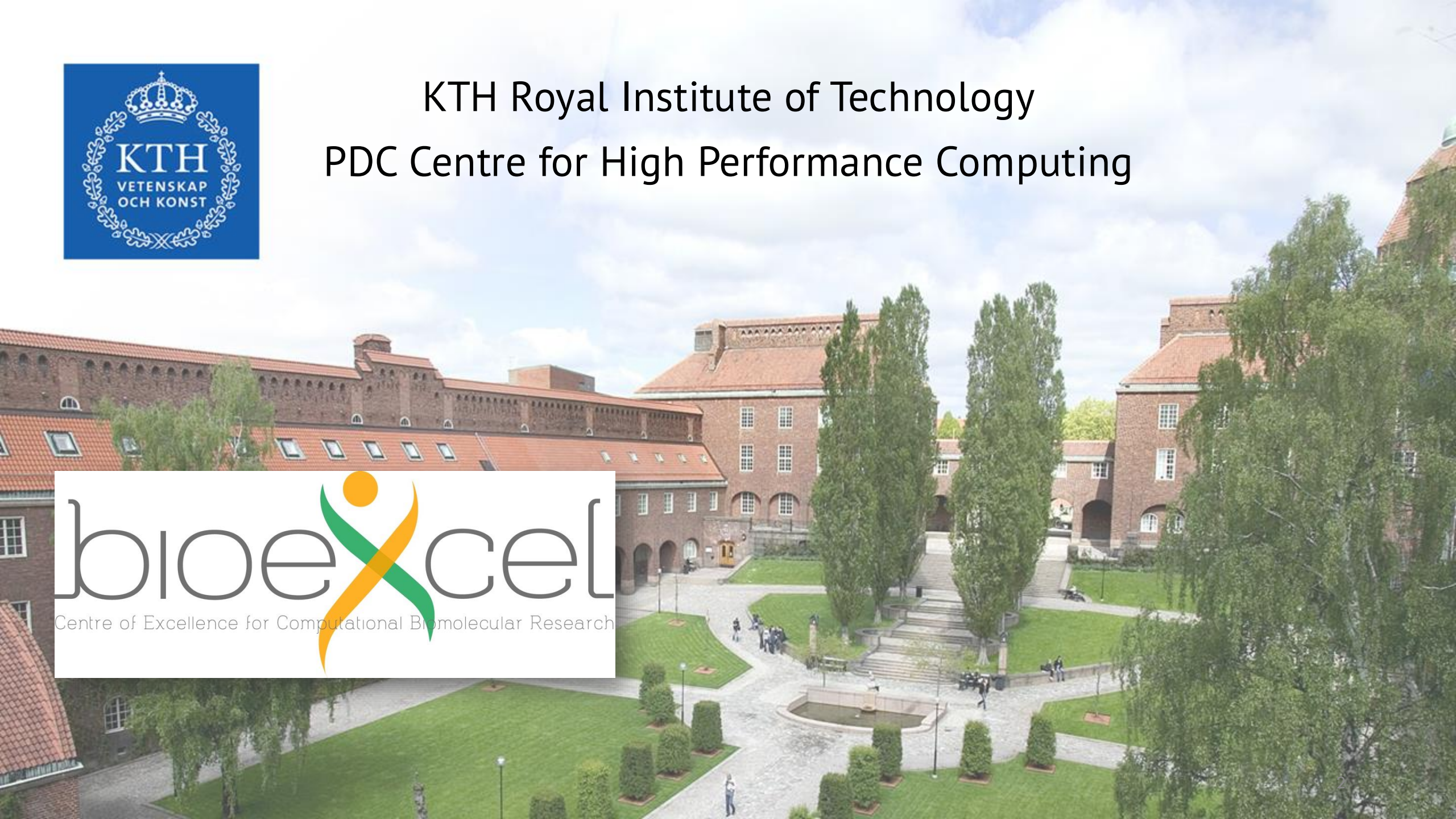
Scalable and Performant Applications for Biomolecular Research: Efficient use of HPC and Cloud

Rossen Apostolov

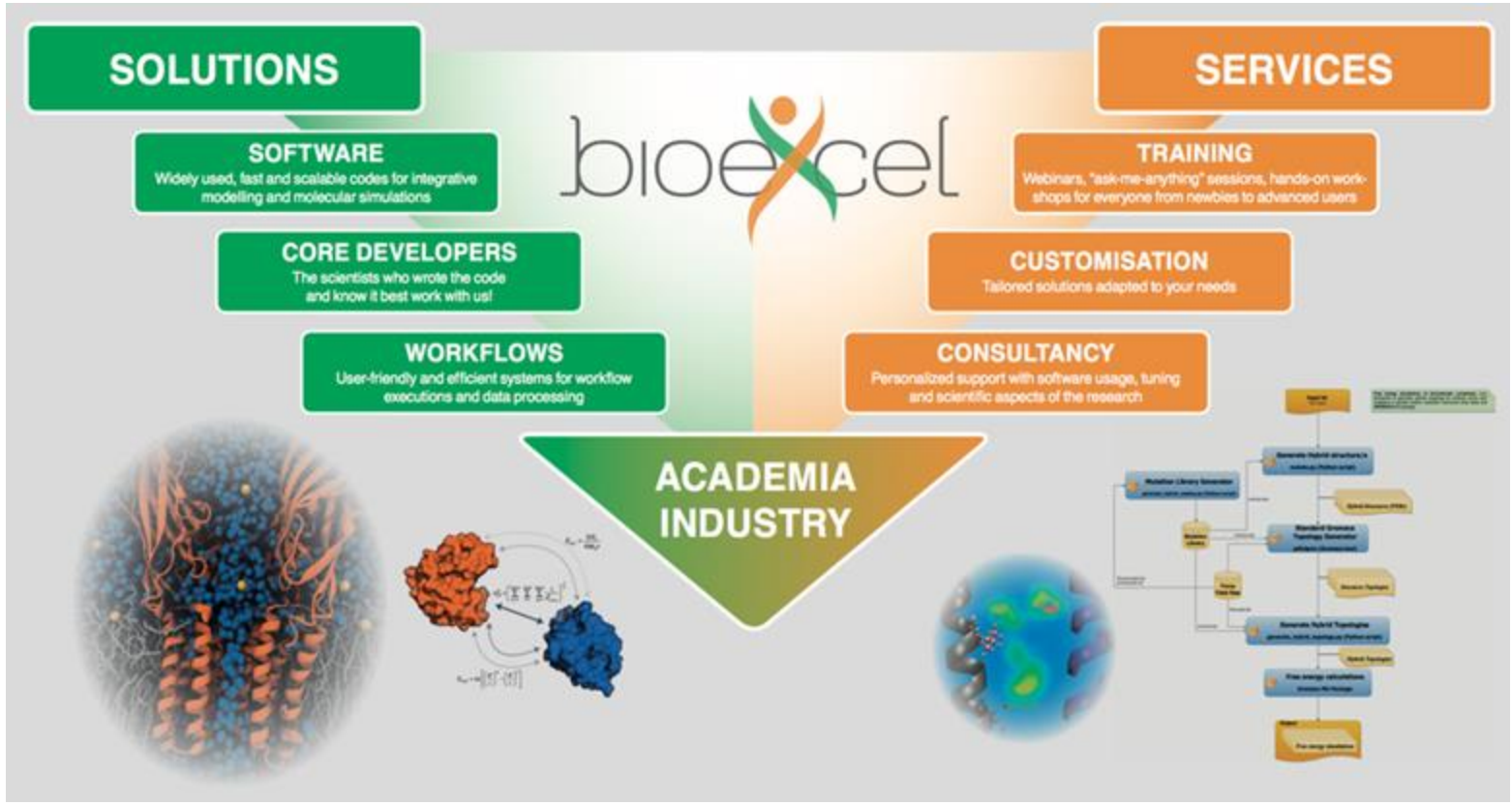
rossen@kth.se

KTH Royal Institute of Technology

PDC Centre for High Performance Computing



BioExcel Centre of Excellence (est. 2015)



Horizon Europe

Funded by the Horizon 2020 Framework Programme of the European Union

EuroHPC Joint Undertaking

- *Provide Life Science researchers with high-quality, user-friendly software*
- *Increase their expertise and skills*
- *Strengthen the community*

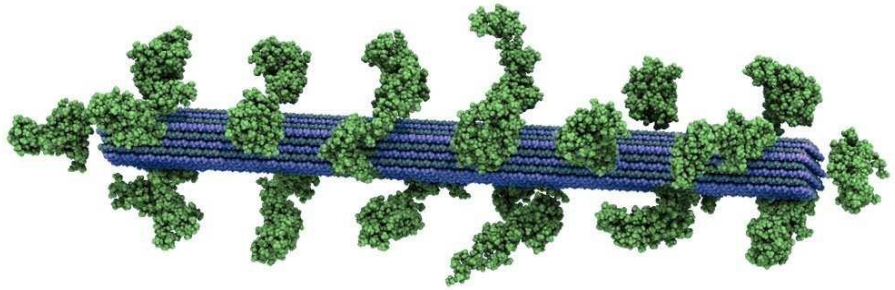
Molecular Dynamics simulations:

GROMACS

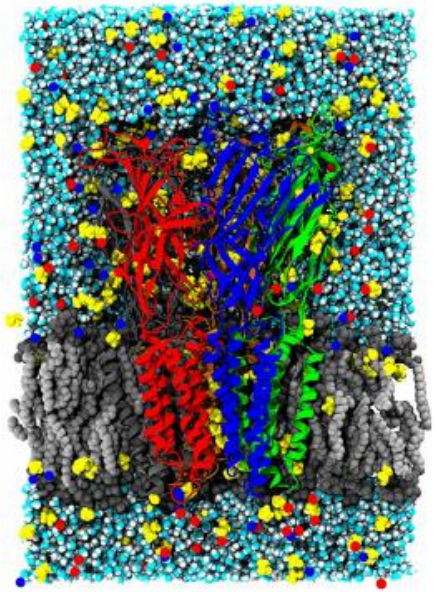
Molecular Dynamics simulations: use-cases

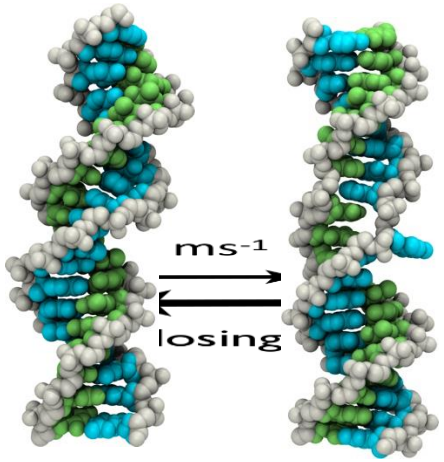
Biomolecular MD

Cellulose + lignocellulose + water: 10^7 particles



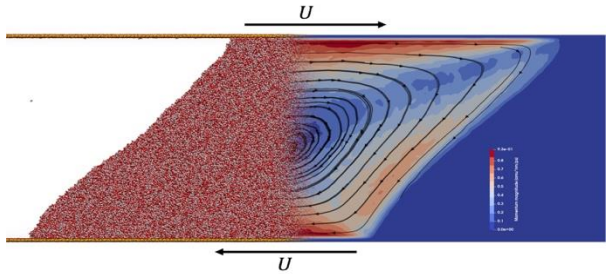
Membrane protein: 10^5 particles



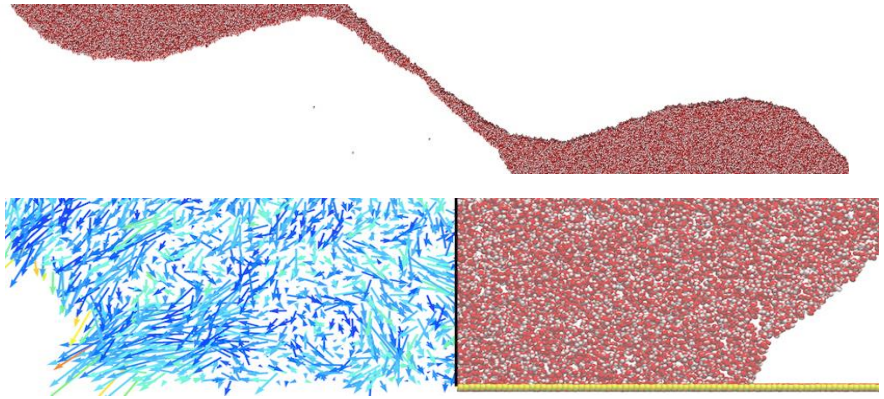


DNA base-pair opening:
10^4 particles

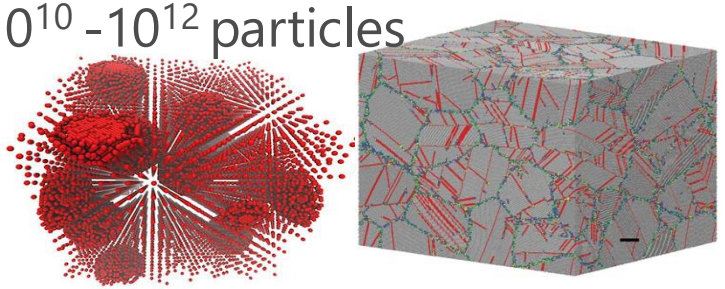
Materials MD



Contact line friction &
wetting dynamics
10^7 -10^9 particles



Nucleation in nano-crystals:
10^{10} -10^{12} particles



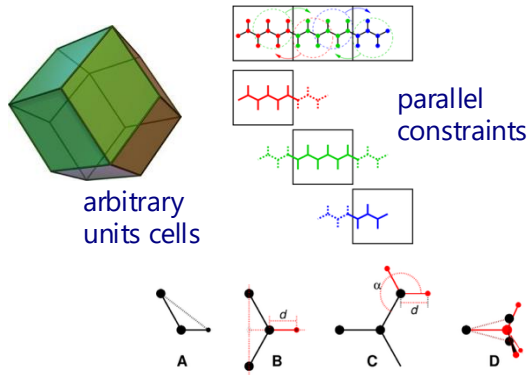
FAST. FLEXIBLE. FREE.

GROMACS

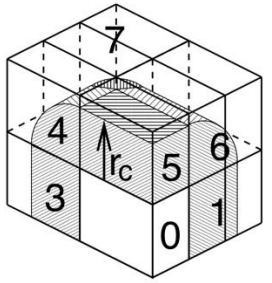
- **Classical MD** code
  - supports all major force-fields
  - broad algorithm support
- **Development**:

  Stockholm Sweden & partners worldwide
- **Large user base**:
  - One of the top HPC codes worldwide

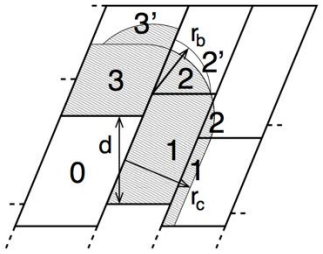
    deployed on most clusters
  - 10k's academic & industry users



virtual interaction sites



Eighth shell
domain
decomposition



Triclinic unit cell with
load balancing and
staggered cell boundaries

- **Open source**: LGPLv2
- **Open development:**
  - code review & bug-tracker: https://gitlab.com/gromacs

FAST. FLEXIBLE. FREE.

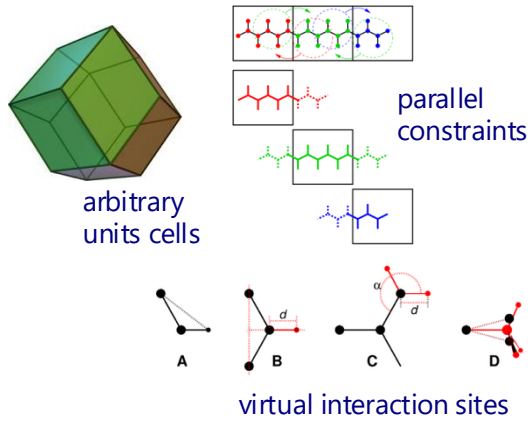
GROMACS

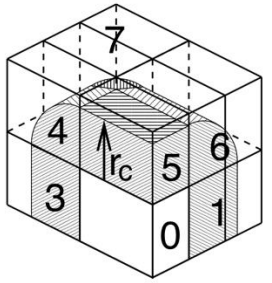
- Focus on **high performance:**

  efficient algorithms & highly-tuned parallel code

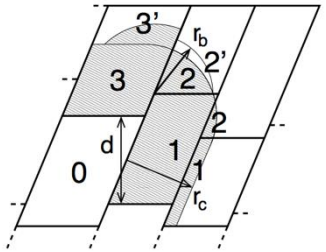
- **Bottom-up performance-oriented design:**
  - absolute performance over "just scaling"
- Focus on **portability**
  - Linux distro integration and CI
  - regular testing on all HPC arch
  - SIMD portability library, GPU abstraction layer
  - open standards-based languages/APIs

- **Modern development workflow**
  - mandatory open code review for >10 years
  - tiered CI testing / verification



