





Biomolecular Simulations@ExaScale


The ExaBioSim project


Sarah Anne Harris


Physics and Astronomy, University of Sheffield, UK

Acknowledgements

HPC Development James Gebbie-Rayet Christopher Woods Charlie Laughton Alan Real	Massive Systems Sarah Harris Adrian Mulholland Syma Khalid Jon Essex	Integration with Experiment Tom Burnley Franca Fraternali Dmitry Nerukh	 CCP BioSim HEC BioSim
Ensemble Computing Antonia Mey Danny Cole Shozeb Haider Edina Rosta Julien Michel Nicolas Foloppe	Multiscale Modelling Oli Henrich Agnes Noy Rosana Collepardo Davide Marenduzzo	Research Software Engineer Rob Welch	 N8 Bede COMPUTATIONALLY INTENSIVE RESEARCH
		Project Support Pirjo Johnson	 archer2
			 xCAL18UR 10

 UK Research and Innovation

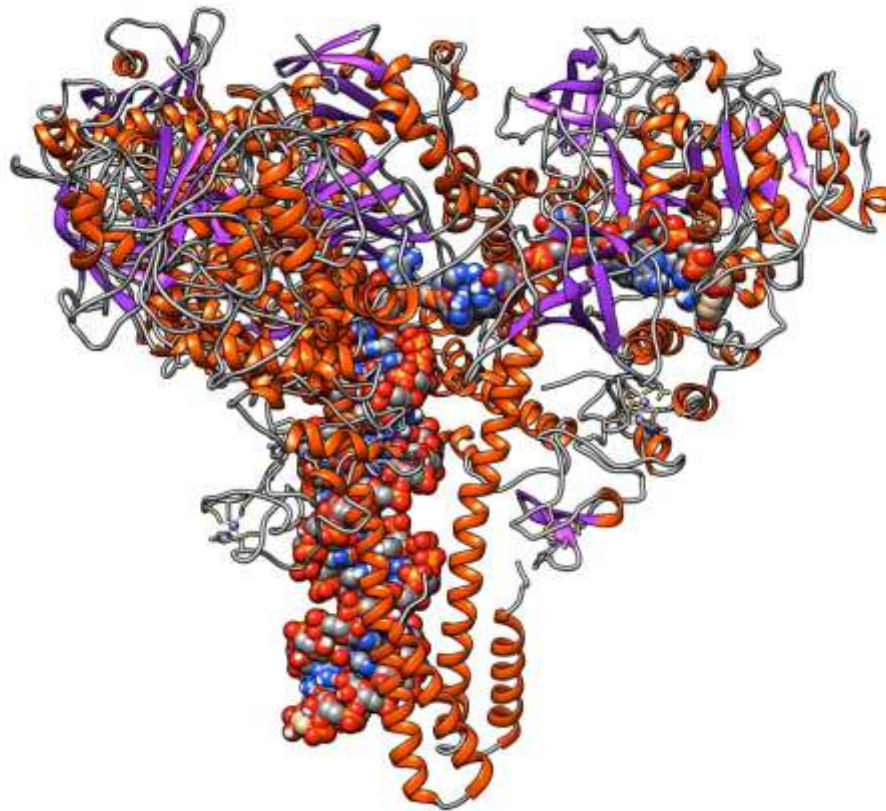
 Science and Technology Facilities Council

 Engineering and Physical Sciences Research Council

Engineering Supercomputing Platforms for Biomolecular Applications
<https://arxiv.org/html/2506.15585v1#S8>

What is a Biomolecular Simulation?

