Grand Challenges

The Collaborative Computational Project Number 5 in Computer Simulation of Condensed Phases (CCP5)

Over the next five to ten years, CCP5 will develop a suite of new processing tools for the simulation of condensed phases via sustainable strategies to enable innovative discoveries

The Community

CCP5 Collaborative Computational Project
Computer Simulation of Condensed Phases

www.ccp5.ac.uk

CCP5 is the UK's Collaborative Computational Project focused on the computer simulation of condensed-phase materials. Established over 40 years ago to strengthen national collaboration, CCP5 provides researchers across universities and public laboratories with free access to an extensive suite of supported and contributed software for non-commercial use. This open-access model fosters strong industrial engagement through workshops, conferences, and dedicated industry-focused events. CCP5's remit spans atomistic to mesoscopic simulations, bringing together diverse expertise to enable world-class research that is both scientifically ambitious and cost-effective. The UK hosts one of the world's most active and successful communities in condensed-phase simulation, and CCP5 plays a central role in coordinating communication, sharing best practices, and supporting emerging research directions. With a strong international profile, CCP5 maintains and develops leading simulation codes, enhances software functionality, and supports the community in delivering innovative science at the forefront of materials modelling.

The Challenge

To advance machine learning, reproducibility, and data infrastructure for next-generation molecular simulation

Molecular modelling now produces data at a scale that exceeds our ability to analyse and exploit it. While machine learning already delivers major advances in force-field development and shows promise across the modelling workflow, current ML models still lack the robustness and transferability needed for real-world materials and chemistry problems. Meanwhile, modern supercomputers generate molecular dynamics trajectories faster than today's predominantly serial analysis tools can process them, creating a widening performance bottleneck. High-throughput screening and complex multi-software workflows are feasible on advanced HPC and cloud platforms, but they generate vast, heterogeneous datasets that must be efficiently analysed, made FAIR, and integrated into reproducible pipelines. Addressing these challenges requires scalable trajectory analysis tools, deeper and more reliable ML integration, interoperable data

infrastructures, and automated end-to-end workflows that keep pace with rapidly expanding simulation capabilities.

The Solution

GPU-enabled analysis, strong validation frameworks, and FAIR data ecosystems together provide the foundation for next-generation, simulation-driven materials discovery

These challenges can be addressed by accelerating the most computationally demanding analysis tools through GPU-native implementations, which are now feasible due to the widespread availability of modern accelerator hardware. Rigorous validation and benchmarking are essential to ensure that data-driven force fields and ML models perform reliably in real-world applications. Equally important are new strategies that improve reproducibility, reusability, and long-term data stewardship. A robust Python-based workflow infrastructure plays a central role here: it enables automation, data provenance, and the seamless management and sharing of complex high-throughput simulation pipelines. Although such infrastructures have traditionally supported quantum mechanical workflows, they are now being extended to classical molecular dynamics, allowing researchers to run large simulation campaigns, track the full history of each result, and store outputs in FAIR-compliant formats. By building on existing CCP5 tools, the community can streamline the use of multiple MD codes, unify data handling, and ensure scalable, reproducible, and GPU-accelerated analysis for next-generation molecular simulation.

The Outcome

An end-to-end ecosystem of interoperable, GPU-ready tools with FAIR data foundations that supports rigorous structure–function analysis and the computational development of advanced materials.

A suite of GPU-accelerated post-processing tools will be available by 2027, enabling trajectory analysis and data handling to scale with modern high-performance computing architectures. This will be complemented by horizontal interoperability across major molecular simulation codes, providing a consistent foundation for integrating workflow managers and ensuring that data generated within the community adheres to agreed FAIR principles. Together, these developments will position CCP5 to continue serving as the central hub for the UK's world-leading activities in functional and advanced materials design through molecular simulation.

More Information

CoSeC
www.CoSeC.ac.uk
CoSeC@stfc.ac.uk

CCP5
www.ccp5.ac.uk
alin-marin.elena@stfc.ac.uk