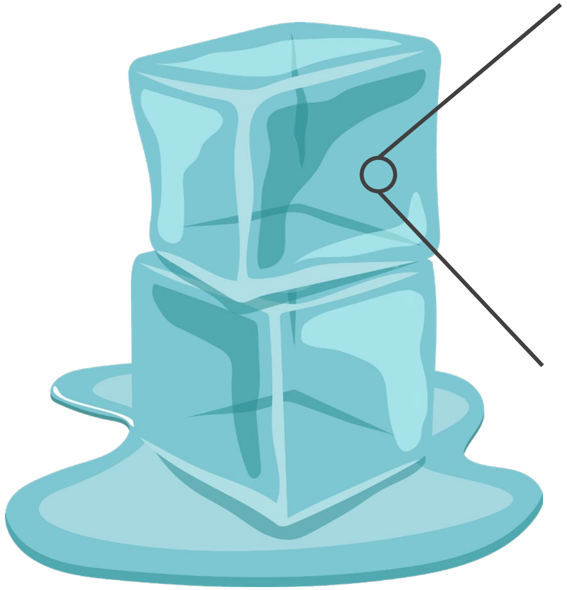
Chemical Sciences

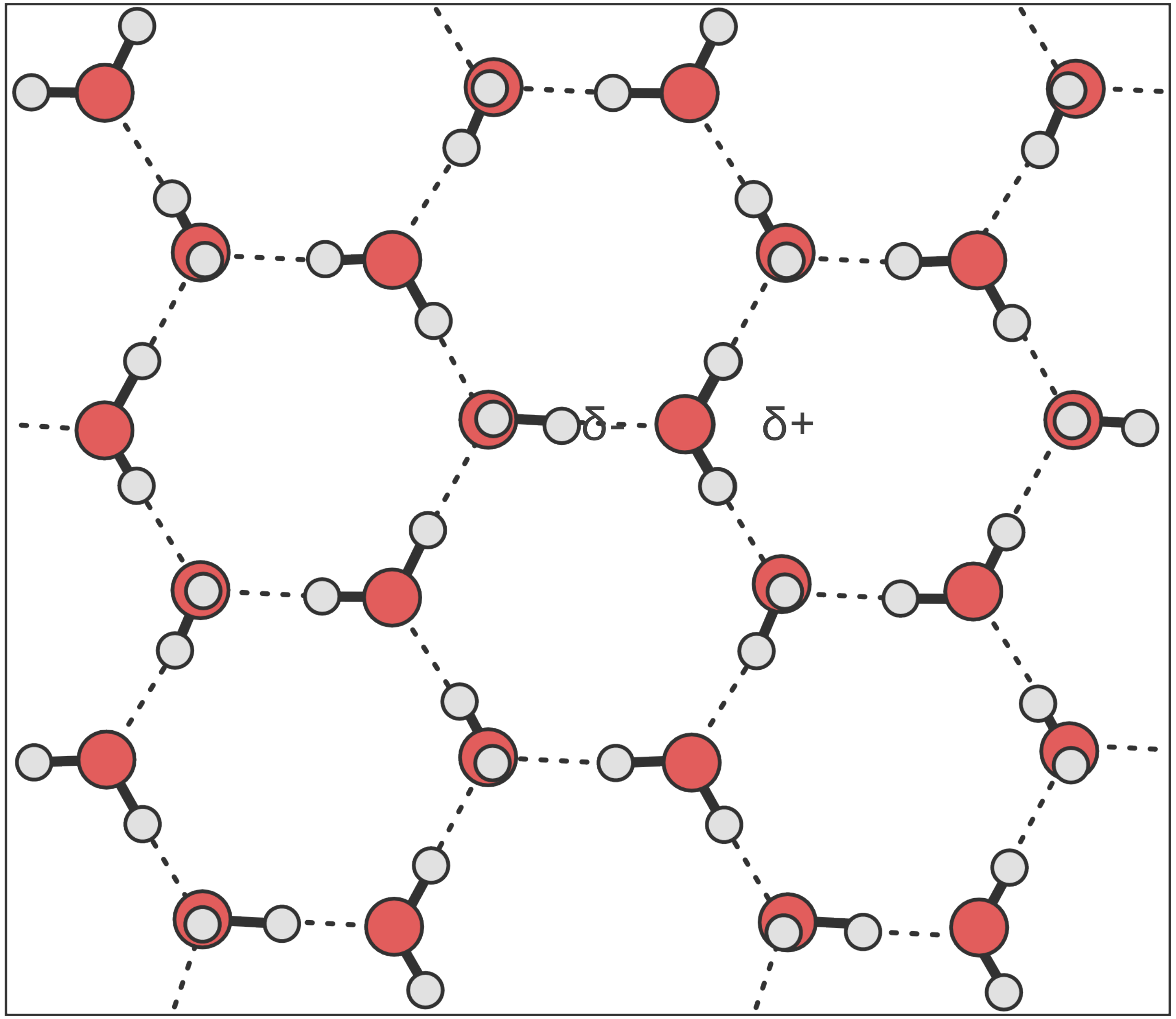
Pushing the limits of computational chemistry: the phase diagram of water

Massimo Bocus, Tom Braeckevelt, Pieter Dobbelaere, Arnout Maet,
Wim Temmerman, Sander Vandenhaute, Siebe Vanlommel, Jelle Vekeman, Veronique Van Speybroeck*

Center for Molecular Modeling, Ghent University

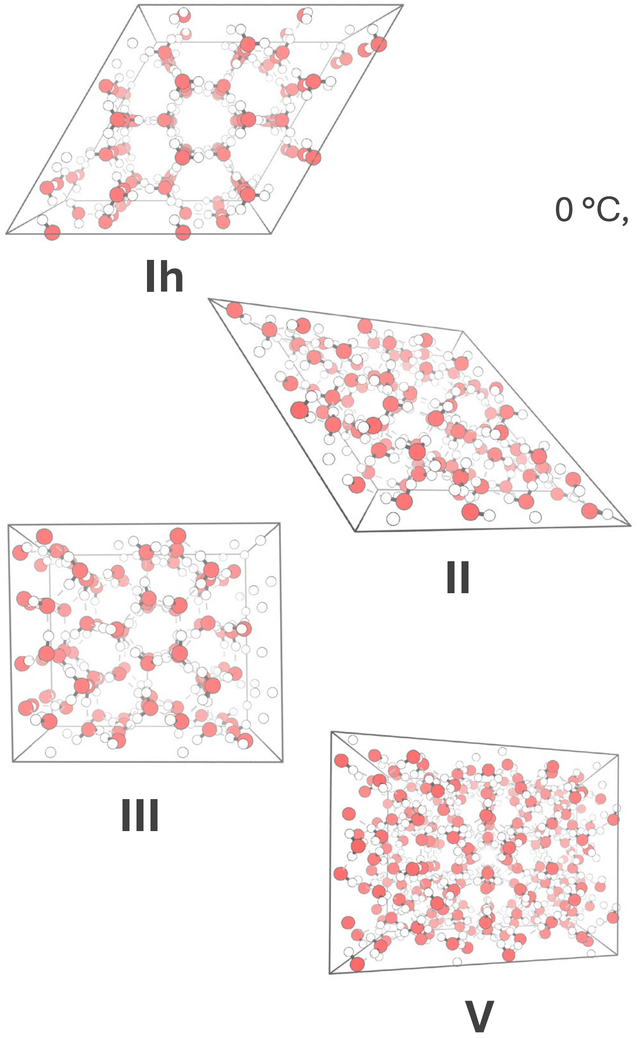
massimo.bocus@ugent.be, veronique.vanspeybroeck@ugent.be

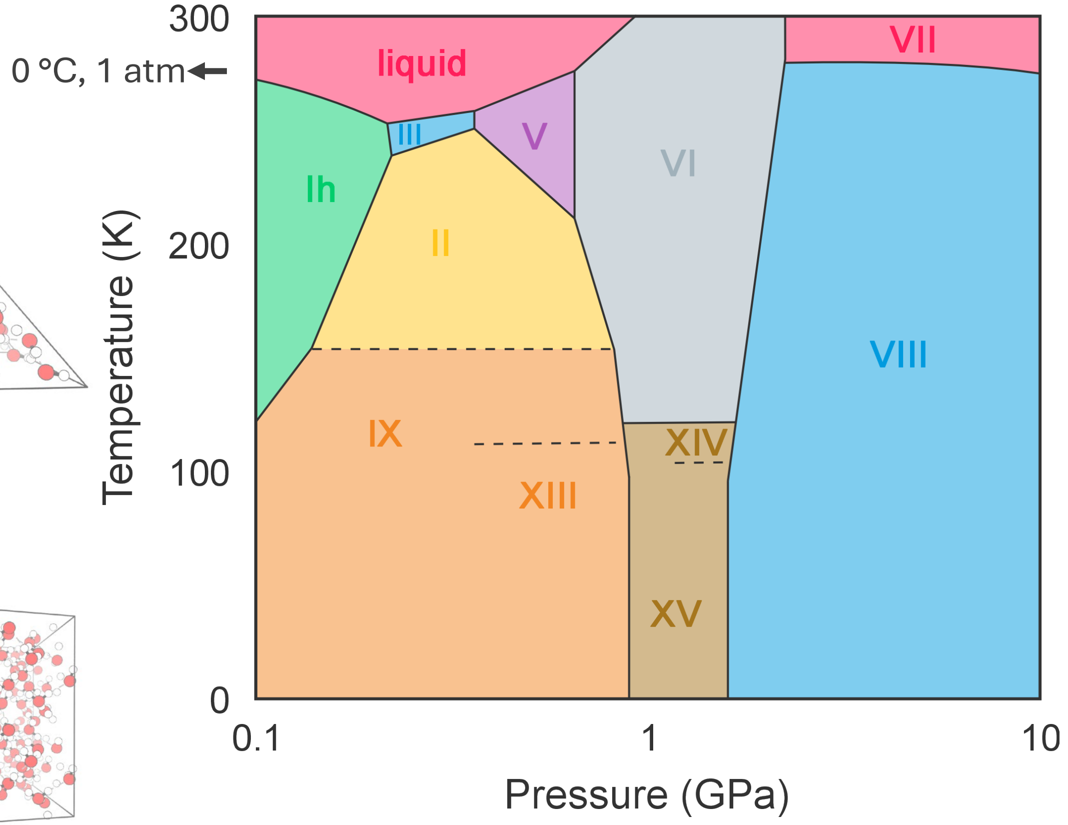




Ice **Ih**, i.e. “just ice”

The phase diagram of water is amazingly complex





Most lattices have a **proton-ordered** and a **proton-disordered** phase (e.g. **III** → **IX**)

Figure adapted from Hansen Nat. Comm. **12**, 3161 (2021)

Building a phase diagram requires the calculation of Gibbs' free energies

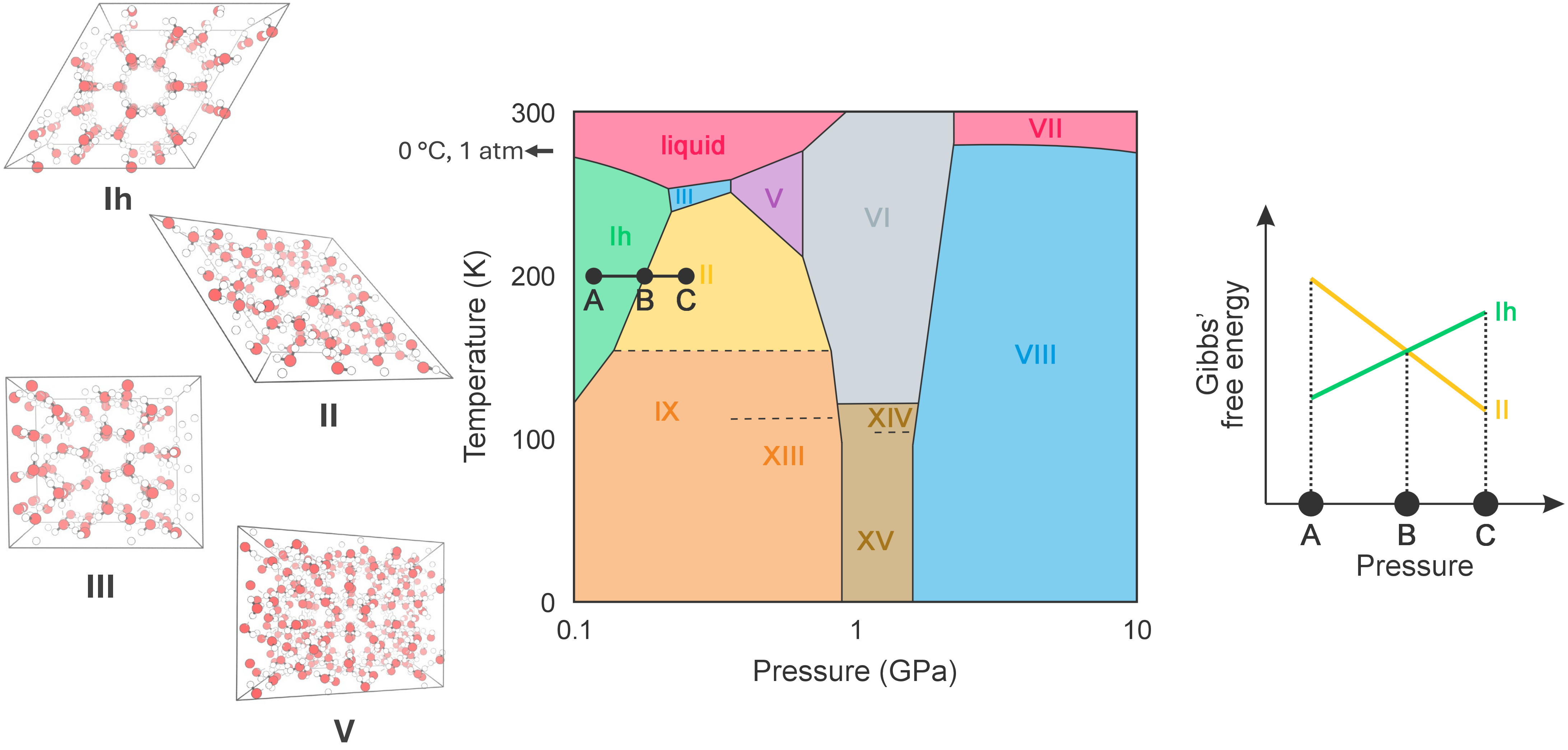
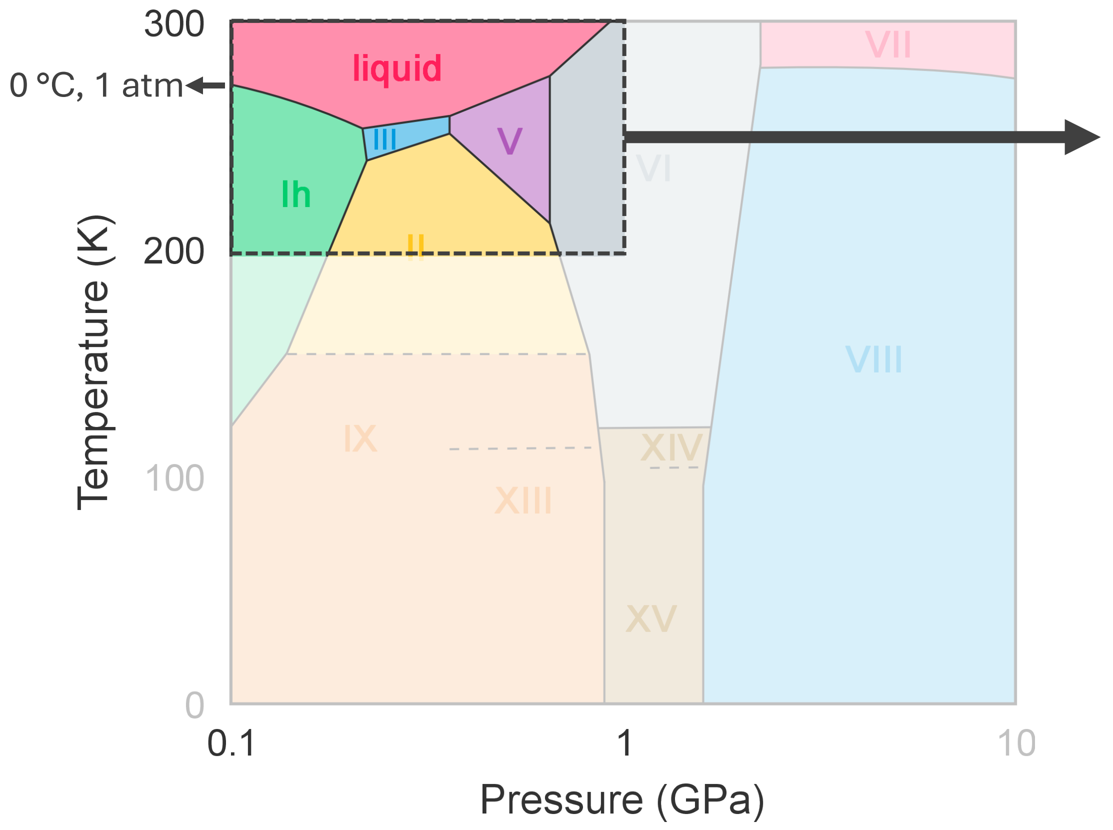
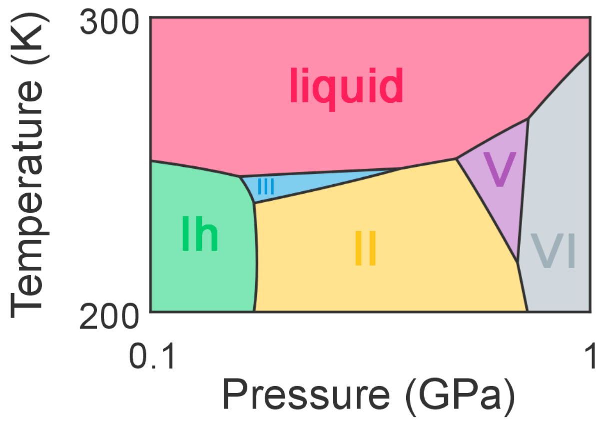


Figure adapted from Hansen Nat. Comm. **12**, 3161 (2021)

The state-of-the-art covers a “limited” range of temperatures and pressures





From quantum MB-pol simulations

Bore & Paesani Nat. Comm. **14**, 3349 (2023)

Our goal:
Obtaining an **accurate phase diagram** of water from **first-principles** over an unprecedented broad range of temperatures and pressures.