SARS-CoV-2
mRTC



The “computational microscope” `

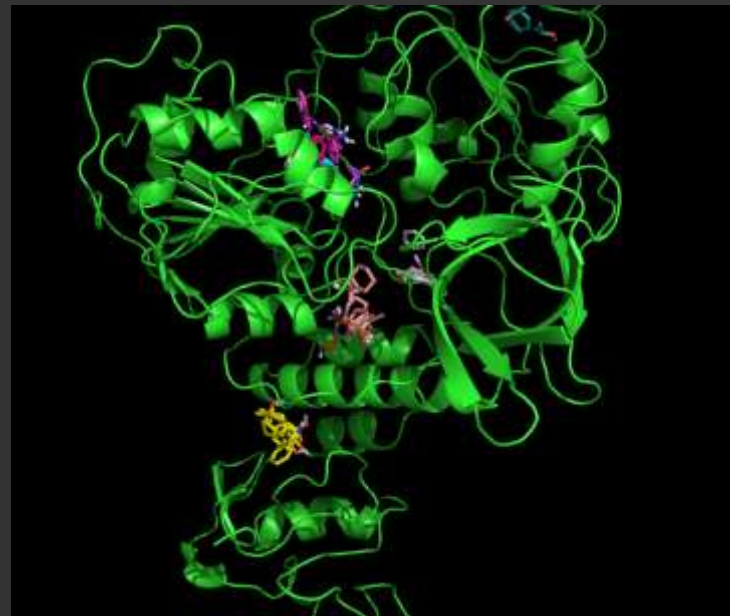
Dror, Dirks, Grossman, Xu & Shaw, Annu. Rev. Biophys. 2012;41:429-52.

Atomistic MD for Drug Design

Codes are:

AMBER (US)
GROMACS (EU)
NAMD (US)
LAMMPS (US)
openMM (US)

All codes run well on
(Nvidia) GPUs but have
multi-physics functionality
that is less well optimised.



Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

[View Article Online](#)
Free Access

[Check for updates](#)

One has DOI: 10.1039/c9sc02775a

All publication charges for this article have been paid for by the Royal Society of Chemistry

Modelling the active SARS-CoV-2 helicase complex as a basis for structure-based inhibitor design†

Dénes Berta,^a Magd Badaoui,^a Sam Alexander Martino,^a Pedro J. Buques,^a Andrei V. Pislakov,^a Nadia Elghobashi-Meinhardt,^a Geoff Wells,^a Sarah A. Harris,^a Elisa Frezza,^a and Edina Rosta^a

Community Support



Our community DO NOT install their own codes.

HECBioSim have specialist support who do this on behalf of the community

Advantages:

1. Avoids duplicating effort
2. Code properly installed and optimised for efficiency
3. Core expertise easy to find
4. Prevents multiple copies of the same software
5. HPC configured to benefit community

Disadvantages:

1. Not how all communities work
2. Requires intervention for bespoke codes
3. Support staff require domain knowledge



ExaBioSim Benchmarking Tools

Open Access benchmarking suite incorporating both test systems and performance monitoring tools

The screenshot displays the GitHub repository page for **HPCBioSim/hpcbench**. The repository is categorized under **Tools**. The main content area lists several benchmarks, including **hpcbench**, **spglib**, **LIGAND**, **MANIFEST**, **REACH**, and **wrapup**. The **About** section states: "A set of benchmarks for biomolecular simulation tools." The **Current support** section lists supported simulators: **GROMACS** and **AMBER**.

20k atoms
3NIR Crambin

61k atoms
1WDN Glutamine-Binding Protein

465k atoms
hEGFR Dimer of 1IVO and 1NQL

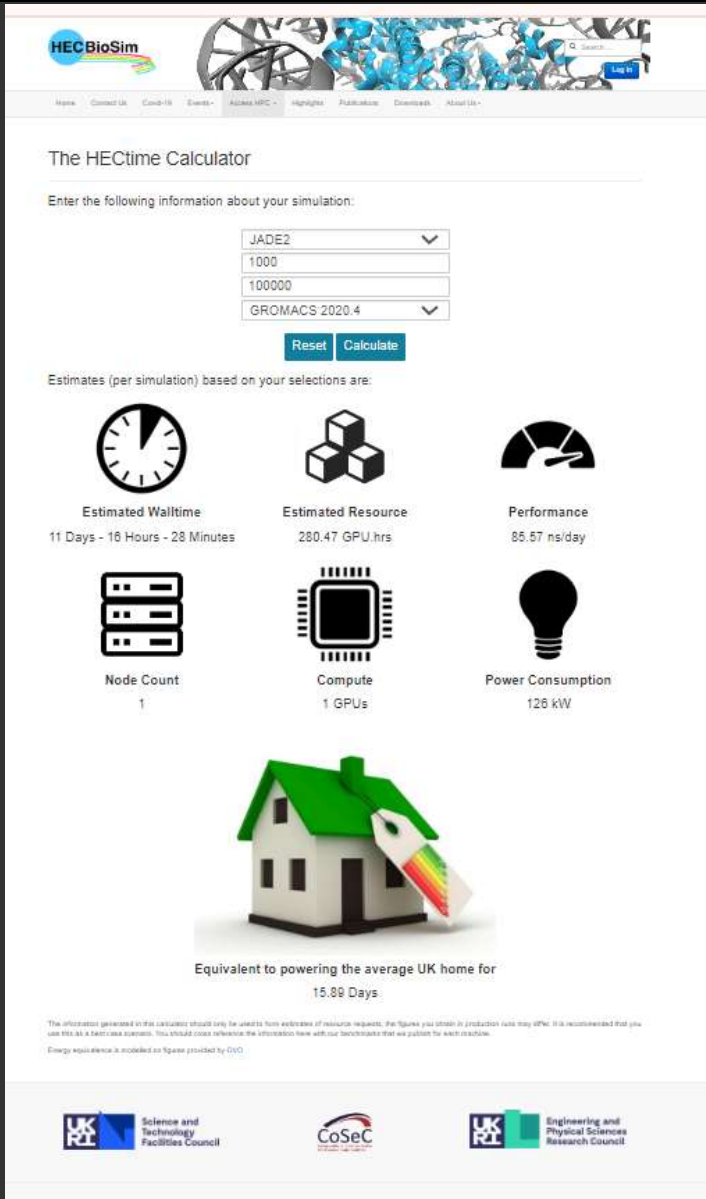
1.4M atoms
Two hEGFR Dimers of 1IVO and 1NQL

1.4M atoms
Two hEGFR tetramers of 1IVO and 1NOL



<https://github.com/HECBioSim/hpcbench>

HECBioSim Benchmarking Tools



The HECtime Calculator

Enter the following information about your simulation:

JADE2
1000
100000
GROMACS 2020.4

Reset Calculate

Estimates (per simulation) based on your selections are:

- Estimated Walltime**
11 Days - 16 Hours - 28 Minutes
- Estimated Resource**
280.47 GPU.hrs
- Performance**
85.57 ns/day
- Node Count**
1
- Compute**
1 GPUs
- Power Consumption**
126 kW

Equivalent to powering the average UK home for
15.89 Days

The information generated in this calculator should only be used to form estimates of resource requests, the figures you obtain in production runs may differ. It is recommended that you use HEC as a first point of enquiry. You should cross reference the information here with our benchmarks that we publish for each workflow.

Energy equivalence is modelled on figures provided by OVO

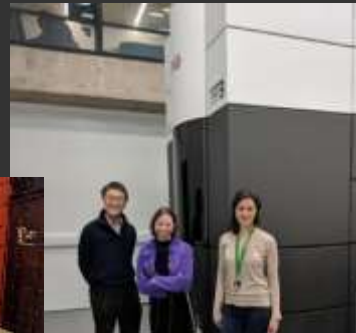
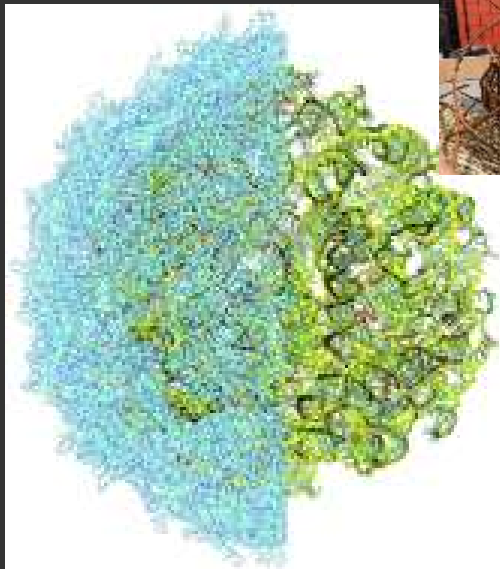
Science and Technology Facilities Council
CoSeC
Engineering and Physical Sciences Research Council

Benchmark suite helps us support the HECBioSim community to apply for compute time

But our community are diverse....

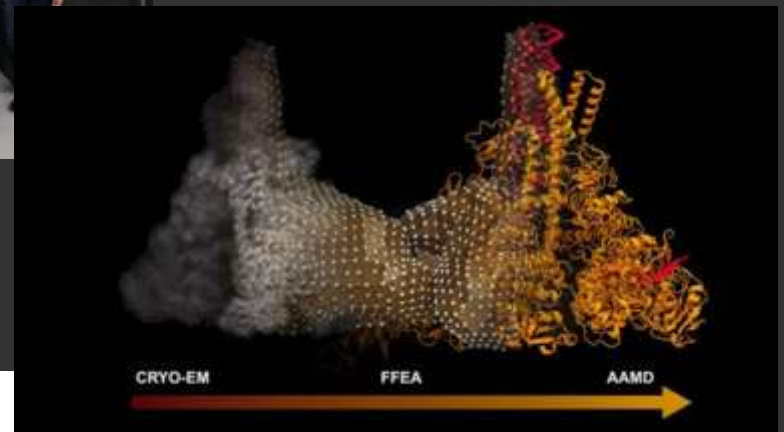
Large systems

> 2 millions atoms
Too big for GPU



Interface with
Experiments

Cryo-EM



Faraday Discussions

Click here to view the discussion

PAPER

Reconstruction and validation of entire virus model with complete genome from mixed resolution cryo-EM density

Wladimir S. Farafonov^{1,2,3}, Richard Stock^{1,2} and Dmitry Nersisyan^{1,2}



View Article Online

DOI: 10.1039/C9FD00000A

The International Union of Pure and Applied Chemistry
IUPAC
© The International Union of Pure and Applied Chemistry
https://doi.org/10.1039/C9FD00000A

Special Issue Paper

Intelligent resolution: Integrating Cryo-EM with AI-driven multi-resolution simulations to observe the severe acute respiratory syndrome coronavirus-2 replication-transcription machinery in action

Arka Tripathi^{1,2,3}, Debas Gupte^{1,2}, Michael Saka^{1,2}, Zongyi Li^{1,2}, Alexander Brice^{1,2,4}, Martin Ziegler^{1,2}, Hong Ma^{1,2}, Anika Chel^{1,2}, David Clark^{1,2}, David J. Hardy^{1,2}, Eam Bannister^{1,2}, Lei Huang^{1,2}, John McAlister^{1,2}, Michael Emami^{1,2}, Hyeonwang Yoo^{1,2}, Junyi Xie^{1,2}, Ananthu Thirumangalakudi^{1,2}, Ganesh Ram^{1,2}, Jeeva Lee^{1,2}, Neel Tinku^{1,2}, Geoffrey Wells^{1,2}, Venkatesh Mysore^{1,2}, Thomas Giblin^{1,2}, James Phillips^{1,2}, S. Chitra Chennambhotla^{1,2}, Ian Foster^{1,2,3}, Rick Stevens^{1,2}, Ananta Anandhan^{1,2}, Venkatesh Venkatesh^{1,2}, John E. Stone^{1,2}, Ewald Tajkhorshid^{1,2}, Sarah A. Harris^{1,2}, and Arvind Ramakrishnan^{1,2}

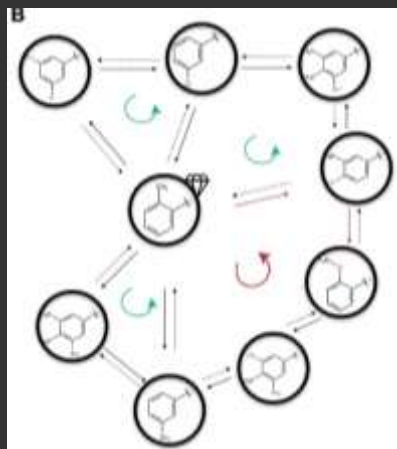
Big issues with
data transfer

Ensemble Calculations

Complex workflows involving very large numbers of “small” simulations with complex dependencies and unpredictable in-flight decision-making steps.

Hardware solutions

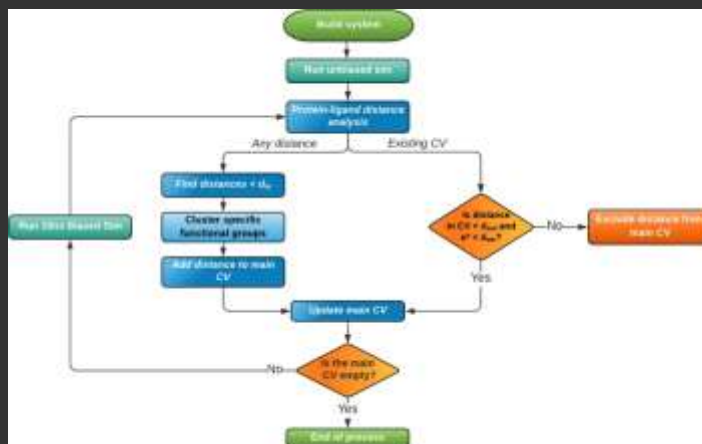
Enhanced Sampling



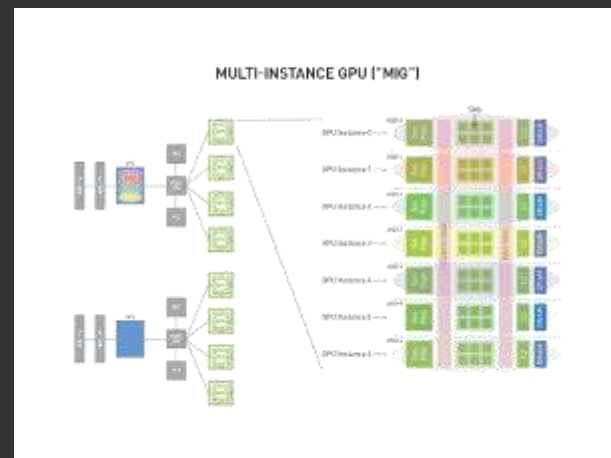
Mey et al., *Living J. Comp. Sci.* 2020, 2(1), 18378

FEP

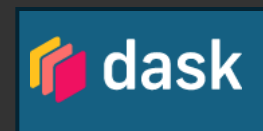
Free energy
perturbation



Badaoui et al *J. Chem. Theory Comput.* 2022, 18, 4, 2543–2555



Software solutions

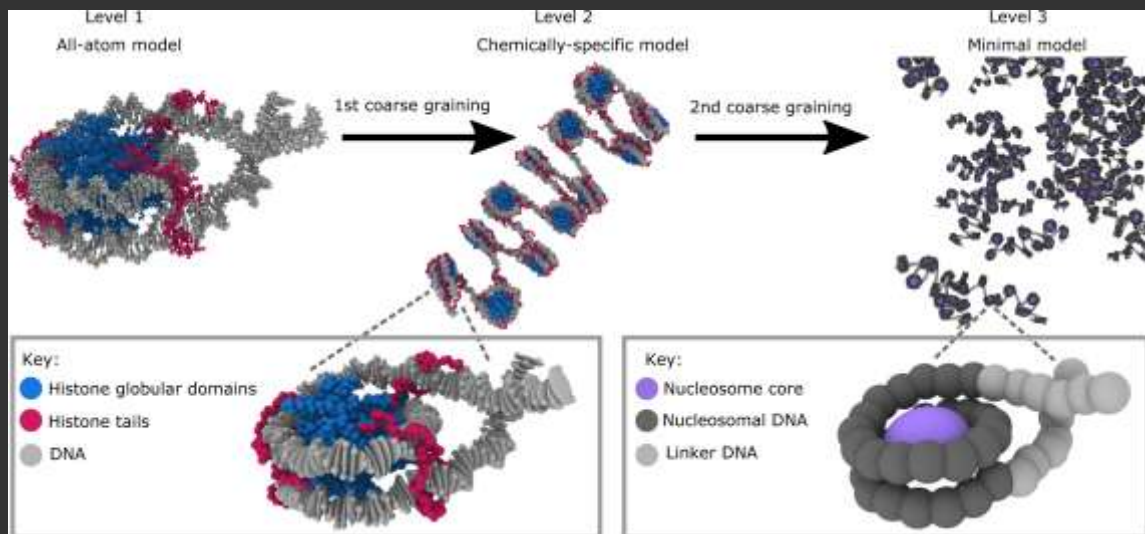


<https://bitbucket.org/claughton/crossflow>

Multiscale: From QM to continuum



QM codes - high
memory node
requirements –
mainly CPU codes



Coarse-grained
simulations – CPU
only

Supercomputer says “NO”



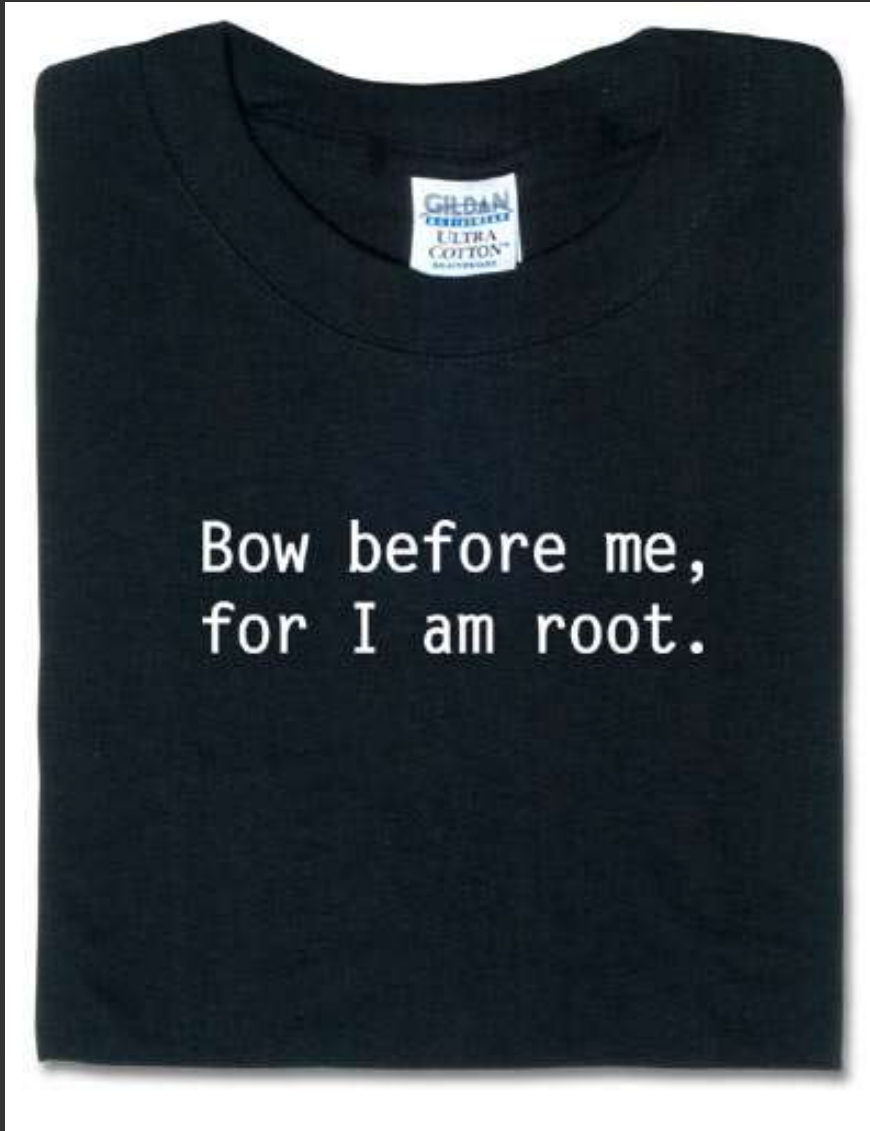
Large systems

> 2 millions atoms
Too big for GPU

Cryo-EM

Transferring large
data sets is a
problem

SysAdmin says “NO”



Many HPC setups
have a limit on the
number of jobs in
the queue and/or
very short queues

ExaBioSim “Wish List” for Exascale

Heterogeneous architectures ~ a bit of everything
makes everyone happy
(CPU+GPU+BigMemory)

Talk to your users when deciding how to
configure your system (e.g. flexible queue
times/large job numbers/internet access etc)

RTPs and users should make our conversations
more effective.

HPC should be
designed for the users
you have ~ not those
you wish you had.

Anton specialist MD HPC

ACM DIGITAL LIBRARY

Home Browse Alert Sign In

Journals Magazines Proceedings Books SIGs Conferences People

Search ACM Digital Library

Conference Proceedings Upcoming Events Authors Affiliations Award Winners

Home > Conferences > SC > Proceedings > SC '21 > Anton 3: twenty microseconds of molecular dynamics simulation before lunch

RESEARCH ARTICLE OPEN ACCESS

Anton 3: twenty microseconds of molecular dynamics simulation before lunch

Authors: David E. Shaw, Peter J. Adams, Deep Jaiswal, Albert A. Bock

SC '21: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis • November 2021 • Article No. 1 • Pages 1–11 • <https://doi.org/10.1145/3458817.3487367>

Published: 13 November 2021



MD performance

