



**Project Acronym:** MDDB

**Project title:** Molecular Dynamics Data Bank. The European Repository for Biosimulation Data

**Call:** HORIZON-INFRA-2022-DEV-01

**Topic:** HORIZON-INFRA-2022-DEV-01-01- Research Infrastructure Concept Development

**Project Number:** 101094651

**Project Coordinator:** Institute for Research in Biomedicine (IRB Barcelona)

**Project start date:** 01/03/2023

**Project end date:** 28/02/2026

## Deliverable 5.3: First policy brief

**Work Package:** WP 5 – Project management and communication

**Lead beneficiary:** IRB-CERCA

**Dissemination level:** PUBLIC

Due date: 31/03/2024

Actual submission date: 28/03/2024



Funded by  
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

## Document history

Version	Contributor(s)	Partner	Date	Comments
0.1	Lejla Bektic	IRB	15/03/2024	First draft
0.2	Adam Hospital, Modesto Orozco	IRB	22/03/2024	Revision and contribution to the document
0.3	Magnus Lundborg	KTH	27/03/2024	Revision
1.0			28/03/2024	Final version

## Table of contents

1	INTRODUCTION: CALL FOR FAIR UNIFIED DATABASE .....	5
2	ALIGNMENT WITH EXISTING INITIATIVES.....	5
3	NATIONAL ROADMAPS TO ESFRI ROADMAP.....	5
4	KEY STRATEGIES.....	6
5	CONCLUSIONS.....	6

## Executive summary

The purpose of this document is to provide an initial policy brief describing the necessity for MDDB, delineating its core objectives, and elucidating its alignment with existing initiatives in both the Europe and global contexts. Recognising the importance of understanding and actively engaging with national roadmaps alongside the ESFRI framework, this brief emphasises the strategic significance of MDDB's establishment and its role as an integral component of the European research infrastructure landscape. Furthermore, it presents key strategies aimed at realising project goals and advancing the development of MDDB.

## 1 INTRODUCTION: CALL FOR FAIR UNIFIED DATABASE

MDDDB aspires to establish a research infrastructure tailored for numerous communities, including those engaged in biomolecular simulation as well as the neighbouring areas such as, structural biology, biophysics, medicinal chemistry, biotechnology, bioinformatics, genomics and personalised medicine. Despite their reliance on molecular dynamics simulations, there exists a glaring deficiency in software interoperability across different domains. Consequently, users find themselves repeatedly conducting identical calculations to generate simulations, a process prone to inefficiency and redundancy.

Our objective is to unify these diverse communities under a common platform employing FAIR (Findable, Accessible, Interoperable, Reusable) data principles, fostering collaboration towards the enhancement of protocols, ensuring seamless interoperability, refining analysis pipelines, customising solutions to accommodate various user needs, and creating a data-culture in the field. In essence, our mission is deeply interdisciplinary, aiming to break down barriers between scientific domains and drive collaborative advancements in research and innovation.

## 2 ALIGNMENT WITH EXISTING INITIATIVES

MDDDB recognises the importance of aligning with established initiatives to leverage existing resources, avoid duplication of efforts, and foster a cohesive scientific ecosystem. In the realm of biomolecular data repositories, numerous databases play vital roles in storing and disseminating critical biological data, focused on sequence (UniProt) or 3D structure (PDBe, EMDB, and AFDB). MDDDB is poised to complement these databases by addressing the unmet needs of biomolecular dynamics behaviour repository, thereby bridging the gap in the research infrastructure landscape and enriching the Life Science Data ecosystem.

Through its consortium members, including the representatives from EMBL-EBI, MDDDB maintains direct communication with prominent Life Science data repositories. Furthermore, several partners actively participate in initiatives such as EOSC-life, ensuring collaboration with existing infrastructures at both the European and global level, as well as, BioExcel CoE, leveraging their expertise to discern the most suitable strategies and offer best practices guidelines in the field.

## 3 NATIONAL ROADMAPS TO ESFRI ROADMAP

As MDDDB embarks on its mission to establish itself as a cornerstone for biomolecular simulation research in Europe, it recognizes the significance of aligning its objectives with both national and European strategic priorities. Within the expansive framework of European research infrastructures, understanding and engaging with national roadmaps represent essential steps in our journey towards integration with the broader ESFRI roadmap.

We acknowledge the pivotal role of national roadmaps in shaping research infrastructure investments and priorities within individual countries. These roadmaps serve as invaluable guides, offering insights into the scientific needs, technological capabilities, and socio-economic priorities at the national level. By exploring these national roadmaps and aligning with the objectives outlined in the ESFRI roadmap, MDDDB seeks to contribute to pan-European efforts to advance scientific research and innovation. We recognize that inclusion in the ESFRI roadmap not only opens doors to European funding and resources but also validates our contributions to the broader scientific community.

## 4 KEY STRATEGIES

- **Implementing good practices:** To strengthen the reliability and integrity of data available in MDDB, we will establish clear guidelines and protocols for collection of simulation data, alongside stringent quality check procedures to ensure data completeness and consistency.
- **Defining standards of interoperability:** Compatibility and seamless integration of data into workflows and pipelines of other databases is crucial for collaboration across different communities and strengthening the impact of MDDB. We are actively working on defining interoperable data formats and seamless integration of MD data into the Life Science Data ecosystem.
- **Establishing clear guidelines for access:** To guarantee fair and equal access to MDDB, we will establish transparent guidelines and procedures for users seeking to deposit trajectory, access to data, post-analysis, training, and other services, e.g. deployment of a new federated node and help for the deposition of new simulations.
- **Addressing data storage and analysis capacity:** We are working on a distributed database policy and actively seeking commitment of computing centres and funding agencies to ensure sustainability of MDDB and mitigate the technical obstacles related to data storage and analysis capacity.
- **Engaging stakeholders:** Collaboration with stakeholders, spanning academia, industry, and government agencies is essential for the success and sustainability of MDDB. We are actively engaging stakeholders in the planning, implementation, and evaluation of our activities to ensure that MDDB effectively addresses the needs and priorities of all stakeholders, thereby maximising its impact and relevance.
- **Involving national communities and HPC centres:** As HPC centres play an important role in creating large data centres with powerful computational facilities, our aim is to engage national communities connected to these centres to actively participate in building, expanding, and maintaining local infrastructures coordinated by MDDB.
- **Monitoring and evaluation:** Regular monitoring and evaluation will be conducted to assess the impact and effectiveness of MDDB. Feedback from users and stakeholders will be solicited and used to inform continuous improvement efforts. This way, we can enhance MDDB's functionality and relevance and keep it responsive to the evolving needs of the community.

## 5 CONCLUSIONS

In conclusion, MDDB aims to transcend disciplinary boundaries and catalyse collaborative advancements in molecular dynamics simulations and its related fields. Serving as a FAIR unified platform, it holds significant potential to foster transformative discoveries and innovation while addressing pressing societal challenges such as, development of new biological therapies, advancement of personalised medicine, and training of AI models for high-throughput prediction of bioactivities. Leveraging insights gained from national roadmaps and actively engaging with ESFRI initiatives, MDDB strives to position itself as a leading research infrastructure of strategic importance at the European level.

By adhering to the principles outlined in this policy brief and collaborating with stakeholders, we will maximise the impact of MDDB and contribute to the advancement of science and technology on a global scale.