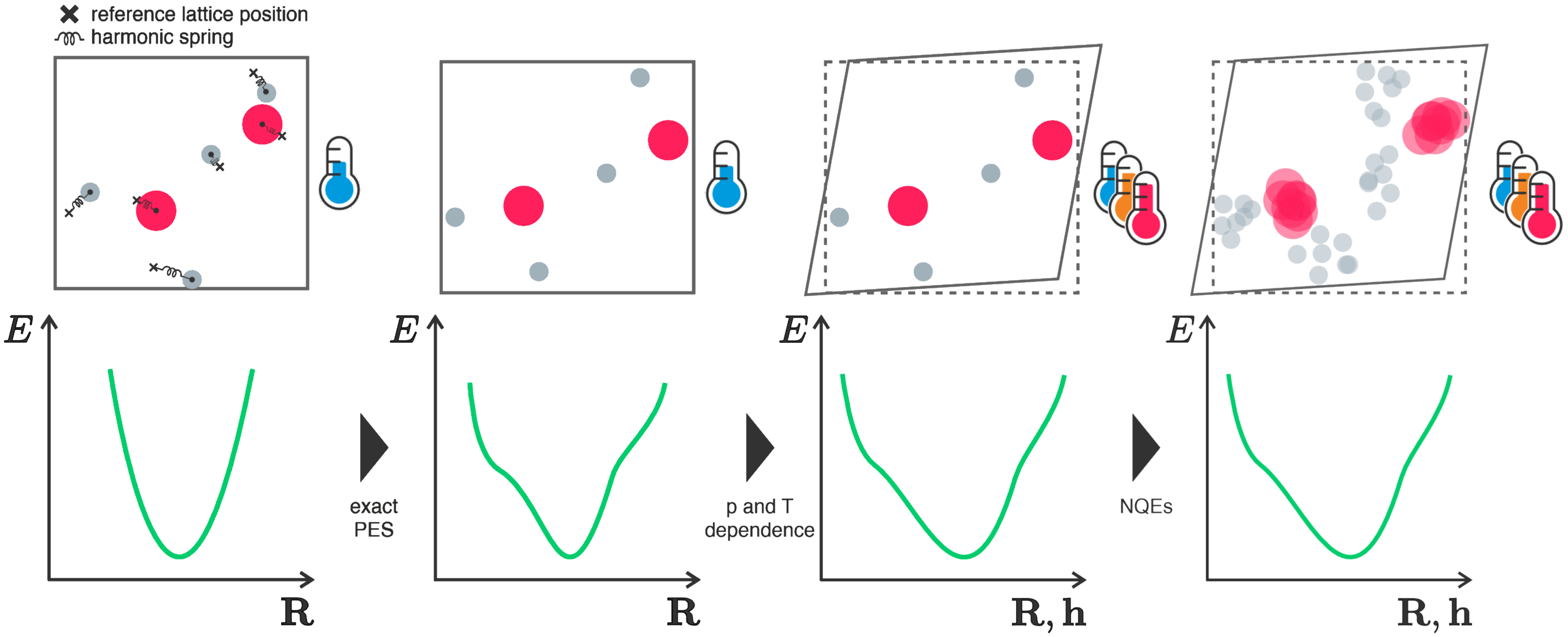
Outline

Thermodynamic integration to compute exact Gibbs' free energies

Machine learning potentials are necessary to perform the simulations

The theoretical phase diagram of water

The central problem: an absolute reference for the Gibbs' free energy is unknown



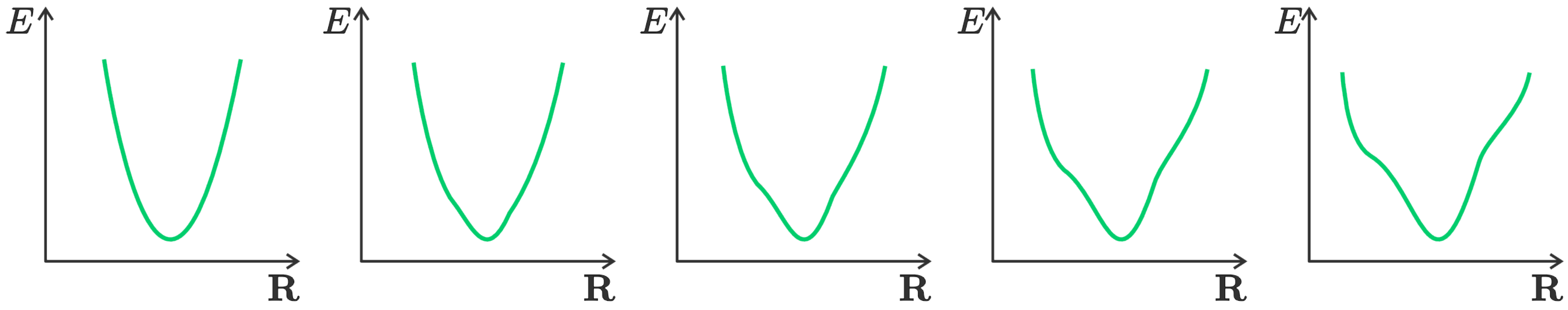
Provide an initial analytical **reference ΔG** between phases

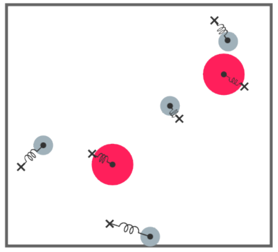
Correct the ΔG accounting for the **true potential energy surface**

Correct the ΔG for **high temperatures and isobaric conditions**

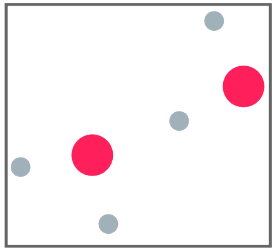
Include **nuclear quantum effects** to obtain the **exact ΔG**

Thermodynamic integration to compute the free energy difference between two states





$$E = (1-\lambda)E_{\mathrm{harm}} + \lambda E_{\mathrm{real}}$$



Each point requires a **long molecular dynamics** to converge the ensemble average.

Typical values are 15 MDs of more than 100 ps per correction.

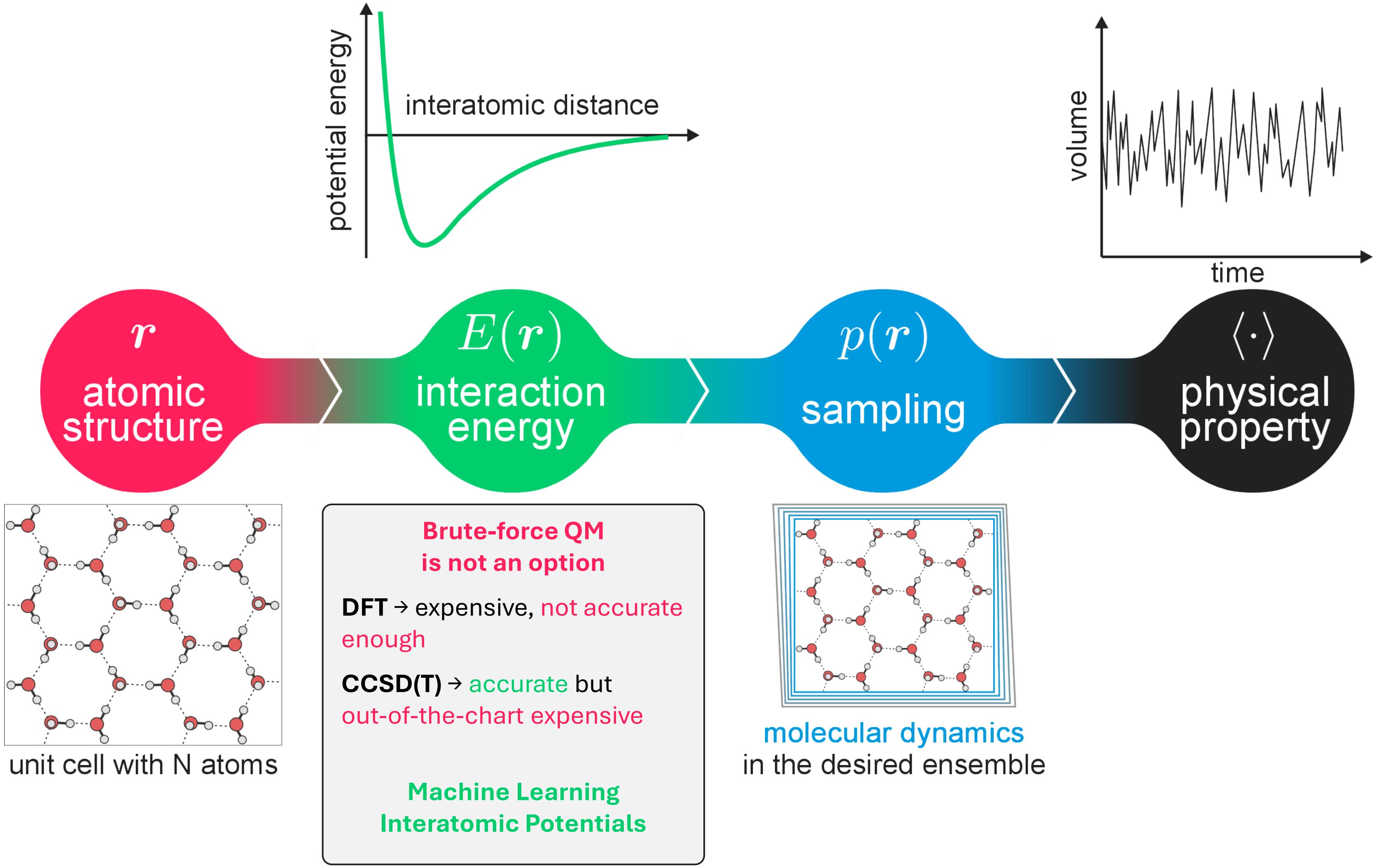
Outline

Thermodynamic integration to compute exact Gibbs' free energies

Machine learning potentials are necessary to perform the simulations

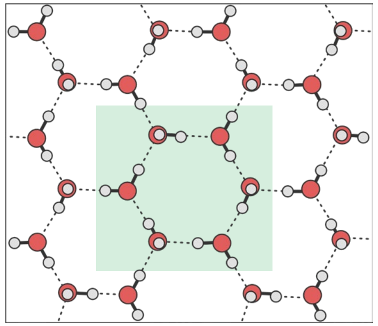
The theoretical phase diagram of water

Computing the interaction energy is the bottleneck of molecular modeling

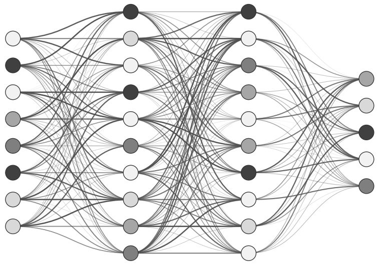


Van Speybroeck Phil. Trans. R. Soc. A **381**, 20220239 (2023)

MLPs can learn an ab initio PES



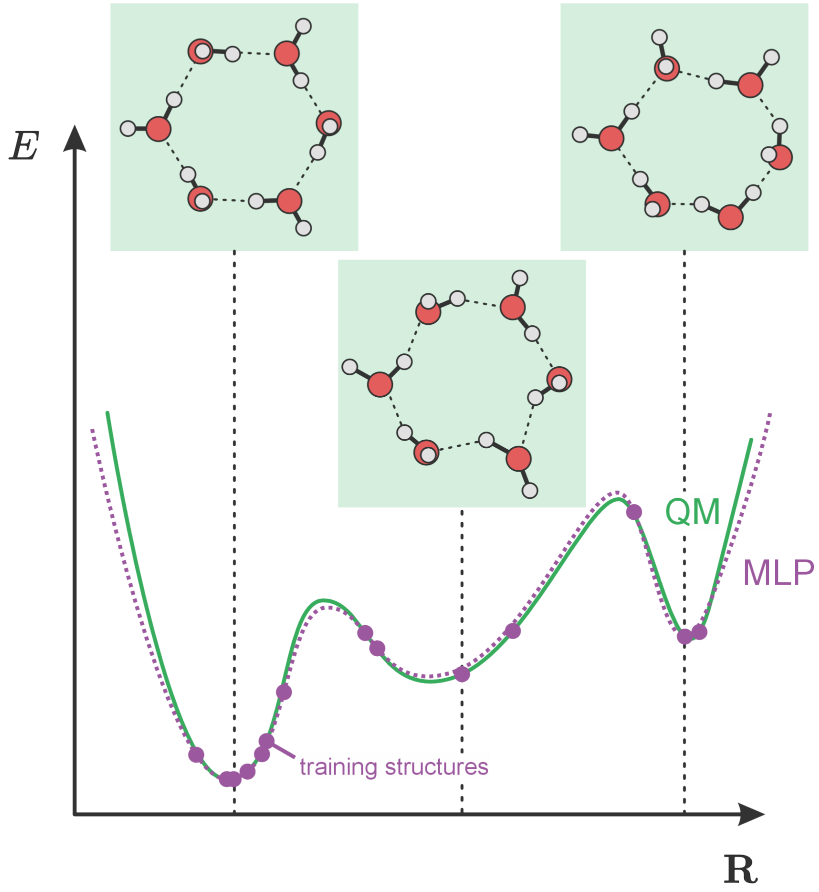
input structure



graph neural network (MACE)
trained on QM-evaluated structures

total energy

$$E = \sum_i E_i$$



- Training structures from **active learning** and literature.
- Reference level of theory **revPBE-D3**/TZV2P (GGA DFT).
- In progress: level of theory enhancement to CCSD(T).

Cheng et al. Proc. Natl. Acad. Sci. U. S. A. **116**, 1110-1115 (2019)

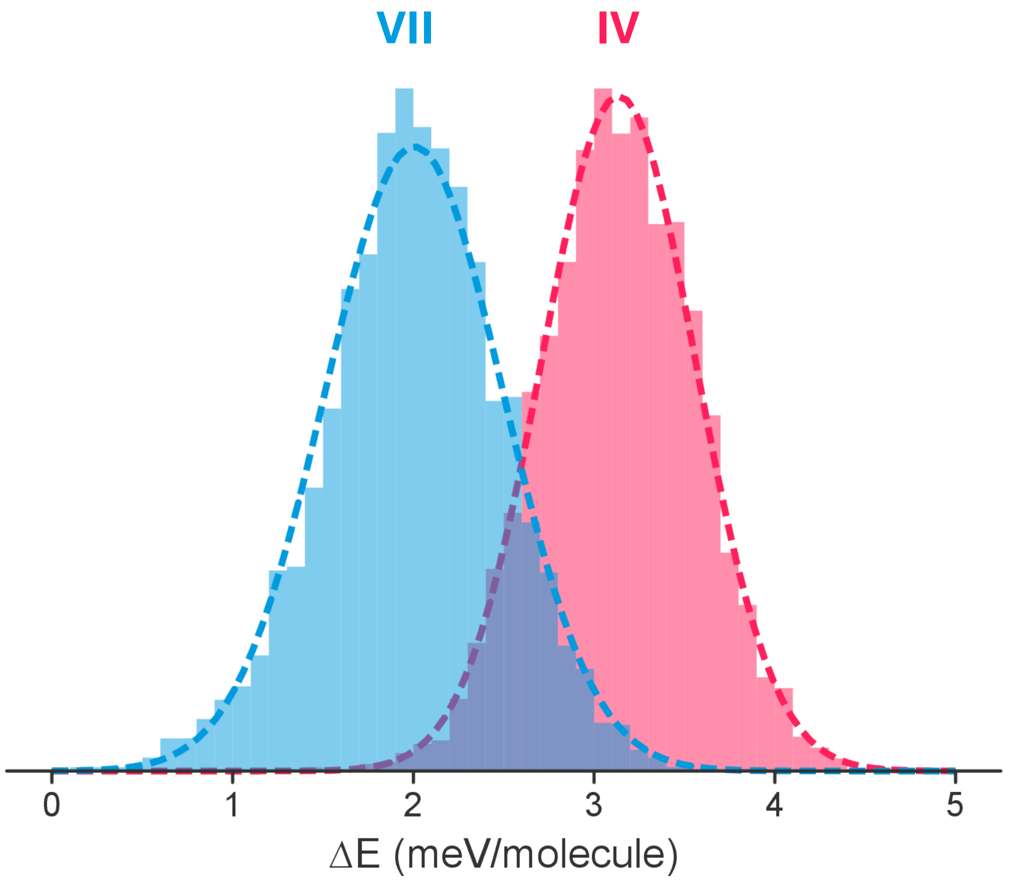
Outline

Thermodynamic integration to compute exact Gibbs' free energies

Machine learning potentials are necessary to perform the simulations

The theoretical phase diagram of water

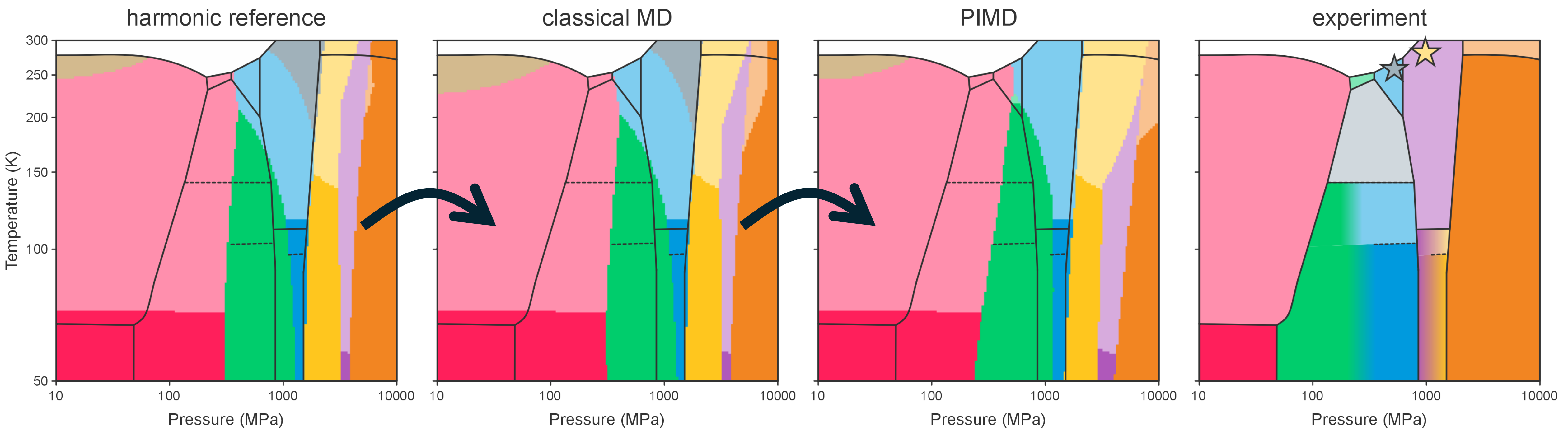
Intermezzo: how proton disorder is taken into account