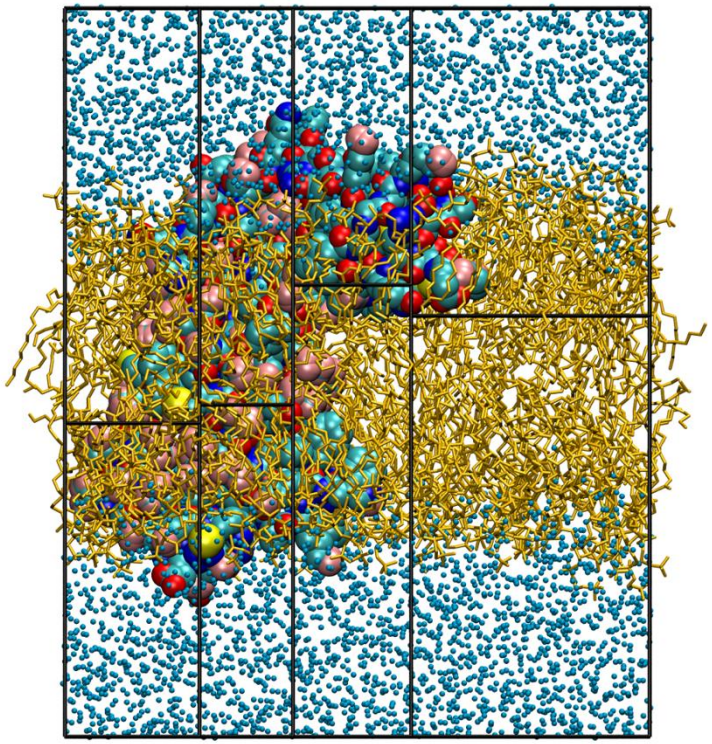


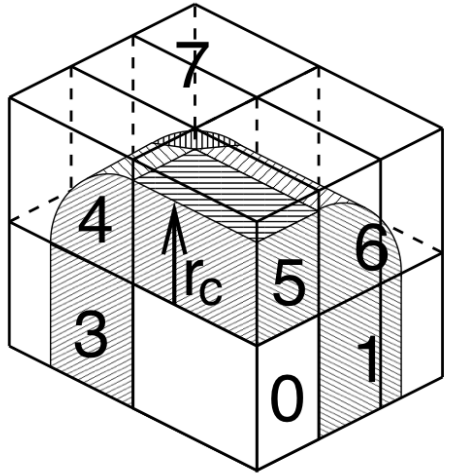
Eighth shell domain decomposition



Triclinic unit cell with load balancing and staggered cell boundaries

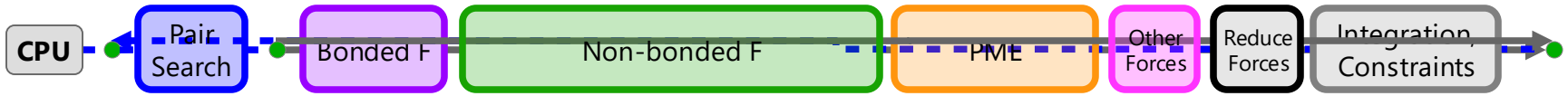
- **Multi-level hierarchical parallelization:** target hardware levels individually
  - Intra-node:
    - OpenMP multi-threading
      - static loop schedule, cache optimized work decomposition, sparse reduction
    - SIMD C++ library abstraction: 14 SIMD flavors supported
    - GPU abstraction layer: CUDA, OpenCL, SYCL backends
    - thread-MPI: pthreads-based MPI for ease of use
  - Inter-node:
    - MPI:
      - SPMD / MPMD
      - Direct GPU communication (GPU-aware on all platforms)
      - hierarchical ensemble parallelization
    - NVSHMEM – codesign work in progress
  - multi-level load balancing:
    - dynamic load balancing
    - task balancing



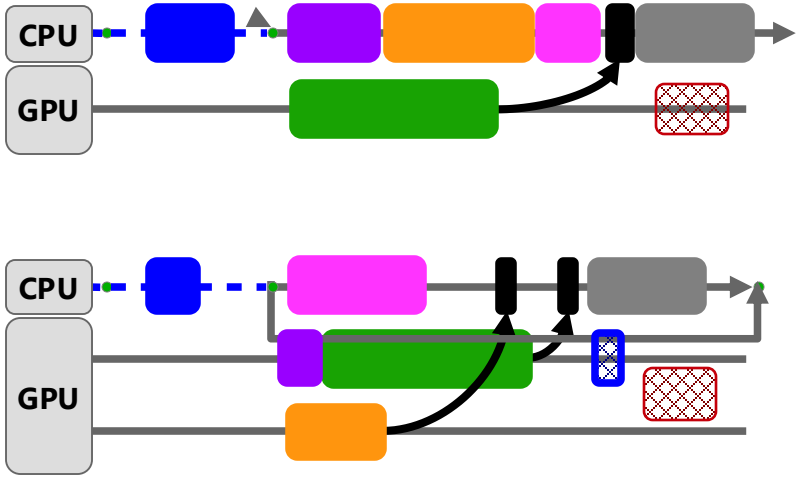


GROMACS parallelization schemes

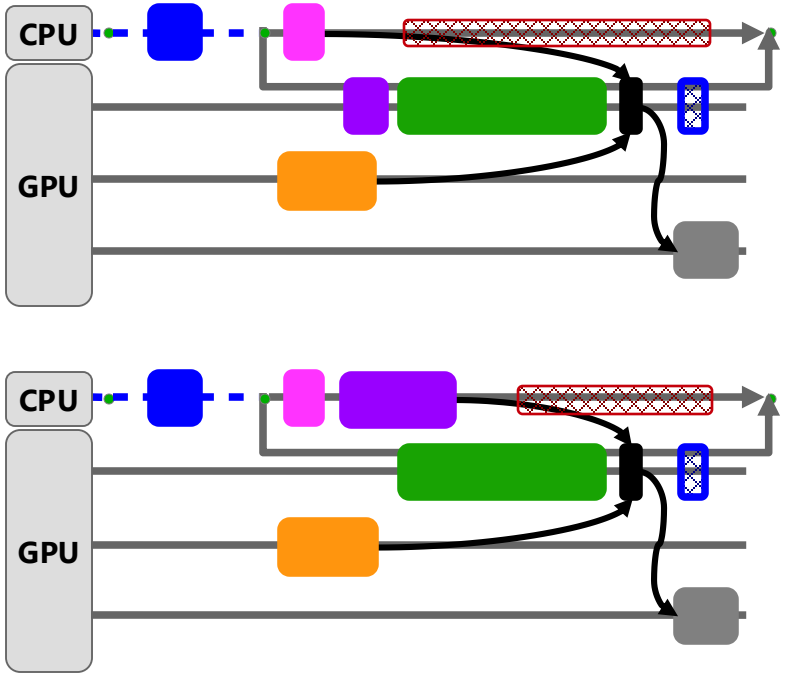
Homogeneous scheme



Force offload parallelization



GPU-resident parallelization

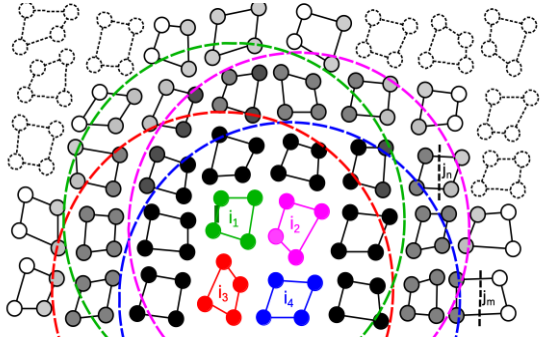


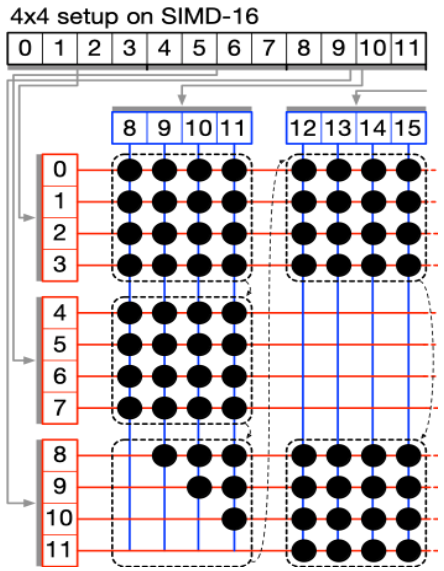
Szilárd Páll et al., J. Chem. Phys. 153, 134110 (2020) doi:10.1063/5.0018516

Learn more: http://bioexcel.eu/webinars

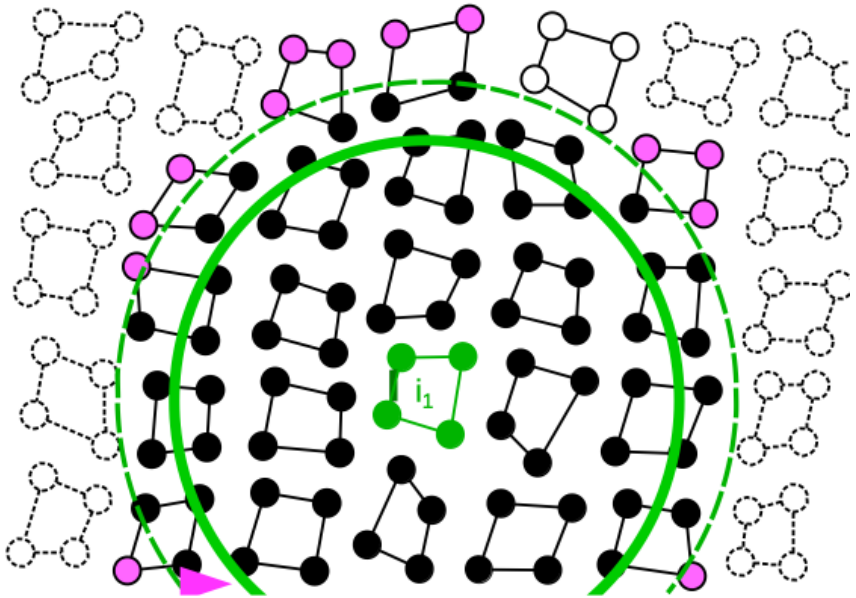
Algorithm redesign for modern architectures

Cluster pair-interaction algorithm for SIMD/SIMT

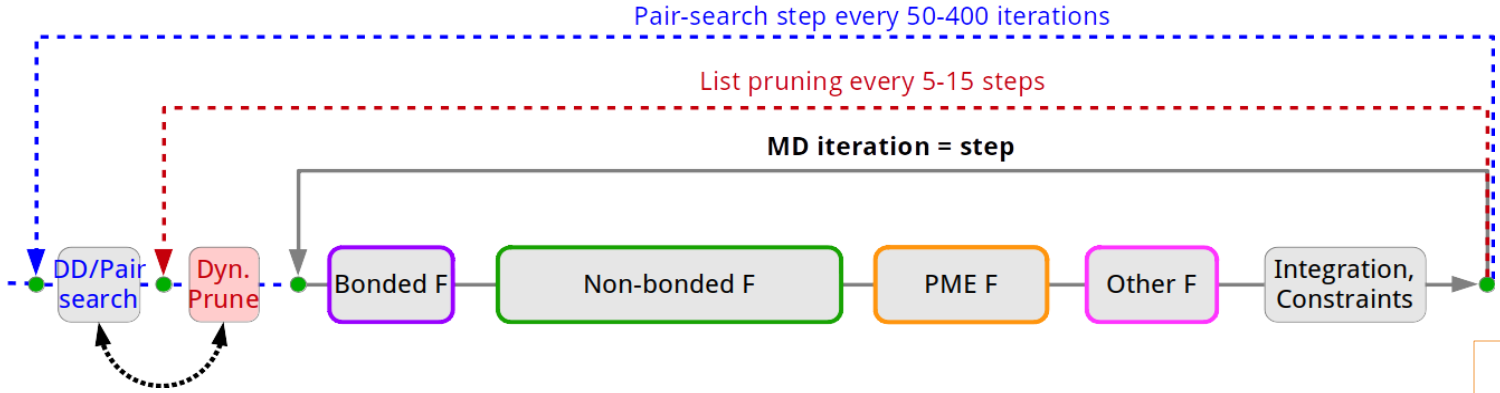




Accuracy-based automated list buffer improves SIMD algorithm parallel efficiency



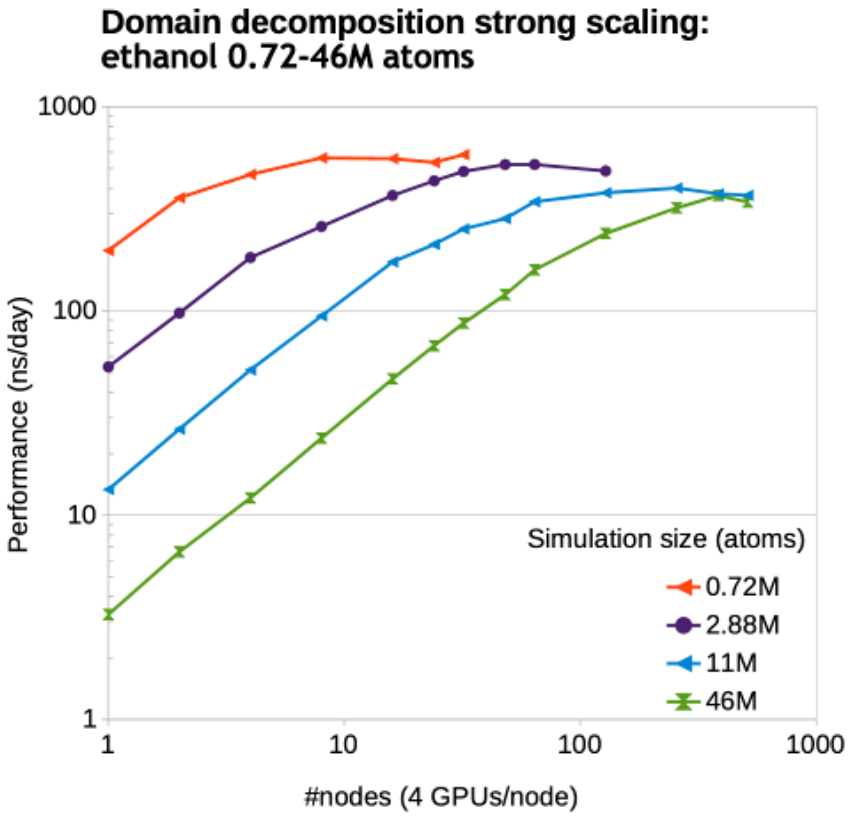
Dual pair list with dynamic pruning

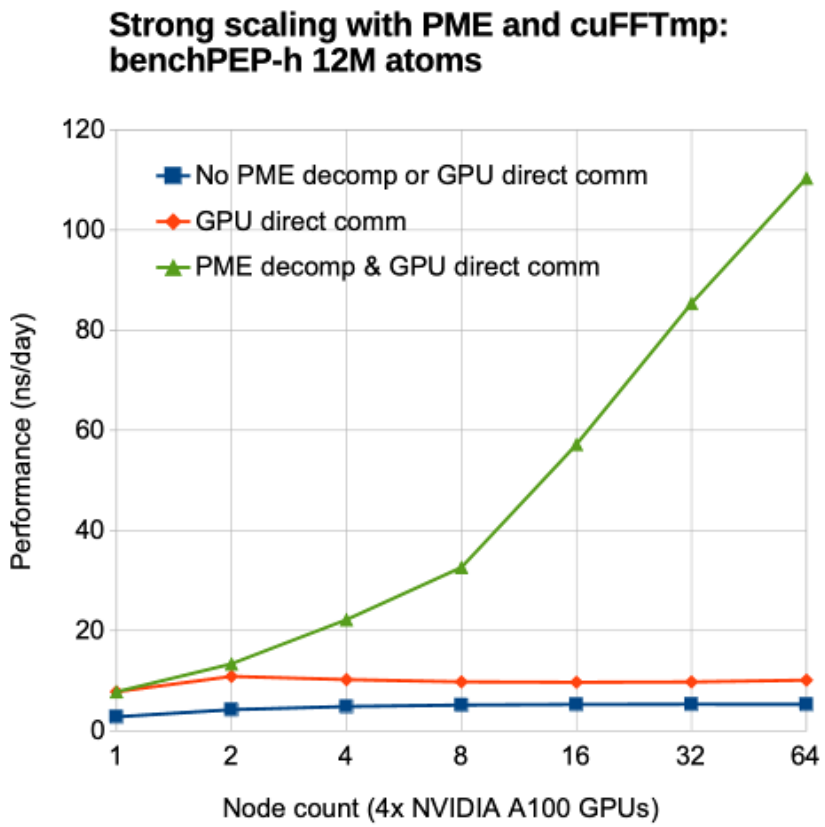


Páll, Szilárd, and Berk Hess, https://doi.org/10.1016/j.cpc.2013.06.003

Long-term readiness efforts

Direct GPU communication with proven strong scaling



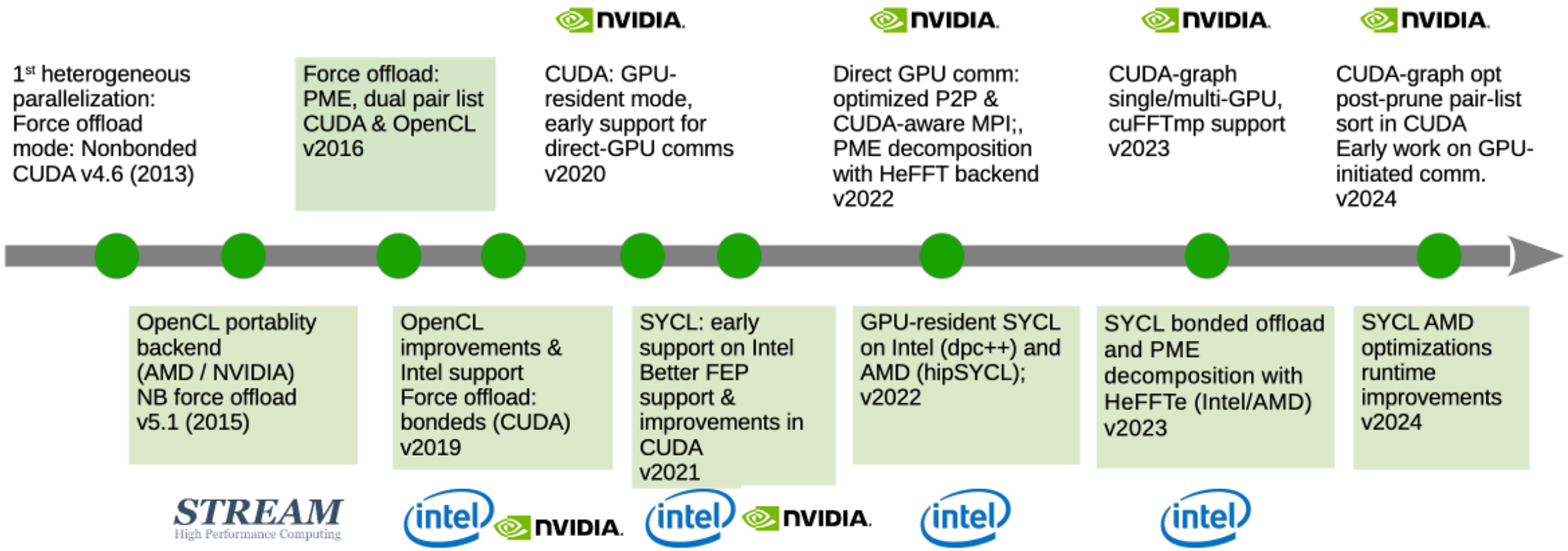


Recent & ongoing codesign activities

- NVIDIA collaboration:
  - >10 years of collaboration
  - Recent years: focus on codesign in anticipation of heterogeneous HPC in exascale
  - Direct GPU communication algorithm / implementation redesign
  - More efficient GPU scheduling: P2P multi-GPU, graph scheduling (ongoing)
  - Energy efficiency (ongoing)
  - Time-to-solution vs energy-to-solution
  - Considering from low-level algorithmic tuning to node-level optimization

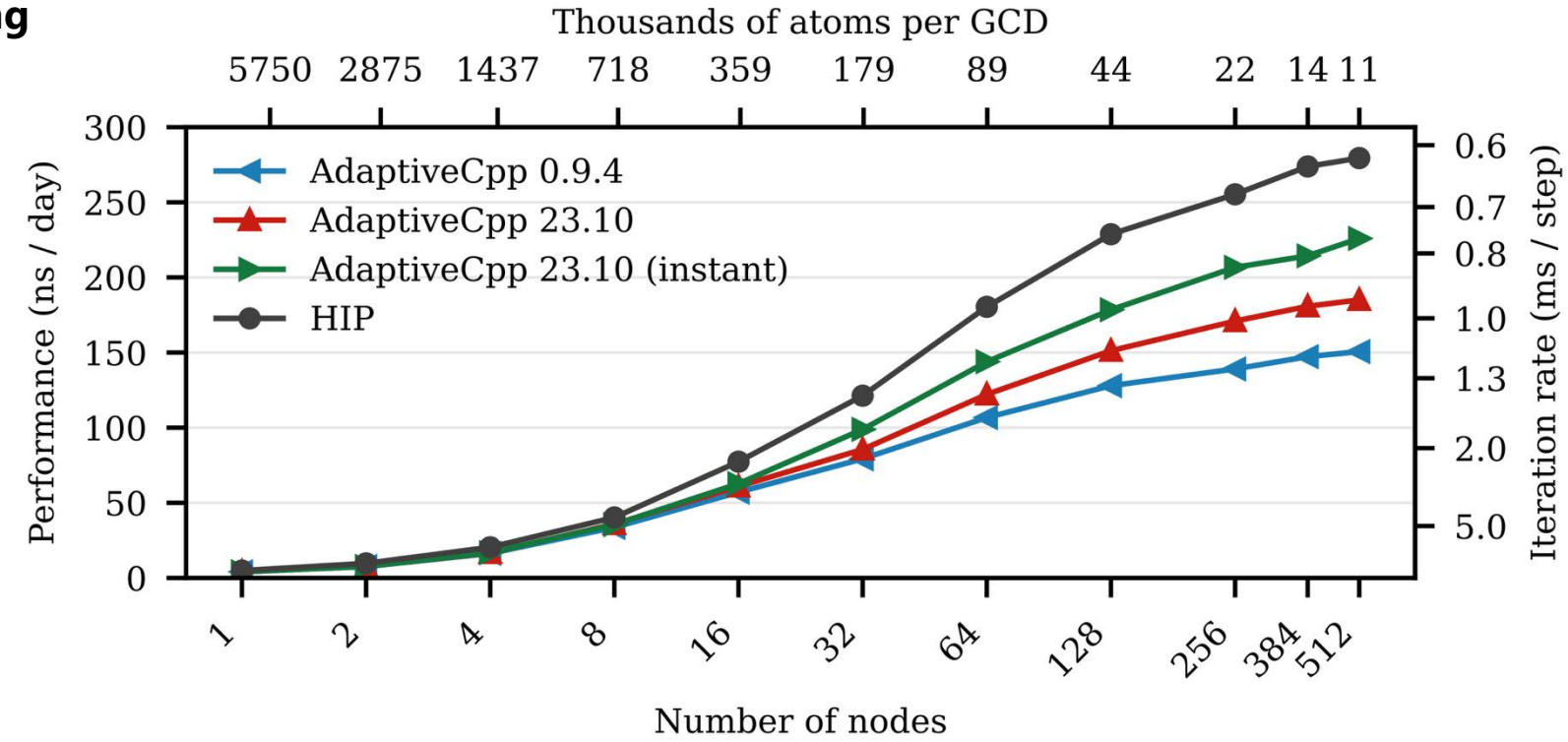
- Intel OneAPI CoE: development of SYCL port
  - as a new portability backend (replacing OpenCL)
  - SYCL in production on Intel and AMD GPU systems (ORNL Frontier, LUMI)
  - Collaboration on future programming models and standards: (hipSYCL, OneAPI/IntelLLVM)

Evolution of GPU hardware & API support



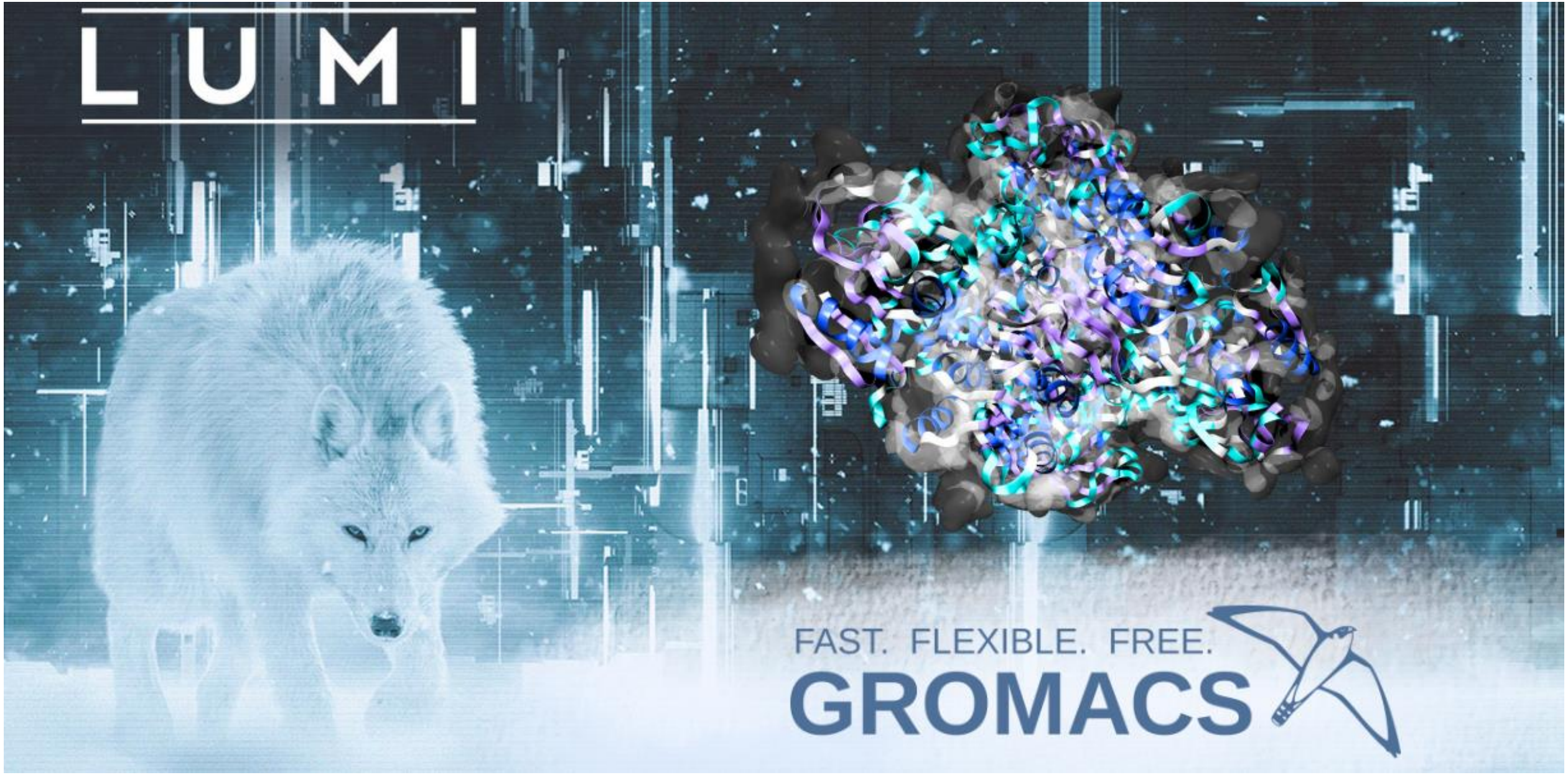
SYCL on AMD systems: strong scaling

GROMACS SYCL vs HIP fork scaling on Cray EX235a (LUMI-G)



Strong scaling of domain decomposition on up to 512 LUMI-G nodes

- parallel efficiency with ACPP instant submission on par with HIP fork
- absolute performance only ~15-20% from HIP fork (mainly due to compute kernels)



Rasmus Kronberg, CSC: https://bioexcel.eu/efficient-molecular-dynamics-simulations-with-gromacs-on-lumi

LUMI – Europe's Fastest Supercomputer

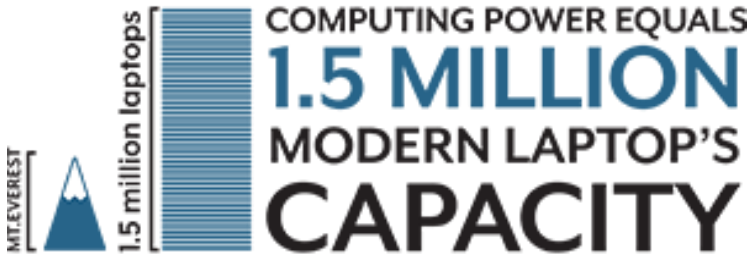
1 SYSTEM

375 Petaflop/s

SUSTAINED COMPUTING POWER

100%

HYDROELECTRIC ENERGY



COMPUTING POWER EQUALS

1.5 MILLION

MODERN LAPTOP'S

CAPACITY

LUMI – Europe's Fastest Supercomputer

CPU partition (Mahti)

Core Complex Die (CCD)

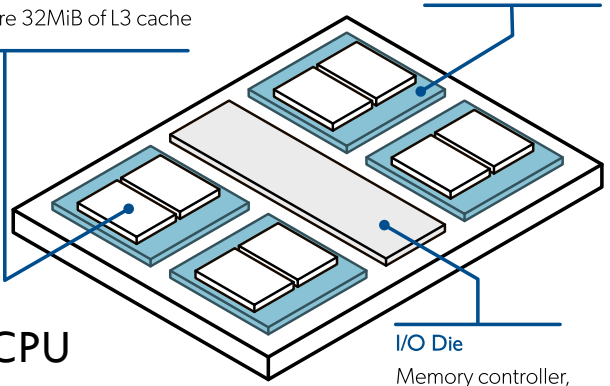
8 CPU cores with two-way SMT

- 32KiB of L1 cache
- 512KiB of L2 cache

Cores in a CCD share 32MiB of L3 cache

NUMA node

4 NUMA nodes per socket, 2 CCDs per NUMA node



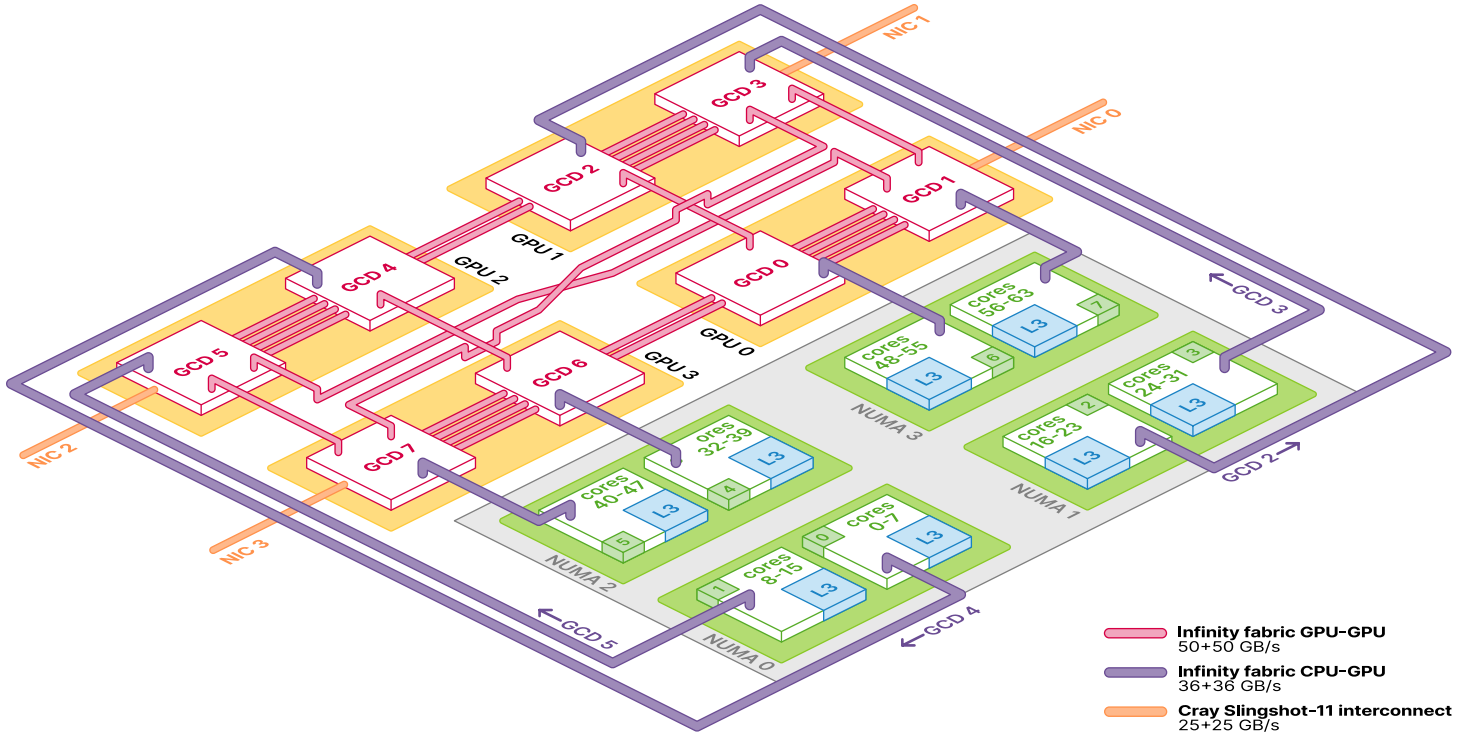
I/O Die

Memory controller, security features, PCIe and Infinity Fabric links

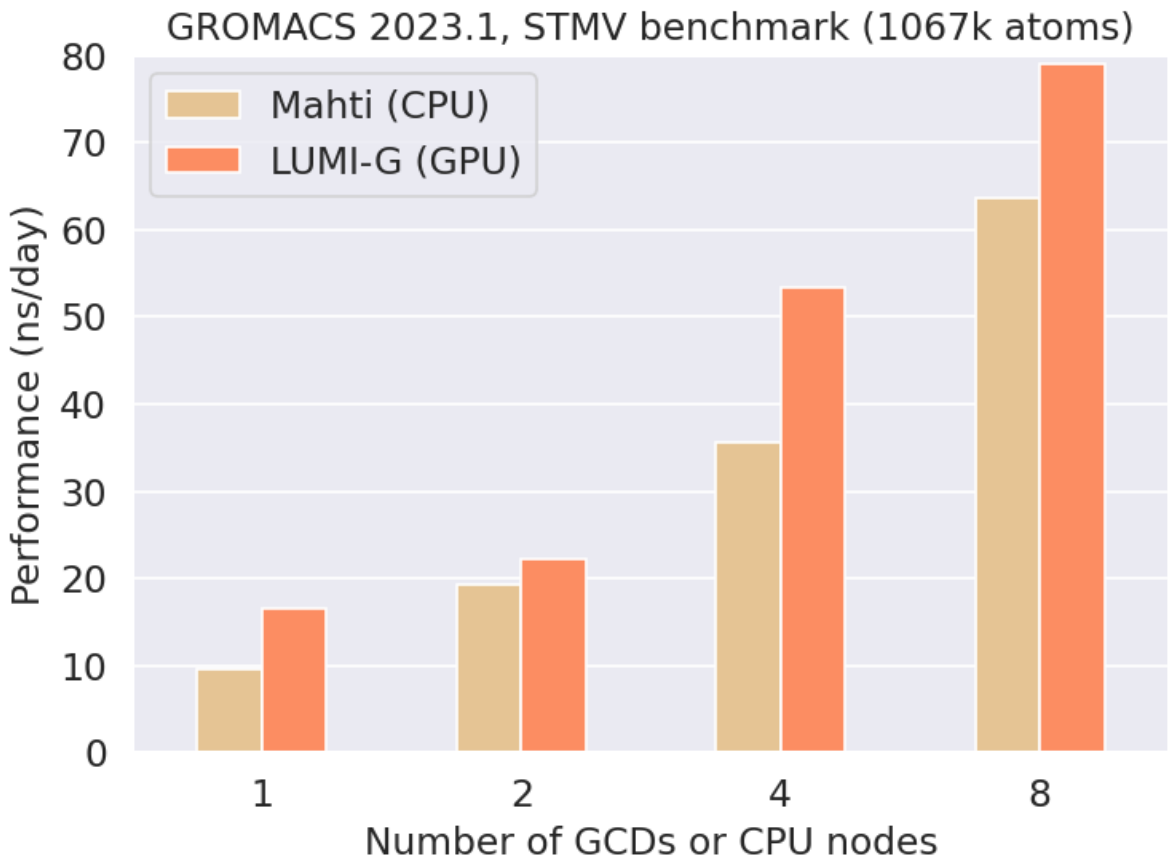
8 x CCDs / CPU

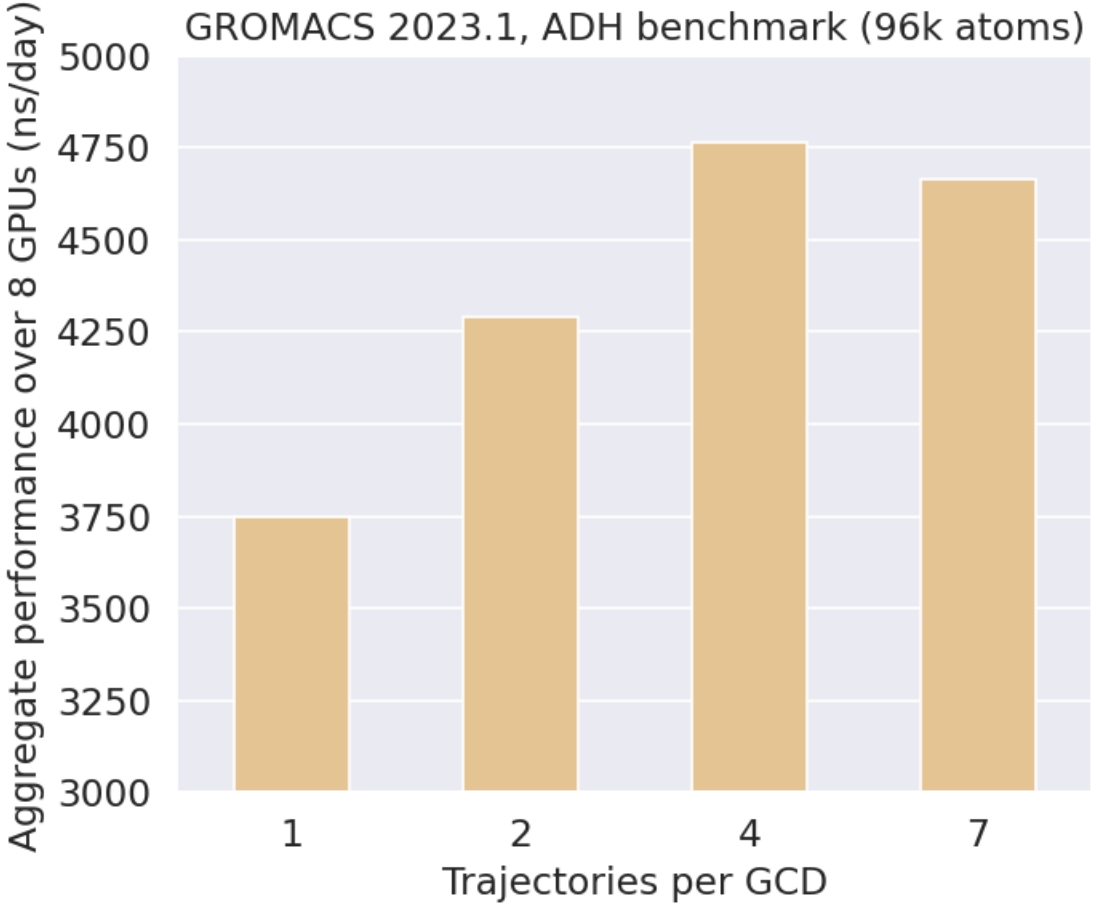
2 CPUs / Node

GPU partition (LUMI-G)



https://docs.lumi-supercomputer.eu/





Rasmus Kronberg, CSC: https://bioexcel.eu/efficient-molecular-dynamics-simulations-with-gromacs-on-lumi

GROMACS on LUMI

- Most systems run well on a single GCD, performance typically better than on one 128 core CPU node
- Large systems (several 100 000 – 1M atoms) are usually able to scale to multiple GPUs (mind the proper binding of CPUs and GPUs)
- GPU utilization can be maximized for small systems (less than 100 000 atoms) by running multiple independent trajectories per GCD (multidir feature)
- 100 microseconds per: either 1) 42 LUMI-G nodes (<2% of the total LUMI-G) or 2) 560 CPU-only nodes (40% of Mahti)

GROMACS developers team

Berk Hess
Mark Abraham
Paul Bauer
Szilard Pall
Aleksei Yupinov
Sander Pronk
Magnus Lundborg
Viveca Lindahl

Christian Wennberg
Christian Blau
Artem Zhmurov
David van der Spoel
Carsten Kutzner
Justin Lemkul
Roland Schulz
Sebastian Wingbermühle

Joe Jordan
Cathrine Bergh
Kevin Boyd
Alexey Shvetsov
Eric Irrgang
Pascal Merz
Tatjana Shugaeva
Farzaneh Jalalypour

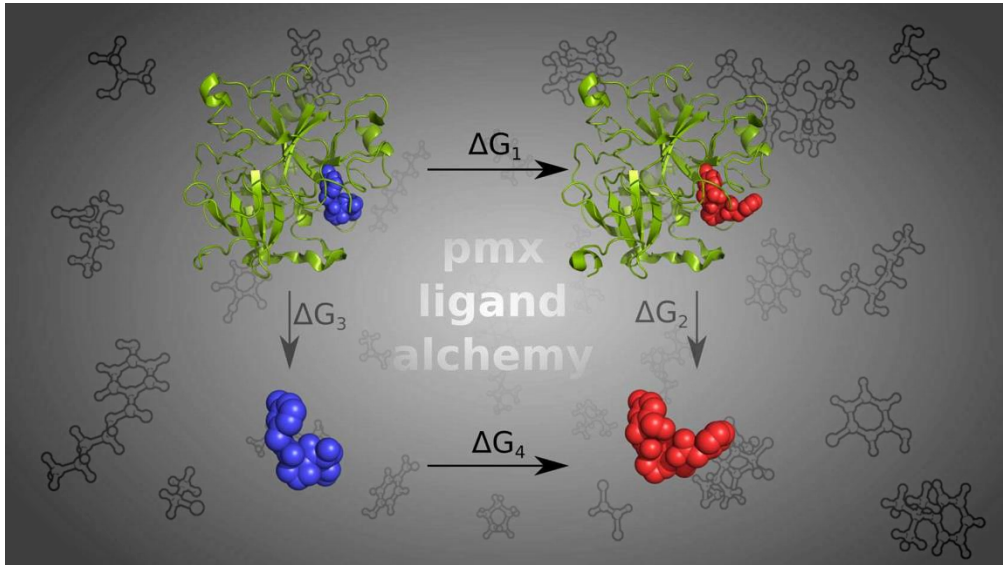
And many, many others (old & new) – many of which did not even code!

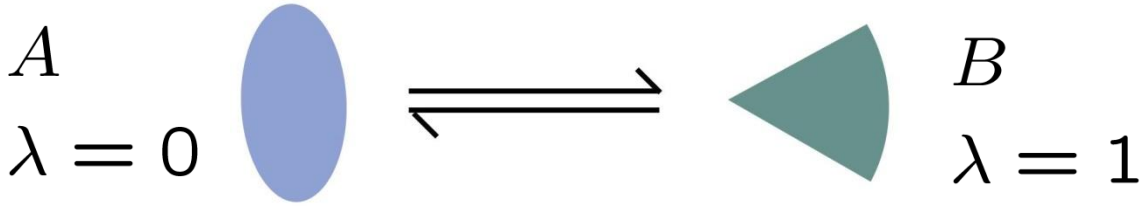
Stiftelsen för Strategisk Forskning
erc European Research Council
Sjunde Ramprogrammet
National Institutes of Health
PRACE
SeRC Swedish e-Science Research Centre
Vetenskapsrådet
bioexcel Center of Excellence for Computational Biomolecular Research

Free Energy Calculations:

GROMACS + PMX

Free Energy Calculation for Drug Screening

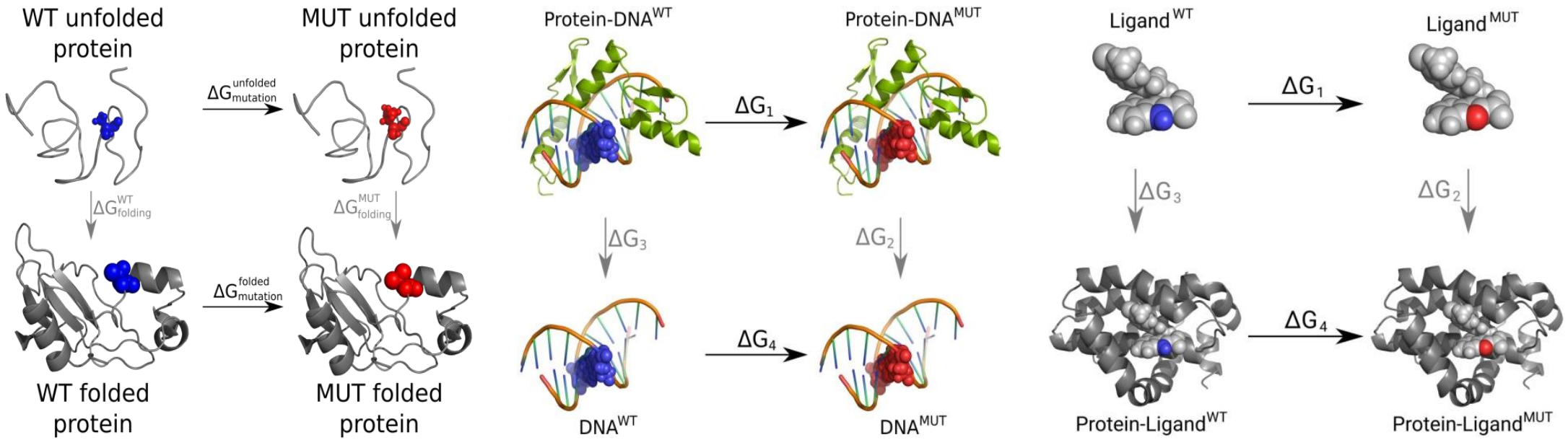




$$\Delta G = \int_0^1 \langle \frac{\delta H(r)}{\delta \lambda} \rangle_\lambda d\lambda$$

Bert de Groot et al.

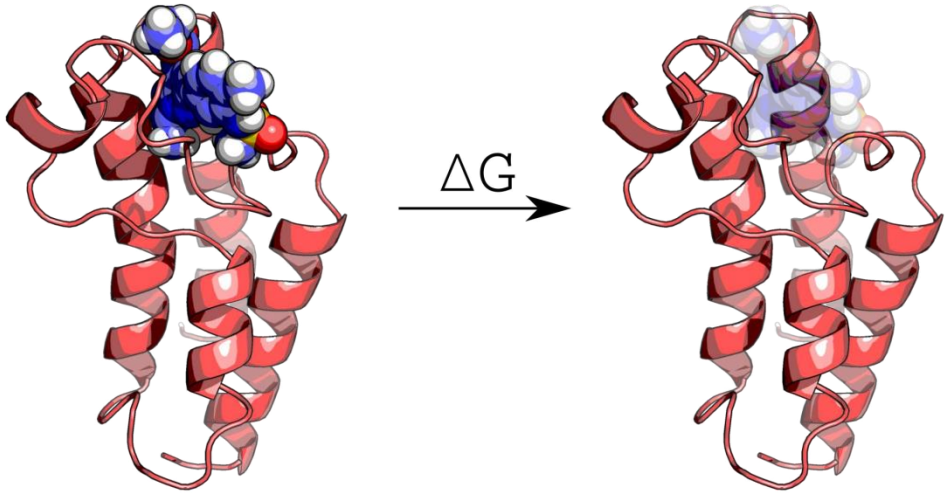
Alchemical transformations & Free Energy Calculations



Amino acid mutations | Nucleotide mutations | Ligand binding

Bert de Groot et al.

Absolute protein-ligand binding ΔG



Large scale absolute protein-ligand binding free energies

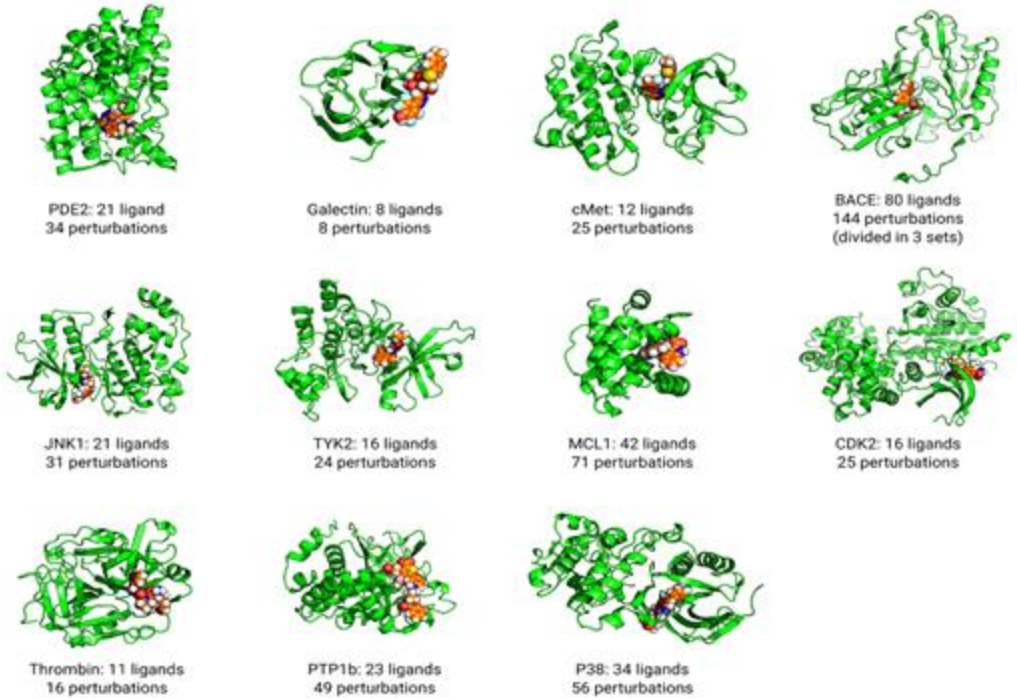
Chem. Sci, 2021

On-demand rapid large-scale drug screening with PMX: "Days, not months!" with HPC

A large-scale protein-ligand relative binding free energy scan: 10.1039/C9SC03754C

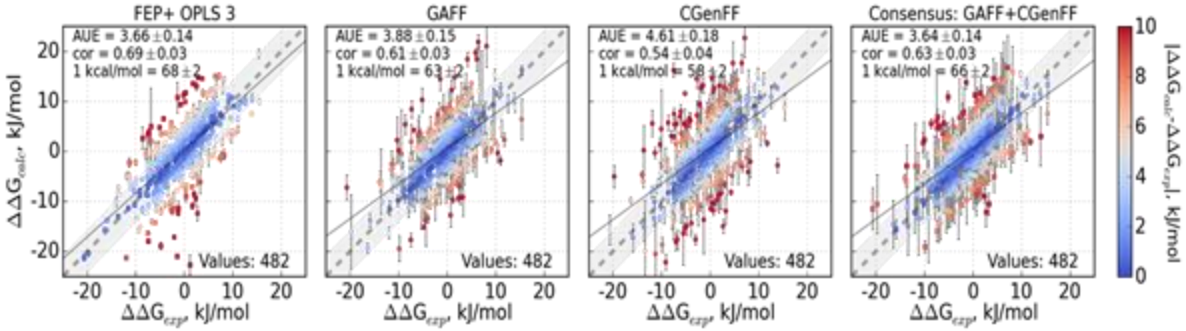
This study showcases scaling achieved with the non-equilibrium free energy calculation protocol using GROMACS+PMX.

482 relative binding free energy estimations in 11 protein-ligand systems using 2 molecular mechanics force fields.

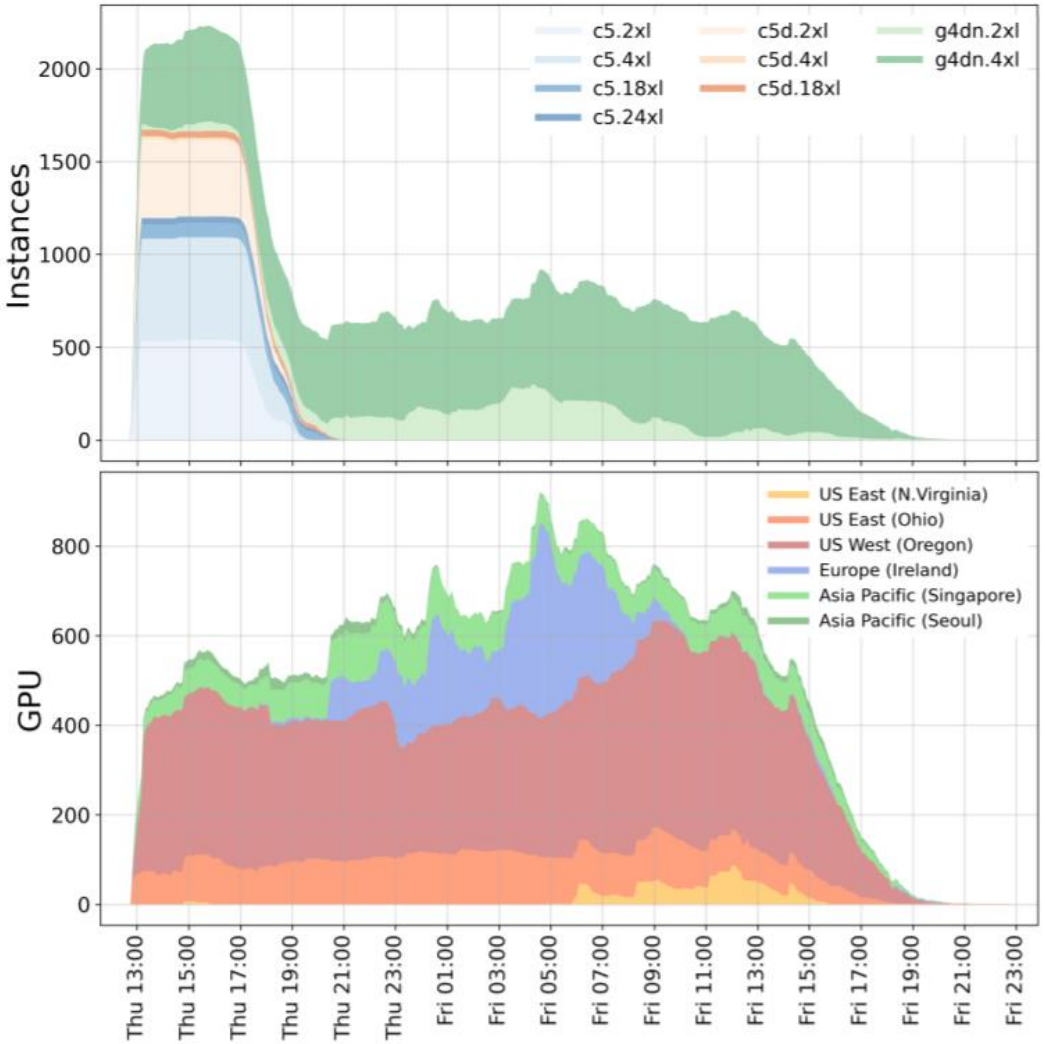


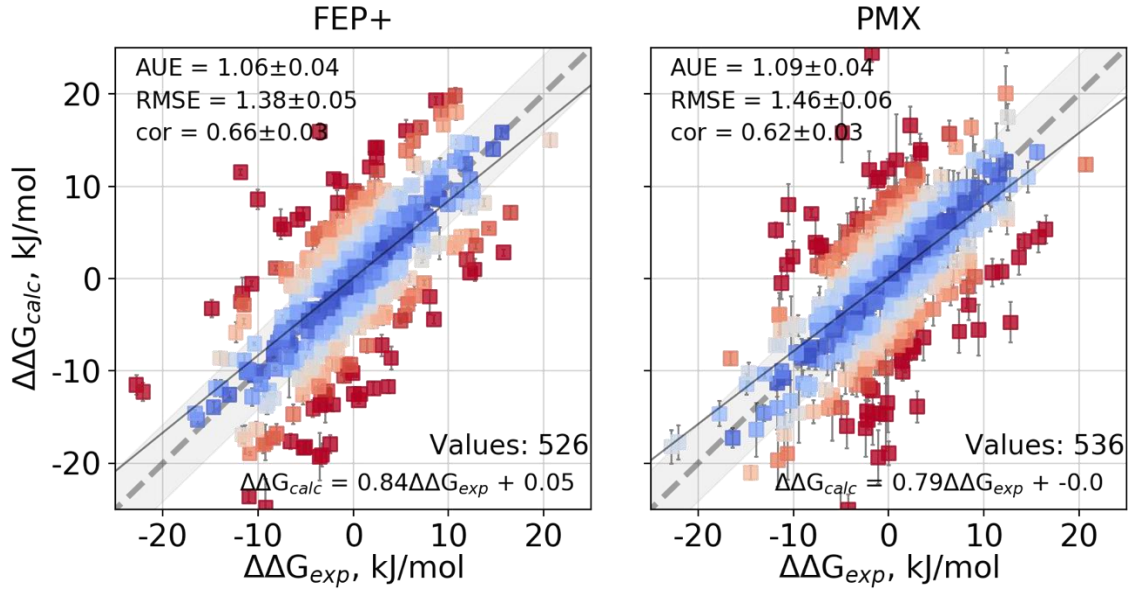
Similar case used for benchmarking

- Raven@MPCDF (**512 nodes**)
  - 1 node = Intel Xeon Cascade Lake-AP (96 cores)
- **480 nodes (~46k cores)**
- **~3.4 million core hours** in **3 days**



High-throughput screening: 5k ΔΔG in a day





Kutzner et al. JCIM 2022

PMX developers team

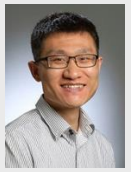
Membrane channels:

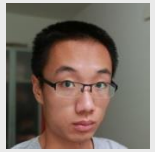
Andrei Mironenko

Wojciech Kopec

Csaba Daday

Ruo-Xu Gu





Tom Lam

Chengong Hui

Thomas Baukrowitz
Adam Lange
Ulrich Zachariae
Kornelius Zeth
Claudia Steinem
Thomas Jansen
Markus Zweckstetter

SFB 803

Dynamon

TI PHARMA

Free energy calculations:

Bert De Groot

Martin Paleico

Vytautas Gapsys

Yuriy Khalak

Matteo Aldeghi

Kai Tittmann
David Mobley
Hannah Baumann
Boehringer Ingelheim
Janssen pharmaceuticals
AstraZeneca

bioexcel

Boehringer Ingelheim

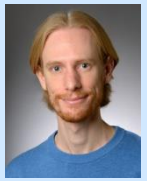
Alexander von Humboldt Stiftung/Foundation

Deutsche Forschungsgemeinschaft DFG

Collective dynamics/ aggregation:

Martin Werner

Dirk Matthes



Benjamin Eltzner

Axel Munk
Christian Griesinger
Markus Zweckstetter
Loren Andreas

Alexander von Humboldt Stiftung/Foundation

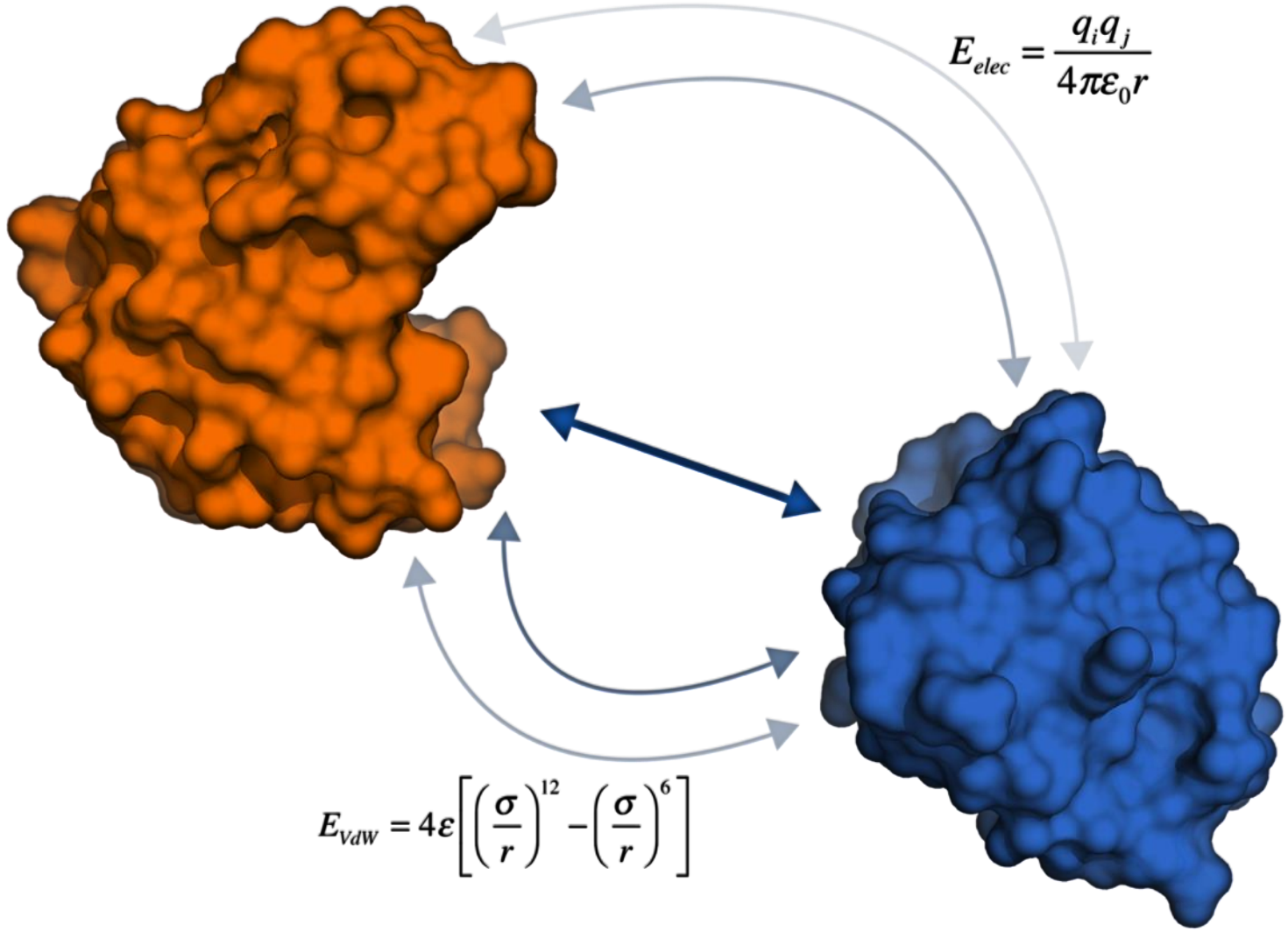
IMPRS

HADDOCK

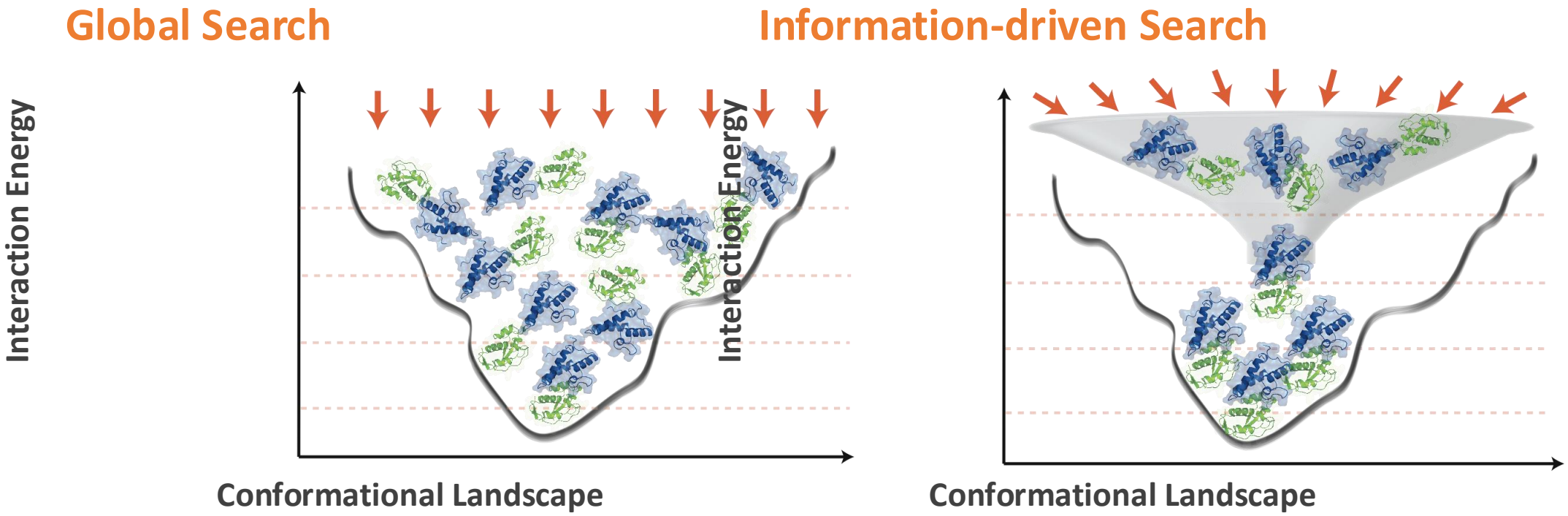
for

Integrative Modelling

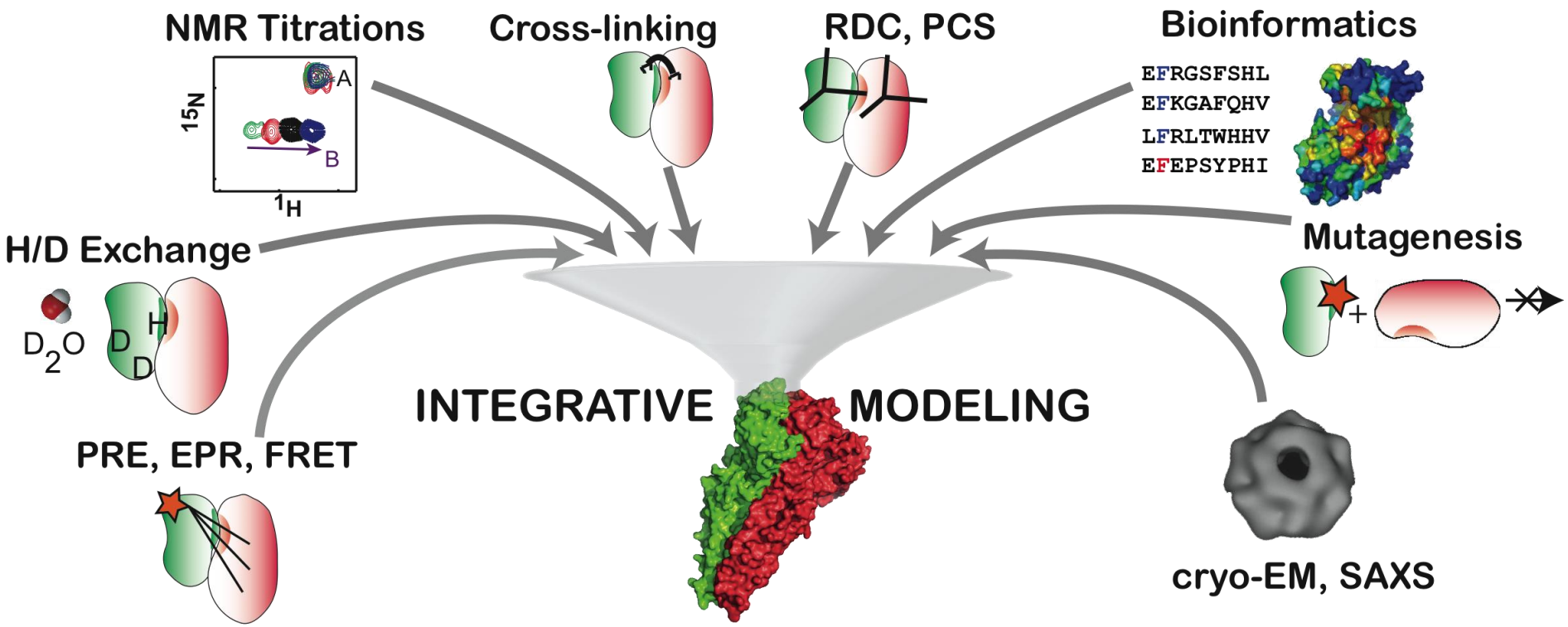
Molecular Docking



Data Integration during Sampling



What is Integrative Modeling?



HADDOCK:
An integrative modeling platform

Incorporates ambiguous and low-resolution data to aid the docking

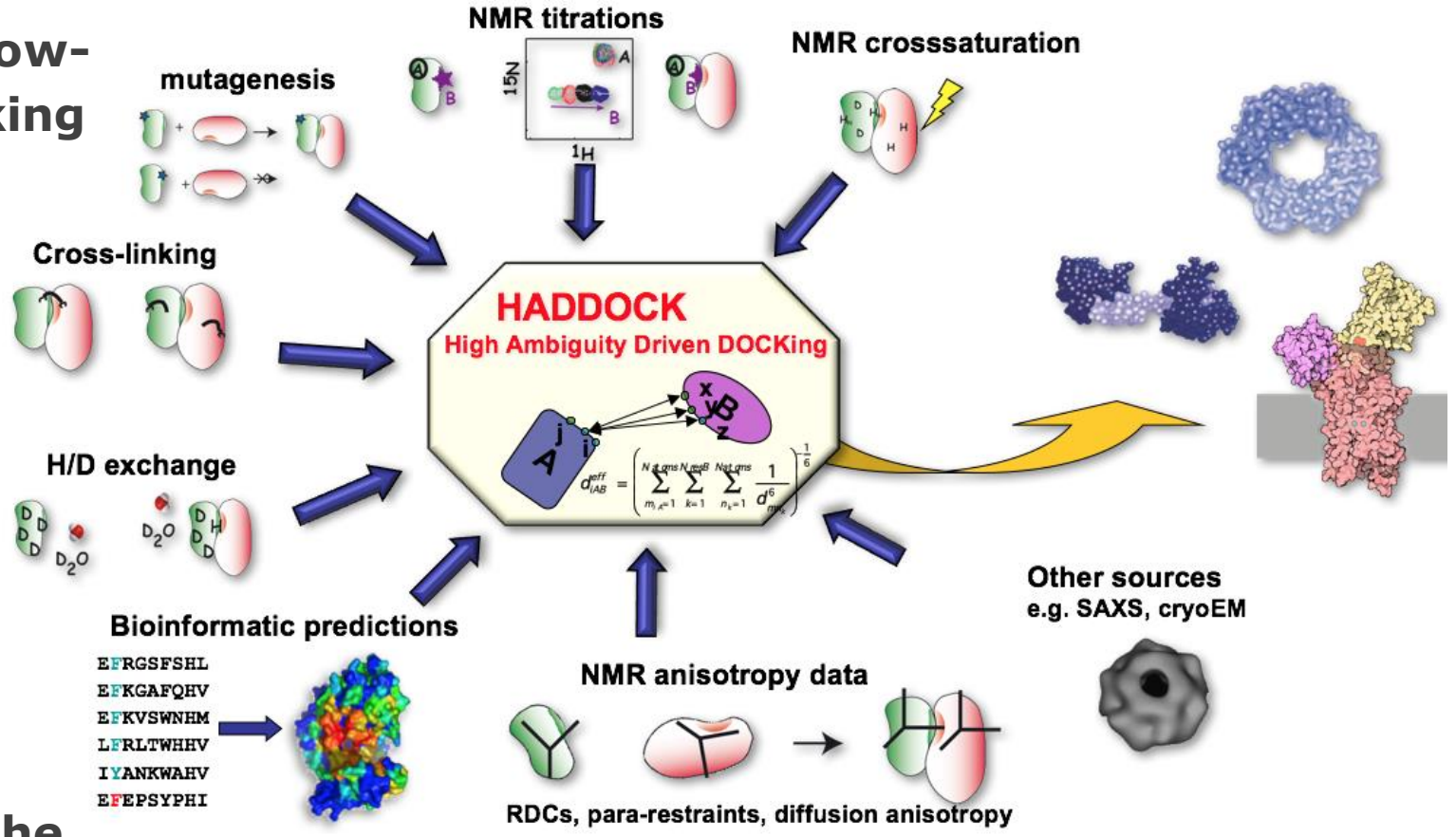
Capable of docking up to 20 molecules (2.4 version)

Symmetries can be leveraged

Allows for flexibility at the interface

Final flexible refinement in explicit solvent

Consistent performance over the years in CAPRI



Dominguez, Boelens & Bonvin. *JACS* 125, 173 (2003).

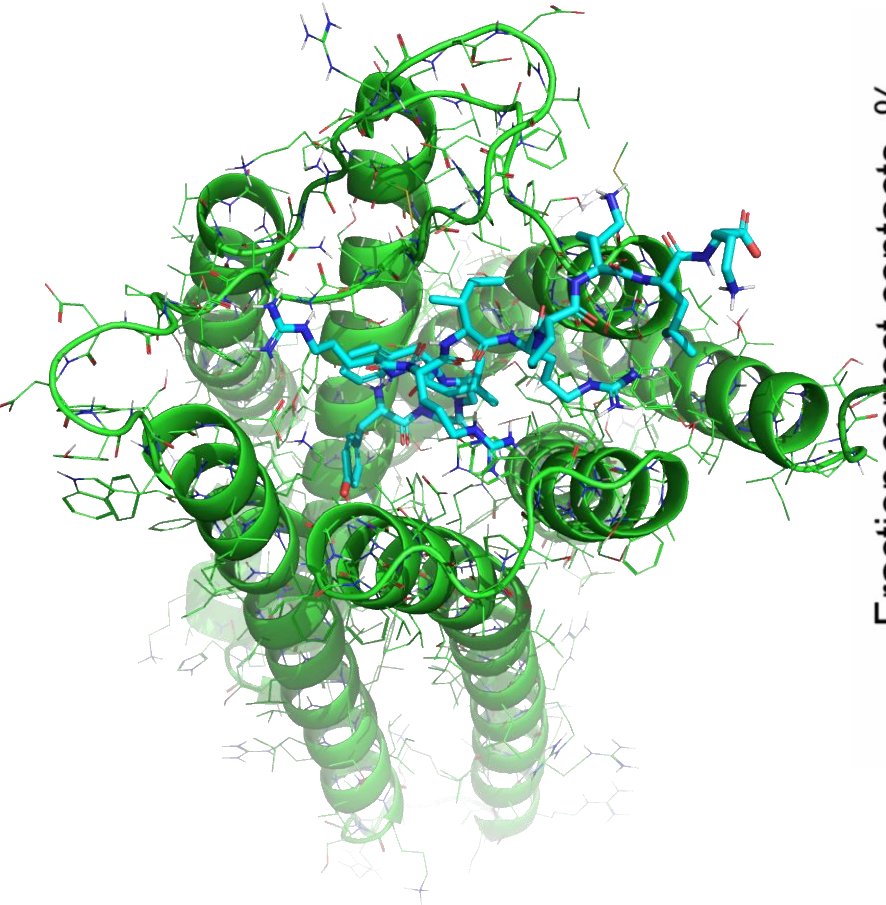
http://www.bonvinlab.org/software

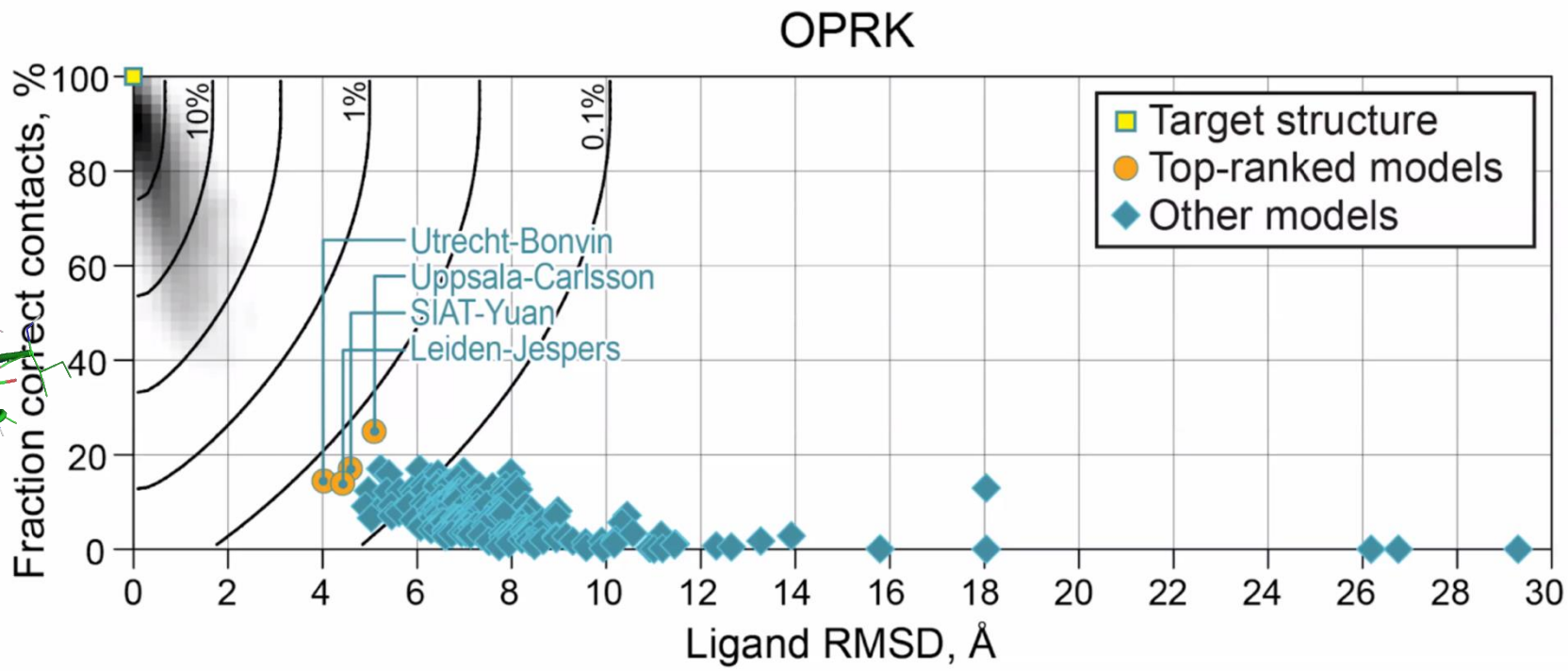
Siri Van Keulen

Manon Réau

GPCRdock 2021

Shape-restrained protocol put to the test in the blind GPRCdock experiment



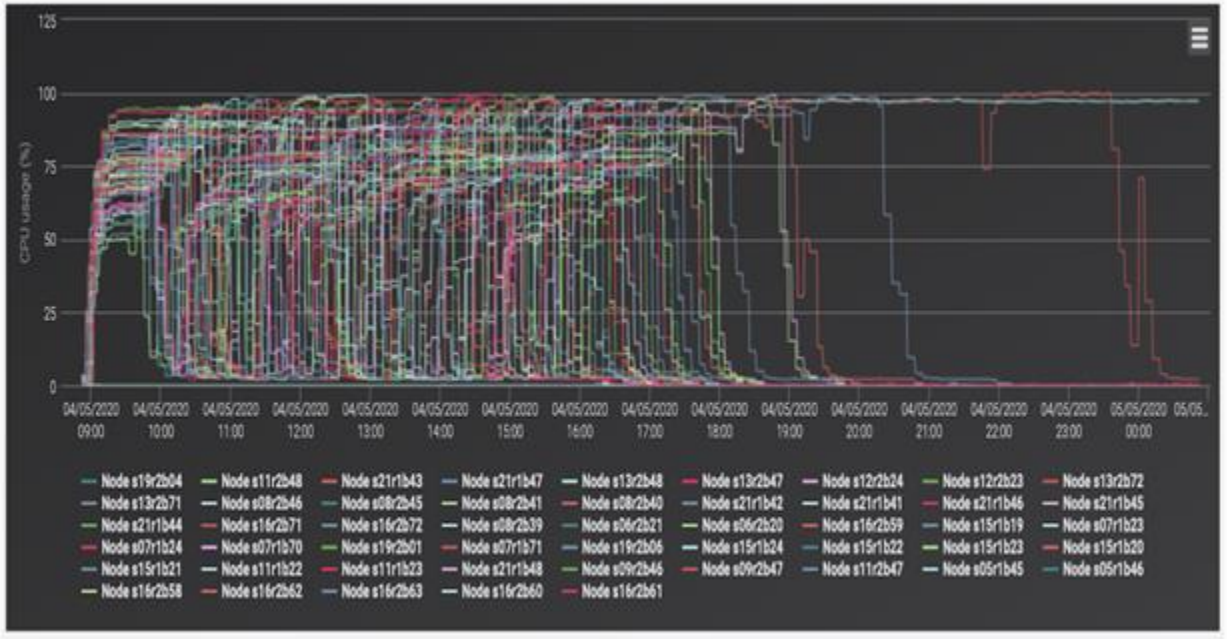


https://gpcrdock2021.org

Large-scale modelling

Demonstration on MareNostrum BSC

- One bioinformatics workflow
- 50 nodes - 2400 cores
- Completed in ~16 hours,
- Generated ~175GB of data.



HADDOCK server at a glance:

- **>21,000 jobs executed** 80% of which on EOSC HTC resources
- **Number of users: >18,500** (from >110 different countries)

HADDOCK 3
High-Ambiguity Driven Docking

- **Modular** re-write in Python 3
- Testing & **Continuous Integration** enabled
- **Free and Open Source** (Apache 2)
- Supports **basic docking and scoring scenarios**
- Ready for **custom workflows** and integration of third-party software

HADDOCK developer team / Alexandre Bonvin

€€

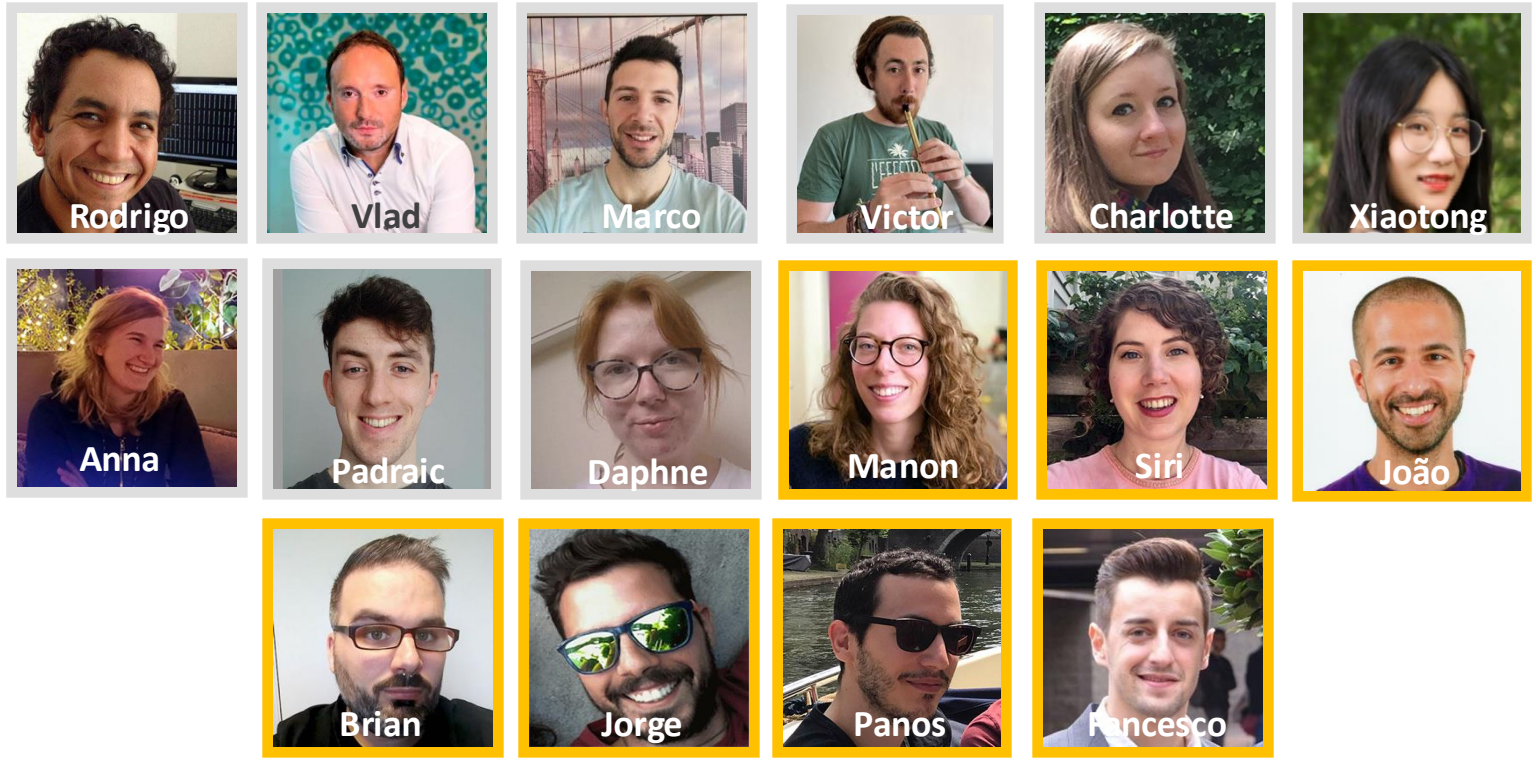
NWO

VICI
TOP-PUNT

WeNMR
West-Life
EGI-Engage
INDIGO-Datacloud
BioExcel CoE
EOSC-Hub
EGI-Ace

bioexcel
Centre of Excellence for Computational Biomolecular Research

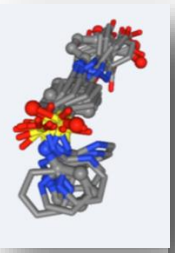
former (some) and current CSB group@UU



bonvinlab.org/people

Automation Workflows:

BioExcel Building Blocks (BioBB)



BioActive Compounds pipeline: a platform designed to efficiently generate **bioactive conformers** and speed up the **drug discovery process**.

http://mmb.irbbarcelona.org/BCE/

Sanja Zivanovic et al., J. Chem. Theory Comput. 2020, 16, 10, 6586–6597

Input Ligand
3-letter Code
SMILES
PDB File

Generate structure
OpenBabel

Ligand Parameterization

Step 1. Ligand Parameterization.
Adding Hydrogens, energetically minimizing the system, and generating parameters.

- **Adding Hydrogens** — OpenBabel
- **Energetically minimizing system** — Obminimize
- **Ligand parameterization** — ACPype

MD System Setup

Step 2. MD System Setup.
Building topology and solvent box.

- **Building GROMACS Topology** — GROMACS
- **Building System Box** — GROMACS editconf
- **Solvating the System** — GROMACS genbox

Hamiltonian Replica Exchange Simulations (Equilibration Phase)

Step 3. Hamiltonian Replica Exchange Simulations (Equilibration Phase).
Energetically equilibrating the system (ligand + solvent), for N replicas.

Generate Hamiltonian REMD equilibration configuration files — GROMACS + plumed

Replica 1	Replica 2	Replica N
Energy equilibration — GROMACS grompp	Energy equilibration — GROMACS grompp	Energy equilibration — GROMACS grompp
Equilibration MD run — GROMACS mdrun	Equilibration MD run — GROMACS mdrun	Equilibration MD run — GROMACS mdrun

Hamiltonian Replica Exchange Simulations (Production Phase)

Step 4. Hamiltonian Replica Exchange Simulations (Production Phase).
Running Hamiltonian REMD simulations for N replicas.

Generate Hamiltonian REMD equilibration configuration files — GROMACS + plumed

Replica 1	Replica 2	Replica N
Energy equilibration — GROMACS grompp	Energy equilibration — GROMACS grompp	Energy equilibration — GROMACS grompp
Equilibration MD run — GROMACS mdrun	Equilibration MD run — GROMACS mdrun	Equilibration MD run — GROMACS mdrun

Trajectory Clusterization

Step 5. Trajectory Clusterization.
Extract representative conformations from Hamiltonian REMD generated trajectory.

- **Trajectory Clusterization** — GROMACS cluster
- **Trajectory Clusterization post-process** — Dihedral-clustering

Quantum Mechanics

Step 6. Quantum Mechanics.
Computing QM energy using 2 levels of theory: Hartree Fock (HF/3-21G) and B3LYP (B3LYP/6-31G).

- **QM Hartree Fock HF/3-21G** — Gaussian
- **QM B3LYP B3LYP/6-31G*** — Gaussian

Output
- *QM energy conformer 1*
- *QM energy conformer 2*
- *QM energy conformer 3*
- *QM energy conformer N*

GROMACS
FAST. FLEXIBLE. FREE.

Amber18

AmberTools19

OpenMM

ADDOCK
High-Ambiguity Driven Docking

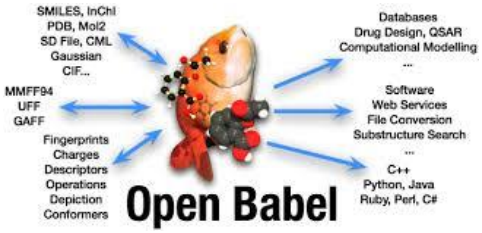
NAMD
Scalable Molecular Dynamics

VMD
Visual Molecular Dynamics

CPMD
QM/MM

CP2K

ACPYPE



RDKit

pmx: generate hybrid protein structure and topology
Computational Biomolecular Dynamics Group

Gaussian, Inc.

Bio3D

MD ANALYSIS

AutoDock 4

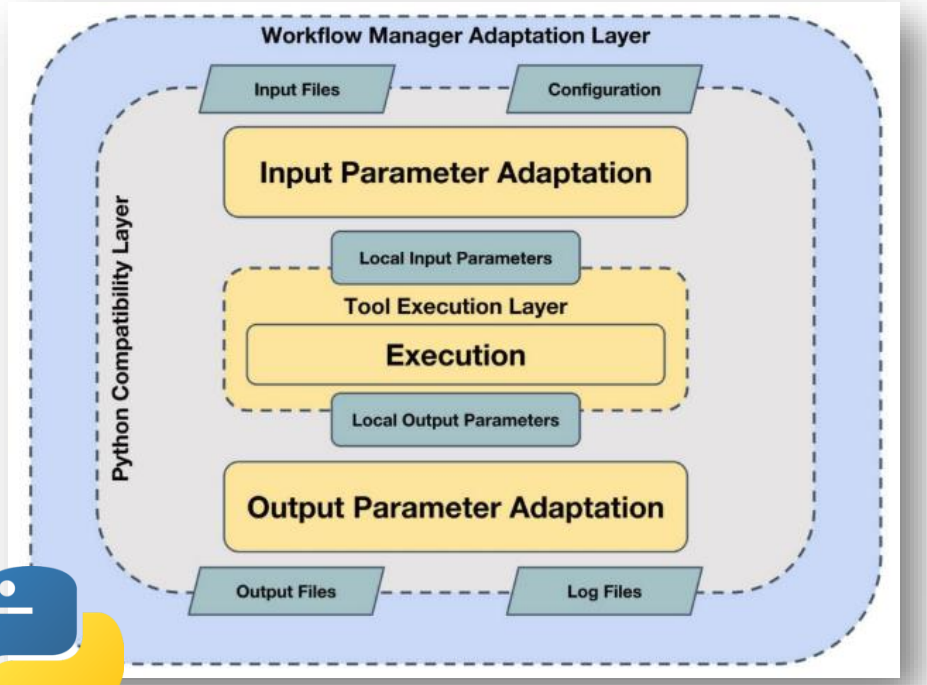
Modeller
Program for Comparative Protein Structure Modelling by Satisfaction of Spatial Restraints

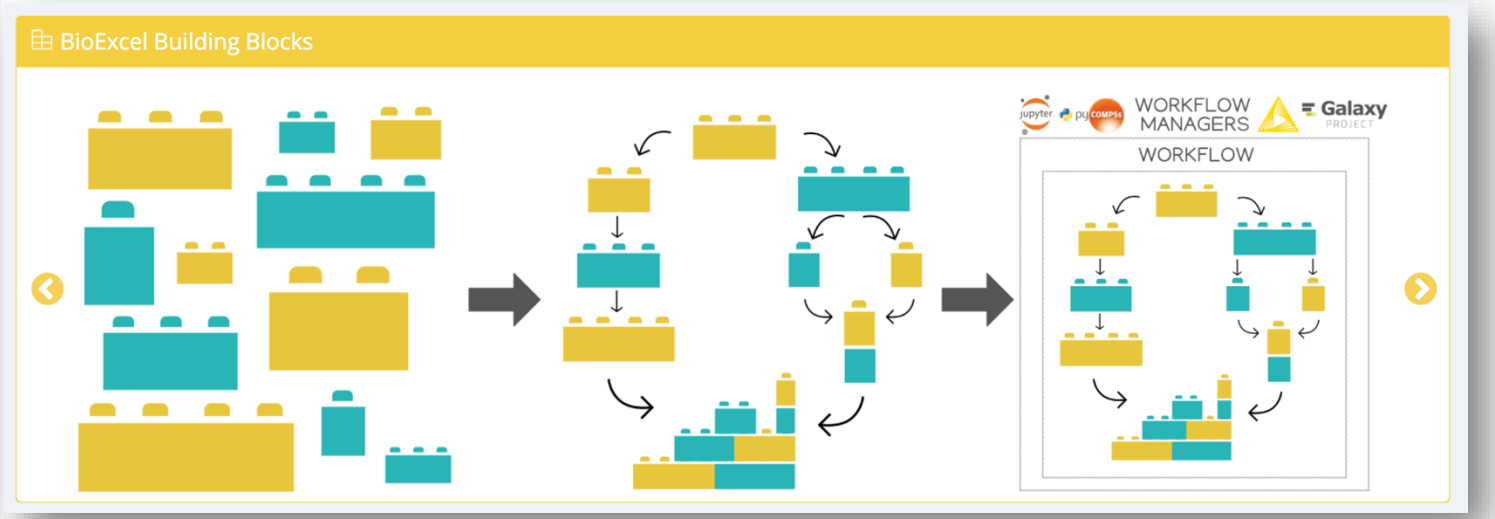
TensorFlow

scikit learn

And many many more…

BioExcel Building Blocks: BioBB



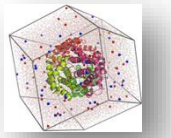


1) Building Blocks *2) Workflows* *3) Workflow Managers*

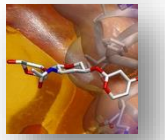
https://mmb.irbbarcelona.org/biobb/

BioBB Modules (17, Release 2023.2)

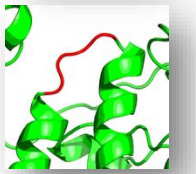
biobb_common
Common auxiliar functions



biobb_gromacs
Molecular Dynamics GROMACS

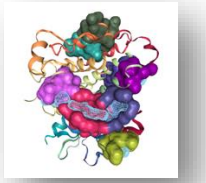


biobb_amber
Molecular Dynamics AMBER



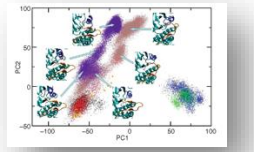
biobb_model
Molecular Modelling

biobb_ml
Machine learning

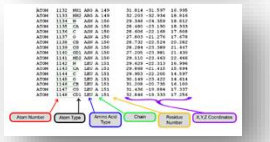


biobb_vs
Virtual Screening

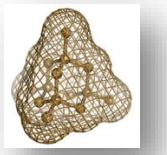
biobb_io
Biological databases



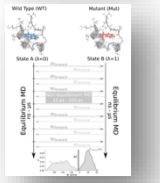
biobb_analysis
MD trajectories analysis



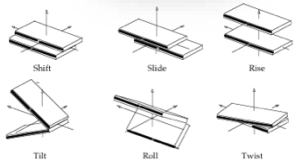
biobb_structure_utils
Modify or extract information from PDB



biobb_chemistry
Chemoinformatics functionalities



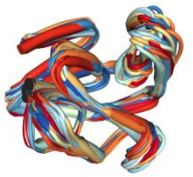
biobb_pmx
Free energy calculations



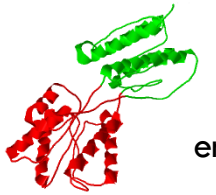
biobb_dna
Nucleic Acids MD
Trajectory analyses

biobb_cmip
Molecular Interaction Potentials

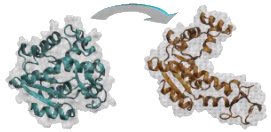
biobb_cp2k
Quantum Mechanics



biobb_flexdyn
NMA-based conformational ensemble generation



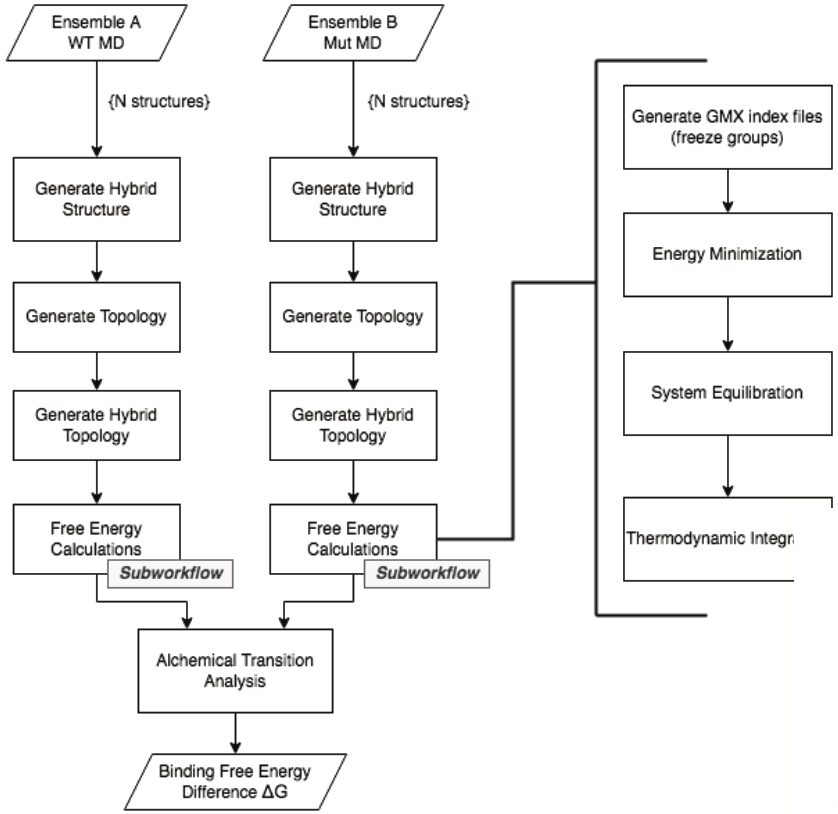
biobb_flexserv
Coarse-Grained conformational ensemble generation and flexibility analysis



biobb_godmd
Coarse-Grained conformational transitions

biobb Release 2023.2

And many others to come…




```python
from biobb_common.configuration import settings
from biobb_md.gromacs.pdb2gmx import pdb2gmx
from biobb_md.gromacs.editconf import editconf
...
def main(config, system):
    ...
    # Load configuration
    conf = settings.ConfReader(config, system)
    mut_prop = conf.get_prop_dic()
    mut_paths = conf.get_paths_dic()
    ...
    # Workflow definition
    for structure in conf.properties['input_structures'].split(','):
        pdb2gmx(**mut_paths["step1_pdb2gmx"], properties=mut_prop["step1_pdb2gmx"])
        editconf(**mut_paths["step2_editconf"], properties=mut_prop["step2_editconf"])
        solvate(**mut_paths["step3_solvate"], properties=mut_prop["step3_solvate"])
        grompp(**mut_paths["step4_grompp"], properties=mut_prop["step4_grompp"])
        mdrun(**mut_paths["step5_mdrun"], properties=mut_prop["step5_mdrun"])
```

WMS adapter

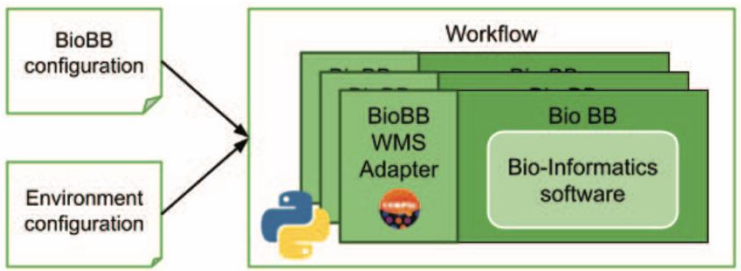
```python
...
@task(pdb_path=FILE_IN, gro_path=FILE_OUT,top_zip_path=FILE_OUT)
```

BioBB definition

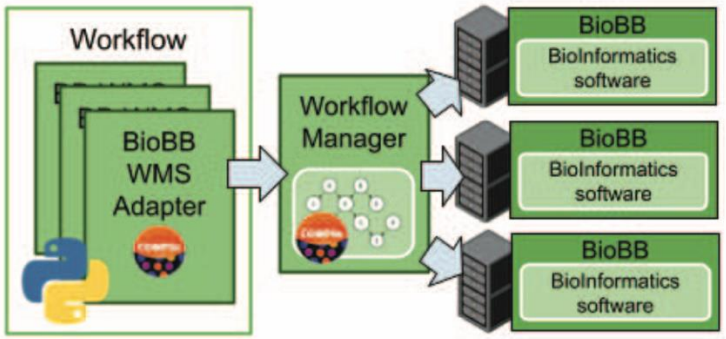
```python
def pdb2gmx(pdb_path, gro_path, top_zip_path, properties):
```

BioBB implementation

```python
""" pdb2gmx biobb implementation """
```

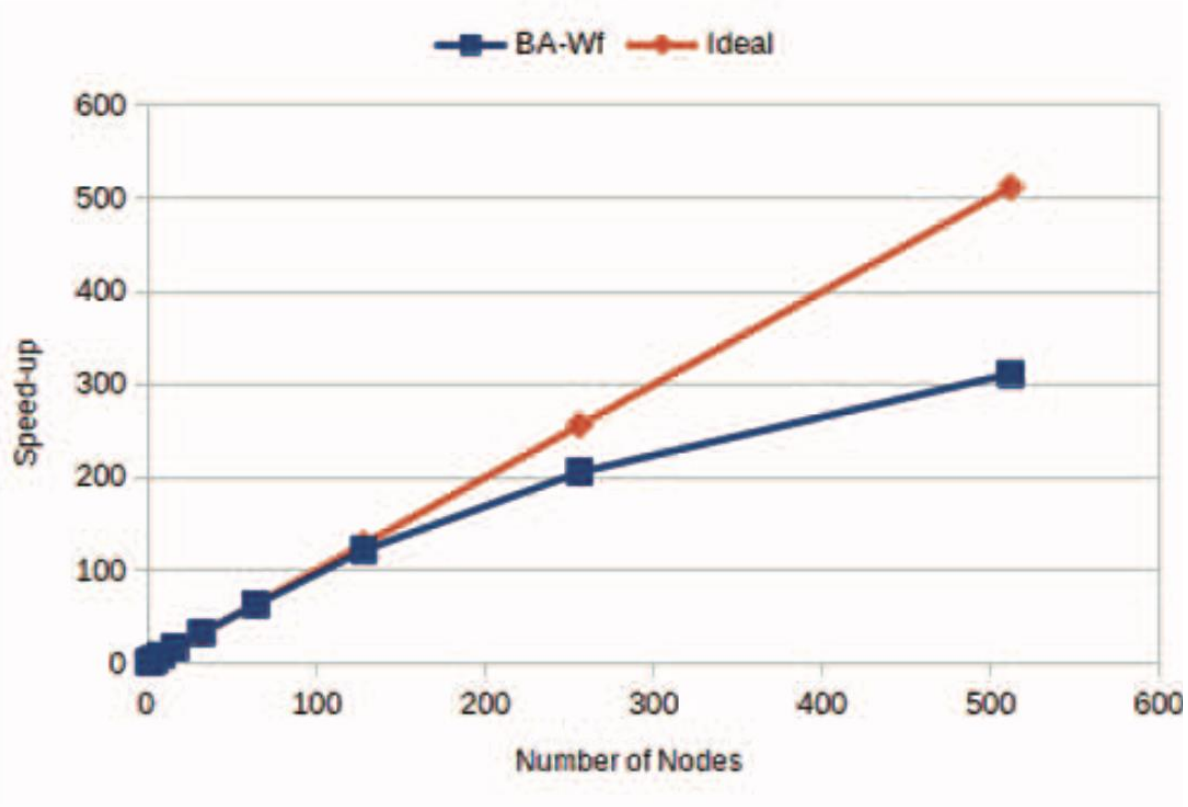


(a) Development

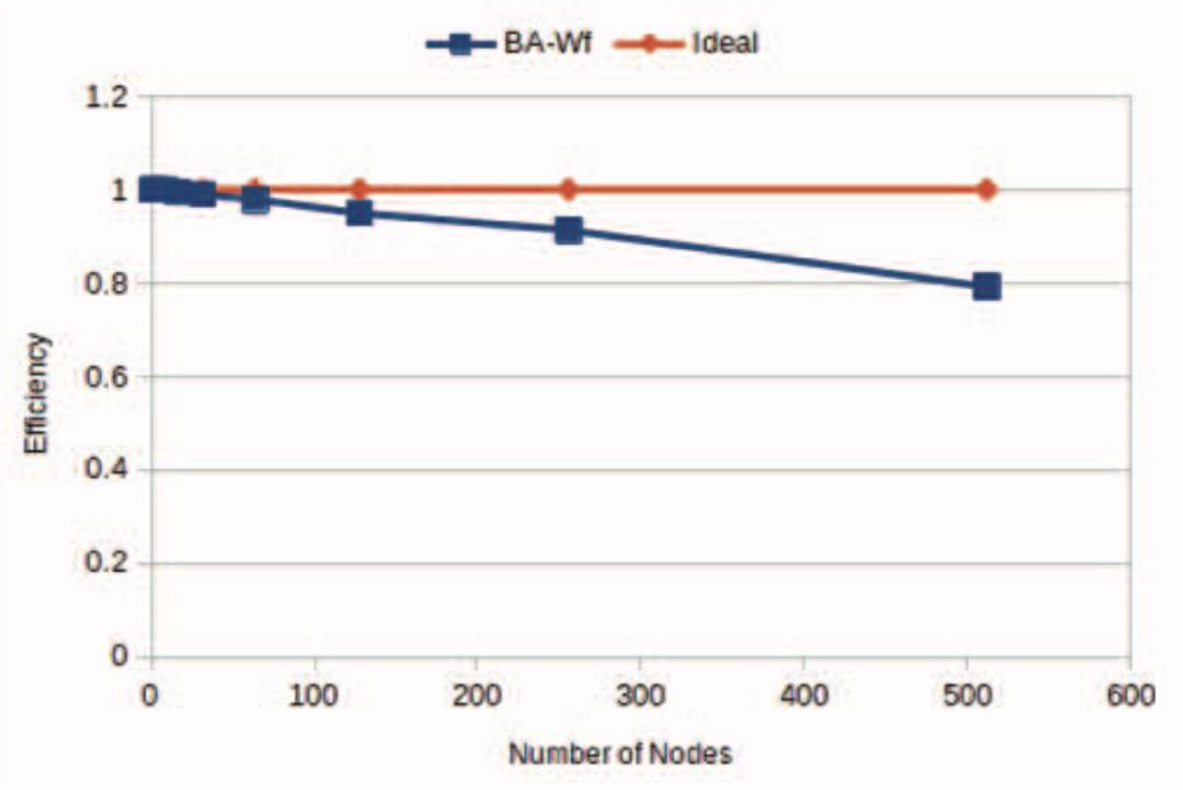


(b) Execution

WF scalability on Discoverer Supercomputer (Bulgaria)



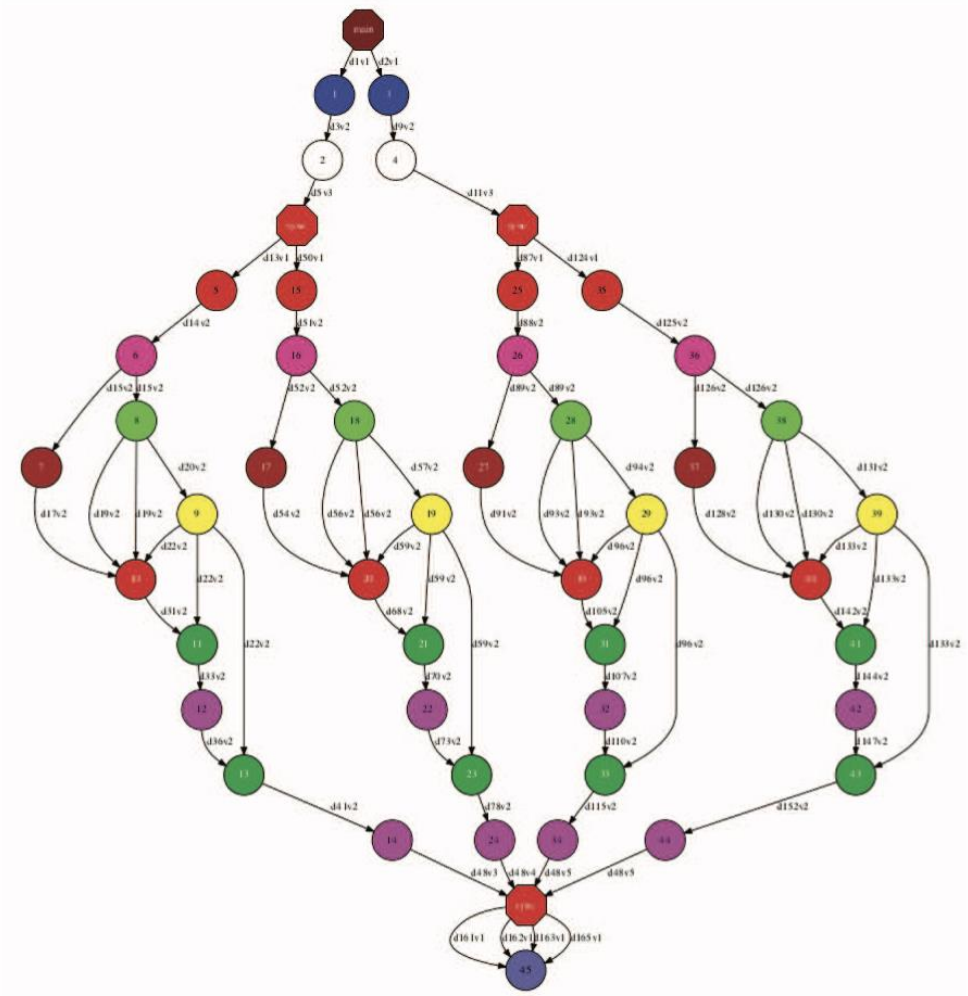
Strong scaling

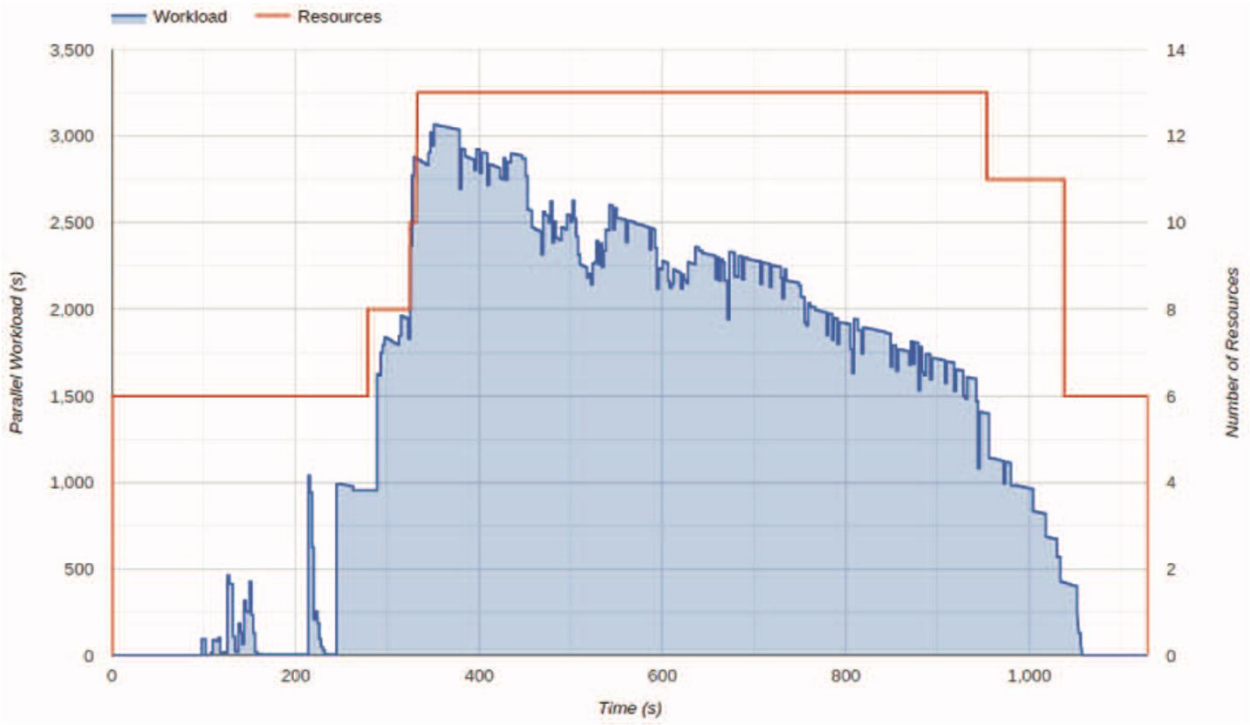


Weak scaling

Jorge Ejarque et al., 2022 IEEE 18th International Conference on e-Science

Malleability



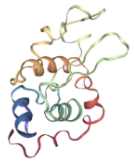


Jorge Ejarque et al., 2022 IEEE 18th International Conference on e-Science

Demonstration Workflows Collection

GROMACS PROTEIN MD SETUP

2023.1



This tutorial aims to illustrate the process of setting up a simulation system containing a protein, step by step, using the BioExcel Building Blocks library (biobb). The particular example used is the Lysozyme protein (PDB code 1AKI).

WorkflowHub | Jupyter Notebook | CWL | Python | Galaxy

Launch | Jupyter Notebook * | Galaxy | BioBB Workflows

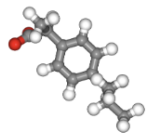
View tutorial | Open Github repository | Open documentation

gmx md protein

(*) Binder for biobb is a small installation and to promote fair use of our resources, one user is allowed to run only one notebook server at a time. Launching a new notebook server should stop the previous one. Users cannot see the notebooks run by other users, but please avoid entering secret data to the notebooks.

AUTOMATIC LIGAND PARAMETERIZATION

2023.1



This tutorial aims to illustrate the process of ligand parameterization for a small molecule, step by step, using the BioExcel Building Blocks library (biobb). The particular example used is the Ibuprofen small compound (3-letter code IBP, Drugbank code DB01050), a non-steroidal anti-inflammatory drug (NSAID) derived from propionic acid and it is considered the first of the propionics.

WorkflowHub | Jupyter Notebook | CWL | Python | Galaxy

Launch | Jupyter Notebook * | Galaxy | BioBB Workflows

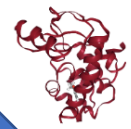
View tutorial | Open Github repository | Open documentation

gmx ligand

(*) Binder for biobb is a small installation and to promote fair use of our resources, one user is allowed to run only one notebook server at a time. Launching a new notebook server should stop the previous one. Users cannot see the notebooks run by other users, but please avoid entering secret data to the notebooks.

GROMACS PROTEIN-COMPLEX MD SETUP

2023.1



This tutorial aims to illustrate the process of setting up a simulation system containing a protein in complex with a ligand, step by step, using the BioExcel Building Blocks library (biobb). The particular example used is the T4 lysozyme L99A/M102Q protein (PDB code 3HTB), in complex with the 2-propylphenol small molecule (3-letter Code JZ4).

WorkflowHub | Jupyter Notebook | CWL | Python | Galaxy

Launch | Jupyter Notebook * | Galaxy | BioBB Workflows

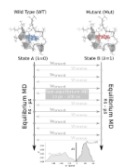
View tutorial | Open Github repository | Open documentation

gmx ligand md protein

(*) Binder for biobb is a small installation and to promote fair use of our resources, one user is allowed to run only one notebook server at a time. Launching a new notebook server should stop the previous one. Users cannot see the notebooks run by other users, but please avoid entering secret data to the notebooks.

MUTATION FREE ENERGY CALCULATIONS

2023.1



This tutorial aims to illustrate how to compute a fast-growth mutation free energy calculation, step by step, using the BioExcel Building Blocks library (biobb). The particular example used is the Staphylococcal nuclease protein (PDB code 1STN), a small, minimal protein, appropriate for a short tutorial.

WorkflowHub | Jupyter Notebook | Python

Launch | Jupyter Notebook *

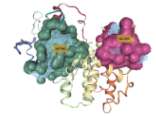
View tutorial | Open Github repository | Open documentation

free_energy gmx md

(*) Binder for biobb is a small installation and to promote fair use of our resources, one user is allowed to run only one notebook server at a time. Launching a new notebook server should stop the previous one. Users cannot see the notebooks run by other users, but please avoid entering secret data to the notebooks.

PROTEIN-LIGAND DOCKING (FPOCKET)

2023.1



This tutorial aims to illustrate the process of protein-ligand docking, step by step, using the BioExcel Building Blocks library (biobb). The particular example used is the Mitogen-activated protein kinase 14 (p38-α) protein (PDB code 3HEC), a well-known Protein Kinase enzyme, in complex with the FDA-approved Imatinib, (PDB Ligand code STI, DrugBank Ligand Code DB00619), a small molecule kinase inhibitor used to treat certain types of cancer.

WorkflowHub | Jupyter Notebook | CWL | Python | Galaxy

Launch | Jupyter Notebook * | Galaxy | BioBB Workflows

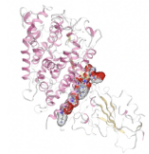
View tutorial | Open Github repository | Open documentation

docking ligand protein

(*) Binder for biobb is a small installation and to promote fair use of our resources, one user is allowed to run only one notebook server at a time. Launching a new notebook server should stop the previous one. Users cannot see the notebooks run by other users, but please avoid entering secret data to the notebooks.

MOLECULAR INTERACTION POTENTIALS

2023.2



This tutorial aims to illustrate the process of computing classical molecular interaction potentials from protein structures, step by step.

Launch | Jupyter Notebook *

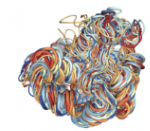
View tutorial | Open Github repository | Open documentation

ligand protein

(*) Binder for biobb is a small installation and to promote fair use of our resources, one user is allowed to run only one notebook server at a time. Launching a new notebook server should stop the previous one. Users cannot see the notebooks run by other users, but please avoid entering secret data to the notebooks.

PROTEIN CONFORMATIONAL ENSEMBLES GENERATION

2023.1



This tutorial aims to illustrate the process of generating protein conformational ensembles from 3D structures and analysing its molecular flexibility, step by step, using the BioExcel Building Blocks library (biobb). Workflow included in the ELIXIR 3D-Bioinfo Implementation Study: Building on PDBe-KB to chart and characterize the conformation landscape of native proteins

WorkflowHub | Jupyter Notebook | CWL | Python | Galaxy

Launch | Jupyter Notebook * | Galaxy

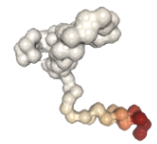
View tutorial | Open Github repository | Open documentation

protein

(*) Binder for biobb is a small installation and to promote fair use of our resources, one user is allowed to run only one notebook server at a time. Launching a new notebook server should stop the previous one. Users cannot see the notebooks run by other users, but please avoid entering secret data to the notebooks.

MACROMOLECULAR COARSE-GRAINED FLEXIBILITY

2023.2



This tutorial aims to illustrate the process of generating protein conformational ensembles from 3D structures and analysing its molecular flexibility, step by step.

WorkflowHub | Jupyter Notebook | CWL | Python | Galaxy

Launch | Jupyter Notebook * | Galaxy

View tutorial | Open Github repository | Open documentation

protein

(*) Binder for biobb is a small installation and to promote fair use of our resources, one user is allowed to run only one notebook server at a time. Launching a new notebook server should stop the previous one. Users cannot see the notebooks run by other users, but please avoid entering secret data to the notebooks.

BioBB developers team

Adam Hospital

Federica Battistini

Genís Bayarri

Modesto Orozco

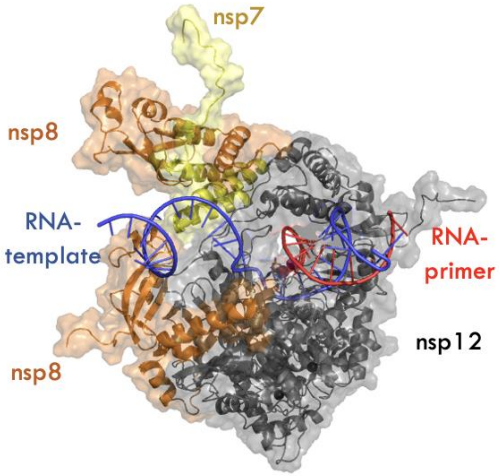
Pau Andrio

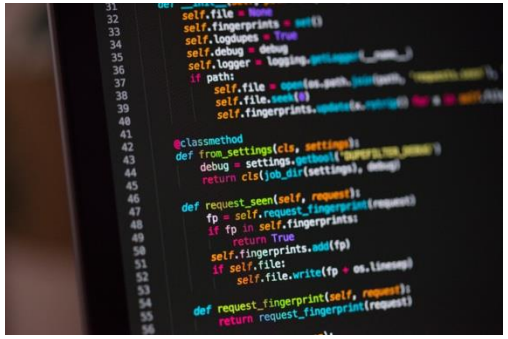
Josep Ll. Gelpí

BioExcel Center of Excellence, funded from the European Union's Horizon 2020 Framework Programme for Research and Innovation under Specific Grant Agreements No. 675728, 823830, 101093290 (BioExcel-1, BioExcel-2 and BioExcel-3).

Solving Grand Challenges requires Global Community Efforts

Infrastructure | **Tools & Data** | **Expertise** | **Community**



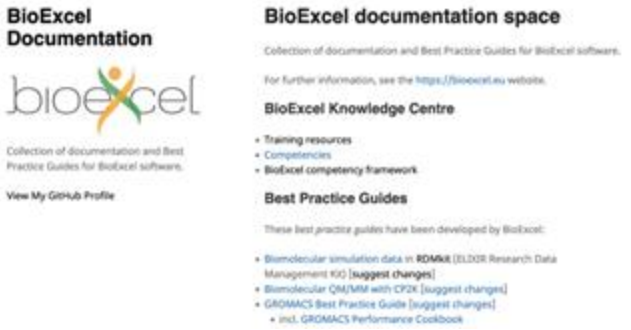


BioExcel Services

Support Forums: ask.bioexcel.eu

Webinar series: bioexcel.eu/webinars

Documentation: docs.bioexcel.eu



Training: bioexcel.eu/training

bioexcel

GROMACS · HADDOCK · PMX · BioBB

AMBASSADOR PROGRAM

Ambassador Council members

Wim VRANKEN
Belgium

Vrije Universiteit Brussel (VUB)
@wim-vranken-8b5a713

Anela IVANOVA
Bulgaria

Faculty of Chemistry and Pharmacy,
Sofia University "St. Kliment Ohridski"
@anela-ivanova-369384259

Gordan HORVAT
Croatia

Department of Chemistry,
Faculty of Science,
University of Zagreb
@gordan-horvat-a2a71423

Hector MARTINEZ-SEARA
Czech Republic

Institute of Organic Chemistry,
Czech Academy of Sciences

Kresten LINDORFF-LARSEN
Denmark

Linderstrøm Lang Centre for Protein Science,
Department of Biology,
University of Copenhagen
@lindorff
@LindorffLarsen

Luca MONTICELLI
France

Institut national de la santé et de la recherche médicale (INSERM),
and Molecular Microbiology and Structural Biochemistry (MMSB),
CNRS & University of Lyon

Zoe COURNIA
Greece

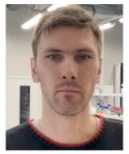
Biomedical Research Foundation
Academy of Athens
@zoecournia
@zoecournia @BRFAA_IIBEAA

Damien THOMPSON
Ireland

Department of Physics,
University of Limerick

Attilio VARGIU
Italy

Department of Physics,
University of Cagliari
@attilio-vittorio-vargiu-ab28a335
@AttilioVargiu @univca



Raitis BOBROVS
Latvia

Latvian Institute of Organic Synthesis
@osi_lv

Luka FILIPOVIC
Montenegro

University of Donja Gorica (UDG),
and NCC Montenegro
@luka-filipovic-me

Nathalie REUTER
Norway

Department of Chemistry,
University of Bergen
@reuter_nathalie @cbu_bergen @UiB

Artur GÓRA
Poland

Biotechnology Centre,
Silesian University of Technology
@_artgora @tunnelinggroup

Irina MOREIRA
Portugal

Center for Neuroscience and Cell Biology
& Faculty of Sciences and Technology,
University of Coimbra
@irina-moreira-a25a4a1
@IrinaSMoreira @UnivdeCoimbra
@FCTUC @DCV_UC @cnc_uc

Vlad COJOCARU
Romania

Computational Structural Biochemistry Group,
Institute of Advanced Studies in Science and Technology,
Babes-Bolyai University
@cojocarulab

Katarina NIKOLIĆ
Serbia

Department of Pharmaceutical Chemistry,
Faculty of Pharmacy,
University of Belgrade
@katarina-nikolic-30866728

Franci MERZEL
Slovenia

National Institute of Chemistry,
Ljubljana
@kemijski

David DE SANCHO
Spain

University of the Basque Country &
Donostia International Physics Center,
San Sebastian
@DIPCehu @ehu_kimika

Ambassador Program Events

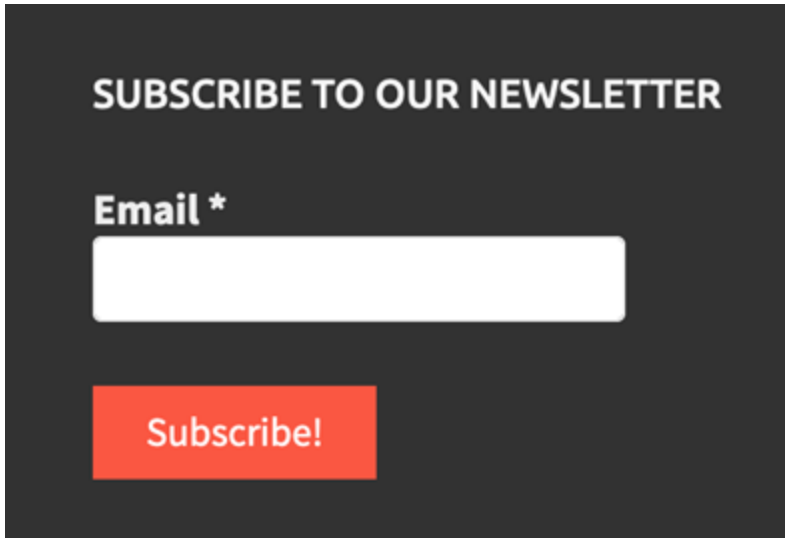
Training event	Lead	Organizers	Date
Carpathian Edition	Slovakia	Austria, Czechia, Hungary	Autumn 2023 (18-19 October)
Atlantic Edition	Portugal	France, Ireland, Spain	Autumn 2024 (26-28 November)
Balkan Edition	Bulgaria	Montenegro, North Macedonia, Romania, Serbia	Spring 2025 (21-22 May)
Adriatic Edition		Croatia, Slovenia	Autumn 2025
Aegean Edition		Cyprus, Greece, Türkiye	Spring / Autumn 2026
Aurora Edition		Denmark, Finland, Norway	Spring / Autumn 2026

Social media

Newsletter	> 2000 subscribers
Twitter @BioExcelCoE	> 3000 followers
YouTube @BioExcelCoE	74 webinars some > 7000 views
LinkedIn	> 1000 followers
Events	> 300
Trained researchers	> 2000

Contact

http://bioexcel.eu



Co-funded by
the European Union

EuroHPC
Joint Undertaking

Utrecht
University

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación