MDDB: Molecular Dynamics Data Bank. The European Repository for Biosimulation Data

Description

After decades of development and a Nobel prize, Molecular Dynamics (MD) has reached maturity. It is no longer an exotic technique used by a small group of theoreticians but rather a method extensively used by a very large community of users. Millions of supercomputer hours are devoted to collecting trajectories, thus producing a deluge of simulation data that the community is unable to handle. A poor tradition of data sharing and the lack of appropriate infrastructures to do so lead to the loss of data after limited analysis that most likely revealed only a small fraction of the information contained. Sparse initiatives to build trajectory repositories have encountered difficulties related to: i) the lack of trust of the community in the reliability of the data deposited; ii) the lack of interoperable standards and simulation ontologies; iii) uncertainties regarding the database technology required; v) difficulties of the users to interact in an open manner with the data; and vi) disconnection of the MD-field with neighboring communities.

The MDDB projects intends to design a European-scale repository of MD simulation (and associated analysis tools), which will: i) optimize computational resources; ii) favor the analysis (and meta-analysis) of trajectories for many different perspectives and fields; iii) guarantee a fast and efficient interchange of information between groups; and iv) facilitate the integration of the MD simulation field into neighboring communities. The overall result will be a more efficient use of MD and the integration of the MD field intomainstream biology and chemistry research. Molecular dynamics, Simulation databases, interoperability, big data, data standards, good practices, FAIR.

Barcelona Supercomputing Center - Centro Nacional de Supercomputación