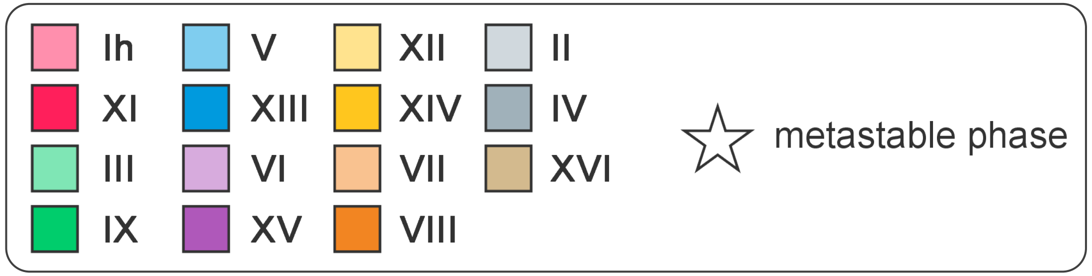


Proton-disordered structures are **generated randomly** and **optimized** to retrieve their energy w.r.t. the proton-ordered phase

$$S \propto \sum e^{-\beta \Delta E}$$

The computed phase diagram is in semi-quantitative agreement with the experiment

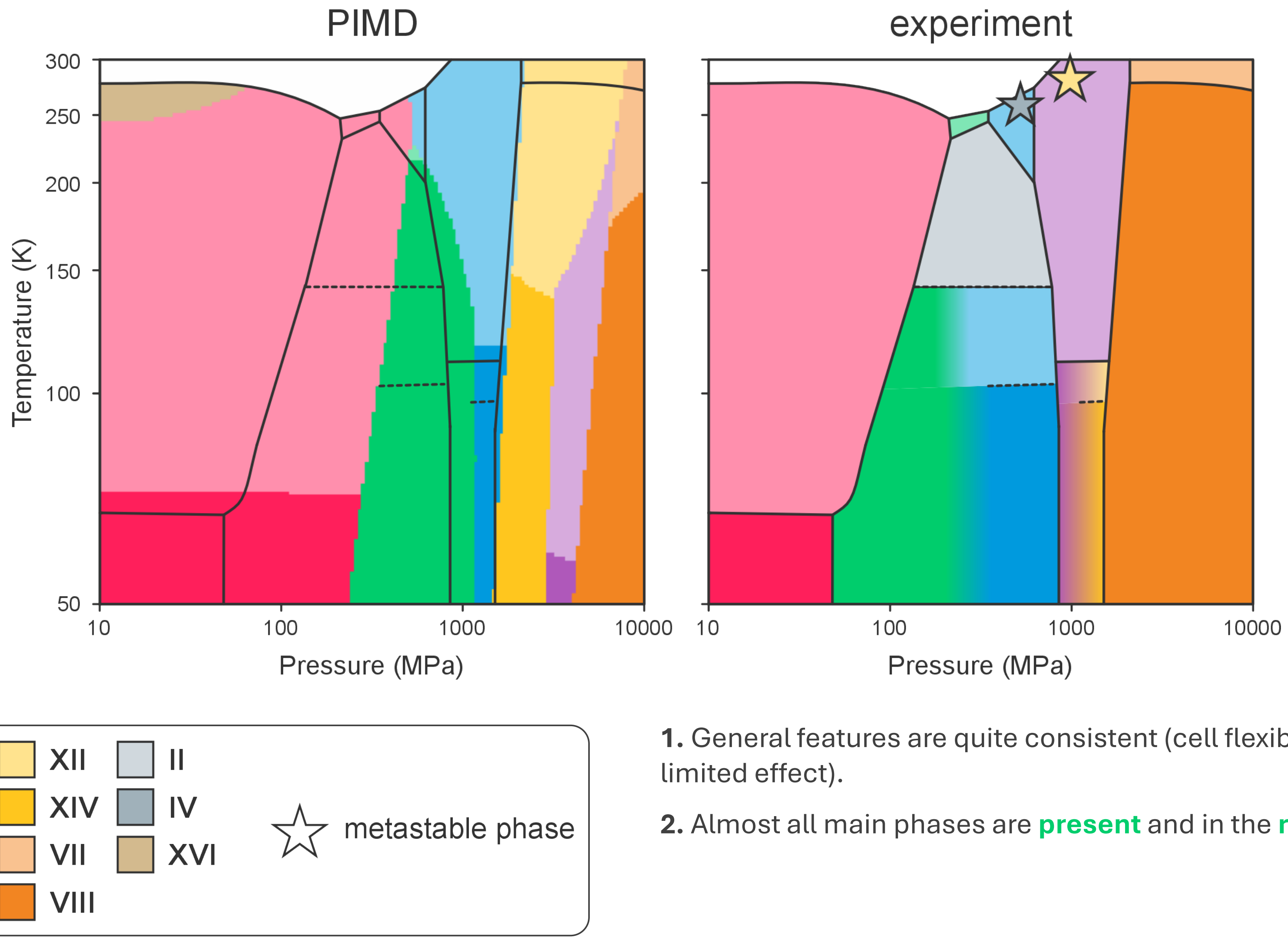




1. General features are quite consistent (cell flexibility and NQEs have limited effect).

Experimental diagram from Hansen Nat. Comm. **12**, 3161 (2021)

The computed phase diagram is in semi-quantitative agreement with the experiment

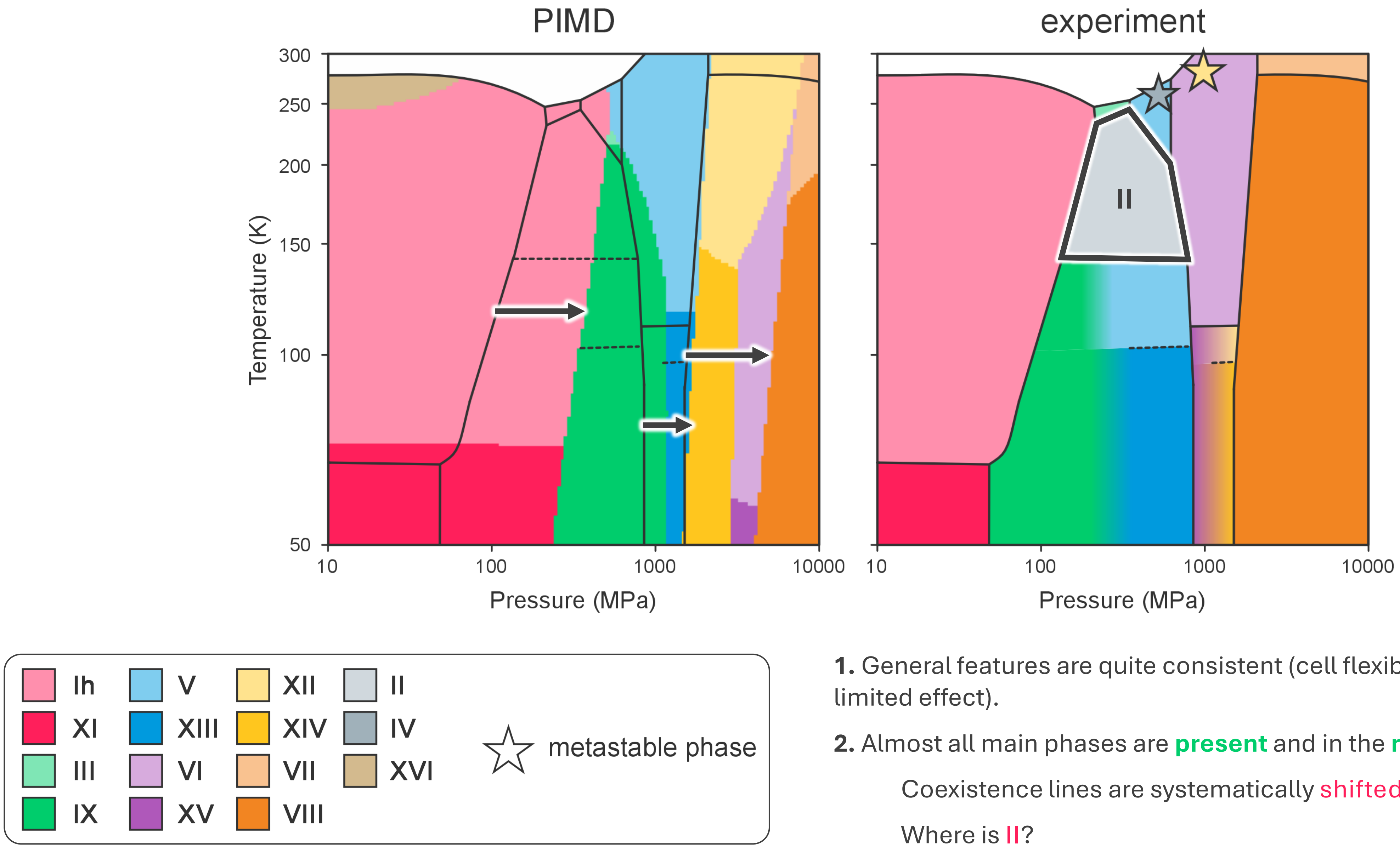


1. General features are quite consistent (cell flexibility and NQEs have limited effect).

2. Almost all main phases are **present** and in the **right location**.

Experimental diagram from Hansen Nat. Comm. **12**, 3161 (2021)

The computed phase diagram is in semi-quantitative agreement with the experiment



1. General features are quite consistent (cell flexibility and NQEs have limited effect).

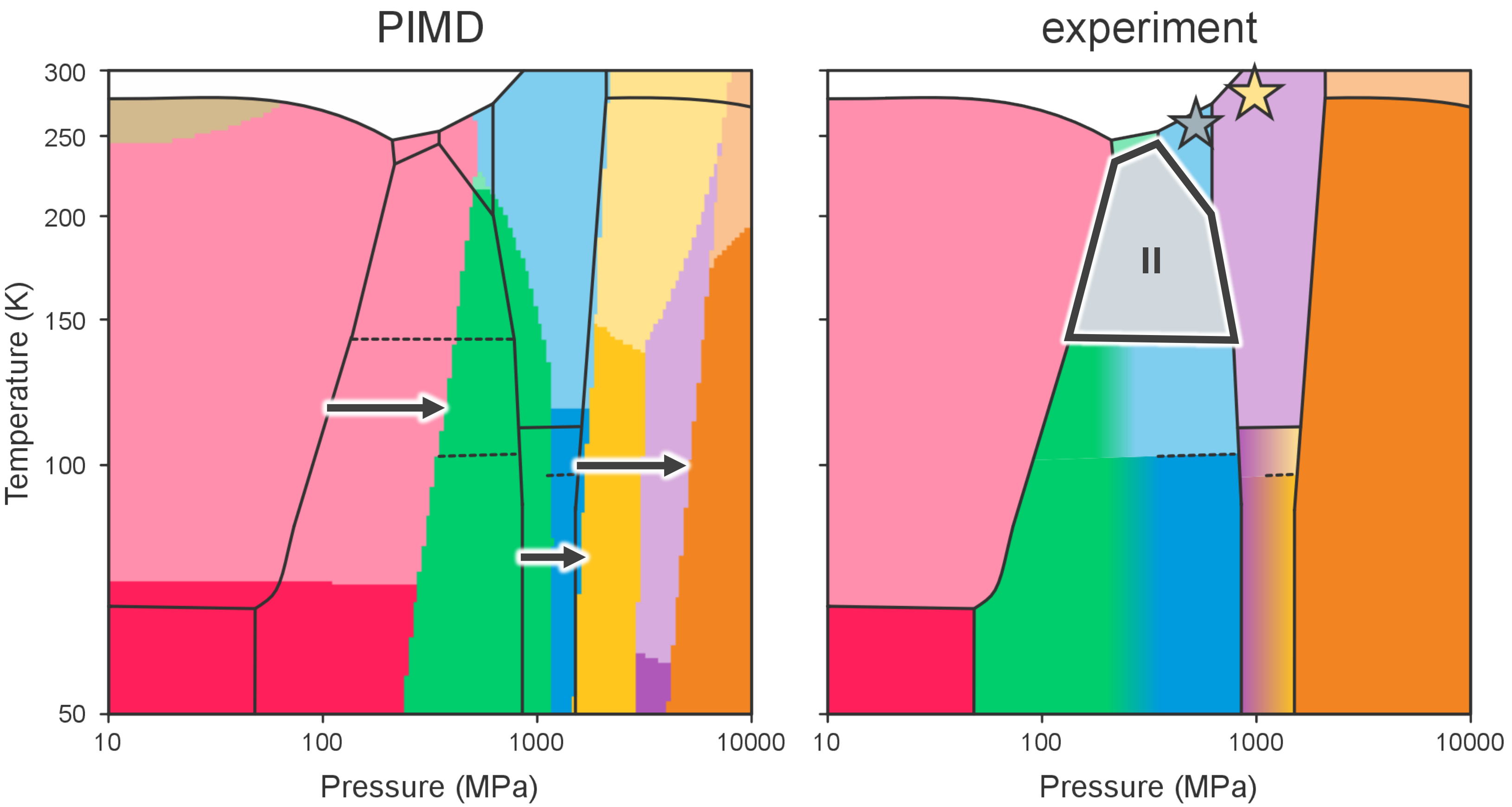
2. Almost all main phases are present and in the right location.

Coexistence lines are systematically shifted to high pressures.

Where is II?

Experimental diagram from Hansen Nat. Comm. **12**, 3161 (2021)

The computed phase diagram is in semi-quantitative agreement with the experiment



Train on **CCSD(T)** data to improve the agreement with experiment.

1. General features are quite consistent (cell flexibility and NQEs have limited effect).

2. Almost all main phases are **present** and in the **right location**.

Coexistence lines are systematically shifted to high pressures.

Where is II?

Experimental diagram from Hansen Nat. Comm. **12**, 3161 (2021)

CENTER FOR MOLECULAR MODELING

Contributors (in alphabetical order):

Massimo Bocus
Tom Braeckevelt
Pieter Cnudde
Pieter Dobbelaere
Mieke De Schepper
Arnout Maet
Jenna Mancuso
Siddarth Ravichandran
Berd Schmidt
Wim Temmerman
Sander Vandenhaute
Siebe Vanlommel
Jelle Vekeman

Principal investigator:
Veronique Van Speybroeck

EuroHPC Joint Undertaking

fwo

VLAAMS SUPERCOMPUTER CENTRUM

Vlaanderen is supercomputing

GHENT UNIVERSITY

CoSSPIA

Computational Screening of Sustainable Polymers for Industrial Applications

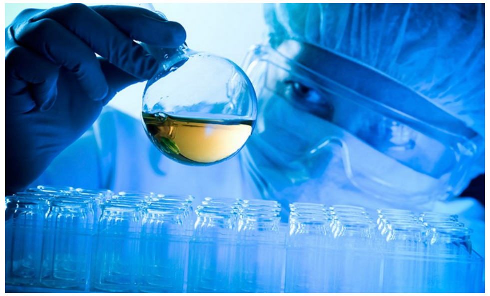
NEXTMOL™

30/09/2025
EuroHPC User Days 2025

Dr. Javier Ortín-Fernández, Application Engineer

Challenges for the chemical industry

Industry has a continuous need to improve its products.

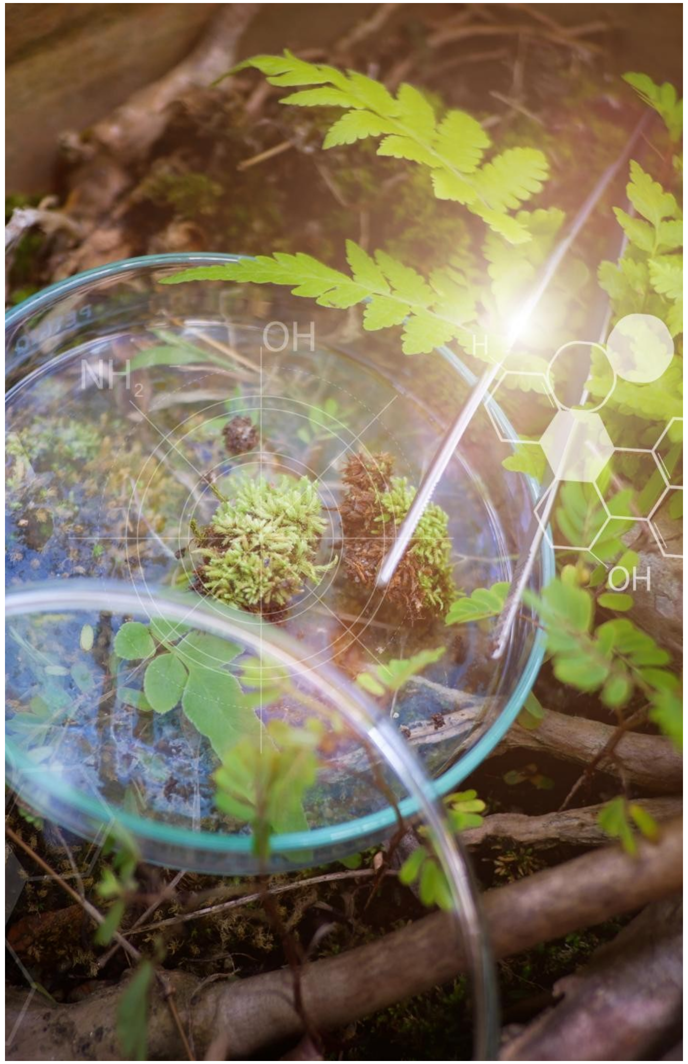


Performance drivers:
- Improvement of already existing products
- Competition within the market

Sustainability drivers:
- Sustainable raw materials
- Environmentally and health-safe
- Biodegradable

Cost drivers:
- Cheap raw materials
- Simple synthesis

Sustainable polymers are hard to develop



Home Care

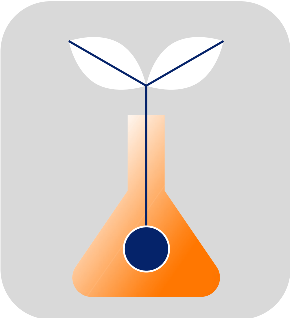
Persona
l
Care

Packaging & Plastics

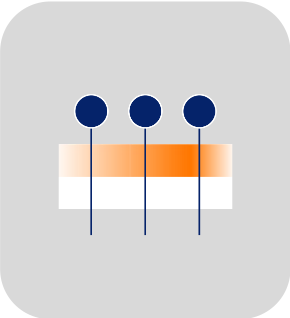
Adhesives & Sealants

Food Additives

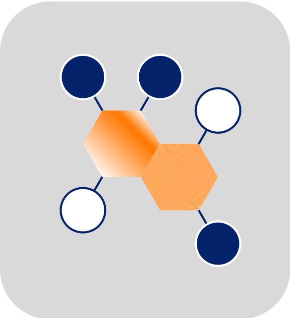
Need for sustainable alternatives



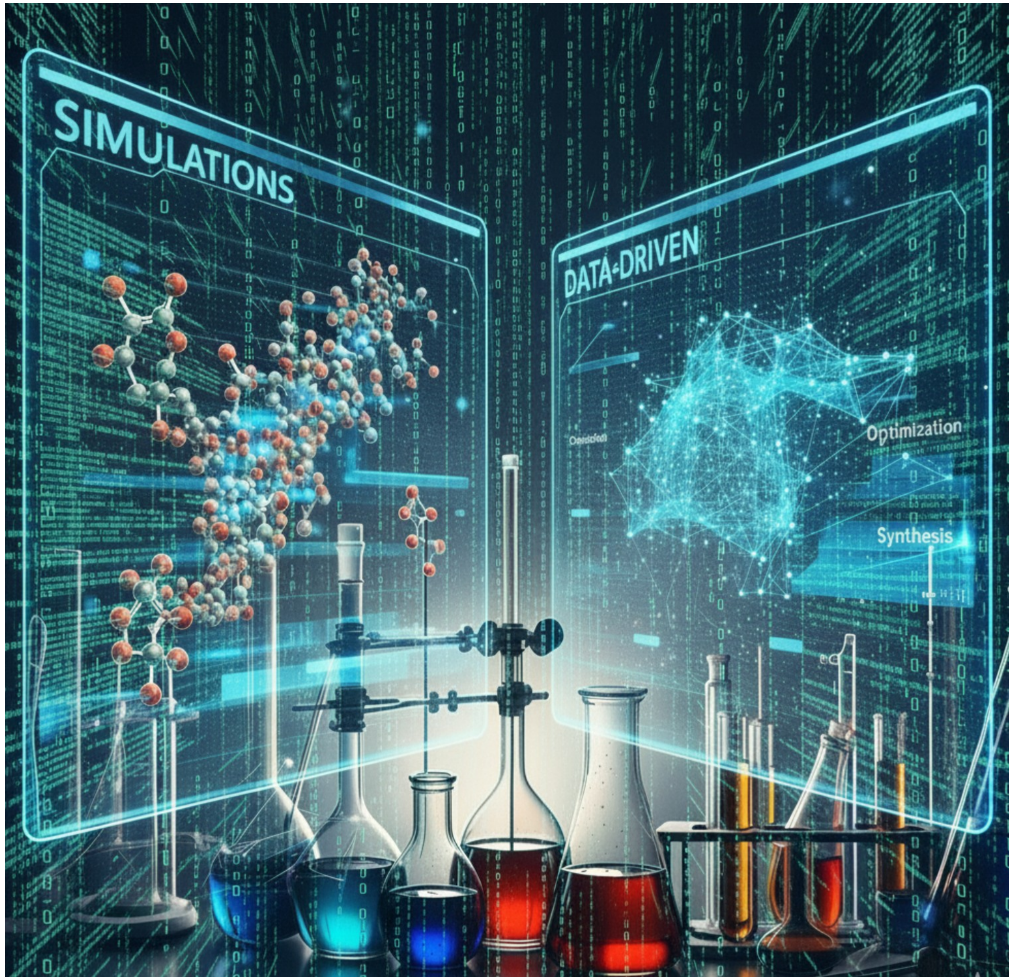
Challenges in polymer testing



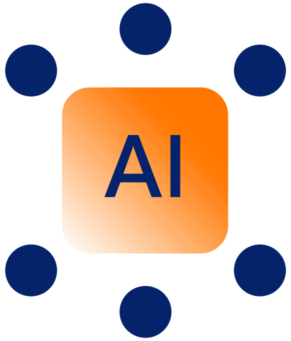
Design limitations



Our solution: digital pipeline to develop polymers



Molecular Dynamics



Artificial Intelligence

Scalable: assessment of thousands of polymer candidates

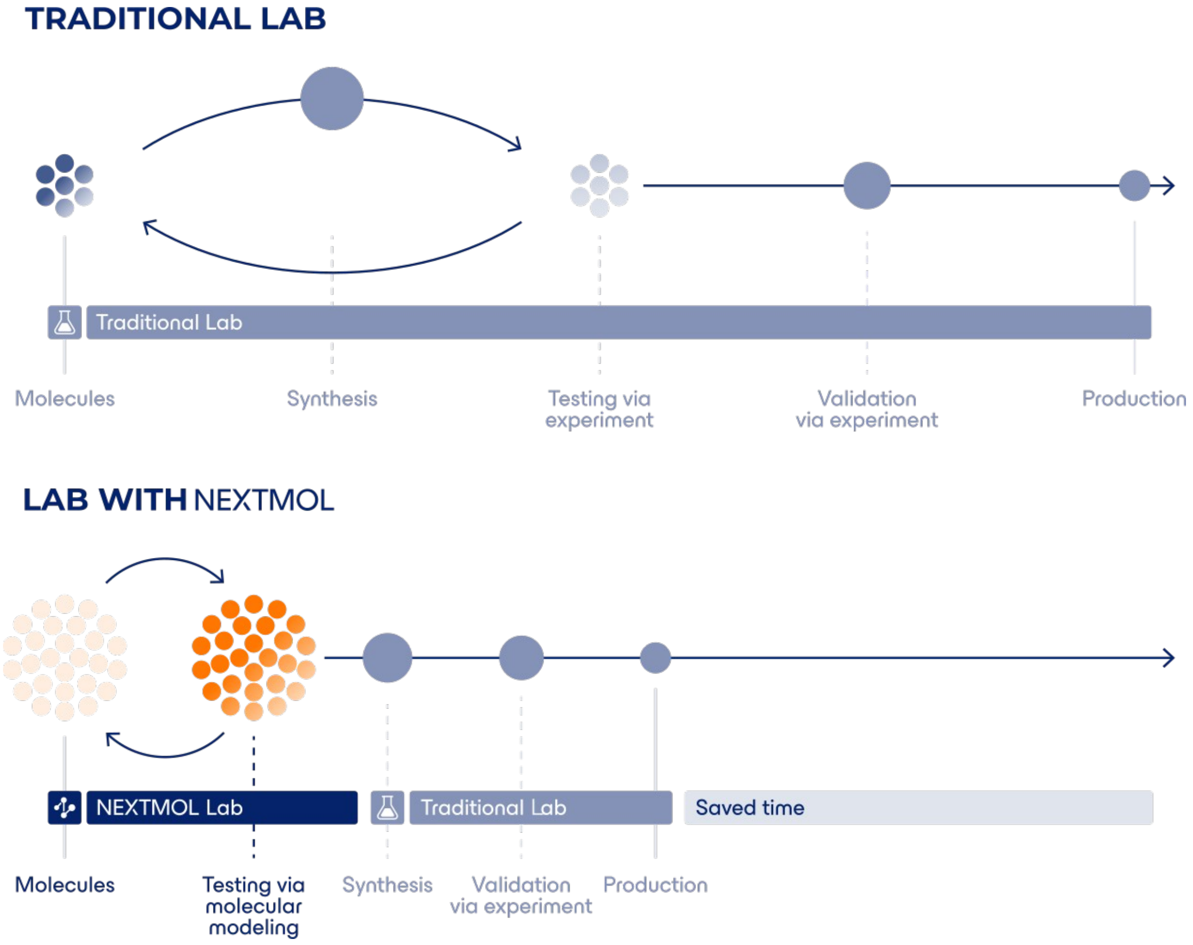
Consistent: reproducible and standardized predictions.

Digital: entirely *in-silico*.

Sustainable: Minimizes chemical waste and resource use.

Efficient: Reduces R&D costs while speeding up innovation

Digital lab vs traditional lab



The "traditional" (experimental) lab

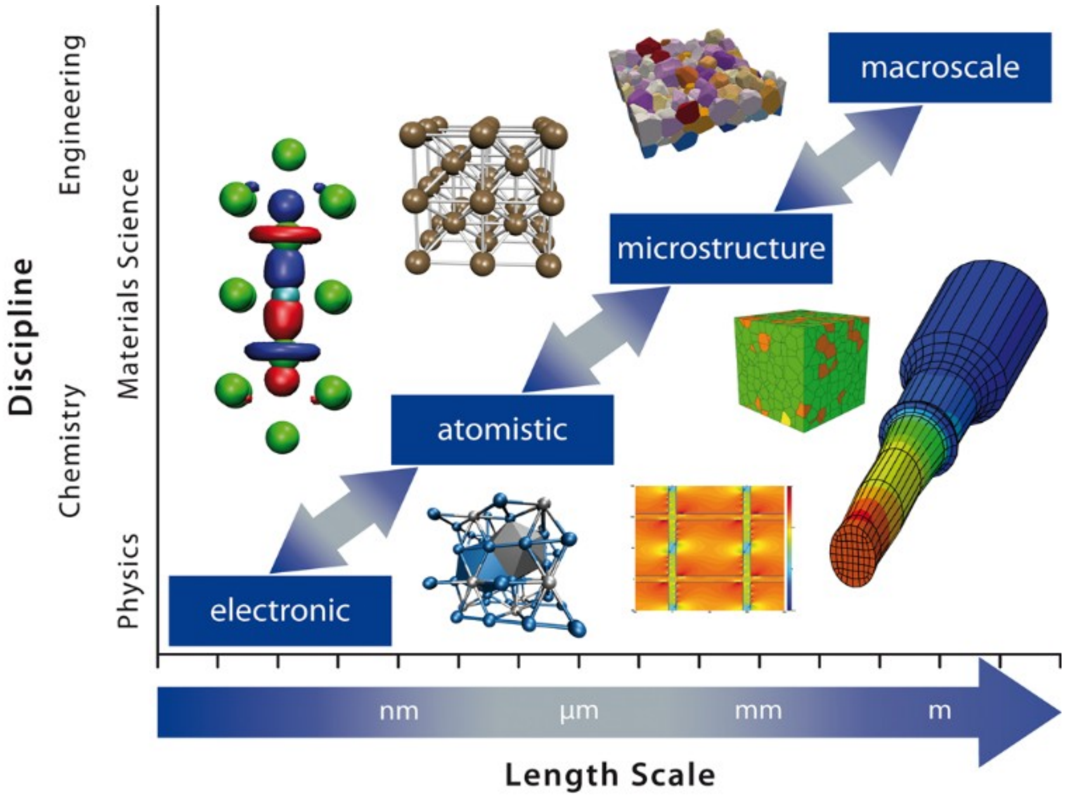
- Synthesis and experimental characterization and testing
- If the test is not successful, the same process has to be iterated.
- Inefficient, slow, tedious and error-prone.

The digital lab

- All testing is digital. Only the most promising candidates are synthesized and tested experimentally.
- More trial systems in less time, making the process faster, cheaper and more efficient.

The first ingredient: computational modeling

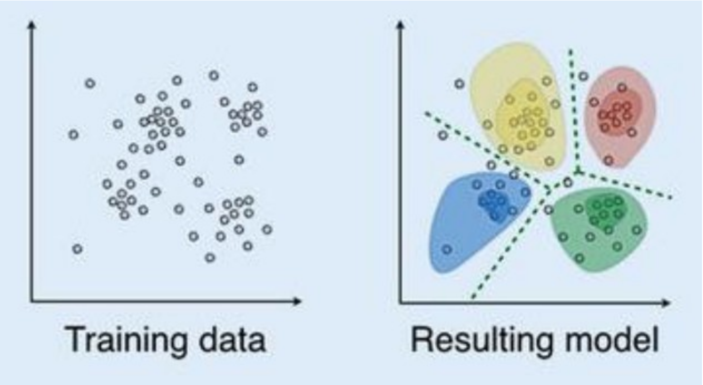
Modeling is a multiscale problem that spans more than ten orders of magnitude.
For CoSSPIA we used atomistic modeling based on classical **Molecular Dynamics**.



The second ingredient: data-driven ML

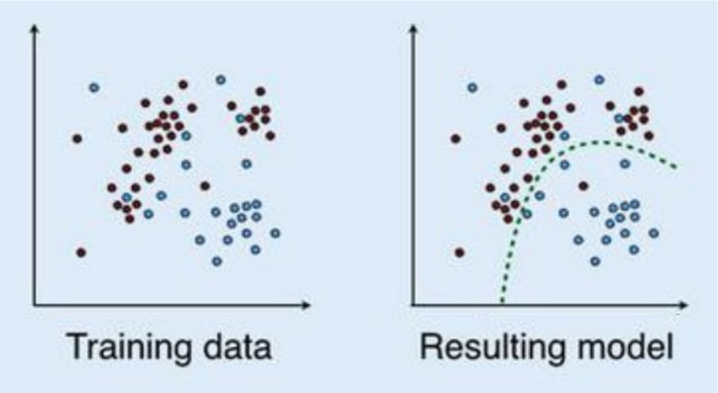
Unsupervised Learning

- No labeled data required.
- Chemical space insights: visualize and understand relevant regions.
- Smart prioritization & exploration: focus computational (and experimental) work on key areas of interest.



Supervised Learning

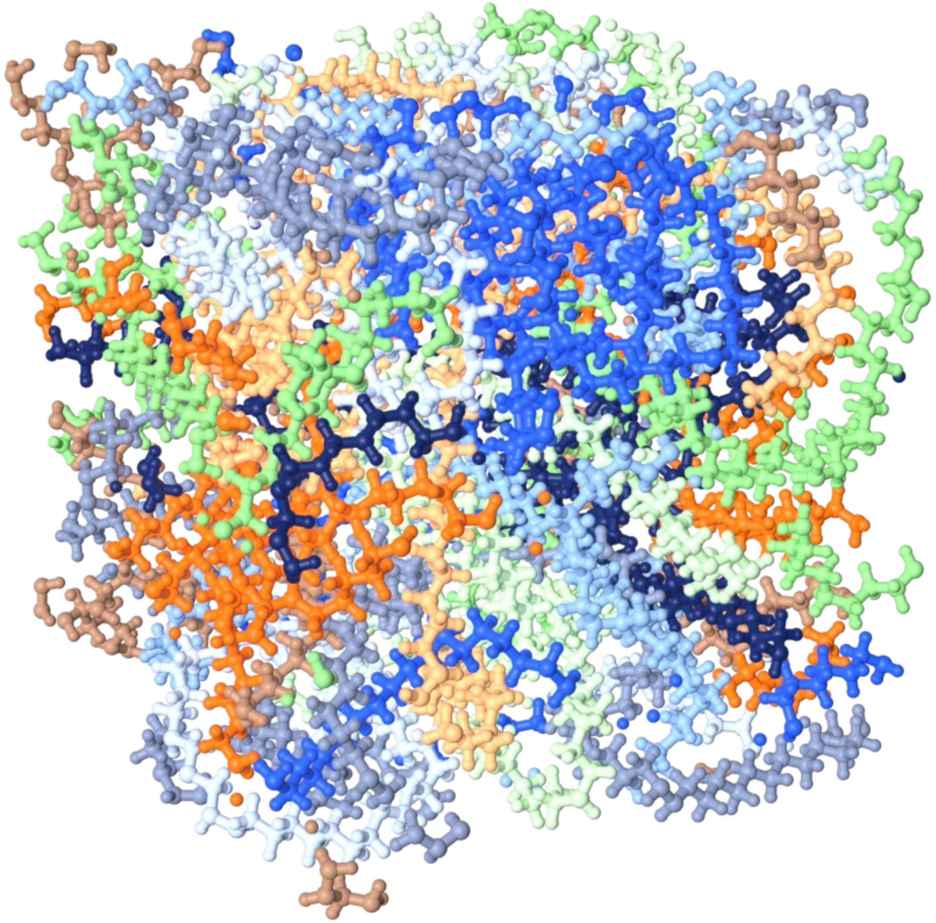
- Predictive: discover well-performing candidates.
- Structure–activity insights: identify key relationships.
- Main challenge: large, high-quality datasets.



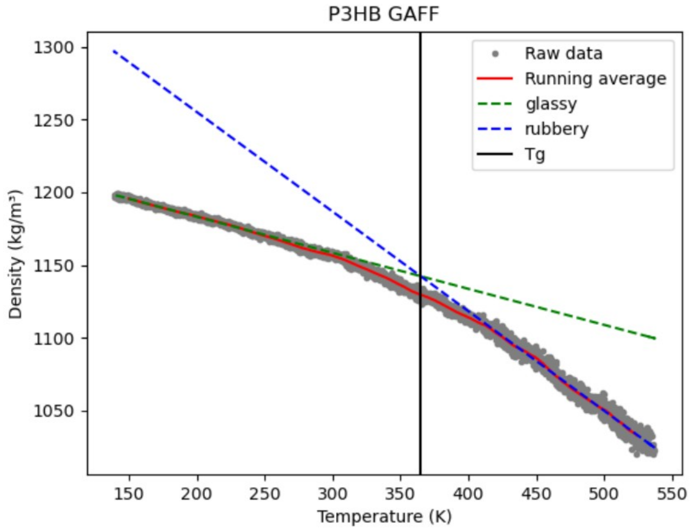
Computational descriptors

Within CoSSPIA we focused on the following physico-chemical descriptors of polymers:

- **Glass transition temperature (Tg),**
- Viscosity,
- Young's modulus,
- Solubility,
- Interfacial tension.

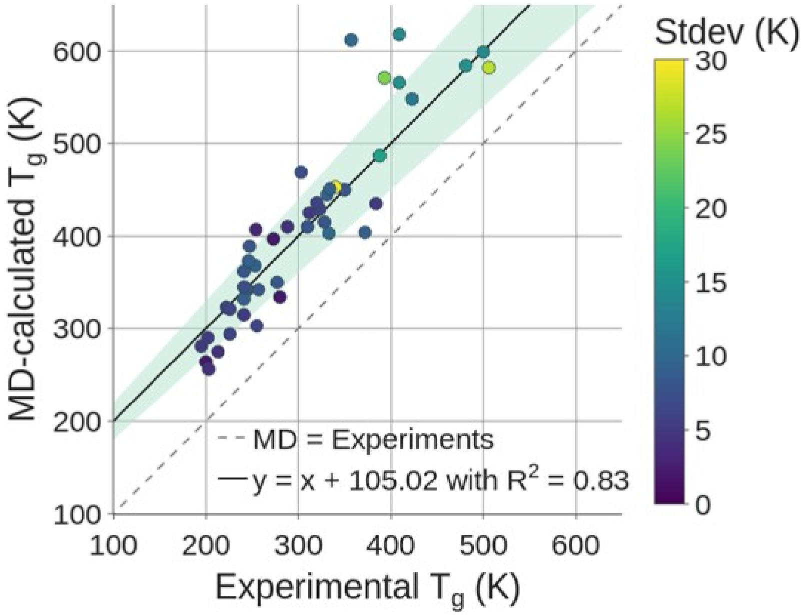


MD simulation to predict the Tg of polymers



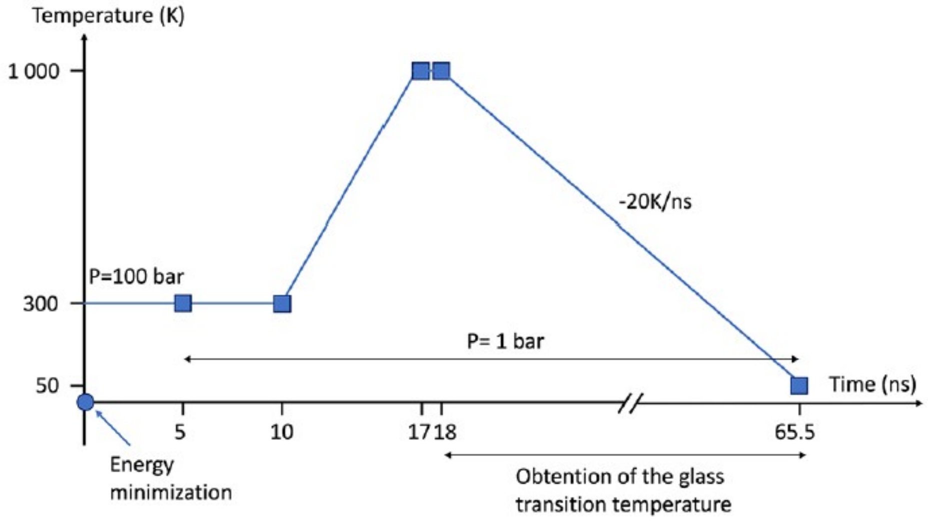
Four replicas
per polymer

Annealing Molecular
Dynamics Simulations



The change in the slope
indicates the Tg

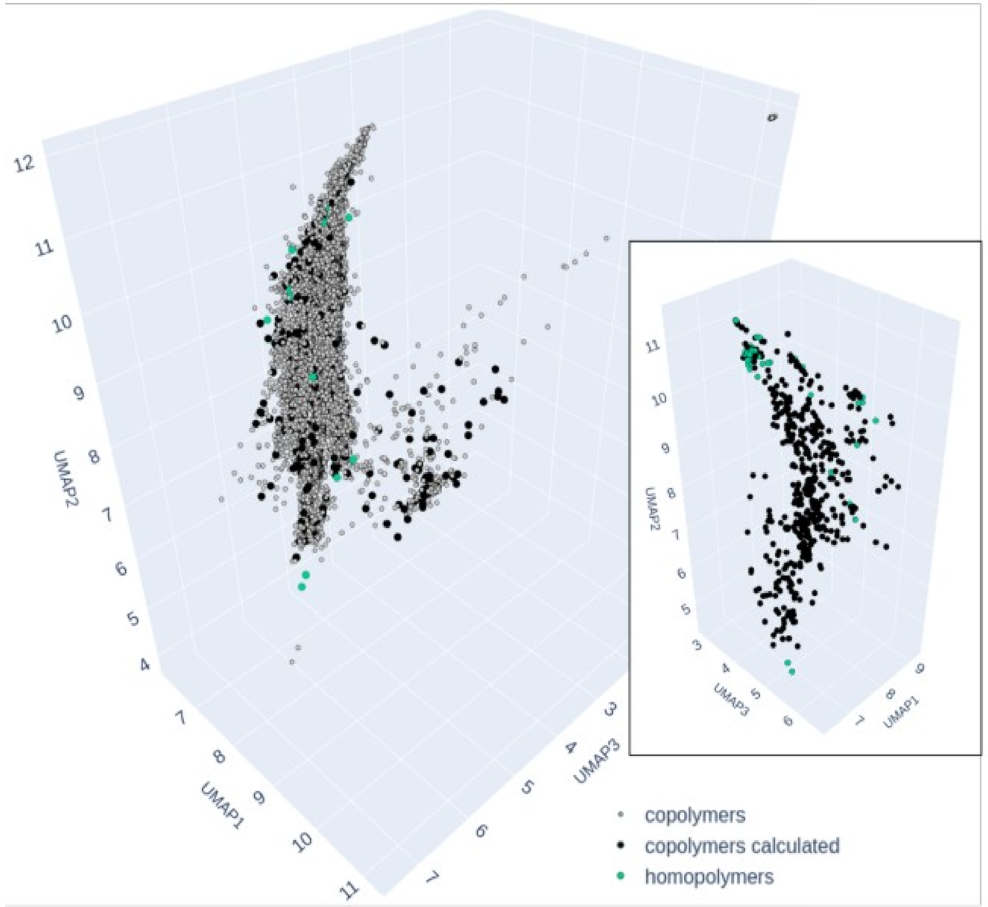
Corresponds to transition
from glassy to rubbery



Simulations reproduce
the experimental trends

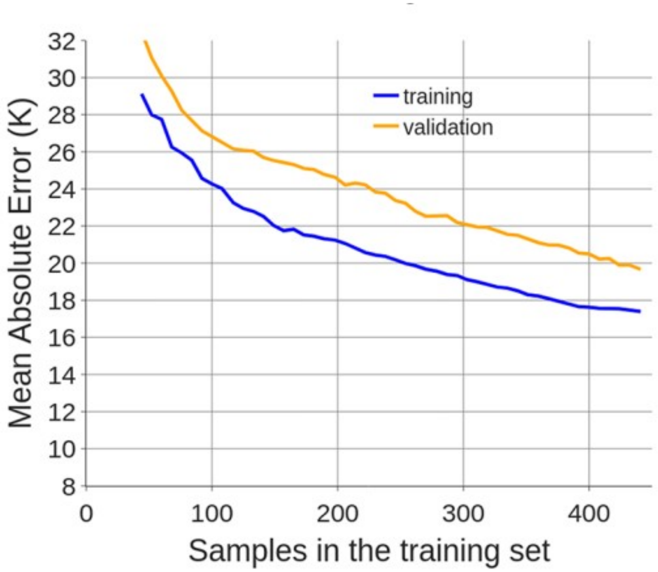
They display a constant
offset

HT screening to generate data and train an ML model

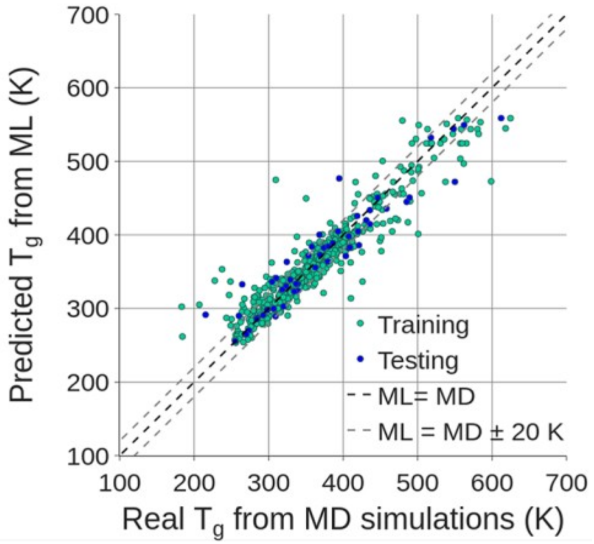


~15.000 polymers generated,
~500 polymers selected for MD

Low MAE,
small ovefitting



Model generates
almost identical
results to MD

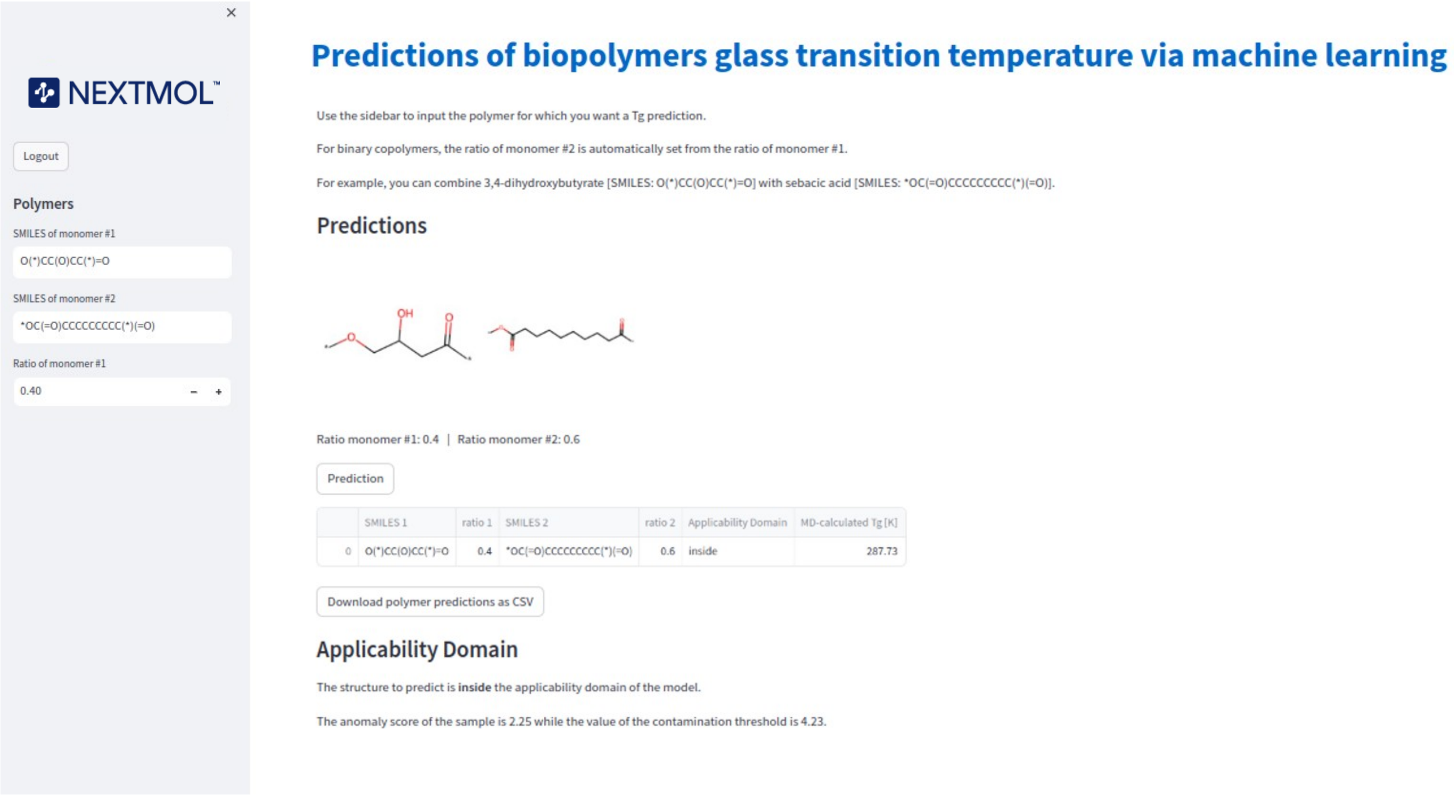


Glass transition temperature of polymers

- Model works for homopolymers and binary copolymers.
- Input: SMILES code of monomers and mixing ratio in copolymers.
- Almost instantaneous prediction.
- Allows rapid screening of thousands to millions of polymers in short time.
- The tool is accessible via a simple and intuitive web app:

  https://biopolymer-ml-pub.nextmol.com/

- Predictive ML models for other properties are under preparation.



Conclusions

- **Industry Challenge:** Need for high-performance, sustainable, cost-effective polymers.
- **Problem:** Traditional polymer development is slow, costly, inefficient, and limited by experimental capacity constraints.
- **Our solution:** Digital pipeline combining Molecular Dynamics (MD) and Machine Learning (ML) for *in-silico* polymer design.
- **Benefits:**
  - Scalable, predictive, and sustainable approach.
  - Reduces R&D costs and chemical waste.
  - Accelerates innovation.
- **Key achievements of CoSSPIA:**
  - High-throughput screening of thousands of polymers.
  - Database of physico-chemical descriptors.
  - ML models that can predict key physico-chemical descriptors for millions of polymers.
- **Impact:** Enables rapid, large-scale screening and supports circular economy goals.

Designing the next molecules the world needs!

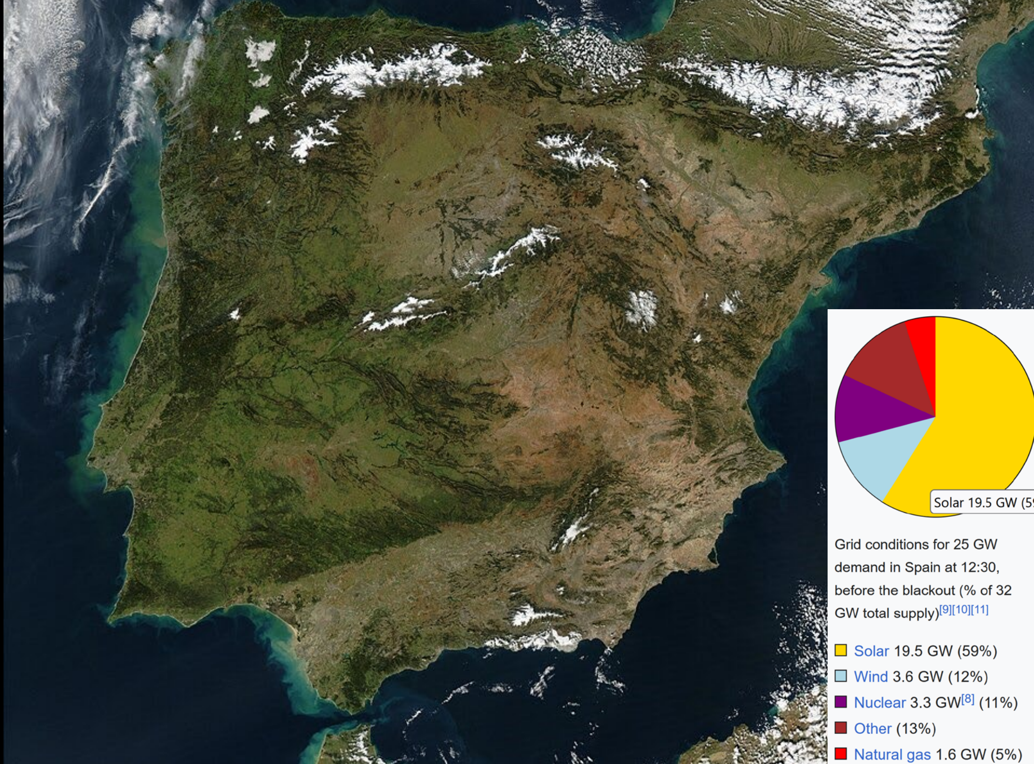
javier.ortiz@nextmol.com

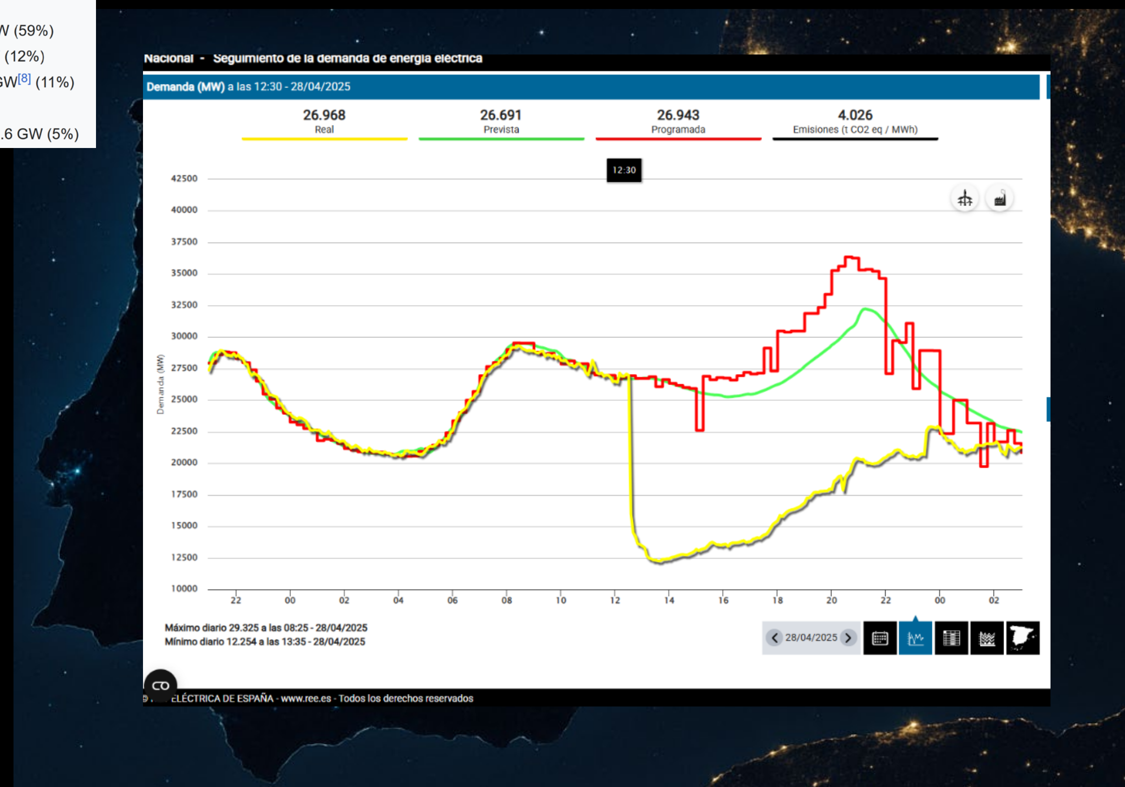
nextmol.com

NEXTMOL™

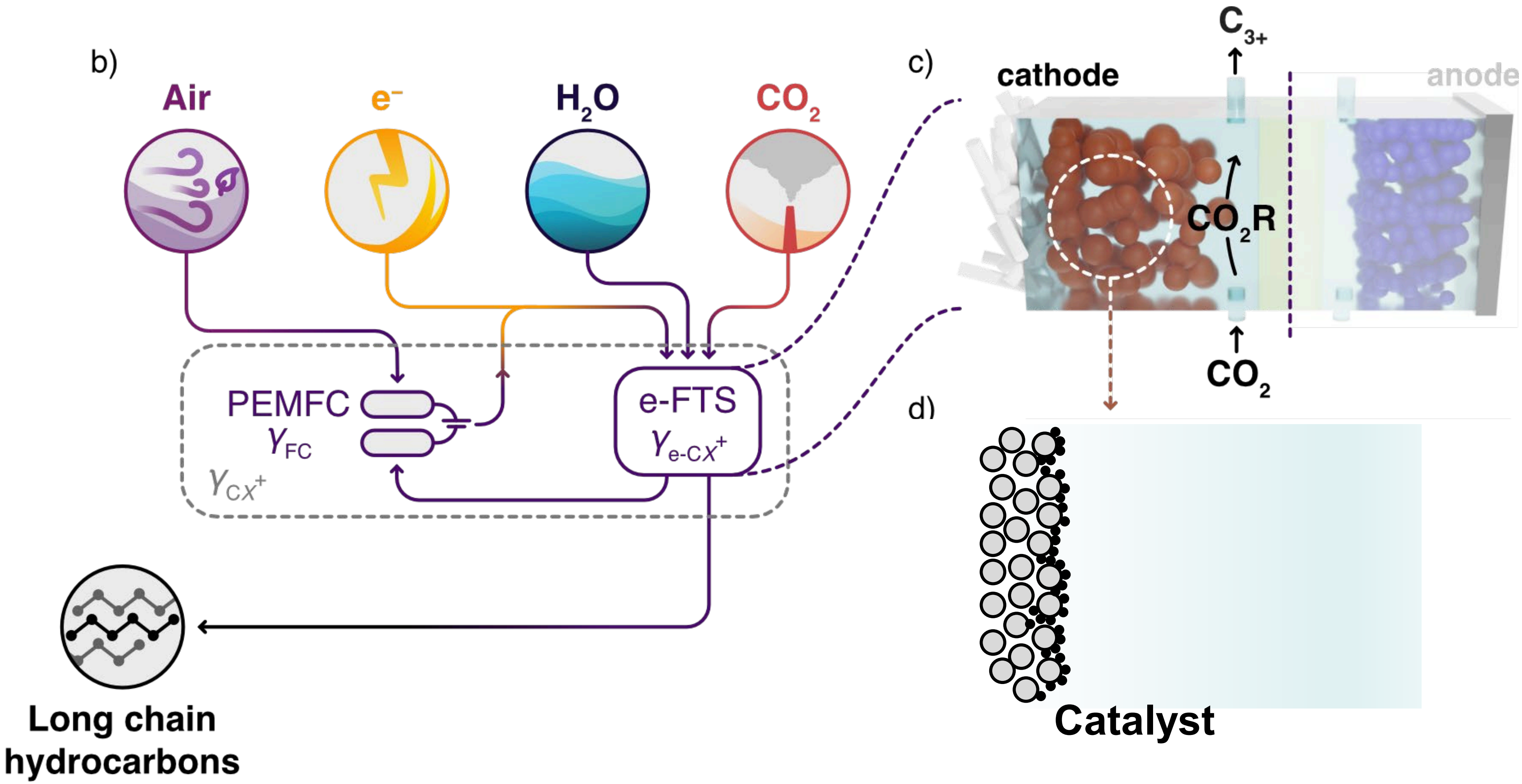
Advanced modeling of Materials for Energy transformations

Núria López


Solar 19.5 GW (59
Grid conditions for 25 GW demand in Spain at 12:30, before the blackout (% of 32 GW total supply)[9][10][11]
Solar 19.5 GW (59%)
Wind 3.6 GW (12%)
Nuclear 3.3 GW[8] (11%)
Other (13%)
Natural gas 1.6 GW (5%)


Nacional - Seguimiento de la demanda de energía eléctrica
Demanda (MW) a las 12:30 - 28/04/2025
26.968
Real
26.691
Prevista
26.943
Programada
4.026
Emisiones (t CO2 eq / MWh)
12:30
Demanda (MW)
42500
40000
37500
35000
32500
30000
27500
25000
22500
20000
17500
15000
12500
10000
22 00 02 04 06 08 10 12 14 16 18 20 22 00 02
Máximo diario 29.325 a las 08:25 - 28/04/2025
Mínimo diario 12.254 a las 13:35 - 28/04/2025
28/04/2025
ELÉCTRICA DE ESPAÑA - www.ree.es - Todos los derechos reservados

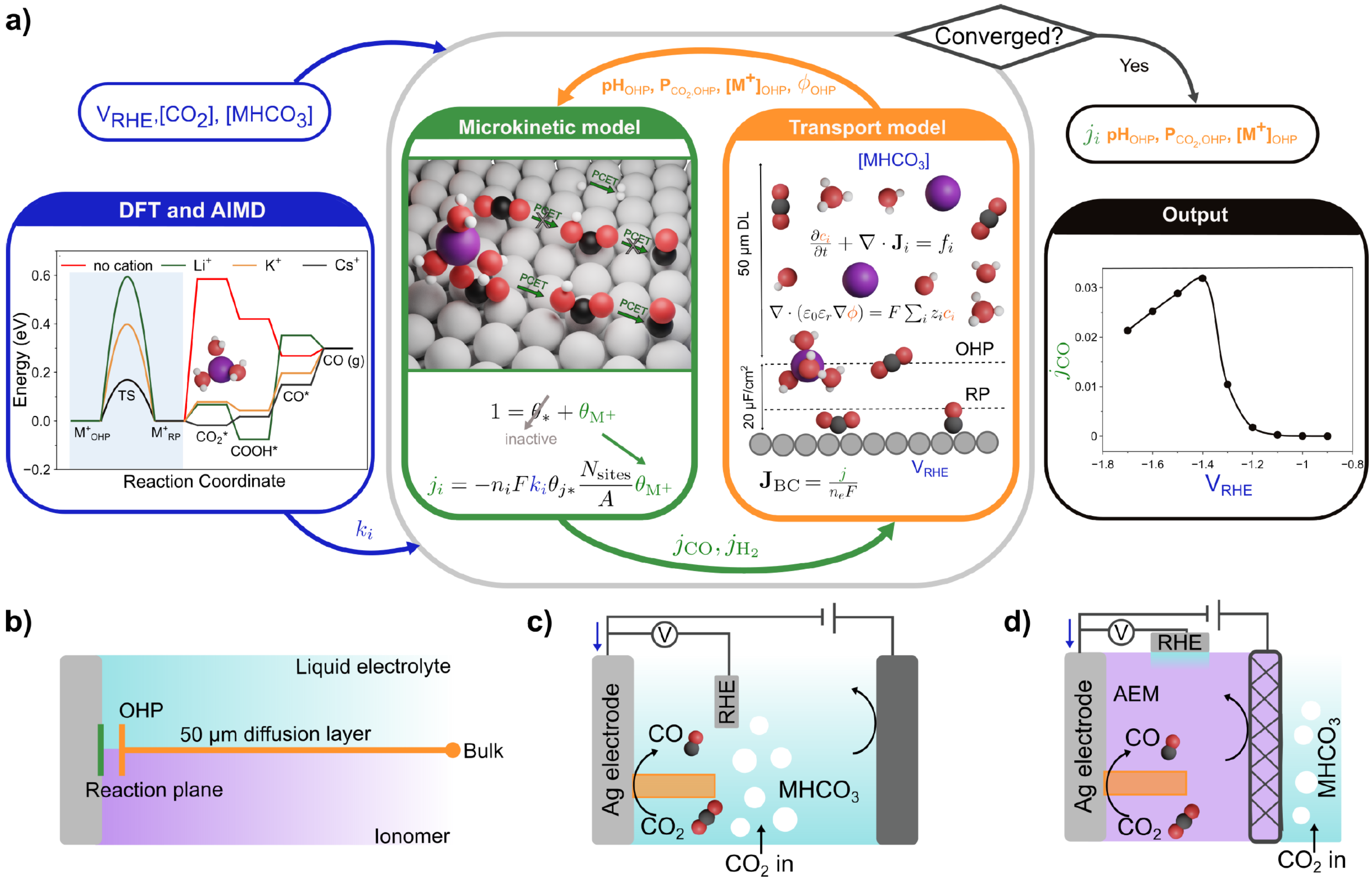
Power-to-X



b)
Air
e⁻
H2O
CO2
PEMFC
γFC
e-FTS
γe-CX+
γCX+
Long chain
hydrocarbons
c)
cathode
anode
C3+
CO2R
CO2
d)
Catalyst

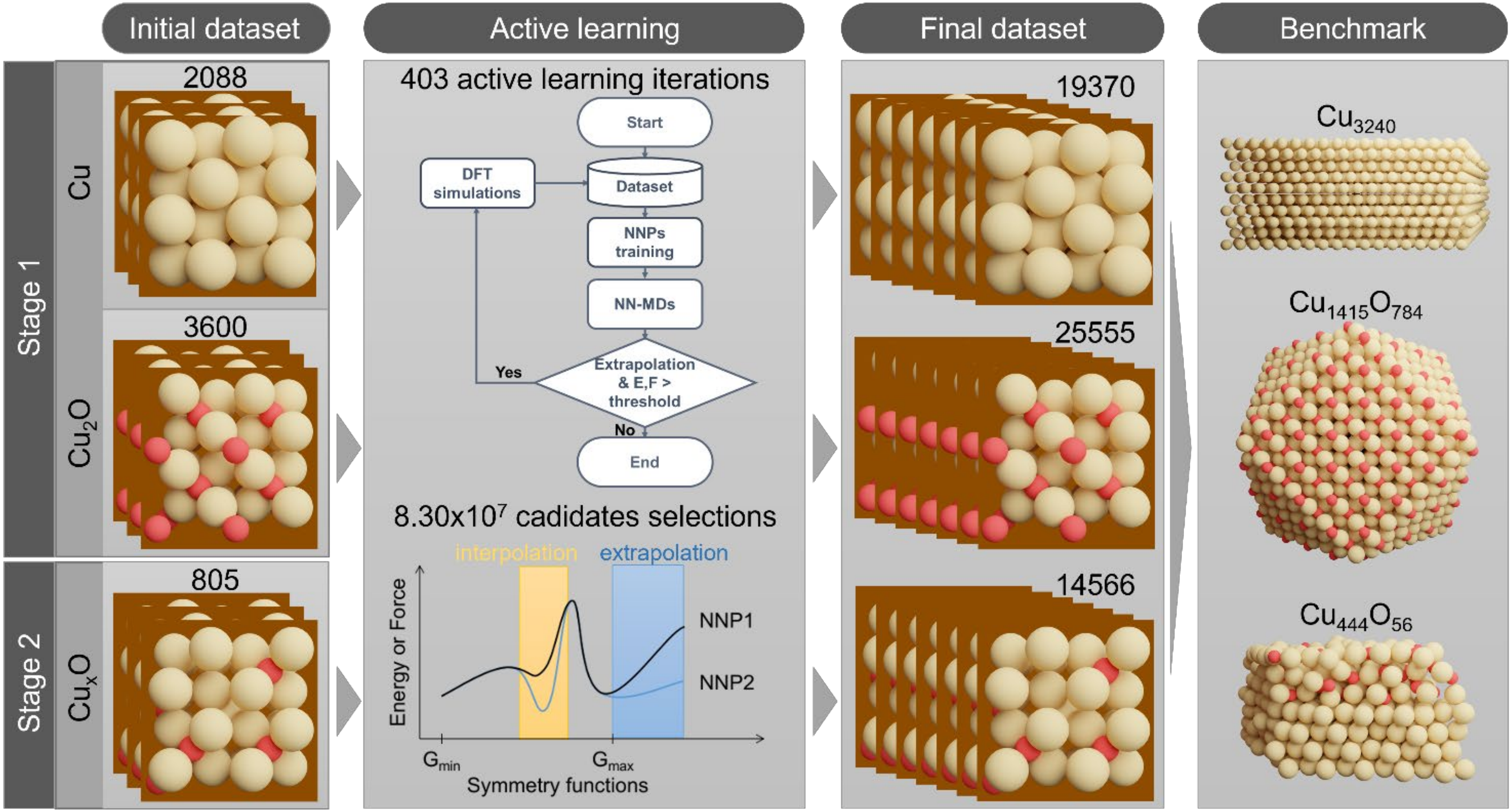
Copper

Cation effects beyond DFT



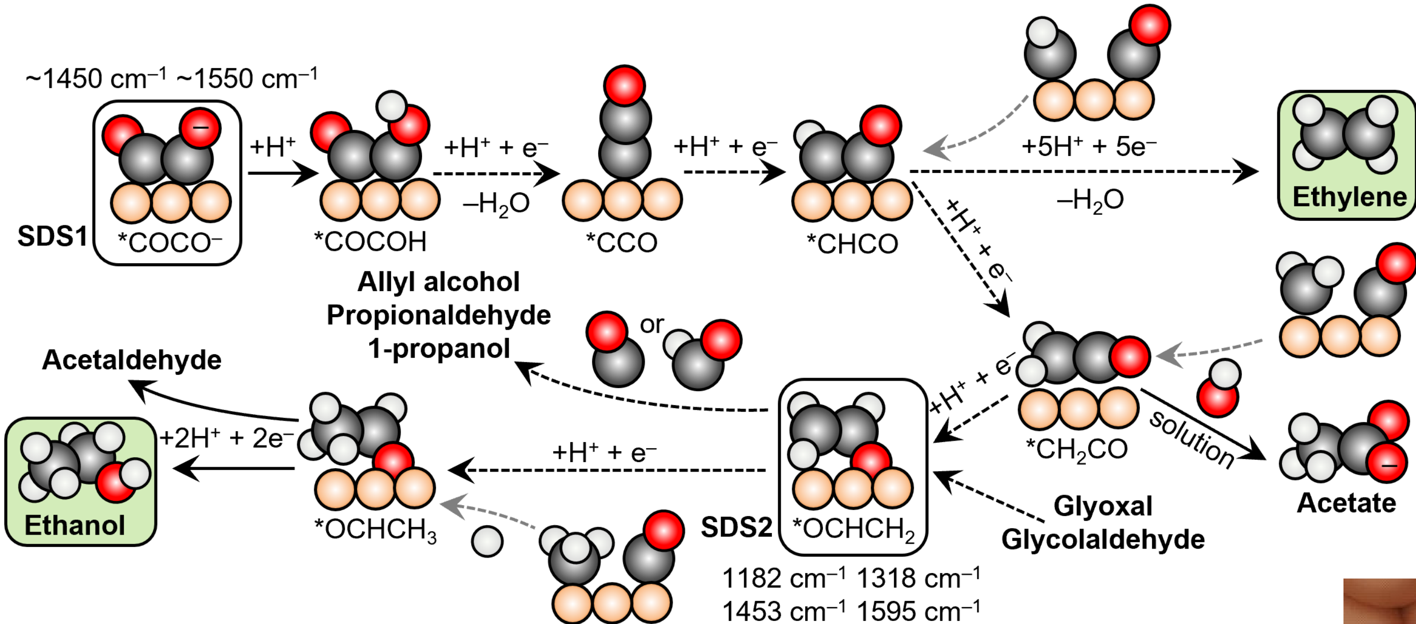
Dynamics models OD-Cu

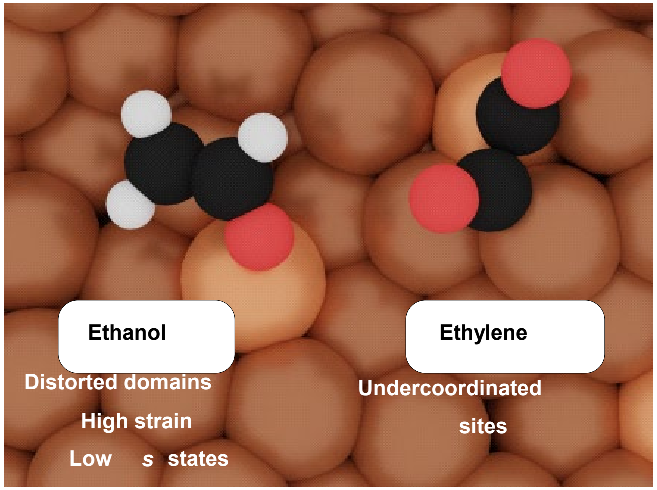
Dynamics with ML potentials



Nature Catal. (2024)

Reactivity OD-Cu

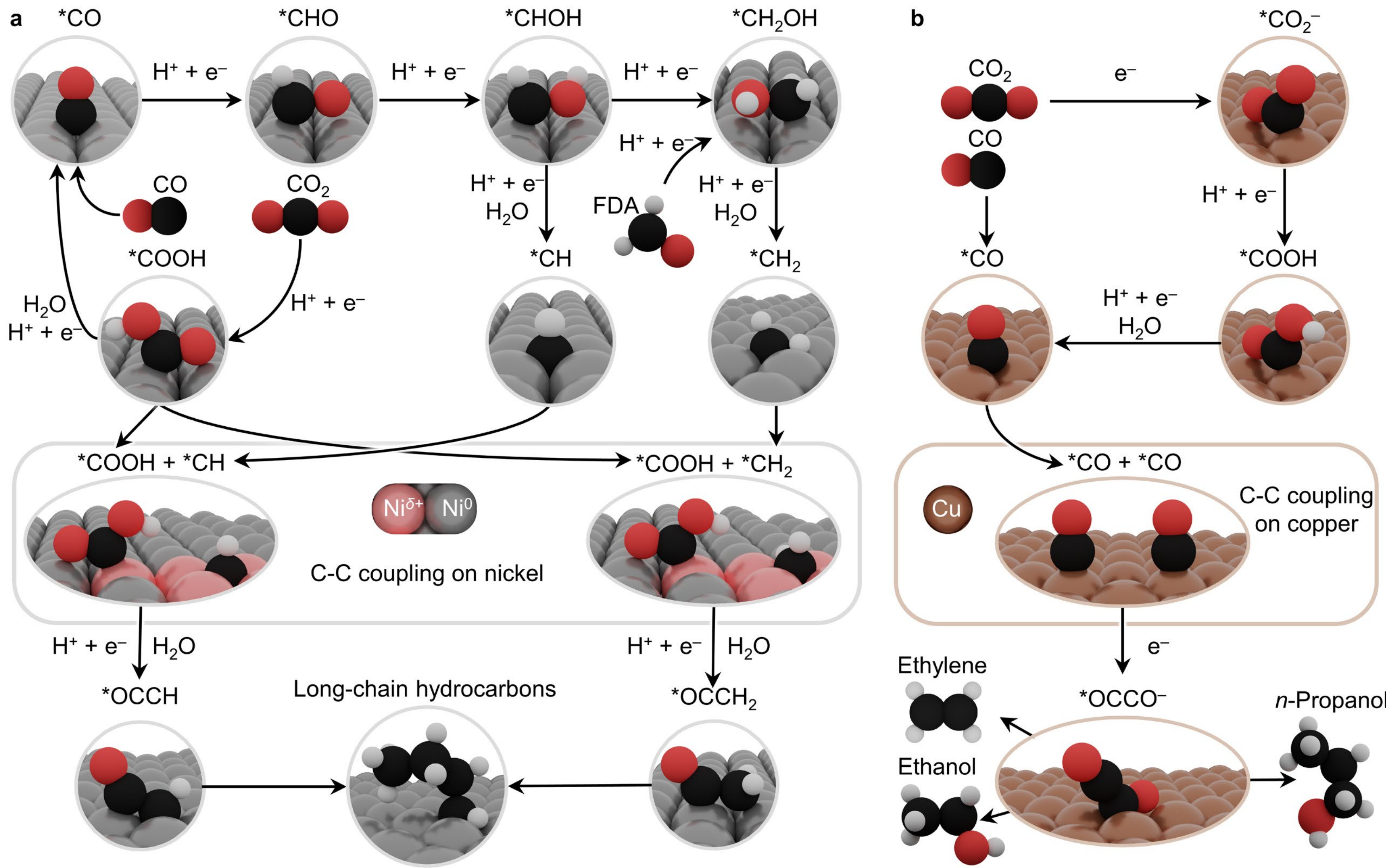




FRITZ-HABER-INSTITUT
MAX-PLANCK-GESELLSCHAFT

Nature Energ. (2024)

Up to C6

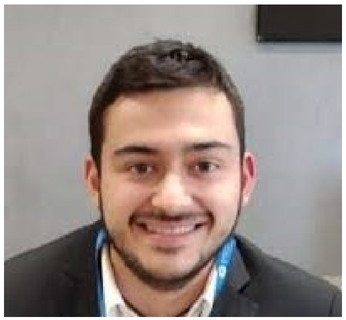


a
*CO
*CHO
*CHOH
*CH$_2$OH
H$^+$ + e$^-$
H$^+$ + e$^-$
H$^+$ + e$^-$
H$^+$ + e$^-$
CO
CO$_2$
H$^+$ + e$^-$
H$_2$O
FDA
H$^+$ + e$^-$
H$_2$O
*COOH
*CH
*CH$_2$
H$_2$O
H$^+$ + e$^-$
H$^+$ + e$^-$
*COOH + *CH
*COOH + *CH$_2$
Ni$^{\delta+}$
Ni^0
C-C coupling on nickel
H$^+$ + e$^-$
H$_2$O
H$^+$ + e$^-$
H$_2$O
*OCCH
Long-chain hydrocarbons
*OCCH$_2$
b
*CO$_2^-$
CO$_2$
e$^-$
CO
H$^+$ + e$^-$
*CO
*COOH
H$^+$ + e$^-$
H$_2$O
*CO + *CO
Cu
C-C coupling on copper
e$^-$
Ethylene
*OCCO$^-$
n-Propanol
Ethanol

Federico Dattila

Zan Lian

B. Roldan-Cuenya
S. Haussener
J. Pérez-Ramírez
B. S. Yeo
M. T. M. Koper
R. Buonsanti
M. Lingenfelder
P. Garcia de Arquer

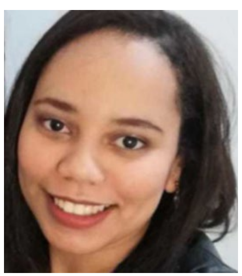


Enric Ibañez-Alè

Rodrigo Garcia-Muelas

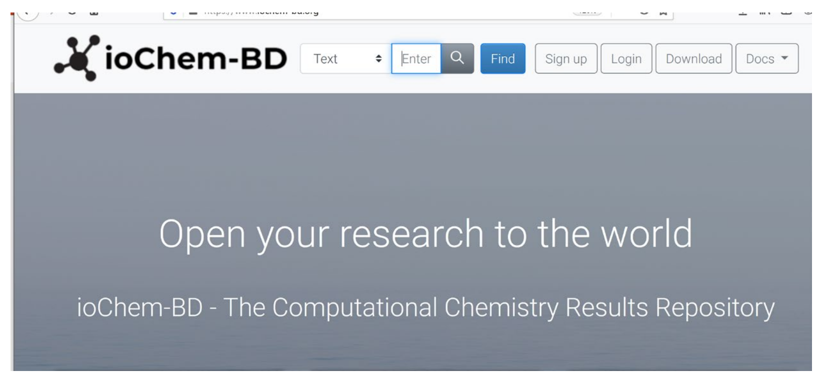
Ranga R. Seemakurthi

EuroHPC
Joint Undertaking



Hind Benzidi

Pol Sanz-Berman



DESIGNING POINT-DEFECTS IN LOW-DIMENSIONAL MATERIALS WITH QUANTUM CHARACTERISTICS

MARCO GOVONI

Department of Physics, Computer Science, Mathematics
University of Modena and Reggio Emilia, Italy

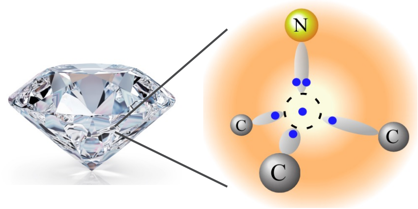
EuroHPC User Days 2025
Copenhagen (Denmark), September 30th, 2025

OVERVIEW

Goals & Opportunities

Optically active spin defects in semiconductors are interesting platforms for the development of solid-state quantum technologies

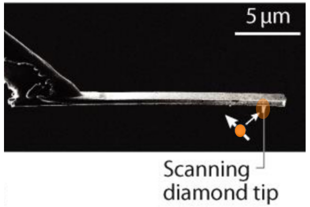
- **two-level system** in material
- operates at **room temperature**



NV$^-$ in diamond

G. Kucsko et al. Nature 500, 54 (2013)
B. Hensen, et al. Nature 526, 682 (2015)
S. J. Whiteley, et al. Nature Physics 15, 490 (2019)
G. Wolfowicz, et al., Nat. Rev. Mater. 6, 906 (2021)

Quantum sensing & metrology



Nanoscale sensors

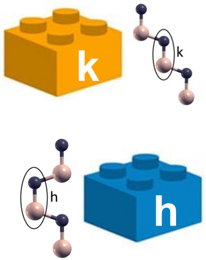
Nat. Commun. 15, 4722 (2024)

Quantum communication

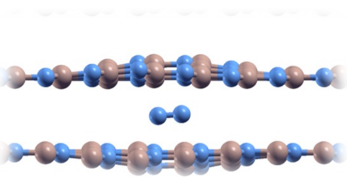
Single-photon emitters for quantum internet

Nature 526, 682 (2015)

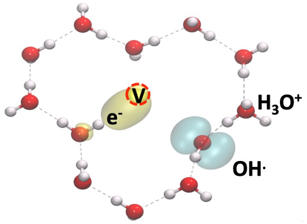
Silicon Carbide



Nitrides



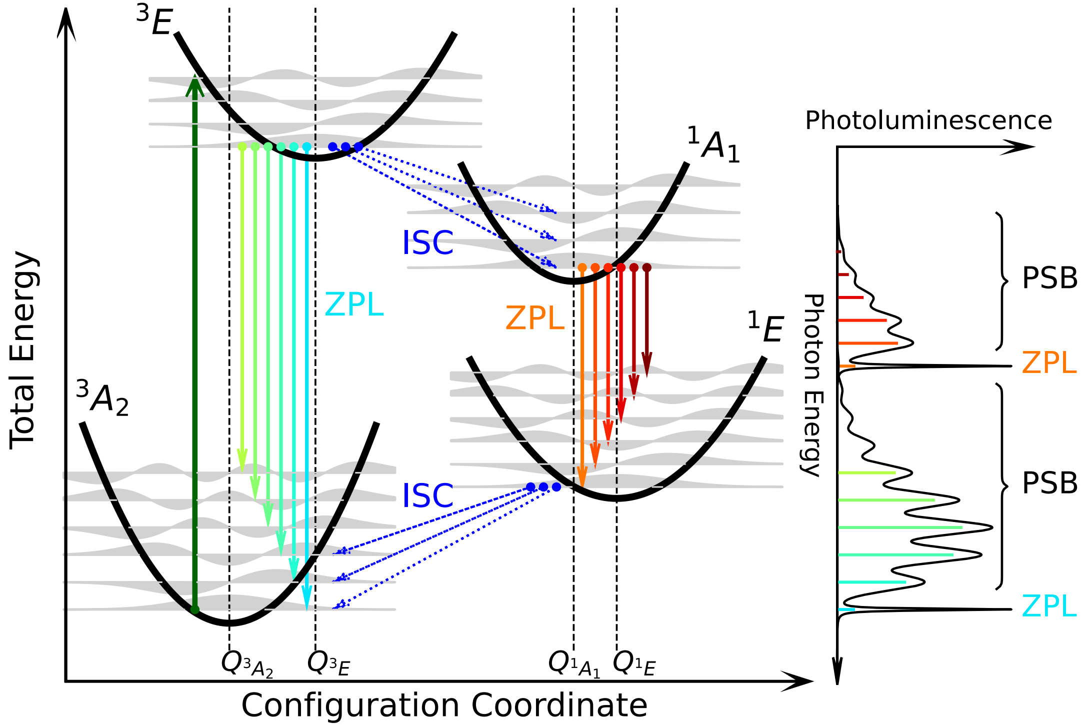
Ice



EuroHPC Joint Undertaking

UNIMORE UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

COMPUTATIONAL CHALLENGES



(3N atom positions)

To simulate photo-luminescence we compute:

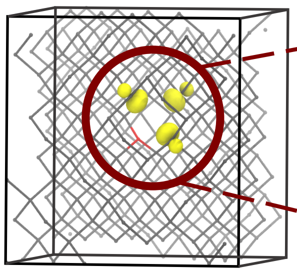
1) Excitation energies
→ **multiconfigurational excited states**

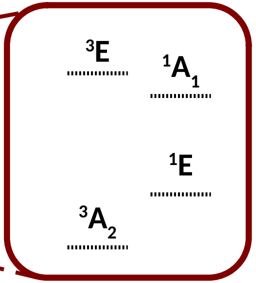
2) Optimization of atom positions in excited potential energy surfaces (PES)
→ **Forces for excited states**

EXCITATION ENERGIES

Spin defect

Many-body spectrum





Density Functional Theory (DFT)

$$|X\rangle = \underbrace{\overbrace{\sum_{vc} A^X_{vc}|vc\rangle}^{\textbf{TD-DFT / BSE}} + \sum_{vv'cc'} A^X_{vv'cc'}|vv'cc'\rangle + \cdots}_{\textbf{Full Configuration Interaction (FCI)}}$$

single double n-ple

Full Configuration Interaction (FCI)

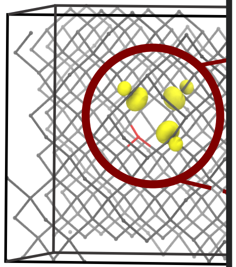
https://www.quantum-espresso.org/

https://west-code.org/

Walker, Saitta, Gebauer, Baroni, Phys. Rev. Lett. 96, 113001 (2006)
Rocca et al., J. Chem. Phys. 128 154105 (2008), J. Chem. Phys. 113 164109 (2010), Phys. Rev. B 85 045116 (2012)
Nguyen, Ma, Govoni, Gygi, Galli, Phys. Rev. Lett. 122, 237402 (2019)
Bockstedte, Schütz, Garratt, Ivady, Gali, npj Comput Mater *3*, 31 (2018)
Ma, Govoni, Galli, npj Comput Mater 6, 85 (2020)
Ma, Sheng, Govoni & Galli, JCTC 17, 2116 (2021)
Sheng, Vorwerk, Govoni, Galli, JCTC 18, 3512 (2022)
Jin, Yu, Govoni, Xu, Galli, JCTC 19, 8689 (2023)
Yu, Jin, Galli, Govoni, JCTC 20, 10899 (2024)
Chen, Yu, Jin, Govoni, Galli, JCTC 21 (2025)

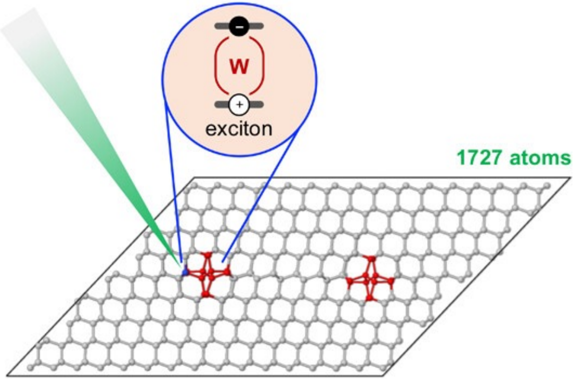
EXCITA

Spin defe



Density Functiona Theory (DFT)

Large scale MBPT calculation



NV$^-$@dislocation in diamond

G_0W_0-BSE without empty states

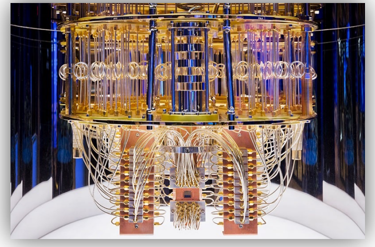
5 min on 64 GPU nodes

Yu, Jin, Galli, Govoni, JCTC 20, 10899 (2024)

Hybrid classical/quantum computing

CPU/GPU

QPU



Huang, Govoni, Galli, PRX Quantum 3, 010339 (2022)
Huang, Sheng, Govoni, Galli, JCTC 19, 1487 (2023)

$'cc'\rangle + \cdots$

n-ple

on (FCI)

t. 96, 113001 (2006)
J. Chem. Phys. 113
B 85 045116 (2012)
122, 237402 (2019)
t Mater 3, 31 (2018)
t Mater 6, 85 (2020)
TC 17, 2116 (2021)
TC 18, 3512 (2022)
JCTC 19, 8689 (2023)
, JCTC 20, 10899 (2024)
Chen, Yu, Jin, Govoni, Galli, JCTC 21 (2025)

t_z

EXCITED STATES FORCES

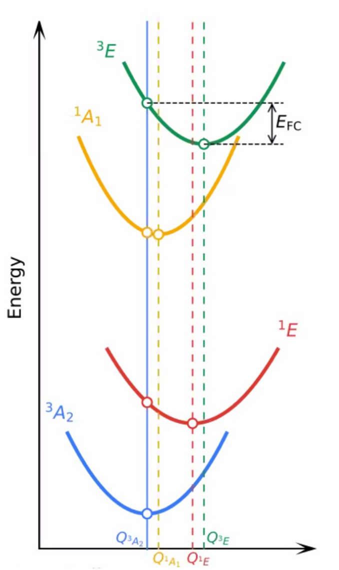
- TDDFT w/ Tamm-Dancoff approx. $(D + K^{1e} - K^{1d})X_I = \boldsymbol{\omega_I} X_I$

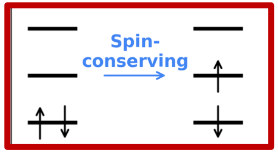
- **<u>TDDFT analytical forces</u>** on nuclei $\boldsymbol{\nabla}_R \boldsymbol{\omega_I} = \int dr\, \boldsymbol{\nabla}_R V_{ext}(\boldsymbol{r})\, [\Delta\rho^a(\boldsymbol{r}) + \Delta\rho^z(\boldsymbol{r})]$

Density variation obtained from the virtual manyfold

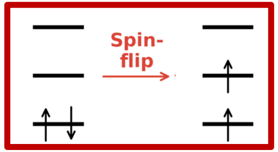
Handy–Schaefer Z-vector correction

We use the generalized Lagrangian framework to obtain TDDFT/BSE forces





Triplet excited states



Singlet excited states

Hutter, J. Chem. Phys 118, 3928 (2003)
Jin, Yu, Govoni, Xu, Galli, JCTC 19, 8689 (2023)

EXCITE

- TDDFT w
- **TDDFT a**

$\rho^{z}(\boldsymbol{r})]$

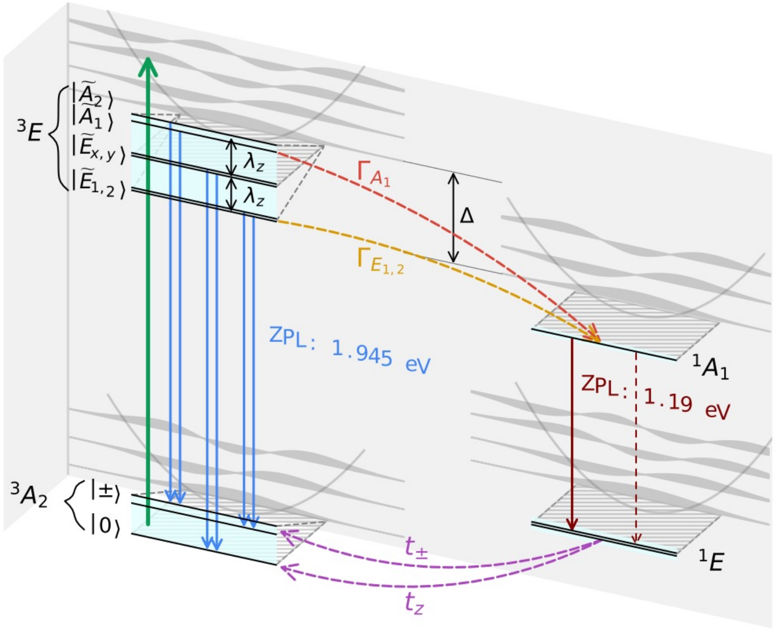
ndy–Schaefer
ector correction

We use
Lagrangi
obtain TD

Triplet
xcited
states

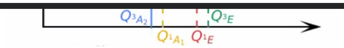
Singlet
xcited
states

Intersystem crossing rates



Jin et al., PRL 135, 036401 (2025)

Hutter, J. Chem. Phys
Jin, Yu, Govoni, Xu, Galli, JCTC 19, 8689 (2023)



UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

CODES

- Density functional theory (DFT) calculations with periodic boundary conditions

QUANTUMESPRESSO

- KS-DFT with plane-waves
- ONCV pseudopotentials
- Density Functional Perturbation Theory

MAX
DRIVING THE EXASCALE TRANSITION

https://www.quantum-espresso.org/

Giannozzi, et al., J. Phys. Condens. Matter 21, 395502 (2009)
Giannozzi, et al., J. Phys.:Condens.Matter 29, 465901 (2017)
Giannozzi, et al., J. Chem. Phys. 152, 154105 (2020)

- Excited states for large systems

⟨ WEST! ⟩

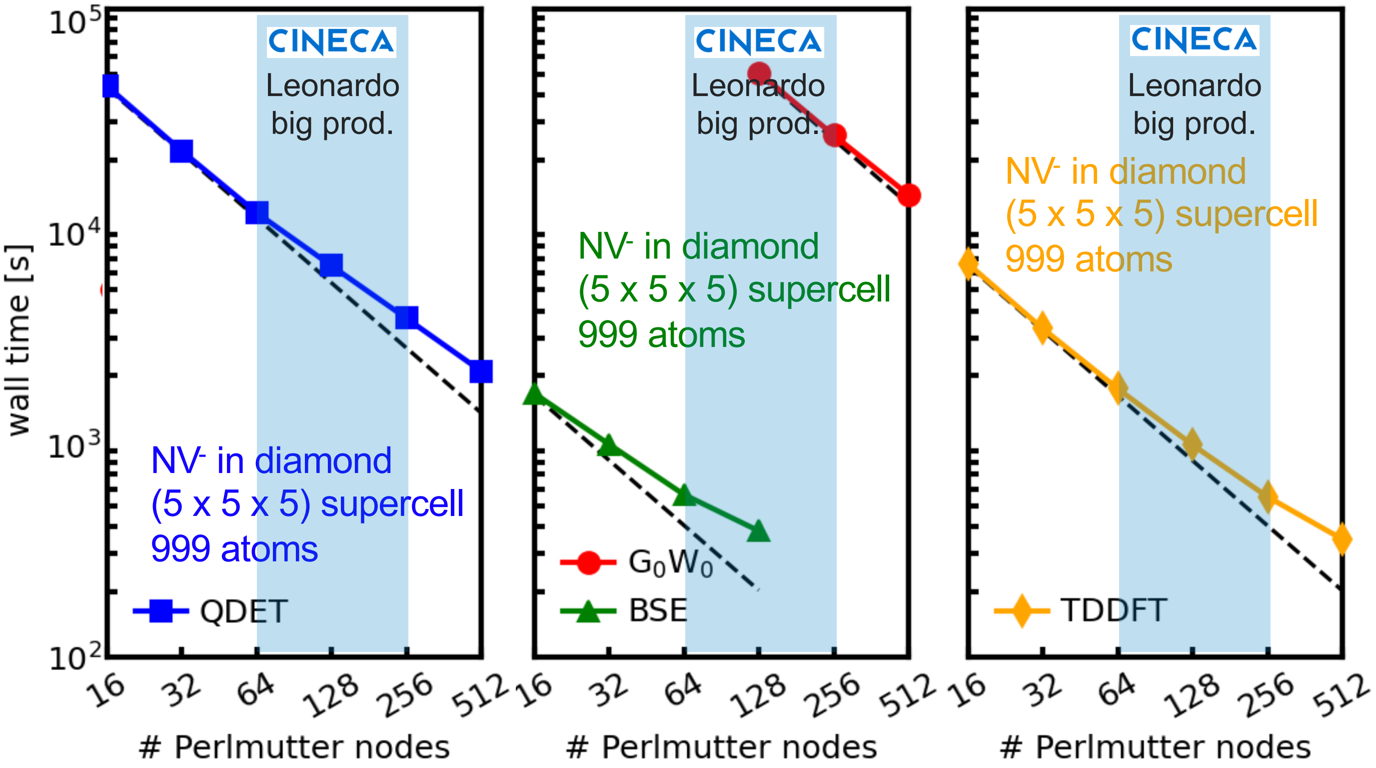
- Many-body perturb. Theory: GW-BSE
- Time-dependent DFT (w/ spin-flip)
- Quantum Embedding (FCI-in-DFT)

MICCoM

https://west-code.org/

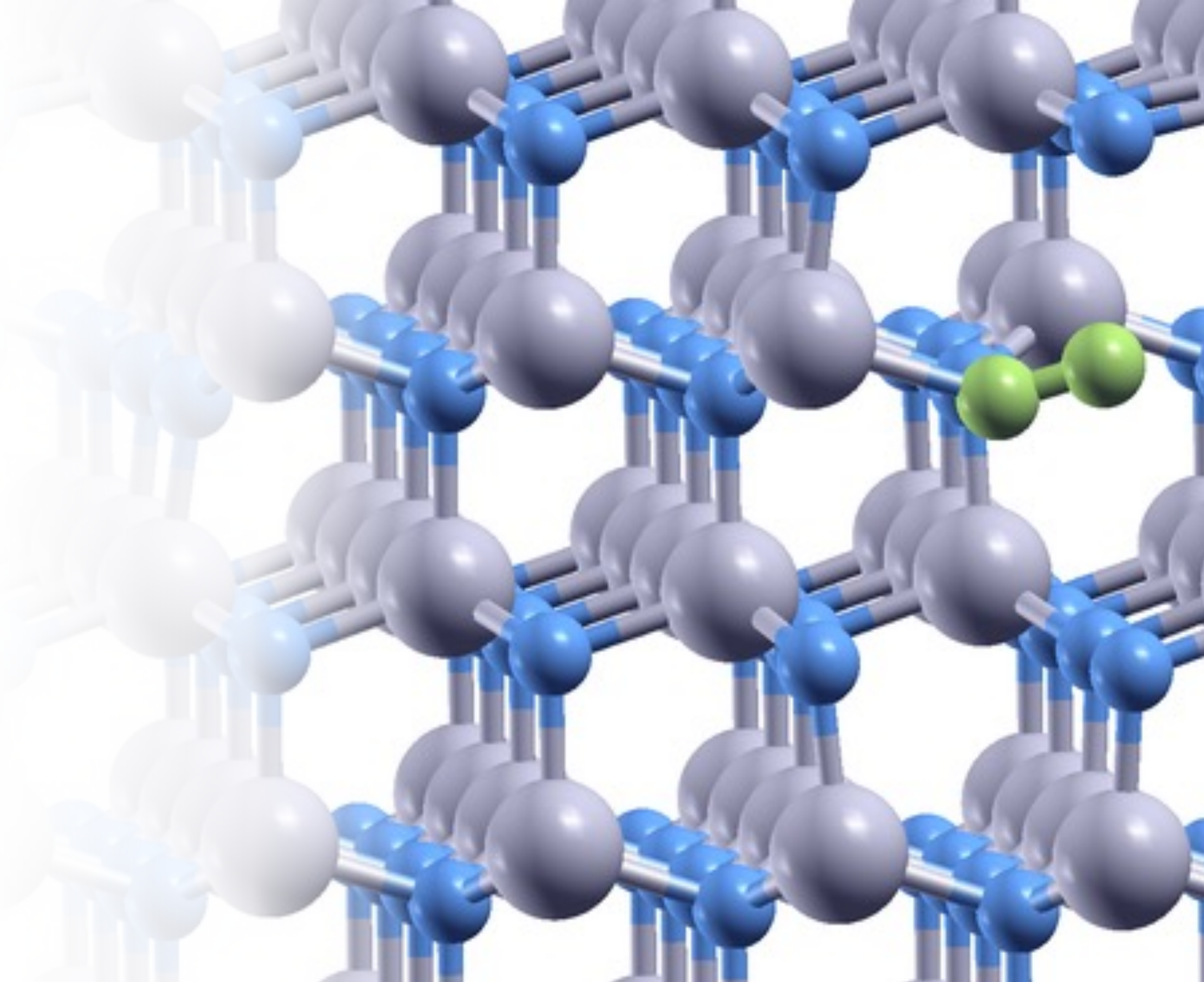
Govoni, Galli, JCTC 11, 2680 (2015)
Yu, Jin, Galli, Govoni, JCTC 20, 10899 (2024)
Chen, Yu, Jin, Govoni, Galli, JCTC 21 (2025)

PERFORMANCE ON GPUS



NERSC/Perlmutter node: 1 AMD EPYC Milan (64 cores) + 4 NVIDIA A100 GPUs
CINECA/Leonardo: 1 Intel Ice Lake (32 cores) + 4 NVIDIA A100 GPUs

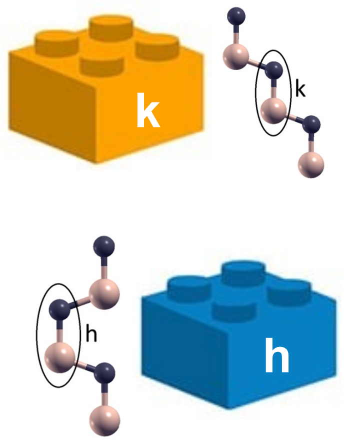
Results from the EuroHPC allocation

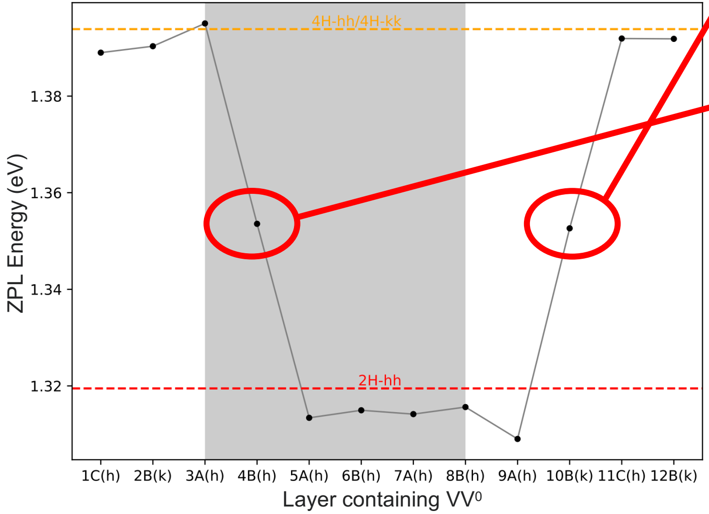


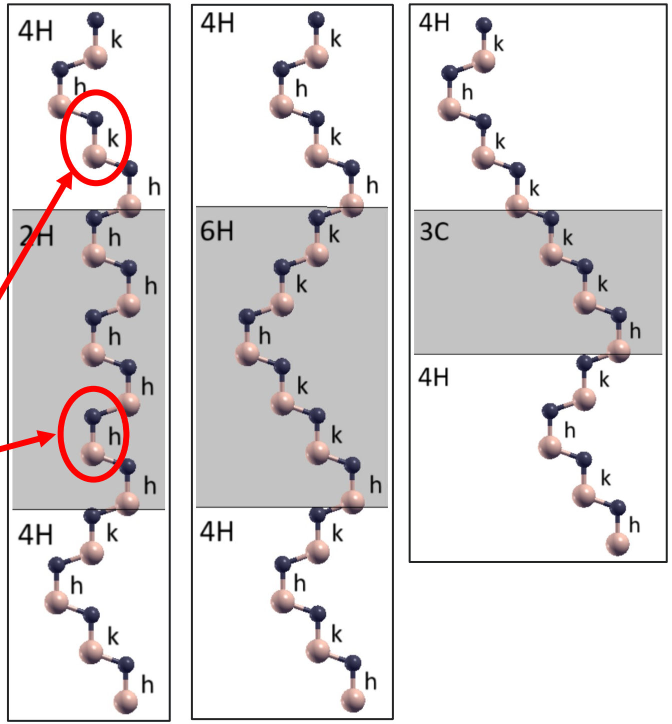
EuroHPC
Joint Undertaking

EFFECT OF LOCAL ENVIRONMENT IN SILICON CARBIDE

- Silicon carbide exhibits polytypism, can it be leveraged to tailor defect properties?
- Developed a model based on local environment to describe the influence of heteropolytypism on VV^0



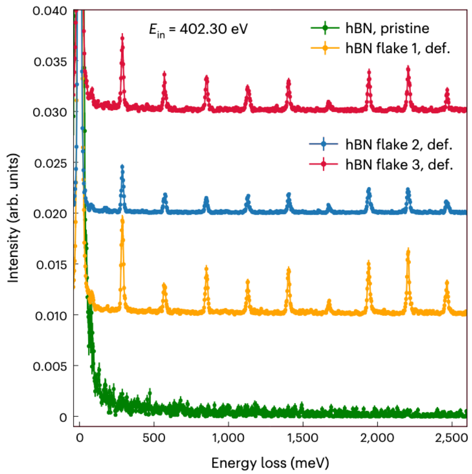




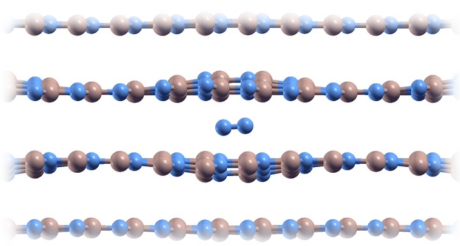
Experimental collab.: Heremans (Argonne National Lab), manuscript in preparation

SINGLE PHOTON EMISSION IN NITRIDE MATERIALS

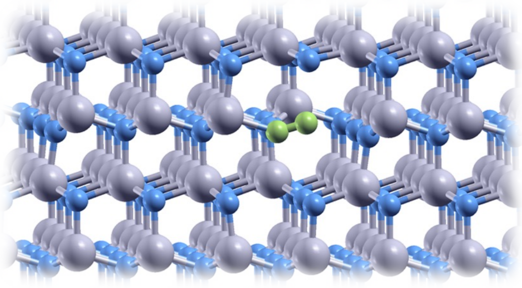
- Several quantum technologies rely on quantum emitters capable of producing single photons
- Vibrational modes of molecular-like defects influence single photon emission in hBN, GaN, AlN



Pelliciari et al., Nature Materials 23, 1230 (2024)



Explained microscopic origin of excitation patterns in hBN, reconciling RIXS and PL spectra



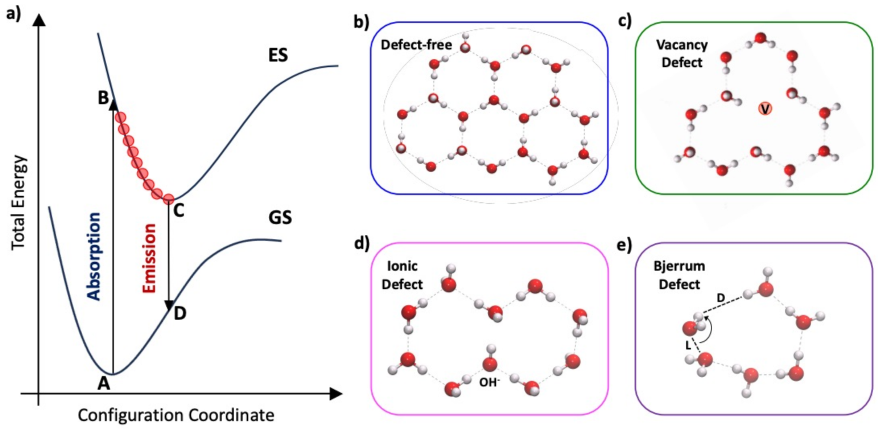
Proposed a similar mechanism to explain excitation patterns in wurtzite lattices (AlN, GaN)

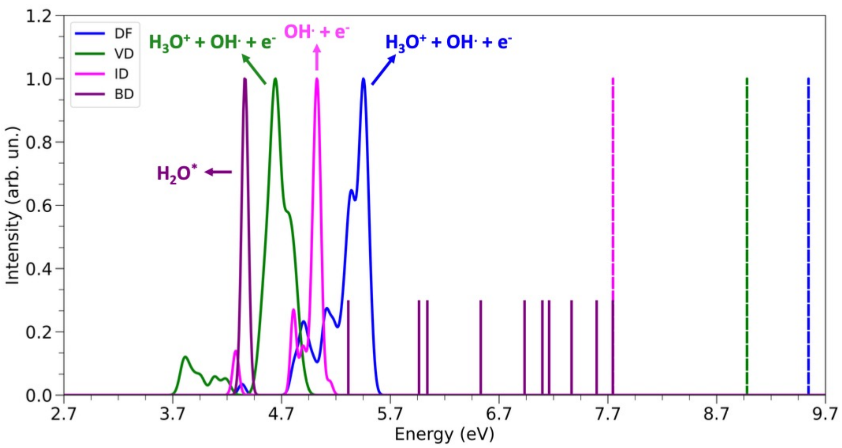
Experimental collaborators: Pelliciari (Brookhaven National Lab), Grosso (City U. New York), manuscript in preparation

UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

PHOTOCHEMISTRY OF ICE

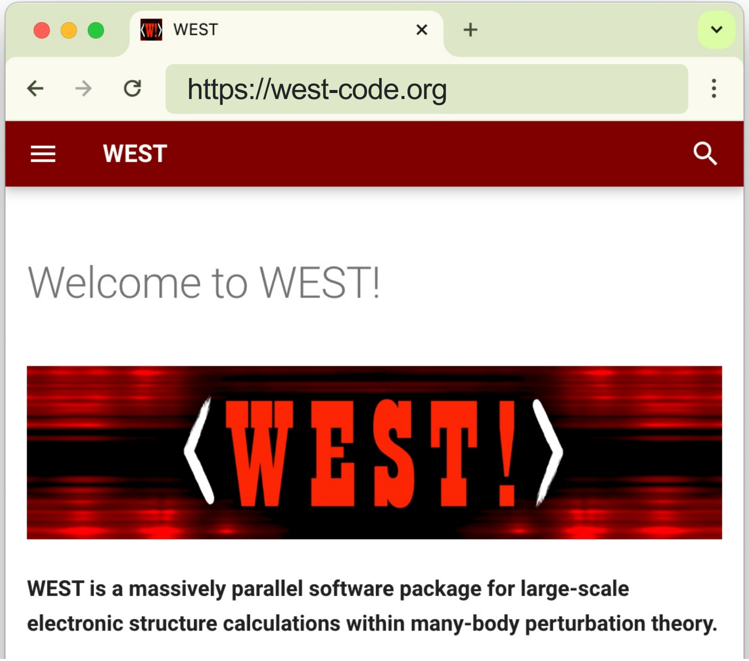
- UV light drives fundamental atmospheric & planetary processes in ice
- Vacancies, ionic species, and orientational defects influence UV absorption and emission, controlling the formation of photoproducts in ice





M. Monti, Y. Jin, G.D. Miron, A. Kundu, M. Govoni, G. Galli, A. Hassanali, under review, arXiv:2506.16568

GET STARTED WITH THE WEST CODE



- **New release! GPU-enabled TDDFT/BSE/Embedding + TDDFT forces**
  v6.2.1 (July 2025)
- **Website:** https://west-code.org
- **Git repository:** https://github.com/west-code-development
- **Tutorials:** https://west-code.org/doc/West/latest/tutorial

ACKNOWLEDGMENTS

- Collaborators (theory/code):
  - Giulia Galli (U. Chicago)
  - Yu Jin (U. Chicago)
  - Victor Yu (Argonne National Lab)
  - Ali Hassanali (ICTP)
  - Marta Monti (ICTP)
- Collaborators (experimental):
  - Joe Heremans (Argonne National Lab)
  - Jonathan Pelliciari (Brookhaven National Lab)
  - Gabriele Grosso (CUNY)
- Computational facilities used:
  - CINECA/Leonardo, NERSC/Perlmutter

Funding

U.S. DEPARTMENT OF ENERGY

MICCoM
http://miccom-center.org

FONDO ITALIANO PER LA SCIENZA
MUR

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani

THANK YOU

EuroHPC
Joint Undertaking

UNIMORE
UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA