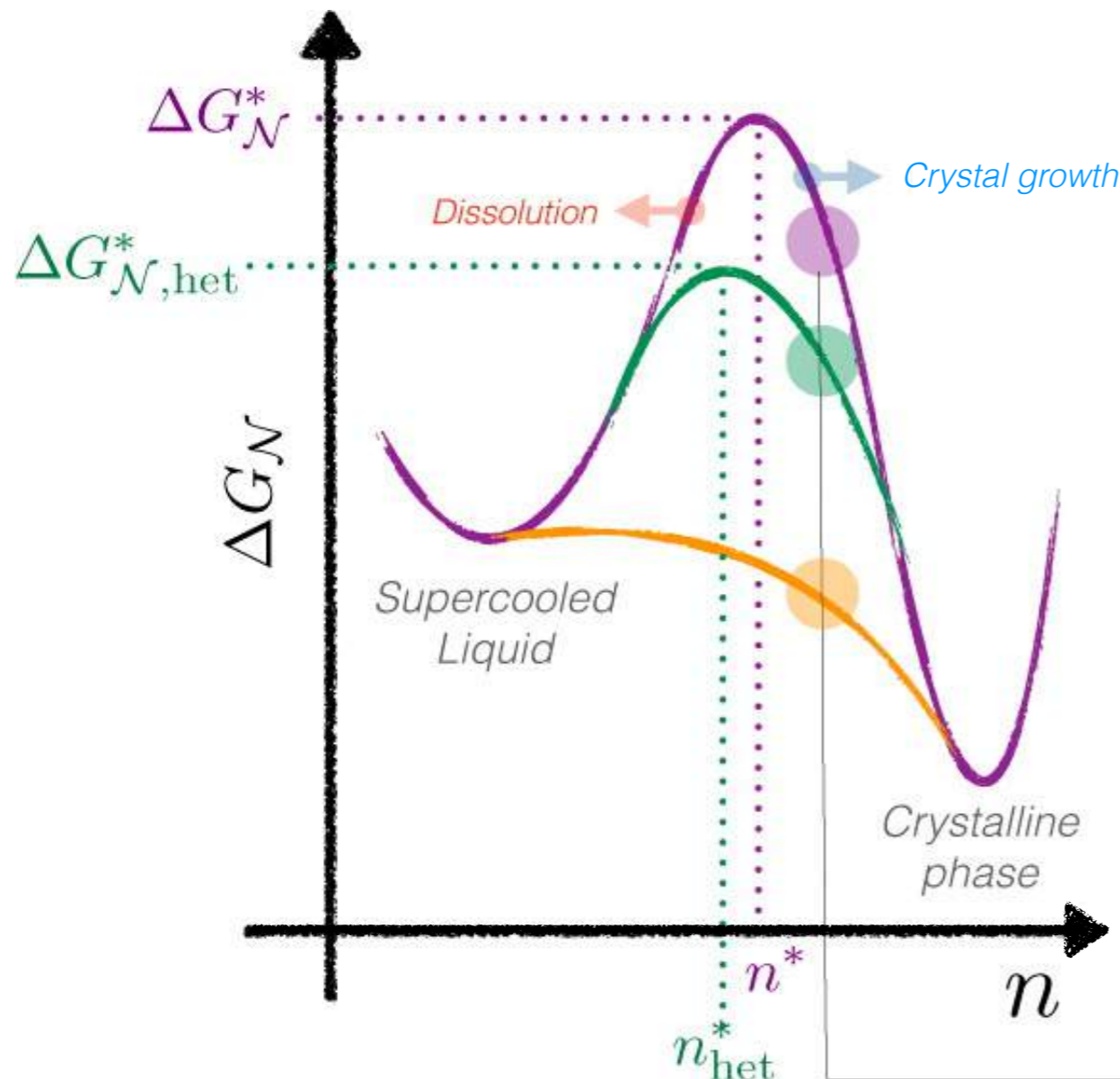
$$n^* = \frac{32\pi\rho_C}{3} \frac{\gamma_{\mathcal{S}}^3}{\Delta\mu_{\mathcal{V}}^3} \Rightarrow \textit{Critical nucleus size}$$

$$\Delta G_{\mathcal{N}}^* = \frac{16\pi}{3} \frac{\gamma_{\mathcal{S}}^3}{\Delta\mu_{\mathcal{V}}^2} \Rightarrow \textit{Free energy barrier}$$

$$\mathcal{J} = \mathcal{J}_0 \exp\left(-\frac{\Delta G_{\mathcal{N}}^*}{k_{\mathrm{B}} T}\right) \Rightarrow \textit{Nucleation rate}$$

Kinetic prefactor   *Free energy barrier*

Experiments

WARWICK
THE UNIVERSITY OF WARWICK

Nucleation time & length scales:
nanometers & nanoseconds

	Electron Microscopy	Scanning Probe Microscopy	X-Ray Experiments	Optical Microscopy	Chamber Experiments	Spectroscopy	Calorimetry
min	Environmental SEM	AFM	Powder diffraction	High-speed imaging	Flow chamber	FTIRS	DSC
	cryo-TEM SMRT-TEM		Ambient pressure XPS	Confocal microscopy	Cloud chamber	Analytical Ultracentrifugation	
ps			Fs x-ray scattering	Standard microscopy			

temporal resolution

nm   µm   mm   cm

spatial resolution

There are colloids…

Experiments have no access to e.g. the critical nucleus size

The *only* quantity that we can (hope to) compare between experiments and simulations:
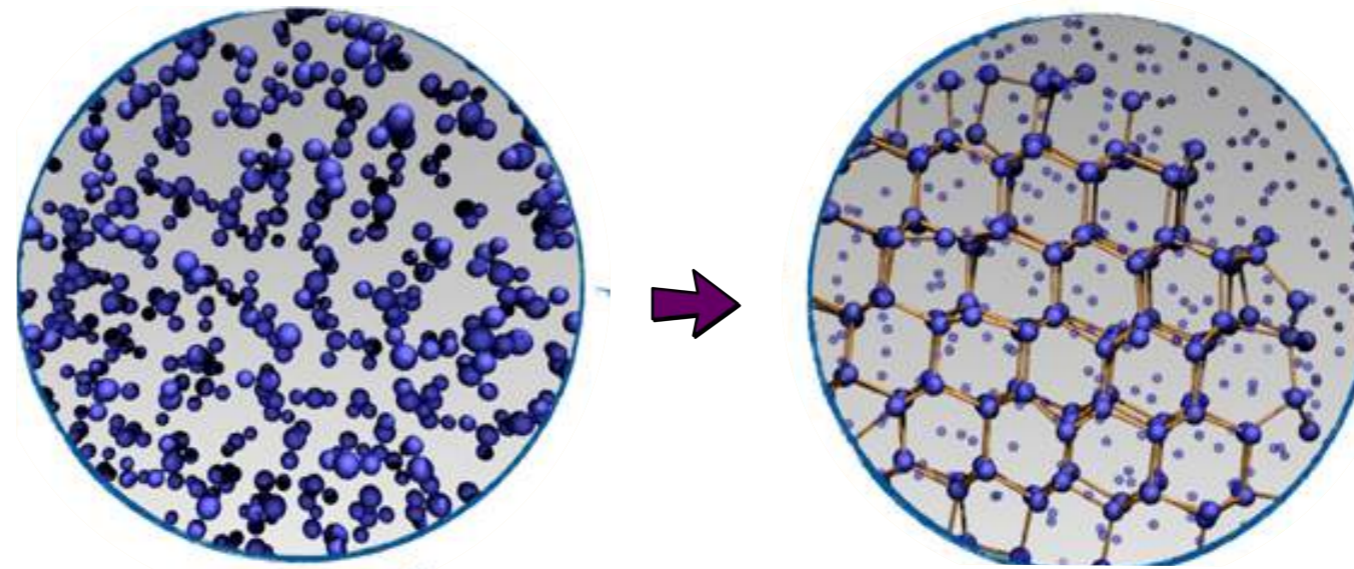
The nucleation rate

$$\mathcal{J} = \mathcal{J}_0 \exp\left(-\frac{\Delta G_N^*}{k_B T}\right)$$

Simulations

Molecular simulations could help!
nanometers & nanoseconds sounds about right…

Classical molecular dynamics (MD) simulations
Rely on classical/empirical/analytical force fields/interatomic potentials
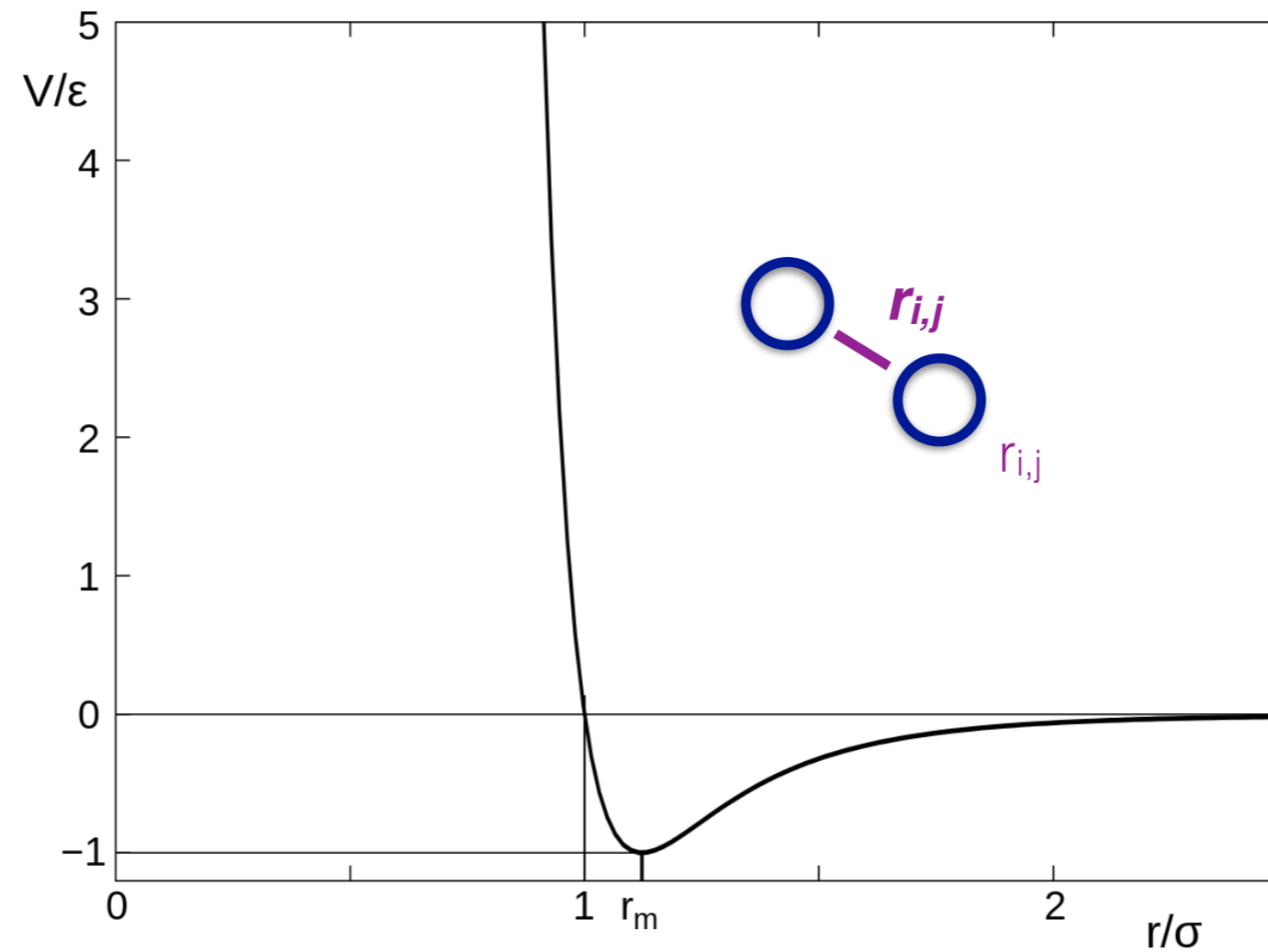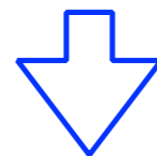
The timescale problem

seconds → nanoseconds

The timescale problem

Nothing   Nucleation

Simulations: **no way**
Experiments: OK!

Simulations: OK!
Experiments: **no way**
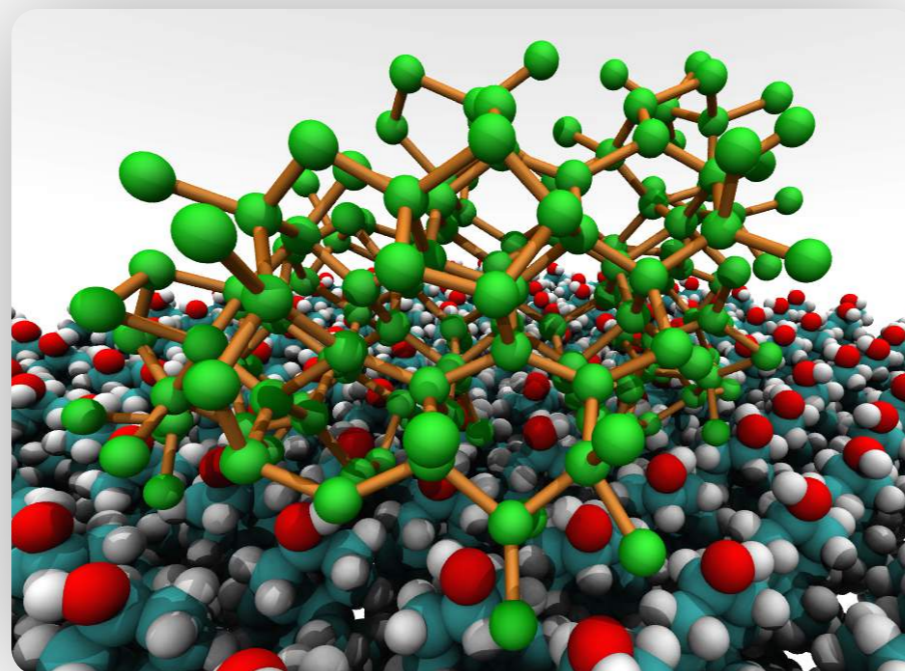
The computational problem(S)

Molecular simulation of crystal nucleation

Microscopic understanding:
- Mechanics
- Thermodynamics
- Kinetics

HOWEVER

- The timescale problem: nucleation is a *rare event* [it means you need statistics]
- Classical force fields are often not good enough (think heterogeneous crystal nucleation)

Sosso, G.C. et al. (2016). Crystal Nucleation in Liquids: Open Questions and Future Challenges in Molecular Dynamics Simulations. Chem. Rev. 116, 7078–7116.

Brute force molecular dynamics

WARWICK
THE UNIVERSITY OF WARWICK

Sometimes, brute force molecular dynamics simulations *are* an option:

- Hard spheres and Lennard-Jones liquids
- "Simple" liquids (metallic liquids, typically strong supercooling and *massive* computational effort)
- Coarse grained simulations (famously, mW water)
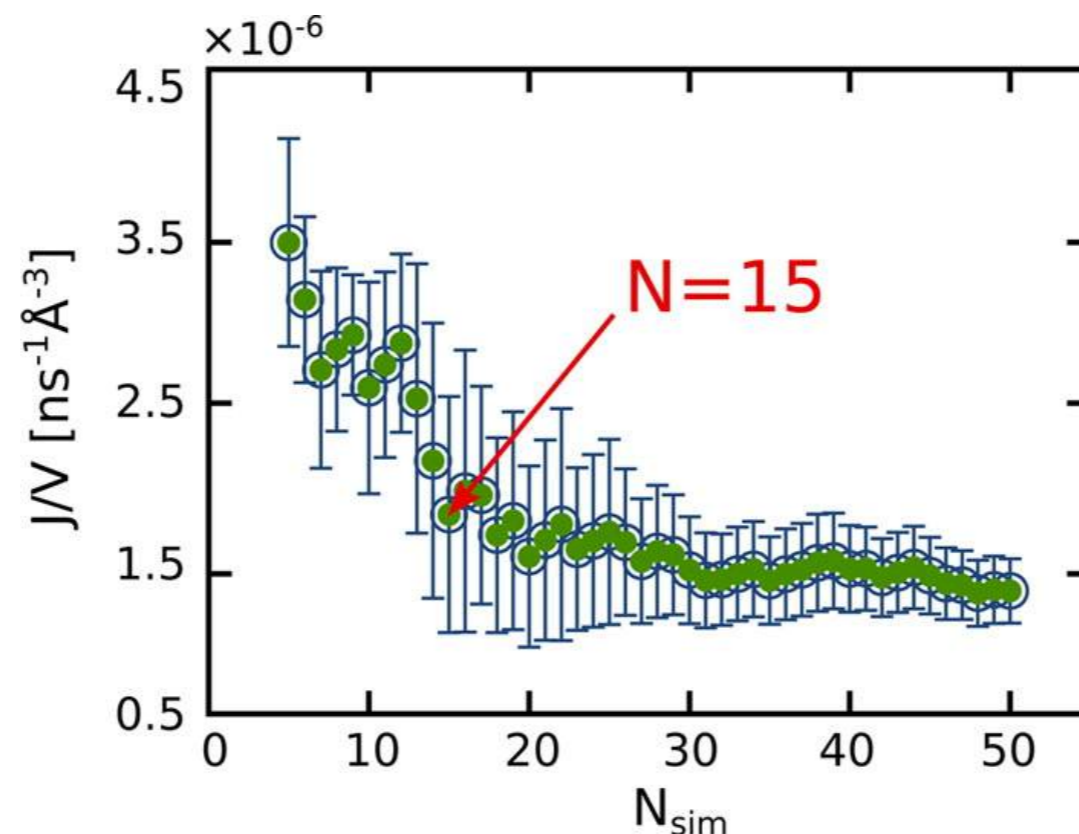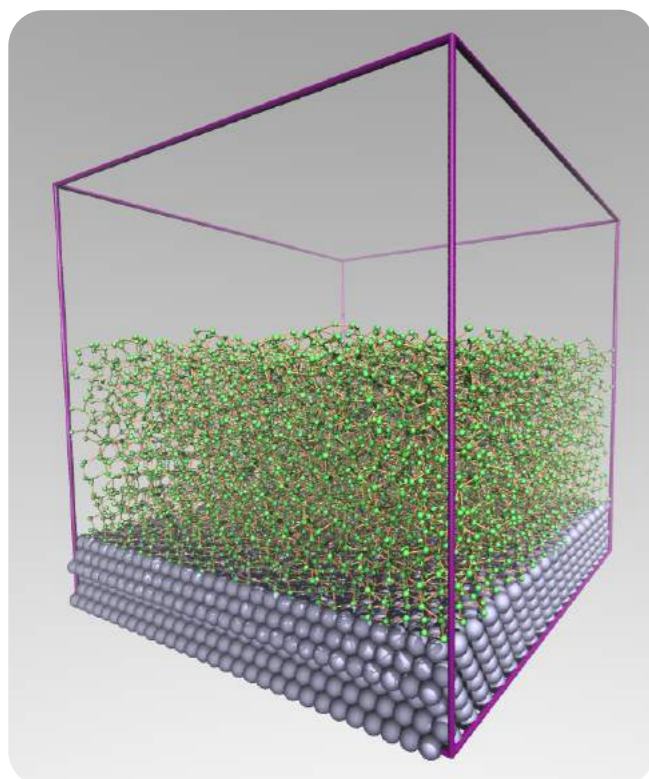
• **Time**

The system must be allowed to evolve in time until spontaneous fluctuations lead to a nucleation event.

• **Size**

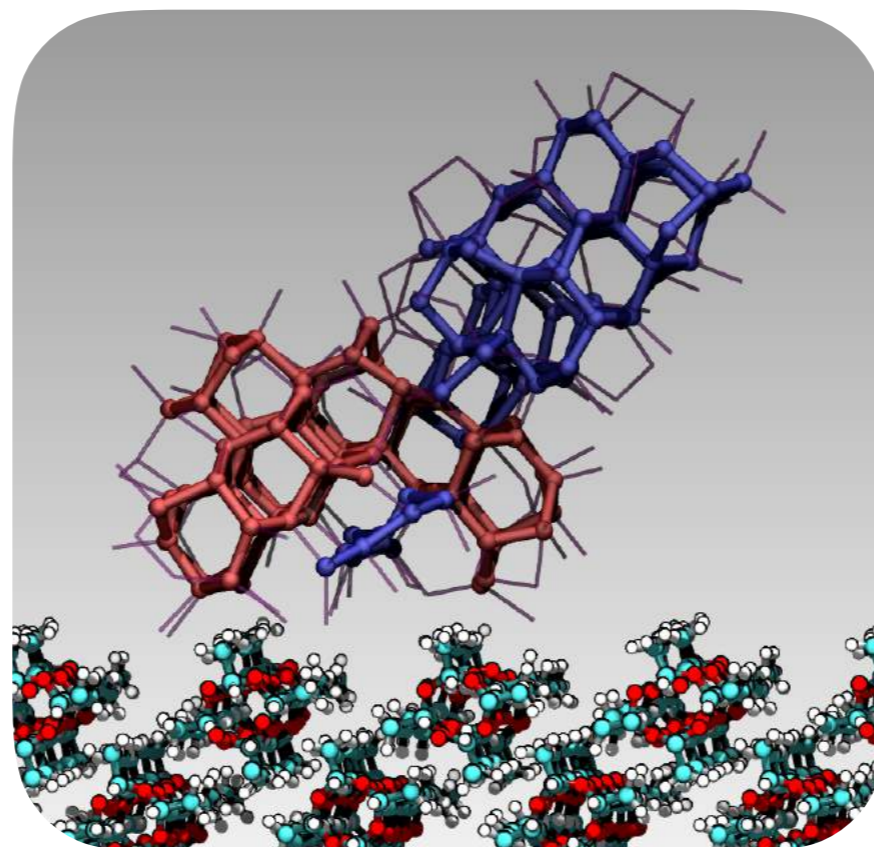The system size must be significantly larger than the critical nucleus.

• **Statistics**

Significant statistics of nucleation events must be collected.

Complex systems
*(heterogenous nucleation,
nucleation from solution…)*

NO WAY

Fitzner, M., Sosso, G.C., Cox, S.J., and Michaelides, A. (2015).
J. Am. Chem. Soc. 137, 13658–13669.

Enhanced sampling

Complex systems/scenarios require…


Enhanced sampling techniques

1. **Free energy methods**
2. **Path sampling methods**
3. **(Seeded molecular dynamics)**

What are we looking for?

- Speed up simulations (so that we can observe nucleation event**S**)
- Avoid tempering with the natural evolution of the system (dynamics & mechanism)
- Get the microscopic mechanism and the kinetics of nucleation (nucleation rate)

Order parameters

Order parameters

You *assume* you can describe nucleation using one
(or a few. Or a lot.) order parameter(s)
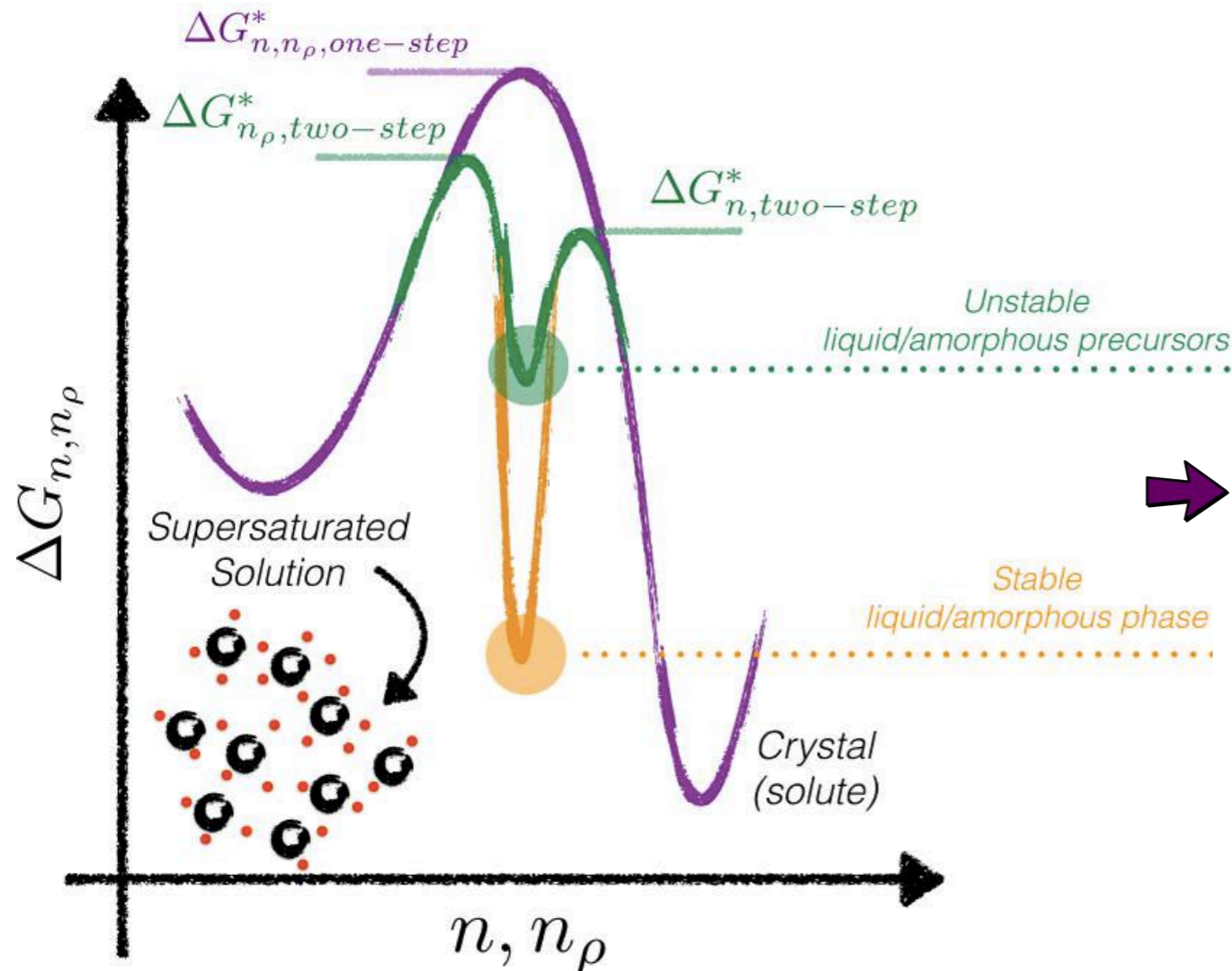
→ *i.e. coarse graining the free energy surface*

e.g. Ice nucleation:
The order parameter is the number of water molecules within the largest ice nucleus (a plethora of options exist!)

In reality…
- More than one structural degree of freedom
- Density (nucleation from solution)
- The substrate (heterogeneous nucleation)
- Two-step nucleation

Free energy methods

WARWICK
THE UNIVERSITY OF WARWICK
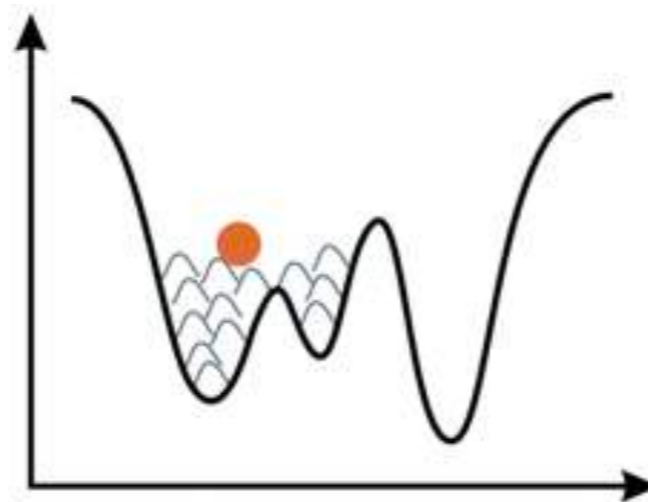
The usual suspects:
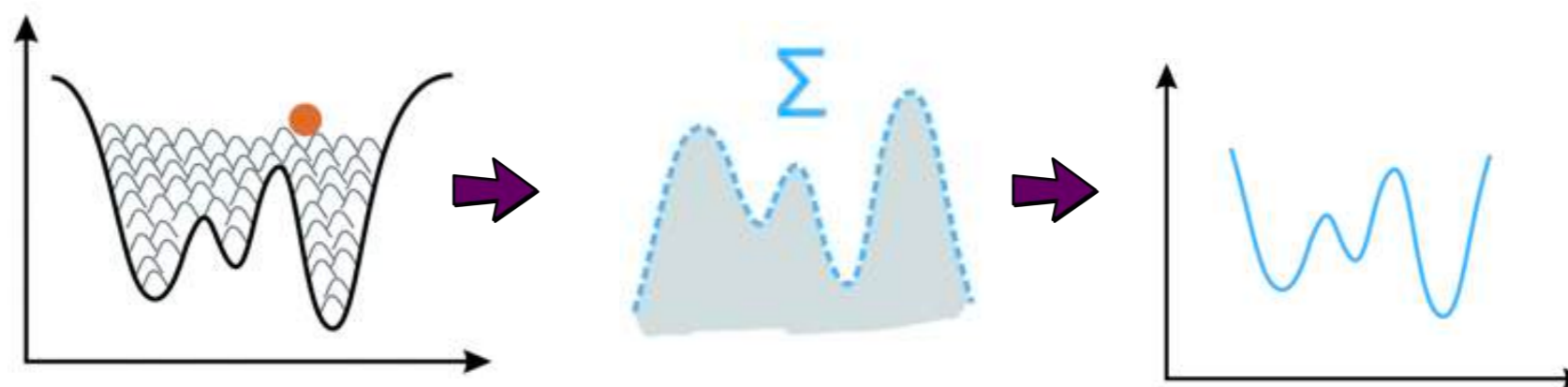- Umbrella sampling
- Metadynamics

The idea:
Add an external bias potential,
driving the system on top of the free energy barrier

*Free energy
(unknown!)*

*Laio, A., and Parrinello, M. (2002).
Proc Natl Acad Sci U S A 99, 12562–12566.*

*Order
parameter*

Σ

What do you get:
- The free energy barrier
- The critical nucleus size
- The mechanism? *Maybe…*
- The dynamics (kinetics prefactor)? *Maybe…*

Free energy methods ++

What you *don't* get:
The kinetics, i.e. the nucleation rate

There are ways to get rates starting from the free energy surface:
- Bennett-Chandler (transition state theory-based methods) [incredibly costly]
- *Rates from metadynamics* [works for either simple or *fast* systems]
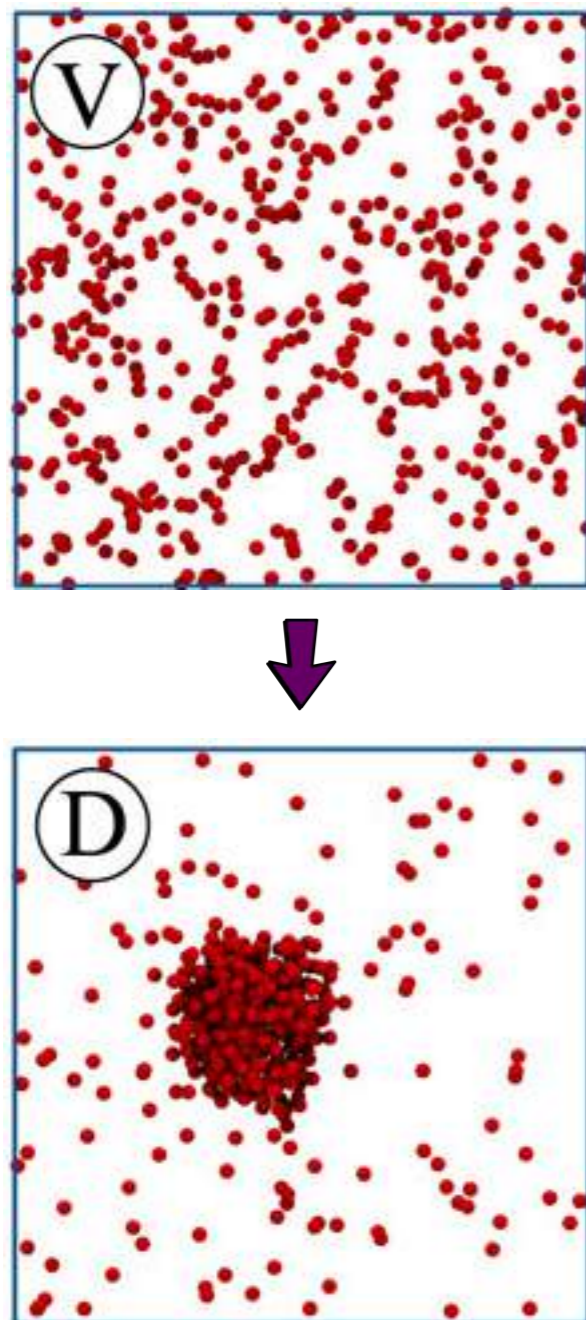
Nucleation of a liquid droplet from vapour (Lennard-Jones)

Salvalaglio, M., et al (2016). The Journal of Chemical Physics 145, 211925.
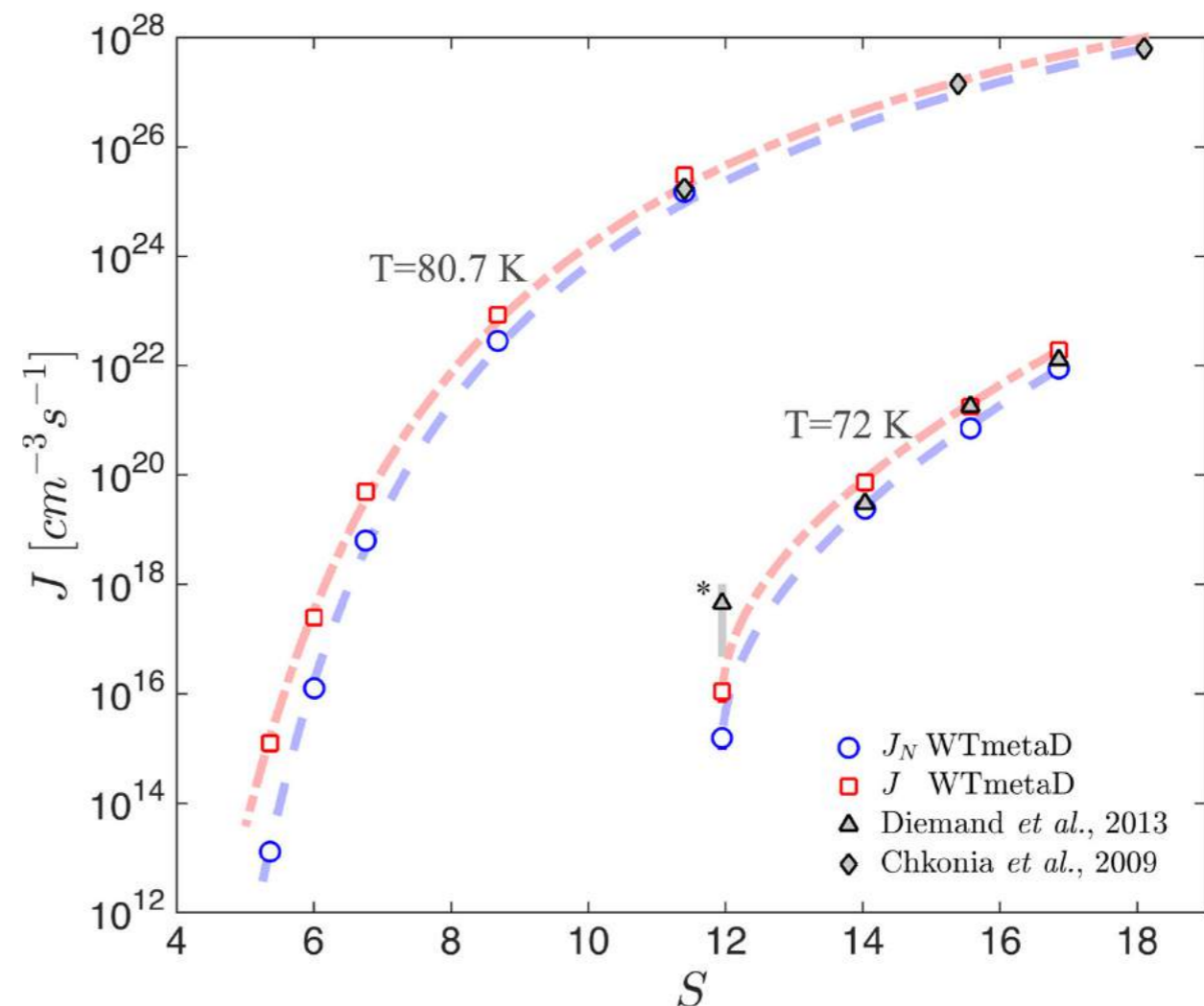
Works if *no bias is added on top of the free energy barrier*

Path sampling methods

The usual suspects:
- (Transition path sampling)
- Transition interface sampling
- Forward flux sampling

> Allen, R.J., Frenkel, D., and Wolde, P.R. ten (2006).
> The Journal of Chemical Physics 124, 194111.

The idea:

- The path(S) from A (liquid) to B (crystal) are described in terms of an order parameter
  [e.g. ice nucleation: λ is the number of water molecules in the largest ice nucleus]
- We want to know the probability P(B|A) of going from A (e.g. water) to B (e.g. ice)

$$P(\lambda_B | \lambda_A)$$
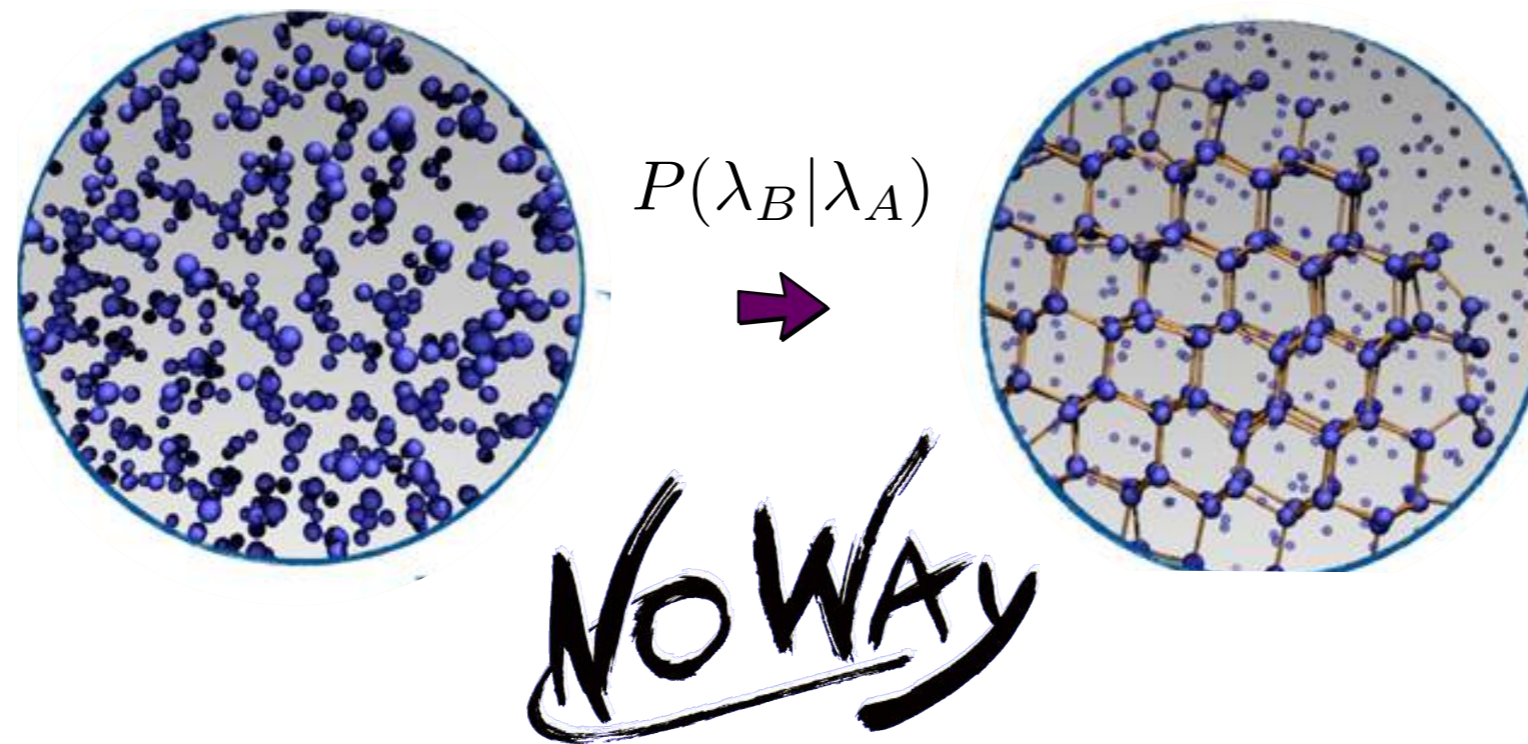
This probability is (usually) incredibly small

Forward flux sampling

Divide the path into a series of interfaces
*Each interface has an different, increasing value of λ*
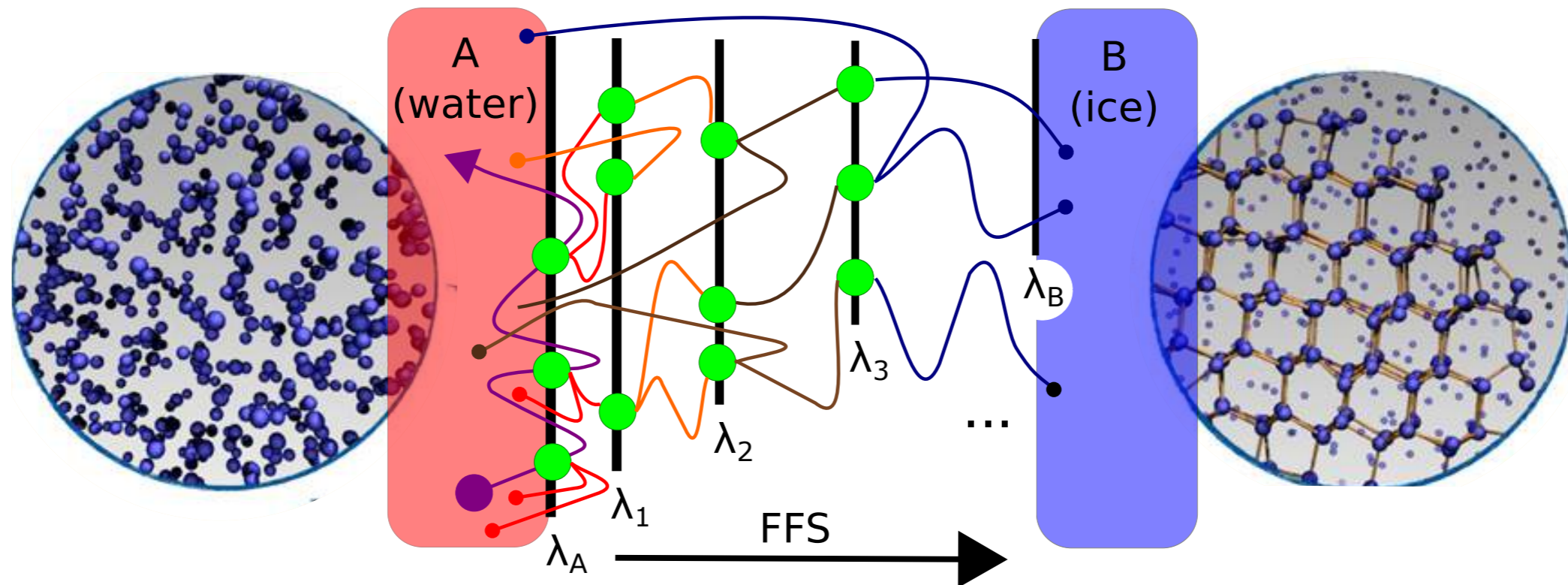
$$P(\lambda_B | \lambda_A) = \prod_{i=1}^{N_\lambda} P(\lambda_i | \lambda_{i-1})$$

If you choose your interfaces close enough, you can (with some effort) compute each of the

$$P(\lambda_i | \lambda_{i-1}) \text{ ?}$$

Forward flux sampling

WARWICK
THE UNIVERSITY OF WARWICK

We start by looking at the natural fluctuations of the system
Long unbiased molecular dynamics run

$$\Phi_{\lambda_A} = \frac{N.\text{ of crossings}}{\text{Simulation time} \cdot \text{ Volume}}$$
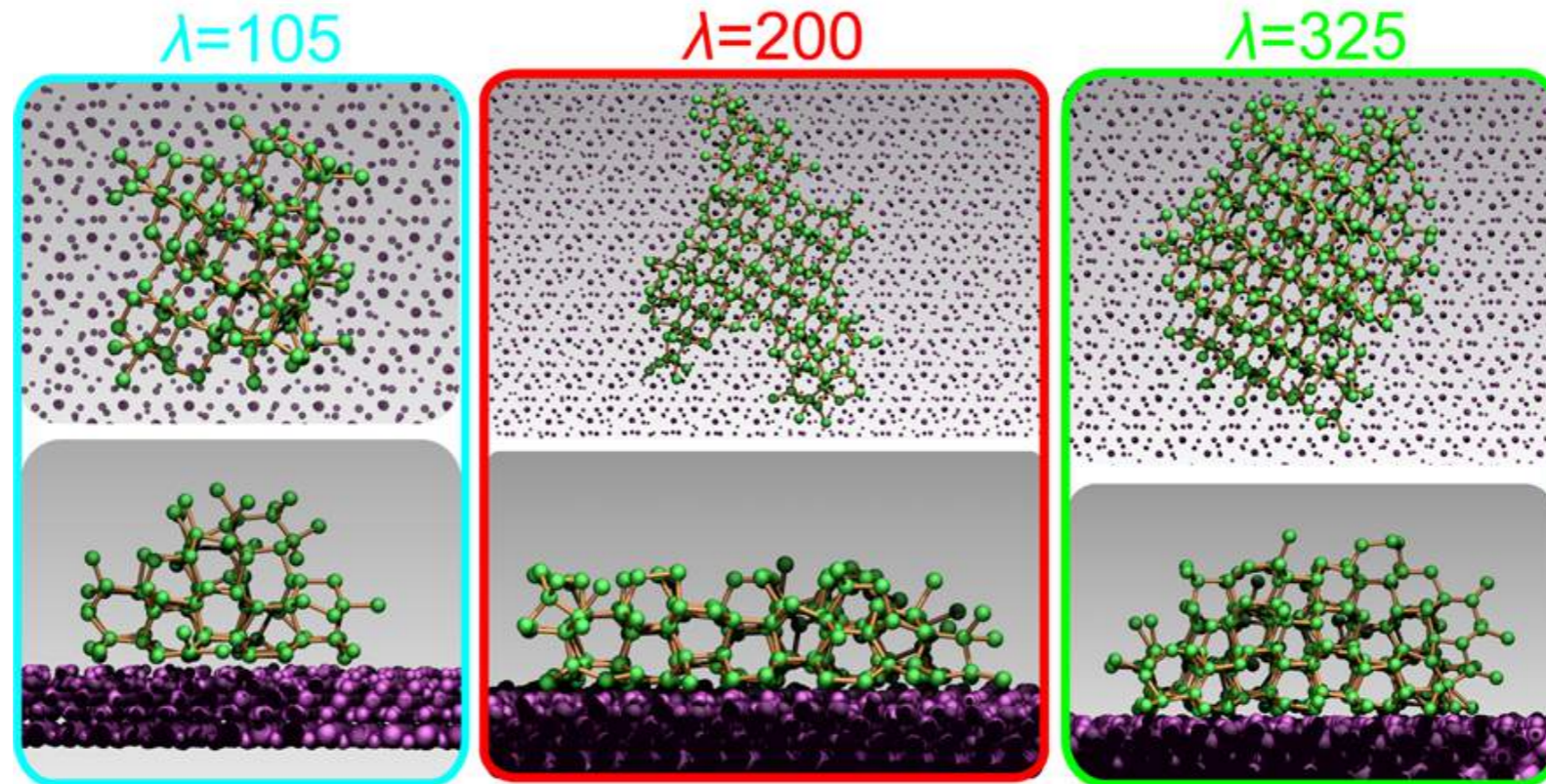
Initial *flux*
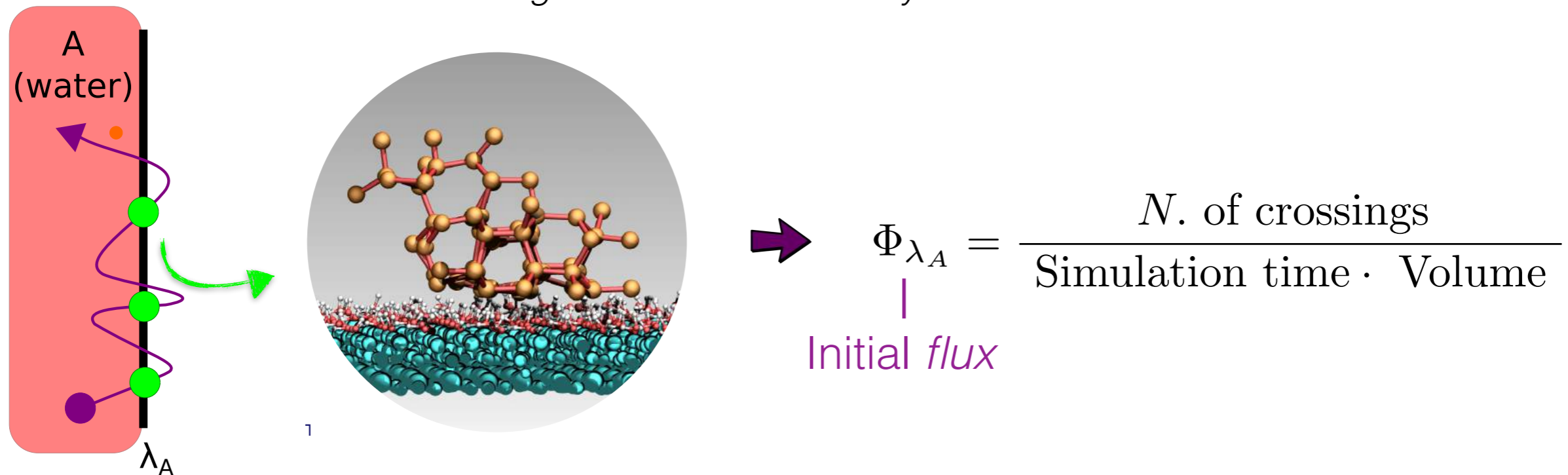
- At each interface λ_i we shoot a (large) number of trial molecular dynamics runs
- Those that reach the next interface (λ_{i+1}) are used as starting point to reach the following interface, and so on…

$$P(\lambda_i | \lambda_{i-1})$$

Forward flux sampling
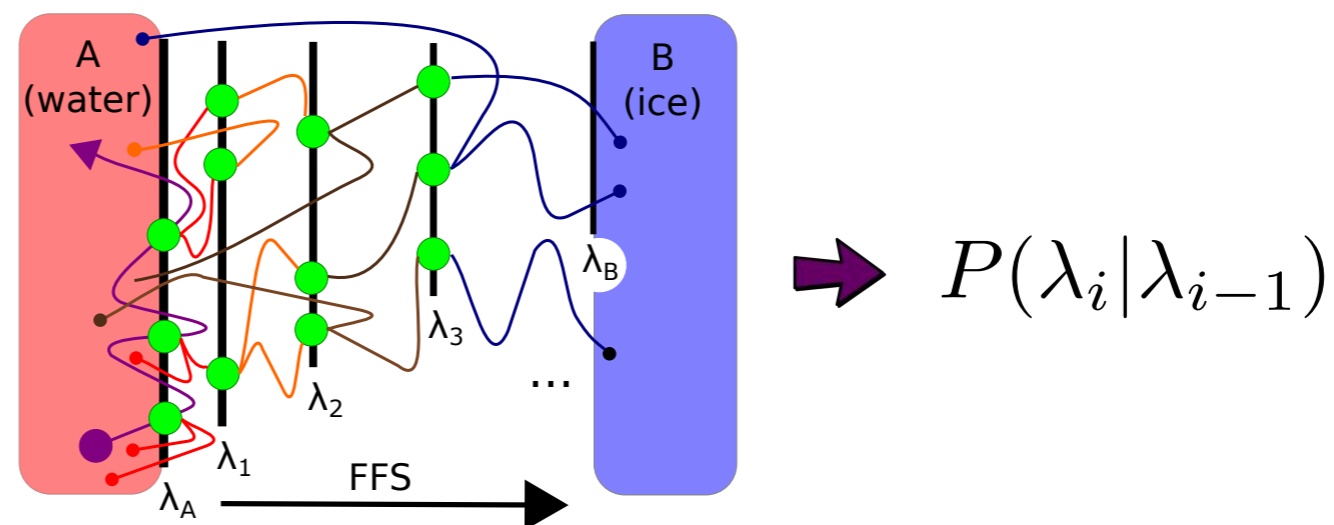
What do you get?

The nucleation mechanism

λ=105   λ=200   λ=325

Sosso, G.C., Li, T., Donadio, D., Tribello, G.A., and Michaelides, A. (2016). J. Phys. Chem. Lett. 7, 2350–2355.

The nucleation rate

$$\mathcal{J} = \Phi_{\lambda_A} \prod_{i=1}^{N_\lambda} P(\lambda_i | \lambda_{i-1})$$

The free energy barrier? *Maybe…*
The critical nucleus size? *Maybe…*

What you don't get:
The dynamics (kinetics prefactor)

Seeded MD

- A crystalline seed of a given size is inserted *beforehand* into the system
- Run different molecular dynamics simulations: same starting point, different temperatures

If the nucleus is (on average) neither growing or melting, it is critical at that temperature
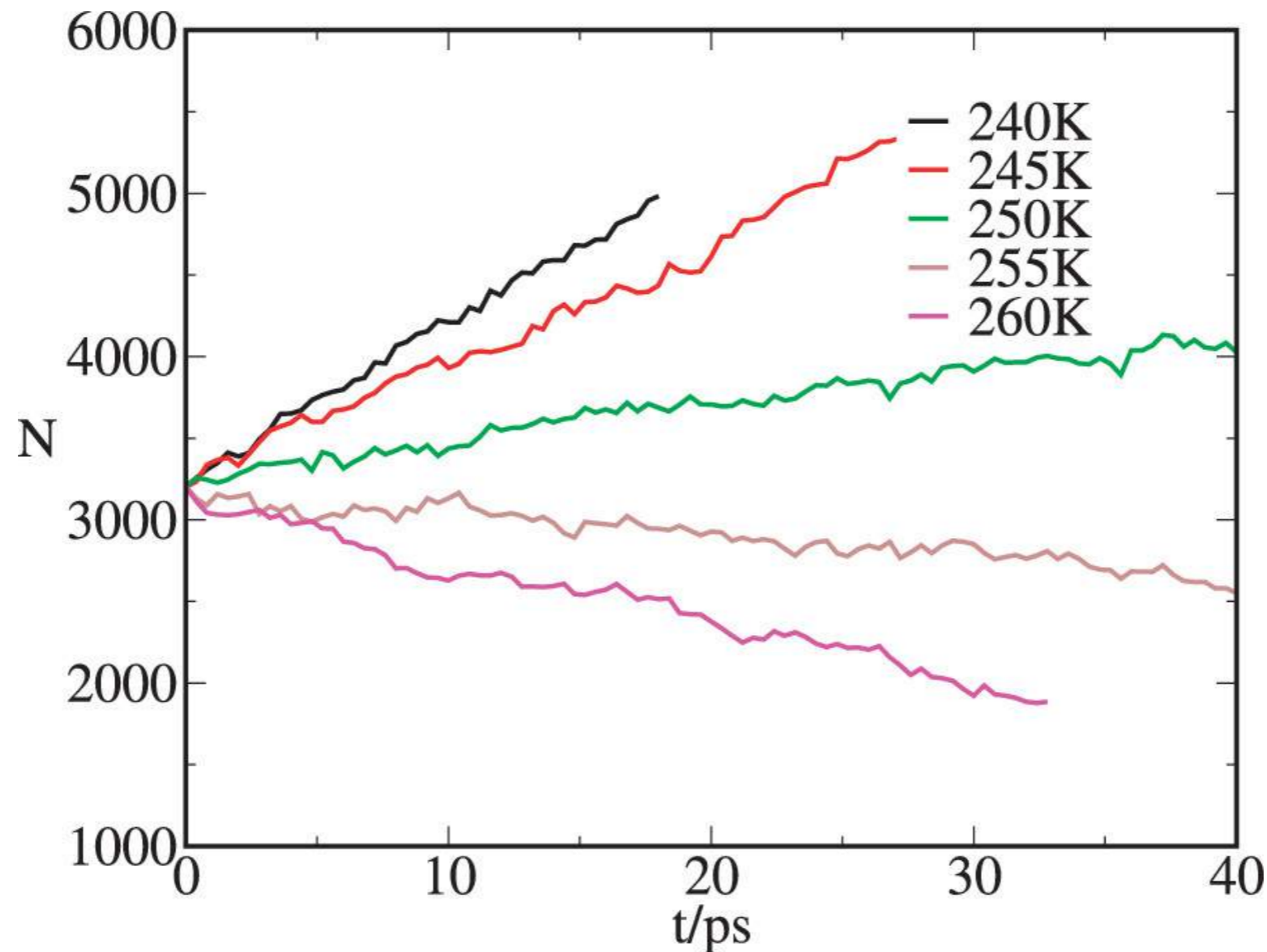
Espinosa, J.R., Vega, C., Valeriani, C., and Sanz, E. (2016).
The Journal of Chemical Physics 144, 34501.

This is *much* faster then either e.g. metadynamics or forward flux sampling
But you have to know what sort of crystal is actually nucleating

Seeded MD

So now you have the critical nucleus size N*…

Use CNT to compute the nucleation rate. You need:
- The kinetic prefactor (attachment rate)
- The free energy barrier

$$\mathcal{J} = \mathcal{J}_0 \exp\left(-\frac{\Delta G^*_\mathcal{N}}{k_\mathrm{B}T}\right) \Rightarrow \textit{Nucleation rate}$$

Kinetic prefactor   *Free energy barrier*

The kinetic prefactor is computed by looking at the *attachment rate*

The free energy barrier is computed from N*, the density of the crystalline phase, and
the free energy difference between the liquid and the crystal [thermodynamic integration]

It works!

HOWEVER

- Error propagation
- You *assume* CNT throughout!

Embracing the failure
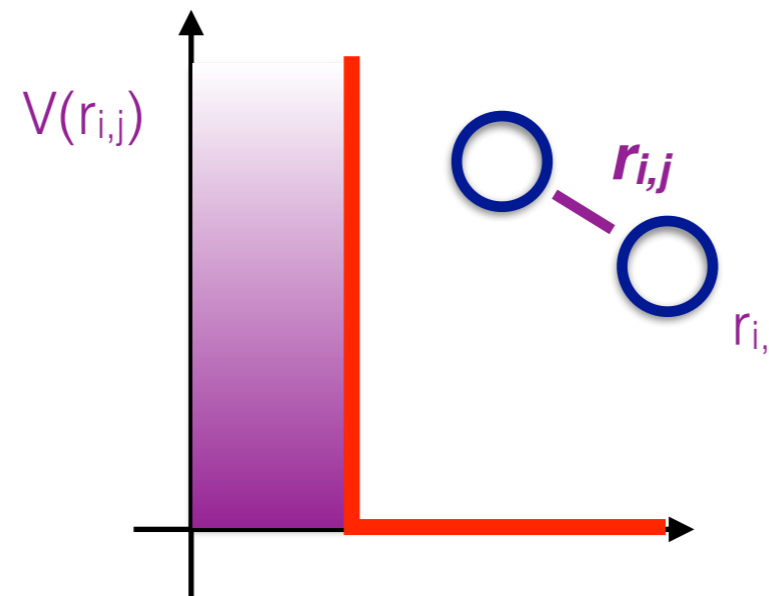
WARWICK
THE UNIVERSITY OF WARWICK

We want to compare the experimental nucleation rates with those
computed via molecular simulations

We now have the whole arsenal of enhanced sampling methods at our disposal
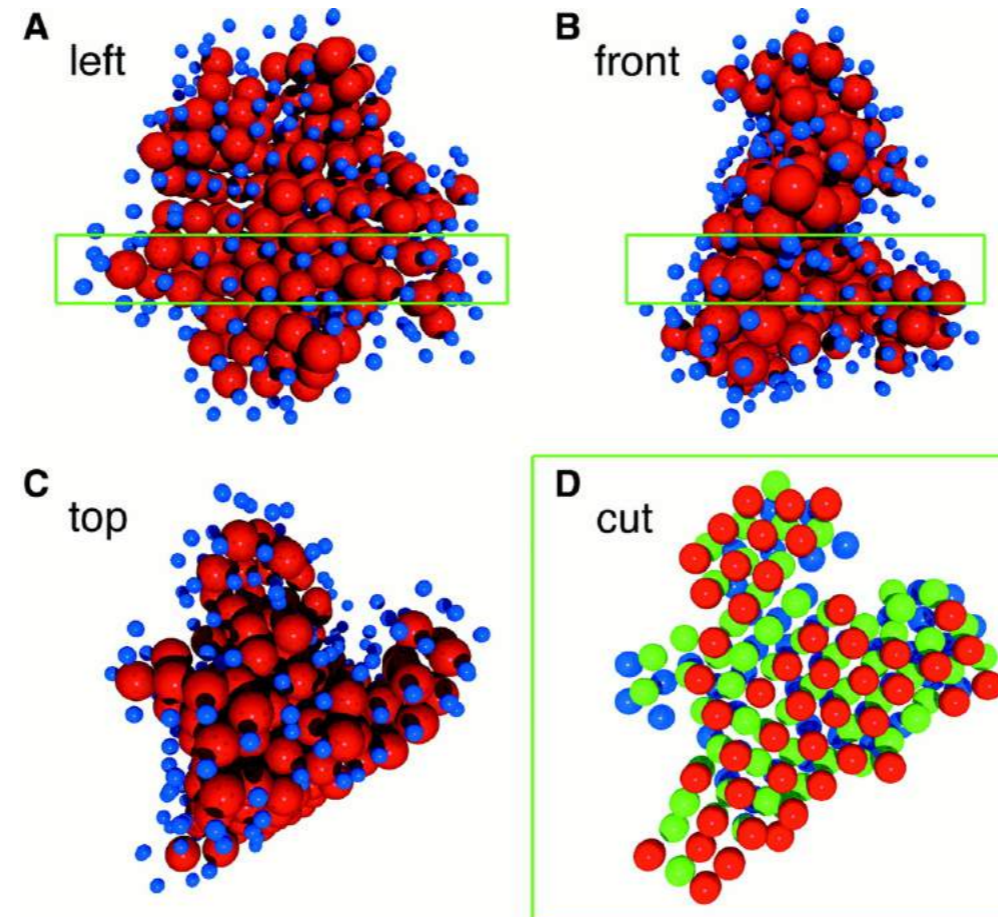
We choose the simplest system:
Colloids

Simulations:
Simple(ST) potential: Hard spheres [computationally *very* fast]

$V(r_{i,j})$

$r_{i,j}$

$r_{i,j}$

We can deploy the whole arsenal … unbiased MD/MC, US, FFS, KMC…

Embracing the failure

Experiments:
Colloidal particles can be imaged in real time and space (e.g. PMMA, confocal microscopy)

Gasser, U., et al. (2001).
Real-Space Imaging of Nucleation and Growth in Colloidal Crystallization.
*Science 292, 258–262.*

Nucleation Rates - Experimental vs Computational
This is *so* simple it *has* to work!

Embracing the failure

Palberg, T. (2014).
J. Phys.: Condens. Matter 26, 333101.

Nucleation rate

$J(2a)^5 / D_s^{long}$

Φ

Volume fraction
*proportional to the **number density***
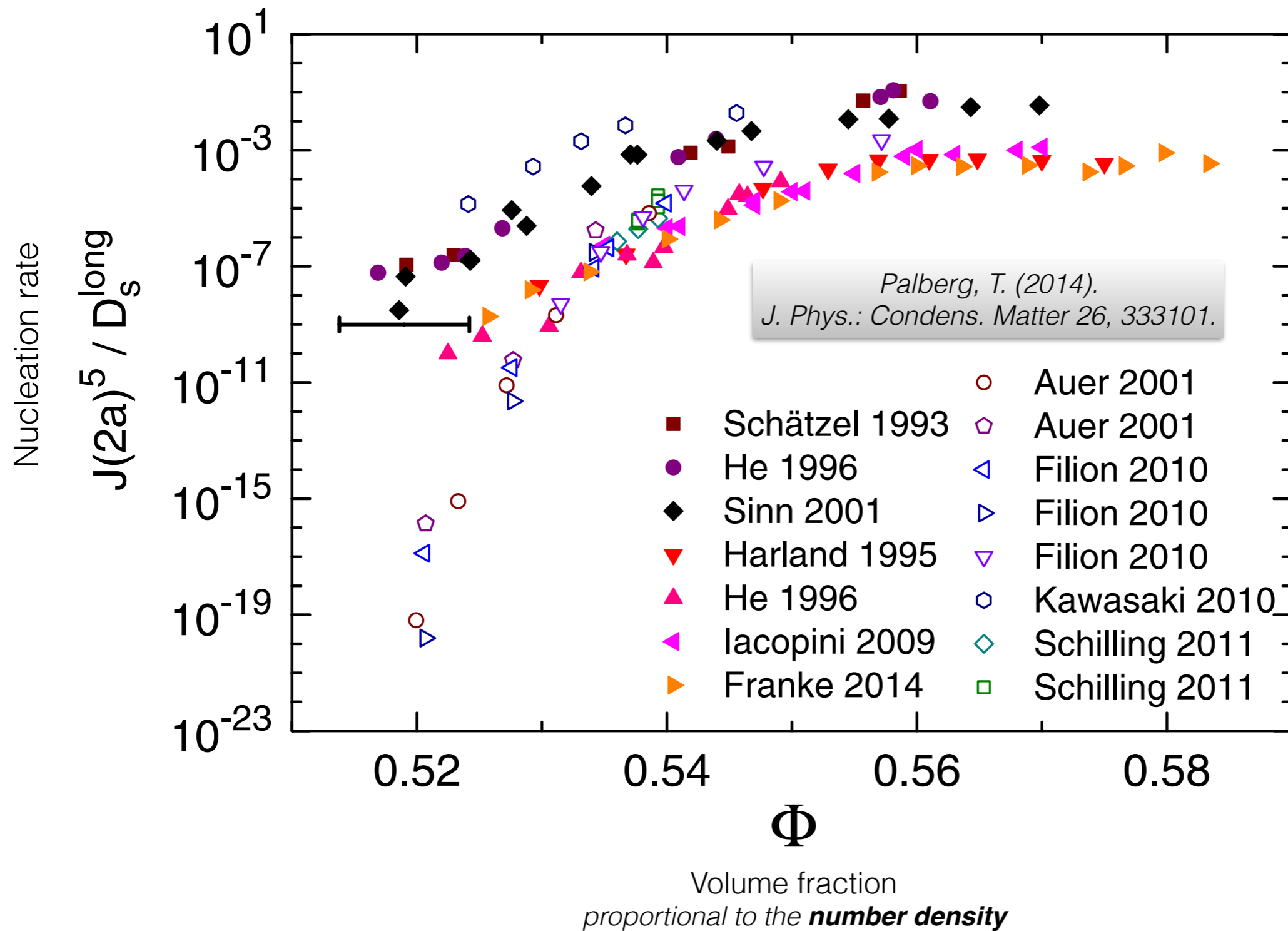
Why?
- It is *not* about the computational methods
- Chances are this is about *the model…*

Embracing the failure

Water into ice
homogeneous nucleation

Sosso, G.C., et al. (2016). Chem. Rev. 116, 7078–7116.
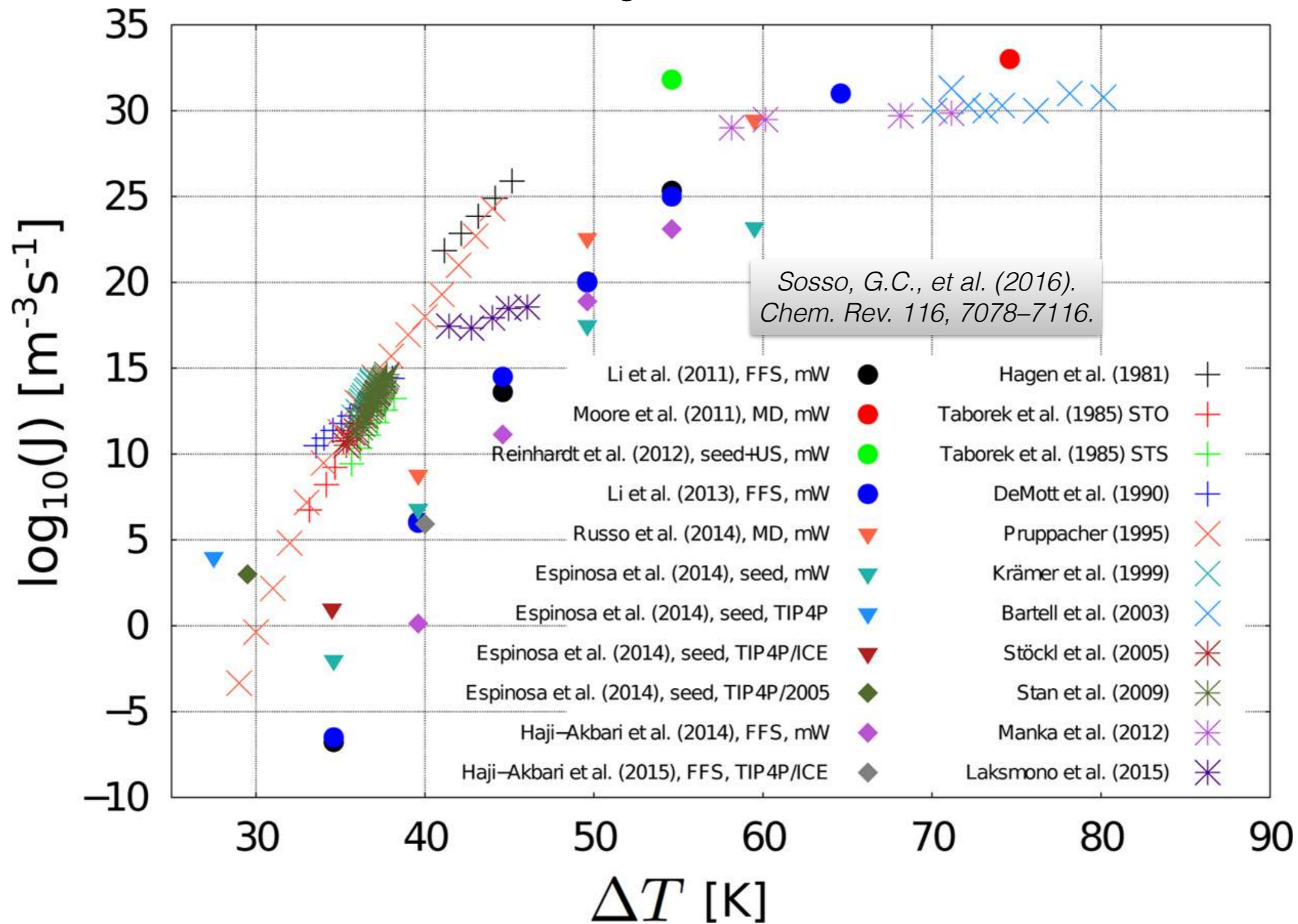
Why? Models *and* Methods…

Blaming the methods

Seeded MD vs FFS:
Inconsistent (five orders of magnitude for mW at strong supercooling)

In principle, FFS simulations are the most accurate/reliable option

HOWEVER

Some FFS simulations at *exactly* the same conditions are inconsistent

Blaming the models

The discrepancy between mW and TIP4P is expected

Tiny error in the reproducing the thermodynamical properties lead to enormous errors in the nucleation rates

$\Delta\mu$ (TIP4P$_{Ice}$) is about %20 smaller than $\Delta\mu$ (exp)
⇨ 9 orders of magnitude!

The interfacial free energy is equally cumbersome [stacking disordered ice]

- We need better models
- It's not about absolute numbers
- It's about *relative trends*

Recap

- Crystal nucleation is important

- It is a challenge for experiments *and* simulations

- The computational options:

    - Brute force molecular dynamics

    - Free energy methods

    - Path sampling methods

    - Seeded MD

- Nucleation Rates: Exp. vs Sim.

    - We fail even when considering colloids

    - Methods and models are both to be blamed

    - It is all about relative trends
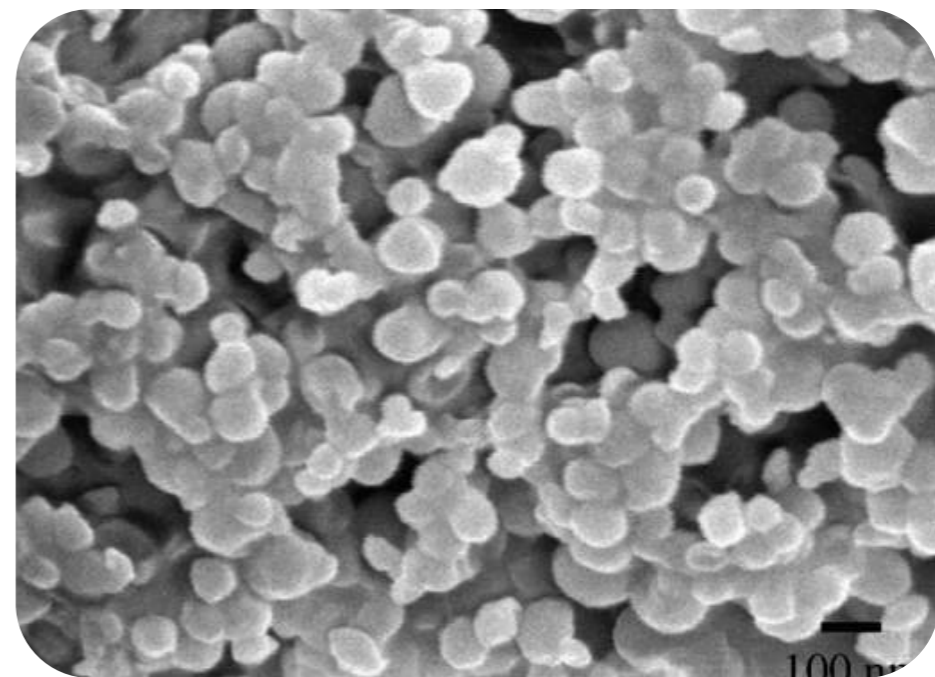
My research

Atomistic simulations of
heterogeneous ice nucleation

Why do we care?

Heterogeneous
Ice Nucleation

Mineral
Dust
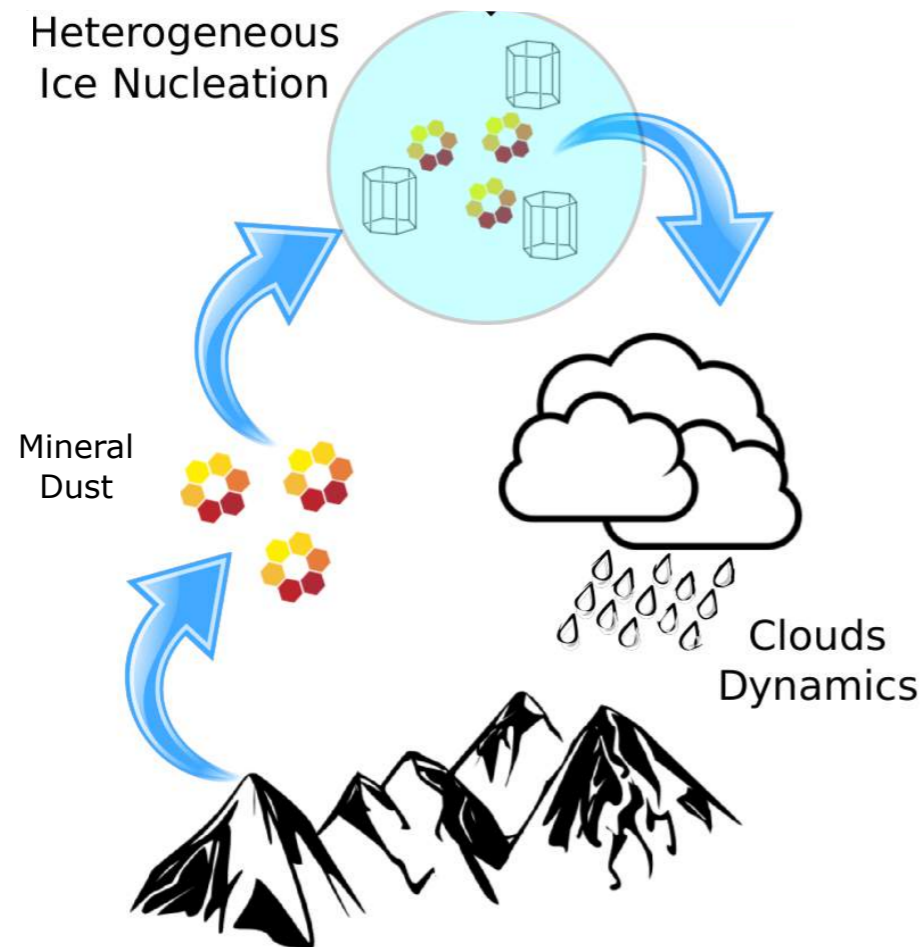
Clouds
Dynamics

Atmospheric science:

Clouds formation and dynamics

(climate change)

Cryobiology:

Intracellular freezing

(cryotherapy and
cryopreservation)

100 n

Ice formation on top of lipid bilayers
Cryobiology, 55, 210 (2007)

Heterogeneous Ice Nucleation

It is surprisingly difficult to freeze pure water
One has to go to -40 ^{0}C…

Ice on Earth forms thanks to heterogeneous nucleation

water

θ

ice

solid
impurity

Ice nucleating agents

WARWICK
THE UNIVERSITY OF WARWICK
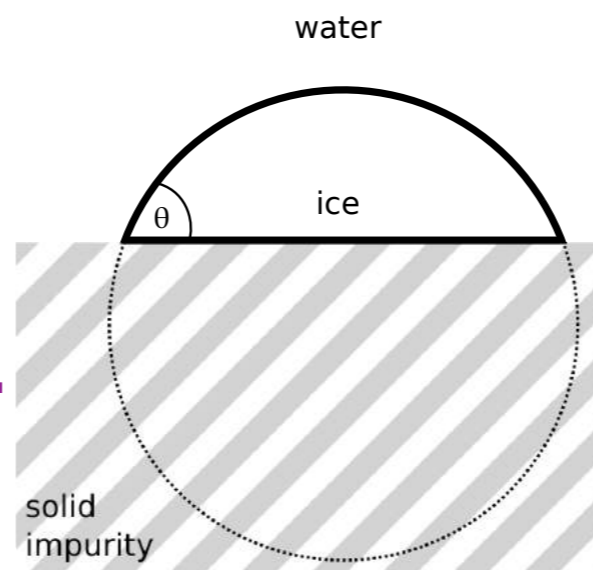
Atmospheric Science

- Mixed-phase (ice and water) clouds: form @ ~ -15 °C
- What sort of impurities can help the formation of ice?
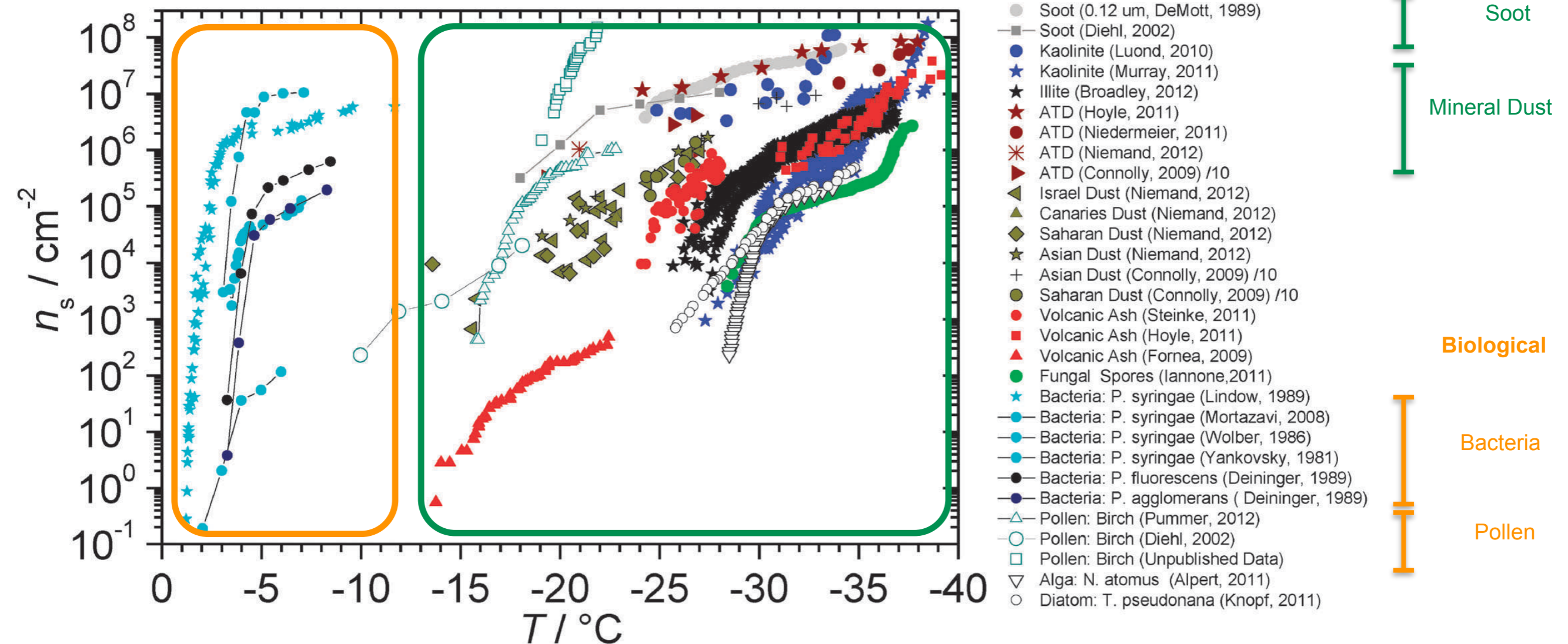
Murray, B.J., O'Sullivan, D., Atkinson, J.D., and Webb, M.E. (2012).
Chem. Soc. Rev. 41, 6519–6554.

What is it that makes a certain material capable of promoting the formation of ice?

Model systems

WARWICK
THE UNIVERSITY OF WARWICK

Simple systems ➡ General trends

Coarse grained (mW) water

- Computationally fast
- Fast water dynamics even at strong

on top of
ideal FCC crystals (LJ particles, frozen)

*Fitzner, M., Sosso, G.C., Cox, S.J., and Michaelides, A. (2015).
J. Am. Chem. Soc. 137, 13658–13669.*
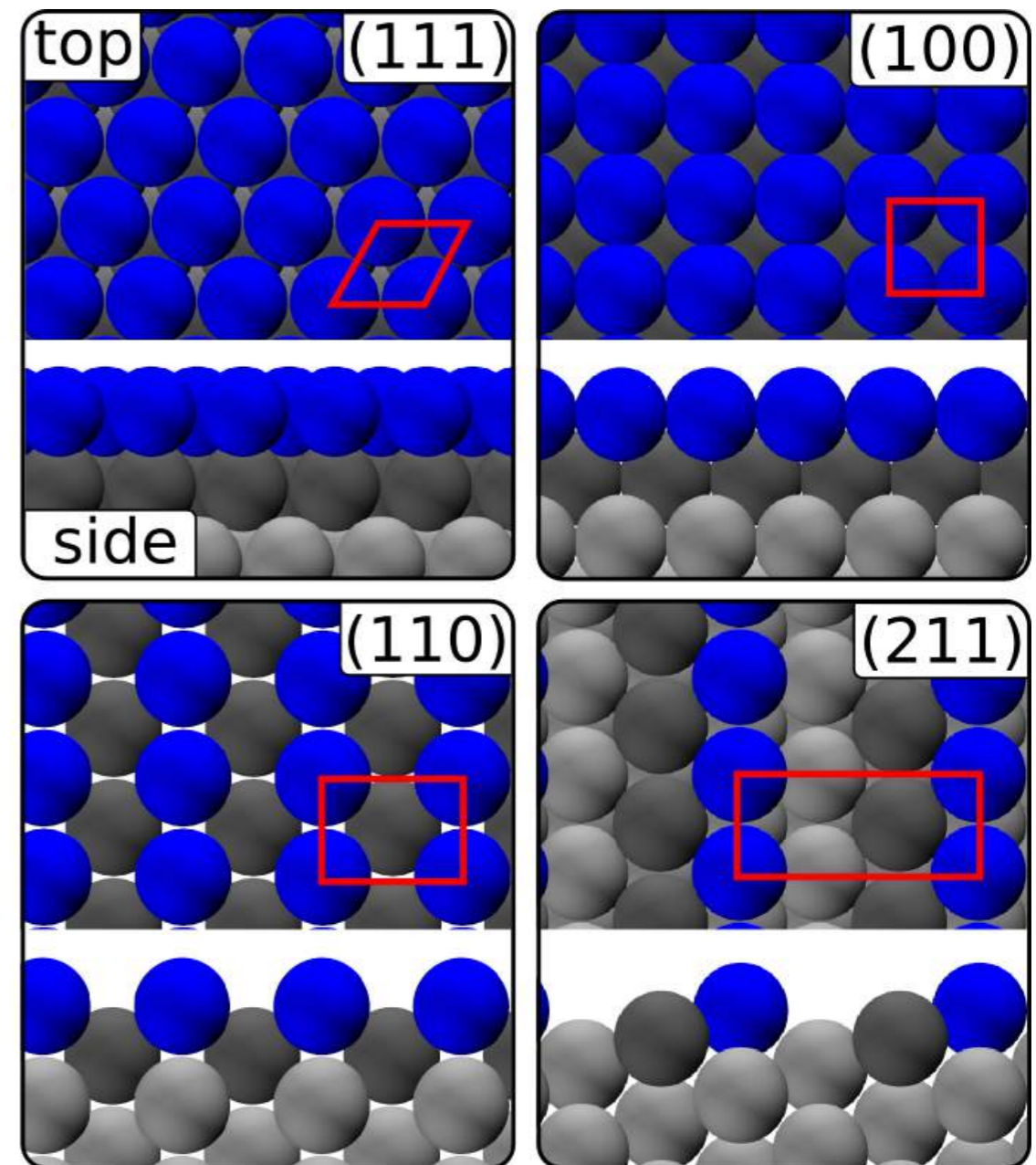
- (111), (100), (110) and (211) surfaces (surface morphology)
- Different lattice parameters afcc [3.52 - 4.66 Å] (surface morphology)
- Different water-surface interaction (LJ potential) strength Eads [0.2-12 kcal/mol] (hydrophobicity)

Model systems

Brute force molecular dynamics simulations:
Nucleation rate from survival probability

$$P_{liq}(t) = 1 - \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} \Theta(t - t_n^{(i)})$$

γ = 1.12

$$P_{\mathrm{liq}}(t) = \exp[-(J \cdot t)^{\gamma}]$$

Model systems
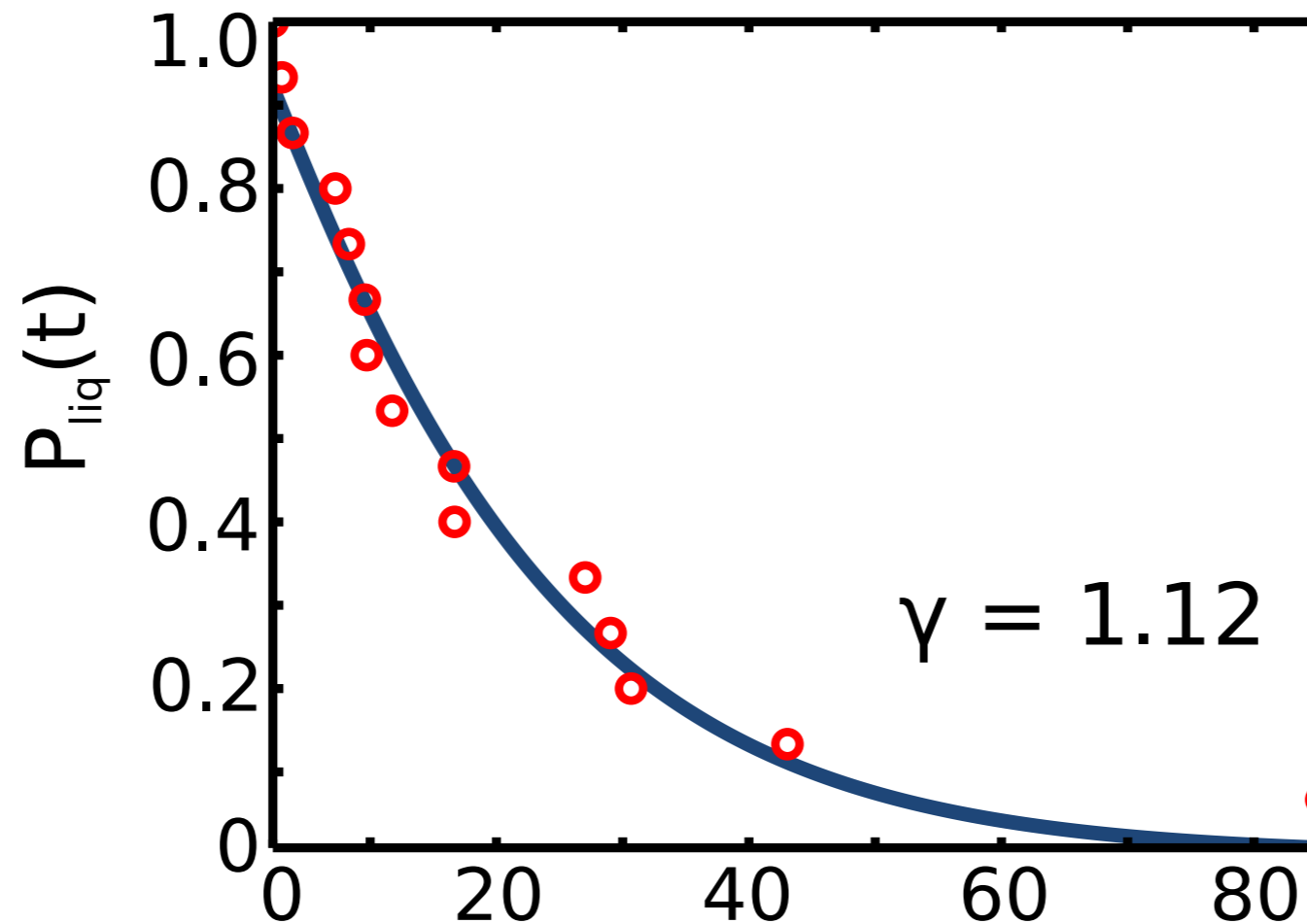
Simple models, complex behaviours

(211) face

Model systems
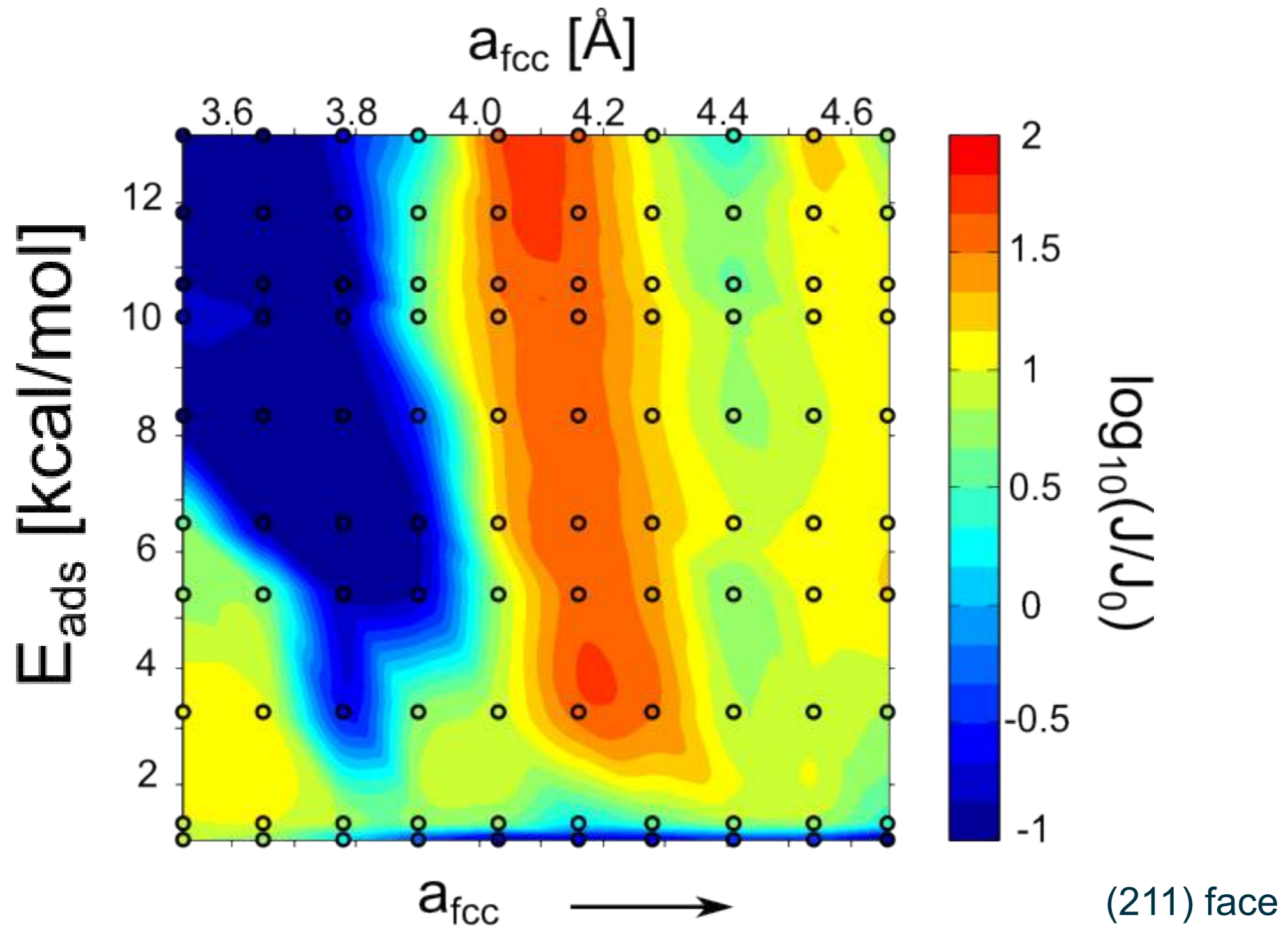
Different ice faces on top of the very same surface
e.g. the (211) face
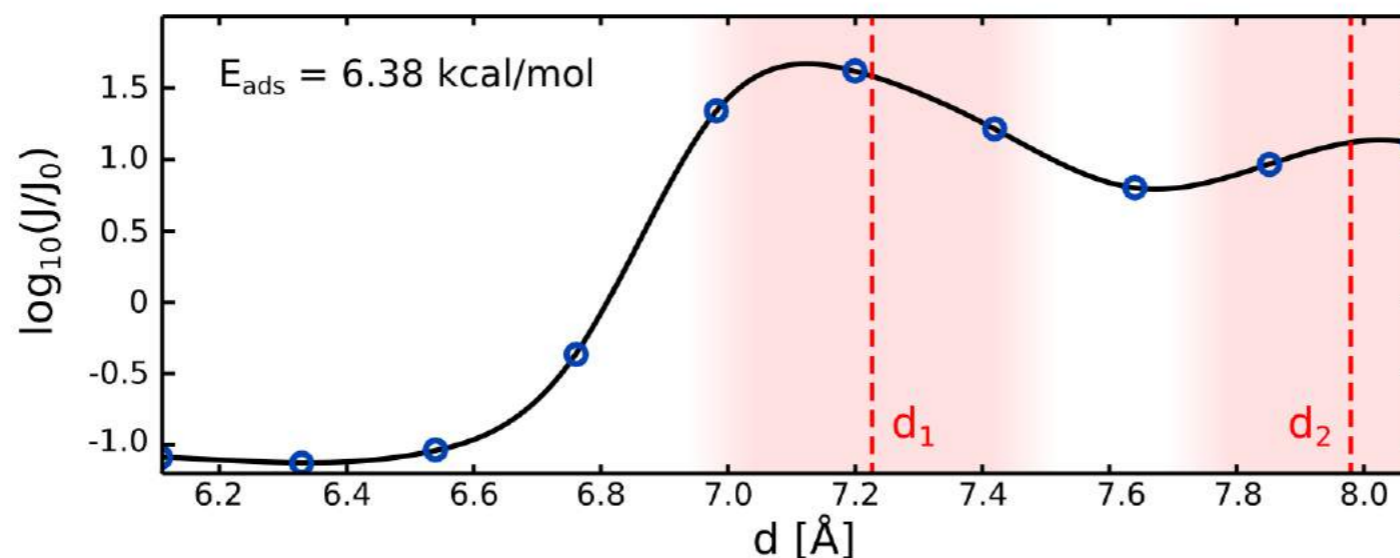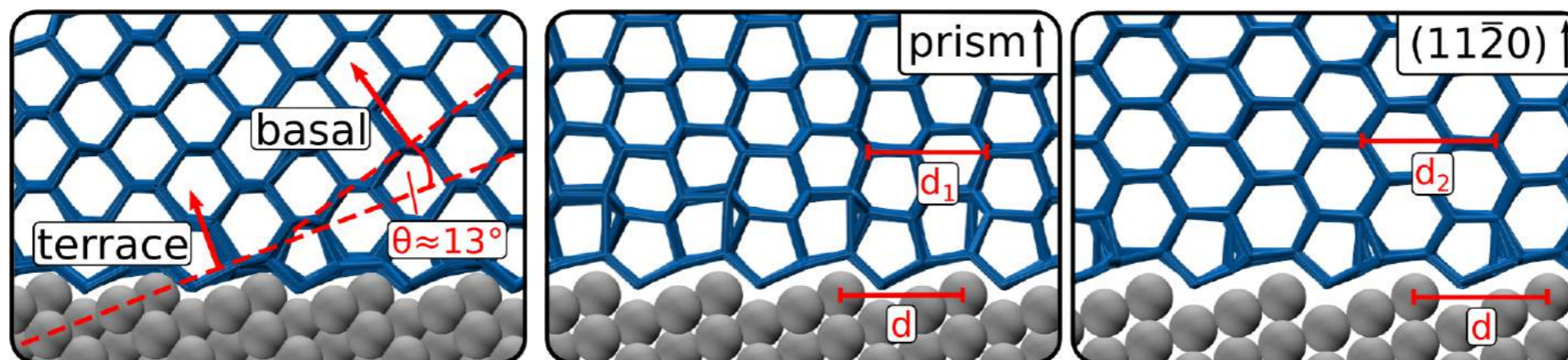
Different a_fcc ➡ Different templating effects ➡ Different ice faces

Each surface has its own story to tell

Realistic systems

Realistic systems:
- Complex interactions (hydrogen bonding)
- Flexibility of the surface
- Nucleation sites (defects…)
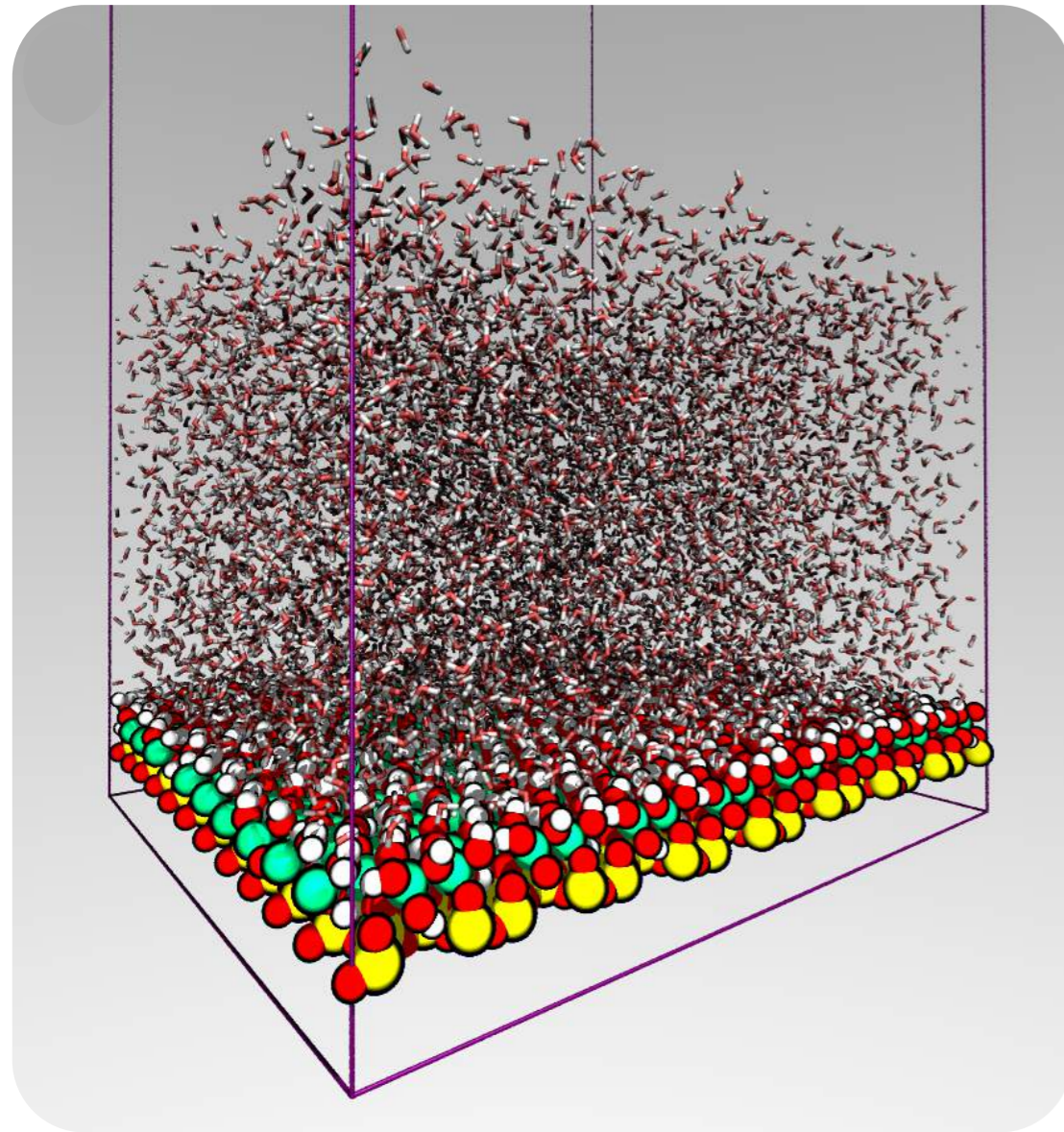
- Layered aluminosilicate [$Al_2Si_2O_5(OH)_4$] : it's a clay
- (Siloxane or) Hydroxylated (001) surface
- The hexagonal arrangement of -OH groups promotes ice formation

- *H. R. Pruppacher and J. D. Klett, Microphysics Of Clouds And Precipitation (1997)*
- *S. J. Cox et al., Farad. Discuss. 167, 389 (2014)*

Experiments:
Kaolinite is quite effecting in promoting ice nucleation in clouds

Murray, B.J., et al. Nature Geosci 3, 233–237.

MetaD vs FFS

WARWICK
THE UNIVERSITY OF WARWICK

Atomistic (e.g. TIP4P/Ice) simulations of
heterogeneous ice nucleation [kaolinite and more…]

~~Metadynamics~~

Forward Flux Sampling

Metadynamics:
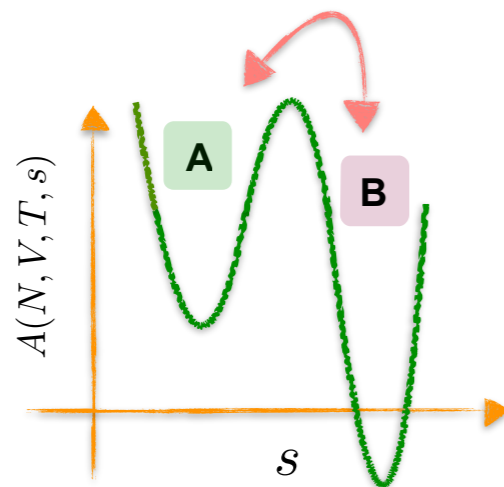
✗ Very sensitive to the choice of the CV(s)

✗ Sub-regions of the system have to be biased

✗ You often end up with the wrong polymorph

✗ Info about kinetics cannot be easily obtained

✗ Massive hysteresis

✗ Simple CVs are not enough

✗ Computationally expensive

Forward Flux Sampling:
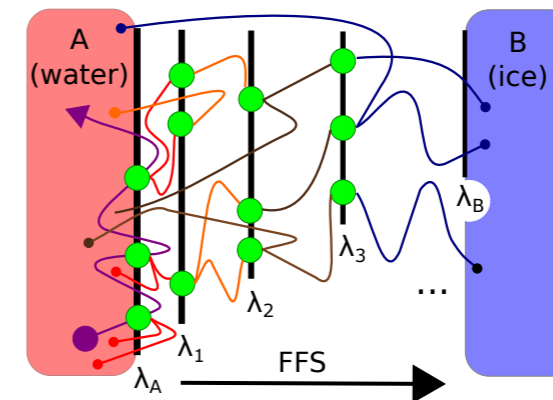
✓ Less sensitive to the choice of the CV(s)

✓ The whole system can be considered

✓ You get the right polymorph

✓ Info about kinetics come for free

✓ No hysteresis

✗ Simple CVs are not enough

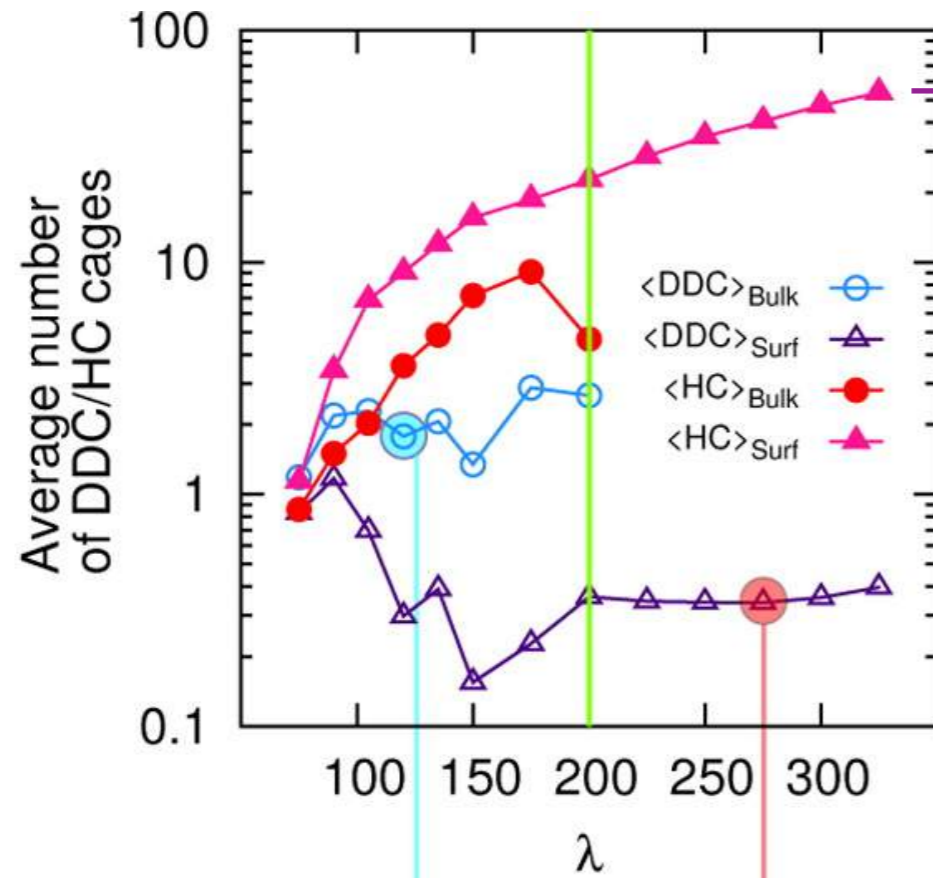✗ Computationally **awfully** expensive

Mechanism
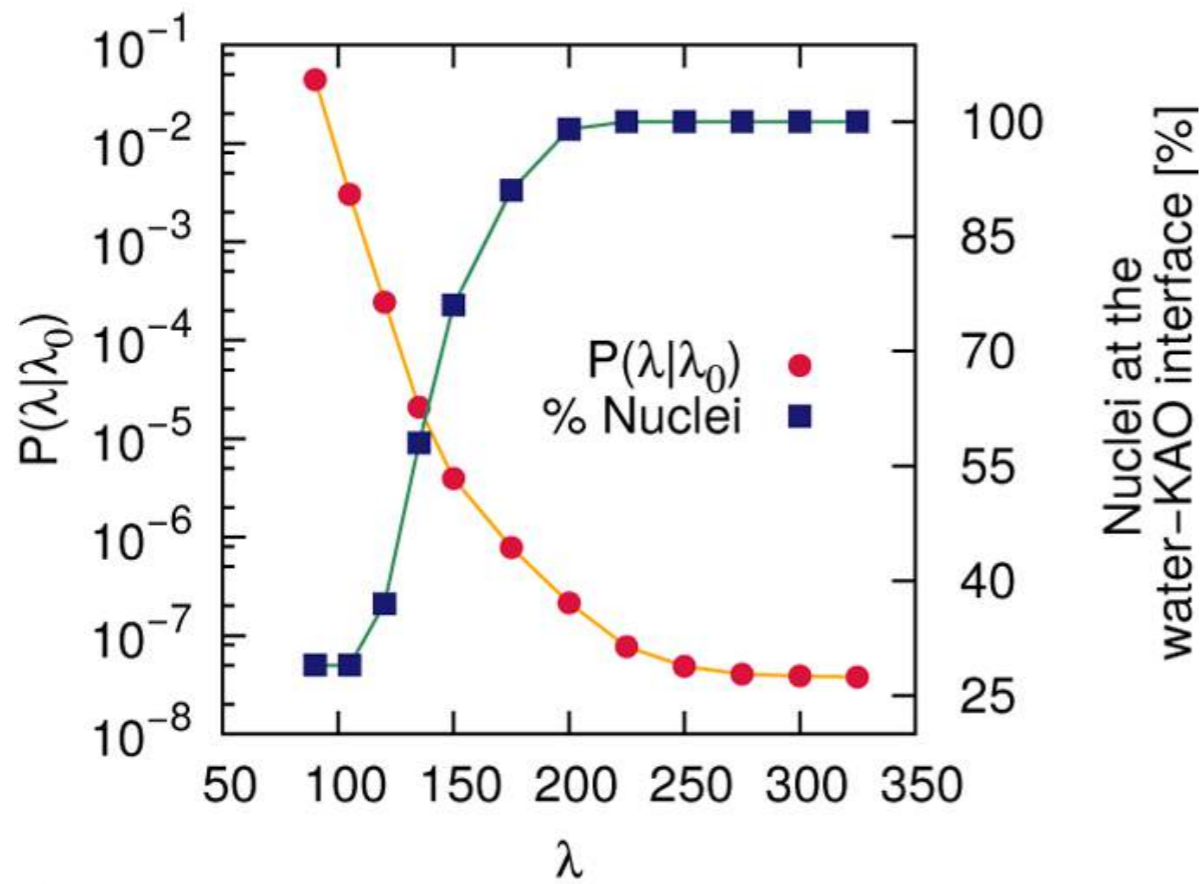
Ice nuclei nucleate at
the kaolinite-water interface

Sosso, G.C., Li, T., Donadio, D., Tribello, G.A., and Michaelides, A. (2016).
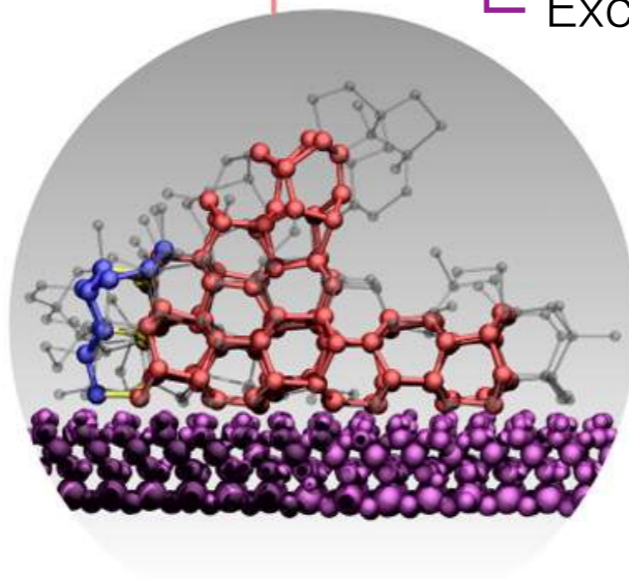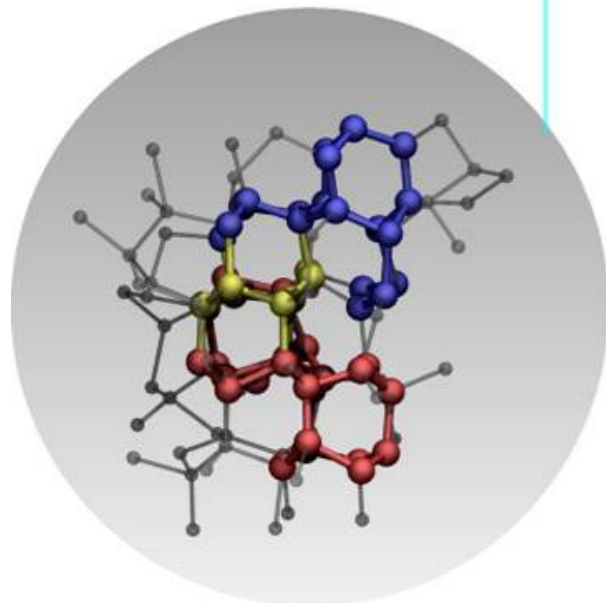*J. Phys. Chem. Lett. 7, 2350–2355.*

Clear templating effect
of the -OH groups

λ=105   λ=200   λ=325

Mechanism
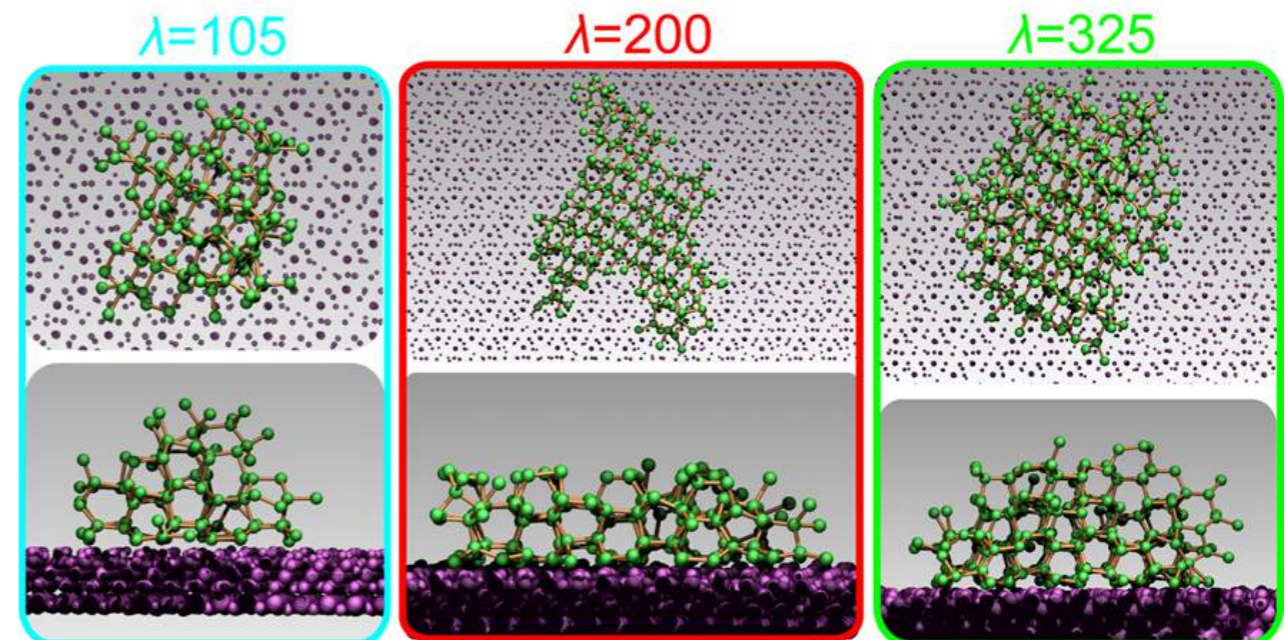
Homogeneous ice nucleation:
At strong supercooling, *stacking disordered ice* forms
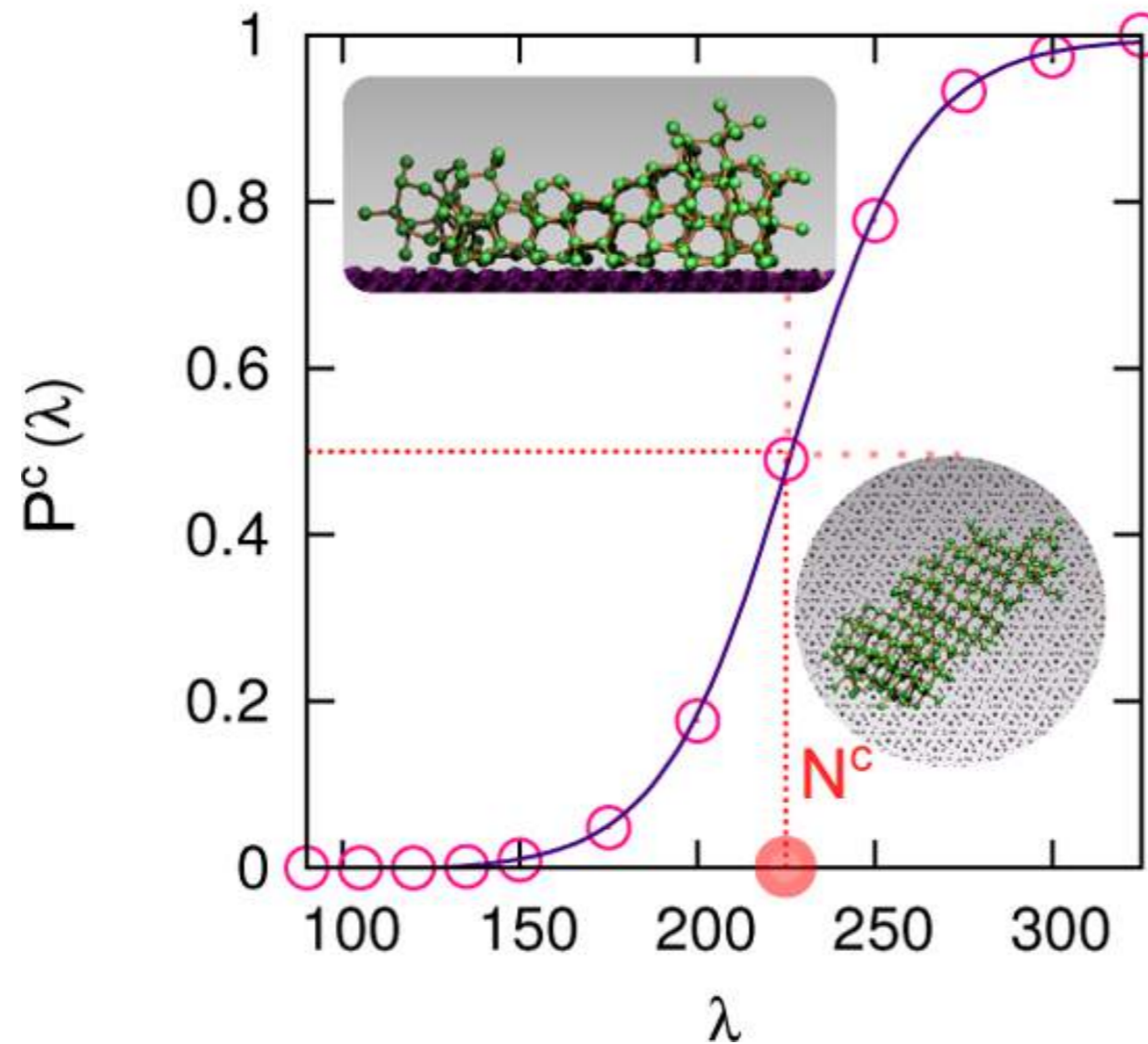
Exclusively *hexagonal ice* exposing the primary prism face
to the -OH pattern of kaolinite

The critical nucleus

The critical nucleus (N*)

P^C = Committor probability
(would a nucleus of a certain size λ melt into the liquid or grow into the crystal?)

- N*$_{KAOLINITE}$ is ~ 225, less than one half of N*$_{HOMOGENEOUS}$
- N* is not shaped as a spherical cap (as Classical Nucleation Theory would assume)

The nucleation rate

$J_{KAOLINITE}$ is about
20 orders of magnitude larger than
$J_{HOMOGENEOUS}$

Can we compare our result with the experimental number?

- Supercooling is too strong (~42 K means homogenous freezing)
- Experimental nucleation rates are usually scattered along several orders of magnitude
- We are probing the ideal, defects-free (001) hydroxylated surface of kaolinite. Experiments deal instead with kaolinite particles exposing different faces, most likely characterised by defects.

Forget about absolute numbers:
we should look for **relative trends**

- We can use the same method to compare the ice nucleating ability of different materials (here: kaolinite versus the homogeneous case)
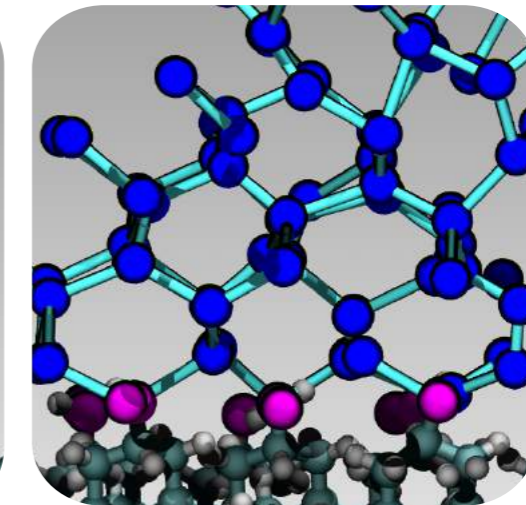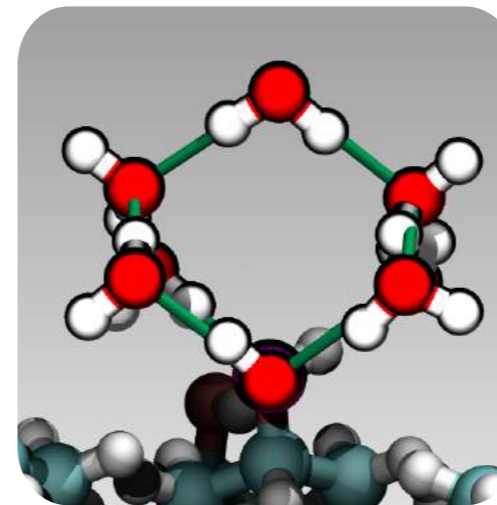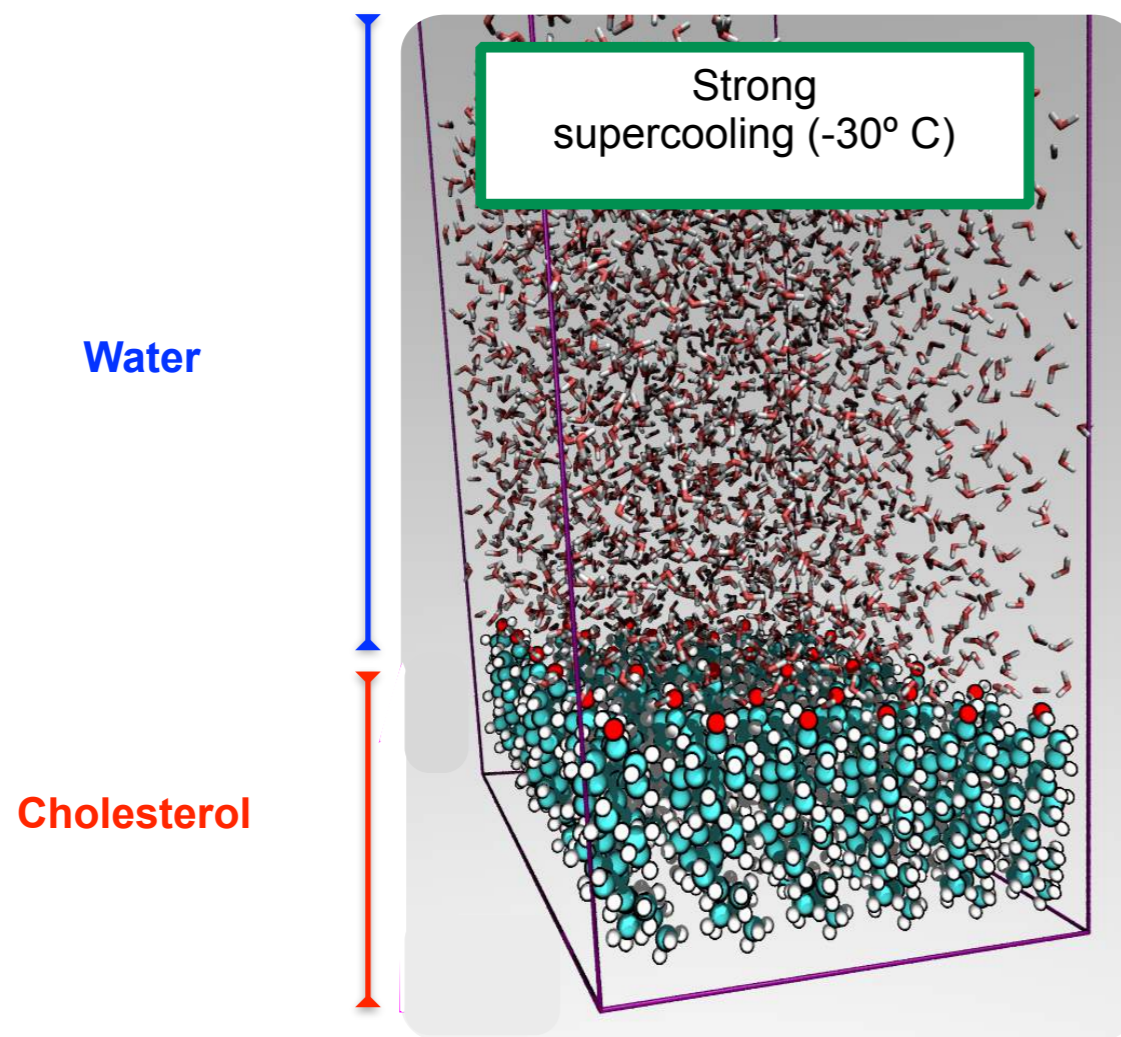- We can get insight into the molecular mechanism of ice formation
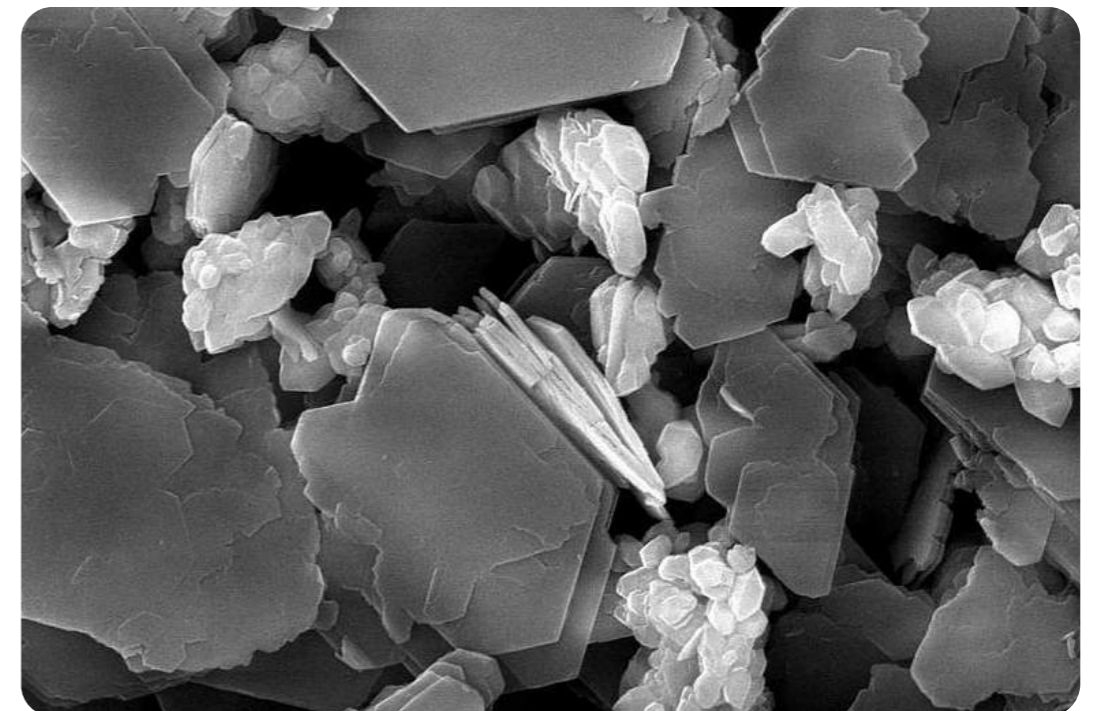
Ice on cholesterol

Ice nucleation on cholesterol crystals

Frozen Droplets Experiments (T. Whale, Leeds)

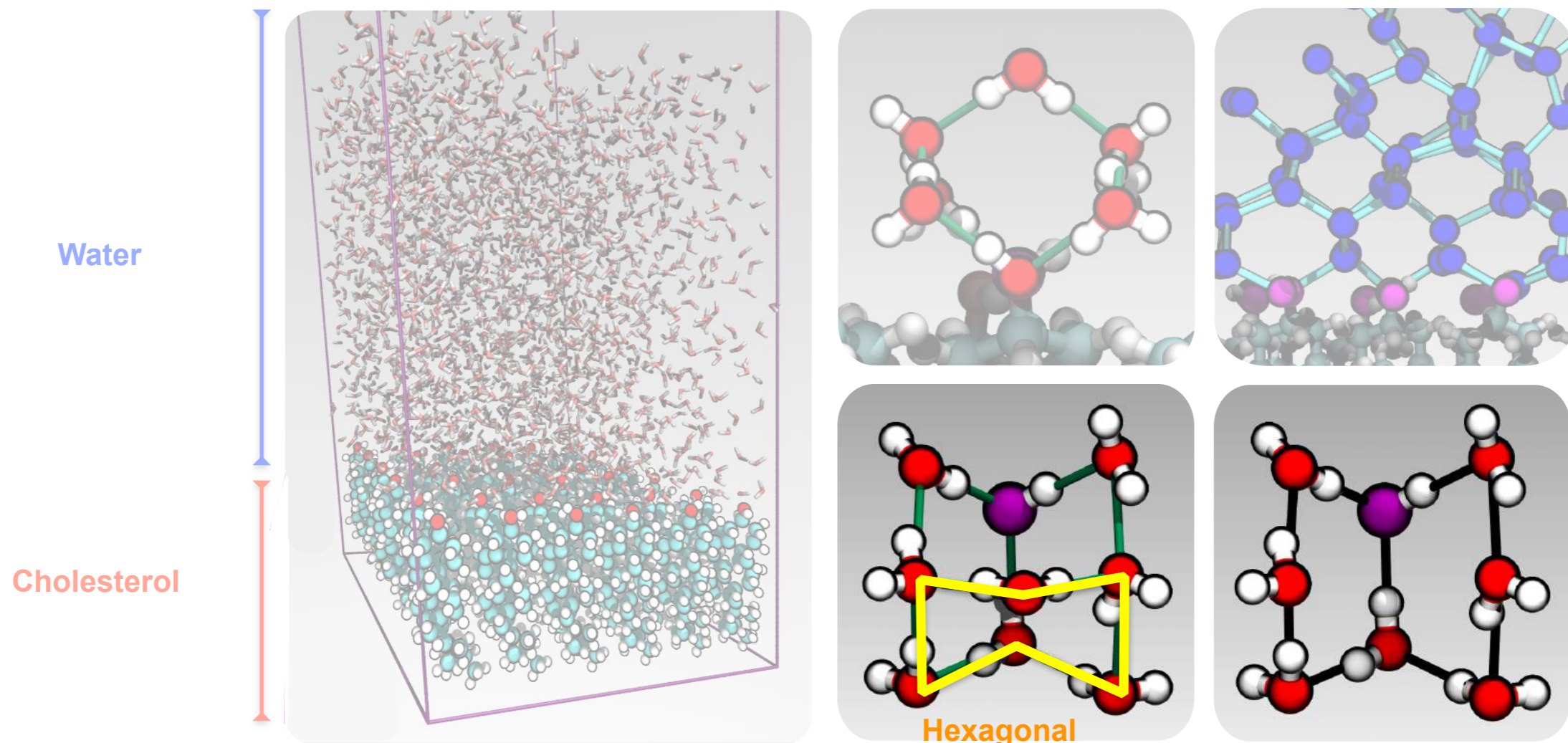**Cholesterol crystals: spectacularly good (-30/-2 °C) ice nucleating agents**

WHY?

Water

Strong supercooling (-30º C)

Cholesterol

Force Fields
Water = TIP4P/Ice
Cholesterol = CHARMM_36

Molecular Dynamics Simulations
Hydrogen-bonded H_2O/-OH cages
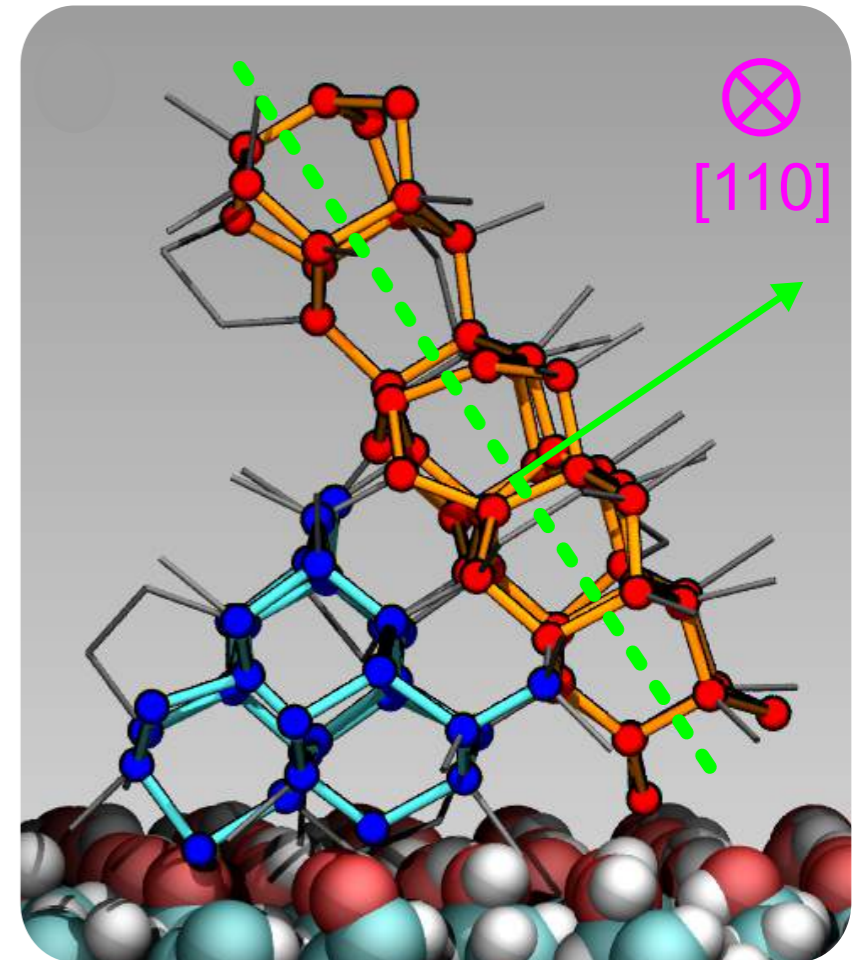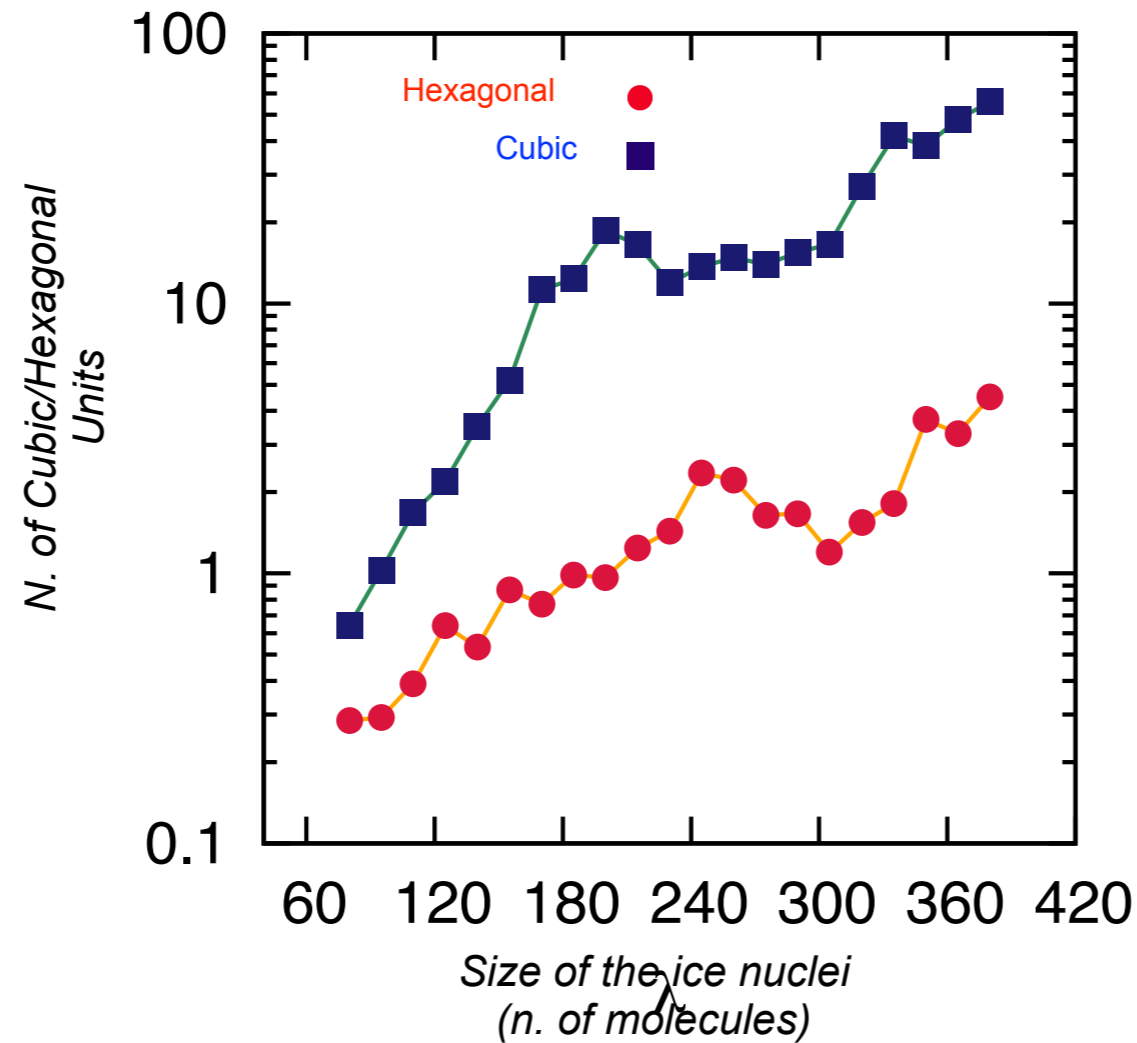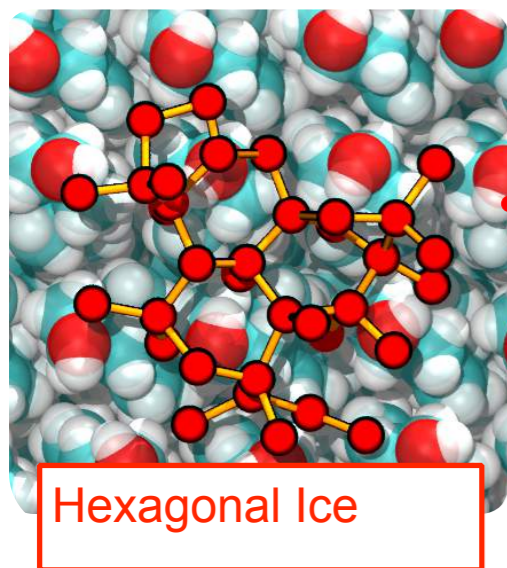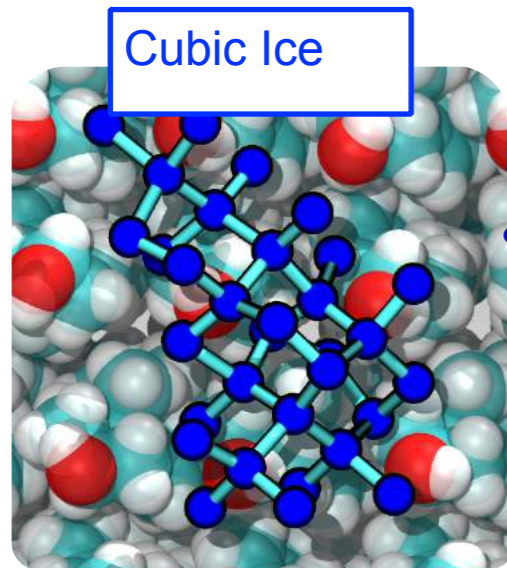
Frozen Droplets Experiments (T. Whale, Leeds)

Cholesterol crystals: spectacularly good (-30/-2 °C) ice nucleating agents

WHY?

Water

Cholesterol

Hexagonal

Force Fields
Water = TIP4P/Ice
Cholesterol = CHARMM_36

Molecular Dynamics Simulations
Pentagonal rings

Ice on cholesterol

WARWICK
THE UNIVERSITY OF WARWICK

Cubic Ice

Hexagonal Ice

Hexagonal ●
Cubic ■

N. of Cubic/Hexagonal Units

Size of the ice nuclei
(n. of molecules)

⊗
[110]

Forward Flux Sampling Simulations
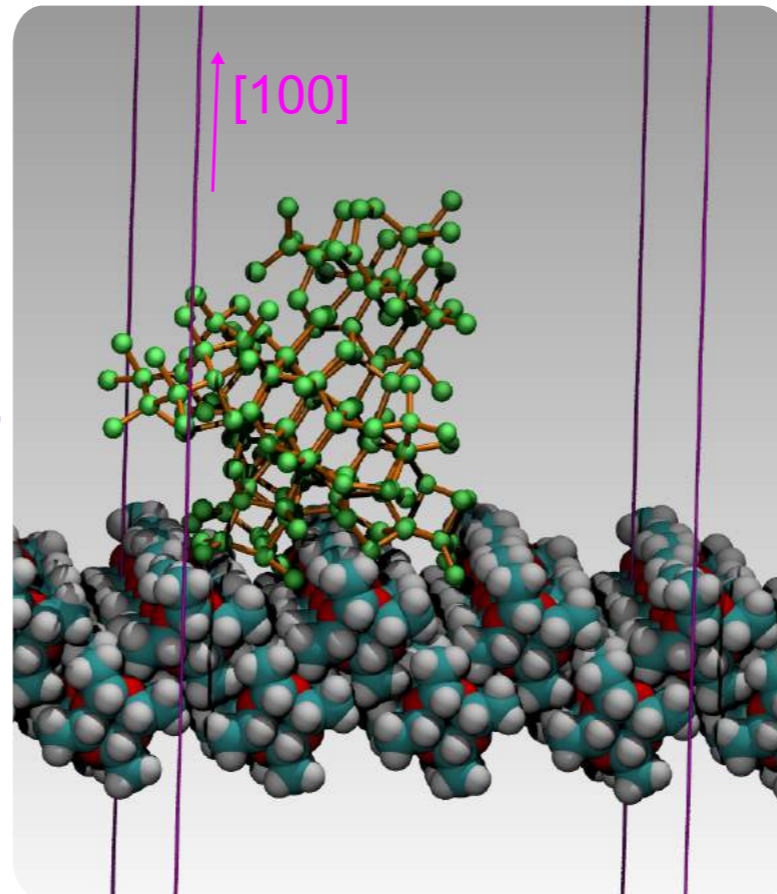
- **One surface, two ice polymorphs (hexagonal and cubic)**
- **The ice nucleating ability of cholesterol crystals: a Game of Temperatures**

It's not enough!

www.slido.com || #Y690

We have a (costly, but accurate) computational framework to investigate the heterogenous nucleation of ice

HOWEVER

The nucleation rates of ice on:
- Kaolinite
- Cholesterol
- Metaldehyde

[100]

are basically identical
(at strong supercooling, all good INA do a similarly good job…)

This is not good enough!
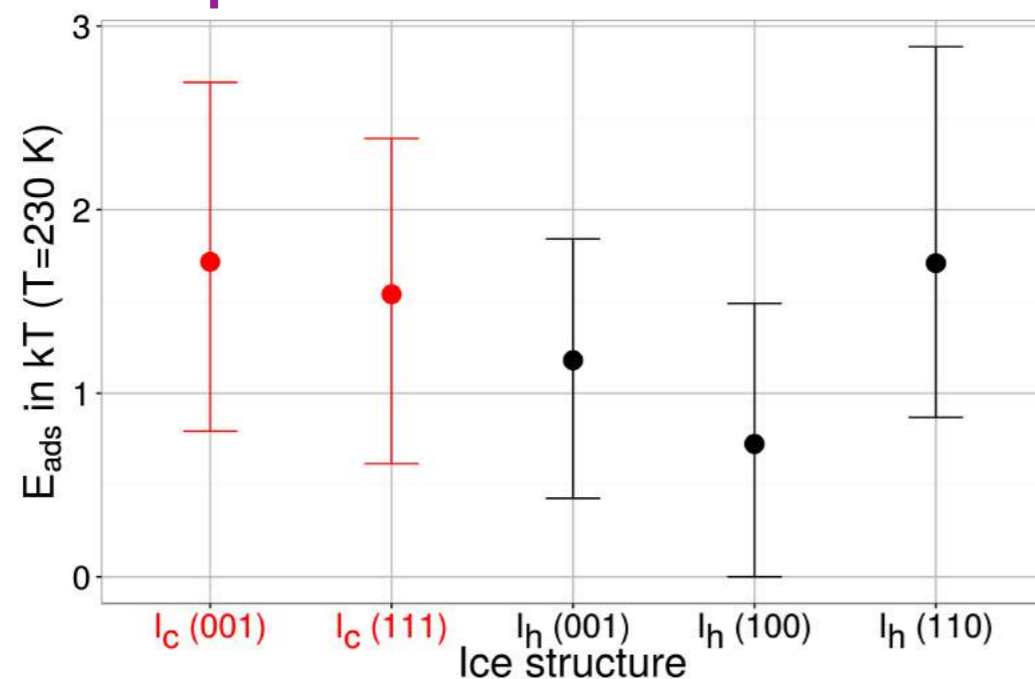(remember, we are interested in *trends*!)

Heterogeneous seeded MD

Hseeded MD

WARWICK
THE UNIVERSITY OF WARWICK

Same basic idea of seeded MD
(homogeneous case)

*The case of
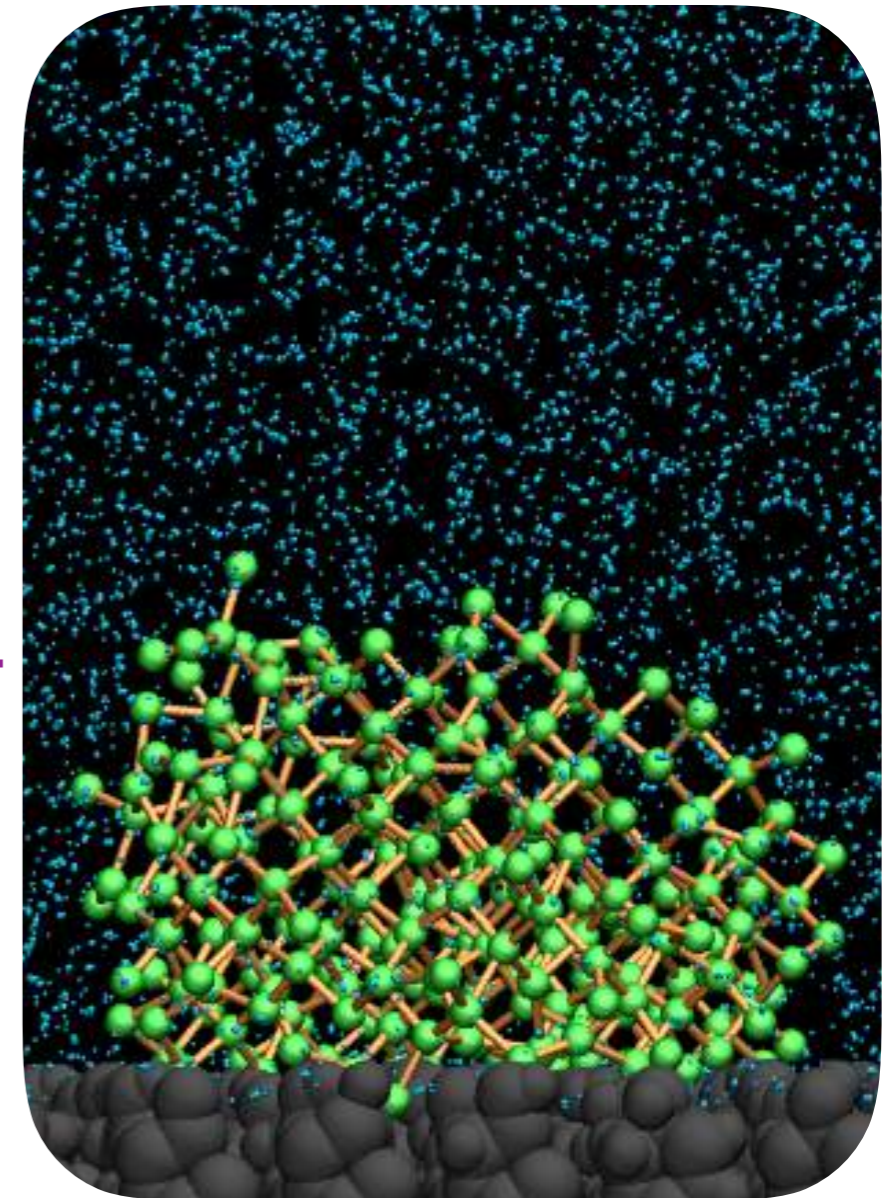ice on cholesterol*

HOWEVER

- *Which seed* do we choose?
- How do we put the seed in contact with the surface?

⬇

- "Adsorption energy" of the first overlayer (we can do better!)
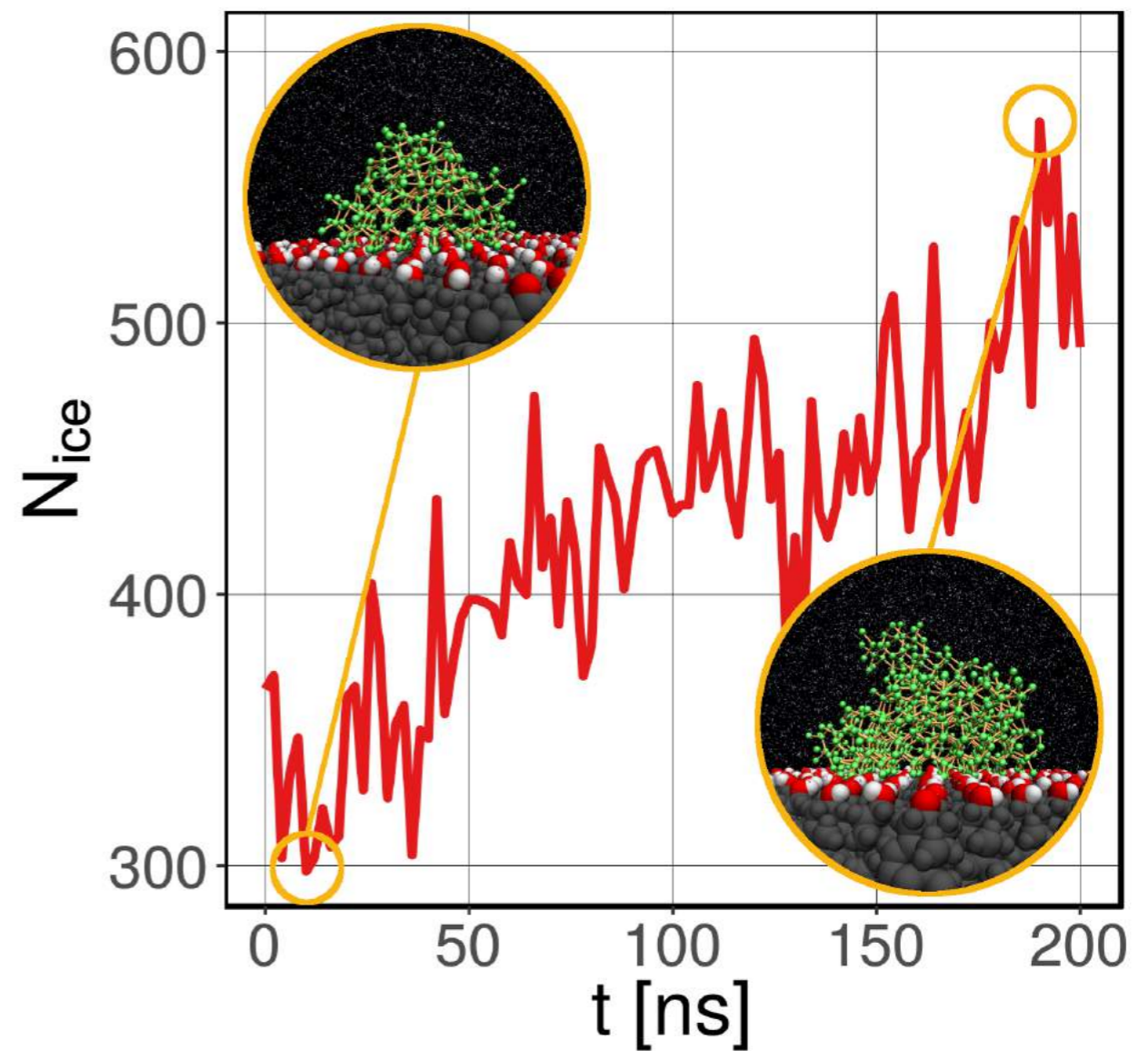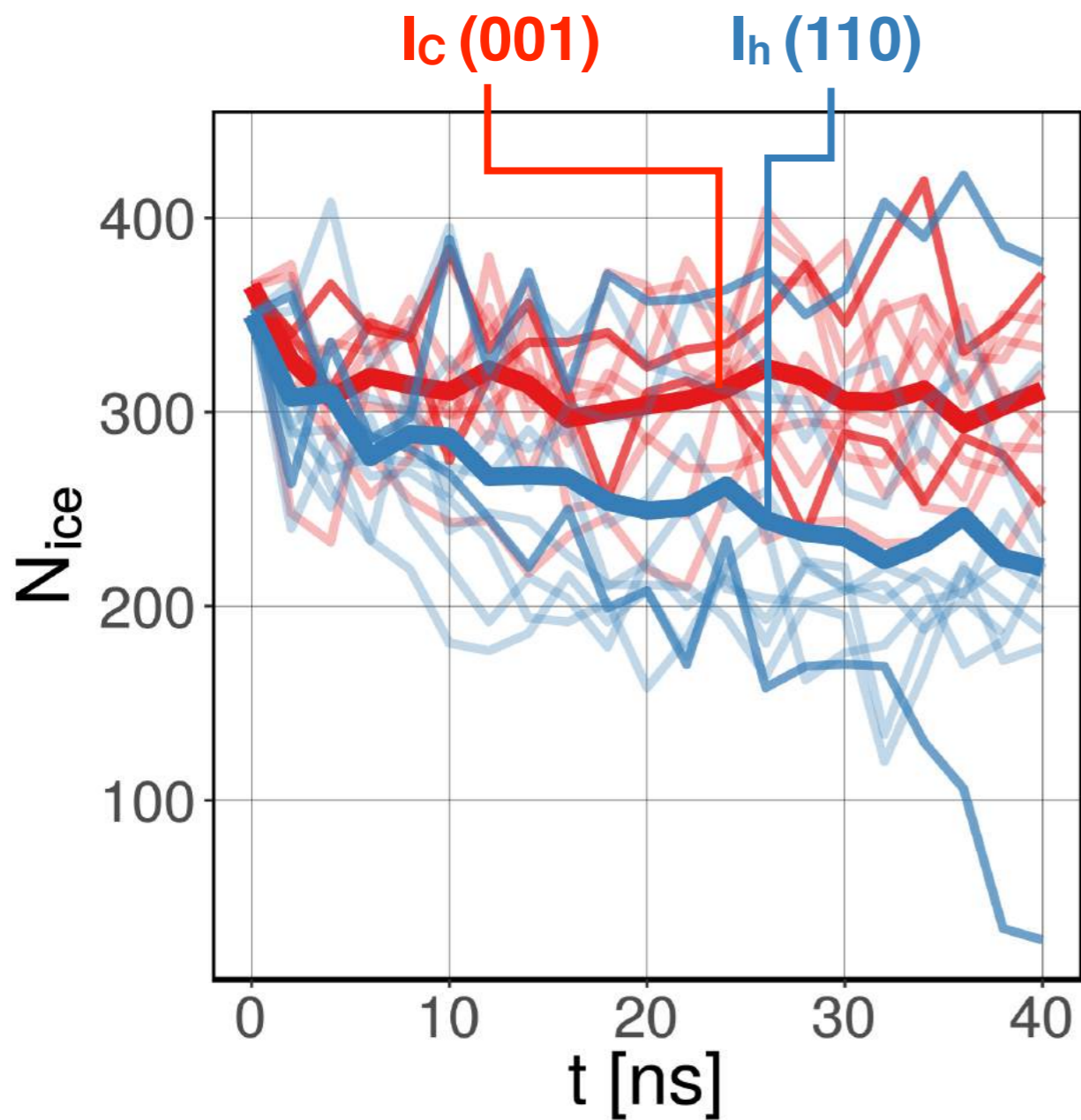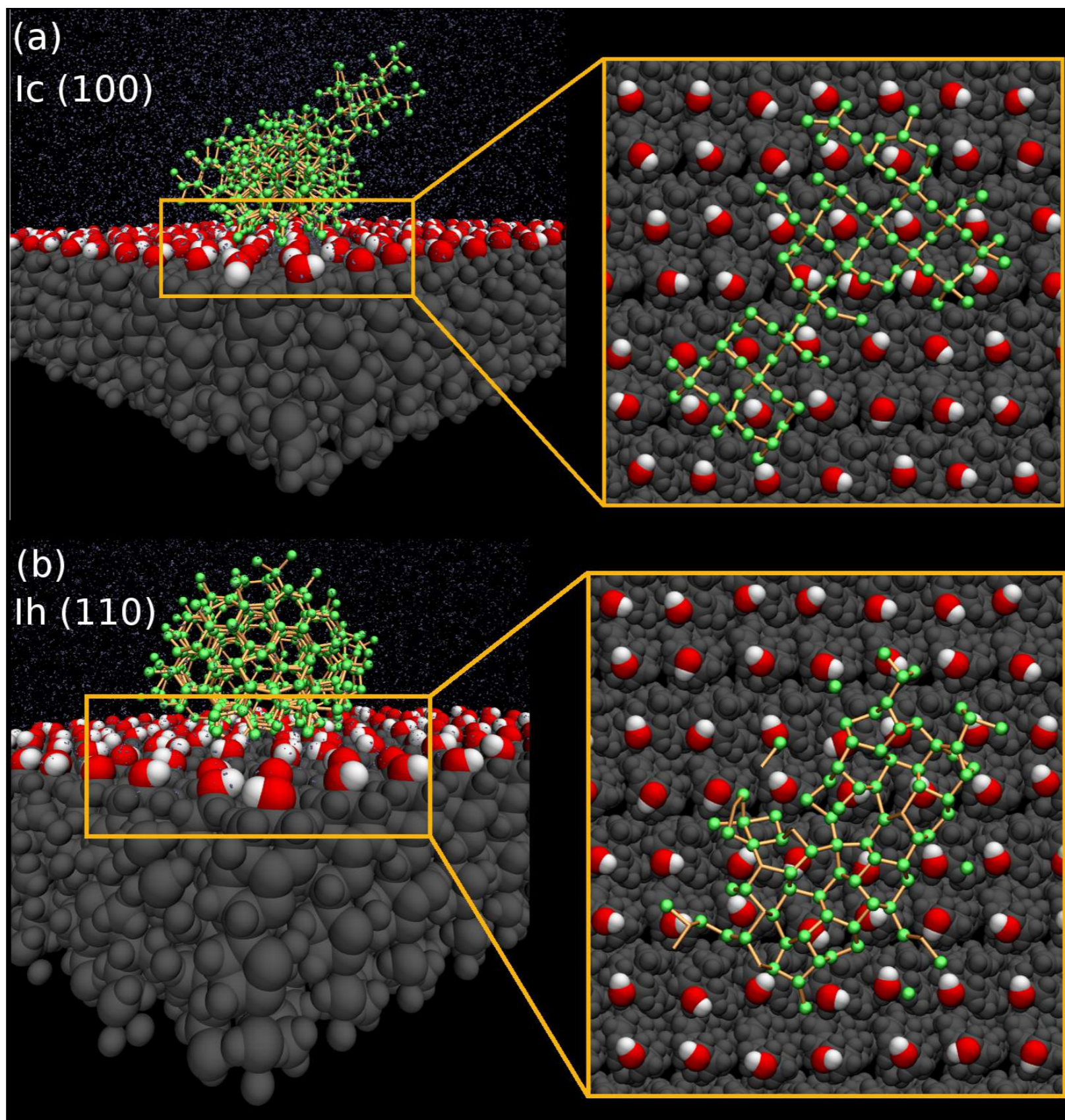- Random structure search to build the hydrogen bond network

⬇

Consistent with FFS!

Dealing with different ice polytypes and surfaces

HseededMD

Conclusions

- Molecular simulations have the potential to provide unique insight into crystal nucleation (and growth)

- The nucleation rate is the key quantity we need to build a bridge between experiments and simulations

- Predicting nucleation rates for realistic systems is challenging. We have a few options in terms of enhanced sampling methods (they are improving fast), but the accuracy of the force fields is the most pressing issues (building force fields is tedious, difficult, it does not pay in terms of publications, and nobody wants to fund it)

- Heterogeneous nucleation is even more challenging. Tailoring seeded MD methods is a possible way forward, but expertise from e.g. crystal structure prediction people is needed

- Ice in biological matter is a challenge for molecular simulations - but it's worth it!

Shameless advertisement: this is where I am going - join in!

Thanks

WARWICK
THE UNIVERSITY OF WARWICK

People

- Gareth Tribello (Queen's University Belfast)
- Davide Donadio (University of California Davis)
- Tianshu Li (George Washington University)
- Martin Fitzner & Angelos Michaelides (University College London)
- Thomas Whale & Ben Murray (University of Leeds)
- Alexei Kiselev (Karlsruhe Institute of Technology)
- Ellen Backus & Mischa Bonn (Max Planck Institute for Polymer Research)
- […]

Resources
[…]

CSCS

Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre